RELATIVE HERMITIAN MATRICES

Magnus R. Hestenes

RELATIVE HERMITIAN MATRICES

Magnus R. Hestenes

1. Introduction. The purpose of the present paper is to develop a spectral theory for an arbitrary $m \times n$ dimensional matrix A, which is analogous to that given in the hermitian case and which reduces to the usual spectral theory when A is hermitian. The theory is centered around the triple product AB^*C of matrices of the same dimension. Here B^* is the transpose of B in the field of real numbers and the conjugate transpose of B in the field of complex numbers. The matrix T will be said to be elementary in case $T = TT^*T$. Elementary matrices play the role of units and in case of vectors are unit vectors. Given an elementary matrix T and a matrix A of the same dimension the matrix TA^*T can be considered to be the conjugate transpose of A relative to T. If $A = TA^*T$, then A is hermitian relative to T. The polar decomposition theorem for matrices implies that to each matrix A there is a unique elementary matrix R such that A is hermitian relative to R, AR^* is non-negative hermitian in the usual sense, and R has the same null space as A. Every elementary matrix T relative to which A is hermitian is of the form $T = T_0 + R_1 - R_2$, where $R_1 + R_2 = R$ and T_0, R_1, R_2 are mutually *-orthogonal. Two matrices A and B are *-orthogonal in case $AB^* = 0$ and $A^*B = 0$. A matrix B will be called a section of A if B and $A - B$ are *-orthogonal.

If A is hermitian relative to an elementary matrix T, it is shown below that A and T can be written as sums of sections

$$A = A_1 + \cdots + A_q , \qquad T = T_1 + \cdots + T_q$$

such that $A_i = \lambda_i T_i$, where λ_i is a real number. Moreover these sections can be chosen so that $\lambda_i \neq \lambda_j, (i \neq j)$. If in this event the decomposition is unique. If $AT^* \geqq 0$, then $\lambda_i \geqq 0$. If in addition A and T have the same null space then $\lambda_i > 0$. In the event T is the identity, this result gives the usual spectral representation of hermitian matrices.

A matrix A will be said to be normal relative to an elementary matrix T in case $A = AT^*T = TT^*A, AA^*T = TA^*A$. In this event the spectral decomposition theorem described above holds, the coefficients λ_i being complex instead of real.

In the development of the theory the concept of *-commutativity of two matrices plays a significant role. The matrices A and B will be said to *-commute (see § 4 below) in case $AB^* = BA^*$ and $A^*B = B^*A$. If A and B *-commute, there is an elementary matrix T relative to

which they are hermitian. Moreover $A, B\ T$ can be written as sums of section A_i, B_i, T^i such that $A_i = \lambda_i T_i$ and $B_i = \mu_i T_i$, where λ_i and μ_i are real. If $\mathfrak{A}$ is a linear class of $m \times n$ dimensional matrices that are hermitian relative to an elementary matrix T in $\mathfrak{A}$ (including T itself) and have the property that the product AB^*C is invariant under permutations A, B and C, then this class forms an algebra with AT^*B as the product of A and B. The elements A of $\mathfrak{A}$ are all matrices expressible in the form $A = \lambda_i T_i + \cdots + \lambda_q T_q$, where $T_1, \cdots, T_q$ are suitably chosen sections of T. Throughout the paper the reciprocal A^{*-1} of A^* plays a role analogous to that of A itself.

The main results obtained in the present paper can be extended to a closed operator A from a Hilbert space $\mathfrak{N}$ to a second Hilbert space $\mathfrak{N}'$. Whenever it is convenient to do so, the theorems are stated so as to be valid for operators in Hilbert space. The terminology used has been chosen so as to make this transition as simple as possible. The extension to Hilbert spaces yields a common spectral theory for the gradient of a function and the divergence of a vector field.

2. Terminology and notations. Throughout the following pages matrices will be denoted by capital letters $A, B, C, P, Q, R, \cdots$. The elements can be considered to be real or complex. The conjugate transpose of A will be denoted by A^*. It will be convenient to consider all matrices to be square, since this can be obtained by the addition of zero-elements. However, this is not essential. The paper is written so as to be valid for rectangular matrices, the equality of the dimensions of two or more matrices being implied by the condition that the operations used should be well defined.

Occasionally we shall use column vectors and row vectors. A column vector will be denoted by $x, y, z, \cdots$. Row vectors are conjugate transposes of column vectors. If $x^*x = 1$ then x is a unit vector. Given two vectors x and y then $A = yx^*$ is a matrix of rank 1. Every matrix of rank 1 is represented in the form $A = \lambda yx^*$, where λ is a real number and x, y, are unit vectors. In fact λ can be taken to be positive. The greek letters, $\alpha, \beta, \gamma, \mu, \cdots$ appearing in the text normally denote real numbers.

A matrix is *hermitian* if $A^* = A$. A hermitian matrix A is *non-negative*, written $A \geqq 0$, if $x^*Ax \geqq 0$ for every column vector x. If $A \geqq 0$, there is a unique matrix $B \geqq 0$ such that $B^2 = A$. The matrix B will be called the *square root* of A. A matrix E will be called a *pro-*

---

[1] E. H. Moore, *General Analysis*, Part I, Mem. Philos. Soc. (1935), p. 197. See also Penrose, "A generalized inverse of matrices", *Proc.* Cambridge Philos. Soc. **51** (1953), 406–413; M. R. Hestenes, "Inversion of matrices by biorthogonalization and related results," *J. Soc. Ind. Appl. Math.* **6** (1958), p. 84; J. von Neumann, On regular rings, Proc. Nat. Acad. Sci. U. S. A. **22** (1936), 707–713.

jection if $E = E^* = E^2$. The identity matrix will be denoted by I. The null space of the matrix A will be denoted by $\mathfrak{N}_A$.

To each matrix A there is a unique matrix B such that

$$A = AA^*B = AB^*A = BA^*A \,, \qquad B = BB^*A = BA^*B = AB^*B \,.$$

The matrix B is the reciprocal of A^* in the sense of E. H. Moore[1] and will be called the *-reciprocal of A. It is also the conjugate transpose of the *reciprocal A^{-1} of A. If A is nonsingular, A^{-1} is the inverse of A. We shall accordingly use the symbols A^{*-1}, A^{-1*} for the *-reciprocal of A. The matrices

(2.1) $$E = A^{-1}A = A^*A^{*-1} \,, \qquad E' = AA^{-1} = A^{*-1}A^*$$

are projections and satisfy the relations

(2.2) $$E'A = AE = A \,, \qquad E'A^{*-1} = A^{*-1}E = A^{*-1} \,.$$

They will be called the *projections associated with* A. It should be noted that the reciprocal of A^*A is $A^{-1}A^{*-1}$ and that the reciprocal of AA^* is $A^{*-1}A^{-1}$. If A is hermitian then $A^{*-1} = A^{-1}$. If A is nonsingular then A^{-1} is the inverse of A.

A matrix R will be said to be an *elementary matrix* in case $RR^*R = R$. It is easily seen that R is elementary if and only if $R = R^{*-1}$ or equivalently if and only if $R^* = R^{-1}$. If R is elementary so also is R^*. A projection is an elementary matrix. If R is a hermitian elemetary matrix, then

$$E_+ = \frac{1}{2}(R^2 + R) \,, \qquad E_- = \frac{1}{2}(R^2 - R)$$

are projections such that

$$R = E_+ - E_- \,, \quad E_+E_- = E_-E_+ = 0 \,, \quad R^2 = E_+ + E_- \,.$$

Conversely, the difference of two projections that are orthogonal is a hermitian elementary matrix.

A matrix A will be said to be *hermitan relative to an elementary matrix* R if $A = RA^*R$. The following result is fundamental.

THEOREM 2.1. *Suppose that* A *is hermitian relative to an elementary matrix* R. *Then* A^{*-1} *is hermitian relative to* R *and* A^*, A^{-1} *are hermitian relative to* R^*. *Moreover,* $\mathfrak{N}_R \subset \mathfrak{N}_A$ *and* $\mathfrak{N}_{R^*} \subset \mathfrak{N}_{A^*}$. *The matrices* A *and* R *satisfy the further relations*

(2.3a) $$A = RR^*A = AR^*R = RA^*R \,, \quad RA^*A = AA^*R = AR^*A$$

(2.3b) $$A^*R = R^*A \,, \quad AR^* = RA^* \,.$$

(2.3c) $$(A^*R)^2 = A^*A \,, \quad (AR^*)^2 = AA^* \,.$$

(2.3d) $A^{*-1}A^*R = RA^*A^{*-1} = AR^*A^{*-1} = AA^{-1}R = RA^{-1}A = A^{*-1}R^*A$.

Since $A = RA^*R$ we have $A^* = R^*AR^*$ and $\mathfrak{N}_R \subset \mathfrak{N}_A, \mathfrak{N}_{R^*} \subset \mathfrak{N}_{A^*}$. Moreover

$$RR^*A = RR^*RA^*R = RA^*R = A = AR^*R$$
$$AR^*A = RA^*RR^*RA^*R = RA^*RA^*R = AA^*R = RA^*A .$$

It follows that (2.3a) holds. The relations (2.3b) and (2.3c) follow from the computations

$$A^*R = R^*AR^*R = R^*A , \quad RA^* = RR^*AR^* = AR^*$$
$$(A^*R)^2 = A^*RR^*A = A^*A , \quad (AR^*)^2 = AR^*RA^* = AA^* .$$

It is easily verified that $RA^{-1}R$ is the reciprocal $R^*AR^* = A^*$. Consequently $A^{*-1} = RA^{-1}R$, that is, A^{*-1} is hermitian relative to R. Similarly A^{-1} is hermitian relative to R^*. The relations (2.3d) follow from (2.3b) and the relations $AA^{-1} = A^{*-1}A^*$, $A^{-1}A = A^*A^{*-1}$.

COROLLARY. *Suppose $A = RA^*R$, $R = RR^*R$ and set $P = A^*R$, $Q = RA^*$.   Then*

(2.4) $A = RP = QR , \quad P = R^*QR , \quad Q = RPR^* .$

The matrix P is nonnegative if and only if Q is nonnegative.   Moreover $\mathfrak{N}_R = \mathfrak{N}_A$ if and only if

(2.5) $R = A^{*-1}R^*A$

and hence if and only if $\mathfrak{N}_{R^*} = \mathfrak{N}_{A^*}$.

In view of this result we define a matrix A to be *nonnegative hermitian relative to R* in case $A = RA^*R$ and $A^*R \geqq 0$.

THEOREM 2.2. *Given a matrix A there is a unique elementary matrix R such that A is nonnegatively hermitian relative to R and such that $R = A^{*-1}R^*A$.   Moreover*

(2.6) $R^*R = A^{-1}A = A^*A^{*-1} , \quad RR^* = AA^{-1} = A^{*-1}A^* .$

Let P be the square root of A^*A. The matrix $R = A^{*-1}P$ has the properties described in the theorem. Clearly $A^*R = A^*A^{*-1}P = P \geqq 0$. Moreover

$$RR^*R = A^{*-1}P^2A^{-1}R = A^{*-1}(A^*AA^{-1}A^{*-1})P = A^{*-1}P = R ,$$
$$RA^*R = RP = A^{*-1}P^2 = A^{*-1}A^*A = A ,$$
$$A^{*-1}R^*A = A^{*-1}PA^{-1}A = A^{*-1}P = R ,$$
$$R^*R = R^*A^{*-1}R^*A = A^{-1}A = A^*A^{*-1} ,$$
$$RR^* = A^{*-1}R^*AR^* = A^{*-1}A^* = AA^{-1} .$$

The uniqueness of R follows from the uniqueness of P as the square

root of A^*A. This proves the theorem.

If R is chosen as described in Theorem 2.2 then the formula (2.4) for A in terms of $P = A^*R$ and R is called the polar decomposition[2] of A.

The matrix R described in Theorem 2.2 will be called the *elementary matrix associated with* A.

COROLLARY. *Let R be the elementary matrix associated with A and let S and T be elementary matrices such that*

$$A = SS^*A = AT^*T$$

Then $V = TR^*S$ is the elementary operator associated with $B = TA^*S$
This follows because

$$VB^*V = TR^*SS^*AT^*TR^*S = TR^*AR^*S = TA^*S = B$$
$$B^*V = S^*AT^*TR^*S = S^*(AR^*)S \geqq 0 \ .$$
$$B^{*-1}V^*B = B^{*-1}S^*RT^*TA^*S = B^{*-1}S^*RA^*S$$
$$= TA^{-1}SS^*AR^*S = TA^{-1}AR^*S = TR^*S = V \ .$$

As a further result we have

THEOREM 2.3. *Let R be the elementary matrix associated with A. If A is normal so also is R. If A is nonnegative hermitian, then R is a projection. If A is hermitian, then R is hermitian and is the difference of two orthogonal projections.*

If A is normal, its associated projections E, E' coincide. By virtue of (2.6) we have $RR^* = R^*R$ and R is normal. If A is nonnegative hermitian, then $R = E$. If A is hermitian, let P be the square root of A^2. Then $A = RP = PR = PR^* = R^*P$. Consequently $R = R^*$, as was to be proved.

The following result is of interest.

THEOREM 2.4. *Let R be the elementary matrix associated with A. There exists a unique pair of matrices B, C such that*

$$B + C = R \ , \qquad A = BR^*C^{*-1} = C^{*-1}R^*B$$

and having R as their associated elementary matrix. The matrices B and C are defined by the formulas

$$B^{-1} = A^{-1} + R^* \ , \qquad C^{-1} = A^* + R^*$$

and satisfy the relations

$$BB^*C = BC^*B = CB^*B \ , \qquad CC^*B = CB^*C = BC^*C$$

$$B^{-1}RC^{-1} = C^{-1}RB^{-1} = B^{-1} + C^{-1}$$
$$A^{*-1} = B^{*-1}R^*C = CR^*B^{*-1} , \qquad A^* = B^*RC^{-1} = C^{-1}RB^* ,$$
$$A^{-1} = B^{-1}RC^* = C^*RB^{-1} .$$

Since no direct use of this result will be made, its proof will be omitted.

3. ∗-orthogonality. Two matricies A and B will be said to be *∗-orthogonal* in case

$$(3.1) \qquad\qquad A^*B = B^*A = 0 , \qquad AB^* = BA^* = 0 .$$

Consider now two matrices A and B and let

$$(3.2) \qquad E = A^{-1}A , \quad E' = AA^{-1} , \quad F = B^{-1}B , \quad F' = BB^{-1}$$

be the associated projections. We have the following

LEMMA 3.1. *Two matrices A and B are ∗-orthogonal if and only if*

$$(3.3) \qquad\qquad EF = FE = 0 \qquad E'F' = F'E' = 0 .$$

Moreover two matrices A and B are ∗-orthogonal if and only if their associated elementary matrices R and S are ∗-orthogonal. Finally A is ∗-orthogonal to B if and only if A^{-1*} is ∗-orthogonal to B.

If (3.1) holds, then

$$EF = (A^{-1}A)(B^*B^{-1*}) = A^{-1}(AB^*)B^{-1*} = 0$$
$$E'F' = (A^{-1*}A^*)(BB^{-1}) = 0 .$$

Hence (3.3) holds. The converse follows from the relations $A = AE = E'A$, $B = BF = F'B$. The last two statements in the lemma follow from the first.

LEMMA 3.2. *Let A and B be ∗-orthogonal matrices and set $C = A + B$. Then*

$$(3.4) \quad C^* = A^* + B^* , \quad C^{-1} = A^{-1} + B^{-1} , \quad C^{*-1} = A^{*-1} + B^{*-1} .$$

The rank of C is the sum of the ranks of A and B. The elementary matrix associated with C is the sum $T = R + S$ of the elementary matrices R, S associated with A, B, respectively. The matrix C is elementary if and only if A and B are elementary.

By the use of Lemma 3.1 it is seen that

$$A^{-1}B = B^{-1}A = 0 , \quad AB^{-1} = BA^{-1} = 0 .$$

It follows that

$$(A^{-1} + B^{-1})C = A^{-1}A + B^{-1}B = E + F$$

$$C(A^{-1} + B^{-1}) = AA^{-1} + BB^{-1} = E' + F' \, .$$

In view of (3.3) the matrices $G = E + F$ and $G' = E' + F'$ are projections. Moreover, setting $C^{-1} = A^{-1} + B^{-1}$ we have

$$G'C = CG = C \, , \quad C^{-1}G' = GC^{-1} = C^{-1} \, .$$

The matrix C^{-1} is therefore reciprocal of C and the relations (3.4) hold. To show that $T = R + S$ is the elementary matrix associated with C observe that $R^*B = S^*A = 0$, $BR^* = AS^* = 0$, by Theorem 3.1. Hence

$$T^*C = R^*A + S^*B = A^*R + B^*S = C^*T \geqq 0 \, , \quad CT^* = TC^* \geqq 0 \, .$$
$$T^*T = R^*R + S^*S = E + F = G = C^{-1}C \, , \qquad TT^* = G' = CC^{-1} \, ,$$

as was to be proved. The remaining statements in the lemma are easily established.

A matrix A will be said to be a *section* of a matrix C if there is a matrix B *-orthogonal to A such that $C = A + B$. By virtue of the last lemma the elementary matrix R associated with a section A of C is a section of the elementary matrix T belonging to C. A section of an elementary matrix is elementary.

LEMMA 3.3. *Let E, E' be the projections associated with a matrix A and let F and F' be projections such that $F'A = AF$. Then $EF = FE$, $E'F' = F'E'$. Moreover $A_1 = AF$ is a section of A.*

Since $AE = A$ it follows that $AFE = F'AE = F'A = AF$. Consequently

$$EFE = A^{-1}AFE = A^{-1}AF = EF \, .$$

This implies that $EF = FE$. Similarly $E'F' = F'E'$. Observe that

$$FA^{-1} = FEA^{-1} = EFA^{-1} = A^{-1}AFA^{-1} = A^{-1}F'AA^{-1} = A^{-1}F'E' = A^{-1}F'$$
$$FA^{-1}AF = FEF = EF \, , \qquad F'AA^{-1}F' = F'E'F' = E'F' \, .$$

Consequently $A_1^{-1} = FA^{-1} = A^{-1}F'$ is the reciprocal of $A_1 = AF = F'A$. The projections $E_1 = EF$, $E_0 = E - E_1$ are orthogonal as are $E_1' = E'F'$, $E_0' = E' - E_1'$. Consequently $A_0 = AE_0 = A - A_1 = E_0'A$ is *-orthogonal to A_1. Since $A = A_0 + A_1$ it follows that A_1 is a section of A, as was to be proved.

LEMMA 3.4. *A matrix B is a section of A if and only if $A^*B = B^*B$, $BA^* = BB^*$.*

Let

$$F = B^{-1}B \, , \qquad F' = BB^{-1} \, .$$

If $A^*B = B^*B$, $BA^* = BB^*$

$$A^*F' = A^*BB^{-1} = B^*BB^{-1} = B^*$$

$$FA^* = B^{-1}BA^* = B^{-1}BB^* = B^* .$$

Consequently $B = AF = F'A$ and B is a section of A, by Lemma 3.3. The converse is immediate.

LEMMA 3.5. *Let R be an elementary matrix and set $E = R^*R$. Let $S = RF$, where F is a projection. Then S is a section of R if and only if $EF = FE$. If r is the rank of R then R is expressible as the sum $R = R_1 + \cdots + R_r$ of r *-orthogonal sections of rank 1.*

If S is a section of R then $R^*S = S^*S = EF = EFE$. Consequently $EF = FE$. Conversely if $EF = FE$ then

$$S^*S = FR^*RF = FEF = EF = R^*RF = R^*S$$
$$SS^* = RF^2R^* = RFR^* = SR^* .$$

It follows from Lemma 3.4 that S is a section of R.

In order to prove the last statement in the theorem suppose that $R \neq 0$ and choose a unit vector x such that $Ex = x$. Then $E_1 = xx^*$ is a projection that commutes with E. Hence $R_1 = RE_1$ is a section of R of rank 1. Moreover $R - R_1$ is a section of R of rank $r - 1$ and is *-orthogonal to R_1. If $R - R_1 \neq 0$ it has a section R_2 of rank 1. Clearly R_2 is *-orthogonal to R_1 and $R - R_1 - R_2$ is a section of R of rank $r - 2$. By a repetition of this argument it is seen that R is expressible as the sum of r *-orthogonal sections, as was to be proved.

4. *-commutativity. Given two matrices A and B the products A^*B and AB^* can be considered to be two types of *-products of A and B. If these *-products are unaltered upon interchanging A and B, that is, if

(4.1) $$A^*B = B^*A , \qquad AB^* = BA^* ,$$

then A and B will be said to *-*commute*. It should be noted that A and B *-commute if and only if A^*B and AB^* are hermitian in the usual sense. As a first result we have

LEMMA 4.1. *If A and B *-commute and*

(4.2) $$E = A^{-1}A , \quad E' = AA^{-1} , \quad F = B^{-1}B , \quad F' = BB^{-1} ,$$

then $F'A = AF, E'B = BE$ and $EF = FE, E'F' = F'E'$.

For if (4.1) holds then, since $F'B = BF = B$, we have

$$F'AF = B^{*-1}B^*AF = B^{*-1}A^*BF = B^{*-1}A^*B = B^{*-1}B^*A = F'A$$
$$F'AF = F'AB^*B^{*-1} = F'B^*AB^{*-1} = B^*AB^{*-1} = AB^*B^{*-1} = AF .$$

Consequently $AF = F'A$. Similarly $E'B = BE$. In view of Lemma 3.3 the relations $EF = FE, E'F' = F'E'$ hold.

THEOREM 4.1. *Two matrices A and B $*$-commute if and only if they are expressible in the form $A = A_0 + A_1$, $B = B_0 + B_1$, where A_0 is $*$-orthogonal to A_1 and B, B_0 is $*$-orthogonal to B_1 and A, A_1 and B_1 $*$-commute and have the same associated projections*

$$(4.3) \qquad E_1 = A_1^{-1}A_1 = B_1^{-1}B_1 , \qquad E_1' = A_1 A_1^{-1} = B_1 B_1^{-1} .$$

If A and B $*$-commute, then, by Lemmas 4.1 and 3.3, the matrices $A_1 = AF$, $B_1 = BE$ are sections of A and B respectively and have $E_1 = EF$, $E_1' = E'F'$ as their associated projections. Moreover $A_0 = A - A_1$ has $E - EF$, $E' - E'F'$ as its projections and hence is $*$-orthogonal to B and A_1 and hence also to B_1 and $B_0 = B - B_1$. Similarly B_0 is $*$-orthogonal to B_1, A, A_0 and A_1. The converse is immediate and the lemma is proved.

COROLLARY. *Suppose that A and B $*$-commute. Then A and A^{*-1} $*$-commute with B and B^{*-1}. Moreover A^*, A^{-1} $*$-commute with B^* and B^{-1}.*

We shall see later that their associated matrices R, S $*$-commute with $A, B, R,$ and S.

THEOREM 4.2. *Let R be the elementary matrix associated with a matrix A and let S be an elementary matrix that $*$-commutes with A. Then S $*$-commutes with R. Moreover A, R, S are expressible uniquely as sums and differences*

$$(4.4) \quad A = A_0 + A_+ + A_- , \quad R = R_0 + R_+ + R_- , \quad S = S_0 + R_+ - R_-$$

of $*$-orthogonal matrices such that the matrices R_0, R_+, R_- are the elementary matrices associated with A_0, A_+, A_- respectively and such that S_0 is $*$-orthogonal to A_0 and R_0. Conversely if A, R, S can be decomposed in this manner the A and R $*$-commute with S.

By virtue of the last theorem A and S can be expressible uniquely as the sum of $*$-orthogonal sections $A = A_0 + A_1$, $S = S_0 + S_1$ such that A_1 and S_1 have the same associated projections, S_0 being $*$-orthogonal to A and A_0 being $*$-orthogonal to S. The elementary matrix R associated with A is expressible in the form $R = R_0 + R_1$, where R_0 and R_1 are the elementary matrices associated with A_0 and A_1 respectively. In view of these remarks we can restrict ourselves to the case in which $A_0 = 0$, $S_0 = 0$ and $R_0 = 0$. Then

$$E = A^{-1}A = R^*R = S^*S , \qquad E' = AA^{-1} = RR^* = SS^* .$$

Since A^*S is self-adjoint, its associated elementary matrix T is the difference $T = E_+ - E_-$ of two orthogonal projections E_+ and E_- whose sum is E. The matrix A^*ST is nonnegative and self-adjoint. It follows from Theorem 2.1 that $R = ST$. The matrices $R_+ = RE_+$, $R_- = RE_-$ are $*$-orthogonal elementary matrices such that

$$R = RE = R_+ + R_- , \qquad S = RT = R_+ - R_- .$$

Since AR^* and AS^* are hermitian it follows that the matrices

$$AR_+^* = \frac{1}{2}A(R^* + S^*) , \qquad AR_-^* = \frac{1}{2}A(R^* - S^*)$$

are hermitian and orthogonal. Moreover they are nonnegative because of the relations

$$0 \leqq AR^* = AR_+^* + AR_-^* .$$

The elementary matrices R_+ and R_- are accordingly the elementary matrices associated with $A_+ = AE_+$ and $A_- = AE_-$ respectively. It is clear that A_+ and A_- are $*$-orthogonal and that $A = A_+ + A_-$. The matrices A, R, S are therefore expressible in the form (4.4). The converse is immediate and the theorem is established.

COROLLARY 1. *Two elementary matrices R and S $*$-commute if and only if there exist mutually matrices R_0, R_+, R_-, S_0 such that $R = R_0 + R_+ + R_-, S = S_0 + R_+ - R_-$. Moreover this decomposition is unique.*

COROLLARY 2. *If the matrix S appearing in Theorem 4.2 is of rank 1 then the decomposition (4.4) of A takes the simpler form*

$$(4.5) \qquad\qquad A = \mu S + A_0$$

where μ is a real number and A_0 is $*$-orthogonal to S.

For in this case two of the matrices S_0, R_+, R_- are zero since S has rank 1. If $S = S_0$, then (4.5) holds with $\mu = 0$. If $S = R_+$, then $R_- = A_- = 0$ and A_+ is of rank 1. Since S^*A_+ is a nonegative hermitian matrix of rank 1 it follows that A_+ is of the form $A_+ = \mu S$, where $\mu > 0$. If $S = -R_-$, then A_- is of the form $A_- = \mu S$ with $\mu < 0$.

COROLLARY 3. *If $S_1, \cdots, S_r$ are r mutually $*$-orthogonal elementary matrices of rank 1 that $*$-commute with A, then A is expressible in the form*

$$(4.6) \qquad\qquad A = \mu_1 S_1 + \cdots + \mu_r S_r + A_0$$

where $\mu_1, \cdots, \mu_r$ are real numbers and A_0 is $*$-orthogonal to each S_i $(i = 1, \cdots, r)$ and hence to $S_1 + \cdots + S_r$.

This result follows from Corollary 2 by induction. At the kth step one applies Corollary 2 with A replaced by $A - \mu_1 S_1 - \cdots - \mu_{k-1}S_{k-1}$ and with $S = S_k$.

THEOREM 4.3. *If a matrix A $*$-commutes with every section of an elementary matrix S than A is expressible in the form*

$$A = \mu S + A_0$$

where μ is a real number and A_0 is $*$-orthogonal to S.

If S has rank 1, the theorem holds by virtue of Corollary 2 to Theorem 4.2. If S is of rank $r > 1$, then, by Lemma 3.5, S is expressible in the form $S = S_1 + \cdots + S_r$ where $S_1, \cdots, S_r$ are mutually $*$-orthogonal elementary matrices. Consequently A is expressible in the form (4.6). It remains to show that $\mu_1 = \mu_2 = \cdots = \mu_r$. To this end choose unit vectors x_i and y_i such that $S_i = x_i y_i^*$. Then for $i \neq j$ the vector x_i is orthogonal to x_j and y_i is orthogonal to y_j. Let α and β be two nonnull real numbers such that $\alpha^2 + \beta^2 = 1$ and set

$$x = \alpha x_i + \beta x_j , \qquad y = \alpha y_i + \beta y_j .$$

Then $T = xy^*$ is a section of S and is $*$-orthogonal to S_k if $k \neq i$, $k \neq j$. The matrix

$$A^* T = (\mu_i S_i^* + \mu_j S_j^*)T = (\mu_i \alpha y_i + \mu_j \beta y_j)y^*$$

is hermitian if and only if $\mu_i = \mu_j$, that is T $*$-commutes with A if and only if $\mu_i = \mu_j$. This completes the proof of the theorem.

THEOREM 4.4. *A matrix A $*$-commutes with an elementary matrix S and has no nonnull section $*$-orthogonal to S if and only if $A = SA^*S$.*

This result is easily established. The condition that $A = SA^*S$, when $S = I$ is the condition that A be hermitian. Accordingly one can consider the condition $A = SA^*S$ to be an extension of the concept of a matrix being hermitian.

In the complex domain we have the following:

COROLLARY. *If A is a matrix and S is an elementary matrix such that $SS^*A = AS^*S = A$, then $B = (1/2)(A + SA^*S)$ and $C = (1/2i)(A - SA^*S)$ $*$-commute with S. Moreover, $A = B + iC$.*

5. Principal values and principal sections of matrices. Here and elsewhere the symbol $\| y \|$ denotes the length or norm of the vector y. By the norm $\| A \|$ of a matrix A will be meant the least upper bound of the quantity $\| Ax \|$ for all unit vectors x. If R is the elementary matrix associated with A, then $\| A \|$ is equal to the least upper bound of $\| Ax \|$ subject to the condition $\| Rx \| = 1$. As is well known there is a unit vector x such that $\| Ax \| = \| A \|$. For such a vector x we have $\| Rx \| = 1$ also. It is well known that $\| A \| = \| A^* \|$. If $A \neq 0$ then $\| A^{-1} \| = \| A^{*-1} \|$ is equal to the least number m such that $\| Ax \| \geqq (1/m) \| Rx \|$.

THEOREM 5.1. *Let R be the elementary matrix associated with A. Given a positive number λ there exists a unique decomposition*

(5.1) $$A = A_+ + A_0 + A_- , \qquad R = R_+ + R_0 + R_-$$

on A and R into mutually $*$-orthogonal sections such that $R_+, R_0, R_-,$
$R_+ - R_-$ are the elementary matrices associated with $A_+, A_0, A_-, A - \lambda R$
respectively. Moreover $A_0 = \lambda R_0$ and

(5.2)
$$\| A_+ x \| > \lambda \| R_+ x \| \qquad whenever \ R_+ x \neq 0$$
$$\| A_- x \| < \lambda \| R_- x \| \qquad whenever \ R_- x \neq 0 \ .$$

If $\lambda \geqq \| A \|$, then $A_+ = R_+ = 0$. If $1/\lambda \leqq \| A^{*-1} \|$ then $A_- = R_- = 0$.

In order to prove this result let $B = A - \lambda R$ and let S be the associated elementary matrix. Since R $*$-commutes with A and R, it $*$-commutes with B and hence also with S. Since S $*$-commutes with B and R it follows that S $*$-commutes with A. Applying Theorem 4.2 to A, R, S and to B, R, S it is seen that they are expressible as sums

$$A = A_+ + A_0 + A_- , \qquad R = R_+ + R_0 + R_-$$
$$B = B_+ + B_0 - B_- , \qquad S = R_+ + S_0 - R_- \ .$$

of $*$-orthogonal matrices such that R_+ is the elementary matrix associated with A_+ and B_+, R_- is the elementary matrix associated with A_- and B_-, R_0 is the elementary matrix associated with A_0. It is clear that $B_0 = S_0 = 0$ since every matrix that is $*$-orthogonal to A and R is also $*$-orthogonal to B and S. From the relation $A = B + \lambda R$ it follows that

$$A_0 = \lambda R_0 , \quad A_+ = B_+ + \lambda R_+ , \quad A_- = B_- + \lambda R_- \ .$$

Consequently

$$R_+^* A_+ = R_+^* B_+ + \lambda E_+ , \qquad R_-^* A_- + R_-^* B_- = \lambda E_- \ .$$

Since these matrices are nonnegative and hermitian, it is seen that (5.2) holds. The last statement in the theorem follows from the relations (5.2).

THEOREM 5.2. *A nonnull matrix A and its associated elementary matrix R have unique decompositions of the form*

(5.3)
$$A = \lambda_1 R_1 + \cdots + \lambda_k R_k , \qquad R = R_1 + \cdots + R_k$$

into $*$-orthogonal sections, where $\lambda_1, \cdots, \lambda_k$ are distinct positive numbers.

In order to prove this result let $\lambda_1 = \| A \|$. By virtue of the last theorem the matrices A and R are expressible as sums

(5.4)
$$A = \lambda_1 R_1 + B , \qquad R = R_1 + S$$

of $*$-orthogonal sections with $\| B \| < \lambda_1$. If $B \neq 0$, choose $\lambda_2 = \| B \|$ and, by Theorem 5.1, again, B and S are sums

$$B = \lambda_2 R_2 + C , \qquad S = R_2 + T$$

of sections. Proceeding in this manner one obtains the representation

described in the theorem.

The numbers $\lambda_1, \cdots, \lambda_k$ appearing in the last theorem will be called the *principal values* of A and the matrices $A_1 = \lambda_1 R_1, \cdots, A_k = \lambda_k R_k$ will be called the *principal sections* of A. The rank of A_i will be called the *multiplicity* of λ_i as a principal value of A. It is easily seen that a number $\lambda > 0$ is a principal value of A of multiplicity m if and only if it is an eigenvalue of R^*A (or AR^*) of multiplicity m. Similarly a number $\lambda > 0$ is a principal value of A if and only if λ^2 is an eigenvalue of A^*A (or of AA^*) and the multiplicities are the same. It is easily seen that a number λ is a principal value if and only if the equation $Ax = \lambda Rx$ has a solution x with $Rx \neq 0$.

It follows from the last theorem that the principal values of A are the norms of the nonnull sections of A and in particular the norms of the principal sections of A. A section B of rank 1 of the principal section $A_i = \lambda_i R_i$ of A is expressible in the form $B = \lambda_i y x^*$, where x and y are unit vectors. A vector of the form $\alpha x\ (\alpha \neq 0)$ will be called a *principal vector* of A and a vector of the form $\beta y\ (\beta \neq 0)$ will be called a *reciprocal principal vector* of A corresponding to the principal value λ_i. It is easily seen that x is an eigenvector of A^*A and that y is an eigenvector of AA^*.

THEOREM 5.3. *A matrix A is normal if and only if its principal sections are normal. A matrix A is hermitian if and only if its principal sections are hermitian. A matrix A is hermitian and nonnegative if and only if its principal sections are hermitian and nonnegative.*

In order to prove this result let A and R be represented in the form (5.3) with $\lambda_1 > \lambda_2 > \cdots > \lambda_k$. Let

$$B = A - \lambda_k R = (\lambda_1 - \lambda_k)R_1 + \cdots + (\lambda_{k-1} - \lambda_k)R_{k-1}.$$

If A is normal so also are R and B. It follows that $S = R_1 + R_2 + \cdots + R_{k-1}$, the elementary matrix belonging to B, is also normal. This implies that $R_k = R - S$ and $A_k = \lambda_k R_k$ are normal. The same argument applied to B shows that R_{k-1} and $A_{k-1} = \lambda_{k-1}R_{k-1}$ are normal. It follows that each principal section $A_i = \lambda_i R_i\ (1 \leq i \leq k)$ of A is normal whenever A is normal. Conversely if $A_1, \cdots, A_k$ are normal so also is A. This proves the first statement in the theorem. The second statement can be proved similarly. The third statement is an easy consequence of the second and the concept of nonnegativeness.

THEOREM 5.4. *Let X and Y be elementary matrices such that the relation*

$$YY^*A = AXX^* = A$$

holds for a given matrix A. Then $B = Y^*AX$ has the same principal

values as those of A and have the same multiplicities. If R is the associated elementary matrix for A, then $S = Y^*RX$ is the associated elementary matrix for B. There exist elementary matrices X and Y such that B is a nonnegative diagonal matrix and $R = YX^*$.

Setting $S_i = Y^*R_iX$, $S = S_1 + \cdots + S_k$ it is found that $B = \lambda_1 S_1 + \cdots + \lambda_k S_k$. The ranks of S_i and R_i coincide and S_i $*$-commutes with S_j $(i \neq j)$. The last statement can be obtained by selecting a maximal set of mutually orthogonal principal vectors $x_1, \cdots, x_r$ of A of unit length and setting $y_h = Rx_h$ $(h = 1, \cdots, r)$. Let X be the matrix whose first r column vectors are $x_1, \cdots, x_r$ and the remaining vectors are null vectors. The matrix $Y = RX$ has $y_1, \cdots, y_r$ as its first r column vectors. It is easily seen that X and Y are elementary matrices of rank r having the properties described in the theorem. In fact the nonzero elements of $B = Y^*AX$ are the principal values of A. One could restrict X to have only r columns if one so desires. One could modify X and Y so as to be nonsingular. In this event we would have $R = YEX^*$, where $E = R^*R$. In either event the column vectors of $C = AX$ are mutually orthogonal and the lengths of the nonnull column vectors of C are the principal values of A. This fact can be used to devise a modified Jacobi method for finding the principal values of A. A discussion of a method of this type will be given by the author in a forthcoming paper.

6. Further properties of $*$-commutativity. Throughout the present section let A denote a given matrix and let R be its associated elementary matrix. Let

$$(6.1) \qquad A = \lambda_1 R_1 + \cdots + \lambda_k R_k , \qquad R = R_1 + \cdots + R_k$$

be its decomposition into principal sections, given in Theorem 5.2. As before we set

$$(6.2) \quad A_i = \lambda_i R_i , \quad E_i = R_i^* R_i , \quad E_i' = R_i R_i^* , \quad E = R^*R , \quad E' = RR^* .$$

The first result to be established is given in the following.

THEOREM 6.1. *If a matrix B $*$-commutes with A, then it $*$-commutes with every matrix of the form*

$$(6.3) \qquad C = \nu_1 R_1 + \cdots + \nu_k R_k$$

where $\nu_1, \cdots, \nu_k$ are real numbers. In particular B $*$-commutes with R and with each principal section A_j $(j = 1, \cdots, k)$. The matrix B is expressible uniquely as the sum

$$(6.4) \qquad B = B_0 + B_1 + \cdots + B_k$$

of $*$-orthogonal sections such that B_i is $*$-orthogonal to A_j $(j \neq i)$ and

$B_i \, (i > 0)$ *-commutes with A_i.

In order to prove this result we may suppose that the principal sections of A have been ordered so that $\lambda_1 < \lambda_2 < \cdots < \lambda_k$. Recall that, by virtue of the corollary to Theorem 4.1, the matrix B not only *-commutes with A but also with

$$A^{*-1} = (1/\lambda_1)R_1 + \cdots + (1/\lambda_k)R_k \,.$$

It follows that B *-commutes with

$$C_2 = A - \lambda_1^2 A^{*-1} = \lambda_{22}R_2 + \cdots + \lambda_{k2}R_k$$

where $\lambda_{i2} = \lambda_i - \lambda_1(\lambda_1/\lambda_i) > 0 \; (i = 2, \cdots, k)$. Moreover $\lambda_{i2} < \lambda_{j2} \, (i < j)$, as one readily verifies. Using the recursion formula

$$C_{j+1} = C_j - \lambda_{jj}^2 C_j^{*-1} \,, \qquad \lambda_{j+1,i} = \lambda_{ji} - \frac{\lambda_{jj}^2}{\lambda_{ji}} \qquad (i = j+1, \cdots, k)$$

one obtains matrices of the form

$$C_j = \lambda_{jj}R_j + \cdots + \lambda_{kj}R_k \qquad\qquad (j = 2, \cdots, k)$$

that *-commute with B. Moreover each R_i is a linear combination of $C_1 = A, C_2, \cdots, C_k$. It follows that B *-commutes with each of the sections $R_1, \cdots, R_k$ of R. Consequently B *-commutes with any matrix C of the form (6.3).

In order to prove that B is of the form (6.4) it follows from Theorem 4.1 with A replaced A_i that B is expressible in the form $B = C_i + B_i$ where C_i is *-orthogonal to A_i and $B_0 = B - (B_1 + \cdots + B_k)$ is *-orthogonal to each B_j and $A_j \, (j = 1, \cdots, k)$ and hence also to A. This completes the proof of the theorem.

THEOREM 6.2. *If a matrix B *-commutes with every section of A, then B is expressible in the form*

(6.5) $$B = B_0 + \mu_1 R_1 + \cdots + \mu_k R_k$$

*where $\mu_1, \cdots, \mu_k$ are real numbers and B_0 is *-orthogonal to A.*

Let $B_0, B_1, \cdots, B_k$ be the sections of B given in (6.4). Since every section of A_i and hence every section of R_i *-commutes with B and hence with B_i it follows from Theorem 4.3 that B_i is of the form $B_i = \mu_i R_i + B_{i0}$, where μ_i is a real number and B_{i0} is *-orthogonal to R_i. It is clear from the definition of B_i that B_{i0} must be zero. This proves the theorem.

THEOREM 6.3. *A matrix B *-commutes with A if and only if it *-commutes with R and $AR^*B = BR^*A$.*

If B *-commutes with A, then B *-commutes with R and

$$AR^*B = AB^*R = BA^*R = BR^*A \,.$$

Conversely, suppose that B *-commutes with R and that $AR^*B = BR^*A$. Then $A^*RB^* = B^*RA^*$ and

$$A^*B = R^*RA^*B = R^*AR^*B = R^*BR^*A = B^*RR^*A = B^*A$$
$$AB^* = RR^*AB^* = RA^*RB^* = RB^*RA^* = BR^*RA^* = BA^* ,$$

as was to be proved.

COROLLARY. *If A is a positive definite hermitian matrix, then B *-commutes with A if and only if B is hermitian and $AB = BA$.*

This result is immediate since $R = I$ for a positive definite matrix.

It should be observed that if A is hermitian but not definite, then there are nonhermitian matrices that *-commute with A. For example the matrices

$$A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

*-commute even though B is not hermitian. However A is elementary and $B = AB^*A$, that is, B is hermitian relative to A.

THEOREM 6.4. *Let A be a matrix and let T be an elementary matrix such that $TA^*T = A$. Let B be a matrix that *-commutes with T. Then B *-commutes with A if and only if $AT^*B = BT^*A$.*

The proof is similar to that of the last theorem and will be omitted.

THEOREM 6.5. *Given a matrix B that *-commutes with A there exists a set of mutually *-orthogonal elementary matrices $T_1, \cdots, T_q$ with the property that A and B are expressible in the form*

(6.6) $$A = \alpha_1 T_1 + \cdots + \alpha_q T_q , \qquad B = \beta_1 T_1 + \cdots + \beta_q T_q$$

where $\alpha_1, \cdots, \alpha_q$ are equal numbers, $\beta_1, \cdots, \beta_q$ are equal numbers and $\alpha_i = \alpha_j, \beta_i = \beta_j$ holds only in case $i = j$. If $\alpha_i \neq 0$, then $|\alpha_i|$ is a principal value of A. Similarly if $\beta_i \neq 0$ then $|\beta_i|$ is a principal value.

It is clear that T_i may be replaced by $-T_i$ in the theorem. The matrices $T_1, \cdots, T_q$ will be uniquely determined if, for example, one requires that $\beta_i \geqq 0$ and that $\alpha_i > 0$ if $\beta_i = 0$.

Let

(6.7) $$A = \lambda_1 R_1 + \cdots + \lambda_k R_k , \qquad B = \mu_1 S_1 + \cdots + \mu_m S_m$$

be the decompositions of A and B respectively into principal sections. Recall that by virtue of Corollary 1 to Theorem 4.2 the matrices R_i and S_j are expressible as sums

$$R_i = R_{ij0} + R_{ij+} + R_{ij-} , \qquad S_j = S_{ij0} + R_{ij+} - R_{ij-} .$$

Let $T_1, \cdots, T_p$ be all non-null elementary matrices R_{ij+} and R_{ij-} obtained

in this manner. Adjoin to these the maximal nonnull section of each R_i that is *-orthogonal to S and the maximal nonnull section of each S_j that is *-orthogonal to R. The elementary matrices $T_1, \cdots, T_q$ obtained in this manner are *-orthogonal and have the property that each R_i and S_j is expressible uniquely in the form

$$R_i = \rho_{i1} T_1 + \cdots + \rho_{iq} T_q , \qquad S_j = \sigma_{j1} T_1 + \cdots + \sigma_{jq} T_q$$

where $\rho_{ik} = 0$ if T_k is *-orthogonal to R_i, $\rho_{ih} = 1$ if T_h is a section of R_i, $\rho_{ih} = -1$ if $-T_h$ is a section of R_i, $\sigma_{jh} = 0$ if T_h is *-orthogonal to S_j, $\sigma_{jh} = 1$ if T_h is a section of S_j and $\sigma_{jh} = -1$ if $-T_h$ if a section of S_j. Combining this result with (6.7) one obtains (6.6). The last statement of the theorem follows from the construction just made.

As a consequence of this result we have

THEOREM 6.6. *Let T be an elementary matrix and let A be a matrix satisfying the condition $A = TA^*T$. Then A and T can be represented uniquely as the sum*

$$(6.7) \qquad A = \alpha_1 T_1 + \cdots + \alpha_q T_q , \qquad T = T_1 + \cdots + T_q$$

*of mutually *-orthogonal matrices such that $\alpha_i \neq \alpha_j$ $(i \neq j)$.*

This result follows from the last theorem with $B = T$ and the condition that $\beta_i \geqq 0$. Since no nonnull section of A is *-orthogonal to T we have $\beta_i = 1$, and the theorem follows.

If T is the identity then A is hermitian and $\alpha_1, \cdots, \alpha_q$ are the eigenvalues of A. The rank of T_i is the multiplicity of α_i as an eigenvalue of A. This result suggests that we call $\alpha_1, \cdots, \alpha_q$ the principal values or eigenvalues of A relative to T, the rank of T_i being the multiplicity of α_i.

As an extension of the last theorem we have

THEOREM 6.7. *Let T be an elementary matrix and let A and B be *-commutative matrices such that $A = TA^*T$ and $B = TB^*T$. There is a unique decomposition*

$$T = T_1 + \cdots + T_q$$

into sections such that A and B are representable in the form

$$A = \alpha_1 T_1 + \cdots + \alpha_q T_q , \qquad B = \beta_1 T_1 + \cdots + \beta_q T_q$$

where $\alpha_i = \alpha_j, \beta_i = \beta_j$ holds only in case $i = j$.

The proof of this result can be made by a simple modification of the proof of the last two theorems and will be omitted.

In the complex domain we have the following

COROLLARY 1. *Let T be an elementary matrix and let C be a*

matrix such that $TT^*C = CT^*T = C$ *and* $TC^*C = CC^*T$. *Then* C *and* T *have unique decompositions*

$$C = \gamma_1 T_1 + \cdots + \gamma_q T_q, \qquad T = T_1 + \cdots + T_q$$

in sections, where $\gamma_1, \cdots, \gamma_q$ *are distinct complex numbers.*

For, by the corollary to Theorem 4.4, the matrix C is expressible in the form $C = A + iB$, where A and B $*$-commute with T. From the relation $TC^*C = CC^*T$ it is found that $AT^*B = BT^*A$ and hence that A and B $*$-commute. The corollary follows from the last theorem with $\gamma_j = \alpha_j + i\beta_j$ $(j = 1, \cdots, q)$.

If $T = I$, the result described in the corollary yields the spectral decomposition for normal matrices.

By the use of an argument like that given in the proof of Theorem 5.4 one obtains the further result described in the following

COROLLARY 2. *Let* A *and* B *be* $*$-commutative matrices and let T *be an elementary matrix such that* $A = TA^*T$ *and* $B = TB^*T$. *There exist elementary matrices* X *and* Y *such that*

$$YY^*T = TXX^* = T$$

and such that Y^*AX, Y^*BX *are diagonal matrices. If* $C = A + iB$, *then* Y^*CX *is also a diagonal matrix.*

7. Certain classes of matrices.

Let $\mathscr{S}(A)$ be the class of matrices B that $*$-commute with A and have no non-null section that is $*$-orthogonal to A. Let $\mathscr{C}(A)$ be all matrices B such that $\mathscr{S}(A)$ is a subclass of $\mathscr{S}(B)$. It is clear that A is in $\mathscr{C}(A)$. We have the following

THEOREM 7.1. *Let* R *be the elementary matrix associated with* A *and let*

$$A = \lambda_1 R_1 + \cdots + \lambda_k R_k, \qquad R = R_1 + \cdots + R_k$$

be the decomposition of A *into principal sections. The class* $\mathscr{C}(A)$ *consists of all matrices* B *that are expressible in the form*

$$(7.1) \qquad\qquad B = \mu_1 R_1 + \cdots + \mu_k R_k$$

where $\mu_1, \cdots, \mu_k$ *are real numbers. If* B *is in* $\mathscr{C}(A)$ *so also is* B^{*-1} *and its associated elementary matrix* S.

This result follows from Theorems 6.1 and 6.2.

COROLLARY. *If* B *is in* $\mathscr{C}(A)$ *then* $\mathscr{C}(B) \subset \mathscr{C}(A)$. *Moreover* $\mathscr{C}(B) = \mathscr{C}(A)$ *if and only if* B *has the same number of distinct principal values as* A.

As a further result we have

THEOREM 7.2. *If* B, C, D *are matrices in* $\mathscr{C}(A)$ *so also is* $M = BC^*D$.

In fact

(7.2) $BC^*D = BD^*C = CB^*D = CD^*B = DB^*C = DC^*B$.

The relations (6.2) follow from the fact that B, C, D *-commute with each other. Observe that $M^* = B^*CD^*$. If N is a matrix in $\mathscr{E}(A)$ then

$$N^*M = N^*BC^*D = B^*NC^*D = B^*CN^*D = B^*CD^*N = M^*N .$$

Similarly $NM^* = MN^*$. This proves the theorem.

In view of the formula (7.1) one obtains the following

THEOREM 7.3. *Let*

(7.3) $B = \mu_1 R_1 + \cdots + \mu_k R_k$, $C = \nu_1 R_1 + \cdots + \nu_k R_k$

be two matrices in $\mathscr{E}(A)$. *Then*

$$\alpha B + \beta C = (\alpha\mu_1 + \beta\nu_1)R_1 + \cdots + (\alpha\mu_k + \beta\nu_k)R_k$$

is in $\mathscr{E}(A)$ *for every pair of real numbers* α *and* β. *If we define the product of B and C by the formula* $B{\cdot}C = BR^*C$, *then*

$$B{\cdot}C = \mu_1\nu_1 R_1 + \cdots + \mu_k\nu_k R_k$$

is in $\mathscr{E}(A)$ *and the usual laws of algebra hold. In particular* $B{\cdot}R = R{\cdot}B = B$. *Given a polynomial*

$$p_m(\lambda) = \alpha_0 + \alpha_1\lambda + \cdots + \alpha_m\lambda^m$$

with real coefficients set

$$p_m(A, R) = \alpha_0 A^{(0)} + \alpha_1 A^{(1)} + \cdots + \alpha_{\hat{n}} A^{(n)} ,$$

where $A^{(0)} = R, A^{(1)} = A, A^{(h)} = A{\cdot}A^{(h-1)}$. *Then the polynomial*

$$p_m(A, R) = p_m(\lambda_1)R_1 + \cdots + p_m(\lambda_k)R_k$$

in A relative to R is in $\mathscr{E}(A)$. *Conversely every matrix B in* $\mathscr{E}(C)$ *is expressible as a real polynominal in A relative to R of degree* $\leq k - 1$. *There is a unique polynomial* $p_k(\lambda)$ *with leading coefficient* $\alpha_k = 1$ *such that* $p_k(A, R) = 0$.

The first three statements in the lemma are immediate. The matrix B is given by the relative polynomial $p_{k-1}(A, R)$ whose coefficients $\alpha_0, \alpha_1, \cdots, \alpha_{k-1}$ are given by the solutions of the equations

$$\alpha_0\lambda_1^h + \alpha_1\lambda_2^h + \cdots + \alpha_{k-1}\lambda_k^h = \mu_h \qquad (h = 0, 1, \cdots, k - 1) .$$

Finally the polynomial $p_k(\lambda)$ described in the last statement in the theorem is the polynomial of degree k whose roots are $\lambda_1, \lambda_2, \cdots, \lambda_k$.

COROLLARY. *On the class* $\mathscr{E}(A)$ *the norm* $\| B \|$ *satisfies the rela-*

tion $\|B \cdot C\| \leqq \|B\| \|C\|$.

As a final property of the classes $\mathscr{C}(A)$ we have

THEOREM 7.4. *If B is in $\mathscr{S}(A)$ there is a matrix C such that A and B are in $\mathscr{C}(C)$.*

This result follows from Theorem 6.5 with

$$C = T_1 + 2T_2 + 3T_3 + \cdots qT_q .$$

Let T be a given elementary matrix and let A be a matrix in $\mathscr{S}(T)$. Let $\mathscr{S}(A, T)$ be all matrices in $\mathscr{S}(T)$ that $*$-commute with A. Let $\mathscr{C}(A, T)$ be all matrices B in $\mathscr{S}(A, T)$ such that $\mathscr{S}(A, T) \subset \mathscr{S}(B, T)$. If T is the elementary matrix R associated with A then $\mathscr{C}(A, T) = \mathscr{C}(A)$. Let

$$A = \alpha_1 T_1 + \cdots + \alpha_q T_q , \qquad T = T_1 + \cdots T_q$$

be the decomposition of A and T given in Theorem 6.6. Then, as is easily seen, the class $\mathscr{C}(A, T)$ consists of all matrices of the form

$$B = \beta_1 T_1 + \cdots + \beta_q T_q$$

where $\beta_1, \cdots, \beta_q$ are real numbers. If we set

$$C = T_1 + 2T_2 + \cdots + qT_q$$

then the class $\mathscr{C}(A, T)$ coincides with the class $\mathscr{C}(C)$. Consequently the results stated above are applicable to the class $\mathscr{C}(A)$. If T is the identity, then A is hermitian and $\mathscr{C}(A, T)$ consists of all hermitian matrices that commute with every hermitian matrix that commutes with A.

Consider now an elementary matrix T and let $\mathfrak{N}(T)$ be the class of all matrices A such that $TT^*A = AT^*T = A$ and $AA^*T = TA^*A$. Given a matrix A in $\mathfrak{N}(T)$ let $\mathfrak{M}(A, T)$ be the class of all matrices B in $\mathfrak{N}(T)$ such that $AT^*B = BT^*A$ and $AB^*T = TB^*A$. If B is in $\mathfrak{M}(A, T)$ then $BA^*T = TA^*B$ also. Moreover, TB^*T is in $\mathfrak{M}(A, T)$. Let $\mathscr{B}(A, T)$ be the class of all matrices B such that $\mathfrak{M}(B, T) \supset \mathfrak{M}(A, T)$. In view of Corollary 1 to Theorem 6.7 the matrices A and T are expressible uniquely in the form

$$A = \alpha_1 T_1 + \cdots + \alpha_q T_q , \qquad T = T + \cdots + T_q ,$$

where $\alpha_1, \alpha_2, \cdots, \alpha_q$ are distinct complex numbers. It is not difficult to show a matrix B is in $\mathscr{B}(A, T)$ if and only if it is expressible in the form

$$B = \beta_1 T_1 + \cdots + \beta_q T_q ,$$

where $\beta_1, \cdots, \beta_q$ are complex numbers. If B and C are in $\mathscr{B}(\dot{A}, T)$ so

also are $\alpha B + \beta C$, where α and β are complex numbers. Moreover, the product $B \cdot C = BT^*C$ is in $\mathscr{B}(A, T)$. If polynomials of A relative to T are defined as before, but with complex coefficients, it is seen that the class $\mathscr{B}(A, T)$ is made up of all polynomials of A relative to T of degree $\leqq q - 1$. Again we have the relation $\| B \cdot C \| \leqq \| B \| \| C \|$. These results generalize the corresponding theory for normal matrices.

University of California, Los Angeles

PACIFIC JOURNAL OF MATHEMATICS

Pacific Journal of Mathematics

Vol. 11, No. 1 November, 1961