AN IDENTITY FOR NORMAL-LIKE OPERATORS

BY
T. I. SEIDMAN

ABSTRACT

For $\lambda \varepsilon \sigma(A)$ (A a bounded linear operator on a Hilbert space) with λ a boundary point of the numerical range, the 'spectral theory' for λ is 'just as if A were normal'. If A is *normal-like* (the smallest disk containing $\sigma(A)$ has radius $r = \inf_z \| A - z \|$), then also $\sup \{ \| Ax \|^2 - | \langle x.Ax \rangle |^2 : \| x \| = 1 \} = r^2$.

1. For any bounded subset S in the plane, it is easy to see that there is a unique closed disk $D_{q,r} = \{p: | p - q | \leqq r\}$ containing S with minimal radius $r = r(S)$. For any bounded normal operator A acting on a Hilbert space H

(1) $$\inf \{ \| A - z \| : z \in C\} = r(\sigma(A)).$$

We will call any bounded linear operator A *normal-like* if it satisfies (1). If we write

$$A_z = A - z, \quad n(A_*) = \inf_z \{ \| A_z \|^2 \}$$

(we will show that the inf in computing n_* is attained uniquely) then (1) becomes

$$n_*(A) = r^2(\sigma(A)).$$

For any bounded linear operator $A: H \to H$, define $\gamma(A;x)$ for $x \in H$ with $\| x \| = 1$ by

$$\gamma(A;x) = \| Ax \|^2 - | \langle Ax, x \rangle |^2.$$

Direct computation shows that $\gamma(A_z;x) = \gamma(A;x)$ where, as above, $A_z x = Ax - zx$ for scalar z. It is then immediate that $\gamma(A) \leqq n_*(A)$, where

$$\gamma(A) = \sup \{ \gamma(A;x) : x \in H, \| x \| = 1\}.$$

Recently, Z. Nehari pointed out to the author the inequality $\gamma(A) \leqq n_*(A)$ and raised the question: For which bounded linear operators A does one have

(2) $$\sup \{ \| Ax \|^2 - | \langle x, Ax \rangle |^2 : x \in H, \| x \| = 1\} = n_*(A)$$

* This research was partially supported by Air Force Contract AF–AFOSR–62–414.
Received October 10, 1968 and in revised form February 2, 1969

(i.e., $\gamma(A) = n_*(A)$)? V. J. Mizel showed [2] that (2) holds for hermitian matrices. The object of this paper is to show that (2) holds for normal-like operators on any Hilbert space H. It should be noted that (2) certainly does not characterize normal-like operators—direct computation shows that (2) holds for *every* 2×2 matrix A (i.e., H 2-dimensional) and there is reason to believe that this may be the case for infinite-dimensional H as well. Thus, this paper must be viewed as a partial result in this direction. Although a lemma in the present context, Theorem 1 has been so dignified as it seems of some interest in its own right.

2. Let A be any bounded linear operator in the Hilbert space H. Since the function $z \to \phi(z) = \| A_z \|^2$ from C to R is continuous and as we may clearly restrict our attention, in computing $\inf_z\{\phi(z)\}$, to the compact set $\{z : |z| \leqq 2\|A\|\}$, the infimum will certainly be attained. We show that ϕ is strictly convex from which it follows that the inf is *uniquely* attained; let $z(A)$, then, be the unique complex number such that $n_*(A) = \| A_{z(A)} \|^2$.

LEMMA 1. *The function* $z \to \| A - z \|^2$ *is a strictly convex function of* z.

Proof. Given $x \in H$ with $\|x\| = 1$, $u, v \in C$, $t, s \in (0,1)$ with $t + s = 1$, let $x_u = Ax - ux$, $x_v = Ax - vx$, $\phi(z;x) = \| A_z x \|^2$. Then

$$\|x_u\|^2 + \|x_v\|^2 - 2\,\mathrm{Re}\,\langle x_u, x_v\rangle = \|x_u - x_v\|^2 = |u - v|^2$$

and

$$\begin{aligned} \phi(tu + sv; x) &= \| t x_u + s x_v \|^2 \\ &= t\phi(u;x) + s\phi(v;x) - ts|u - v|^2 . \end{aligned}$$

Thus, since $\phi(z) = \sup\{\phi(z;x) : x \in H,\ \|x\| = 1\}$, one has

$$\phi(tu + sv) \leqq t\phi(u) + s\phi(v) - ts|u - v|^2$$

so that ϕ is strictly convex. ‖

Let $\sigma_*(A)$ be the set of spectral points of A on the boundary of the numerical range; i.e., $\sigma_*(A) = \sigma(A) \cap \partial w(A)$ where $w(A) = \{\langle Ax, x\rangle : x \in H,\ \|x\| = 1\}$. We show that the 'spectral theory' for points in $\sigma_*(A)$ is just as in the case of a normal operator. More precisely, we have the following results.

THEOREM 1. *Let* $\lambda_0 \in \sigma_*(A)$. *Then*

(a) λ_0 *is in the approximate point spectrum of* A,

(b) *if* $\{x_n\}$ *is an approximate eigenvector* (*aev*) *of* A *then it is also an aev of* A^*,

(c) *if* $\lambda \neq \lambda_0$ *is also in the approximate point spectrum of* A *and* $\{x_n\}$, $\{y_n\}$ *are aev's associated with* λ_0, λ *respectively then they are 'ultimately orthogonal'*,

(d) *the index of* λ_0 *is* 1 *so, if* $\|x_n\| = 1$ *and* $\|(A - \lambda_0)^m x_n\| \to 0$ *for some* $m \geqq 2$, *then* $\|(A - \lambda_0)x_n\| \to 0$ *and* $\{x_n\}$ *is an aev of* A *associated with* λ_0.

Proof. As $\sigma(A) \subseteq \overline{w(A)}$, λ_0 must be a boundary point of $\sigma(A)$ and so is in the approximate point spectrum of A. Now $w(A)$ is a convex set in C and, as every real-linear functional on C is of the form $z \to l_\theta(z) = \operatorname{Re}\theta z$ with $\|l_\theta\| = |\theta|$, there is, by the Hahn-Banach Theorem, a complex number θ with $|\theta| = 1$ such that $\operatorname{Re}\theta z \geqq \operatorname{Re}\theta\lambda$ for $z \in \overline{w(A)}$. Replacing the operator A by $\theta(A - \lambda_0)$, there is no loss of generality in assuming, for the remainder of this proof, that $\lambda_0 = 0$ and $\theta = 1$ so $0 \in \sigma_*(A)$ and

$$(*) \qquad \operatorname{Re}\langle Ax, x\rangle \geqq 0, \qquad x \in H.$$

If, now, $\|x\| = 1$, we have (setting $\alpha = 1/2\|A\|$ and $y = x - \alpha A^*x$ so $\|y\| \leqq 3/2$

$$\begin{aligned} 0 &\leqq \operatorname{Re}\langle Ay, y\rangle = \operatorname{Re}\langle Ax, y\rangle - \alpha\operatorname{Re}\langle AA^*x, x\rangle + \alpha^2\operatorname{Re}\langle AA^*, x, A^*x\rangle \\ &\leqq \|Ax\|\,\|y\| - \alpha\|A^*x\|^2 + \alpha^2\|A\|\,\|A^*x\|^2 \\ &\leqq 3\|Ax\|/2 - \alpha(1 - \alpha\|A\|)\|A^*x\|^2 = (3\|Ax\| - \alpha\|A^*x\|^2)/2 \end{aligned}$$

whence

$$(**) \qquad \|A^*x\| \leqq \sqrt{6\|A\|.\|Ax\|}, \qquad \text{if } \|x\| = 1.$$

This gives (b) on Putting $x = x_n$. If $\|x\| = \|y\| = 1$ with $\|Ax\|$, $\|A_\lambda y\| \leqq \varepsilon$, then

$$\begin{aligned} |\lambda\langle y, x\rangle| &= |\langle Ay, x\rangle - \langle A_\lambda y, x\rangle| \\ &\leqq \|y\|\,\|A^*x\| + \|A_\lambda y\|\,\|x\| \leqq \sqrt{6\|A\|}\varepsilon + \varepsilon. \end{aligned}$$

Thus if $\|x_n\| = \|y_n\| = 1$ and $\|Ax_n\| \to 0$, $\|A_\lambda y_n\| \to 0$ with $\lambda \neq 0$, one has $\langle y_n, x_n\rangle \to 0$ so the aev's $\{x_n\}$, $\{y_n\}$ are ultimately orthogonal. Suppose, for some $m \geqq 2$, $\|A^m x_n\| \to 0$. If $\|A^{m-1}x_n\| \nrightarrow 0$, then (taking a subsequence if necessary) we could set $y_n = A^{m-2}x_n/\|A^{m-2}x_n\|$ and have $\|y_n\| = 1$, $\|Ay_n\| \geqq \beta > 0$ and $\|A^2 y_n\| \to 0$. With $w_n = y_n - \alpha Ay_n$ $(\alpha > 0)$ we now would have $\|w_n\| \leqq 1 + \alpha\|A\|$ and

$$\begin{aligned} 0 &\leqq \operatorname{Re}\langle Aw_n, w_n\rangle \\ &= \operatorname{Re}\langle Ay_n, y_n\rangle - \alpha\|Ay_n\|^2 - \alpha\operatorname{Re}\langle A^2y_n, w_n\rangle \\ &\leqq \|A\| - \alpha\beta^2 + \alpha\|A^2y_n\|(1 + \alpha\|A\|) \end{aligned}$$

which would be a contradiction for $\alpha > \|A\|/\beta^2$ and n so large that the third term is negligible. Thus, $\| A^m x_n \| \to 0$ for $m \geqq 2$ implies $\| A^{m-1} x_n \| \to 0$ and (d) follows by induction on m. ‖

We shall need the following results for bounded closed sets S in the plane.

LEMMA 2. (a) *There is a unique minimal center* $q(S)$ *such that the closed disk* $D(S) = \{p: |p - q(S)| \leqq r(S)\}$ *contains* S.

(b) *Letting* S^* *be the convex hull of* S, *one has* $q(S) \in S^*$.

(c) *Letting* $S^0 = S \cap \partial D(S)$, *one has* $D(S) = D(S^0)$.

REMARKS. It is interesting to note that these results hold for spaces more general than the plane. (a) has already been observed and holds in any uniformly convex Banach space; c,f., [1]. (b) is known to hold in any 2-dimensional Banach space in which $D(S)$ is uniquely defined and in any Hilbert space (in fact, for dimensions greater than 2 it is known to characterize Hilbert space). (c) holds in uniformly convex Banach spaces if S is compact; it asserts that $D(S)$ is determined by 'extreme points' of S. For proofs, we refer the reader to, e.g., the preliminary version [3] of this paper.

3. We are now ready to prove (2) for normal-like operators A on H. It may be remarked that (2) can be shown equivalent to the existence of an aev $\{x_n\}$ of P associated with $\| P \|$ such that $\langle x_n, Ux_n \rangle \to 0$, where UP is the polar decomposition of $[A - z(A)]$.

THEOREM 2. *Let* A *be a bounded normal-like linear operator on the Hilbert space* H. *Then* (2) *holds* (i.e., $\gamma(A) = n_*(A) = r^2(\sigma(A))$.

Proof. Without loss of generality we may assume $z(A) = 0$ and $r^2(\sigma(A)) = n_*(A) = 1$ so that $D(\sigma(A))$ is just the unit disk D_1 centered at 0. Letting $\sigma_0 = \{z \in \sigma(A): |z| = 1\}$, it follows from the above and from (c) of Lemma 2 that $D(\sigma_0) = D_1$ and from (b) of Lemma 2 that 0 is in the convex hull of σ_0. Thus, there is a finite subset $\{\lambda_1, \cdots, \lambda_K\}$ of σ_0 and positive real numbers $\{c_1, \cdots, c_K\}$ such that

$$0 = \Sigma_k c_k^2 \lambda_k, \quad \Sigma_k c_k^2 = 1.$$

Clearly, each λ_k is not only a boundary point of $\sigma(A)$ but, as $w(A) \subseteq \{z: z \leqq \| A \|\}$ which is just D_1 by the definitions of $n_*(A)$ and of $z(A)$, each λ_k lies in $\partial w(A)$. By Theorem 1 it now follows that, for any $\varepsilon > 0$, there are $\{x_1, \cdots, x_K\}$ in H such that

$$\| x_k \| = 1, \quad \| (A - \lambda_k) x_k \| \leqq \varepsilon, \quad |\langle x_j x_k \rangle| \leqq \varepsilon \quad \text{for} \quad j \neq k.$$

Now set

$$x = x_\varepsilon = \sum_k c_k x_k, \quad y = \sum_k c_k \lambda_k x_k.$$

Then

$$\left| 1 - \| x \|^2 \right| = \left| 1 - \sum_k c_k^2 \| x_k \|^2 - \sum_{j \neq k} c_j c_k \langle x_j, x_k \rangle \right|$$

$$\leqq \sum_{j \neq k} c_j c_k \left| \langle x_j, x_k \rangle \right| \leqq K(K-1)\varepsilon,$$

$$\| Ax - y \| \leqq \sum_k c_k \| (A - \lambda_k) x_k \| \leqq K\varepsilon,$$

and

$$\| y \|^2 = \sum_k c_k \left| \lambda_k \right|^2 \| x_k \|^2 - \sum_{j \neq k} c_j c_k \left| \lambda_j \lambda_k \right| \left| \langle x_j, x_k \rangle \right|$$

$$\geqq 1 - \varepsilon \sum_{j \neq k} c_j c_k \geqq 1 - K(K-1)\varepsilon.$$

Thus, since $1 \geqq \| Ax \| \geqq \| y \| - \| Ax - y \|$, we have $\| Ax_\varepsilon \|^2 \to 1$ as $\varepsilon \to 0$. At the same time,

$$\left| \langle x, y \rangle \right| \leqq \sum_k c_k^2 \lambda_k \| x_k \|^2 \Big| + \sum_{j \neq k} c_j c_k \left| \lambda_k \right| \left| \langle x_j, x_k \rangle \right|$$

$$= 0 + \sum_{j \neq k} c_j c_k \left| \langle x_j, x_k \rangle \right| \leqq K(K-1)\varepsilon$$

so that

$$\left| \langle x, Ax \rangle \right| \leqq \left| \langle x, y \rangle \right| + \left| \langle x, Ax - y \rangle \right| \leqq K^2 \varepsilon.$$

We have $\gamma(A) \leqq n_*(A) = 1$ and the estimates above for $x = x_\varepsilon$ show that $\gamma(A; x_\varepsilon) \to 1$ as $\varepsilon \to 0$ so $\gamma(A) = 1 = n_*(A)$. ‖

I should like to acknowledge the referee's observation that (2) may be formulated as a 'minimax principle', restating it in the form

$$(2') \qquad \sup_x \inf_z \| A_z x \| = \inf_z \sup_x \| A_z x \|$$

(taken for $x \in H$ with $\| x \| = 1$ and $z \in C$). This follows on noting that

$$\inf_z \| A_z x \|^2 = \| Ax \|^2 - \left| \langle Ax, x \rangle \right|^2 = \gamma(A; x)$$

($\| x \| = 1$). The validity of (2′) might now be considered for operators on a general Banach space.

References

1. V. Klee, Math. Reviews **13** (1952), 661.
2. V. J. Mizel, Personal communication.
3. T. I. Seidman, *On an inequality for operators on Hilbert space*, Report 68–7, Carnegie-Mellon University (1968).

Carnegie-Mellon University,
Pittsburg, Pennsylvania