PACIFIC JOURNAL OF MATHEMATICS
Vol. 39, No. 3, 1971

CONVEXITY PROPERTIES OF A GENERALIZED NUMERICAL RANGE

JOHN DE PILLIS

A numerical range $W_n(A)$ of a bounded linear operator A on Hilbert space $\mathscr{H}$ is defined to be the set of complex numbers $W_n(A) = \{\mathrm{tr}(AM)\colon \text{dimension } M = n\}$ where M runs over all orthogonal n-dimensional projections on $\mathscr{H}$, and $\mathrm{tr}(\cdot)$ is the trace functional. It is known that $W_n(A)$ is always convex (the Hausdorff-Toeplitz theorem tells us that $W_1(A)$ is convex). In what follows, we replace the trace functional by the more general elementary symmetric functions, and derive certain convexity results.

The classical Haudorff-Toeplitz theorem has it that for any bounded linear operator A on Hilbert space $\mathscr{H}$, the numerical range

$$W(A) = \{\langle Ax, x\rangle\colon \|x\| = 1\}$$

is a convex subset of the Complex plane (cf. [4], [9], [10]).

Let P_x denote the orthogonal projection P_x; $y \to \langle y, x\rangle x$ onto the one-dimensional subspace spanned by x. Then $\langle Ax, x\rangle$ can be shown to equal $\mathrm{tr}(AP_x)$, the trace of the operator AP_x (equivalently $\langle Ax, x\rangle = \mathrm{tr}(P_xAP_x)$, the trace of the compression of A to the space $\mathrm{sp}[x]$ spanned by x.)

In light of the above interpretation, it is natural to ask whether the set

$$W_n(A) = \{\mathrm{tr}(AM)\colon \text{dimension } M = n\}\ , \tag{1.1}$$

where M runs over all n-dimensional orthogonal projections on $\mathscr{H}$, is convex; as a convenient ambiguity, we use the symbol M to represent both the n-dimensional subspace M and the orthogonal projection on $\mathscr{H}$ whose range is M. This question seems to have been raised first by Halmos [5], and consequently answered in the affirmative, by C. A. Berger [1], [6]. The convexity of $W_n(A)$ when A is normal was proved by R. C. Thompson [11, Theorem 2], which appeared almost simultaneously with Berger's thesis [1].

In this paper, we extend the notion of n^{th} order numerical range (1.1) by replacing the linear trace functional $\mathrm{tr}(\cdot)$ by the more general elementary symmetric functions

$$E_1(\cdot) = \mathrm{tr}(\cdot), E_2(\cdot), \cdots, E_r(\cdot), \cdots, E_n(\cdot) = \text{determinant}(\cdot)\ ,$$

defined on the compressions of the operator A to n-dimensional subspaces of $\mathscr{H}$. That is we shall study the set complex of scalars

$$W_{r,n}(A) = \{E_r(AM)\colon M \subset \mathscr{H},\ \text{dimension}(M) = n\} . \tag{1.2}$$

First, a quasi-convexity is shown for the set (1.2) for $n = r$ and with A replaced by $A + z$ for large complex z (Theorem 5.2). Then, we present a convexity statement (Theorem 6.1) for those linear operators which exhibit a certain kind of strong convergence (Definition 3.4).

It is somewhat surprising that the badly nonlinear elementary symmetric functions allow for any convexity properties relative to (1.2) at all, but we are able to prove as our principal result, the following property of $W_{r,n}(A)$. (See Theorem 6.2.)

THEOREM. *Let A be any bounded linear operator on Hilbert space $\mathscr{H}$ and let r and n be any positive integers such that $1 \leqq r \leqq n$. Then for $E_r(AM)$ and $E_r(N)$ in $W_{r,n}(A)$, the entire line segment*

$$\{\lambda E_r(AM) + (1-\lambda)E_r(AN)\colon 0 < \lambda < 1\}$$

is also in $W_{r,n}(A)$ provided the n-dimensional subspaces M and N are mutually orthogonal. More specifically, for every pair of mutually orthogonal n-dimensional subspaces, M and N, and for every $\lambda \in (0, 1)$, there exists an n-dimensional subspace U_λ such that

$$\lambda E_r(AM) + (1-\lambda)E_r(AN) = E_r(AU_\lambda) . \tag{1.3}$$

Due to the constructive nature of our proof, we are able to show the interesting fact that U_λ may be chosen once and for all so that (1.3) remains valid for each $r = 1, 2, \cdots, n$.

We proceed to the development of these results now.

2. Preliminaries. Throughout, $\mathscr{H}$ will be a Hilbert space (finite or infinite dimensional) with inner product $\langle \cdot, \cdot \rangle$. For each $r = 1, 2, \cdots$, we construct the vector space

$$\Lambda^r \mathscr{H} = \text{sp}\,[(x_1 \wedge x_2 \wedge \cdots \wedge x_r)],\ x_1, x_2, \cdots, x_r \in \mathscr{H}$$

spanned by all decomposable r-vectors $x_1 \wedge x_2 \wedge \cdots \wedge x_r$, where the vectors $x_1, x_2, \cdots, x_r$ run over $\mathscr{H}$. We use the wedge $\wedge$ to deonte the exterior (Grassmann) product and the symbol

$$\text{sp}\,[x],\ x \in \mathscr{A}$$

denotes the vector subspace generated by all x in $\mathscr{A}$. (See Vala [12].)

An *inner product* may be defined on $\Lambda^r \mathscr{H}$ by requiring that for decomposable vectors $x_1 \wedge x_2 \wedge \cdots \wedge x_r$, and $y_1 \wedge y_2 \wedge \cdots \wedge y_r$ in $\Lambda^r \mathscr{H}$,

$$(2.1) \qquad \langle x_1 \wedge x_2 \wedge \cdots \wedge x_r, y_1 \wedge y_2 \wedge \cdots \wedge y_r \rangle = \det(\langle x_i, y_j \rangle) ,$$

the determinant of the $r \times r$ matrix with ij^{th} entry $\langle x_i, y_j \rangle$. (See [8, Ch. XVI].)

Let A be any bounded linear operator on $\mathscr{H}$. Then a (bounded) linear operator $C_r(A)$, the r^{th} compound of A, is defined on $\Lambda^r \mathscr{H}$ by the equation

$$(2.2) \qquad C_r(A)(x_1 \wedge x_2 \wedge \cdots \wedge x_r) = Ax_1 \wedge Ax_2 \wedge \cdots \wedge Ax_r ,$$

for all $x_1, x_2, \cdots, x_r$ in $\mathscr{H}$.

As a notational convenience, we shall introduce $Q_{n,r}$, the $\binom{n}{r}$-element set of order-preserving functions σ sending $\{1, 2, \cdots, r\}$ into $\{1, 2, \cdots, r, \cdots n\}$, where $1 \leqq r \leqq n$. More exactly,

$$Q_{n,r} = \{\sigma: \{1, 2, \cdots, r\} \longrightarrow \{1, 2, \cdots, r, \cdots, n\}\}$$

where $1 \leqq \sigma(1) < \sigma(2) < \cdots < \sigma(r) \leqq n$.

As an immediate use of this set $Q_{n,r}$, we set down the following useful property:

PROPOSITION 2.1. *Let M be an n-dimensional subspace of $\mathscr{H}$ having* o.n. *basis $\mathscr{E} = \{e_1, e_2, \cdots, e_n\}$. Then for each integer r, $1 \leqq r \leqq n$, the $\binom{n}{r}$-element set*

$$\mathscr{E}_r = \{e_{\sigma(1)} \wedge e_{\sigma(2)} \wedge \cdots \wedge e_{\sigma(r)}: \sigma \in Q_{n,r}\}$$

is o.n. *in $\Lambda^r \mathscr{H}$. Moreover, if M is the orthogonal projection on $\mathscr{H}$ with range M spanned by $\mathscr{E}$, then $C_r(M)$ is the orthogonal projection on $\Lambda^r \mathscr{H}$ with range spanned by the* o.n. *set $\mathscr{E}_r$.*

Proof. The fact that $\mathscr{E}_r$ is o.n. in $\Lambda^r \mathscr{H}$ is immediate from (2.1). Similarly, by extending $\mathscr{E}$ to an o.n. basis for all of $\mathscr{H}$, the second assertion of our proposition follows from (2.1) and (2.2). (See also de Pillis [3].) This ends the proof.

Given the n-element set $\{x_1, x_2, \cdots, x_n\}$ in $\mathscr{H}$. Then for each $\sigma \in Q_{n,r}$, $\underset{\sim}{x}_\sigma$ will be that vector in $\Lambda^r \mathscr{H}$ defined by

$$(2.3) \qquad \underset{\sim}{x}_\sigma = x_{\sigma(1)} \wedge x_{\sigma(2)} \wedge \cdots \wedge x_{\sigma(r)} .$$

Now let A be any bouded linear operator on $\mathscr{H}$. Let M be (the orthogonal projection onto) a finite dimensional subspace of $\mathscr{H}$. Then for each $r = 1, 2, \cdots, n$ $E_r(AM)$, the r^{th} elementary symmetric function of AM is defined by

$$(2.4) \qquad E_r(AM) = \operatorname{tr}(C_r(AM)) ,$$

the trace of the operator $C_r(AM)$ on $\Lambda^r \mathscr{H}$.

The existence of the trace presents no problem since M is finite dimensional (hence, $C_r(AM)$ is finite dimensional on $\Lambda^r\mathscr{H}$). In fact we offer the following more explicit form for $E_r(AM)$:

PROPOSITION 2.2. *For A a bounded linear operator on $\mathscr{H}$, and for M an n-dimensional subspace with* o.n. *basis* $\{e_1, e_2, \cdots, e_n\}$, *we have the equality*

$$E_r(AM) = \sum_{\sigma \in Q_{n,r}} \langle C_r(AM)\underset{\sim}{e}_\sigma, \underset{\sim}{e}_\sigma\rangle . \tag{2.5}$$

Proof. The trace of any bounded operator B on Hilbert space may be written as

$$\sum_{i \in I} \langle Be_i, e_i\rangle \text{ where } \{e_i\}_{i \in I}$$

is any o.n. basis of $\mathscr{H}$. Since $C_r(AM) = C_r(A)C_r(M)$ (see 2.2)) is zero on all but a finite-dimensional subspace $C_r(M)$ of $\Lambda^r\mathscr{H}$, it suffices to consider the sum above relative to the finite o.n. basis $\{\underset{\sim}{e}_\sigma : \sigma \in Q_{n,r}\}$ of $C_r(M)$ (see Porposition 2.1). That is,

$$E_r(AM) = \operatorname{tr}(C_r(AM)) = \sum_{\sigma \in Q_{n,r}} \langle C_r(AM)\underset{\sim}{e}_\sigma, \underset{\sim}{e}_\sigma\rangle$$

where $\{e_1, e_2, \cdots, e_n\}$ is an o.n. basis of M, and $\{\underset{\sim}{e}: \sigma \in Q_{n,r}\}$ is an o.n. basis of $C_r(M)$. The proof is done.

REMARK. It is to be noted that

$$E_r(AM) = \sum_{\sigma \in Q_{n,r}} \langle C_r(AM)\underset{\sim}{e}_\sigma, \underset{\sim}{e}_\sigma\rangle = \sum_{\sigma \in Q_{n,r}} \langle C_r(A)\underset{\sim}{e}_\sigma, \underset{\sim}{e}_\sigma\rangle , \tag{2.6}$$

so that the appearance of $M = \operatorname{sp}[e_1, e_2, \cdots e_n]$ becomes superfluous in (2.5). To see this, observe that $C_r(AM) = C_r(A)C_r(M)$ (from (2.2)), and that $C_r(M)\underset{\sim}{e}_\sigma = \underset{\sim}{e}_\sigma$.

REMARK. Let us write $\operatorname{tr}(AM)$ in the form

$$\operatorname{tr}(AM) = \sum_{i=1}^{n} \langle Ae_i, e_i\rangle \tag{2.7}$$

where $\{e_1, e_2, \cdots, e_n\}$ is any o.n. basis of M. Then the extended Hausdorff-Toeplitz theorem of Berger [1], [6] tells us that the set $W_n(A)$ of all such n-term sums, is convex. Can we then conclude that the set $W_{r,n}(A)$ of all $\binom{n}{r}$-term sums (2.6) is also convex (by replacing A in (2.7) with $C_r(A)$ and replacing e_i with $\underset{\sim}{e}_\sigma$) ? The answer is the convexity $W_{r,n}(A)$ cannot be so deduced from the convexity of $W_n(A)$ since in (2.7), *all* n-element o.n. sets $\{e_1, e_2, \cdots, e_n)$ of $\mathscr{H}$ are permitted. Contrariwise, in considering $W_{r,n}(A)$ given in (1.2), the set of all sums of the form (2.6), only those o.n. sets $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_R\}$, $R = \binom{n}{r}$, are permitted where each $\underset{\sim}{e}_i$ is a *decomposable* unit vector of $\Lambda^r\mathscr{H}$. Recall that $\Lambda^r\mathscr{H}$ is the span of all decomposable vectors $x_1 \wedge x_2 \wedge \cdots \wedge x_r$ so that in general, not every unit vector is decomposable.

Now let M be an r-dimensional subspace of $\mathscr{H}$ with o.n. basis $\{e_1, e_2, \cdots, e_r\}$. Let A be a bounded linear operator on $\mathscr{H}$ and let z be complex. Then

$$(2.8)' \qquad \langle C_r(A+z)e_1 \wedge e_2 \wedge \cdots \wedge e_r, e_1 \wedge e_2 \wedge \cdots e_r\rangle$$

is the r^{th} degree polynomial

$$(2.8) \qquad \sum_{j=0}^{r} \sum_{\sigma \in Q_{r,j}} z^{r-j}\langle C_j(A)\underset{\sim}{e}_\sigma, \underset{\sim}{e}_\sigma\rangle :$$

For $j = 0$, we define $\sum_{\sigma \in Q_{r,0}}\langle C_0(A)\underset{\sim}{e}_\sigma, \underset{\sim}{e}_\sigma\rangle$ to be the number 1. The equivalence of (2.8)′ and (2.8) follows from the orthogonality of the set $\{e_1, e_2, \cdots, e_r\}$ along with use of (2.1) and (2.2).

Finally, we define Berger bases for a pair of r-dimensional subspaces M and N in $\mathscr{H}$, after C. A. Berger, who proved [1], [6], that such pairs of bases always exist for any such M and N.

DEFINITION 2.3. The orthonormal bases $\{x_1, x_2, \cdots, x_r\} \subset M$ and $\{y_1, y_2, \cdots, y_r\} \subset N$ will be called Berger bases (relative to M and N) if and only if $\langle x_i, y_j\rangle = 0$ whenever $i \neq j$. (Note: No constraint is placed on $\langle x_i, y_i\rangle$.)

3. Convergence properties of A.

In what follows, A will be a fixed bounded linear operator on Hilbert space $\mathscr{H}$ and x is a unit vector in $\mathscr{H}$.

DEFINITION 3.1. The unit vector $x_z \in \mathscr{H}$ is defined for each sufficiently large complex z by the conditions that

$$(1) \qquad (A+z)x_z = \langle (A+z)x_z, x\rangle x \,,$$

and

$$(2) \qquad \langle x_z, x\rangle > 0 \,.$$

REMARK. Since A is bounded, sufficiently large complex z may always be found so that $A + z$ is invertible; that is, so that a unit vector x_z always exists which is sent by $A + z$ to a scalar multiple of the fixed unit vector x. We note that if Condition (1) above obtains for some unit vector x_z, then it obtains equally well for the vector $\underline{w} \cdot x_z$, where $\underline{w}$ is any complex scalar of modulus one. Hence, Condition (2) is presented to allow a unique x_z; as we shall now see, x_z has the further property that $x_z \to x$ as $z \to \infty$. In fact, we present a stronger statement of convergence in the following lemma:

LEMMA 3.2. *Given the bounded linear operator A and unit vectors x and x_z defined in Definition 3.1. Then*

$$z(1 - |\langle x_z, x\rangle|^2) \longrightarrow 0 \quad as \quad z \longrightarrow \infty .$$

Consequently, $x_z \to x$ as $z \to \infty$.

Proof. We write the expression $\langle (A+z)x_z, x_z\rangle$ in two ways:

$$\langle (A+z)x_z, x_z\rangle = \langle Ax_z, x_z\rangle + z . \tag{3.1}$$

Substituting $\langle (A+z)x_z, x\rangle x$ for $(A+z)x_z$, yields

$$\langle (A+z)x_z, x_z\rangle = \langle Ax_z, x\rangle\langle x, x_z\rangle + z\,|\langle x_z, x\rangle|^2 . \tag{3.2}$$

Subtraction of (3.2) from (3.1) gives us

$$z(1 - |\langle x_z, x\rangle|^2) = \langle Ax_z, x\rangle\langle x, x_z\rangle - \langle Ax_z, x_z\rangle . \tag{3.3}$$

From (3.3) it follows that $\langle x_z, x\rangle \to 1$ as $z \to \infty$; to see this, divide through by z and let z go to infinity. This, in turn, implies that

$$x_z \longrightarrow x \quad as \quad z \longrightarrow \infty . \tag{3.4}$$

Again from (3.3) we may deduce that $z\,(1-|\langle x_z, x\rangle|^2) \to 0$ as $z \to \infty$; use the continuity of both A and the inner product along with (3.4) to show that the right-hand side of (3.3) goes to zero as $z \to \infty$. The lemma is proved.

For later convenience, the following notations are introduced:

DEFINITION 3.3. Let $\{x_1, x_2, \cdots, x_r\}$ be any finite set spanning subspace M of Hilbert space $\mathscr{H}$. Let A be the compression to M of some bounded operator on $\mathscr{H}$. (That is, $A = P_M A P_M$, where P_M is the orthogonal projection onto M.) By x_z, we mean that unit vector of M defined in Definition 3.1 relative to x_r. That is,

$$(A+z)x_z = \langle (A+z)x_z, x_r\rangle x_r ,$$

and

$$\langle x_z, x_r\rangle > 0 .$$

We then define the vectors $\underset{\sim}{x}_z$ in $\Lambda^r\mathscr{H}$, and $\underset{\sim}{x}_j$ in $\Lambda^j\mathscr{H}$, for $j = 1, 2, \cdots, r$ by the following equations:

$$\underset{\sim}{x}_z = x_1 \wedge x_2 \wedge \cdots \wedge x_{r-1} \wedge x_z , \tag{a}$$

and

(b) $$\underset{\sim}{x}_j = x_1 \wedge x_2 \wedge \cdots \wedge x_j\,, \quad \text{for each } j = 1, 2, \cdots, r\,.$$

We shall have occasion to use the notion of an operator A having power r on a vector x, the definition of which follows now.

DEFINITION 3.4. Let A be a bounded linear operator on Hilbert space M. Let x be a unit vector of M and let x_z be the unit vector defined above (Definition 3.1). Then A is said to have power r on x if

$$\lim_{z \to \infty} z^r(1 - \langle x_z, x\rangle^2) = 0\,.$$

A simple restatement of Lemma 3.2, using the terminology of Definition 3.4, is the proposition:

PROPOSITION 3.5. *For any bounded linear operator A on $\mathscr{H}$ and for any unit vector $x \in \mathscr{H}$, A has power one on x.*

One important instance occurs where A has power r on x for all integer values of $r = 1, 2, \cdots$. Specifically, we have

PROPOSITION 3.6. *Suppose x is an eigenvector for A. Then A has unbounded power on x. That is, for all $r = 1, 2, \cdots, m, \cdots$,*

$$z^r(1 - |\langle x_z, x\rangle|^2) \longrightarrow 0\,.$$

Proof. Observe that for all z, $x_z = x$ whenever x is an eigenvector of A.

4. **Induction hypothesis.** In the following section, we shall refer to and hence extend the induction hypothesis which we now present.

Induction Hypothesis. Given a bounded linear operator A on $\mathscr{H}$. Given n-dimensional subspaces M and N of $\mathscr{H}$ with Berger bases $\{x_1, x_2, \cdots, x_n\}$ and $\{y_1, y_2, \cdots, y_n\}$, respectively. Let the o.n. set $\{u_1, u_2, \cdots, u_n\}$ be defined by requiring of each $u_i \in \operatorname{sp}[x_i, y_i]$, that

$$\lambda\langle Ax_i, x_i\rangle + (1-\lambda)\langle Ay_i, y_i\rangle = \langle Au_i, u_i\rangle$$

for arbitrary but fixed $\lambda \in (0, 1)$. Then

$$\lambda\langle C_j(A)\underset{\sim}{x}_\sigma, x_\sigma\rangle + (1-\lambda)\langle C_j(A)\underset{\sim}{y}_\sigma, \underset{\sim}{y}_\sigma\rangle = \langle C_j(A)\underset{\sim}{u}_\sigma, \underset{\sim}{u}_\sigma\rangle$$

for each $j = 1, 2, \cdots, r-1$, and for each $\sigma \in Q_{n\ j}$.

REMARK 4.1. In the "earliest" possible case, $j = 1$, $r = 2$, the induction hypothesis reduces to the classical Hausdorff-Toeplitz theorem relative to the Berger bases for M and N. In what follows, we shall prove that under certain restrictions on A, or on the subspaces M and N, the induction hypothesis may be extended from the cases $j = 1, 2, \cdots, r-1$ to the cases $j = 1, 2, \cdots, r-1, r$.

REMARK 4.2. The fixed o.n. sets $\{x_i\}$, $\{y_i\}$ and $\{u_i\}$ meet all the conditions described above if the operator A is replaced by $A + z$, where z is any complex scalar (i.e., z is a complex scalar multiple of the identity operator).

5. A quasi-convexity result.

We now extend our induction hypothesis to a convexity result involving the operator $A + z$ (Theorem 5.2).

REMARK 5.1. Let F be some scalar-valued function on the unit sphere of $\mathscr{H}$, which has convex range. That is, for all $\lambda \in (0, 1)$, and for all x, y in $\mathscr{H}$ such that $\| x \| = \| y \| = 1$, there exists u, $\| u \| = 1$ such that

$$\lambda F(x) + (1-\lambda)F(y) = F(u) . \tag{5.1}$$

Now let $p_1(\theta, z)$ and $p_2(\theta, z)$ be two complex valued functions of z, whose values have the same argument θ for each z; the common argument, θ, necessarily depends on z. Then

(5.2)
$$\lambda p_1(\theta, z)F(x) + (1-\lambda)p_2(\theta, z)F(y) = (\lambda p_1(\theta, z) + (1-\lambda)p_2(\theta, z))F(u_z) ,$$

where u_z is a unit vector depending on z.

To verify (5.2), divide both sides by $\lambda p_1(\theta, z) + (1-\lambda)p_2(\theta, z)$. The fact that $p_1(\theta, z)$ and $p_2(\theta, z)$ each factor as $e^{i\theta}$ times positive scalars, reveals the lefthand side of (5.2) as a convex combination of $F(x)$ and $F(y)$.

THEOREM 5.2. (*A quasi convexity theorem*) *Suppose the induction hypothesis obtains for each* $j = 1, 2, \cdots, r-1$. *Let* M *and* N *be* r-*dimensional subspaces of* $\mathscr{H}$ *with Berger bases* $\{x_1, x_2, \cdots, x_r\}$ *and* $\{y_1, y_2, \cdots, y_r\}$, *respectively. Then for every neighborhood* $U(\infty)$ *of infinity, there exists a* $z \in U(\infty)$ *such that*

$$\lambda\langle C_r(A+z)\underset{\sim}{x}_z, \underset{\sim}{x}_z\rangle + (1-\lambda)\langle C_r(A+z)\underset{\sim}{y}_z, \underset{\sim}{y}_z\rangle = \langle C_r(A+z)\underset{\sim}{u}_z, \underset{\sim}{u}_z\rangle$$

where $\underset{\sim}{x}_z$ *and* $\underset{\sim}{y}_z$ *are defined above* (Definition 3.3) *and*

$$\underset{\sim}{u}_z = u_1 \wedge u_2 \wedge \cdots \wedge u_{r-1} \wedge u_z ,$$

where u_z *is a unit vector of* sp $[x_z, y_z]$.

Proof. Recall that x_z is that unit vector such that

$$(A+z)x_z \in \operatorname{sp}[x_r] ,$$

where $\langle x_z, x_r\rangle > 0$. (Definition 3.3.) Thus, $(A+z)x_z$ is orthogonal to sp $[x_1, x_2, \cdots, x_{r-1}]$. From the definition (2.1) of the inner product on $\Lambda^r\mathscr{H}$, we have

$$\begin{aligned} &\lambda\langle C_r(A+z)\underset{\sim}{x}_z, \underset{\sim}{x}_z\rangle \\ (5.3)\quad &= \lambda\langle(A+z)x_1 \wedge \cdots \wedge (A+z)x_{r-1} \wedge (A+z)x_z, x_1 \wedge \cdots \wedge x_{r-1} \wedge x_z\rangle \\ &= \lambda\langle C_{r-1}(A+z)\underset{\sim}{x}_{r-1}, \underset{\sim}{x}_{r-1}\rangle\langle(A+z)x_z, x_z\rangle . \end{aligned}$$

Similarly, we may write

$$(5.4)\quad (1-\lambda)\langle C_r(A+z)\underset{\sim}{y}_z, \underset{\sim}{y}_z\rangle = (1-\lambda)\langle C_{r-1}(A+z)\underset{\sim}{y}_{r-1}, \underset{\sim}{y}_{r-1}\rangle\langle(A+z)y_z, y_z\rangle.$$

We add (5.3) and (5.4) to obtain

$$\begin{aligned} &\lambda\langle C_r(A+z)\underset{\sim}{x}_z, \underset{\sim}{x}_z\rangle + (1-\lambda)\langle C_r(A+z)\underset{\sim}{y}_z, \underset{\sim}{y}_z\rangle \\ (5.5)\quad &= \lambda\langle C_{r-1}(A+z)\underset{\sim}{x}_{r-1}, \underset{\sim}{x}_{r-1}\rangle\langle(A+z)x_z, x_z\rangle \\ &\quad + (1-\lambda)\langle C_{r-1}(A+z)\underset{\sim}{y}_{r-1}, \underset{\sim}{y}_{r-1}\rangle\langle(A+z)y_z, y_z\rangle . \end{aligned}$$

We assert that for each neighborhood $U(\infty)$ of infinity, a $z \in U(\infty)$ may be found so that the (complex) arguments of the two monic $(r-1)$st degree polynomials,

$$\langle C_{r-1}(A+z)\underset{\sim}{x}_{r-1}, \underset{\sim}{x}_{r-1}\rangle \quad \text{and} \quad \langle C_{r-1}(A+z)\underset{\sim}{y}_{r-1}, \underset{\sim}{y}_{r-1}\rangle$$

(see (2.8)), agree. In fact, their quotient, call it f, is an analytic function in a neighborhood of infinity; moreover, $f(z)$ converges to one at infinity. By the open mapping theorem for analytic functions (c.f. [2: pg 175], [7]), open neighborhoods (of analyticity) of infinity will be sent by f to open neighborhoods of the positive number one. Therefore, every neighborhood of infinity will be sent by f to open neighborhoods of the positive number one. Thus, every neighborhood of infinity contains a z such that $f(z) > 0$; moreover, $f(z)$ is as near to one as we please. That is, every neighborhood $U(\infty)$ of infinity contains a z for which the arguments of

$$(5.6)\quad \langle C_{r-1}(A+z)\underset{\sim}{x}_{r-1}, \underset{\sim}{x}_{r-1}\rangle \quad \text{and} \quad \langle C_r(A+z)\underset{\sim}{y}_{r-1}, \underset{\sim}{y}_{r-1}\rangle$$

are equal. Replace the symbols $p_1(\theta)$ and $p_2(\theta)$ in (5.2), Remark 5.1, by the polynomials of (5.6), and replace $F(x)$ and $F(y)$ by $\langle(A+z)x_z, x_z\rangle$

and $\langle (A+z)y_z, y_z\rangle$, to obtain the following equivalent expression for the right hand side of (5.5):

(5.7)

$$\left[\underset{a(z)}{\lambda\langle C_{r-1}(A+z)\underset{\sim}{x}_{r-1}, \underset{\sim}{x}_{r-1}\rangle} + \underset{b(z)}{(1-\lambda)\langle C_{r-1}(A+z)\underset{\sim}{y}_{r-1}, \underset{\sim}{y}_{r-1}\rangle}\right]\langle (A+z)u_z, u_z\rangle ,$$

$$= [a(z) + b(z)]\langle (A+z)u_z, u_z\rangle$$

where u_z in $\mathrm{sp}\,[x_z, y_z]$ is a unit vector satisfying the equation

$$\frac{a(z)}{a(z)+b(z)}\langle (A+z)x_z, x_z\rangle + \frac{b(z)}{b(z)+b(z)}\langle (A+z)y_z, y_z\rangle = \langle (A+z)u_z, u_z\rangle .$$

This, in turn, is equivalent to the equation

$$(5.8)\qquad \frac{a(z)}{a(z)+b(z)}\langle Ax_z, x_z\rangle + \frac{b(z)}{a(z)+b(z)}\langle Ay_z, y_z\rangle = \langle Au_z, u_z\rangle$$

for $a(z)$ and $b(z)$ defined in (5.7).

We note that the quotients $a(z)/(a(z) + b(z))$ and $b(z)/(a(z) + b(z))$ approach the values λ and $(1-\lambda)$, respectively as $z \to \infty$. (To see this, divide numerators and denominators by z^{r-1} and pass to the limit as $z \to \infty$.) Since $x_z \to x_r$ and $y_z \to y_r$ (Lemma 3.2) as $z \to \infty$, we have

$$(5.9)\qquad \langle Au_z, u_z\rangle \longrightarrow \langle Au_r, u_r\rangle \quad \text{as} \quad z \longrightarrow \infty .$$

(Recall that u_r is a unit vector of $\mathrm{sp}\,[x_r, y_r]$ satisfying the equation $\lambda\langle Ax_r, x_r\rangle + (1-\lambda)\langle Ay_r, y_r\rangle = \langle Au_r, u_r\rangle$.)

We couple the statement of our induction hypothesis (see (2.8) and Remark 4.2) with (5.7) to replace the left-hand factor of (5.7) by the $r-1$ degree polynomial

$$(5.10)\ \langle C_{r-1}(A+z)u_1\wedge\cdots\wedge u_{r-1}, u_1\wedge\cdots\wedge u_{r-1}\rangle = \langle C_{r-1}(A+z)\underset{\sim}{u}_{r-1}, \underset{\sim}{u}_{r-1}\rangle ,$$

where each unit vector u_i in $\mathrm{sp}\,[x_i, y_i]$, $i = 1, 2, \cdots, r-1$ has the property that

$$\lambda\langle Ax_i, x_i\rangle + (1-\lambda)\langle Ay_i, y_i\rangle = \langle Au_i, u_i\rangle .$$

This yields the simplified form

$$(5.11)\qquad \langle C_{r-1}(A+z)\underset{\sim}{u}_{r-1}, \underset{\sim}{u}_{r-1}\rangle\langle (A+z)u_z, u_z\rangle$$

which is equivalent to (5.5) and to (5.7).

Notice that $(A+z)u_z$ is a certain linear combination of the vectors x_r and y_r. This is so because $u_z \in \mathrm{sp}\,[x_z, y_z]$ and $(A+z)$ sends x_z into $\mathrm{sp}\,[x_r]$ and also sends y_z into $\mathrm{sp}\,[y_r]$ by definition (Definitions 3.1 and 3.3). Thus, $(A+z)u_z$ is orthogonal to $\mathrm{sp}\,[u_1, u_2, \cdots, u_{r-1}]$;

orthogonality is guaranteed by the fact that

$$\{x_1, x_2, \cdots, x_r\} \quad \text{and} \quad \{y_1, y_2, \cdots, y_r\}$$

are Berger bases for M and N, respectively (Definition 2.3). Accordingly, we may write (5.11) as follows:

(5.12)
$$\langle C_r(A+z)u_1 \wedge \cdots \wedge u_{r-1} \wedge u_z, u_1 \wedge \cdots \wedge u_{r-1} \wedge u_z \rangle = \langle C_r(A+z)\underset{\sim}{u}_z, \underset{\sim}{u}_z \rangle .$$

If we combine (5.12) with (5.5), we obtain the final equality

(5.13)
$$\lambda\langle C_r(A+z)\underset{\sim}{x}_z, \underset{\sim}{x}_z \rangle + (1-\lambda)\langle C_r(A+z)\underset{\sim}{y}_z, \underset{\sim}{y}_z \rangle = \langle C_r(A+z)\underset{\sim}{u}_z, \underset{\sim}{u}_z \rangle .$$

Statement (5.13) above completes the proof of the theorem.

6. A convexity result for $W_{r,n}(A)$. In this section, we combine the quasi-convexity result (Theorem 5.2) with the notion of A having sufficient integer power on certain x (Definition 3.4) to obtain a convexity theorem for A (Theorem 6.1). As a consequence, we obtain our main convexity result (Theorem 6.2) which holds for arbitrary but fixed linear operator A so long as the subspace M and N are mutually orthogonal.

THEOREM 6.1. *Let M and N be r-dimensional subspaces of Hilbert space $\mathscr{H}$ having* o.n. *Berger bases $\{x_1, x_2, \cdots, x_r\}$ and $\{y_1, y_2, \cdots, y_r\}$, respectively. Suppose the bounded linear operator A, when restricted to each of these subspaces has power r on x_r and y_r; that is,*

$$\text{(i)} \qquad \lim_{z\to\infty} z^r(1-|\langle x_z, x_r\rangle|^2) = \lim_{z\to\infty} z^r(1-|\langle y_z, y_r\rangle|^2) = 0 .$$

If an o.n. *set $\{u_1, u_2, \cdots, u_r\}$ satisfies the equations*

$$\text{(ii)} \qquad \lambda\langle Ax_i, x_i\rangle + (1-\lambda)\langle Ay_i, y_i\rangle = \langle Au_i, u_i\rangle,\ i = 1, 2, \cdots, r .$$

for arbitrary but fixed $\lambda \in (0, 1)$, then necessarily,

$$\text{(6.1)} \qquad \lambda\langle C_r(A)\underset{\sim}{x}_r, \underset{\sim}{x}_r\rangle + (1-\lambda)\langle C_r(A)\underset{\sim}{y}_r, \underset{\sim}{y}_r\rangle = \langle C_r(A)\underset{\sim}{u}_r, \underset{\sim}{u}_r\rangle$$

whenever

$$\text{(iii)} \qquad \lim_{z\to\infty} z^r(1 - |\langle u_z, u_r\rangle|^2) = 0$$

where the unit vector u_z satisfies the equation

$$\text{(iv)} \qquad \lambda\langle Ax_z, x_z\rangle + (1-\lambda)\langle Ay_z, y_z\rangle = \langle Au_z, u_z\rangle .$$

Proof. Consider the function

$$\varphi(z) = \lambda\langle C_r(A+z)\underset{\sim}{x}_z, \underset{\sim}{x}_z\rangle + (1-\lambda)\langle C_r(A+z)\underset{\sim}{y}_z, \underset{\sim}{y}_z\rangle - \langle C_r(A+z)\underset{\sim}{u}_z, \underset{\sim}{u}_z\rangle .$$

Now $x_z \in \mathrm{sp}\,[x_1, x_2, \cdots, x_r]$ must be of the form

$$x_z = \sum_{i=1}^{r-1} a_i x_i + \langle x_z, x_r\rangle x_r$$

for certain scalars $a_1, a_2, \cdots, a_{r-1}$. Thus

$$\begin{aligned}\underset{\sim}{x}_z &= x_1 \wedge \cdots \wedge x_{r-1} \wedge x_z \\ &= \sum_{i=1}^{r-1} a_i(x_1 \wedge \cdots \wedge x_{1-1} \wedge x_i) + \langle x_z, x_r\rangle(x_1 \wedge \cdots \wedge x_{r-1} \wedge x_r) \\ &= 0 + \langle x_z, x_r\rangle \underset{\sim}{x}_r ,\end{aligned}$$

where each term in the summation involving $a_1, a_2, \cdots, a_{r-1}$ equals zero due to repetitions of x_i in the Grassmann product

$$x_1 \wedge \cdots \wedge x_{r-1} \wedge x_i .$$

We substitute $\langle x_z, x_r\rangle \underset{\sim}{x}_r$ for $\underset{\sim}{x}_z$ into our expression for $\phi(z)$ to obtain

(6.2)
$$\varphi(z) = \lambda\langle C_r(A+z)\underset{\sim}{x}_r, \underset{\sim}{x}_r\rangle\alpha_z + (1-\lambda)\langle C_r(A+z)\underset{\sim}{y}_r, \underset{\sim}{y}_r\rangle\beta_z - \langle C_r(A+z)\underset{\sim}{u}_r, \underset{\sim}{u}_r\rangle\gamma_z$$

where

$$\text{(6.3)} \qquad \alpha_z = |\langle x_z, x_r\rangle|^2, \ \beta_z = |\langle y_z, y_r\rangle|^2 \ \text{ and } \ \gamma_z = |\langle u_z, u_r\rangle|^2 .$$

If we substitute the explicit polynomial expressions for each of the three inner products in (6.2) (see (2.8)), we obtain the following expression for $\varphi(z)$:

(6.4)
$$\sum_{j=0}^{r} \sum_{\sigma \in Q_{r,j}} z^{r-j} [\lambda\langle C_j(A)\underset{\sim}{x}_\sigma, \underset{\sim}{x}_\sigma\rangle\alpha_z + (1-\lambda)\langle C_j(A)\underset{\sim}{y}_\sigma, \underset{\sim}{y}_\sigma\rangle\beta_z - \langle C_j(A)\underset{\sim}{u}_\sigma, \underset{\sim}{u}_\sigma\rangle\gamma_z] .$$

Our induction hypothesis allows us to replace $\langle C_j(A)\underset{\sim}{u}_\sigma, \underset{\sim}{u}_\sigma\rangle$ in (6.4) by

$$\lambda(C_j(A)\underset{\sim}{x}_\sigma, \underset{\sim}{x}_\sigma\rangle + (1-\lambda)\langle C_j(A)\underset{\sim}{y}_\sigma, \underset{\sim}{y}_\sigma\rangle ,$$

at least for the cases $j = 1, 2, \cdots, \cdots, r-1$, and for all $\sigma \in Q_{r,j}$. Effecting this substitution, (6.4) yields the following form for $\varphi(z)$:

$$\text{(6.5)} \qquad \begin{aligned}\varphi(z) = {} & z^r[\lambda(\alpha_z - \gamma_z) + (1-\lambda)(\beta_z - \gamma_z] \\ & + \sum_{j=1}^{r-1} z^{r-j}\underset{\sim}{X}_j(\alpha_z - \gamma_z) + \sum_{i=1}^{r-1} z^{r-j}\underset{\sim}{Y}_j(\beta_z - \gamma_z) \\ & \qquad\qquad \swarrow\text{—(constant term)} \\ & + \lambda\langle C_r(A)\underset{\sim}{x}_r, \underset{\sim}{x}_r\rangle\alpha_z + (1-\lambda)\langle C_r(A)\underset{\sim}{y}_r, \underset{\sim}{y}_r\rangle\beta_z - \langle C_r(A)\underset{\sim}{u}_r, \underset{\sim}{u}_r\rangle\gamma_z ,\end{aligned}$$

where for each $j = 1, 2, \cdots, r-1$,

$$\underset{\sim}{X}_j = \lambda \sum_{\sigma \in Q_{r,j}} \langle C_j(A)\underset{\sim}{x}_\sigma, \underset{\sim}{x}_\sigma\rangle , \quad \text{and} \quad \underset{\sim}{Y}_j = (1-\lambda) \sum_{\sigma \in Q_{r,j}} \langle C_j(A)\underset{\sim}{y}_\sigma, \underset{\sim}{y}_\sigma\rangle .$$

Assumptions (i) and (iii) together guarantee that for $k = 1, 2, \cdots, r$, the quantities

$$z^k(\alpha_z - \gamma_z) , \quad \text{and} \quad z^k(\beta_z - \gamma_z)$$

tend to zero as $z \to \infty$. (Write $\alpha_z - \gamma_z = (1-\gamma_z) - (1-\alpha_z)$, and $\beta_z - \gamma_z = (1-\gamma_z) - (1-\beta_z)$.)

We then constrain the growth of z in accordance with Theorem 5.2, so that $\varphi(z) = 0$ as $z \to \infty$. From hypothesis (i), α_z and $\beta_z \to 1$ as $z \to \infty$; from hypothesis (iii), $\gamma_z \to 1$ as $z \to \infty$. Thus, we may conclude that the "constant" term of $\varphi(z)$ (see (6.5)), which tends to

$$\lambda\langle C_r(A)\underset{\sim}{x}_r, \underset{\sim}{x}_r\rangle + (1-\lambda)\langle C_r(A)\underset{\sim}{y}_r, \underset{\sim}{y}_r\rangle - \langle C_r(A)\underset{\sim}{u}_r, \underset{\sim}{u}_r\rangle ,$$

approaches the value zero. The proof of the theorem is done.

We now present our main result which, as a corollary to Theorem 6.1, holds for all bounded linear operators A on $\mathscr{H}$, and all $r = 1, 2, \cdots, n$, provided the n-dimensional subspaces M and N are orthogonal.

THEOREM 6.2. *Let M and N be n-dimensional subspaces of $\mathscr{H}$ such that M is orthogonal to N. Then for any bounded linear operator A, and for any $\lambda \in (0, 1)$, there exists an n-dimensional subspace U_λ in $M + N$ such that for each $r = 1, 2, \cdots, n$,*

$$\lambda E_r(AM) + (1-\lambda)E_r(AN) = E_r(AU_\lambda) .$$

Note that U_λ does not depend on r.

Proof. Let

$$\mathscr{X} = \{x_1, x_2, \cdots, x_r, \cdots, x_n\}$$

and

$$\mathscr{Y} = \{y_1, y_2, \cdots, y_r, \cdots, y_n\}$$

be o.n. bases of M and N, respectively, which triangularize the compression operators A: $M \to M$ and A: $N \to N$. That is, let x_n be an eigenvector for the (finitedimensional) compression of A to M. Choose x_{n-1} as a unit eigenvector of A restricted to the orthogonal complement of x_n in M; consequently, $A(x_{n-1}) \in \text{sp}\,[x_{n-1}, x_n]$. Similarly, x_{n-2} is a unit eigenvector of A restricted to the orthogonal complement of $\text{sp}\,[x_{n-1}, x_n]$, and so on, until for each $j = 1, 2, \cdots, n$, (6.6a) $$A(x_j) \in \mathrm{sp}\,[x_j, x_{j+1}, \cdots, x_n]\ .$$

By the same reasoning on N, $\{y_1, y_2, \cdots, y_n\}$ is an o.n. basis of N such that

(6.6b) $$A(y_j) \in \mathrm{sp}\,[y_j, y_{j+1}, \cdots, y_n]\ .$$

Due to the orthogonality of M and N, $\mathscr{X}$ and $\mathscr{Y}$ are Berger bases. Moreover, corresponding r-element subsets $\{x_{\sigma(1)}, x_{\sigma(2)}, \cdots, x_{\sigma(r)}\}$ and $\{y_{\sigma(1)}, y_{\sigma(2)}, \cdots, y_{\sigma(r)}\}$ for each $\sigma \in Q_{n,r}$, are Berger bases for their respective linear spans, which we denote as follows:

$$M_\sigma = \mathrm{sp}\,[x_{\sigma(1)}, x_{\sigma(2)}, \cdots, x_{\sigma(r)}]$$

$$N_\sigma = \mathrm{sp}\,[y_{\sigma(1)}, y_{\sigma(2)}, \cdots, y_{\sigma(r)}]$$

for each $\sigma \in Q_{n,r}$.

Note that $x_{\sigma(r)}$ and $y_{\sigma(r)}$ are eigenvectors for (the compression of) A on the spaces M_σ and N_σ, respectively: This follows from (6.6a) and (6.6b). Thus, A has unbounded power on $x_{\sigma(r)}$ and $y_{\sigma(r)}$ when restricted to M_σ and N_σ (Proposition 3.6). Consequently, for x_z and y_z of Definition 3.3 (see also Lemma 3.2), $1 - |\langle x_z, x_{\sigma(r)}\rangle|^2 = 1 - |\langle y_z, y_{\sigma(r)}\rangle|^2 = 0$, since $x_z = x_{\sigma(r)}$ and $y_z = y_{\sigma(r)}$. Moreover, the u_z defined in hypothesis (iv) of Theorem 6.1, can be chosen to equal $u_{\sigma(r)}$, where

$$\lambda\langle Ax_{\sigma(r)}, x_{\sigma(r)}\rangle + (1-\lambda)\langle Ay_{\sigma(r)}, y_{\sigma(r)}\rangle = \langle Au_{\sigma(r)}, u_{\sigma(r)}\rangle\ .$$

This implies that $z^k[1 - |\langle u_z, u_{\sigma(r)}\rangle|^2] = 0$ so that all hypothesis of Theorem 6.1 are fulfilled by *each* of the orthogonal subspaces M_σ and N_σ as σ runs over $Q_{n,r}$. Therefore, we may write for each $\sigma \in Q_{n,r}$, that

(6.7) $$\lambda\langle C_r(A)\underset{\sim}{x}_\sigma, \underset{\sim}{x}_\sigma\rangle + (1-\lambda)\langle C_r(A)\underset{\sim}{y}_\sigma, \underset{\sim}{y}_\sigma\rangle = \langle C_r(A)\underset{\sim}{u}_\sigma, \underset{\sim}{u}_\sigma\rangle\ ,$$

where $\{u_1, u_2, \cdots, u_r, \cdots, u_n\}$ is an o.n. set satisfying the equations

$$\lambda\langle Ax_i, x_i\rangle + (1-\lambda)\langle Ay_i, y_i\rangle = \langle Au_i, u_i\rangle,\ i = 1, 2, \cdots, n\ .$$

If we sum each side of (6.7) over all $\sigma \in Q_{n\,r}$ (see Proposition 2.2 and (2.6)), we obtain

$$\lambda E_r(AM) + (1-\lambda)E_r(AN) = E_r(AU_\lambda)\ ,$$

where $U_\lambda = \mathrm{sp}\,[u_1, u_2, \cdots, u_r, \cdots, u_n]$. The theorem is proved.

REMARK. It is an open question as to whether $W_{r\,n}(A)$ ks always convex. It may be conjectured that if $2r > \dim \mathscr{H}$, then convexity is automatic since, in this case, every vector of $\Lambda^r\mathscr{H}$ is automatically decomposable. (See Remarks following Proposition 2.2.)

References

1. C. A. Berger *Normal dilations*, Ph.D. Thesis, Cornell University, 1963.
2. H. Cartan, *Theory of of analytic functions*, Addison Wesley (1963).
3. J. de Pillis, *Linear operators and their partitioned matrices* Arch. der Math., **22** (1971), 79-84.
4. W. F. Donoghue, Jr., *On the numerical range of a bounded operator*, Mich. Math. J., (1957), 261-263.
5. P. R. Halmos, *Numerical ranges and normal dilations*, Acta Szeged, vol. **25**, pp. 1-5.
6. ———, *A hilbert space problem book*, Van Nostrand, Princeton, N. J. (1967).
7. M. Heins, *Complex function theory*, Academic Press (New York, 1968).
8. S. MacLane and G. Birkhoff, *Algebra*, Macmillan (New York, 1967).
9. R. Raghavendran, *Toeplitz-Hausdorff theorem on numerical ranges*, Proc. Amer. Math. Soc., **20** (1969), 214-285.
10. M. H. Stone, *Hausdorff's theorem concerning Hermitian forms*, Bull. Amer. Math. Soc., **36** (1930), 259-261.
11. R. C. Thompson, *A note on normal matrices*, Can. J. Math., **15** (1963), 220-225.
12. K. Vala, *Sur la puissance exterieure d'un espace lineaire*, Ann. Acad. Sci. Fenn. A. I. 233 (1956).

Received October 10, 1969 and in revised form September 23, 1970. Part of this work was performed under the auspices of the U. S. Atomic Energy Commission and part was supported by NSF Grant GP-8758.

BROOKHAVEN NATIONAL LABORATORY
UPTON, LONG ISLAND, NEW YORK
AND
UNIVERSITY OF CALIFORNIA, RIVERSIDE