Canad. Math. Bull. Vol. **18** (3), 1975

EP OPERATORS AND GENERALIZED INVERSES

BY

STEPHEN L. CAMPBELL AND CARL D. MEYER

ABSTRACT: The relationship between properties of the generalized inverse of A, $A^\dagger$, and of the adjoint of A, A^*, are studied. The property that $A^\dagger A$ and $AA^\dagger$ commute, called (E4), is investigated. (E4) generalizes the property of A being EP_r. A canonical form and a formula for $A^\dagger$ are given if a matrix A is (E4). Results are in a Hilbert space setting whenever possible. Examples are given.

1. A bounded linear operator A on a complex Hilbert space $\mathscr{h}$ is called an EP operator if its range, $R(A)$, is closed and $R(A)=R(A^*)$. This concept was introduced for matrices by Schwerdtfeger in [14] and has been studied in detail by several authors [1], [9], [10], [11], [12], *et al.* For bounded linear operators A with closed range, the generalized inverse of A, $A^\dagger$, is defined to be the bounded operator $A_1^{-1}P$ where A_1 is the restriction of A to $R(A^*)$ and P is the orthogonal projection onto $R(A)$. Some equivalent definitions and properties of $A^\dagger$ are given in [2], [5], and [13].

There are some interesting relationships between normal operators, EP operators, and generalized inverses. One phenomenon which frequently occurs is that if one obtains a statement which characterizes a normal operator and replaces the adjoint operation (*) by the generalized inverse operation (†), then the resulting statement, which shall be referred to as the dual statement, is a characterization for EP operators. For example, consider the statement "A is a normal operator if and only if $A^*A=AA^*$." It is easy to verify that the dual statement "A is an EP operator if and only if $A^\dagger A=AA^\dagger$" is valid. Other relations of this type may be found in [10] and [12] where finite complex matrices were considered.

Our purpose in this paper is to explore further relationships between EP operators, generalized inverses, normal operators, and binormal operators. In particular, we completely characterize operators for which $A^\dagger A$ and $AA^\dagger$ commute by giving a block decomposition.

2. We begin by establishing some of the properties of (*) that we wish to exploit. All operators are bounded linear operators with closed ranges. Except possibly for those occurring as blocks in block matrices, they map a given Hilbert space into itself. $B(\mathscr{h})$ is the set of all bounded linear operators on $\mathscr{h}$ with closed range. If $X, Y\in B(\mathscr{h})$, then $[X, Y]=XY-YX$.

Received by the editors September 28, 1973 and, in revised form, February 6, 1974.

AMS 1970 Subject Classifications: Primary 15A09, Secondary; 15A21, 47A65.

Key Words and Phrases: Generalized inverse, EP operators, EP_r matrices, canonical form.

https://doi.org/10.4153/CMB-1975-061-4 Published online by Cambridge University Press

THEOREM 1. *Suppose that $A \in B(\mathscr{h})$. Then each of the following implies the next.*

(N1) A is normal.

(N2) A^2 is normal.

(N3) $[A^*A, AA^*]=0$.

Proof. (N1)$\Rightarrow$(N2) is clear. To show that (N2)$\Rightarrow$(N3), assume that A^2 is normal. Fuglede's Theorem asserts that if B is normal and $[C, B]=0$, then $[C^*, B]=0$. (See [8, pp. 88–89] for a discussion of Fuglede's Theorem). Thus $[A^*, A^2]=0$ and $[A, A^{*2}]=0$ since $[A, A^2]=0$ and $[A^*, A^{*2}]=0$. (N3) now follows since

$$(AA^*)(A^*A) = AA^{*2}A = A^{*2}A^2 = A^*A^2A^* = (A^*A)(AA^*).$$

Operators satisfying (N3) were studied in [4] where they were referred to as binormal operators. (N3) also appears in the work of Embry [7]. Examples exist to show that none of the implications in Theorem 1 reverse. Condition (N2) may be replaced by

(N4) $[A, A^*A]=0$.

Operators satisfying (N4) have been studied by Brown [3]. Implicitly contained in [3] is the fact that if the underlying Hilbert space is finite dimensional, then (N4) is equivalent to (N1).

3. The next result is the (†) analogue of Theorem 1. It is interesting that some of the same examples used to show the implications are irreversible will work for Theorem 2 as for Theorem 1.

THFOREM 2. *Suppose that $A \in B(\mathscr{h})$. Then each of the following implies the next.*

(N1) *A is normal.*

(E1) *A is* EP.

(E2) *A^2 is* EP.

(E3) $(A^\dagger)^2=(A^2)^\dagger$.

(E4) $[A^\dagger A, AA^\dagger]=0$.

Furthermore all the implications are proper.

Proof. (N1)$\Rightarrow$(E1) is well known. That (E1)$\Rightarrow$(E2) is also clear. If $B, C \in B(\mathscr{h})$ and $BCB=B$, then C is called a (1)-inverse for B. If $BCB=B$ and $CBC=C$, then C is called a $(1, 2)$-inverse for B. Clearly $B^\dagger$ is a $(1, 2)$-inverse for B. We will show later in Theorem 3 that (E4) is equivalent to $A^{\dagger 2}$ being a (1) or $(1, 2)$-inverse for A^2. Then (E3) will imply (E4). We will now show (E2)$\Rightarrow$(E3). Suppose that A^2 is EP. Let $N(A^2)$ denote the null space of A^2. Then relative to the orthogonal decomposition $\mathscr{h}=R(A^2)\oplus N(A^2)$, we have $A^2=\begin{bmatrix} T & 0 \\ 0 & 0 \end{bmatrix}$ where T is invertible. Let $A=\begin{bmatrix} B & C \\ D & E \end{bmatrix}$. Then $[A, A^2]=0$ and the invertibility of T imply that $C=0$ and $D=0$. Hence $E^2=0$. But $A^\dagger=\begin{bmatrix} B^{-1} & 0 \\ 0 & E^\dagger \end{bmatrix}$ and $E^{\dagger 2}=0$ so that $A^{\dagger 2}=A^{2\dagger}$.

https://doi.org/10.4153/CMB-1975-061-4 Published online by Cambridge University Press

The same proof shows that if A^n is EP, then $A^{\dagger n}=A^{n\dagger}$.

EXAMPLE 1. Let $A=\begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}$. Then A^2 is EP while A is not. Thus (E2)$\not\Rightarrow$(E1).

EXAMPLE 2. Let

$$A=\begin{bmatrix}0 & 1 & 0\\ 0 & 0 & 1\\ 0 & 0 & 0\end{bmatrix}.$$

Then $A^\dagger=A^*$ and $A^{2\dagger}=A^{2*}$. Hence $A^{2\dagger}=A^{2*}=A^{*2}=A^{2\dagger}$. But $R(A^2)$ and $R(A^{*2})$ are unequal and hence A^2 is not EP. Thus (E3)$\not\Rightarrow$(E2).

EXAMPLE 3. Let

$$A=\begin{bmatrix}\begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix} & 0 & \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\\ \begin{bmatrix}0 & 0\\ 1 & 0\end{bmatrix} & 0 & \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\\ 0 & 0 & 0\end{bmatrix} \quad \text{where} \quad 0=\begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix}.$$

Then

$$A^\dagger=\begin{bmatrix}\begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix} & \begin{bmatrix}0 & 1\\ -1 & 0\end{bmatrix} & 0\\ 0 & 0 & 0\\ \begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix} & \begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix} & 0\end{bmatrix}.$$

A direct calculation shows that $A^2A^{\dagger 2}$ is not self-adjoint. Hence $A^{\dagger 2}\neq A^{2\dagger}$. However, it can be verified that $[AA^\dagger, A^\dagger A]=0$. Thus (E4)$\not\Rightarrow$(E3).

Before proceeding we need the following well-known lemma [15, p. 58].

LEMMA 1. *If $P\in B(\mathscr{h})$, $P^2=P$, and $\|P\|\leq 1$, then P is a self-adjoint projection.*

One use of condition (E4) is the following.

THEOREM 3. *Let $A\in B(\mathscr{h})$. Then $A^{\dagger 2}$ is a* (1)*-inverse for A^2 if and only if $[A^\dagger A, AA^\dagger]=0$.*

Proof. Suppose $A^{\dagger 2}$ is a (1)-inverse for A^2. Then $A^2A^{\dagger 2}A^2=A^2$. Multiplying on the right and left by $A^\dagger$ gives $(A^\dagger A^2A^\dagger)(A^\dagger A^2A^\dagger)=A^\dagger A^2A^\dagger$. But $\|A^\dagger A^2A^\dagger\|\leq \|A^\dagger A\|\,\|AA^\dagger\|=1$. Thus $(A^\dagger A)(AA^\dagger)=A^\dagger A^2A^\dagger$ is self-adjoint by Lemma 1. Since $A^\dagger A$ and $AA^\dagger$ are self-adjoint, and their product is self-adjoint, we have $[A^\dagger A, AA^\dagger]=0$ as desired. Now suppose $A^\dagger AAA^\dagger=AA^\dagger A^\dagger A$. Multiplying on the right and left by A yields the desired result.

Note that if $A^{\dagger 2}$ is a (1)-inverse for A^2, then it is a (1, 2)-inverse. Thus if (E4) is known to hold one may calculate (1, 2)-inverses for A^2 from $A^\dagger$ with little additional work.

4. EP operators have the advantage of a simple canonical form, $\begin{bmatrix}T & 0\\ 0 & 0\end{bmatrix}$, and an easily computed (†), $\begin{bmatrix}T^{-1} & 0\\ 0 & 0\end{bmatrix}$. Unfortunately, even in the finite dimensional

https://doi.org/10.4153/CMB-1975-061-4 Published online by Cambridge University Press

case, EP matrices form a very restrictive class. Not all matrices are even similar to an EP matrix.

As we will see shortly, operators satisfying (E4) also have a nice canonical form and a (†) that is fairly easy to compute. However, they admit a much greater degree of variety, for every matrix is similar to a matrix satisfying (E4). This can be easily seen by observing that every block in the Jordan form is (E4). In fact, the Jordan form can be written $\begin{bmatrix} J_1 & 0 \\ 0 & J_2 \end{bmatrix}$ where J_1 is invertible, if it is present (hence (E1)), and J_2 is (E3).

5. A condition implying (E4) will now be established.

THEOREM 4. *Let* $A \in B(\mathscr{h})$. *If* $[A^*A, AA^*]=0$, *then* $[A^\dagger A, AA^\dagger]=0$.

Proof. Let $\sigma(A)$ be the spectrum of A. It is known that $\sigma(A^*A)$ and $\sigma(AA^*)$ have the same nonzero elements. Let f be the function defined on the real line by $f(0)=0$ and $f(\lambda)=1$ if $\lambda\neq 0$. Using the spectral theorem for self-adjoint operators as found in [6, Chapter X], we can define $f(C)$ for any self-adjoint operator C. Then $f(C)$ is a self-adjoint projection onto the closure of $R(C)$. Suppose that $[A^*A, AA^*]=0$. Then $[f(A^*A), f(AA^*)]=0$. But $R(A^*A)=R(A^*)$ and $R(AA^*)=R(A)$. Hence $A^\dagger A=f(A^*A)$, $AA^\dagger=f(AA^*)$, and the result follows.

Notice that $A=\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ satisfies (E1), and hence (E4), but not (N3) so that the converse of Theorem 4 is not valid. In general, the (E) conditions are geometric statements about ranges, whereas the (N) conditions convey information on the entries of a matrix of A, that is, quantitative information. Thus one would not expect conditions of the (E) type to imply any of the (N) type without additional quantitative assumptions.

The next section is concerned primarily with the matrix case and will contain applications of some of our earlier ideas.

6. To determine whether A is EP it is necessary in principle, to determine $R(A)$ and $R(A^*)$ or calculate $A^\dagger$. In applications using matrices this can be a time consuming process if the matrices are large. To find $A^\dagger A$ or $AA^\dagger$ requires finding only one of the ranges. With the exception of parts 2 and 3, which are included for completeness, the next theorem contains several potentially useful reformulations of EP for matrices. The theorem is motivated by a result of Embry [7] which states that if both A^*A and AA^* commute with $A+A^*$, then A is normal.

THEOREM 5. *Suppose that A is an $n\times n$ complex matrix. Then the following conditions are equivalent.*

(1) *A is* EP.
(2) $[A^\dagger A, A+A^\dagger]=0$.
(3) $[AA^\dagger, A+A^\dagger]=0$.
(4) $[A^\dagger A, A+A^*]=0$.

https://doi.org/10.4153/CMB-1975-061-4 Published online by Cambridge University Press

(5) $[AA^{\dagger}, A+A^*]=0$.
(6) $[A, A^{\dagger}A]=0$.
(7) $[A, AA^{\dagger}]=0$.

Proof. Clearly (1) implies the rest. To show that (2)⇒(1), assume that $A^{\dagger}A(A+A^{\dagger})=(A+A^{\dagger})A^{\dagger}A$. Then $A^{\dagger}A^2+A^{\dagger}-A^{\dagger 2}A=A$. Hence $R(A)\subseteq R(A^{\dagger})=R(A^*)$. Thus $R(A)=R(A^*)$ and A is EP. The proofs of (3)⇒(1), (4)⇒(1), and (5)⇒(1) are similar. Suppose then (6) so that $A(A^{\dagger}A)=(A^{\dagger}A)A$ or $A=A^{\dagger}A^2$.

Then $AA^{\dagger}=(A^{\dagger}A)(AA^{\dagger})$. Since $A^{\dagger}A$ and $AA^{\dagger}$ are self-adjoint projections of the same rank, this implies $A^{\dagger}A=AA^{\dagger}$ and A is EP. The final implication (7)⇒(1) is similar to (6)⇒(1).

EXAMPLE 4. Suppose $\{e_0, e_1, \ldots\}$ is an orthonormal basis for a separable, infinite dimensional, Hilbert space. Let S be the bilateral shift defined by $Se_i=e_{i+1}$ and extended linearly. Let P be the projection onto the subspace spanned by $\{e_1, e_2, \ldots\}$. It is easy to verify that $S^*=S^{\dagger}$ and $SS^*=P$ while $S^*S=I$, the identity Thus S is not EP since $SS^{\dagger}\neq S^{\dagger}S$. But S satisfies (2), (4), and (6). S^* is not EP and satisfies (3), (5), and (7). The assumption of finite dimensionality was thus crucial to all parts of Theorem 5.

The dual to Embry's result is valid even if H is infinite dimensional.

THEOREM 6. *Let* $A\in B(\mathscr{h})$. *If* $[A^{\dagger}A, A+A^{\dagger}]=0$ *and* $[AA^{\dagger}, A+A^{\dagger}]=0$, *then* A *is* EP.

Proof. Suppose $[A^{\dagger}A, A+A^{\dagger}]=0$. Then $A^{\dagger}A^2+A^{\dagger}=A+A^{\dagger 2}A$. Multiplication on the right by $A^{\dagger}$ gives $(A^{\dagger}A)(AA^{\dagger})=AA^{\dagger}$. Thus $R(A)\subseteq R(A^*)$. Similarily $[AA^{\dagger}, A+A^{\dagger}]=0$ gives $R(A^*)\subseteq R(A)$ and hence A is EP.

In [10] it was shown for matrices, that if $R(A^2)=R(A)$ and $A^{\dagger 2}=A^{2\dagger}$, then A is EP. We now improve this result.

THEOREM 7. *Let* A *be an* $n\times n$ *complex matrix. If* $R(A^2)=R(A)$ *and* $[A^{\dagger}A, AA^{\dagger}]=0$, *then* A *is* EP.

Proof. Suppose that $[A^{\dagger}A, AA^{\dagger}]=0$, that is, that the projections onto $R(A)$ and $N(A)$ commute. Thus $\mathbb{C}^n=R(A)\cap R(A^*)\oplus R(A)\cap N(A)\oplus N(A^*)\cap N(A)\oplus N(A^*)\cap R(A^*)$. But if $R(A^2)=R(A)$, then $R(A)\cap N(A)=\{0\}$. Since $\mathbb{C}^n$ is finite dimensional, we also have $R(A^{*2})=R(A^*)$ and hence $R(A^*)\cap N(A^*)=\{0\}$. Thus $N(A)$ is perpendicular to $R(A)$ and A is EP.

If $A=S^*$, S as in Example 4, then $R(A^2)=R(A)$ and $A^{2\dagger}=A^{\dagger 2}$ but A is not EP. Thus finite dimensionality was needed both for Theorem 7 and the original result in [10].

Not all the results of [10] can be improved by the substitution of one of our conditions. For example, it was shown in [10] that the matrix A is normal if and only if A is EP and $[A^{\dagger}, A^*]=0$. We cannot weaken A is EP to any of (E2), (E3), or

(E4). $A=\begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}$ satisfies (E2), and $[A^\dagger, A^*]=0$, while A is not normal.

7. The condition $[A^*, A^\dagger]=0$ is different from our (E) conditions in the quantitative way discussed earlier. We give two examples to show that it is actually independent of (E4).

EXAMPLE 5. Let

$$A=\begin{bmatrix} 0 & \begin{bmatrix}1 & 1\\ 0 & 1\end{bmatrix} & 0\\ 0 & 0 & 0\\ \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} & 0 & 0\end{bmatrix}.$$

Then

$$A^\dagger=\begin{bmatrix} 0 & 0 & \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\\ \begin{bmatrix}1 & -1\\ 0 & 1\end{bmatrix} & 0 & 0\\ 0 & 0 & 0\end{bmatrix} \quad \text{and} \quad [AA^\dagger, A^\dagger A]=0,$$

but $[A^\dagger, A^*]\neq 0$.

EXAMPLE 6. Let $A=\begin{bmatrix}1/\sqrt{2} & 1/\sqrt{2}\\ 0 & 0\end{bmatrix}$. Then $A^*=A^\dagger$ and hence $[A^*, A^\dagger]=0$. But $[A^\dagger A, AA^\dagger]\neq 0$. Notice also that $R(A^2)=R(A)$.

8. The proof of Theorem 7 suggests that matrices satisfying (E4) must have a nice standard form.

THEOREM 8. *If A is an $n\times n$ matrix and $[A^\dagger A, AA^\dagger]=0$, then there exists a unitary matrix U, and matrices A_{11}, A_{12}, A_{31}, A_{32} such that*

$$A=U\begin{bmatrix}A_{11} & A_{12} & 0 & 0\\ 0 & 0 & 0 & 0\\ A_{31} & A_{32} & 0 & 0\\ 0 & 0 & 0 & 0\end{bmatrix}U^*=U\begin{bmatrix}I & 0 & 0 & 0\\ 0 & 0 & I & 0\\ 0 & I & 0 & 0\\ 0 & 0 & 0 & I\end{bmatrix}\begin{bmatrix}A_{11} & A_{12} & 0 & 0\\ A_{31} & A_{32} & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{bmatrix}U^*$$

where $\begin{bmatrix}A_{11} & A_{12}\\ A_{31} & A_{32}\end{bmatrix}$ is invertible and each I is an identity matrix.

Proof. Suppose $[A^\dagger A, AA^\dagger]=0$. Consider A as a linear transformation on $\mathbb{C}^n$ with the standard basis. Since $[A^\dagger A, AA^\dagger]=0$ we have

$$\mathbb{C}^n=R(A)\cap R(A^*)\oplus R(A^*)\cap N(A^*)\oplus R(A)\cap N(A)\oplus N(A)\cap N(A^*).$$

Pick an orthonormal basis for each summand and combine to get an orthonormal basis for $\mathbb{C}^n$. Let U^*AU be the matrix of A relative to this new basis. Relative to the decomposition of $\mathbb{C}^n$ we have $U^*AU=[A_{ij}]$, $1\le i,j\le 4$, a 4×4 block matrix. An easy computation shows that $A_{ij}=0$ except possibly for A_{11}, A_{12}, A_{31}, and A_{32}. We note that $\dim[R(A^*)\cap N(A^*)]=\dim[R(A)\cap N(A)]$ since rank A^*-

https://doi.org/10.4153/CMB-1975-061-4 Published online by Cambridge University Press

rank A^{*2}=rank A−rank A^2. That $\begin{bmatrix} A_{11} & A_{12} \\ A_{31} & A_{32} \end{bmatrix}$ is invertible follows from dim $R(A)$=dim $R(A^*)$ and the decomposition.

If (E4) is satisfied and A is an $n\times n$ matrix, then $A^\dagger$ can be calculated from U and $\begin{bmatrix} C_{11} & C_{13} \\ C_{21} & C_{23} \end{bmatrix}=\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}^{-1}$. In fact,

$$A^\dagger = U\begin{bmatrix} C_{11} & 0 & C_{13} & 0 \\ C_{21} & 0 & C_{23} & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}U^*.$$

The standard form in Theorem 8 has several uses. In addition to providing a way of calculating $A^\dagger$, it can also be useful in producing examples and counter examples. The standard form was used in constructing Examples 3 and 5.

If $\mathscr{h}$ is allowed to be infinite dimensional, then we still get the first block form of Theorem 8, but $\begin{bmatrix} A_{11} & A_{12} \\ A_{31} & A_{32} \end{bmatrix}$ may no longer be invertible or square.

References

1. T. S. Baskett and I. J. Katz, *Theorems on products of* EP_r *matrices*, Linear Algebra and Its Applications, **2** (1969), 87–103.

2. F. J. Beutler, *The operator theory of the pseudoinverse I: bounded operators*, J. Math. Anal. Appl. **10** (1965), 451–470.

3. Arlen Brown, *On a class of operators*, Proc. Amer. Math. Soc. (1953), 723–728.

4. Stephen L. Campbell, *Linear operators for which T*T and TT* commute*, Proc. Amer. Math. Soc., **34** (1972), 177–180.

5. C. A. Desoer and B. H. Wahlen, *A note on pseudoinverses*, SIAM J. Appl. Math., **11** (1963), 442–447.

6. Nelson Dunford and Jacob T. Schwartz, *Linear Operators Part II*, Interscience Publishers, New York, 1963.

7. Mary R. Embry, *Conditions implying normality in Hilbert space*, Pacific J. Math., **18** (1966), 457–460.

8. Paul R. Halmos, *A Hilbert Space Problem Book*, D. Van Nostrand Co., Princeton, New Jersey, 1967.

9. I. J. Katz and M. H. Pearl, *On* EP_r *and normal* EP_r *matrices*, J. of Research of the N.B.S., Series B, **70B** (1966), 47–75.

10. Carl D. Meyer, *Some remarks on* EP_r *matrices and generalized inverses*, Linear Algebra and Its Applications, **3** (1970), 275–278.

11. Martin H. Pearl, *On normal and* EP_r *matrices*, Mich. Math. J., **6** (1959), 1–5.

12. Martin H. Pearl, *On generalized inverses of matrices*, Proc. Cambridge Phil. Soc. **62** (1966), 673–677.

13. W. V. Petryshyn, *On generalized inverses and on the uniform convergence of* $(I-BK)^n$ *with applications to iterative methods*, J. Math. Anal. Appl., **18** (1967), 417–439.

14. H. Schwerdtfeger, *Introduction to Linear Algebra and the Theory of Matrices*, P. Noordhoff N.V., Groningen-Holland, 1961.

15. Kôsaku Yosida, *Functional Analysis*, Springer Verlag, New York, 1968.

Department of Mathematics
North Carolina State University
Raleigh, North Carolina 27607