JOURNAL OF FUNCTIONAL ANALYSIS 33, 332–338 (1979)

Almost Commuting Matrices*

CARL PEARCY AND ALLEN SHIELDS

Department of Mathematics, University of Michigan, Ann Arbor, Michigan 48109

Communicated by Peter D. Lax

Received April 25, 1978

It is shown that if A and B are $n \times n$ complex matrices with $A = A^*$ and $\| AB - BA \| \leqslant 2\epsilon^2/(n-1)$, then there exist $n \times n$ matrices A' and B' with $A' = A'^*$ such that $A'B' = B'A'$ and $\| A - A' \| \leqslant \epsilon$, $\| B - B' \| \leqslant \epsilon$.

We consider operators (that is, bounded linear transformations) on a finite-dimensional Hilbert space $\mathscr{H}$ (over the complex numbers) with the usual operator norm

$$\| T \| = \sup\{\| Tf \| : f \in \mathscr{H}, \| f \| \leqslant 1\}.$$

We obtain the following result.

THEOREM 1. *Let $\mathscr{H}$ be an n-dimensional complex Hilbert space and let A and B be operators on $\mathscr{H}$, with A self-adjoint. Let $\epsilon > 0$ be given. If*

$$\| AB - BA \| \leqslant \frac{2\epsilon^2}{n-1}, \tag{2}$$

then there exist two operators A' and B' on $\mathscr{H}$, with A' self-adjoint, such that $A'B' = B'A'$ and

$$\| A - A' \| < \epsilon, \qquad \| B - B' \| \leqslant \epsilon. \tag{3}$$

Furthermore, if B is self-adjoint then B' can be chosen to be self-adjoint also.

Note the asymmetry in (3), with strict inequality for one term but not for the other. The proof will show that we could also have the weak inequality for A and the strict inequality for B. (*See the comment at the end of the* proof.)

This theorem gives a quantitative answer (in case A is self-adjoint) to a problem raised by Rosenthal [9] about whether a pair of $n \times n$ matrices that "almost commute" is close to another pair that do commute. More precisely,

* This work was supported in part by NSF.

0022-1236/79/090332-07$02.00/0

the following is a slight reformulation of his question. Let $\mathscr{H}$ be an n-dimensional complex Hilbert space. Then for each $\epsilon > 0$, does there exist $\delta = \delta(\epsilon, n) > 0$ such that if A, B are operators on $\mathscr{H}$ with $\| AB - BA \| < \delta$, then there are operators A', B' on $\mathscr{H}$ that commute and for which (3) holds? He suggested that perhaps one should require in addition that

$$\| A \| \leqslant 1, \| B \| \leqslant 1. \tag{4}$$

Our result gives an affirmative answer in case A is self-adjoint and shows that in this case $\delta(\epsilon, n) \geqslant 2\epsilon^2/(n-1)$; we do not know what the largest possible δ is.

An affirmative answer (when (4) is assumed, but with no assumption of self-adjointness) was first given by Luxemburg and Taylor [6], using nonstandard analysis. "Standard" proofs were then given, but not published, by P. R. Halmos and W. Kahan. Halmos' proof later appeared as Lemma 1 in [1]. J. Deddens, in an unpublished preprint [4], showed that if A, B are both self-adjoint, and if $\| AB - BA \| \leqslant \epsilon^2/n^2$, then a pair of commuting self-adjoint operators A', B' can be found that satisfy (3). (Our work and his work were done independently of one another at about the same time—namely, early in 1972.) His proof is simpler than ours; it uses the Hilbert-Schmidt norm as a tool. This norm is easier to work with, but in passing to the regular norm some precision is lost. (His proof works even if B is not assumed to be self-adjoint; in this case B' will usually not be self-adjoint.) Note that neither our result nor Deddens' result requires (4). On the other hand, to establish either of our results it would be sufficient to establish it in the special case when (4) is assumed, the general case then would follow upon division by a constant. Thus (4) really plays no role in this problem (at least if one of the operators is self-adjoint), although it *is* needed in the compactness arguments used by the authors referred to above.

After the proof of the theorem we discuss some open problems. We wish to thank Arlen Brown, P. R. Halmos, and W. Kahan for helpful discussions of this circle of ideas.

We require three lemmas, the first of which is due to Schur (see VI, p. 13, in [11]). He considered infinite matrices (viewed as linear transformations on (l^2)) but his proof works for finite rectangular matrices as well.

LEMMA 1. *Let $T = (t_{ij})$ be given and let $\{f_i\}, \{g_j\} \subset L^2(I)$, where I is a subinterval (possibly infinite) of the real line. Assume that $\| f_i \| \leqslant \alpha, \| g_j \| \leqslant \beta$ for all i, j, and let $v_{ij} = (f_i, g_j)$. Then*

$$\| (v_{ij} t_{ij}) \| \leqslant \alpha\beta \| T \|.$$

The matrix with entries $(v_{ij} t_{ij})$ is sometimes called the *Schur product* of the matrices (v_{ij}) and (t_{ij}). Von Neumann suggested the name "Hadamard product" because of the analogous product of two power series, but the name "Schur product" seems much more appropriate since Hadamard never discussed this concept, whereas Schur's paper [11] contains a number of very useful theorems concerning it. The next lemma is an easy consequence of Lemma 1.

LEMMA 2. *Let $T = (t_{ij})$ $(1 \leqslant i \leqslant u, 1 \leqslant j \leqslant v)$ be a complex rectangular matrix. Let $a_1, \dots, a_u$ and $b_1, \dots, b_v$ be real numbers, with $a_i - b_j \geqslant d > 0$ for all i, j. Then*

$$\left\|\left(\frac{1}{a_i - b_j} t_{ij}\right)\right\| \leqslant \frac{1}{d} \| T \|.$$

Proof. Let c be the number midway between the smallest value of a_i and the largest value of b_j. Thus $a_i - c \geqslant d/2$ and $c - b_j \geqslant d/2$ for all i, j. In $L^2(0, \infty)$ let $f_i(x) = \exp(-(a_i - c)x)$, $g_j(x) = \exp(-(c - b_j)x)$. Then $\| f_i \|^2$, $\| g_j \|^2 \leqslant 1/d$, and $v_{ij} = (f_i, g_i) = 1/(a_i - b_j)$. The result now follows from Lemma 1. (Note: instead of using Lemma 1 the proof could also be based on a theorem of Rosenblum [8].)

The final lemma was conjectured (and established in a weaker form) by the authors and proved by S. Schanuel [10] by an elementary but very clever argument.

LEMMA 3. *Let $d_1, \dots, d_n$ be non-negative numbers. Then a subset $\{i(1), i(2), \dots, i(m)\}$ of the integers $\{1, \dots, n\}$ can be found such that $i(1) < \cdots < i(m)$, and*

$$\text{(i)} \qquad d_1 + \cdots + d_{i(1)-1} < 1, \qquad d_{i(1)+1} + \cdots + d_{i(2)-1} < 1, \dots;$$

$$\text{(ii)} \qquad \frac{1}{d_{i(1)}} + \frac{1}{d_{i(2)}} + \cdots + \frac{1}{d_{i(m)}} \leqslant n.$$

We allow the empty subset ($m = 0$). In this case (ii) is vacuously satisfied, and (i) will be satisfied provided $d_1 + \cdots + d_n < 1$. We also allow the full set ($m = n$). In this case (i) is vacuously satisfied. In general we regard the subset $\{d_{i(1)}, \dots, d_{i(m)}\}$ as being "removed" from the ordered set $d_1, \dots, d_n$, and then condition (i) requires that the sum in each "block" that remains must be less than one, while condition (ii) requires that the sum of the reciprocals of the removed elements must not exceed n.

Proof of Theorem 1. Since A is self-adjoint we may choose an orthonormal basis relative to which the matrix for A has diagonal form, with diagonal entries $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n$. Let (b_{ij}) denote the matrix for B with respect to this basis. We shall identify A and B with their matrix representations. A calculation shows that

$$AB - BA = ((\lambda_i - \lambda_j)b_{ij}) \quad (1 \leqslant i, j \leqslant n).$$

We divide the numbers $\lambda_1, ..., \lambda_n$ into disjoint consecutive blocks by a procedure to be described later. We require two properties of this division, the second of which will be given later (after (12)). The first property is that in each block the difference between the largest and smallest number shall be less than 2ϵ. We introduce some notation. Let s denote the number of blocks and let the integers $i(1) < i(2) < \cdots < i(s)$ denote the final indices in the blocks (thus $i(s) = n$). The blocks then are

$$1 \leqslant j \leqslant i(1), i(1) + 1 \leqslant j \leqslant i(2), ..., i(s-1) + 1 \leqslant j \leqslant i(s), \tag{6}$$

and our first condition is that

$$\lambda_1 - \lambda_{i(1)} < 2\epsilon, \lambda_{i(1)+1} - \lambda_{i(2)} < 2\epsilon, ... \tag{7}$$

In each block we take the average value of the largest and smallest number, and we form a new sequence $\{\mu_j\}$ from these averages:

$$\mu_1 = \cdots = \mu_{i(1)} = \frac{\lambda_1 + \lambda_{i(1)}}{2}, \qquad \mu_{i(1)+1} = \cdots = \mu_{i(2)} = \frac{\lambda_{i(1)+1} + \lambda_{i(2)}}{2}, \ldots .$$

Thus $|\mu_i - \lambda_i| < \epsilon$ for all i. Let A' denote the diagonal matrix formed from the $\{\mu_i\}$; then clearly $\| A - A' \| < \epsilon$.

Using these blocks the matrix B is partitioned into a block matrix (B_{ij}) $(1 \leqslant i, j \leqslant s)$. Let B' denote the diagonal block matrix formed by keeping only the diagonal terms $B_{11}, B_{22}, ...$ and replacing all the off-diagonal terms by 0. Note that if B is self-adjoint, then so is B'. We illustrate the case $s = 3$:

$$B = \begin{pmatrix} B_{11} & B_{12} & B_{13} \\ B_{21} & B_{22} & B_{23} \\ B_{31} & B_{32} & B_{33} \end{pmatrix}, \qquad B' = \begin{pmatrix} B_{11} & 0 & 0 \\ 0 & B_{22} & 0 \\ 0 & 0 & B_{33} \end{pmatrix}. \tag{8}$$

Then $A'B' = B'A'$, since the entries for A' are constant in each diagonal block. Thus to complete the proof we must show that the blocks can be chosen so that $\| B - B' \| \leqslant \epsilon$.

We may write $B - B' = B_1 + B_2 + \cdots + B_{s-1}$, where B_1 is formed from B by replacing B_{11} by 0, and by replacing B_{ij} by 0 whenever both $i > 1$ and $j > 1$. Likewise, B_2 is formed from B by replacing the first row and the first column of B by 0, by replacing B_{22} by 0, and by replacing B_{ij} by 0 whenever both $i > 2$ and $j > 2$. The matrices $B_3, ..., B_{s-1}$ are defined similarly. Thus

$$\| B - B' \| \leqslant \| B_1 \| + \cdots + \| B_{s-1} \|. \tag{9}$$

We illustrate the case $s = 3$ as in (8):

$$B_1 = \begin{pmatrix} 0 & B_{12} & B_{13} \\ B_{21} & 0 & 0 \\ B_{31} & 0 & 0 \end{pmatrix}, \qquad B_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & B_{23} \\ 0 & B_{32} & 0 \end{pmatrix}.$$

Note that each matrix B_k is the sum of two matrices, $B_k = B_k' + B_k''$, where B_k' contains the non-zero column of B_k, and B_k'' contains the non-zero row. (Thus B_1' is formed from B_1 by replacing $B_{12}, B_{13}, \ldots$ by 0, and similarly for B_1'', etc.). Also it is easy to see that

$$\| B_k \| = \max(\| B_k' \|, \| B_k'' \|). \tag{10}$$

Now let $AB - BA = C = (C_{ij})$ $(1 \leqslant i, j \leqslant s)$ where (C_{ij}) is the block form of C. For $1 \leqslant k \leqslant s - 1$, let C_k, C_k', C_k'' be formed from C in the same manner that the corresponding matrices were formed from B. It can be shown that $\| C_k'' \| \leqslant \| C \|$. (Consider a unit vector where C_k'' attains its norm.) Hence, by taking adjoints, one sees that $\| C_k' \| \leqslant \| C \|$. Thus by (2),

$$\| C_k' \| \leqslant \frac{2\epsilon^2}{n-1},$$

and similarly for C_k''. From (5) we see that the entries in C_1' are of the form

$$c_{ij} = (\lambda_i - \lambda_j) b_{ij}, \; i(1) + 1 \leqslant i \leqslant n, \; 1 \leqslant j \leqslant i(1).$$

Analogous statements can be made for C_1'' and for C_k', C_k'', $2 \leqslant k \leqslant s - 1$. Let

$$\alpha_1 = \lambda_{i(1)} - \lambda_{i(1)+1}, \; \alpha_2 = \lambda_{i(2)} - \lambda_{i(2)+1}, \ldots, \tag{11}$$

Hence $| \lambda_i - \lambda_j | \geqslant \alpha_1$ whenever $i(1) + 1 \leqslant i$ and $j \leqslant i(1)$.

Applying Lemma 2 with $T = C_1'$, $a_i = \lambda_i$ $(i(1) + 1 \leqslant i \leqslant n)$, $b_j = \lambda_j$ $(1 \leqslant j \leqslant i(1))$, we have

$$\| B_1' \| \leqslant \frac{1}{\alpha_1} \frac{2\epsilon^2}{n-1}.$$

We obtain the same estimate for B_1'', and hence, by (10), for B_1. Similarly, we have

$$\| B_k \| \leqslant \frac{1}{\alpha_k} \frac{2\epsilon^2}{n-1} \qquad (k \leqslant s - 1).$$

Thus by (9),

$$\| B - B' \| \leqslant \left(\frac{1}{\alpha_1} + \cdots + \frac{1}{\alpha_{s-1}} \right) \frac{2\epsilon^2}{n-1}. \tag{12}$$

To complete the proof we must arrange that the term in parentheses doesn't exceed $(n - 1)/2\epsilon$. This is the second property that we require when we divide $\{\lambda_i\}$ into blocks.

Note that in proving the theorem it would be sufficient to establish it for one fixed value ϵ_0 of ϵ; the general case would follow from this by multiplying (2) and (3) by suitable constants. Let us take $\epsilon_0 = \frac{1}{2}$. Then our two conditions on the division into blocks become: (i) the difference between the first and last λ in each block must be less than 1, (ii) $\sum \alpha_i^{-1} \leqslant n - 1$, where the numbers α_i are defined by (11), that is, α_i is the difference between the last element in the ith block and the first element in the $(i + 1)$-st block. It will be convenient to reformulate these conditions by introducing the differences

$$d_1 = \lambda_1 - \lambda_2 \,, \; d_2 = \lambda_2 - \lambda_3 \,, \ldots, \; d_{n-1} = \lambda_{n-1} - \lambda_n \,.$$

These are non-negative numbers, and our two conditions now become

$$\text{(i)} \qquad d_1 + \cdots + d_{i(1)-1} < 1, \qquad d_{i(1)+1} + \cdots + d_{i(2)-1} < 1, \ldots;$$

$$\text{(ii)} \qquad \frac{1}{d_{i(1)}} + \frac{1}{d_{i(2)}} + \cdots + \frac{1}{d_{i(s-1)}} \leqslant n - 1$$

by virtue of (6), (7), and (11). It follows from Lemma 3 that the indices $i(1), \ldots, i(s - 1)$ can be chosen so as to satisfy these two conditions. This completes the proof of the theorem.

COROLLARY 1. *Let $\mathscr{H}$ be an n-dimensional complex Hilbert space and let T be an operator on $\mathscr{H}$. If*

$$\| T^*T - TT^* \| \leqslant \frac{\epsilon^2}{n - 1} \,,$$

then there is a normal operator N on $\mathscr{H}$ with $\| T - N \| < \epsilon$.

Proof. Let $T = A + iB$, with A, B self-adjoint. A calculation shows that $T^*T - TT^* = 2i(AB - BA)$. Hence equation (2) is satisfied with ϵ replaced by $\epsilon/2$. Thus there are commuting self-adjoint operators A', B' that are within distance $\epsilon/2$ of A, B respectively. *Let* $N = A' + iB'$. Then N is normal and $\| T - N \| < \epsilon$.

There are several open problems that we would like to mention. First, does the theorem hold without the assumption that A is self-adjoint? Two results in this direction were obtained by Bernstein (see [2; 3, Theorem 2.2]) who showed that if A, B are $n \times n$ matrices that "almost commute" then they have a common "almost eigenvector," and there is an orthonormal basis with respect to which they both "almost have" upper triangular form. He does not require self-adjointness, but his hypotheses are very restrictive in their dependence on n.

The second problem concerns the dependence on n. It is still an open question whether, for each $\epsilon > 0$, there exists $\delta = \delta(\epsilon) > 0$ such that for all n, if A, B are $n \times n$ complex matrices with $\| AB - BA \| \leqslant \delta$, then there exist commuting $n \times n$ matrices A', B' with $\| A - A' \| \leqslant \epsilon$, $\| B - B' \| \leqslant \epsilon$. One could ask the analogous question about the corollary. If the answer to this question is affirmative, then one could pass to the limit and have an analogous result for compact operators on infinite-dimensional Hilbert space.

It is known that such a result is not true for non-compact operators. The following example was shown to us by Halmos; it is a slight modification of one given by Bastian and Harrison in [1].

Let $\{e_i\}_{i=1}^{\infty}$ be an orthonormal basis for a separable, infinite-dimensional Hilbert space, and let T_n be the weighted shift operator defined by $T_n e_i = (i/n)^{1/2} e_{i+1}$ $(1 \leqslant i \leqslant n)$, $T_n e_i = e_{i+1}$ $(i > n)$. A calculation shows that $\| T_n^* T_n - T_n T_n^* \| = 1/n$. On the other hand, T_n differs from the unweighted unilateral shift by a compact operator, and hence, by a lemma of Halmos (see [1, Lemma 2]), $\| T_n - N \| \geqslant 1$ for every normal operator N. For some other results in the infinite-dimensional case, see [5, 7].

The last problem we shall mention is the following. (It is probably easier than the others.) As noted at the beginning of the paper, if we limit our attention to operators of norm at most one, then for each $\epsilon > 0$ there exists $\delta = \delta(\epsilon, n)$ such that if $\| AB - BA \| \leqslant \delta$, then there exist commuting operators A', B' for which $\| A - A' \| \leqslant \epsilon$, $\| B - B' \| \leqslant \epsilon$. We take $\tilde{\delta}(\epsilon, n)$ to be the maximum admissable value of such δ. Is it true that $\tilde{\delta}(\epsilon, n) \geqslant \tilde{\delta}(\epsilon, n+1)$ for all ϵ and n?

References

1. J. J. Bastian and K. J. Harrison, Subnormal weighted shifts and properties of normal operators, *Proc. Amer. Math. Soc.* **42** (1974), 475–479.
2. A. R. Bernstein, Almost eigenvectors for almost commuting matrices, *SIAM J. Appl. Math.* **21** (1971), 232–235.
3. A. R. Bernstein, Invariant subspaces for certain commuting operators on Hilbert space, *Ann. of Math.* **95** (1972), 253–260.
4. J. A. Deddens, Normal plus compact, unpublished preprint circulated 1972.
5. B. E. Johnson and J. P. Williams, The range of a normal derivation, *Pacific J. Math.* **58** (1975), 105–122.
6. W. A. J. Luxembourg and R. F. Taylor, Almost commuting matrices are near commuting matrices, *Indag. Math.* **32** (1970), 96–98.
7. R. Moore, An asymptotic Fuglede theorem, *Proc. Amer. Math. Soc.* **50** (1975), 138–142.
8. M. Rosenblum, On the operator equation $BX - XA = Q$, *Duke Math. J.* **23** (1956), 263–269.
9. P. Rosenthal, Are almost commuting matrices near commuting matrices?, *Amer. Math. Monthly* **76** (1969), 925–926.
10. S. H. Schanuel, A combinatorial problem of Pearcy and Shields, *Proc. Amer. Math. Soc.* **65** (1977), 185–186.
11. J. Schur, Bemerkungen zur Theorie der beschränkten Bilinearformen mit unendlich vielen Veränderlichen, *J. Reine Angew. Math.* **140** (1911), 1–28.