PACIFIC JOURNAL OF MATHEMATICS
Vol. 102, No. 2, 1982

THE OPERATOR EQUATION $AX - XB = C$ WITH NORMAL A AND B

ALLEN SCHWEINSBERG

A necessary and sufficient condition for the solvability of $AX-XB=C$ was given by W. E. Roth for finite matrices and by M. Rosenblum for selfadjoint operators A and B on a Hilbert space. Here the result is extended to include normal operators and finite rank operators on Hilbert space.

1. Introduction. In [6] W. E. Roth proved for finite matrices over a field that $AX - XB = C$ is solvable for X if and only if the matrices $\begin{bmatrix} A & C \\ 0 & B \end{bmatrix}$ and $\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$ are similar. A considerably briefer proof has been given by Flanders and Wimmer [4]. In [5] Rosenblum showed that the result remains true when A and B are bounded selfadjoint operators on a complex, separable Hilbert space. In the present paper the theorem is extended to include finite rank operators and normal operators on Hilbert space. We give an example to show that normality cannot be weakened to quasinormality. Finally, when $A = B$ the following is true, even in the absence of normality: if $\begin{bmatrix} A & C \\ 0 & A \end{bmatrix}$ and $\begin{bmatrix} A & 0 \\ 0 & A \end{bmatrix}$ are similar, then C is a commutator.

2. The normal case. We begin with a lemma.

LEMMA 1. *If $\begin{bmatrix} Q & R \\ S & T \end{bmatrix}$ is an invertible operator acting in the usual way on the direct sum of Hilbert spaces $\mathscr{H}_1 \oplus \mathscr{H}_2$, then $SS^* + TT^*$ is invertible on $\mathscr{H}_2$.*

Proof. Let $W = \begin{bmatrix} Q & R \\ S & T \end{bmatrix}\begin{bmatrix} Q & R \\ S & T \end{bmatrix}^*$. Now $SS^* + TT^*$ is bounded below. For if $\|f_n\| = 1$ and $\lim \|(SS^* + TT^*)f_n\| = 0$, we would have

$$\begin{aligned} \lim \| W^{1/2}(0 \oplus f_n)\|^2 &= \lim (W(0 \oplus f_n),\ (0 \oplus f_n)) \\ &= \lim ((SS^* + TT^*)f_n, f_n) = 0\ , \end{aligned}$$

contradicting the invertibility of $W^{1/2}$. Since $SS^* + TT^*$ is Hermitian and bounded below, it is invertible.

This lemma is also a direct consequence of [3, Corollary 1].

THEOREM 1. *Let A and B be bounded normal operators on* *complex Hilbert spaces* $\mathscr{H}_1$ *and* $\mathscr{H}_2$ *respectively. Then* $AX - XB = C$ *has a solution* X *if and only if* $\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$ *and* $\begin{bmatrix} A & C \\ 0 & B \end{bmatrix}$ *are similar operators on* $\mathscr{H}_1 \oplus \mathscr{H}_2$.

Proof. Since $\begin{bmatrix} I & -X \\ 0 & I \end{bmatrix}\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}\begin{bmatrix} I & X \\ 0 & I \end{bmatrix} = \begin{bmatrix} A & AX - XB \\ 0 & B \end{bmatrix}$, half of the theorem is immediate.

Conversely, assume similarity. Then $\begin{bmatrix} Q & R \\ S & T \end{bmatrix}\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} = \begin{bmatrix} A & C \\ 0 & B \end{bmatrix}\begin{bmatrix} Q & R \\ S & T \end{bmatrix}$ where $\begin{bmatrix} Q & R \\ S & T \end{bmatrix}$ is invertible. Consequently $QA - AQ = CS$, $RB - AR = CT$, $SA = BS$, and $TB = BT$. The Putnam-Fuglede theorem implies $AS^* = S^*B$ and $T^*B = BT^*$. Also $BSS^* = SAS^* = SS^*B$, so B commutes with both SS^* and TT^*. Now

$$\begin{aligned} C(SS^* + TT^*) &= (QA - AQ)S^* + (RB - AR)T^* \\ &= (QAS^* + RBT^*) - (AQS^* + ART^*) \\ &= (QS^* + RT^*)B - A(QS^* + RT^*)\,. \end{aligned}$$

The lemma shows that $SS^* + TT^*$ is invertible. Since the inverse commutes with B, we arrive at $C = AX - XB$ for $X = -(QS^* + RT^*)(SS^* + TT^*)^{-1}$, and the proof is complete.

This theorem does not hold for all operators. The following example, similar to one given in [5], shows that the normality hypothesis cannot even be weakened to quasinormality. Let $A = U$, the unilateral shift, $B = 0$, and let $P = I - UU^*$. Then we have the similarity

$$\begin{bmatrix} U^* & -I \\ P & U \end{bmatrix}\begin{bmatrix} U & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} U & I \\ 0 & U^* \end{bmatrix} = \begin{bmatrix} U & I \\ 0 & 0 \end{bmatrix},$$

but clearly for no X does $I = UX - X0$.

Not surprisingly, the order of the diagonal entries is critical. The theorem holds for $A = 0$ and $B = U$, because

$$C = 0(-CU^*) - (-CU^*)U \text{ for every } C\,.$$

Later we will see that the theorem is true for $A = B = U$, as well.

In spite of the example given above, the normality hypothesis can be weakened somewhat. For example A and B need only be similar to normal operators, as can be seen from the next theorem.

THEOREM 2. *Let* $\mathscr{C}$ *be the collection of pairs of operators* (A, B) *for which* $AX - XB = C$ *has a solution* X *if and only if* $\begin{bmatrix} A & C \\ 0 & B \end{bmatrix}$ *and* $\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$ *are similar. Suppose* $(A, B) \in \mathscr{C}$. *Then*

(i) $(A_1, B_1) \in \mathscr{C}$ *if* A_1 *and* B_1 *are similar to* A *and* B *respectively.*

(ii) $(B^*, A^*) \in \mathscr{C}$.

(iii) $(A^{-1}, B^{-1}) \in \mathscr{C}$ *if* A *and* B *are invertible.*

(iv) $(A + \lambda I, B + \lambda I) \in \mathscr{C}$ *for all complex* λ.

Proof. (i) If $S^{-1}A_1S = A$, $T^{-1}B_1T = B$, and

$$R^{-1}\begin{bmatrix} A_1 & 0 \\ 0 & B_1 \end{bmatrix} R = \begin{bmatrix} A_1 & C \\ 0 & B_1 \end{bmatrix},$$

then

$$R^{-1}\begin{bmatrix} S & 0 \\ 0 & T \end{bmatrix}\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}\begin{bmatrix} S^{-1} & 0 \\ 0 & T^{-1} \end{bmatrix} R = \begin{bmatrix} S & 0 \\ 0 & T \end{bmatrix}\begin{bmatrix} A & S^{-1}CT \\ 0 & B \end{bmatrix}\begin{bmatrix} S^{-1} & 0 \\ 0 & T^{-1} \end{bmatrix}.$$

Since $(A, B) \in \mathscr{C}$ we get $S^{-1}CT = AX - XB$ for some X. Then $C = SAXT^{-1} - SXBT^{-1} = A_1(SXT^{-1}) - (SXT^{-1})B_1$.

(ii) If $\begin{bmatrix} B^* & C \\ 0 & A^* \end{bmatrix}$ is similar to $\begin{bmatrix} B^* & 0 \\ 0 & A^* \end{bmatrix}$, then $\begin{bmatrix} A & C^* \\ 0 & B \end{bmatrix}$ is similar to $\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$. Thus, $C^* = AX - XB$, i.e., $C = B^*(-X^*) - (-X^*)A^*$.

(iii) If $\begin{bmatrix} A^{-1} & C \\ 0 & B^{-1} \end{bmatrix}$ and $\begin{bmatrix} A^{-1} & 0 \\ 0 & B^{-1} \end{bmatrix}$ are similar, then their inverses, $\begin{bmatrix} A & -ACB \\ 0 & B \end{bmatrix}$ and $\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$ are similar. So $-ACB = AX - XB$, or $C = A^{-1}X - XB^{-1}$.

(iv) This is clear, since the relevant equations remain valid if A and B are replaced by $A - \lambda I$ and $B - \lambda I$.

3. The finite rank case. First we observe a lemma which in general would be false without the assumption of finite dimensionality.

LEMMA 2. *If* $\begin{bmatrix} A_1 & 0 \\ 0 & 0 \end{bmatrix}$ *and* $\begin{bmatrix} B_1 & 0 \\ 0 & 0 \end{bmatrix}$ *are similar operators on* $\mathscr{H}_1 \oplus \mathscr{H}_2$ *and* $\mathscr{H}_1$ *has finite dimension, then* A_1 *and* B_1 *are similar.*

Proof. Let $A = \begin{bmatrix} A_1 & 0 \\ 0 & 0 \end{bmatrix}$, $B = \begin{bmatrix} B_1 & 0 \\ 0 & 0 \end{bmatrix}$, and $k = \dim \mathscr{H}_1$. Denote nullity and rank by ν and ρ respectively. If $\lambda \neq 0$ then $\nu(A_1 - \lambda I)^n = \nu(A - \lambda I)^n = \nu(B - \lambda I)^n = \nu(B_1 - \lambda I)^n$ for all natural numbers n. Also

$$\nu(A_1^n) = k - \rho(A_1^n) = k - \rho(A^n) = k - \rho(B^n) = k - \rho(B_1^n) = \nu(B_1^n).$$

Since $\nu(A_1 - \lambda I)^n = \nu(B_1 - \lambda I)^n$ for all λ and n, A_1 and B_1 are similar.

THEOREM 3. *Let* A *and* B *be finite rank operators on complex Hilbert spaces. Then* $AX - XB = C$ *has a solution* X *if and only* *if* $\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$ *is similar to* $\begin{bmatrix} A & C \\ 0 & B \end{bmatrix}$.

Proof. Observe that each finite rank operator A on a Hilbert space $\mathscr{H}$ is unitarily equivalent to one of the form $\begin{bmatrix} A_1 & 0 \\ 0 & 0 \end{bmatrix}$ where A_1 operates on a finite dimensional space. Simply write $\mathscr{H}$ as $\mathscr{H}_1 \oplus \mathscr{H}_2$ where $\mathscr{H}_1$ is finite dimensional and contains the ranges of A and A^*. In light of Theorem (2i) we may assume that A and B are already in this form. Then the assumption that $\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$ is similar to $\begin{bmatrix} A & C \\ 0 & B \end{bmatrix}$ becomes

$$\begin{bmatrix} A_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & B_1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \text{ is similar to } \begin{bmatrix} A_1 & 0 & C_1 & C_2 \\ 0 & 0 & C_3 & C_4 \\ 0 & 0 & B_1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Equality of ranks requires that $C_4 = 0$, $C_2 = A_1X_2$ and $C_3 = -X_3B_1$ for some X_2 and X_3. Next observe the similarity

$$\begin{bmatrix} I & 0 & 0 & X_2 \\ 0 & I & X_3 & 0 \\ 0 & 0 & I & 0 \\ 0 & 0 & 0 & I \end{bmatrix} \begin{bmatrix} A_1 & 0 & C_1 & A_1X_2 \\ 0 & 0 & -X_3B_1 & 0 \\ 0 & 0 & B_1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} I & 0 & 0 & -X_2 \\ 0 & I & -X_3 & 0 \\ 0 & 0 & I & 0 \\ 0 & 0 & 0 & I \end{bmatrix} = \begin{bmatrix} A_1 & 0 & C_1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & B_1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

After row and column interchanges we obtain the similarity of

$$\begin{bmatrix} A_1 & 0 & 0 & 0 \\ 0 & B_1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \text{ and } \begin{bmatrix} A_1 & C_1 & 0 & 0 \\ 0 & B_1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Lemma 2 implies the similarity of $\begin{bmatrix} A_1 & 0 \\ 0 & B_1 \end{bmatrix}$ and $\begin{bmatrix} A_1 & C_1 \\ 0 & B_1 \end{bmatrix}$, and Roth's original theorem for finite matrices implies $C_1 = A_1X_1 - X_1B_1$ for some X_1. Finally

$$\begin{bmatrix} C_1 & C_2 \\ C_3 & C_4 \end{bmatrix} = \begin{bmatrix} A_1X_1 - X_1B_1 & A_1X_2 \\ -X_3B_1 & 0 \end{bmatrix} = \begin{bmatrix} A_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} X_1 & X_2 \\ X_3 & 0 \end{bmatrix} - \begin{bmatrix} X_1 & X_2 \\ X_3 & 0 \end{bmatrix} \begin{bmatrix} B_1 & 0 \\ 0 & 0 \end{bmatrix},$$

i.e., $C = AX - XB$.

3. The case $A = B$. If we set $A = B$ in Theorem 1 then it becomes a statement about commutators in the range of the derivation $\Delta_A(X) = AX - XA$.

THEOREM 4. *Let A be a bounded normal operator on a complex Hilbert space. Then* $\begin{bmatrix} A & 0 \\ 0 & A \end{bmatrix}$ *and* $\begin{bmatrix} A & C \\ 0 & A \end{bmatrix}$ *are similar if and only if C is in the range of the derivation Δ_A.*

Once again normality cannot be weakened even to quasinormality, because the theorem fails for $A = \begin{bmatrix} U & 0 \\ 0 & 0 \end{bmatrix}$ where U is the unilateral shift. Note that, as in the earlier example,

$$\begin{bmatrix} U & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & U & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \text{ is similar to } \begin{bmatrix} U & I & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & U & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

which, via a row and column interchange, is similar to

$$\begin{bmatrix} U & 0 & 0 & I \\ 0 & 0 & 0 & 0 \\ 0 & 0 & U & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

But clearly $\begin{bmatrix} 0 & I \\ 0 & 0 \end{bmatrix}$ is not in the range of Δ_A.

If $A = B = U$ then all is well.

THEOREM 5. *Let U be the unilateral shift. Then $UX - XU = C$ if and only if* $\begin{bmatrix} U & C \\ 0 & U \end{bmatrix}$ *is similar to* $\begin{bmatrix} U & 0 \\ 0 & U \end{bmatrix}$.

Proof. As usual, one direction is immediate. Suppose though that $\begin{bmatrix} U & C \\ 0 & U \end{bmatrix}$ is similar to $\begin{bmatrix} U & 0 \\ 0 & U \end{bmatrix}$. Then, with procedure and notation as in the proof of Theorem 1, U commutes with S and T, operator entries in the invertible matrix $\begin{bmatrix} Q & R \\ S & T \end{bmatrix}$. Invertibility implies the existence of operators Y and Z for which $SY + TZ = I$. Let $f \in \mathscr{H}$. Then $f = Y^*S^*f + Z^*T^*f$. So $\|f\| \leqq \|Y^*\| \|S^*f\| + \|Z^*\| \|T^*f\| \leqq (\|Y^*\|^2 + \|Z^*\|^2)^{1/2}(\|S^*f\|^2 + \|T^*f\|^2)^{1/2}$. Since S and T commute with the shift, they must be analytic Toeplitz operators. The inequality $\|S^*f\|^2 + \|T^*f\|^2 \geqq \varepsilon^2 \|f\|^2$ for $\varepsilon > 0$ implies the existence of additional analytic Toeplitz operators W and X for which $SW + TX = I$ ([1]; Theorem 6.3). So, again as in the proof of Theorem 1, $C = C(SW + TX) = (QU - UQ)W + (RU - UR)X = (QW + RX)U - U(QW + RX)$, and the proof is complete.

In general, while similarity of $\begin{bmatrix} A & C \\ 0 & A \end{bmatrix}$ and $\begin{bmatrix} A & 0 \\ 0 & A \end{bmatrix}$ need not imply that C is a commutator in the range of Δ_A, it is true that C must be some commutator.

THEOREM 6. *Let A be a bounded operator on a complex Hilbert space $\mathscr{H}$. If $\begin{bmatrix} A & C \\ 0 & A \end{bmatrix}$ is similar to $\begin{bmatrix} A & 0 \\ 0 & A \end{bmatrix}$, then C is a commutator.*

Proof. If $\mathscr{H}$ is finite dimensional then of course Roth's theorem implies the present one. Suppose $\dim \mathscr{H} = \aleph_0$. Recall that in [2] Brown and Pearcy characterized commutators on infinite dimensional, separable Hilbert spaces as all those operators not of the form $\lambda I + K$ where λ is a nonzero scalar and K is compact. Thus it suffices to show that $\begin{bmatrix} A & \lambda I + K \\ 0 & A \end{bmatrix}$ cannot be similar to $\begin{bmatrix} A & 0 \\ 0 & A \end{bmatrix}$. Suppose on the contrary that

$$S^{-1}\begin{bmatrix} A & 0 \\ 0 & A \end{bmatrix}S = \begin{bmatrix} A & \lambda I + K \\ 0 & A \end{bmatrix}.$$

Then $S^{-1}\begin{bmatrix} A^n & 0 \\ 0 & A^n \end{bmatrix}S = \begin{bmatrix} A^n & \lambda n A^{n-1} + K_n \\ 0 & A^n \end{bmatrix}$ where K_n is compact, for $n = 0, 1, 2, \cdots$. In fact if f is any analytic function, then

$$S^{-1}\begin{bmatrix} f(A) & 0 \\ 0 & f(A) \end{bmatrix}S = \begin{bmatrix} f(A) & \lambda f'(A) + \hat{K} \\ 0 & f(A) \end{bmatrix}$$

where $\hat{K}$ is compact. Now, letting $\| \ \|_1$ be the Calkin algebra norm, we see that

$$\|\lambda f'(A)\|_1 = \|\lambda f'(A) + \hat{K}\|_1 \leqq \left\| \begin{bmatrix} f(A) & \lambda f'(A) + \hat{K} \\ 0 & f(A) \end{bmatrix} \right\|_1$$
$$\leqq \|S^{-1}\|_1 \|S\|_1 \|f(A)\|_1 .$$

In particular let $f(z) = e^{\alpha z}$. Then $\|\alpha\lambda e^{\alpha A}\|_1 \leqq \|S^{-1}\|_1 \|S\|_1 \|e^{\alpha A}\|_1$. For large $|\alpha|$ this implies $\|e^{\alpha A}\|_1 = 0$, i.e., $e^{\alpha A}$ is compact. But $e^{\alpha A}$ is invertible and cannot be compact.

Finally suppose $\dim \mathscr{H} > \aleph_0$. Then again the noncommutators are those of the form $\lambda I + K$ where $\lambda \neq 0$ and K belongs to a certain proper closed ideal [2; Theorem 4]. The preceding argument with obvious modifications handles this case as well, and the proof is complete.

ACKNOWLEDGMENT. The author is grateful to P. Rosenthal for helpful conversations while the latter was a Distinguished Visiting Professor at Bucknell University.

References

1. W. B. Arveson, *Interpolation problems in nest algebras*, J. Functional Anal., **20** (1975), 208-233.
2. A. Brown and C. Pearcy, *Structure of commutators of operators*, Ann. Math., **82** (1965), 112-127.
3. P. A. Fillmore and J. P. Williams, *On operator ranges*, Adv. in Math., **7** (1971), 254-281.
4. H. Flanders and H. K. Wimmer, *On the matrix equations* $AX-XB=C$ *and* $AX-YB=C$, SIAM J. Appl. Math., **32** (1977), 707-710.
5. M. Rosenblum, *The operator equation* $BX-XA=Q$ *with selfadjoint* A *and* B, Proc. Amer. Math. Soc., **20** (1969), 115-120.
6. W. E. Roth, *The equations* $AX-YB=C$ *and* $AX-XB=C$ *in matrices*, Proc. Amer. Math. Soc., **3** (1952), 392-396.

Received August 28, 1979.

BUCKNELL UNIVERSITY
LEWISBURG, PA 17837