PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 115, Number 1, May 1992

A SHARP ESTIMATE IN AN OPERATOR INEQUALITY

R. McEACHIN

(Communicated by Palle E. T. Jorgensen)

ABSTRACT. Let $\mathscr{H}$ and $\mathscr{K}$ be Hilbert spaces, and suppose $A \in \mathscr{B}(\mathscr{H})$ and $B \in \mathscr{B}(\mathscr{K})$ are selfadjoint operators with $\operatorname{dist}(\sigma(A), \sigma(B)) \geq \delta > 0$. It is known that for any $Q \in \mathscr{B}(\mathscr{K}, \mathscr{H})$ we must have $\frac{\pi}{2}\|AQ - QB\| \geq \delta\|Q\|$. In this paper we give examples proving that $\pi/2$ is sharp in this inequality.

1. INTRODUCTION

In 1983 Bhatia, Davis, and McIntosh (see [3]) proved the following

Theorem. *Suppose $\mathscr{H}$ and $\mathscr{K}$ are Hilbert spaces, and let $A \in \mathscr{B}(\mathscr{H})$ and $B \in \mathscr{B}(\mathscr{K})$ be selfadjoint operators with* $\operatorname{dist}(\sigma(A), \sigma(B)) \geq \delta > 0$. *Then there is some constant c (independent of A and B) such that $c|||AQ-QB||| \geq \delta|||Q|||$ for any $Q \in \mathscr{B}(\mathscr{K}, \mathscr{H})$ and any unitary invariant norm $|||\cdot|||$.*

We will say a norm is unitary invariant, or *ui*, if $|||UTV||| = |||T|||$ whenever U and V are unitary. Let c be the smallest constant such that, with A, Q, and B as above, the inequality

$$(*) \qquad c|||AQ - QB||| \geq \delta|||Q|||$$

holds for arbitrary *ui* norms. In [3] the authors establish $1.22 < c < 2$, and in [7] Sz.-Nagy proves $c \leq \pi/2$. (This also follows from his earlier paper [6].) The purpose of this paper is to prove $c = \pi/2$. To do this we give $n \times n$ matrices A_n, Q_n, and B_n such that A_n and B_n are selfadjoint and $\delta\|Q_n\|/\|A_nQ_n - Q_nB_n\| \to \pi/2$ as $n \to \infty$. Here $\|\cdot\|$ is the operator norm, and since this is unitary invariant it follows that $\pi/2$ is sharp in $(*)$.

This inequality can be viewed as geometrical in nature. As seen from [3, 4], it is an important tool in the analysis of the spectral variation of selfadjoint or normal operators. It is well known that if a selfadjoint matrix S undergoes a small selfadjoint perturbation H then its eigenvalues cannot change very much. Results of this nature, for example, are surveyed in [1, Chapter 3]. Eigenvectors, however, can rotate through a large angle. For instance, two operators very close to I, and thus close to each other, could still have eigenspaces far apart. But suppose $\sigma(S)$ can be divided into two or more parts separated from each other.

Received by the editors November 12, 1990.
1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 47A55, 47B15; Secondary 15A45.

Then for $T = S + H$, $\sigma(T)$ must have corresponding separated parts when $|||H|||$ is small, and a spectral subspace for one part of $\sigma(T)$ must be almost orthogonal to a spectral subspace for a different part of $\sigma(S)$. In this case, then, a small perturbation cannot rotate a spectral projection through a large angle. While bounds on the variation of $\sigma(S)$ are classical, bounds on the variation of spectral projections are newer and are obtained as a corollary to the $AQ - QB$ inequality. Furthermore, this inequality and its corollaries can be generalized to yield important bounds on the variation of both the spectrum and the spectral subspaces of a normal operator. These results are chiefly developed in [3, 4], and the discussions in [1, Chapter 4; 2] explain their interconnections very well.

In the examples below, if n is understood we let (a_{ij}) denote the $n \times n$ matrix whose entry in the ith row and jth column is a_{ij}. The following discussion was first presented in my doctoral dissertation at the University of Illinois. At this time I wish to thank my advisor, Professor I. D. Berg, as well as Professor Chandler Davis at the University of Toronto for their assistance.

2. The examples

For each n, let $A_n = \operatorname{diag}(2n, 2n-2, 2n-4, \dots, 2)$ and let $B_n = \operatorname{diag}(1, 3, 5, \dots, 2n-1)$. For $1 \le i \le n$, set $\lambda_i = (2n - 2i + 2) \in \sigma(S)$ and $\mu_i = (2i-1) \in \sigma(T)$. Then $\delta = \operatorname{dist}(\sigma(A), \sigma(B)) = 1$, and for any $Q_n = (q_{ij})$ we have $A_nQ_n - Q_nB_n = ((\lambda_i - \mu_j)q_{ij})$. Now set $Q_n = ((\lambda_i - \mu_j)^{-2})$, so that $C_n := A_nQ_n - Q_nB_n = ((\lambda_i - \mu_j)^{-1}) =$

$$\begin{bmatrix} \frac{1}{2n-1} & \frac{1}{2n-3} & \frac{1}{2n-5} & \cdots & 1 \\ \frac{1}{2n-3} & \frac{1}{2n-5} & \frac{1}{2n-7} & \cdots & -1 \\ \frac{1}{2n-5} & \frac{1}{2n-7} & \frac{1}{2n-9} & \cdots & -\frac{1}{3} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & -1 & -\frac{1}{3} & \cdots & \frac{-1}{2n-3} \end{bmatrix}.$$

The following three propositions yield values for $\lim_n \|C_n\|$ and $\lim_n \|Q_n\|$.

Proposition 1. *For* $\mathbf{x} = (\dots, -1/5, -1/3, -1, 1, 1/3, 1/5, \dots) \in \ell_2(Z)$ *and* S *the right shift* $\mathbf{e}_i \mapsto \mathbf{e}_{i+1}$, *we have* $\langle S^n\mathbf{x}, \mathbf{x}\rangle = 0$ *for all* $n \ne 0$.

Proof. It is sufficient to consider $n > 0$. Let

$$\mathbf{x} = \left(\dots, -\frac{1}{3}, -1, 1, \frac{1}{3}, \dots, \frac{1}{2n-1}, \frac{1}{2n+1}, \frac{1}{2n+3}, \dots\right),$$
$$S^n\mathbf{x} = \left(\dots, \frac{-1}{2n+3}, \frac{-1}{2n+1}, \frac{-1}{2n-1}, \frac{-1}{2n-3}, \dots, -1, 1, \frac{1}{3}, \dots\right).$$

Note that there are n positions where the entries have opposite signs and in all other positions the entries have the same sign. Then $\langle S^n\mathbf{x}, \mathbf{x}\rangle$ is the sum of n negative terms and two (identical) series of positive terms. The negative terms contribute

$$\begin{aligned}\sum_{k=0}^{n-1} \frac{1}{2k+1} \cdot \frac{1}{2k-(2n-1)} &= \frac{-1}{2n} \sum_{k=0}^{n-1} \left[\frac{1}{2k+1} - \frac{1}{2k-(2n-1)} \right] \\ &= \frac{-1}{2n} \left[\sum_{k=0}^{n-1} \frac{1}{2k+1} + \sum_{m=0}^{n-1} \frac{1}{2m+1} \right] \qquad (m = n-k+1) \\ &= \frac{-1}{n} \sum_{k=0}^{n-1} \frac{1}{2k+1}.\end{aligned}$$

On the other hand, the positive terms sum to

$$2\sum_{k=0}^{\infty} \frac{1}{2k+1} \cdot \frac{1}{2n+2k+1} = \frac{1}{n} \sum_{k=0}^{\infty} \left[\frac{1}{2k+1} - \frac{1}{2n+2k+1} \right].$$

This is a telescoping sum after cancellation it reduces to $n^{-1} \sum_{k=0}^{n-1} 1/(2k+1)$, and so $\langle S^n \mathbf{x}, \mathbf{x} \rangle = 0$ as desired. □

Proposition 2. *For* C_n *as above,* $\|C_n\| \nearrow \pi/2$ *as* $n \to \infty$.

Proof. Consider the doubly infinite matrix

$$T = \begin{bmatrix} & \vdots & \vdots & \vdots & \vdots & \vdots & \\ \dots & 1/9 & 1/7 & 1/5 & 1/3 & 1 & \dots \\ \dots & 1/7 & 1/5 & 1/3 & 1 & -1 & \dots \\ \dots & 1/5 & 1/3 & 1 & -1 & -1/3 & \dots \\ \dots & 1/3 & 1 & -1 & -1/3 & -1/5 & \dots \\ \dots & 1 & -1 & -1/3 & -1/5 & -1/7 & \dots \\ & \vdots & \vdots & \vdots & \vdots & \vdots & \end{bmatrix}.$$

Each column has norm $(2(1+1/3^2+1/5^2+\cdots))^{1/2} = \pi/2$, and by Proposition 1 the columns are mutually orthogonal. Thus T is a multiple of a unitary operator on $\ell_2(Z)$ and $\|T\| = \pi/2$. Since C_n is a submatrix of T for any n, $\|C_n\| \le \pi/2$. Furthermore, $\|C_n\|$ increases because C_n is a submatrix of C_{n+1}. By considering the lengths of the columns of C_n, it is immediate that $\|C_n\| \nearrow \pi/2$ as $n \to \infty$. □

Proposition 3. *For* Q_n *as above,* $\|Q_n\| \nearrow \pi^2/4$ *as* $n \to \infty$.

Proof. We know that $\frac{\pi}{2}\|C_n\| \ge \|Q_n\|$, and so by Proposition 2 we have $\pi^2/4 \ge \|Q_n\|$ for each n. Furthermore $\|Q_n\|$ increases because each Q_n is a submatrix of Q_{n+1}. We will show that for any $\varepsilon > 0$ there is some n such that $\|Q_n\| > \pi^2/4 - \varepsilon$. To this end define $s_n = 2(1 + 1/3^2 + 1/5^2 + \cdots + 1/(2n-1)^2)$, and note that $\{s_n\} \nearrow \pi^2/4$ as $n \to \infty$. For each n let $\mathbf{x}$ be the n-vector $[1, 1, 1, \dots, 1]^T$ and $\mathbf{y} = Q_n \mathbf{x} =$

$$\left[\sum_{k=1}^{n} \frac{1}{(2k-1)^2}, \sum_{k=1}^{n} \frac{1}{(2k-3)^2}, \sum_{k=1}^{n} \frac{1}{(2k-5)^2}, \dots, \sum_{k=1}^{n} \frac{1}{(2k-2n+1)^2} \right]^T.$$

If y_i is the ith component of $\mathbf{y}$ observe that $y_2, y_n > s_1$. Similarly, $y_3, y_{n-1} > s_2$, $y_4, y_{n-2} > s_3$, and so on. If $\varepsilon > 0$, choose p large enough so that $s_p > \pi^2/4 - \varepsilon$. Then take r so large that $\sqrt{r}/\sqrt{r+2p+3} > s_p/s_{p+1}$ and set $n = r + 2p + 3$. Then all coordinates of $\mathbf{y}$ except the first $p+2$ and the last $p+1$ are larger than s_{p+1}; that is, the r components $y_{p+3}, y_{p+4}, \dots, y_{p+r+2}$ are all greater than s_{p+1}, and so $\|\mathbf{y}\| > \sqrt{r}s_{p+1}$. Since $\|\mathbf{x}\| = \sqrt{n}$, $\|Q_n\| > \sqrt{r}s_{p+1}/\sqrt{n} = \sqrt{r}s_{p+1}/\sqrt{r+2p+3} > s_p > \pi^2/4 - \varepsilon$. This establishes our proposition. □

We now observe that since $\delta = 1$, $\lim_n \|Q_n\| = \pi/2$, and $\lim_n \|C_n\| = \lim_n \|A_nQ_n - Q_nB_n\| = \pi^2/4$, it follows that $\delta\|Q_n\|/\|A_nQ_n - Q_nB_n\| \to \pi/2$ as $n \to \infty$. Therefore $c \geq \pi/2$, where c is the constant in $(*)$. Since $c \leq \pi/2$ is known, the value $\pi/2$ is sharp.

3. Concluding remarks

In this section we indicate an alternate derivation of $\lim_n\|C_n\|$ and $\lim_n\|Q_n\|$. First, note that $Q_n = C_n \circ C_n$, where $A \circ B$ denotes the Schur product of A and B. (That is, if $A = (a_{ij})$ and $B = (b_{ij})$ then $A \circ B = (a_{ij}b_{ij})$.) Then $\lim_n \|C_n\|$ and $\lim_n \|Q_n\|$ are easily deducible from the fact that, for T as above, $\|T\| = \pi/2$ and $\|T \circ T\| = \pi^2/4$.

To prove $\|T\| = \pi/2$, let $\mathscr{H} = L^2(-1/2, 1/2)$, and define

$$f(x) = \begin{cases} i & \text{if } x \in (-1/2, 0], \\ -i, & \text{if } x \in (0, 1/2). \end{cases}$$

Then let $T_f \colon \mathscr{H} \to \mathscr{H}$ be the linear map with $T_fh = fh$. If we choose the orthonormal basis $\{\exp(2\pi inx)\}_{n=-\infty}^{\infty}$ for $\mathscr{H}$, the matrix of T_f in this basis is $M = (a_{ij})_{i,j=-\infty}^{\infty}$, where

$$a_{ij} = \begin{cases} 0 & \text{if } i-j \text{ is even}, \\ -2/\pi(i-j) & \text{if } i-j \text{ is odd}. \end{cases}$$

One can show that $\|T\| = \frac{\pi}{2}\|M\|$ and $\|M\| = \|f\|_{L^\infty} = 1$. Therefore, $\|T\| = \pi/2$. The proof for $\|T \circ T\|$ is similar. Define

$$g(x) = \begin{cases} 4x+1 & \text{if } x \in (-1/2, 0], \\ -4x+1 & \text{if } x \in [0, 1/2). \end{cases}$$

Then T_g has matrix $M \circ M$, and in the same way as before, $\|T \circ T\| = \frac{\pi^2}{4}\|M \circ M\|$ and $\|M \circ M\| = \|g\|_{L^\infty} = 1$. Therefore $\|T \circ T\| = \pi^2/4$.

With a little more work, one can now conclude that $\|C_n\| \to \pi/2$ and $\|Q_n\| \to \pi^2/4$ as $n \to \infty$.

References

1. R. Bhatia, *Perturbation bounds for matrix eigenvalues*, Pitman Research Notes in Math., no. 162, Longman, Essex, 1987.
2. R. Bhatia, Ch. Davis, and P. Koosis, *An extremal problem in Fourier analysis with applications to operator theory*, J. Funct. Anal. **82** (1989), 138–150.
3. R. Bhatia, Ch. Davis, and A. McIntosh, *Perturbation of spectral subspaces and solution of linear operator equations*, Linear Algebra Appl. **52–53** (1983), 45–67.
4. Ch. Davis and W. M. Kahan, *The rotation of eigenvectors by a perturbation.* III, SIAM J. Numer. Anal. **7** (1970), 1–46.

5. R. McEachin, *Analysis of an inequality concerning perturbation of self-adjoint operators*, Doctoral disseration, Univ. of Illinois, Urbana, IL, October 1990.
6. B. Sz.-Nagy, *Über die Ungleichung von H. Bohr*, Math. Nachr. **9** (1953), 225–259.
7. B. Sz.-Nagy, *Bohr inequality and an operator equation*, Operators in indefinite metric spaces, scattering theory, and other topics (Bucharest, 1985), pp. 321–327; Oper. Theory: Adv. Appl., vol. 24, Birkhauser, Basel and Boston, MA, 1987.

DEPARTMENT OF MATHEMATICAL SCIENCES, INDIANA UNIVERSITY-PURDUE UNIVERSITY AT FORT WAYNE, FORT WAYNE, INDIANA 46805

Current address: Department of Mathematics, University of the West Indies, Mona, Kingston 7, Jamaica