NONSYMMETRIC PROJECTIONS IN HILBERT SPACE

V. J. MIZEL[1] AND M. M. RAO[2]

0. Introduction. An initial investigation into the kind of operators which can be obtained as the difference of two projections led to the study presented below. In this paper a characterization is given for the general (not necessarily symmetric) bounded linear idempotent operator, or projection, on Hilbert space. These results are applied to the investigation of a projection problem and to a "weak" ordering of such operators. The paper falls naturally into two parts. In the first we give two theorems and several more or less direct consequences which together provide the characterization. In the second part we apply these results to the investigation and solution of a problem which is of importance in probability and statistics. A sketch of the role of this problem in statistical theory and an examination of how our results fit in with previous conclusions complete the present study.

We mention that Dixmier [3] has done work related to the first part, obtaining results of an entirely different nature from ours. So far as we know the point of view presented here does not appear in the literature.

1. Characterization theorems. We utilize the following compressed notation: "positive" for "positive semi-definite", "s.a." for "self adjoint", "skew" for "skew-adjoint". $A \smile B$ indicates that A and B commute, $A \mid \mathscr{V}$ stands for the restriction of the (always linear) operator A to the subspace $\mathscr{V}$, and $\mathscr{R}_A$, $\mathscr{N}_A$ respectively denote the range and the null space of A. The terminology used below is that of complex Hilbert space but as is made clear in the proofs our results apply to the real case as well. In other respects, notation is mostly patterned after that of Riesz and Sz.-Nagy [9].

THEOREM 1. *An operator P on a Hilbert space $\mathfrak{X}$ is a projection (bounded idempotent linear operator) if, and only if, there exist*
(I) *a bounded s. a. operator S such that $S^2 - S$ is positive*
(II) *a unitary operator U, with $U(\overline{\mathscr{R}}_{S^2-S}) \subset \overline{\mathscr{R}}_{S^2-S}$, whose restriction to $\overline{\mathscr{R}}_{S^2-S}$ satisfies*

(i) $\quad U^2 = -I$

Received February 6, 1961.
Presented to the American Mathematical Society on January 23, 1961.

[1] Supported under contract no. DA-36-061-ORD-490.

[2] Supported under contract no. AF 33(616)-3878.

(ii) $SU = U(I - S)$,

in terms of which the following representation holds:

$$P = S + U(S^2 - S)^{(1/2)} \tag{$*$}$$

Proof.[3] We first recall the relation $\overline{\mathscr{R}}_A = \overline{\mathscr{R}}_{A^2}$ for a bounded s.a. operator A (the 'bar' indicates closure). This is utilized, with $A = (S^2 - S)^{(1/2)}$, both in the present proof and later. To show sufficiency we note that since $S \smile (S^2 - S)$ and $(I - S) \smile (S^2 - S)$ there follows from (II) (ii) the relation $U \smile S^2 - S$ on $\overline{\mathscr{R}}_{S^2-S}$. Thus by the preceding remark $U \smile (S^2 - S)^{(1/2)}$ on $\overline{\mathscr{R}}_{S^2-S}$. In this way we obtain:

$$\begin{aligned}[S + U(S^2 - S)^{(1/2)}]^2 &= S^2 + (SU + US)(S^2 - S)^{(1/2)} \\ &\quad + U^2(S^2 - S) = S + U(S^2 - S)^{(1/2)}.\end{aligned}$$

Therefore idempotence of the bounded operator P is proved.

To prove necessity we decompose the projection P into its s.a. and skew parts, S and W respectively. Then idempotence of P leads to the equation $(S + W)^2 = S + W$, which on rearrangement gives

$$S^2 + W^2 - S = W - SW - WS . \tag{1.1}$$

The left hand side is s.a. while the right hand side is skew. Hence (1.1) is equivalent to the pair of equations

$$S^2 - S = - W^2 = WW^* \tag{1.2}$$

$$SW = W(I - S) . \tag{1.3}$$

From (1.2) we conclude that the polar decomposition of W is of the form

$$W = U(S^2 - S)^{(1/2)} . \tag{1.4}$$

Since W is normal it is well-known ([9], p. 286) that the partially isometric operator U is in fact unitary and that $U \smile (S^2 - S)^{(1/2)}$. Since W is skew we have in addition the relation

$$U(S^2 - S)^{(1/2)} = - [U(S^2 - S)^{(1/2)}]^* = - (S^2 - S)^{(1/2)}U^{-1} . \tag{1.5}$$

These facts ensure that

$$U^2(S^2 - S)^{(1/2)} = - (S^2 - S)^{(1/2)} \tag{1.6}$$

and so by our previous remark yield (II) (i). Finally, (1.3) gives the relation, since $S \smile (S^2 - S)^{(1/2)}$,

$$SU(S^2 - S)^{(1/2)} = U(I - S)(S^2 - S)^{(1/2)} \tag{1.7}$$

which, for the same reason, yields (II) (ii). Q.E.D.

---

[3] We are grateful to Professor P.R.Halmos for bringing to our attention this simplified form of our original proof.

An equivalent form of the preceding result, which explicitly involves the spectral theorem, follows:

THEOREM 1′. *An operator P on a Hilbert space $\mathfrak{X}$ is a projection if, and only if, there exist*

(I′) *a resolution of the identity $\{E_\lambda\}$ varying only for λ in the set $[-\alpha, 0] \cup [1, 1+\alpha]$ for some $\alpha \geqq 0$ and having the property that $\lambda \neq 0$ or 1 is a point of continuous [discontinuous] growth precisely when $1-\lambda$ is such a point,*

(II′) *a unitary operator U satisfying*

(i) $U^2(E_\lambda - E_\mu) = E_\mu - E_\lambda$ *when $[\lambda, \mu]$ and $(0, 1)$ are disjoint*

(ii) $UE_\lambda = (I - E_{(1-\lambda)-})U$ *when* $\lambda \notin [0, 1]$,

in terms of which the following representation holds:

$$(*') \qquad P = \int_{(-\alpha)-}^{1+\alpha} \lambda\, dE_\lambda + \int_{(-\alpha)-}^{1+\alpha} (\lambda^2 - \lambda)^{(1/2)} d\, UE_\lambda \,.$$

REMARK. These hypotheses are essentially direct translations, with use of the spectral theorem, of the hypotheses of Theorem 1. The only feature of this equivalence which is not straightforward is (II′) (ii). The point here is that when a decreasing sequence of polynomials $\{p_n(\lambda)\}$ is such that the operators $\{p_n(S)\}$ converge strongly to the projection $E_{\lambda'}$ then the operators $\{p_n(I-S)\}$ formed of the same polynomials converge strongly to the projection $I - E_{(1-\lambda')-}$.

It is of interest to state explicitly what the implications of Theorem 1′ are for the finite dimensional case. In a finite dimensional vector space the notion of projection is of course a purely algebraic one so that any description involving an inner product is in a sense over-elaborate.

The term "partition of the identity" as used below refers to a family $\{H_1, \cdots, H_m\}$ of idempotent operators whose sum is I and which satisfy $H_iH_j = 0$ for $i \neq j$. The partition is "symmetric" if all the H_i are symmetric, or s.a.

COROLLARY. *An operator P on a finite dimensional vector space $\mathscr{V}$ is a projection if, and only if, there exists for each inner product $\mathscr{I}$ on $\mathscr{V}$*

(I″) *a symmetric partition of the identity $\mathscr{P} = \{G_0, G_1, F_1, \cdots, F_{2k}\}$ and a set of constants $\{\lambda_i\}$ with $\lambda_1 > \cdots > \lambda_k > 1 > 0 > 1 - \lambda_k > \cdots > 1 - \lambda_1$,*

(II″) *an isometry U satisfying*

(i) $U^2F_j = -F_j \qquad j = 1, \cdots, 2k$

(ii) $UF_j = F_{2k-j}U \qquad j = 1, \cdots, 2k$,

in terms of which the following equation holds:

$$(*'') \qquad P = 0 \cdot G_0 + 1 \cdot G_1 + \sum_{j=1}^{k} [\lambda_j F_j + (1 - \lambda_j) F_{2k-j}] + \sum_{j=1}^{k} (\lambda_j^2 - \lambda_j)^{(1/2)} [UF_j + UF_{2k-j}] \, .$$

REMARKS 1. The symmetric partition $\mathscr{P}$, the numbers λ_i, and the isometry U all vary with the choice of the inner product $\mathscr{I}$. Even the integer k depends on this choice. In particular, since there exist inner products in which $\mathscr{R}_P$ and $\mathscr{N}_P$ are orthogonal subspaces so that P is symmetric, we see that $\min_{\mathscr{I}} \{k\} = 0$. On the other hand $\max_{\mathscr{I}} \{k\} = k_m \equiv \min \{\dim \mathscr{R}_P, \dim \mathscr{N}_P\}$. The reasoning is as follows, where we suppose, say, $k_m = \dim \mathscr{N}_P \leqq \dim \mathscr{R}_P$. To each independent family $\{\vec{v}_1, \cdots, \vec{v}_{k'}\} \subset \mathscr{R}_P$ and $\{\vec{w}_1, \cdots, \vec{w}_{k'}\} \subset \mathscr{N}_P$ there are inner products in which the subspaces $\mathscr{V}_j$ spanned by the pairs $\vec{v}_j, \vec{w}_j$, $j = 1, \cdots, k'$, are orthogonal to one another and to mutually orthogonal subspaces $\mathscr{V}_0 \subset \mathscr{R}_P$, and $\tilde{\mathscr{V}}_0 \subset \mathscr{N}_P$, where $\mathscr{V}_0 + \tilde{\mathscr{V}}_0 + \mathscr{V}_1 + \cdots + \mathscr{V}_{k'} = \mathscr{V}$. Each $\mathscr{V}_j$, $j = 1, \cdots, k'$, is then an invariant two-dimensional manifold for the projections P and P^*, and it is a matter of computation to show that these manifolds correspond to distinct sets of values $\{\lambda, 1 - \lambda\}$ provided that the inner product $\mathscr{I}$ is so chosen that the pairs $\vec{v}_j, \vec{w}_j$ determine distinct angles $\theta_j < (\pi/2)$, $j = 1, \cdots, k'$. Thus for each such choice of $\mathscr{I}$ we have $k = k'$. The equation $k_m \leqq \max_{\mathscr{I}} \{k\}$ is in keeping with the fact that k_m as defined is the largest k' value. We omit the proof of the reversed inequality. These matters are related to some work of Seidel [12].

2. Another feature brought out by the corollary is that, no matter which inner product is utilized, the symmetric part of P will have a spectrum symmetrically located outside $[0, 1]$, if the possible eigenvalues $\lambda = 0$, $\lambda = 1$ are excluded. Moreover λ_j and $1 - \lambda_j$ will have equal multiplicity. Theorem 2 will show that these are the only conditions needed on S.

3. A further consequence of the corollary is the following result on canonical forms: if P is an $N \times N$ idempotent matrix then P is unitarily equivalent to a matrix of the following form (I_j denotes the $j \times j$ identity matrix, O_j denotes the $j \times j$ zero matrix):

$$P \longleftrightarrow \begin{pmatrix} O_m & & & & \\ & I_n & & & \\ & & Z_1 & & \\ & & & \ddots & \\ & & & & Z_k \end{pmatrix}$$

where the matrix Z_j is of size $2s_j \times 2s_j$, with $s_j = \dim \mathscr{R}_{F_j} = \dim$

$\mathscr{R}_{F_{2k-j}}$, $j = 1, \cdots, k$, and has the form

$$Z_j = \left(\begin{array}{c|c} \lambda_j I_{s_j} & (\lambda_j^2 - \lambda_j)^{(1/2)} I_{s_j} \\ \hline -(\lambda_j^2 - \lambda_j)^{(1/2)} I_{s_j} & (1 - \lambda_j) I_{s_j} \end{array}\right)$$

It follows from (II) (ii) of Theorem 1 that $S \mid \overline{\mathscr{R}}_{S^2-S}$ and $(I - S) \mid \overline{\mathscr{R}}_{S^2-S}$ are unitarily equivalent operators. A natural question is what further knowledge about a bounded s.a. operator is needed in order to conclude that it is the s.a. part of a projection. This is answered by

THEOREM 2. *A bounded s.a. operator S on a Hilbert space $\mathfrak{X}$ is the s.a. part of a projection if, and only if,*
(I_2) $S^2 - S$ *is positive*
(II_2) $S \mid \overline{\mathscr{R}}_{S^2-S}$ *and* $(I - S) \mid \overline{\mathscr{R}}_{S^2-S}$ *are unitarily equivalent.*

Discussion. We adhere to the practice of using the term "unitary" to refer to an onto isometric operator even for real Hilbert space. With this understanding Theorem 2 is true whether $\mathfrak{X}$ is real or complex. However, since we will utilize the spectral theorem for unitary operators in our proof, it will be necessary in the former case to deal with a complexification $\hat{\mathfrak{X}}$ of $\mathfrak{X}$ and, more particularly, with $\hat{\overline{\mathscr{R}}}_{S^2-S}$, the closed subspace of $\hat{\mathfrak{X}}$ generated by $\overline{\mathscr{R}}_{S^2-S}$. When $\mathfrak{X}$ is itself complex we have $\hat{\overline{\mathscr{R}}}_{S^2-S} = \overline{\mathscr{R}}_{S^2-S}$.

Proof. In one direction the statement is a consequence of Theorem 1. For the other direction we proceed as follows. The hypothesis (II_2) ensures that there is a unitary operator V on $\overline{\mathscr{R}}_{S^2-S}$ such that

$$SV = V(I - S) , \tag{1.8}$$

but it need not be true that $V^2 = -I$ as is required in Theorem 1. Using $\hat{V}$ to denote operator V extended to $\hat{\overline{\mathscr{R}}}_{S^2-S}$ but retaining "S" rather than $\hat{S}$ for the extension of S, we have (argument also applies for complex $\mathfrak{X}$):

$$S\hat{V} = \hat{V}(I - S) . \tag{$\widehat{1.8}$}$$

Moreover, there follows from ($\widehat{1.8}$)

$$\hat{V}^2 \smile S \text{ on } \hat{\overline{\mathscr{R}}}_{S^2-S} . \tag{1.9}$$

Now use of the functional calculus for the unitary operator $\hat{V}^2$ $\left[u(\lambda) \longleftrightarrow u(\hat{V}^2) = \int_0^{2\pi} u(e^{i\phi}) dE_\phi\right]$ permits us to deal with the operator

$\tilde{V} = (-\hat{V}^2)^{(1/2)}$ corresponding to the periodic function $u(e^{i\phi}) = (-e^{i\phi})^{(1/2)} = e^{i(\phi/2+\pi/2)}$, $0 < \phi \leqq 2\pi$ (E_ϕ is taken to be continuous from the right at $\phi = 0$). As is well known ([9], p. 343) $\tilde{V}$ then has the property that $\tilde{V} \smile A$ for every bounded operator A such that $A \smile \hat{V}^2$, and the same property holds for $\tilde{V}^*$. In particular

$$\tilde{V}^* \smile S\,;\ \tilde{V}^* \smile \hat{V}\,;\ \tilde{V}^* \smile \hat{V}^*\,. \tag{1.10}$$

Moreover $\tilde{V}$ and $\tilde{V}^*$ are unitary since $\tilde{V}\tilde{V}^*$ corresponds to the function $u(e^{i\phi})\bar{u}(e^{i\phi})^{(1/2)} = (-e^{i\phi})^{(1/2)}\overline{(-e^{i\phi})^{(1/2)}} = 1$. Finally, it can be easily verified that $u(e^{i\phi}) = (-e^{i\phi})^{(1/2)}$ is a limit of *real* polynomials in $e^{i\phi}$ and $e^{-i\phi}$, so that $\tilde{V}$ and $\tilde{V}^*$ are limits of *real* polynomials in $\hat{V}^2$ and $(\hat{V}^2)^*$. It therefore follows, even when $\mathfrak{X}$ is real, that $\tilde{V} \mid \overline{\mathscr{R}}_{S^2-S}$ and $\tilde{V}^* \mid \overline{\mathscr{R}}_{S^2-S}$ are unitary transformations of this subspace.

Define U on $\overline{\mathscr{R}}_{S^2-S}$ as follows:

$$U = \tilde{V}^*\hat{V}\,. \tag{1.11}$$

Then U is unitary and in addition satisfies:

$$U^2 = (\tilde{V}^*)^2\hat{V}^2 = (-\hat{V}^2)^*\hat{V}^2 = -I, \text{ and} \tag{1.12}$$

$$SU = \tilde{V}^*S\hat{V} = \tilde{V}^*\hat{V}(I - S) = U(I - S)\,. \tag{1.13}$$

Extend U to $\mathfrak{X}$ by choosing an arbitrary unitary operator on $\mathscr{N}_{(S^2-S)|\mathfrak{X}}$ and extending linearly.

The conditions of Theorem 1 are all met by the operator $S + U(S^2 - S)^{(1/2)}$, so a projection whose s.a. part is the given operator S has been constructed.

The class of s.a. parts of projections is a rather large one, as is seen from the following.

COROLLARY. *Let σ_0 denote a compact subset of $(-\infty, 0] \cup [1, \infty)$ which, except possibly for $\lambda = 0$ or $\lambda = 1$, is invariant under the transformation $\lambda \to 1 - \lambda$. Then there exists a projection P, defined on some Hilbert space $\mathfrak{X}$, whose s.a. part S has spectrum σ_0. Furthermore, if T denotes any s.a. operator with spectrum $\sigma_0^- = \sigma_0 \cap (-\infty, 0]$ then the projection P may be chosen so that S^-, the negative part of S, is unitarily equivalent to $-T$.*

REMARK. $\mathfrak{X}$ may be chosen as the product Hilbert space $\overline{\mathscr{R}}_T \times \overline{\mathscr{R}}_T$. S is defined to be $-T$ on one copy of $\overline{\mathscr{R}}_T$ and to be $I + T$ on its orthogonal complement and U is defined on the first copy to be the negative of the canonical mapping between these subspaces while on the orthogonal complement it is taken equal to this mapping.

2. An ordering. We begin this section by introducing a partial ordering whose properties we propose to investigate.

DEFINITION 1. The relation $P_1 \succ P_2$ between two projections on a real Hilbert space $\mathfrak{X}$ signifies that the quadratic form based on the (not necessarily symmetric) operator $A = P_1 - P_2$ is nonnegative: Using inner product notation, this means $(x, Ax) \geqq 0$, for all $x \in \mathfrak{X}$.

On a complex Hilbert space $P_1 \succ P_2$ signifies: $Re\{(x, Ax)\} \geqq 0$; for all $x \in \mathfrak{X}$.

We mention that $\succ$ is a partial ordering in a " weak " sense, for it is transitive but not anti-symmetric:

$$P_1 \succ P_2,\ P_2 \succ P_3 \Rightarrow P_1 \succ P_3 \tag{2.1}$$

$$P_1 \succ P_2,\ P_2 \succ P_1 \not\Rightarrow P_1 = P_2 . \tag{2.2}$$

This ordering further differs from the usual partial ordering for projection operators ([7]) in that the relation $0 \prec P \prec I$ is not universal: it holds only when P is symmetric. Our interest in this analytically—rather than geometrically—motivated ordering arises from considerations in probability which will be discussed later.

For reasons which will soon appear we find it convenient to single out a certain subclass of the projection operators as follows.

DEFINITION 2. $\mathscr{C}$ denotes that class of projections P in the Hilbert space $\mathfrak{X}$ whose members posess the property

(c.c.) $S = (P + P^*)/2$ has a compact (also called completely continuous) negative part.

We now discuss certain consequences of membership in the class $\mathscr{C}$. We find by utilizing the decomposition $S = S^+ - S^-$ that the positiveness of $S^2 - S$ is equivalent to that of $(S^+)^2 - S^+$. Denote by E_+ the symmetric projection onto $\overline{S^+(\mathscr{R}_{S^2-S})}$ and by E_- the symmetric projection onto $\overline{\mathscr{R}}_{S^2-S} \ominus \overline{S^+(\mathscr{R}_{S^2-S})}$. Then on $\overline{\mathscr{R}}_{S^2-S}$: $(S^+)^2 - S^+ = (S^+ - E_+)S^+$, from which it follows that $S^+ - E_+$ is positive on $\overline{(S^+)^{(1/2)}(\mathscr{R}_{S^2-S})} = \overline{S^+(\mathscr{R}_{S^2-S})}$ and therefore that $S^+ - E_+$ is positive on $\overline{\mathscr{R}}_{S^2-S}$. In addition, the above decomposition for S leads on $\overline{\mathscr{R}}_{S^2-S}$ to the relation $I - S = (E_+ - S^+) + (E_- + S^-)$. Therefore we can also conclude that on $\overline{\mathscr{R}}_{S^2-S}$, $(I - S)^- = S^+ - E_+$. Now S and $I - S$ are unitarily equivalent on $\overline{\mathscr{R}}_{S^2-S}$, so it follows in particular that S^- and $(I - S)^-$ are unitarily equivalent on this subspace. Hence, the decomposition

$$S = (S^+ - E_+) + E_+ - S^- , \qquad \text{on } \overline{\mathscr{R}}_{S^2-S} , \tag{2.3}$$

is a decomposition into positive operators, with $S^+ - E_+$ and S^- being unitarily equivalent. On the other hand S behaves on $\mathscr{N}_{S^2-S}$ as a symmetric projection since $Sx = S^2x$ when $x \in \mathscr{N}_{S^2-S}$. Hence $\mathscr{N}_{S^2-S}$ is the direct sum of orthogonal subspaces (possibly trivial) on which S has the eigenvalues $\lambda = 0$ and $\lambda = 1$, respectively.

Now when $P \in \mathscr{C}$, S^- is compact as well as positive so that, with the possible exception of $\lambda = 0$, S^- has a pure point spectrum consisting of distinct positive eigenvalues $\mu_1 > \mu_2 > \cdots$, and the corresponding eigenspaces $\mathscr{V}_1, \mathscr{V}_2, \cdots$ are finite dimensional. In addition, the eigenvalues μ_j occur as successive maxima of the quadratic form (x, S^-x), ([9], p. 233). From the decomposition (2.3) we therefore can conclude that, with the possible exception of $\lambda = 0$ and $\lambda = 1$, S itself has a pure point spectrum consisting of the distinct eigenvalues $-\mu_1 < -\mu_2 < \cdots$ (negative) and $1 + \mu_1 > 1 + \mu_2 > \cdots$ (positive), and the corresponding eigenspaces $\mathscr{V}_1, \mathscr{V}_2, \cdots$ and $\tilde{\mathscr{V}}_1, \tilde{\mathscr{V}}_2, \cdots$ are finite dimensional subspaces of $\bar{\mathscr{R}}_{S^2-S}$. Moreover the eigenvalues $\{-\mu_j\}$ and $\{1 + \mu_j\}$ occur as successive minima and successive maxima, respectively, of the quadratic form (x, Sx). *A* further consequence of the unitary equivalence on $\bar{\mathscr{R}}_{S^2-S}$ of S^- and $S^+ - E_+$ is the relation

$$\dim \mathscr{V}_j = \dim \tilde{\mathscr{V}}_j, \qquad j = 1, 2, \cdots. \tag{2.4}$$

Denoting by S_j the s.a. part of P_j, $j = 1, 2$, and by $\{E_\lambda^{(j)}\}$ the spectral family of S_j, we have

THEOREM 3. *If $P_1, P_2 \in \mathscr{C}$ then the following conditions are equivalent*:

(I_3) $P_1 \succ P_2$

(II_3) $S_1^2 - S_1 = S_2^2 - S_2 = Q$, *with*

(i) $S_1 = S_2$ *on* $\mathscr{R}_Q$

(ii) $S_1(\mathscr{N}_Q) \subset S_2(\mathscr{N}_Q)$

(III_3) $E_\lambda^{(1)} = E_\lambda^{(2)}$ $\lambda \notin [0, 1)$, $\mathscr{R}_{E_0^{(1)}} \subset \mathscr{R}_{E_0^{(2)}}$.

Proof. The proof will be given in the order $(\mathrm{I}_3) \Rightarrow (\mathrm{II}_3) \Rightarrow (\mathrm{III}_3) \Rightarrow (\mathrm{I}_3)$.

a. $(I_3) \Rightarrow (II_3)$. The hypothesis $Re\ \{(x, (P_1 - P_2)x)\} \geqq 0$ for all x is equivalent to

$$(x, S_1x) \geqq (x, S_2x) \text{ for all } x. \tag{2.5}$$

We apply (2.5) to show that the eigenvalues $\{-\mu_j^{(1)}\}$, $\{1 + \mu_j^{(1)}\}$ and eigenspaces $\{\mathscr{V}_j^{(1)}\}$, $\{\tilde{\mathscr{V}}_j^{(1)}\}$ for S_1 are respectively identical to the eigenvalues $\{-\mu_j^{(2)}\}$, $\{1 + \mu_j^{(2)}\}$ and eigenspaces $\{\mathscr{V}_j^{(2)}\}$, $\{\tilde{\mathscr{V}}_j^{(2)}\}$ for S_2.

First, it is an immediate consequence of (2.5) that the following relation holds between the maximum eigenvalues of S_1 and S_2.

$$1 + \mu_1^{(1)} \geqq 1 + \mu_1^{(2)} \tag{2.6$_1$}$$

(simply consider (2.5) when $x \in \tilde{\mathscr{V}}_1^{(2)}$). On the other hand, by taking

$x \in \mathscr{V}_1^{(1)}$ we conclude:

$$(2.6_1') \qquad - \mu_1^{(1)} \geqq - \mu_1^{(2)} .$$

These inequalities ensure that $\mu_1^{(1)} = \mu_1^{(2)}$. What is more, we then observe by means of the argument used in deriving these inequalities that $\tilde{\mathscr{V}}_1^{(2)}$ is a subspace of $\tilde{\mathscr{V}}_1^{(1)}$ and that $\mathscr{V}_1^{(1)}$ is a subspace of $\mathscr{V}_1^{(2)}$. Reference to (2.4) now leads to: $\mathscr{V}_1^{(1)} = \mathscr{V}_1^{(2)}, \tilde{\mathscr{V}}_1^{(1)} = \tilde{\mathscr{V}}_1^{(2)}$, and therefore all the desired relations between $\{-\mu_j^{(1)}\}$, $\{1 + \mu_j^{(1)}\}$, $\{\mathscr{V}_j^{(1)}\}$, $\{\tilde{\mathscr{V}}_j^{(1)}\}$ and $\{-\mu_j^{(2)}\}$, $\{1 + \mu_j^{(2)}\}$, $\{\mathscr{V}_j^{(2)}\}$, $\{\tilde{\mathscr{V}}_j^{(2)}\}$ have been established for the case $j = 1$.

In general, $1 + \mu_j^{(i)}$ $(i = 1, 2)$ is the maximum of the quadratic form $(x, S_i x)$ among unit vectors x orthogonal to the subspaces $\tilde{\mathscr{V}}_1^{(i)}, \cdots, \tilde{\mathscr{V}}_{j-1}^{(i)}$, while $-\mu_j^{(i)}$ is the minimum of the quadratic form among unit vectors x orthogonal to the subspaces $\mathscr{V}_1^{(i)}, \cdots, \mathscr{V}_{j-1}^{(i)}$, so we can reproduce the argument of the preceding paragraph to obtain the inequalities

$$(2.6_j) \qquad 1 + \mu_j^{(1)} \geqq 1 + \mu_j^{(2)}$$

$$(2.6_j') \qquad - \mu_j^{(1)} \geqq - \mu_j^{(2)}$$

(for the first inequality take $x \in \tilde{\mathscr{V}}_j^{(2)}$, for the second take $x \in \mathscr{V}_j^{(1)}$). The desired relations of eigenspaces and eigenvalues then follow in the same manner as before.

The results obtained above lead by application of the spectral theorem to the conclusion $S_1^2 - S_1 = S_2^2 - S_2 = Q$. Since $\{\mathscr{V}_j^{(i)}\}$, $\{\tilde{\mathscr{V}}_j^{(i)}\}$, $i = 1, 2$, span $\overline{\mathscr{R}}_Q$, it remains only to consider the behavior of $S_i (i = 1, 2)$ on $\mathscr{N}_Q$. The fact that S_i is a symmetric projection on $\mathscr{N}_Q$ taken together with (2.5) yields the final relation (II_3) (ii). This completes the proof of a.

b. $(II_3) \Rightarrow (III_3)$. We have seen earlier that the spectrum of S_i $(i = 1, 2)$ lies outside (0, 1). Therefore, since the spectrum of S_i on $\mathscr{N}_Q$ consists at most of the points $\lambda = 0$ and $\lambda = 1$, the relation (II_3) (i) has as an immediate consequence $E_\lambda^{(1)} = E_\lambda^{(2)}$, $\lambda \notin [0, 1]$. The normalization of the $E_\lambda^{(i)}$ as right-continuous families then gives the desired relation: $E_\lambda^{(1)} = E_\lambda^{(2)}$, $\lambda \notin [0, 1)$, i.e., $\lambda = 1$ is included.

As a consequence of the above result we have $E_1^{(1)} - E_{0-}^{(1)} = E_1^{(2)} - E_{0-}^{(2)}$. Since $E_1^{(i)} - E_{0-}^{(i)} = (E_1^{(i)} - E_{1-}^{(i)}) + (E_0^{(i)} - E_{0-}^{(i)})$ we see that the relation (II_3) (ii) for $\mathscr{N}_Q$, which requires that $\mathscr{R}_{E_1^{(1)} - E_{1-}^{(1)}} \supset \mathscr{R}_{E_1^{(2)} - E_{1-}^{(2)}}$, implies $\mathscr{R}_{E_0^{(1)} - E_{0-}^{(1)}} \subset \mathscr{R}_{E_0^{(2)} - E_{0-}^{(2)}}$. Combining this with $E_{0-}^{(1)} = E_{0-}^{(2)}$ yields $\mathscr{R}_{E_0^{(1)}} \subset \mathscr{R}_{E_0^{(2)}}$, as desired.

c. $(III_3) \Rightarrow (I_3)$. The relation $(x, S_i x) = \int_{-\mu_1^{(i)}}^{0} \lambda d(x, E_\lambda^{(i)} x) + \int_{1-}^{1+\mu_1^{(i)}} \lambda \, d(x, E_\lambda^{(i)} x)$, $i = 1, 2$ leads, in view of the hypothesis, to the equation:

$$(x, (S_1 - S_2)x) = 0 \cdot (x, [(E_0^{(1)} - E_{0-}^{(1)}) - (E_0^{(2)} - E_{0-}^{(2)})]x) \\ + 1 \cdot (x, [(E_1^{(1)} - E_{1-}^{(1)}) - (E_1^{(2)} - E_{1-}^{(2)})]x) .$$

Since the hypothesis guarantees $E_1^{(1)} - E_{0-}^{(1)} = E_1^{(2)} - E_{0-}^{(2)}$, as well as $\mathscr{R}_{E_0^{(1)}-E_{0-}^{(1)}} \subset \mathscr{R}_{E_0^{(2)}-E_{0-}^{(2)}}$, we may conclude that $\mathscr{R}_{E_1^{(1)}-E_1^{(1)}} \supset \mathscr{R}_{E_1^{(2)}-E_{1-}^{(2)}}$ which shows that $(x, (S_1 - S_2)x) \geqq 0$ for all x.

This completes c., and with it the proof of the theorem.

REMARKS. The proof given of a. $(I_3) \Rightarrow (II_3)$ was very clearly tied up with the hypothesis, $P_i \in \mathscr{C}$, whereas the other steps in the chain of equivalences are valid without this hypothesis. This brings up the question as to whether the hypothesis is an artificial one tied up only with the particular method of proof given. The following example demonstrates that some such restriction on the operators P_i is necessary in order that the theorem be true.

A Counterexample. Let $\mathfrak{X}$ be the space l_2 of sequences of reals. Denote by e_ν, $\nu = 1, 2, \cdots$ the sequence consisting solely of zeros except for a one in the νth place. Let $\{\gamma_n\}$, $\{\delta_n\}$ denote an arbitrary pair of strictly decreasing real sequences converging to zero and satisfying:

$$\gamma_1 = \delta_1, \quad \gamma_n > \delta_n > \gamma_{n+1>0} , \qquad n = 2, 3, \cdots . \tag{2.7}$$

$$(\text{e.g., } \gamma_1 = \delta_1 = 1, \ \gamma_{n+1} = \frac{1}{2n}, \ \delta_{n+1} = \frac{1}{2n+1}, \ n \geqq 1) .$$

We define operators S_1 and S_2 as follows:

$$S_1 e_\nu = \begin{cases} -\gamma_1 e_\nu & \nu = 4k+1 \\ -\gamma_{k+2} e_\nu & \nu = 4k+2 \\ (1+\gamma_1) e_\nu & \nu = 4k+3 \\ (1+\gamma_{k+2}) e_\nu & \nu = 4k+4 \end{cases} \qquad k \geqq 0 , \tag{2.8}$$

$$S_2 e_\nu = \begin{cases} -\gamma_1 e_\nu & \nu = 4k+1 \\ -\delta_{k+1} e_\nu & \nu = 4k+2 \\ (1+\gamma_1) e_\nu & \nu = 4k+3 \\ (1+\delta_{k+2}) e_\nu & \nu = 4k+4 \end{cases} \qquad k \geqq 0 . \tag{2.9}$$

That is, S_1 and S_2 have the form:

$$S_1(\cdots, x_{4k+1}, x_{4k+2}, x_{4k+3}, x_{4k+4}, \cdots) \\ = (\cdots, -\gamma_1 x_{4k+1}, -\gamma_{k+2}\, x_{4k+2}, (1+\gamma_1)x_{4k+3}, (1+\gamma_{k+2})x_{4k+4}, \cdots)$$

$$S_2(\cdots, x_{4k+1}, x_{4k+2}, x_{4k+3}, x_{4k+4}, \cdots) \\ = (\cdots, -\gamma_1 x_{4k+1}, -\delta_{k+1} x_{4k+2}, (1+\gamma_1)x_{4k+3}, (1+\delta_{k+2})x_{4k+4}, \cdots) .$$

Then S_1, S_2 are bounded symmetric operators with $S_1^2 - S_1$ and $S_2^2 - S_2$

both positive definite. Moreover $I - S_1$ is equivalent to S_1 under the unitary transformation $U_1 : e_{4k+j} \longleftrightarrow e_{4k+\pi_1(j)}$, $k \geqq 0$, with the permutation π_1 given by $\pi_1 = \begin{Bmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{Bmatrix}$, while $I - S_2$ is equivalent to S_2 under the unitary transformation $U_2 : e_{4k+j} \longleftrightarrow e_{\pi_2(4k+j)}$, $k \geqq 0$, with

$$\pi_2(\nu) = \begin{cases} 4k+7 & \nu = 4k+1 \\ 4k & \nu = 4k+2 \\ 4k-3 & \nu = 4k+3 \\ 4k+6 & \nu = 4k+4 \end{cases} \qquad k \geqq 1 ,$$

$$= \begin{cases} 7 & \nu = 1 \\ 3 & \nu = 2 \\ 2 & \nu = 3 \\ 6 & \nu = 4 \end{cases} \qquad (k = 0) .$$

Hence S_1, S_2 are symmetric parts of projections, by Theorem 2. A computation shows that $(x, (S_1 - S_2)x) \geqq 0$ for all $x \in \mathfrak{X}$, whereas $S_1^2 - S_1 \neq S_2^2 - S_2$ and $S_1 \neq S_2$ on $\mathscr{R}_{S_i^2 - S_i} = \mathfrak{X}$ ($\mathscr{N}_{S_i^2 - S_i} = 0$, $i = 1, 2$). Hence for this case the conclusion of Theorem 3 fails—in fact (I_3) holds but both (II_3) and (III_3) are false.

Note that the above example even makes use of operators S_1, S_2 which have, except for $\lambda = 0$ and $\lambda = 1$, pure point spectra.

3. **Convergence of ordered sequences.** We now give a brief discussion of the convergence problem for families of projections which are ordered by the relation $\prec$. In view of the difficulties encountered with Theorem 3 it is not surprising that, in general, an arbitrary family of projections ordered by $\prec$ does not converge. However by imposing further restrictions one arrives at

THEOREM 4. *Let $\{P_n\} \subset \mathscr{C}$ denote a sequence of projections such that $P_n \prec P_{n+1}$, $n = 1, 2, \cdots$ [or else $P_n \succ P_{n+1}$, $n = 1, 2, \cdots$]. Suppose further that $P_n P_m = P_m P_n$, i.e. there is pair-wise commutativity. Then $\{P_n\}$ converges (strongly) to a projection operator P.*

Before proving this result we establish a convenient

LEMMA. *If $P_1, P_2 \in \mathscr{C}$ and $P_1 \succ P_2$ then the following conditions are equivalent:*

(i) $P_1 \smile P_2$

(ii) $W_1 = W_2$, *where* $W_i = \dfrac{1}{2}(P_i - P_i^*)$, $i = 1, 2$

(iii) $U_1 = U_2$ *on* $\overline{\mathscr{R}}_{S_1^2 - S_1^2}$, *where* U_i $(i = 1, 2)$ *denotes the unitary operator appearing in equation* $(*)$ *of Theorem* 1. [$P_2 \succ P_1$ gives the same conclusion.]

Proof of lemma. As usual, denote the s.a. and skew parts of P_i by S_i and W_i, respectively. The given hypothesis leads, by Theorem 3, to the conclusion: $S_1^2 - S_1 = S_2^2 - S_2 = Q$ with $S_1 = S_2$ on $\overline{\mathscr{R}}_Q$, and on $\mathscr{N}_Q$ S_1, S_2 are s.a. projections satisfying $S_1(\mathscr{N}_Q) \supset S_2(\mathscr{N}_Q)$. Since $W_i = U_i(S_i^2 - S_i)^{(1/2)}$, $i = 1, 2$, W_i annihilates $\mathscr{N}_Q$ and so we may restrict our attention to $\overline{\mathscr{R}}_Q$. We show first that (i)$\Longleftrightarrow$(ii). Since $S_1 = S_2 = S$ on this subspace, (ii)$\Rightarrow$(i) is trivial so we only have to consider (i)$\Rightarrow$(ii). Now $P_1P_2 = P_2P_1$ gives

$$(3.1) \qquad SW_2 - W_2S - SW_1 + W_1S = W_2W_1 - W_1W_2 .$$

Since the left and right sides of (3.1) are s. a. and skew respectively, we deduce

$$(3.2) \qquad S(W_2 - W_1) = (W_2 - W_1)S \qquad [\text{and } W_1W_2 = W_2W_1] .$$

Applying the relation (1.3) in the form $W_iS = (I - S)W_i$, $i = 1, 2$, we obtain

$$(3.3) \qquad 2(S - \tfrac{1}{2}I)(W_2 - W_1) = 0 .$$

Since $\lambda = \frac{1}{2}$ is not in the spectrum of S we conclude $W_2 = W_1$, as was to be proved. The proof that (ii)$\Longleftrightarrow$(iii) is an immediate consequence of the representation $W_i = U_i(S_i^2 - S_i)^{(1/2)}$, since $S_1 = S_2$ on $\overline{\mathscr{R}}_Q$. This completes the proof of the lemma.

Proof of Theorem 4. Suppose for definiteness $P_n \prec P_{n+1}$. Then the operators $S_n^2 - S_n$ are all the same. Denote this operator by Q. On $\overline{\mathscr{R}}_Q$ not only do all the $\{S_n\}$ coincide, but according to the lemma the $\{W_n\}$ coincide, too. Therefore on the subspace $\overline{\mathscr{R}}_Q$ we have the relation:

$$(3.4) \qquad P_1 = P_2 = \cdots = P_n = \cdots = P .$$

On the other hand as in the proof of the lemma we see that, restricted to $\mathscr{N}_Q$, the P_n are s.a. projections forming a monotone sequence. Since every monotone sequence of s.a. projections converges to such a projection ([9], p. 268 and 263), we see that on $\mathscr{N}_Q$, $P_n \to P$, whereas on $\overline{\mathscr{R}}_Q$, $P_n = P$. Hence $P_n \to P$ strongly on $\mathfrak{X}$, as was to be proved.

REMARK. According to the lemma, the conditions $P_i \in \mathscr{C}$, $P_1 \succ P_2$ and $P_1 \smile P_2$ together imply that $\overline{\mathscr{R}}_{P_1} \supset \overline{\mathscr{R}}_{P_2}$ and $\mathscr{N}_{P_1} \subset \mathscr{N}_{P_2}$ so that $P_1 \geqq P_2$ in the sense of Lorch [7]. Therefore this result is also a consequence of a result due to Lorch, once the lemma is established.

4. Applications. We consider first in this section one simple application of the preceding work to a problem in probability and statistics. Our results help to clarify the situation.

The operators considered operate on finite dimensional real spaces, and as is customary we consider them as matrices. Let (x, Ax) be the quadratic form of the symmetric, positive operator A on Euclidean n-space, where $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ is an n-vector of random variables (r.v.'s) x_i which are Gaussian distributed. We write $x \sim N[\mu, \Sigma]$ to signify that x is Gaussian distributed with mean vector μ and covariance matrix Σ (i.e. Σ is symmetric and positive definite). The following facts can be directly verified, or may be found, for instance, in [4].

PROPOSITION. If $x \sim N[\mu, \Sigma]$ then the quadratic form (x, Ax) with A positive symmetric, is distributed as a noncentral Chi-square r.v. if, and only if, $A\Sigma$ is a projection. If (x, Ax) is another such form then they are independently distributed if, and only if, $A\Sigma B = 0$. (The nonnegative number $\frac{1}{2}(\mu, A\mu)$ is called the noncentrality parameter.)

As a consequence of Theorem 3 we have

THEOREM 5. *Let A, $B = \sum_{i=1}^{k_1} B_i$, and $C = \sum_{j=1}^{k_2} C_j$ be $n \times n$ symmetric operators. Suppose $x \sim N[\mu, \Sigma]$ and $(x, Ax) = (x, Bx) + (x, Cx)$. Further suppose that (x, Ax) and (x, Bx) are distributed as noncentral Chi-square* r.v.'s. *Then (x, Cx) is also distributed as a noncentral Chi-square* r.v. *independently of (x, Bx) if, and only if, $C\Sigma = \Sigma C \geqq 0$, i.e. is positive. If further $B_iB_{i'} = 0$ and $C_jC_{j'} = 0$, $i \neq i' \varepsilon\{1, \cdots, k_1\}$, $j \neq j' \varepsilon\{1, \cdots, k_2\}$, then all (x, B_ix) and (x, C_jx) are mutually independently distributed as noncentral Chi-square* r.v.'s. (*The noncentrality parameters can be calculated very simply in each case.*)

Proof. Since $(x, Ax) = (x, Bx) + (x, Cx)$, we have

$$A\Sigma = B\Sigma + C\Sigma \ . \tag{4.1}$$

The r.v.'s. (x, Ax) and (x, Bx) are distributed as Chi-square, by hypothesis, so that using the proposition stated above, $A\Sigma$ and $B\Sigma$ are projections. If $C\Sigma = \Sigma C \geqq 0$, so that $C\Sigma$ is symmetric, then by Theorem 3 it follows that $C\Sigma$ is a symmetric projection. By the proposition, (x, Cx) is then distributed as a Chi-square r.v. independently of (x, Bx) since (4.1) also implies, when $C\Sigma$ is a projection, that $B\Sigma C = 0$ ([2], [4]).

Conversely, suppose that $C\Sigma \neq \Sigma C$. Then even if $C\Sigma \geqq 0$, (II_3) of Theorem 3 ensures that $C\Sigma = S + W$ where S is a symmetric projection and W is a skew operator. $W \neq 0$ by our supposition, so (1.2) shows that $C\Sigma$ is not itself a projection. It follows that (x, Cx) is not even distributed as a Chi-square r.v. [Theorem 1.8 of [2]]. This proves the first part. The second part follows from the proposition above and the fact that, under the stated condition, $B_i\Sigma$ and $C_j\Sigma$ are all projections.

Q.E.D.

REMARK 1. If $\Sigma = I$, the above result is a simple corollary of results on orthogonal projections, e.g., Theorem 2, §76 in Halmos [5]. A special case of the above result was proved in an entirely different way in [6].

REMARK 2. Results of the type given in Theorem 5 are useful in extending some "Analysis of Variance" techniques to correlated Gaussian r.v.'s.

As a second application of our results we point out an analogy between our Theorem 1 (or 1′) and somewhat deeper results on averaging (or conditional expectation) operators. There are several studies in this direction and, for instance, reference may be made to the papers [1], [8], [10].

A bounded linear operator A defined on $L^p(\mathscr{S}, \Sigma, \mu)$, where μ is a probability measure, is said to be a *generalized averaging operator* if for f, g in $L^p(\mathscr{S}, \Sigma, \mu)$ we have (Σ is a σ-field on $\mathscr{S}$, here)

$$A(gAf) = (Ag)(Af) \tag{4.2}$$

$$Ae = e \tag{4.3}$$

where e is the identity function on $\mathscr{S}$. If further A is a contraction (i.e. $\|A\| \leqq 1$), then A is an (ordinary) *averaging operator* as considered by the above named authors.

From the definition it follows that A is a projection in either case and, if $1 \leqq p < \infty$ then A is also s.a.[4] whenever it is a contraction, while this latter statement need not be true if A is merely bounded. For an averaging operator, recently Rota [10] has given the following representation: If f is in $L^p(\mathscr{S}, \Sigma, \mu)$, p fixed, and $Af = f'$, then there exists a unique sub σ-field Σ_1 of Σ relative to which f' is the Radon-Nikodým derivative of f. On the other hand, if A is any bounded projection in $L^2(\mathscr{S}, \Sigma, \mu)$ and f is in $L^2(\mathscr{S}, \Sigma, \mu)$, then, without any further restrictions, our result (Theorem 1′) gives $Af = f'$ where

$$f' = \int_{-\alpha-}^{1+\alpha} [\lambda I + (\lambda^2 - \lambda)^{(1/2)} U] df_\lambda\ , \tag{4.4}$$

for some $\alpha \geqq 0$, and where $f_\lambda(s)$ is the image of $f(s)$ under the orthogonal projection E_λ. The further requirement that a bounded projection be an average clearly restricts the spectral family $\{E_\lambda\}$ related to A

[4] i.e., A and its adjoint coincide on ess. bded. functions ([**10**]).

(cf. Theorem 1′) in an essential way.

Because of the unifying influence on some fields of mathematics, particularly probability and ergodic theory, the spectral theory of 'averaging type' operators is of considerable interest. Rota [11] has initiated the study of spectra of operators which satisfy the 'Reynold's Identity' (not all such operators need be projections). On the other hand, the point of view expressed in Theorems 1′ and 3 above constitutes a different attack. It is to be hoped that a specialization to the 'averaging type' operators will contribute to a deeper understanding of their structure. We wish to deal with it separately.

REFERENCES

1. G. Birkhoff, *Moyennes des Fonctions Bornées,* Colloq. International No. 24, du CNRS, Paris, 1950.

2. J. S. Chipman and M. M. Rao, *On the use of idempotent matrices in the treatment of linser restrictions in regression analysis,* Technical Report No 10, University of Minnesota, 1959.

3. J. Diximer, *Position Relative de Deux Variétés Lineaires Fermées dans Un Espace de Hilbert,* Revue Scientifique, **86** (1948), 387-399.

4. F. A. Graybill and G. Marsaglia, *Idempotent matrices and quadratic forms in the general linear hypothesis,* Ann. Math. Stat., **28** (1957), 678-686.

5. P. R. Halmos, *Finite Dimensional Vector Spaces,* 2nd Ed., D. Van Nostrand, 1958.

6. R. V. Hogg and A. T. Craig, *On the decomposition of certain Chi-square variables,* Ann. Math. Stat., **29** (1958), 608-610.

7. E. R. Lorch, *On a calculus of operators in reflexive vector spaces,* Trans. Amer. Math. Soc., **45** (1939), 217-234.

8. S. C. Moy, *Characterization of conditional expectation as a transformation on function spaces,* Pacific J. Math., **4** (1954), 47-63.

9. F. Riesz and B. Sz.-Nagy, *Functional Analysis,* F. Ungar Publishing Co., New York, (Translation) 1955.

10. G.-C. Rota, *On the representation of averaging operators,* Rend. Semi., Mat. Univ. Padova, **30** (1960), 52-64.

11. G.-C. Rota, *Spectral theory of smoothing operators,* Proc. Nat. Acad. Sci., **46** (1960), 863-868.

12. J. Seidel, *Angles and distances in n-dimensional euclidian and noneuclidian geometry, I, II, III,* Konink. Ned. Adad. Wetensch, Series A, **58** (1955), 329-340, pp. 535-541.

CARNEGIE INSTITUTE OF TECHNOLOGY
PITTSBURGH, PENNSYLVANIA