AN ELEMENTARY PROOF OF THE POWER INEQUALITY FOR THE NUMERICAL RADIUS

Carl Pearcy

Let $\mathscr{H}$ be a complex Hilbert space, finite- or infinite-dimensional. With each bounded linear operator A on $\mathscr{H}$ there is associated the nonnegative number

$$w(A) = \sup_{\|x\|=1} |(Ax, x)|,$$

called the *numerical radius* of A. Many properties of the function w are well known and quite elementary. Among these are the following:

(1) $w(A) = 0$ if and only if $A = 0$;

(2) $w(\lambda A) = |\lambda|\, w(A)$ for every scalar λ;

(3) $w(A + B) \leq w(A) + w(B)$;

(4) $\frac{1}{2}\|A\| \leq w(A) \leq \|A\|$;

(5) $r(A) \leq w(A)$, where $r(A)$ denotes the spectral radius of A;

(6) w is not continuous in either the weak or the strong operator topology, but is continuous in the uniform operator topology.

When we inquire into the multiplicative properties of w, the inequality

$$w(AB) \leq w(A)\, w(B) \tag{*}$$

naturally comes to mind. It is easy to see that (*) cannot hold universally, and perhaps somewhat more surprising that (*) can fail for commutative A and B. In fact, the following example, pointed out by Arlen Brown and Allen Shields, shows that (*) can fail when A and B are powers of the same operator. Let N denote the nilpotent operator on a 4-dimensional space whose matrix relative to some orthonormal basis is

$$\begin{pmatrix} 0 & 1 & & \\ & 0 & 1 & \\ & & 0 & 1 \\ & & & 0 \end{pmatrix}.$$

Easy calculations show that $w(N) < 1$, $w(N^2) = 1/2$, and $w(N^3) = 1/2$. Thus (*) fails with $A = N$ and $B = N^2$.

In spite of all this unpleasantness, the following theorem is true.

THEOREM. *If* A *is an operator on* $\mathscr{H}$, *and* n *is a positive integer, then*

Received April 26, 1966.

(H) $$w(A^n) \le [w(A)]^n.$$

Before proceeding to a proof of this inequality, I briefly discuss its background. In [4], Lax and Wendroff proved that if $\mathscr{H}$ is finite-dimensional and A is an operator on $\mathscr{H}$ satisfying $w(A) \le 1$, then A is power bounded; that is, there exists a constant M such that $\|A^n\| \le M$ for all positive integers n. (The constant M depends on the dimension k of $\mathscr{H}$, and it tends to infinity with k.)

P. R. Halmos, in trying to extend this result to infinite-dimensional spaces, conjectured a stronger result—namely (H). (To see that (H) is stronger, note that if $w(A) \le 1$, it follows from (H) and property (4) of w that $\|A^n\| \le 2$ for all positive integers n.) The special case of (H) in which n is a power of 2 was proved in [2], and the special case in which $\mathscr{H}$ is 2-dimensional was proved by Arlen Brown. The arguments for the special cases refused to generalize, however, and the first complete proof of (H) was a beautiful argument given by Charles Berger [1], who used dilation theory. Berger's argument was simplified considerably by Stampfli, and after consideration of Stampfli's simplification, I was able to extract an argument that is essentially elementary. The argument depends upon the following identity, which seems interesting in its own right.

LEMMA. *If* x *is a unit vector,* A *is an operator, and* n *is a positive integer, then*

(**) $$1 - (A^n x, x) = \frac{1}{n}\sum_{j=1}^{n} \|x_j\|^2 \left[1 - w_j\left(\frac{Ax_j}{\|x_j\|}, \frac{x_j}{\|x_j\|}\right)\right],$$

where

$$w_j = e^{2\pi i j/n} \quad \text{and} \quad x_j = \left(\prod_{\substack{k=1\\k\neq j}}^{n} (1 - w_k A)\right)x \qquad (j = 1, 2, \cdots, n).$$

Proof. Note first the two polynomial identities

$$1 - z^n \equiv \prod_{k=1}^{n} (1 - w_k z)$$

and

$$1 \equiv \frac{1}{n}\sum_{j=1}^{n}\prod_{\substack{k=1\\k\neq j}}^{n} (1 - w_k z).$$

Since these identities remain valid with z replaced by A, they imply that

$$\frac{1}{n}\sum_{j=1}^{n} \|x_j\|^2 \left[1 - w_j\left(\frac{Ax_j}{\|x_j\|}, \frac{x_j}{\|x_j\|}\right)\right]$$

$$= \frac{1}{n}\sum_{j=1}^{n} ([1 - w_j A]x_j, x_j) = \frac{1}{n}\sum_{j=1}^{n} ([1 - A^n]x, x_j)$$

$$= \left([1 - A^n]x, \frac{1}{n} \sum_{j=1}^{n} \left[\prod_{\substack{k=1 \\ k \neq j}}^{n} (1 - w_k A) \right] x \right)$$

$$= ([1 - A^n]x, x) = 1 - (A^n x, x).$$

Proof of the theorem. By virtue of property (2) of w, it suffices to assume that $w(A) \leq 1$ and to prove that $w(A^n) \leq 1$. Let x be any unit vector, and let θ be any real number; replacing A in (**) by $e^{i\theta}A$, we obtain the relation

$$1 - e^{in\theta}(A^n x, x) = \frac{1}{n} \sum_{j=1}^{n} \|x_j\|^2 \left[1 - w_j e^{i\theta} \left(\frac{Ax_j}{\|x_j\|}, \frac{x_j}{\|x_j\|} \right) \right].$$

Since $w(A) \leq 1$, the real part of each term on the right-hand side of this equation is nonnegative, so that $\Re[1 - e^{in\theta}(A^n x, x)] \geq 0$. The validity of this inequality for all real θ implies that $|(A^n x, x)| \leq 1$, and the theorem follows.

Remark. After this note was written, Kato [3] obtained a generalization of the above power inequality. Theorem 5 of [3], which Kato mistakenly attributes to Sz.-Nagy, is actually due to Stampfli.

REFERENCES

1. C. Berger, *On the numerical range of powers of an operator* (to appear).
2. S. J. Bernau and F. Smithies, *A note on normal operators*, Proc. Cambridge Philos. Soc. 59 (1963), 727-729.
3. T. Kato, *Some mapping theorems for the numerical range*, Proc. Japan Acad. 41 (1965), 652-655.
4. P. Lax and B. Wendroff, *Difference schemes with high order of accuracy for solving hyperbolic equations*, AEC Research and development report, Courant Institute of Mathematical Sciences (1962), 1-49.

The University of Michigan