THEORY AND APPLICATIONS OF COMPOUND MATRICES

by

ROBERT CHARLES THOMPSON

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ARTS

IN THE DEPARTMENT OF MATHEMATICS

We accept this thesis as conforming to the standard required from candidates for the degree of MASTER OF ARTS.

Members of the DEPARTMENT OF MATHEMATICS.

THE UNIVERSITY OF BRITISH COLUMBIA

April, 1956.

ABSTRACT

If A is an n-square matrix, the p-th compound of A is a matrix whose entries are the p-th order minors of A arranged in a doubly lexicographic order. In this thesis some of the theory of compound matrices is given, including a short proof of the Sylvester-Franke theorem. This theory is used to obtain an extremum property of elementary symmetric functions of the k largest (or smallest) eigenvalues of non-negative Hermitian (n.n.h) matrices. Extensions of theorems due to Weyl and Wielandt are given. The first of these relates elementary symmetric functions of singular values of any matrix A with the same elementary symmetric functions of the eigenvalues. The second gives an extremum property of arbitrary eigenvalues of n.n.h matrices and enables inequalities relating the eigenvalues of A, B with the eigenvalues of $A + B$ to be given (A, B, n.n.h.). Finally, a norm inequality for an arbitrary matrix is given.

ACKNOWLEDGEMENTS

The author wishes to express his thanks to Dr. M.D. Marcus for suggesting the topic of this thesis and for his invaluable advice during its preparation, and also to Dr. B.N. Moyls for his helpful suggestions. He gladly acknowledges his indebtedness to the National Research Council of Canada for the financial assistance that made this study possible.

TABLE OF CONTENTS

1. <u>INTRODUCTION:</u>

The object of this thesis is twofold: to develop and apply the Grassman exterior algebra to eigenvalue problems for finite matrices. The first part of the thesis deals with the construction of the Grassman algebra. This is done by defining a mapping from an n-dimensional vector space to another $\binom{n}{p} = \frac{\lfloor\underline{n}}{\lfloor\underline{p}\;\lfloor\underline{n-p}}$ dimensional vector space that satisfies certain multilinear and alternating properties. We shall prove some of the more important properties of this mapping, including the well known Sylvester-Franke Theorem. Some of the proofs developed here will be shorter than those that are usually given. This material will be found in sections II and III.

The second part of the thesis is concerned with applications of the Grassman algebra. These applications will be concerned with the location of eigenvalues of $n \times n$ matrices A over the complex field; that is, to locating the complex roots of the polynomial equation in λ

$$\det(\lambda I - A) = 0$$

where I is the $n \times n$ unit matrix.

Using the Grassman algebra, certain maximum and minimum characterizations of the eigenvalues of a matrix A will be proved. These maximum and minimum properties will be extensions of earlier properties given by Fan (3).

These results will enable us to give an extension of a theorem due to Wielandt (15). An extension of a theorem due to Minkowski (13) will also be given. For the Minkowski theorem it will be necessary to state, without proof, a recent result due separately to Marcus (11) and Bohnenblust (2). Finally a new result concerning the norm of a matrix will be given; the proof will be made to depend on the extremal properties of eigenvalues proved earlier. All this material will be found in sections IV, V and VI.

Finally, a brief historical survey of the material covered in the thesis will be given in section VII.

Throughout the thesis, the numbers in brackets refer to the references given in the bibliography in section VIII.

II. THE GRASSMAN ALGEBRA.

The chief instrument to be used is the Grassman exterior algebra. Before constructing this, it is necessary to introduce some notation. Let us agree to denote by V_r the r-dimensional vector space consisting of all r-tuples of complex numbers. Let e_i denote the unit r vector whose entries are all zero except for a one in the ith position. Later on, we shall have occasion to consider simultaneously two vector spaces, V_p and $V_{\binom{n}{p}}$. The phrase "unit vectors" will always mean the vectors e_i just mentioned. For convenience we shall use $\overset{p}{\underset{1}{\pi}} \otimes V_r$ to mean $V_r \times V_r \times \ldots \times V_r$, to p factors; that is, the cartesian product of V_r with itself p times. If $p \leq n$ let $Q_{pn} = \left\{ \{i_1, \ldots, i_p\} \,\middle|\, 1 \leq i_1 < \ldots < i_p \leq n \right\}$; that is, the elements of Q_{pn} are sets of p integers chosen from $1, \ldots, n$ and arranged in increasing order. Note that Q_{pn} has $\binom{n}{p}$ elements.

In the sequel it will be necessary to consider a lexicographic ordering of the elements of Q_{pn}. By this is meant, roughly, the scheme used to order the words in a dictionary. If $\{i_1, \ldots, i_p\}$, $\{j_1, \ldots, j_p\}$ are two elements of Q_{pn} then, in the lexicographic ordering, $\{i_1, \ldots, i_p\}$ comes before $\{j_1, \ldots, j_p\}$ if and only if there is an integer $m \leq p$ such that $i_1 = j_1$, $i_2 = j_2$, ..., $i_{m-1} = j_{m-1}$, $i_m < j_m$.

Having disposed of the notational preliminaries, we can now give the first definition.

Definition 2.1: Let $\{x_1, \ldots, x_p\} \in \overset{p}{\underset{1}{\pi}} \otimes V_n$. A multilinear function $f(x_1, \ldots, x_p)$ is a mapping from $\overset{p}{\underset{1}{\pi}} \otimes V_n$ to a vector space W, such that f is a linear in each variable, and

$$f(x_1, \ldots, x_p) = \operatorname{sgn} \pi \, f(x_{\pi(1)}, \ldots, x_{\pi(p)})$$

where π *is a permutation of* $1, \ldots, p$ *and* $\operatorname{sgn} \pi = +1$ *if* π *is even*, -1 *if* π *is odd*.

The first theorem is concerned with the existence of such a multilinear function.

Theorem 2.2. *For each* $p = 1, 2, \ldots, n$, *there exists a multilinear function defined on* $\prod_1^p \otimes V_n$ *to* $V_{\binom{n}{p}}$, *such that the smallest vector space containing range* f *is* $V_{\binom{n}{p}}$.

Proof. Let $x_1, \ldots, x_p$ be any vectors in V_n, where

$$\begin{aligned} x_1 &= (x_{11}, \ldots, x_{1n}) \\ x_p &= (x_{p1}, \ldots, x_{pn}) \end{aligned} \tag{1}$$

the right hand sides of the identities (1) may be thought of as forming a $p \times n$ matrix

$$\begin{pmatrix} x_{11}, & \ldots, & x_{1n} \\ x_{p1}, & \ldots, & x_{pn} \end{pmatrix} \tag{2}$$

Construct $\binom{n}{p}$ numbers as follows: Select any p columns from (2) and form the determinant of the $p \times p$ submatrix so obtained. Clearly this may be done in $\binom{n}{p}$ ways. By arranging these $\binom{n}{p}$ numbers in lexicographic order according to the manner of selection of columns from (2), we can construct an $\binom{n}{p}$ vector. Define this $\binom{n}{p}$ vector to be $f(x_1, \ldots, x_p)$. It is clear that f carries $\prod_1^p \otimes V_n$ into $V_{\binom{n}{p}}$. Because of the elementary properties of determinants, it is also clear that f satisfies the multilinear and alternating properties of Definition 2.1. Only the last statement of the theorem remains to be proved.

Let us consider the set of unit vectors $\{e_1, \ldots, e_n\}$ in V_n. Let $\{i_1, \ldots, i_p\}$ be an element of Q_{pn}, and consider the vector $f\ (e_{i1}, \ldots, e_{ip})$ in $V_{\binom{n}{p}}$. To find the coordinates of this vector we first have to form all pxp submatrices of the pxn matrix in which the jth row consists of zeros except for a one in the i_j column, $j = 1, \ldots, p$. The submatrix formed from the columns $i_1, \ldots, i_p$ will be a unit matrix and therefore have determinant one. Any other submatrix will contain a row of zeros, and so will have determinant zero. Thus $f(e_{i1}, \ldots, e_{ip})$ is an $\binom{n}{p}$ vector with a one in the $\{i_1, \ldots, i_p\}$ position and zeros elsewhere. Consequently, as $\{i_1, \ldots, i_p\}$ ranges over all the elements of Q_{pn}, we obtain the complete set of $\binom{n}{p}$ unit vectors in $V_{\binom{n}{p}}$. This completes the proof.

This theorem supplies the answer to the question of the existence of multilinear functions. Henceforth, the only multilinear function that we shall consider will be the one just defined. Having constructed the multilinear function, it is natural to enquire how the image of a given set of p vectors in V_n is determined by the images of sets of p vectors chosen from a basis set in V_n. The next theorem will provide an answer to this question. First, however, we indicate a use of the Q_{pn} notation introduced earlier:

<u>If</u> $\omega = \{i_1, \dots, i_p\} \in Q_{pn}$ <u>and</u> $x_1, \dots, x_n$ <u>are vectors in</u> V_n, <u>we define</u>

$$x_\omega = f(x_{i1}, \dots, x_{ip}) .$$

In particular, the proof of Theorem 2.2 shows that if $e_\omega = f(e_{i1}, \dots, e_{ip})$, then the set $\left\{ e_\omega \mid \omega \in Q_{pn} \right\}$ is the complete set of unit vectors in $V_{\binom{n}{p}}$.

<u>Theorem 2.3</u>. <u>Let</u> $y_1, \dots, y_p$ <u>belong to</u> V_n, <u>and let</u>

$$y_i = \sum_{j=1}^{n} y_{ij}\, x_j , \qquad (i = 1, \dots, p)$$

<u>where</u> $x_1, \dots, x_n$ <u>also belong to</u> V_n. <u>Then</u>

$$f(y_1, \dots, y_p) = \sum_{\omega \in Q_{pn}} c_\omega\, x_\omega$$

<u>where</u>, <u>if</u> $\omega = \{ i_1, \dots, i_p \} \in Q_{pn}$, <u>then</u>

$$c_\omega = \begin{vmatrix} y_{1i_1} & \cdots & y_{1i_p} \\ & & \\ y_{pi_1} & \cdots & y_{pi_p} \end{vmatrix}$$

<u>and</u>

$$x_\omega = f(x_{i_1}, \dots, x_{i_p}) .$$

Proof. We have $y_i = \sum_{j=1}^{n} y_{ij}\, x_j$, so, using the multilinear properties of f , we get

$$\begin{aligned} f(y_1, \ldots, y_p) &= f\left(\sum y_{1j}x_j, \sum y_{2j}x_j, \ldots, \sum y_{pj}x_j\right) \\ &= \sum_{j_1} y_{1_{j_1}}\, f\left(x_{j_1}, \sum y_{2j}x_j, \ldots, \sum y_{pj}x_j\right) \\ &= \sum_{j_1,\ldots,j_p} y_{1_{j_1}} y_{2_{j_2}}, \ldots, y_{p j_p}\, f(x_{j_1}, \ldots, x_{j_p}) \end{aligned}$$

Now, because of the alternating property of f , it is clear that if two indices in $f(x_{j_1}, \ldots, x_{j_p})$ are the same, then the corresponding value of f is zero. Again, because of the alternating property of f , we can restrict ourselves to summing over those terms in which $1 \leq j_1 < j_2 < \ldots < j_p \leq n$, as follows:

$$f(y_1, \ldots, y_p) = \sum_{1\leq j_1<j_2<\ldots<j_p\leq n} \left(\sum_{\pi} y_{1\pi(j_1)} \cdots y_{p\pi(j_p)} \operatorname{sgn} \pi\right) f(x_{j_i} \ldots x_{j_p})$$

where $\sum_{\pi}$ means summing over all sets of distinct permutation π of $j_1, \ldots, j_p$. Now clearly

$$\sum_{\pi} y_1\pi(y_1) \ldots y_p\pi(j_p) \operatorname{sgn} \pi = \det \begin{pmatrix} y_{1_{j_1}} & & y_{1_{j_p}} \\ y_{p_{j_1}} & & y_{p_{j_p}} \end{pmatrix}$$

so that

$$f(y_1, \ldots, y_p) = \sum_{1\leq j_1<\ldots<j_p\leq n} \det(y_{1_{j_1}} \ldots y_{p_{j_p}}) f(x_{j_1}, \ldots, x_{j_p})$$

or

$$f(y_1, \ldots, y_0) = \sum_{\omega} c_\omega\, x_\omega$$

as required.

Theorems 2.2 and 2.3 give the following interesting result

Theorem 2.4. Let $x_1, \ldots, x_n$ be a basis set for V_n. Then the set $\{x_\omega \mid \omega \in Q_{pn}\}$ is a basis set for $V_{\binom{n}{p}}$

Proof. Because $\{x_1, \ldots, x_n\}$ is a basis set for V_n, we know that any unit vector e_j $(j = 1, \ldots, n)$ can be written as a linear combination of the vectors $x_1, \ldots, x_n$. Let π, ω be elements of Q_{pn}. By Theorem 2.3 it follows that e_π is a linear combination of vectors x_ω, where ω runs over all the elements of Q_{pn}. This holds for every $\pi \in Q_{pn}$. Since the set $\{e_\pi \mid \pi \in Q_{pn}\}$ is the complete set of unit vectors in $V_{\binom{n}{p}}$, it is clear that any vector with $\binom{n}{p}$ coordinates can be written as a linear combination of the vectors of $\{x_\omega \mid \omega \in Q_{pn}\}$. Hence this set of $\binom{n}{p}$ vectors must form a basis for $V_{\binom{n}{p}}$.

Theorem 2.5. Let $x_1, \ldots, x_p$ be vectors in V_n. Then $f(x_1, \ldots, x_p) = 0$ if and only if $\{x_1, \ldots, x_p\}$ is a linearly dependent set.

Proof. Let $x_i = (x_{i_1}, \ldots, x_{i_p})$, $i = 1, \ldots, p$. Then $f(x_1, \ldots, x_p) = 0$ if and only if all $p \times p$ minors of the matrix

$$X = \begin{pmatrix} x_{11} & & x_{1n} \\ x_{p1} & & x_{pn} \end{pmatrix}$$

are zero. This happens if and only if rank $X \leq p - 1$, that is, if and only if $x_1, \ldots, x_p$ are linearly dependent.

Up to this point we have used $f(x_1, \ldots, x_p)$ to denote the multilinear function defined in Theorem 2.2: To conform with usual practice an alternative notation will now be introduced. We write $f(x_1, \ldots, x_p)$ as $x_1 \wedge \ldots \wedge x_p$; that is

$$x_1 \wedge \ldots \wedge x_p = f(x_1, \ldots, x_p) .$$

Henceforth these two notations will be used interchangeably. In this notation, if $\omega = \{ i_1, \ldots, i_p \} \in Q_{pn}$, then

$$x_\omega = x_{i_1} \wedge \ldots \wedge x_{i_p}$$

<u>Theorem 2.6</u>. <u>If</u> $x_1 \wedge \ldots \wedge x_p \, , \, y_1 \wedge \ldots \wedge y_p \in V_{\binom{n}{p}}$, <u>then</u> $(x_1 \wedge \ldots \wedge x_p \, , \, y_1 \wedge \ldots \wedge y_p) = \det((x_i, y_j)_{i,j=1,\ldots,p})$

<u>Remarks</u>: (i) $(\; , \;)$ refers to the usual unitary inner product.

(ii) The Theorem refers only to those vectors in $V_{\binom{n}{p}}$ that are of the form $x_1 \wedge \ldots \wedge x_p$ for suitable $x_1, \ldots, x_p$ in V_n .

<u>Proof</u>. The proof consists of the following calculation. Let

$$x_i = (x_{i1}, \ldots, x_{in})$$
$$y_j = (y_{j1}, \ldots, y_{jn}) \qquad i, j = 1, \ldots, p .$$

Then $\quad (x_i, y_j) = \sum_{r=1}^{n} x_{ir}\, \overline{y}_{jr}$, so

$$\det\big((x_i, y_j)\big) = \begin{vmatrix} \sum_{r=1}^{n} x_{1r}\overline{y}_{1r} & \sum_{r=1}^{n} x_{1r}\, \overline{y}_{2r} , & \ldots, & \sum_{r=1}^{n} x_{1r}\, \overline{y}_{pr} \\ \\ \sum_{r=1}^{n} x_{pr}\, \overline{y}_{1r} & \sum_{r=1}^{n} x_{pr}\, \overline{y}_{2r} , & \ldots, & \sum_{r=1}^{n} x_{pr}\, \overline{y}_{pr} \end{vmatrix}$$

$$= \sum_{r_1=1}^{n} \overline{y}_{1r_1} \begin{vmatrix} x_{1r_1} & \sum_{r=1}^{n} x_{1r}\, \overline{y}_{2r} , & \ldots, & \sum_{r=1}^{n} x_{1r}\, \overline{y}_{pr} \\ \\ x_{pr_1} & \sum_{r=1}^{n} x_{pr}\, \overline{y}_{pr}, & \ldots, & \sum_{r=1}^{n} x_{pr}\, \overline{y}_{pr} \end{vmatrix}$$

$$= \sum_{r_1, r_2=1}^{n} \overline{y}_{1r_1}\, \overline{y}_{2r_2} \begin{vmatrix} x_{1r_1} & x_{1r_2} , & \ldots, & \sum_{r=1}^{n} x_{1r}\, \overline{y}_{pr} \\ \\ x_{pr_1} & x_{pr_2} , & \ldots, & \sum_{r=1}^{n} x_{pr}\, \overline{y}_{pr} \end{vmatrix}$$

$$= \sum_{r_1, r_2, \ldots, r_p} \overline{y}_{1r}\, \overline{y}_{2r_2} \cdots \overline{y}_{pr_p} \begin{vmatrix} x_{1r_1} & x_{1r_p} \\ \\ x_{pr_1} & x_{pr_p} \end{vmatrix}$$

$$= \sum_{1 \leq r_1 < r_2 < \ldots < r_p \leq n} \left(\sum_{\pi} \overline{y}_{1\pi(r_1)} \cdots \overline{y}_{p\pi(r_p)} \operatorname{sgn} \pi \right) \begin{vmatrix} x_{1r_1} & x_{1r_p} \\ \\ x_{pr_1} & x_{pr_p} \end{vmatrix}$$

$$= \sum_{1\le r_1<\cdots<r_p\le n} \begin{vmatrix} x_{1r_1} & & x_{1r_p} \\ & & \\ x_{pr_1} & & x_{pr_p}\end{vmatrix} \overline{\begin{vmatrix} y_{1r_1} & & y_{1r_p} \\ & & \\ y_{pr_1} & & y_{pr_p}\end{vmatrix}}$$

This calculation completes the proof.

<u>Corollary 2.7</u>. <u>Let</u> $\{x_1, \ldots, x_n\}$ <u>be an orthonormal set of vectors in</u> V_n . <u>Then the set</u> $\{X_\omega \mid \omega \in Q_{pn}\}$ <u>is an orthonormal set in</u> $V_{\binom{n}{p}}$.

<u>Proof</u>. Consider $(x_{i_1} \wedge \cdots \wedge x_{i_p}, x_{j_1} \wedge \cdots \wedge x_{j_p}) = \det((x_{i_s} x_{j_t})_{s,t=1,..,p})$

Case (1). Suppose there is an integer m which lies in $\{i_1, \ldots, i_p\}$ but not in $\{j_1, \ldots, j_p\}$. Then the corresponding row of the matrix $((x_{i_s}, x_{j_t}))$ will consist entirely of zeros because of the orthogonality properties of the x's . Hence $(x_{i_1} \wedge \cdots \wedge x_{i_p} , x_{j_1} \wedge \cdots \wedge x_{j_p}) = 0$.

Case (2). If $i_1 = j_1, \ldots, i_p = j_p$, then the matrix $((x_{i_s}, x_{j_t}))$ is a $p \times p$ identity matrix and so has determinant 1. Hence $(x_{i_1} \wedge \cdots \wedge x_{i_p}, x_{i_1} \wedge \cdots \wedge x_{i_p}) = 1$.

III. <u>LINEAR TRANSFORMATIONS ON</u> $V_{\binom{n}{p}}$ <u>TO</u> $V_{\binom{n}{p}}$ <u>INDUCED BY LINEAR TRANSFORMATIONS ON</u> V_n <u>TO</u> V_n .

In this section, linear transformations on $V_{\binom{n}{p}}$ to $V_{\binom{n}{p}}$ that correspond to linear transformations on V_n to V_n are defined, and some of their properties are determined.

Definition 3.1. Let A be a linear transformation on V_n to V_n. Define the linear transformation $C_p(A)$ on $V_{\binom{n}{p}}$ to $V_{\binom{n}{p}}$ by defining its effect on the basis set $\{ e_\omega \mid \omega \in Q_{pn} \}$ to be

$$C_p(A) e_{i_1} \wedge \cdots \wedge e_{i_p} = (Ae_{i_1}) \wedge \cdots \wedge (Ae_{i_p})$$

Since a linear transformation of a vector space is uniquely determined by its effect on a set of basis vectors, Definition 3.1 defines one and only one linear transformation on $V_{\binom{n}{p}}$ to $V_{\binom{n}{p}}$.

Theorem 3.2. Let $y_1 \wedge \cdots \wedge y_p \in V_{\binom{n}{p}}$. Then

$$C_p(A) y_1 \wedge \cdots \wedge y_p = (Ay_1) \wedge \cdots \wedge (Ay_p)$$

Proof. Let A be defined by

$$Ae_i = \sum_{k=1}^{n} a_{ik}\, e_k\,, \quad i = 1, \ldots, n$$

and let

$$y_i = (y_{i1}, \ldots, y_{in}), \quad i = 1, \ldots, p\,.$$

Then, by Theorem 2.3

$$Ae_{j_1} \wedge \cdots \wedge Ae_{j_p} = \sum_{1 \le k_1 < \cdots < k_p \le n} \begin{vmatrix} a_{j_1 k_1} & \cdots & a_{j_1 k_p} \\ & & \\ a_{j_p k_1} & \cdots & a_{j_p k_p} \end{vmatrix} e_{k_1} \wedge \cdots \wedge e_{k_p}$$

and

$$y_1 \wedge \dots \wedge y_p = \sum_{1 \le j_1 < \dots < j_p \le n} \begin{vmatrix} y_{1j_1} & \cdots & y_{1j_p} \\ y_{pj_1} & \cdots & y_{pj_p} \end{vmatrix} e_{j_1} \wedge \dots \wedge e_{j_p}$$

so that, by the definition of $C_p(A)$ as a linear transformation

$$C_p(A) y_1 \wedge \dots \wedge y_p = \sum_{1 \le j_1 < \dots < j_p \le n} \begin{vmatrix} y_{1j_1} & \cdots & y_{1j_p} \\ y_{pj_1} & \cdots & y_{pj_p} \end{vmatrix} (Ae_{j_1}) \wedge \dots \wedge (Ae_{j_p})$$

$$= \sum_{1 \le j_1 < \dots < j_p \le n} \begin{vmatrix} y_{1j_1} & \cdots & y_{1j_p} \\ y_{pj_1} & \cdots & y_{pj_p} \end{vmatrix} \sum_{1 \le K_1 < \dots < K_p \le n} \begin{vmatrix} a_{j_1k_1} & \cdots & a_{j_1k_p} \\ a_{j_pk_1} & \cdots & a_{j_pk_p} \end{vmatrix} e_{k_1} \wedge \dots \wedge e_{k_p}$$

Consequently the $\{i_1, \dots, i_p\}$ coordinate of $C_p(A) y_1 \wedge \dots \wedge y_p$ is

$$\sum_{1 \le j_1 < \dots < j_p \le n} \begin{vmatrix} y_{1j_1} & \cdots & y_{1j_p} \\ y_{pj_1} & \cdots & y_{pj_p} \end{vmatrix} \begin{vmatrix} a_{j_1i_1} & \cdots & a_{j_1i_p} \\ a_{j_pi_1} & \cdots & a_{j_pi_p} \end{vmatrix}$$

To complete the proof, it is necessary to show that this is exactly the $\{i_1, \dots, i_p\}$ coordinate of $(Ay_1) \wedge \dots \wedge (Ay_p)$. We proceed to do this.

From

$$y_i = \sum_{j=1}^{n} y_{ij} e_j$$

we obtain

$$\begin{aligned} Ay_i &= \sum_{j=1}^{n} y_{ij}\, Ae_j \\ &= \sum_{j=1}^{n} y_{ij} \left(\sum_{k=1}^{n} a_{jk}\, e_k \right) \\ &= \sum_{k=1}^{n} \left(\sum_{j=1}^{n} y_{ij}\, a_{jk} \right) e_k \end{aligned}$$

so that the $\{i_1, \ldots, i_p\}$ coordinate of $(Ay_1) \wedge \ldots \wedge (Ay_p)$ is just

$$\begin{vmatrix} \sum_{j=1}^{n} y_{1j}\, a_{ji_1} & \cdots & \sum_{j=1}^{n} y_{1j}\, a_{ji_p} \\ \\ \sum_{j=1}^{n} y_{pj}\, a_{ji_1} & \cdots & \sum_{j=1}^{n} y_{pj}\, a_{ji_p} \end{vmatrix}$$

$$= \sum_{j_1, \ldots, j_p} a_{j_1 i_1} \cdots a_{j_p i_p} \begin{vmatrix} y_{1j_1} & & y_{1j_p} \\ \\ y_{pj_1} & & y_{pj_p} \end{vmatrix}$$

$$= \sum_{1 \leq j_1 < \ldots < j_p \leq n} \begin{vmatrix} y_{1j_1} & & y_{1j_p} \\ \\ y_{pj_1} & & y_{pj_p} \end{vmatrix} \begin{vmatrix} a_{j_1 i_1} & & a_{j_1 i_p} \\ \\ a_{j_p i_1} & & a_{j_p i_p} \end{vmatrix}$$

this completes the proof.

This theorem shows that there is nothing essential in using the basis vectors $e_1, \ldots, e_n$ to define $Cp(A)$; in fact, if $x_1, \ldots, x_n$ is another basis for V_n, we could give an equivalent definition of $Cp(A)$ as

$$Cp(A)\, x_{i_1} \wedge \cdots \wedge x_{i_p} = (Ax_{i_1}) \wedge \cdots \wedge (Ax_{i_p})$$

Before proceeding to the next theorem, it is necessary to give a word of explanation. Suppose $A = (a_{ij})$ is any $n \times n$ matrix. From A, we can construct a new matrix B as follows. Given any set of p rows of A, we fix our attention on the $p \times n$ submatrix of A defined by the given p rows, and use this $p \times n$ submatrix to construct a vector with $\binom{n}{p}$ coordinates, as in Theorem 2.2. That is, we write the $p \times p$ minors of this $p \times n$ submatrix in a row vector form, in lexicographic order according to the manner of selection of columns of A. Thus, for a given set of p rows of A we obtain a vector with $\binom{n}{p}$ coordinates. If this construction is carried out for each possible selection of p rows of A, we will obtain a set of $\binom{n}{p}$ row vectors, each with $\binom{n}{p}$ coordinates. If these row vectors are used as the entries of a column vector and ordered in lexicographic fashion according to the manner of selection of <u>rows</u> of A, we obtain the matrix B. B is said to consist of the $p \times p$ minors of A, arranged in doubly lexicographic order. Note that B is an $\binom{n}{p} \times \binom{n}{p}$ matrix.

Theorem 3.3. If $x_1, \ldots, x_n$ is a basis for V_n, then the representation of $C_p(A)$ relative to the basis x_ω ($\omega \in Q_{pn}$) is a matrix whose entries are the $p \times p$ minors of the representation of A relative to $x_1, \ldots, x_n$, arranged in doubly lexicographic order.

Proof. The method of proof is a now familiar type of calculation. Let

$$Ax_i = \sum_{j=1}^{n} a_{ji}\, x_j \;,\; i = 1, \ldots, n$$

then

$$C_p(A)x_{i_1} \wedge \cdots \wedge x_{i_p} = \Big(\sum_{j=1}^{n} a_{ji_1}\, x_j\Big) \wedge \cdots \wedge \Big(\sum_{j=1}^{n} a_{ji_p}\, x_j\Big)$$

$$= \sum_{j_1, \ldots, j_p} a_{j_1 i_1} \cdots a_{j_p i_p}\, x_{j_1} \wedge \cdots \wedge x_{j_p}$$

$$= \sum_{1 \le j_1 < \cdots < j_p \le n} \Big(\sum_{\pi} a_{\pi(j_1) i_1} \cdots a_{\pi(j_p) i_p}\, \operatorname{sgn} \pi\Big) x_{j_1} \wedge \cdots \wedge x_{j_p}$$

$$= \sum_{1 \le j_1 < \cdots < j_p \le n} \begin{vmatrix} a_{j_1 i_1} & \cdots & a_{j_1 i_p} \\ \\ a_{j_p i_1} & \cdots & a_{j_p i_p} \end{vmatrix} \quad x_{j_1} \wedge \cdots \wedge x_{j_p}$$

This completes the proof.

Because of Theorem 3.2, for any vector $x_{i_1} \wedge \cdots \wedge x_{i_p}$ and any two linear transformations A, B we have

$$\begin{aligned} C_p(AB)x_{i_1} \wedge \dots \wedge x_{i_p} &= (ABx_{i_1}) \wedge \dots \wedge (ABx_{i_p}) \\ &= C_p(A)\ (Bx_{i_1}) \wedge \dots \wedge (Bx_{i_p}) \\ &= C_p(A)\ C_p(B) x_{i_1} \wedge \dots \wedge x_{i_p} \end{aligned}$$

Since this calculation is valid for all basis vectors $x_{i_1} \wedge \dots \wedge x_{i_p}$, it follows that

$$C_p(AB) = C_p(A)\ C_p(B) .$$

Clearly this also holds for the matrix representations of A, B relative to a given basis. Consequently we have proved the following theorem.

__Theorem 3.4.__ __If__ A , B __are linear transformations on__ V_n __to__ V_n , __or the matrix representations of linear transformations relative to a given basis, then__

$$C_p(AB) = C_p(A)\ C_p(B)$$

__Corollary.__ __If__ A^{-1} __exists, then__ $C_p(A)^{-1}$ __exists, and__ $C_p(A)^{-1} = C_p(A^{-1})$.

__Proof.__ From $AA^{-1} = A^{-1}A = I$ it follows that

$$I_{\binom{n}{p}} = C_p(A)\ C_p(A^{-1}) = C_p(A^{-1})\ C_p(A)$$

so that
$$C_p(A^{-1}) = C_p(A)^{-1}$$

The next theorem gives a preliminary result that will be needed in order to determine the eigenvalues of $C_p(A)$.

<u>Theorem 3.5. Let</u> A <u>possess</u> n <u>linearly independant eigenvectors</u> $x_1, \dots, x_n$ <u>belonging to the eigenvalues</u> $\lambda_1, \dots, \lambda_n$ <u>respectively. Then the set of eigenvectors of</u> $C_p(A)$ <u>is the set</u> $\{X\omega \mid \omega \in Q_{pn}\}$ <u>and the set of eigenvalues is the set</u>

$$\left\{ \lambda_\omega = \lambda_{i_1} \cdots \lambda_{i_p} \;\middle|\; \omega = \{i_1, \dots, i_p\} \in Q_{pn} \right\}$$

<u>Proof</u>. We have $Ax_i = \lambda_i x_i$, so

$$\begin{aligned} C_p(A)x_{i_1} \wedge \dots \wedge x_{i_p} &= (Ax_{i_1}) \wedge \dots \wedge (Ax_{i_p}) \\ &= \lambda_{i_1} \cdots \lambda_{i_p} \, x_{i_1} \wedge \dots \wedge x_{i_p} \end{aligned}$$

Since the set $\{x_{i_1} \wedge \dots \wedge x_{i_p}\}$ is a set of $\binom{n}{p}$ linearly independant eigenvectors, $C_p(A)$ can have no other eigenvalues than those stated.

<u>Theorem 3.6. Let</u> A <u>be any linear transformation on</u> V_n <u>to</u> V_n, <u>with eigenvalues</u> $\lambda_1, \dots, \lambda_n$. <u>Then the set of eigenvalues of</u> $C_p(A)$ <u>is the set</u>

$$\left\{ \lambda_\omega = \lambda_{i_1} \cdots \lambda_{i_p} \;\middle|\; \omega = \{i_1, \dots, i_p\} \in Q_{pn} \right\}$$

<u>Proof</u>. Because of the importance of this theorem, two proofs will be given.

<u>First Proof</u>.

Case (1). A has distinct characteristic roots. In this case A possesses a set of n linearly independent eigenvectors, and so the result follows from the previous theorem.

Case (2). Given A , there exists a matrix B whose elements differ arbitrarily little from the elements of A and which has distinct eigenvalues. The theorem is then true for B , by Case (1). Since the compound operator C_p is continuous, and the eigenvalues of a matrix are continuous functions of the entries, if we choose a sequence of matrices B approaching A , and for which the theorem is true, it follows that the theorem is true for A .

Second Proof. This proof depends on the following lemma.

Lemma. *If* T *is a triangular matrix, then* $C_p(T)$ *is triangular*.

Proof. Let us suppose that the lower triangle of T is zero, and attempt to show that the lower triangle of $C_p(T)$ is also zero. To do this, consider an entry in the $(i_1 \ldots i_p , j_1 \ldots j_p)$ position of $C_p(T)$. If this entry is to be below the main diagonal we must have, for some $k < p$,

$$i_1 = j_1, i_2 = j_2, \ldots, i_k = j_k, i_{k+1} > j_{k+1} .$$

This implies

$$i_p > i_{p-1} > \cdots > i_{k+1} > j_{k+1}$$

so that the $(i_1 \ldots i_p, j_1 \ldots j_p)$ entry of $C_p(T)$ is the determinant of the following minor of A , exhibited in a block form:

$$\begin{array}{c|cccc|c|ccc}
 & j_1 & & & j_k & j_{k+1} & j_{k+2} & & j_p \\
\hline
i_1 & t_{i_1 i_1} & & & & & & & \\
 & & \ddots & & & & & & \\
 & & 0 & \ddots & & & & & \\
i_k & & & & t_{i_k i_k} & & & & \\
\hline
i_{k+1} & 0 & \cdots & \cdots & 0 & 0 & & & \\
\hline
i_{k+2} & & & & & 0 & & & \\
 & & 0 & & & \vdots & & & \\
i_p & & & & & 0 & & &
\end{array}$$

Here the numbers down the side and across the top indicate the rows and columns of A used in constructing this minor of A. Because of the form of this matrix, its determinant is zero. This completes the proof of the lemma.

To prove the theorem we use the fact that given A, there always exists U such that $U^{-1} AU = T$, a triangular matrix. Hence

$$C_p(U)^{-1} C_p(A) C_p(U) = C_p(T)$$

where $C_p(T)$ is a triangular matrix whose diagonal entries are precisely the numbers given in the statement of the theorem. The second proof of the theorem is now complete.

Theorem 3.6 enables a simple proof to be given of the well known Sylvester-Franke Theorem.

<u>Theorem 3.7.</u> $\quad |C_p(A)| = |A|^{\binom{n-1}{p-1}}$

<u>Proof</u>. $|C_p(A)|$ is the product of the eigenvalues of $C_p(A)$. In this product each eigenvalue of A occurs exactly $\binom{n-1}{p-1}$ times.

Q.E.D.

By theorem 2.4 we know that if $\{x_1, \ldots, x_n\}$ is a basis set for V_n, then $\{x_\omega \mid \omega \in Q_{pn}\}$ is a basis set for $V_{\binom{n}{p}}$. Using Theorem 3.7, we can complete theorem 2.4 as follows.

<u>Corollary</u>. <u>Let</u> $\{x_1, \ldots, x_n\}$ <u>be a set of vectors in</u> V_n. <u>Then</u> $\{x_\omega \mid \omega \in Q_{pn}\}$ <u>is a basis set for</u> $V_{\binom{n}{p}}$ <u>if and only if</u> $\{x_1, \ldots, x_n\}$ <u>is a basis set for</u> V_n.

<u>Proof.</u> Let A be the matrix whose ith row is the vector x_i. Then $C_p(A)$ is the $\binom{n}{p} \times \binom{n}{p}$ matrix whose successive rows consist of the vectors $x_{i_1} \wedge \cdots \wedge x_{i_p}$, arranged in lexicographic order. Since $|C_p(A)| = |A|^{\binom{n-1}{p-1}}$ it follows that $|C_p(A)| \neq 0$ if and only if $|A| \neq 0$.

The next, and last, theorem in this section gives a resume' of properties of $C_p(A)$ in terms of assumed properties of A.

<u>Theorem 3.8.</u> <u>The following relations are valid.</u>

(1) $C_p(A)^{-1} = C_p(A^{-1})$

(2) $C_p(A^*) = C_p(A)^*$

(3) *If* A *is normal,* $C_p(A)$ *is normal*

(4) *If* A *is Hermitian,* $C_p(A)$ *is Hemitian*

(5) *If* A *is positive definite (or non-negative) Hermitian then* $C_p(A)$ *is positive definite (or non-negative) Hermitian.*

Proof.

(1) This has already been proved

(2) We have, using Theorems 2.7 and 3.2

$$\begin{aligned}
&(C_p(A^*)x_1 \wedge \ldots \wedge x_p, y_1 \wedge \ldots \wedge y_p) \\
&= ((A^*x_1) \wedge \ldots \wedge (A^*x_p), y_1 \wedge \ldots \wedge y_p) \\
&= \det ((A^*x_i, y_j)_{i,j=1,\ldots,p}) \\
&= \det ((x_i, Ay_j)_{i,j=1,\ldots,p}) \\
&= (x_1 \wedge \ldots \wedge x_p, (Ay_1) \wedge \ldots \wedge (Ay_p)) \\
&= (x_1 \wedge \ldots \wedge x_p, C_p(A)\, y_1 \wedge \ldots \wedge y_p)
\end{aligned}$$

In particular it follows that

$$(C_p(A^*)e_\pi, e_\omega) = (e_\pi, C_p(A)e_\omega)$$

for any $\pi, \omega \in Q_{pn}$. Hence

$$C_p(A)^* = C_p(A^*) .$$

(3) If $AA^* = A^*A$, then

$$C_p(A)\, C_p(A^*) = C_p(A^*)\, C_p(A)$$

so that

$$C_p(A)\, C_p(A)^* = C_p(A)^*\, C_p(A)$$

(4) This follows immediately from (2).

(5) Using Theorem 3.6, it follows that if the eigenvalues of A are positive (or non-negative), then the eigenvalues of $C_p(A)$ are positive (or non-negative).

IV. Two Important Results.

In this section two important theorems will be given. Both theorems will be needed later, but only the first will be proved. Because of their dissimilar nature, this section is divided into two subsections.

1. An Extremum Property of Eigenvalues.

The chief result is Theorem 4.2.

Definition 4.1. *Let $1 \leq p \leq k$ be positive integers. The elementary symmetric function of degree p on the k letters $a_1, \ldots, a_k$ is the coefficient of t^{k-p} in*

$$\prod_{i=1}^{k} (t + a_i)$$

and is written as

$$E_p(a_1, \ldots, a_k)$$

or, as will later be seen to be convenient,

$$E_p(a_i \; ; \; i \in \{1, \ldots, k\})$$

Theorem 4.2. *Let* $1 \leq p \leq k \leq n$ *and let* A *be an n-square positive definite Hermitian transformation with eigenvalues* $0 < \alpha_1 \leq \dots \leq \alpha_n$ *. Then*

$$\max \sum_{\omega \in Q_{pk}} (C_p(A)x_\omega, x_\omega) = E_p(\alpha_n, \dots, \alpha_{n-k+1})$$

$$\min \sum_{\omega \in Q_{pk}} (C_p(A)x_\omega, x_\omega) = E_p(\alpha_1, \dots, \alpha_k)$$

where $x_\omega = x_{i_1} \wedge \dots \wedge x_{i_p}$ *if* $\omega = \{i_1, \dots, i_p\} \in Q_{pk}$, *and the max and min are taken over all sets of* k *orthonormal vectors* $\{x_1, \dots, x_k\}$ *in* V_n .

Proof. The proof is in several steps. For notational simplicity we shall let

$$g(x_1, \dots, x_k) = \sum_{\omega \in Q_{pk}} (C_p(A)x_\omega, x_\omega) \quad .$$

(i) First it is clear that a set of maximizing (minimizing) orthonormal vectors for g exists. This is easily seen using a standard continuity argument.

(ii) If $k = n$, then

$$\sum_{\omega \in Q_{pn}} (C_p(A)x_\omega, x_\omega) = \text{trace } C_p(A) = E_p(\alpha_1, \dots, \alpha_n)$$

since the x_ω $(\omega \in Q_{pn})$ are an orthonormal basis in $V_{\binom{n}{p}}$. This completes the proof for the special case $k = n$.

(iii) Let $\{y_1, \dots, y_k\}$ be a maximizing (minimizing) set for $g(x_1, \dots, x_k)$. Consider the linear subspace of V_n of dimension k

$$L = L(y_1, \dots, y_k)$$

spanned by $y_1, \dots, y_k$. Let P be the perpendicular projection onto this space. Consider $PA : L \to L$. Because of the properties of perpendicular projections, $P = P^2 = P^*$ and $Px = x$ for any $x \in L$. If $x, y \in L$, then

$$\begin{aligned}(PAx, y) &= (Ax, Py) = (Ax, y) = (x, Ay) \\ &= (Px, Ay) = (x, PAy)\end{aligned}$$

so that PA is a Hermitian transformation on L to L. Let $u_1, \dots, u_k$ be orthonormal eigenvectors of PA (in L). Then:

$$\begin{aligned}\sum_{\omega \in Q_{pk}} (C_p(A)y_\omega, y_\omega) &= \sum (C_p(A)y_{i_1} \wedge \cdots \wedge y_{i_p}, y_{i_1} \wedge \cdots \wedge y_{i_p}) \\ &= \sum \det((Ay_{i_s}, y_{i_t})_{s,t=1,\dots,p}) \\ &= \sum \det((PAy_{i_s}, y_{i_t})) \\ &= \sum (C_p(PA)y_\omega, y_\omega) \\ &= \text{trace } C_p(PA) \\ &= \sum (C_p(PA)u_\omega, u_\omega) \\ &= \sum \det((PA)u_{i_s}, u_{i_t})) \\ &= \sum \det((Au_{i_s}, u_{i_t})) \\ &= \sum (C_p(A)u_\omega, u_\omega)\end{aligned}$$

(iv) We claim that L is an invariant space of A; that is $A\,L \leq L$. The proof is by contradiction. Since $u_1, \ldots, u_k$ is a basis for L, let us assume that $A\,u_1 \notin L$. Then there exists v belonging to the orthogonal complement of L such that

$$\rho = (Au_1, v) \neq 0$$

Let

$$u_1' = \frac{u_1 - t\,\rho\,v}{\sqrt{1 + t^2\,|\rho|^2}}$$

$$u_j' = u_j \,, \qquad (j = 2, \ldots, k)$$

where t is a real number. It is easy to verify that $\{u_1', \ldots, u_k'\}$ is an orthonormal set. Since $g(u_1, \ldots, u_k)$ is an extremum it follows that $\frac{d}{dt}\,g(u_1', \ldots, u_k')$ must be zero at $t = 0$. Now for $t = 0$

$$\frac{d}{dt}\left(C_p(A)\,\frac{u_1 - t\,\rho\,v}{\sqrt{1 + t^2\,|\rho|^2}} \wedge u_{i_2} \wedge \cdots \wedge u_{i_p},\ \frac{u_1 - t\,\rho\,v}{\sqrt{1 + t^2\,|\rho|^2}} \wedge u_{i_2} \wedge \cdots \wedge u_{i_p}\right)$$

$$= -\rho\,(C_p(A)v \wedge u_{i_2} \wedge \cdots \wedge u_{i_p},\ u_1 \wedge u_{i_2} \wedge \cdots \wedge u_{i_p})$$

$$\quad -\bar{\rho}\,(C_p(A)u_1 \wedge u_{i_2} \wedge \cdots \wedge u_{i_p},\ v \wedge u_{i_2} \wedge \cdots \wedge u_{i_p})$$

$$= -\rho \begin{vmatrix} (Av, u_1) & & & \\ & (Au_{i_2}, u_{i_2}) & & \\ & 0 & \ddots & \\ & & & (Au_{i_p}, u_{i_p}) \end{vmatrix}$$

$$
- \bar{\rho} \begin{vmatrix} (Au_1, v) & & & & \bigcirc \\ & (Au_{i_2}, u_{i_2}) & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & \cdot \; (Au_{i_p}, u_{i_p}) \end{vmatrix}
$$

$$
= - 2 \; |\rho|^2 \; \prod_{\alpha=2}^{p} \; (Au_{i_\alpha}, u_{i_\alpha})
$$

Here we have used the fact that, if $s, t \geq 2$ and $s \neq t$, then

$$
(Au_{i_s}, u_{i_t}) = (PAu_{i_s}, u_{i_t}) = 0
$$

since the u's are orthonormal eigenvectors of PA. Furthermore, $\prod_{\alpha=2}^{p} (Au_{i_\alpha}, u_{i_\alpha}) > 0$ (because A is positive definite), and $\rho \neq 0$, so it clearly follows that $\frac{d}{dt} g(u_1', \ldots, u_k')$ is not zero at $t = 0$. This is a contradiction unless $Au_1 \in L$. Hence $L(y_1, \ldots, y_k)$ is an invariant space of A.

(v). It is now easy to complete the proof. If $\{y_1, \ldots, y_k\}$ is an extremizing set of orthonormal vectors, then $L(y_1, \ldots, y_k)$ is an invariant space of A; let $B = A|L$, the restriction of A to L. Then B is a positive definite Hermitian transformation on a k dimensional space. The eigenvalues of B are some k of the eigenvalues of A. Hence by (ii)

$$\sum_{\omega}(C_p(A)y_\omega, y_\omega) = \sum_{\omega}(C_p(B)y_\omega, y_\omega)$$

$$= E_p(\alpha_{i_1}, \ldots, \alpha_{i_k})$$

for certain eigenvalues $\alpha_{i_1}, \ldots, \alpha_{i_k}$ of A. Thus

$$E_p(\alpha_1, \ldots, \alpha_k) \leq g(x_1, \ldots, x_k) \leq E_p(\alpha_n, \ldots, \alpha_{n-k+1})$$

for any orthonormal vectors $x_1, \ldots, x_k$. Since either equality can be obtained by choosing $x_1, \ldots, x_k$ to be suitable orthonormal eigenvectors of A, the proof is complete.

Two interesting corollaries can be given.

Corollary 1.

$$\min E_p((Ax_1, x_1), \ldots, (Ax_k, x_k))$$

$$= E_p(\alpha_1, \ldots, \alpha_k)$$

where the minimum is taken over all sets of k orthonormal vectors in V_n.

Proof. Because $((Ax_{i_s}, x_{i_t})_{s,t=1, \ldots, p})$ is a positive definite Hermitian matrix, it follows from the Hadamard determinant inequality that

$$(C_p(A)x_\omega, x_\omega) = \det ((Ax_{i_s}, x_{i_t})_{s,t=1,\ldots,p})$$

$$\leq \prod_{s=1}^{p} (Ax_{i_s}, x_{i_s}) ;$$

so that

$$E_p((Ax_1, x_1), \ldots, (Ax_k, x_k)) \geq E_p(\alpha_1, \ldots, \alpha_k)$$

Equality can be attained by use of suitable orthonormal eigenvectors of A, so that the proof is complete.

Corollary 2.

$$\max E_p((Ax_1, x_1), \ldots, (Ax_k, x_k)) = \binom{k}{p} \left(\frac{\alpha_n + \cdots + \alpha_{n-k+1}}{k} \right)^p$$

where the maximum is taken over all sets of k *orthonormal vectors in* V_n.

Proof. Theorem 4.2, in the case that $p = 1$, states

$$\max \sum_{i=1}^{k} (Ax_i, x_i) = E_1(\alpha_n, \ldots, \alpha_{n-k+1})$$

because $C_1(A) = A$. Since for any positive members $a_1, \ldots, a_k$ (9 , p.49)

$$E_p(a_1, \ldots, a_k) \leq \binom{k}{p} \left(\frac{E_1(a_1, \ldots, a_k)}{k} \right)^p$$

it follows that

$$E_p((Ax_1, x_1), \ldots, (Ax_k, x_k)) \leq \binom{k}{p} \left(\frac{(Ax_1, x_1) + \cdots + (Ax_k, x_k)}{k} \right)^p \leq \binom{k}{p} \left(\frac{\alpha_n + \cdots + \alpha_{n-k+1}}{k} \right)^p$$

To complete the proof we require the following lemma.

<u>Lemma 4.3</u>. <u>Let</u> B <u>be any transformation on</u> V_n to V_n. <u>Then there exist</u> n <u>orthonormal vectors</u> $x_1, \ldots, x_n$ <u>such that</u>

$$(Bx_i, x_i) = \frac{\text{trace } B}{n} \qquad i = 1, \ldots, n .$$

<u>Proof</u>. The set $\left\{ (Bx, x) \mid (x, x) = 1 ; x \in V_n \right\}$ is known to be a closed convex set of complex numbers containing the eigenvalues of B, so there exists one vector x_1 with $(Bx_1, x_1) = \frac{\text{trace } B}{n}$. The proof is by induction. Assume there exist k orthonormal vectors $x_1, \ldots, x_k$ with $(Bx_i, x_i) = \frac{\text{tr } B}{n}$, $i = 1, \ldots, k$. Let U be a unitary n-square matrix in which the first k columns are the vectors $x_1, \ldots, x_k$. Then

$$U^* BU = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$$

where B_{11} is a $k \times k$ matrix with diagonal entries $\frac{\text{tr } B}{n}$, and B_{22} is an $(n-k) \times (n-k)$ matrix. We have

$$k \frac{\text{tr } B}{n} + \text{tr } B_{22} = \text{tr } B$$

so that

$$\frac{\text{tr } B_{22}}{n-k} = \frac{\text{tr } B}{n}$$

Now there exists an $(n-k) \times (n-k)$ unitary matrix V such that the $(1, 1)$ entry of $V^* B_{22} V$ is $\frac{\text{tr } B_{22}}{n-k}$. Let $W = I_k \oplus V$. Then $(UW)^* B(UW)$ has $\frac{\text{tr } B}{n}$ as its first $k+1$

diagonal entries, so that the first $k+1$ columns of UW are orthonormal vectors that satisfy $(Bx, x) = \frac{\operatorname{tr} B}{n}$. Hence the induction step is complete and the Lemma is proved.

To complete the proof of Corollary 2, let $x_1, \ldots, x_k$ be orthonormal eigenvectors of A corresponding to $\alpha_n, \ldots, \alpha_{n-k+1}$ respectively. Let B be the restriction of A to $L(x_1, \ldots, x_k)$. Then there exist orthonormal vectors $y_1, \ldots, y_k$ such that

$$(Ay_i, y_i) = (By_i, y_i) = \frac{\alpha_n + \cdots + \alpha_{n-k+1}}{k}$$

Hence

$$E_p((Ay_1, y_1), \ldots, (Ay_k, y_k)) = \binom{k}{p} \left(\frac{\alpha_n + \cdots + \alpha_{n-k+1}}{k}\right)^p$$

By use of a standard continuity argument we may establish

<u>Theorem 4.3</u>. <u>Theorem 4.2 and its two Corollaries remain true if A is a non-negative Hermitian transformation</u>.

Theorems 4.2 and 4.3 and the Corollaries are due to Marcus and McGregor (10). It is interesting to note that proofs of the Corollaries were given before Theorem 4.2 itself was proved.

<u>Definition 4.4</u>. (i) <u>A function of one real variable</u> $f(x)$ <u>defined on $[a, b]$ is said to be convex if for any</u> $x_1, x_2 \in [a, b]$

$$f\left(\frac{x_1 + x_2}{2}\right) \leq \frac{f(x_1) + f(x_2)}{2}$$

(ii) <u>A function</u> f <u>is concave if</u> $-f$ <u>is convex</u>

<u>Lemma 4.5</u>. <u>If</u> f <u>is continuous and convex and if</u> $\sum_{i=1}^{n} \sigma_i = 1$, $\sigma_i \geq 0$, <u>then</u>

$$f\left(\sum_{i=1}^{n} \sigma_i x_i \right) \leq \sum_{i=1}^{n} \sigma_i f(x_i)$$

<u>Proof</u>. This is well known in the theory of convex functions.

If H is any Hermitian matrix with eigenvalues $h_1 \geq \ldots \geq h_n$ corresponding to orthonormal eigenvectors $u_1, \ldots, u_n$, then any vector x has a unique representation in the form

$$x = \sum_{i=1}^{h} (x, u_i) u_i$$

and

$$Hx = \sum_{i=1}^{h} (x, u_i) h_i u_i$$

This remark motivates

<u>Definition 4.6</u>. <u>For any function</u> f <u>defined on the spectrum of the Hermitian matrix</u> H , <u>define</u>

$$f(H)x = \sum_{i=1}^{h} (x, u_i) f(h_i) u_i .$$

<u>Theorem 4.7</u>. <u>If</u> f <u>is a real continuous function of one variable defined on the spectrum of the Hermitian matrix</u> H,

then $f(H)$ *is Hermitian. If* f *is convex (concave) then for any unit vector* x *:*

$$f((Hx, x)) \leq (\geq) (f(H)x, x)$$

Proof. The Hermitian nature of $f(H)$ easily follows from Definition 4.6. If f is convex, then using Lemma 4.4

$$\begin{aligned} f((Hx, x)) &= f\Big(\sum_{i=1}^{h} |(x, u_i)|^2 h_i\Big) \\ &\leq \sum_{i=1}^{h} |(x, u_i)|^2 f(h_i) \\ &= (f(H)x, x) . \end{aligned}$$

Theorem 4.8. If f *is defined on the spectrum of the Hermitian matrix* H *and* $f(xy) = f(x) f(y)$ *then*

$$f(C_p(H)) = C_p(f(H))$$

Proof. By definition

$$\begin{aligned} C_p(f(H))x_1 \wedge \dots \wedge x_p &= f(H)x_1 \wedge \dots \wedge f(H)x_p \\ &= \sum (x_1, u_\alpha) f(h_\alpha) u_\alpha \wedge \dots \wedge \sum (x_p, u_\alpha) f(h_\alpha) u_\alpha \\ &= \sum \det((x_s, u_{i_s})_{s=1,\dots,p}) f(h_{i_1}) \dots f(h_{i_p}) u_{i_1} \wedge \dots \wedge u_{i_p} \\ &= \sum_{\omega} \big((x_1 \wedge \dots \wedge x_p), u_\omega\big) f(h_\omega) u_\omega \\ &= f(C_p(H))x_1 \wedge \dots \wedge x_p \end{aligned}$$

Here we have used Theorem 2.3.

<u>Theorem 4.9</u> <u>Assume</u> <u>the</u> <u>hypotheses</u> <u>of</u> <u>Theorem</u> 4.3. <u>Let</u> $f(x) = x^s$. <u>Then</u>

(a) <u>If</u> $s \geq 1$

$$\max \sum_{\omega \in Q_{pk}} f((C_p(A)x_\omega, x_\omega)) = E_p(f(\alpha_n), \ldots, f(\alpha_{n-k+1}))$$

(b) <u>If</u> $0 \leq s \leq 1$

$$\min \sum_{\omega \in Q_{pk}} f((C_p(A)x_\omega, x_\omega)) = E_p(f(\alpha_1), \ldots, f(\alpha_k))$$

<u>Proof.</u> (a) For $s \geq 1$, $f(x) = x^s$, is a convex function and $f(xy) = f(x)\, f(y)$. Hence, by Theorems 4.7 and 4.8

$$f(C_p(A)) = C_p(f(A))$$
$$f((C_p(A)x_\omega, x_\omega)) \leq (f(C_p(A))x_\omega, x_\omega) .$$

Therefore

$$\sum_{\omega \in Q_{pk}} f((C_p(A)x_\omega, x_\omega)) \leq \sum_{\omega} (C_p(f(A))x_\omega, x_\omega)$$

$$\leq E_p(f(h_n), \ldots, f(h_{n-k+1})) .$$

Equality is attained if $x_1, \ldots, x_k$ are orthonormal eigenvectors corresponding to $\alpha_n, \ldots, \alpha_{n-k+1}$. This proves (a) . (b) is proved similarly.

<u>Corollary.</u> <u>Consider</u> <u>any</u> <u>orthonormal</u> <u>set</u> $\{x_1, \ldots, x_k\}$.

(1) <u>If</u> $0 \leq s \leq 1$,

$$E_p\ ((Ax_1, x_1)^s, \ldots, (Ax_k, x_k)^s)$$

$$\geqq E_p(\alpha_1^s, \ldots, \alpha_k^s)$$

(2) If $s \geq 1$,

$$E_p((Ax_1, x_1)^s, \ldots, (Ax_k, x_k)^s)$$
$$\leq \binom{k}{p} \left(\frac{\alpha_n^s + \cdots + \alpha_{n-k+1}^s}{k} \right)^p$$

Proof. The proofs are along the lines of the proofs of the Corollaries of Theorem 4.2. In (1) above, equality is attained.

2. A New Inequality For Positive Numbers.

Recently Marcus (11) and Bohnenblust (1) have proved the following theorem:

Theorem 4.10. *Let* $1 \leq r \leq k$ *and let* $a_i, b_i, i = 1, \ldots, k$ *be positive numbers then*

$$E_r^{1/r}(a_1+b_1, \ldots, a_k+b_k) \geq E_r^{1/r}(a_1, \ldots, a_k) + E_r^{1/r}(b_1, \ldots, b_k)$$

with equality if and only if the sets (a_i) *and* (b_i) *are proportional.*

The proof will not be given. We will be content to give the following consequence:

Corollary. *Let* A, B *be non-negative n-square Hermitian matrices. For any n-square matrix* M *let* $P_r(M)$ *denote the coefficient of* t^r *in*

$$\det(tI - M) ,$$

$r = 1, \ldots, n$. *Then*

$$|P_r(A+B)|^{1/r} \geq |P_r(A)|^{1/r} + |P_r(B)|^{1/r} .$$

Proof. Let $\alpha_1, \ldots, \alpha_n$; $\beta_1, \ldots, \beta_n$; and $\lambda_1, \ldots, \lambda_n$ denote the eigenvalues of A , B , and $A+B$ respectively. Then

$$|P_r(A)| = E_r(\alpha_1, \ldots, \alpha_n)$$

$$|P_r(B)| = E_r(\beta_1, \ldots, \beta_n)$$

$$|P_r(A+B)| = E_r(\lambda_1, \ldots, \lambda_n) .$$

If $x_1, \ldots, x_n$ are orthonormal eigenvectors of $A+B$, then

$$\begin{aligned}
|P_r(A+B)|^{1/r} &= E_r^{1/r}\,((A+B)x_1,x_1), \ldots, ((A+B)x_n,x_n)) \\
&= E_r^{1/r}\Big((Ax_1,x_1)+(Bx_1,x_1), \ldots, (Ax_n,x_n)+(Bx_n,x_n)\Big) \\
&\geq E_r^{1/r}\,((Ax_1, x_1), \ldots, (Ax_n, x_n)) \\
&\quad + E_r^{1/r}\,((Bx_1, x_1), \ldots, (Bx_n, x_n)) \\
&\geq E_r^{1/r}\,(\alpha_1, \ldots, \alpha_n) + E_r^{1/r}\,(\beta_1, \ldots, \beta_n) \\
&= |P_r(A)|^{1/r} + |P_r(B)|^{1/r}
\end{aligned}$$

using Theorem 4.2., Corollary 1.

This is a generalization of a classical inequality of Minkowski (13):

$$(\det(A+B))^{\frac{1}{n}} \geq (\det A)^{\frac{1}{n}} + (\det B)^{\frac{1}{n}} .$$

V. Applications.

In this section we shall give extensions of two previously known theorems. We state the following Lemma.

Lemma 5.7. (i) *If* $\sigma \geq 0$, $\delta \geq 0$, $\sigma + \delta = 1$, $a \geq 0$, $b \geq 0$, *then*

$$a^{\sigma} b^{\delta} \leq \sigma a + \delta b$$

(ii) *If* $s \geq 1$, $a^s + b^s \leq (a+b)^s$

If $0 \leq s \leq 1$, $a^s + b^s \geq (a+b)^s$

1. Extension of a Theorem of H. Weyl.

The object of this section is to extend a result due to H. Weyl (16) .

Lemma 5.2. *Let* A, B *be any two* n-*square non-negative Hermitian matrices, with eigenvalues* $\alpha_1 \leq \dots \leq \alpha_n$, $\beta_1 \leq \dots \leq \beta_n$ *respectively*. *If* $\sigma + \delta = 1$, $\sigma \geq 0$, $\delta \geq 0$, *and if the eigenvalues of* $\sigma A + \delta B$ *are* $\lambda_1 \leq \dots \leq \lambda_n$, *then for* $1 \leq r \leq k \leq n$

$$E_r(\lambda_1,\dots,\lambda_k) \geq E_r^{\sigma}(\alpha_1,\dots,\alpha_k)\, E_r^{\delta}(\beta_1,\dots,\beta_k)$$

Proof. Let $x_1, \dots, x_k$ be orthonormal eigenvectors of $\sigma A + \delta B$ corresponding to $\lambda_1, \dots, \lambda_k$. Then

$$
\begin{aligned}
E_r^{1/r}(\lambda_1,\ldots,\lambda_k) &= E_r^{1/r}(\ ((\sigma A + \delta B)x_1,\ x_1),\ \ldots) \\
&\geq \sigma\, E_r^{1/r}\ ((Ax_1,\ x_1),\ \ldots\ ) \\
&\quad + \delta\, E_r^{1/r}\ ((Bx_1,\ x_1),\ \ldots\ ) \\
&\geq \sigma\, E_r^{1/r}\ (\alpha_1,\ \ldots,\ \alpha_k) + \delta\, E_r^{1/r}(\beta_1,\ \ldots,\ \beta_k) \\
&\geq E_r^{\sigma/r}\ (\alpha_1,\ \ldots,\ \alpha_k)\ E_r^{\delta/r}(\beta_1,\ \ldots,\ \beta_k)
\end{aligned}
$$

using Theorem 4.2 Corollary 1, Theorem 4.10 and Lemma 5.1.

We can now give the extension of Weyl's result.

Theorem 5.3. *Let* A *be an arbitrary n-square matrix with eigenvalues* λ_i *such that* $|\lambda_1| \geq \ldots \geq |\lambda_n|$. *Let* $\sigma \geq 0$, $\delta \geq 0$, $\sigma + \delta = 1$, *and let* $\sigma A^*A + \delta AA^*$ *have eigenvalues* $\alpha_1^2 \geq \ldots \geq \alpha_n^2$. *Then for* $1 \leq r \leq k \leq n$ and $s \geq 1$

$$E_r(\alpha_1^{2s}\ ,\ \ldots,\ \alpha_k^{2s}) \geq E_r(|\lambda_1|^{2s},\ \ldots,\ |\lambda_k|^{2s}\ ) .$$

Proof. By Schur's Lemma choose orthonormal vectors $x_1,\ \ldots,\ x_k$ such that $(Ax_i,\ x_i) = \lambda_i$ $(i = 1,\ \ldots,\ k)$ and $(Ax_i,\ x_j) = 0$ $(i > j)$. Then, if $f(x) = x^s$,

$$
\begin{aligned}
E_r(\alpha_1^{2s},\ \ldots,\ \alpha_k^{2s}) &\geq \sum_{\omega \in Q_{kr}} (C_r\left[f(\sigma A^*A + SAA^*)\right] x_\omega,\ x_\omega) \\
&\geq \sum_{\omega} f\left[(C_r(\sigma A^*A + \delta AA^*)x_\omega,\ x_\omega)\right] \\
&= \sum f\left[\det(\sigma(A^*Ax_{i_s},x_{i_t}) + \delta(AA^*x_{i_s},\ x_{i_t}))\right] \\
&\geq \sum f\left[\left\{\det(A^*Ax_{i_s},x_{i_t})\right\}^{\sigma}\left\{\det(AA^*x_{i_s},x_{i_t})\right\}^{\delta}\right]
\end{aligned}
$$

$$= \sum f \left[(C_r(A^* A)x_\omega, x_\omega)^\sigma \; (C_r(AA^*)x_\omega, x_\omega)^\delta \right]$$

$$= \sum f \left[(C_r(A)x_\omega, C_r(A)x_\omega)^\sigma \; (C_r(A^*)x_\omega, C_r(A^*)x_\omega)^\delta \right]$$

$$\geq \sum f \left[\left|(C_r(A)x_\omega, x_\omega)\right|^{2\sigma} \; \left|(C_r(A^*)x_\omega, x_\omega)\right|^{2\delta} \right]$$

$$= \sum f \left[\left| \det \begin{pmatrix} \lambda_{i_1} & & \\ & \ddots & \\ 0 & & \lambda_{i_r} \end{pmatrix} \right|^{2\sigma} \left| \det \begin{pmatrix} \overline{\lambda}_{i_1} & & 0 \\ & \ddots & \\ & & \overline{\lambda}_{i_r} \end{pmatrix} \right|^{2\delta} \right]$$

$$= \sum f \left(|\lambda_{i_r}|^2 \cdots |\lambda_{i_r}|^2 \right)$$

$$= E_r \left(|\lambda_1|^{2s}, \ldots, |\lambda_k|^{2s} \right)$$

Here, in succession, we have used Theorem 4.9, Theorem 4.7, Lemma 5.2 (in the special case $r = k = n$) and the fact that $(u, u) \geq |(u, v)|^2$ if v is a unit vector. This completes the proof.

Remarks. This theorem, for the case $k = n$ and any $s \geq 0$ was given by H. Weyl in 1949 (16) . For non-singular matrices, the case $k = n$ was also extended by H. Weyl to negative s .

2. Extension of a recent Theorem of Wielandt.

To simplify the statements of the theorems in this subsection, we introduce some notations. Throughout this subsection we shall assume that S is a given set of p natural numbers $\leq n$ and that $i < j < \ldots < m$ are its elements. By V_i, V_j, ..., V_m will be meant subspaces of V_n with the properties that

$$V_i \subset V_j \subset \dots \subset V_m$$
$$\text{dimension } V_\sigma = \sigma \quad (\sigma \in S) \tag{1}$$

In general, a subscript to a symbol denoting a subspace of V_n will indicate the dimension of the subspace. If $r \leq p$ we define S_{rp} as

$$S_{rp} = \left\{ \omega = \{i_1, \dots, i_r\} \,\middle|\, 1 \leq i_1 < \dots < i_r \leq n; \text{ all } i_j \in S \right\}$$

We can now state Theorem 5.4.

<u>Theorem 5.4. Let S be a set of p natural numbers $\leq n$ and for $1 \leq r \leq p \leq n$ let S_{rp} be as given above. Let A be a non-negative Hermitian operator on V_n such that $\alpha_1 \geq \dots \geq \alpha_n$ and $\beta_1 \leq \dots \leq \beta_n$ are two representations of its eigenvalues. If s, t are any two real numbers such that $0 \leq s \leq 1$, $t \geq 1$, then :</u>

(i) $E_r(B_\sigma;\ \sigma \in S) = \min\limits_V \max\limits_x E_r\left((Ax_\sigma, x_\sigma)^s;\ \sigma \in S\right)$

(ii) $E_1(\alpha_\sigma^t;\ \sigma \in S) = \max\limits_V \min\limits_x E_1\left((Ax_\sigma, x_\sigma)^t;\ \sigma \in S\right)$

(iii) $E_r(B_\sigma^s;\ \sigma \in S) = \min\limits_V \max\limits_x \sum\limits_{\omega \in S_{rp}} (C_r(A)x_\omega, x_\omega)^s$

(iv) $E_r(\alpha_\sigma^t;\ \sigma \in S) = \max\limits_V \min\limits_x \sum\limits_{\omega \in S_{rp}} (C_r(A)x_\omega, x_\omega)^t$

<u>where, if $\omega = \{i_1, \dots, i_r\} \in S_{rp}$, then $x_\omega = x_{i_1} \wedge \dots \wedge x_{i_r}$ and $\max\limits_x$ ($\min\limits_x$) indicates the max (min) over all possible p orthonormal vectors $x_i, x_j, \dots, x_m$ satisfying</u>

$$x_\sigma \in V_\sigma \quad (\sigma \in S) \tag{2}$$

<u>for fixed subspaces</u> (1), <u>and</u> $\min_V$ ($\max_V$) <u>indicates the min</u> (<u>max</u>) <u>over all possible subspaces</u> (1).

<u>Proof</u>. (i): The result is equivalent to the propositions (a) and (b) below.

(a): There is a special sequence $Y_i \subset \ldots \subset Y_m$ of subspaces of V_n such that for every orthonormal set $x_i, x_j, \ldots, x_m$ with $x_\sigma \in Y_\sigma$ ($\sigma \in S$), we have

$$E_r((Ax_\sigma, x_\sigma)^s, \sigma \in S) \leq E_r(B_\sigma^s ; \sigma \in S) . \tag{3}$$

For let $y_1, \ldots, y_n$ be orthonormal eigenvectors of A corresponding to $\beta_1, \ldots, \beta_n$. Define Y_σ to be the subspace spanned by $y_1, \ldots, y_\sigma$. Then $(Ax_\sigma, x_\sigma) \leq \beta_\sigma$ ($x_\sigma \in Y_\sigma$, $(x_\sigma, x_\sigma) = 1$), hence (3) follows.

(b): Let $V_i, V_j, \ldots, V_m$ be given subspaces of V_n satisfying (1). Then there are orthonormal vectors $x_i, \ldots, x_m$ satisfying (2) such that

$$E_r((Ax_\sigma, x_\sigma)^s ; \sigma \in S) \geq E_r(\beta_\sigma^s ; \sigma \in S) \tag{4}$$

The proof is by induction on n.

First, let S contain all natural numbers $\sigma \leq n$. Then choose $x_1, \ldots, x_n$ satisfying (2) but otherwise arbitrary. Then by Theorem 4.2 Corollary 1

$$E_r((Ax_1, x_1)^s, \ldots, (Ax_n, x_n)^s) \geq E_r(\beta_1^s, \ldots, \beta_n^s)$$

Hence (4) holds. This especially applies to $(r=)n=1$.

In what follows we assume there is a natural number $\leq n$ which is not in S. The largest such gap number will be denoted by g. We define f to be the largest number in S which is $< g$ if such exists; if not, we define $f = 0$. In either case $0 \leq f < g \leq n$. If $f > 0$, V_f is defined by hypothesis; if $f = 0$ we define $V_f = 0$.

The simplest case is the one for which $n \notin S$, that is, $g = n$. Then we choose any subspace $\bar{V}_{n-1}$ containing V_f. Define $\bar{A}$ to be the unique Hermitian operator on $\bar{V}_{n-1}$ such that

$$(\bar{A}x, x) = (Ax, x), \quad x \in \bar{V}_{n-1} \tag{5}$$

The eigenvalues $\bar{\beta}_1 \leq \cdots \leq \bar{\beta}_{n-1}$ of $\bar{A}$ are known (2) to satisfy

$$0 \leq \beta_1 \leq \bar{\beta}_1 \leq \beta_2 \leq \bar{\beta}_2 \leq \cdots \leq \beta_{n-1} \leq \bar{\beta}_{n-1} \leq \beta_n \tag{6}$$

By the induction hypothesis there are orthonormal vectors $x_\sigma \in \bar{V}_{n-1}$ such that $x_\sigma \in V_\sigma$ $(\sigma \in S)$ and

$$E_r\,((\bar{A}x_\sigma, x_\sigma)^s\,;\ \sigma \in S) \geq E_r(\,\bar{\beta}_\sigma^{\,s};\ \sigma \in S)$$

In view of (5) and (6) this inequality implies (4). This finishes the case $g = n$.

Now let $n \in S$, that is, $g < n$. We choose orthonormal vectors $y_{g+1}, \ldots, y_n$ of A corresponding to $\beta_{g+1}, \ldots, \beta_n$. Together with V_f they span a subspace of dimension $\leq (n-g) + f < n$. Hence we can choose some subspace $\bar{V}_{n-1}$ such that

$$V_f \subset \bar{V}_{n-1}, \; y_v \in \bar{V}_{n-1}, \; (v = g+1, \ldots, n) \tag{7}$$

Since $g+1, \ldots, n \in S$ we have

$$V_f \subset V_{g+1} \subset \ldots \subset V_n$$

so

$$V_f \subseteq V_{g+1} \cap \bar{V}_{n-1} \subsetneq \ldots \subsetneq V_{n-1} \cap \bar{V}_{n-1} \subsetneq \bar{V}_{n-1}$$

Since the dimension of $V_v \cap \bar{V}_{n-1}$ is at least $v-1$ we can choose subspaces $\bar{V}_g, \ldots, \bar{V}_{n-2}$ such that

$$\bar{V}_g \subset V_{g+1}; \; \ldots \; ; \; \bar{V}_{n-2} \subset V_{n-1} \tag{8}$$

$$V_i \subset \ldots \subset V_f \subset \bar{V}_g \subset \bar{V}_{g+1} \subset \ldots \subset \bar{V}_{n-2} \subset \bar{V}_{n-1} \tag{9}$$

We define as before the operator $\bar{A}$ on $\bar{V}_{n-1}$. By the induction hypothesis applied to $\bar{A}$ and the subspaces (8) there exist orthonormal vectors x_σ $(\sigma \in S)$ such that

$$x_\sigma \in V_\sigma \; (\sigma < g), \; x_\sigma \in \bar{V}_{\sigma-1} \; (\sigma > g) \tag{10}$$

$$E_r\left((\bar{A}x_\sigma, x_\sigma)^s; \sigma \in S\right) \geq E_r(\bar{\beta}_i^s, \ldots, \bar{\beta}_f^s, \bar{\beta}_g^s, \ldots, \bar{\beta}_{n-1}^s)$$

Using (5), (6) and (8) we find that

$$x_\sigma \in V_\sigma \quad (\sigma \in S) \tag{11}$$

$$E_r((Ax_\sigma, x_\sigma)^s; \sigma \in S) \geq E_r(\beta_i^s, \ldots, \beta_f^s, \bar{\beta}_g^s, \ldots, \bar{\beta}_{n-1}^s)$$

Now, by (7) we know that $y_{g+1}, \ldots, y_n$ are eigenvectors of $\bar{A}$ with eigenvalues $\beta_{g+1}, \ldots, \beta_n$. Hence for the largest eigenvalues of $\bar{A}$ we have

$$\beta_n \leq \bar{\beta}_{n-1}; \ \beta_{n-1} \leq \bar{\beta}_{n-2}; \ \cdots ; \ \beta_{g+1} \leq \bar{\beta}_g$$

so that

$$E_r((Ax_\sigma,x_\sigma)^s; \ \sigma \in S) \geq E_r(\beta_i^s,\ldots,\beta_f^s, \ \beta_{g+1}^s,\ldots,\beta_n^s)$$

$$= E_r(\beta_\sigma^s; \ \sigma \in S) \ .$$

This completes the proof of (b) and hence of (i).

Proof of (ii): Obvious modifications to the above proof and the use of Theorem 4.9 Corollary (a) to begin the proof of the analogue to (b) above will yield the result.

Proof of (iii): We have to establish the analogues to (a) and (b). (a) follows easily using the Hodamard determinant inequality. (b) follows without any alterations upon noting that the case in which S contains all natural numbers $\leq n$ is a consequence of Theorem 4.9.

Proof of (iv): The only difficulty here lies in demonstrating the analogue to (a), namely that subspaces $Y_i \subset Y_j \subset \ldots \subset Y_m$ exist such that

$$\sum_{\omega \in S_{rp}} (C_r(A)x_\omega,x_\omega)^t \geq E_r(\alpha_\sigma^t \ ; \ \sigma \in S)$$

for $x_\sigma \in Y_\sigma$ $(\sigma \in S)$. We give the essential fact in detail as follows. If $y_1, \ldots, y_n$ are orthonormal eigenvectors of A corresponding to $\alpha_1, \ldots, \alpha_n$ and Y_σ $(\sigma \in S)$ is the subspace of V_n spanned by $y_1, \ldots, y_\sigma$, and if $\omega = \{i_1, \ldots, i_r\}$ is an element of S_{rp}, then

$$\min \ (C_r(A)x_\omega, \ x_\omega) = \alpha_{i_1} \cdots \alpha_{i_r}$$

where the min is taken over all orthonormal vectors $x_{i_1}, \ldots, x_{i_r}$ with $x_{i_j} \in Y_{i_j}$, $j = 1, \ldots, r$. To see this note that any $x_{i_1} \wedge \ldots \wedge x_{i_r}$ lies in the subspace of $V_{\binom{n}{r}}$ spanned by the set of vectors

$$\left\{ y_{j_1} \wedge \ldots \wedge y_{j_r} \;\middle|\; j_1 \leq i_1;\ j_2 \leq i_2\ ;\ \ldots\ ;\ j_r \leq i_r \right\} \tag{12}$$

This subspace of $V_{\binom{n}{r}}$ is an invariant space of $C_r(A)$ so let B be the restriction of $C_r(A)$ to this subspace. Then the set of eigenvalues of B is exactly the set

$$\left\{ \alpha_{j_1} \ldots \alpha_{j_r} \;\middle|\; j_1 \leq i_1\ ;\ \ldots\ ;\ j_r \leq i_r \right\}$$

and $\alpha_{i_1} \ldots \alpha_{i_r}$ is the smallest eigenvalue of B. Hence for any unit vector z in the space spanned by (11), $(BZ, Z) \geq \alpha_{i_1} \ldots \alpha_{i_r}$, so

$$\min\ (C_r(A)x_\omega,\ x_\omega) = \alpha_{i_1} \ldots \alpha_{i_r}\ .$$

Theorem 5.4 is now completely proved.

We can recast Theorem 5.4 in a slightly different form (Theorem 5.5).

<u>Theorem 5.5</u>. <u>Under the assumptions of Theorem</u> 5.4.

(i) $E_r(\alpha_\sigma^s\ ;\ \sigma \in S) = \min\limits_V\ \max\limits_x\ E_r((Ax_\sigma,\ x_\sigma)^s\ ;\ \sigma \in S)$

(ii) $E_1(\beta_\sigma^t;\ \sigma \in S) = \max\limits_V\ \min\limits_x\ E_1((Ax_\sigma, x_\sigma)^t;\ \sigma \in S)$

(iii) $E_r(\alpha_\sigma^s;\ \sigma \in S) = \min\limits_V\ \max\limits_x\ \sum\limits_{\omega \in S_{rp}} (C_r(A)x_\omega,\ x_\omega)^s$

(iv) $E_r(\beta_\sigma^t\ ;\ \sigma \in S) = \max\limits_V\ \min\limits_x\ \sum\limits_{\omega \in S_{rp}} (C_r(A)x_\omega,\ x_\omega)^t$

where, in (i) for example, for fixed subspaces

$$V_{i-1} \subset V_{j-1} \subset \dots \subset V_{m-1}$$

$$\dim V_{\sigma-1} = \sigma - 1 \quad (\sigma \in S) \tag{13}$$

the max is taken over all orthonormal vectors $x_i, x_j, \dots, x_m$ satisfying

$$x_\sigma \in V_{\sigma-1} \quad (\sigma \in S) \tag{14}$$

and the min is taken over all possible subspaces (13) of V_n .

Proof. (i) Given $S = \{i, j, \dots, m\}$, let $T = \{n-m+1, \dots, n-i+1\}$. Then by Theorem 5.4

$$E_r(\alpha_\sigma^s ; \sigma \in S) = \min_{V_{n-\sigma+1}} \max_{x_\sigma \in V_{n-\sigma+1}} E_r((Ax_{\sigma 1}x_\sigma)^s ; \sigma \in S)$$

Letting $T_{\sigma-1}$ be the orthogonal complement in V_n to $V_{n-\sigma+1} (\sigma \in S)$, we have

$$E_r(\alpha_\sigma^s; \sigma \in S) = \min_{T_{\sigma-1}} \max_{x_\sigma \perp T_{\sigma-1}} E_r((Ax_\sigma,x_\sigma)^s; \sigma \in S)$$

If we relabel $T_{\sigma-1}$ as $V_{\sigma-1}$, (i) will be proved. (ii), ((iii), (iv) follow similarly.

Theorem 5.6. Let A, B, C be non-negative Hermitian operators on V_n such that C = A+B . Let $\alpha_\nu, \beta_\nu, \gamma_\nu$ $(\alpha', \beta_\nu', \gamma_\nu')$ be the respective eigenvalues of A, B, C numbered in decreasing (increasing) order. Then, under the assumptions of Theorem 5.4

$$\text{(i)} \quad E_r^{1/r}(\gamma_\sigma'^{s} ; \sigma \in S) \geq \Big[E_r^{1/r}(\alpha_\sigma' ; \sigma \in S) + E_r^{1/r}(\beta_1', \dots, \beta_p')\Big]^s$$

(ii) $E_r(\gamma_\sigma^t;\ \sigma \in S) \leq \binom{p}{r} \dfrac{(\alpha_i+\alpha_j+\ldots+\alpha_m+\beta_1+\ldots+\beta_p)^{tr}}{p^r}$

Proof.(i) By the proof of Theorem 5.4 there are subspaces $V_i \subset V_j \subset \ldots \subset V_m$ such that

$$E_r^{1/r}(\gamma_\sigma'^{s};\ \sigma \in S) = \max_{\substack{x_\sigma \in V_\sigma \\ (x_\alpha, x_\beta)=\delta_{\alpha\beta}}} E_r^{1/r}((Cx_\sigma, x_\sigma)^s;\ \sigma \in S)$$

Keeping the subspaces V_σ fixed we choose orthonormal vectors z_σ such that $z_\sigma \in V_\sigma$ $(\sigma \in S)$ and such that

$$E_r^{1/r}((Az_\sigma, z_\sigma);\ \sigma \in S) = \max_{x_\sigma \in V_\sigma} E_r^{1/r}((Ax_\sigma, x_\sigma);\ \sigma \in S)$$

Then using Lemma 5.1 and Theorem 4.2 Corollary 1,

$$\begin{aligned} E_r^{1/r}(\gamma_\sigma^{1s};\ \sigma \in S) &\geq E_r^{1/r}((Cz_\sigma, z_\sigma)^s;\ \sigma \in S) \\ &\geq E_r^{s/r}((Cz_\sigma, z_\sigma);\ \sigma \in S) \\ &\geq \Big[E_r^{1/r}((Az_\sigma, z_\sigma);\ \sigma \in S) \\ &\qquad + E_r^{1/r}((Bz_\sigma, z_\sigma);\ \sigma \in S)\Big]^s \\ &\geq \Big[E_r^{1/r}(\alpha_\sigma';\ \sigma \in S) + E_r^{1/r}(\beta_1', \ldots, \beta_p')\Big]^s \end{aligned}$$

(ii) As before we choose subspaces $V_i \subset \ldots \subset V_m$ such that

$$E_r(\gamma_\sigma^t;\ \sigma \in S) = \min_{x_\sigma \in V_\sigma} \sum_{\omega \in S_{rp}} ((C_r(C)x_\omega, x_\omega)^t$$

and orthonormal vectors z_σ such that

$$z_\sigma \in V_\sigma \quad (\sigma \in S)$$

$$E_1((Az_\sigma, z_\sigma);\ \sigma \in S) = \min_{x_\sigma \in V_\sigma} E_1((Ax_\sigma, x_\sigma);\ \sigma \in S) .$$

Then

$$E_r(\ \gamma_\sigma^t\ ;\ \sigma \in S) \leq \sum_{\omega \in S_{rp}} (C_r(C) z_\omega,\ z_\omega)^t$$

$$\leq \sum \prod_{\alpha=1}^{r} (Cz_{i_\alpha},\ z_{i_\alpha})^t$$

$$= E_r\ (\ (\ Cz_\sigma,\ z_\sigma)^t\ ;\ \sigma \in S)$$

$$\leq \binom{p}{r} \left(\frac{\sum_{\sigma \in S} ((A+B) z_\sigma,\ z_\sigma)^t}{p} \right)^r$$

$$\leq \binom{p}{r} \frac{\left(\sum_\sigma (Az_\sigma, z_\sigma) + (Bz_\sigma, z_\sigma)\right)^{tr}}{p^r}$$

$$\leq \binom{p}{r} \frac{\left(\sum_{\sigma \in S} \alpha_\sigma + \beta_1 + \ldots + \beta_p\right)^{tr}}{p^r}$$

Remarks: Theorems 5.4 (ii), 5.5 (ii) and 5.6 (ii) in the special cases $r=s=t=1$ were given by Wielandt in 1955 (15). As generalizations of these special cases, the contents of this subsection are believed to be new. However, the proofs lean heavily on Weilandt's original proof. In passing we may notice that the value of

$$\max \min E_r((Ax_\sigma, x_\sigma);\ \sigma \in S)$$

is as yet unknown.

VI. A Norm Inequality

Definition 6.1. *For any n-square matrix* A *define*
$||A||^2 = \text{trace } AA^*$

Theorem 6.2. *For any Hermitian matrix* H *let* $\lambda_v(H)$ *denote the eigenvalues of* H *in increasing order. Then if* A, B *are any n-square matrices*

$$\begin{aligned} 1/2 \sum_{j=1}^{n} \left[\lambda_j(A+A^*)\lambda_{n-j+1}(B+B^*) + \lambda_j\left(\frac{A-A^*}{i}\right)\lambda_{n-j+1}\left(\frac{B-B^*}{i}\right)\right] \\ \leq ||A+B||^2 - ||A||^2 - ||B||^2 \\ \leq 1/2 \sum_{j=1}^{n}\left[\lambda_j(A+A^*)\lambda_j(B+B^*)+\lambda_j\left(\frac{A-A^*}{i}\right)\lambda_j\left(\frac{B-B^*}{i}\right)\right] \end{aligned}$$

Proof. For any non-negative Hermitian matrix H, Theorem 4.2 states

$$\sum_{j=1}^{k} \lambda_j(H) \leq \sum_{j=1}^{k} (Hx_j, x_j) \leq \sum_{j=1}^{k} \lambda_{n-j+1}(H)$$

if the x's are orthonormal vectors. It can be shown that this inequality holds for any Hermitian matrix H. (In fact, in the special case $p = 1$, the positive definite assumption in Theorem 4.2 is not needed). In what follows we shall let $x_1, \ldots, x_n$ be orthonormal eigenvectors of $A+A^*$ corresponding to $\lambda_1(A+A^*), \ldots, \lambda_n(A+A^*)$, and let $y_1, \ldots, y_n$ be orthonormal eigenvectors of $\frac{A-A^*}{i}$ corresponding to

$\lambda_1\left(\frac{A-A^*}{i}\right), \ldots, \lambda_n\left(\frac{A-A^*}{i}\right)$. Then :

$$2\left\{ ||A+B||^2 - ||A||^2 - ||B||^2 \right\}$$

$$= 2 \quad \text{trace} \quad (AB^* + BA^*)$$

$$= \sum_{j=1}^{n} \left[\left(\left((A+A^*)(B+B^*) + \left(\frac{A-A^*}{i}\right)\left(\frac{B-B^*}{i}\right) \right) x_j, x_j \right) \right]$$

$$= \sum_{j=1}^{n} ((A+A^*)(B+B^*)x_j, x_j) + \sum_{j=1}^{n} \left(\left(\frac{A-A^*}{i}\right)\left(\frac{B-B^*}{i}\right) y_j, y_j \right)$$

$$= \sum_{j=1}^{n} \lambda_j \; (A+A^*)((B+B^*)x_j, x_j) + \sum_{j=1}^{n} \lambda_j \left(\frac{A-A^*}{i}\right)\left(\left(\frac{B-B^*}{i}\right) y_j, y_j\right)$$

$$= \lambda_1(A+A^*) \sum_{j=1}^{n} ((B+B^*)x_j, x_j)$$

$$+ (\lambda_2(A+A^*) - \lambda_1(A+A^*)) \sum_{j=2}^{n} ((B+B^*)x_j, x_j)$$

$$+ \ldots + (\lambda_n(A+A^*) - \lambda_{n-1}(A+A^*)((B+B^*)x_n, x_n)$$

$$+ \text{ a similar decomposition for the second sum}$$

$$\leq \lambda_1 \; (A+A^*) \sum_{j=1}^{n} \lambda_j \; (B+B^*)$$

$$+ (\lambda_2 \; (A+A^*) - \lambda_1(A+A^*)) \sum_{j=2}^{n} \lambda_j \; (B+B^*)$$

$$+ \ldots \text{ etc.}$$

$$= \sum_{j=1}^{n} \left[\lambda_j(A+A^*)\lambda_j \; (B+B^*) + \lambda_j \left(\frac{A-A^*}{i}\right) \lambda_j \left(\frac{B-B^*}{i}\right) \right]$$

The other inequality is proved in a similar fashion.

Various inequalities can be obtained as special cases of Theorem 6.2. If we let $B = -A$ we obtain

Corollary. For any n-square matrix A

$$\frac{1}{4}\sum_{j=1}^{n}\left[\lambda_j(A+A^*)\lambda_{n-j+1}(A+A^*) + \lambda_j\left(\frac{A-A^*}{i}\right)\lambda_{n-j+1}\left(\frac{A-A^*}{i}\right)\right]$$

$$\leq ||A||^2$$

$$\leq \frac{1}{4}\sum_{j=1}^{n}\left[\lambda_j^2(A+A^*) + \lambda_j^2\left(\frac{A-A^*}{i}\right)\right] .$$

VII. Historical Survey.

The material contained in sections II and III concerning compound matrices has been known for some time. The Sylvester-Franke theorem was first proved about 1850. The proof given in this thesis is believed to be one of the shortest given. Other recent proofs have been given by Tornheim (14) and Flanders (7).

The contents of sections IV and V have interesting histories. Weyl's theorem, mentioned in section V.1, was given in 1949 (16). Shortly thereafter, in a paper devoted to Weyl's theorem, Fan (3) gave the special case $p = 1$ of Corollaries 1 and 2 of Theorem 4.2. Other special cases of these Corollaries were later given by Fan in problem form (5, 6). It was not until 1955 that Theorem 4.2 itself was conjectured and proved by Marcus and McGregor (10).

The useful inequality Theorem 4.10 was first given in 1955, proofs are due (separately) to Marcus (11) and Bohnenblust (1).

The origin of the contents of section V.2 is a well known minimax principle of Courant (2) (see also (8)). Wielandt's (1955) theorem (15) is really an extension of the earlier Courant principle. Further extensions have been given in this thesis.

The contents of section VI are believed to be new.

VIII. Bibliography

1. H.F. Bohnenblust, Inequalities between symmetric functions, to appear.
2. R. Courant, D. Hilbert, Methoden der mathematischen Physik, Berlin, 1931. p. 28.
3. Ky Fan, On a theorem of Weyl concerning the eigenvalues of linear transformations.I, Proc. Nat. Acad. Sci. U.S.A. vol. 35(1949) pp. 652-655.
4. , On a theorem of Weyl concerning the eigenvalues of linear transformations, .II, Proc. Nat. Acad. Sci. U.S.A., vol. 36. (1950), pp. 31-34.
5. , Problem 4429, Amer. Math. Monthly, proposed vol 58 (1951) p. 194, solution vol. 60(1953) pp. 48.
6. , Problem 4430, Amer. Math. Monthly, proposed vol 58 (1951) p. 194, solution vol. 60(1953 p. 50.
7. H. Flanders, A note on the Sylvester-Franke theorem, Amer. Math. Monthly, vol. 60 (1953), pp. 543-5.
8. P. Halmos, Finite Dimensional Vector Spaces, Annals of Mathematics Studies No. 7, Princeton, 1953 p. 157.
9. G. Hardy, J. Littlewood, G. Polya, Inequalities, Cambridge, 1934.
10. M. Marcus and J.L. McGregor, Extremal properties of Hermitian matrices, to appear in Can. J. Math.
11. M. Marcus, Some inequalities for Hermitian matrices, to appear.

12. M. Marcus, Symmetric functions of singular values, to appear.

13. H. Minkowski, Diskontinuitatsberich fur arithmetische Aquivalenz, Jounr. fur Math., vol. 129 (1905) pp. 220-274.

14. L. Tornheim, The Sylvester-Franke theorem, Amer. Math. Monthly, vol. 59, (1952) pp. 389-391.

15. H. Wielandt, An extremum property of sums of eigenvalues, Proc. Amer. Math. Soc., vol. 6 (1955), pp. 106-110.

16. H. Weyl, Inequalities between the two kinds of eigenvalues of linear transformations, Proc. Mat. Aca. Sci. U.S.A., vol. 35 (1949) pp. 408-411.