COMMUTATORS AND COMPRESSIONS

BY
J. H. ANDERSON AND J. G. STAMPFLI*

ABSTRACT
We present a short proof of the known theorem that every operator, not of the form $\lambda I +$ compact, where $\lambda \neq 0$, is a commutator.

In [2] Brown, Pearcy, and Halmos showed that every compact operator is a commutator. In [3] Brown and Pearcy showed that operators not of the form "scalar plus compact" also are commutators. The purpose of this note is to present one simple proof from which both of these results follow.

We are grateful to A. Brown, P. Fillmore, C. Pearcy, and J. P. Williams for their comments and suggestions.

In the sequel, $\mathfrak{H}$ will denote a separable, infinite dimensional Hilbert space, and $\mathfrak{B}(\mathfrak{H})$ will denote the algebra of all bounded linear operators acting on $\mathfrak{H}$. An operator T is a commutator if $T = AB - BA$, for $A, B \in \mathfrak{B}(\mathfrak{H})$. The class of operators of the form $\lambda + K$ where λ is a nonzero complex number and K is compact will be denoted by $\mathfrak{G}$. When we designate an operator as an $n \times n$ operator matrix, it is understood that all entries map to and from infinite dimensional spaces, except in a few obvious cases.

Definition 1. Let $T \in \mathfrak{B}(\mathfrak{H})$. Then the essential numerical range of T, $W_e(T)$ is defined to be

$$\bigcap_K W(T+K)^-$$

where the intersection is taken over all compact operators K, $W(T)$ denotes the usual numerical range and $(\ )^-$ denotes closure.

A detailed discussion of the essential numerical range can be found in [5]. We will need only a few elementary properties.

* The second author gratefully acknowledges the support of the National Science Foundation.

Received June 25, 1971

LEMMA 2. *Let $T \in \mathfrak{B}(\mathfrak{H})$. Then the following are equivalent.*

a) $\lambda \in W_e(T)$.

b) *There is an orthonormal set $\{e_n\}$ such that $(Te_n, e_n) \to \lambda$.*

c) *There is a decomposition of $\mathfrak{H}$ as $\mathfrak{H}_1 \oplus \mathfrak{H}_2$ such that*

$$T = \left[\begin{array}{c|c} \begin{matrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ 0 & & \ddots \end{matrix} & * \\ \hline * & * \end{array}\right]$$

where $\lambda_i \to \lambda$.

PROOF. Clearly, $(c) \Rightarrow (b) \Rightarrow (a)$, so it suffices to show $(a) \Rightarrow (c)$. Hence, suppose $\lambda \in W_e(T)$. Then $\lambda \in W(T)^-$ so there is a unit vector $f_1 \in \mathfrak{H}$ such that $\lambda_1 = (Tf_1, f_1)$ and $|\lambda - \lambda_1| < 1$. Let $M_1 = \mathrm{clm}\{f_1, Tf_1, T^*f_1\}$. Let P_1 be the projection onto M_1 and let $1 - P_1 = Q_1$. Note that

$$F = \mu P_1 - P_1TP_1 - P_1TQ_1 - Q_1TP_1$$

is compact for any complex μ and, therefore, $\lambda \in W(T + F)^- = W(\mu P_1 + Q_1TQ_1)^-$. If we take $\mu \in W(Q_1TQ_1 \mid Q_1\mathfrak{H})$, the last statement and the convexity of $W(Q_1TQ_1 \mid Q_1\mathfrak{H})$ implies that $\lambda \in W(Q_1TQ_1 \mid Q_1\mathfrak{H})^-$. Thus it follows that there is a unit vector $f_2 \in M_1^\perp$ such that $\lambda_2 = (Tf_2, f_2)$ and $|\lambda - \lambda_2| < \frac{1}{2}$. Having chosen $f_1, \cdots, f_n$, we set $M_n = \mathrm{clm}\{f_1, \cdots, f_n, Tf_1, \cdots, Tf_n, T^*f_1, \cdots, T^*f_n\}$. Let P_n be the projection onto M_n, and let $1 - P_n = Q_n$. By the same argument as before, $\lambda \in W(\mu P_n + Q_nTQ_n)^-$ and it follows that there is a unit vector $f_{n+1} \in M_1^\perp$ such that $\lambda_{n+1} = (Tf_{n+1}, f_{n+1})$ and $|\lambda - \lambda_{n+1}| < 1/(n+1)$. Thus, we obtain an orthonormal sequence $\{f_n\}$, such that $(Tf_n, f_n) = 0$ if $n \neq m$ and $(Tf_n, f_n) \to \lambda$. If we set $\mathfrak{H}_1 = \mathrm{clm}\{f_n\}$ and $\mathfrak{H}_2 = \mathfrak{H} \ominus \mathfrak{H}_1$, it is easy to check that T has the desired form.

COROLLARY. *Let $\alpha, \beta \in W_e(T)$. Then T can be represented as*

$$\left[\begin{array}{c|c} \begin{matrix} \alpha_1 & & & & 0 \\ & \beta_1 & & & \\ & & \alpha_2 & & \\ 0 & & & \beta_2 & \\ & & & & \ddots \end{matrix} & * \\ \hline * & * \end{array}\right] \quad on\ \mathfrak{H}_1 \oplus \mathfrak{H}_2 = \mathfrak{H},$$

where $\{\alpha_i\}$ and $\{\beta_i\}$ are complex sequences tending to α and β respectively. Note that we may take $\mathfrak{H}_2$ to be infinite dimensional.

The corollary is an obvious consequence of the proof of a) $\to$ c) in Lemma 2.

LEMMA 3. *Let* $T \in \mathfrak{B}(\mathfrak{H})$. *Then* T *is compact if and only if* $(Te_n, e_n) \to 0$ for *every orthonormal set* $\{e_n\}$.

PROOF. Note that $(Te_n, e_n) \to 0$ if and only if $(\mathrm{Re}Te_n, e_n) \to 0$ and $(\mathrm{Im}Te_n, e_n) \to 0$. Thus, we may assume that T is self-adjoint. Furthermore, every self-adjoint operator is the difference of two positive operators acting on orthogonal Hilbert spaces. Hence, it suffices to consider only positive operators. But if $T \geqq 0$ and $(Te_n, e_n) \to 0$ for every orthonormal set $\{e_n\}$, then it is obvious from the spectral theorem that T is compact. (See [6] chap. 2, §7). The reverse implication is obvious.

COROLLARY. $T \in \mathfrak{G}$ *if and only if* $W_e(T) = \{\lambda\}$ *for some non-zero complex* λ.

PROOF. $W_e(T) = \{\lambda\}$ if and only if $W_e(T - \lambda) = \{0\}$, and by the lemma $W_e(T - \lambda) = \{0\}$ if and only if $T \in \mathfrak{G}$.

REMARK. Lemma 3 also follows from the fact that the essential numerical radius is a norm on the Calkin algebra.

LEMMA 4. *Let* $a \neq b$ *be complex numbers. Then for* k *sufficiently large* $0 \in W(QAQ^{-1})$ *where*

$$A = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \quad \text{and} \quad Q = \begin{pmatrix} 1 & k \\ 0 & -1 \end{pmatrix}.$$

In fact any $k \geqq 1 + 2|a + b|/|a - b|$ will suffice.

PROOF. Since

$$QAQ^{-1} = \begin{pmatrix} a & k(a - b) \\ 0 & b \end{pmatrix},$$

$W(QAQ^{-1})$ is the ellipse with foci at a, b and minor axis $k(a - b)$, and it is clear that the assertion is true for large k. The last statement follows from an easy argument which we omit.

The next theorem first appeared in [3]. The proofs seem to have little in common.

THEOREM 1. *If* T *is not compact and if* T *is not in* $\mathfrak{G}$, *then there is an invertible operator* R *such that*

$$RTR^{-1} = \begin{pmatrix} 0 & * \\ * & * \end{pmatrix}$$

in an appropriate basis.

PROOF. By the corollary to Lemma 2 and Lemma 3, T has the form

$$\left[\begin{array}{ccccc|c} \alpha_1 & & & & & \\ & \beta_1 & & & & \\ & & \alpha_2 & & 0 & * \\ & & & \beta_2 & & \\ 0 & & & & \cdot & \\ & & & & \cdot & \\ & & & & \cdot & \\ \hline & & * & & & * \end{array}\right]$$

where $\{\alpha_n\} \to \alpha$ and $\{\beta_n\} \to \beta$. Let $\mathfrak{M}_n$ be the 2-dimensional subspace of $\mathfrak{H}$ such that the compression of T to $\mathfrak{M}_n$ is

$$\begin{pmatrix} \alpha_n & 0 \\ 0 & \beta_n \end{pmatrix}.$$

By selecting a subsequence, we may assume that

$$k = 1 + 2 \sup_n |\alpha_n + \beta_n| / |\alpha_n - \beta_n| < \infty.$$

Choose Q as in Lemma 4 and set

$$R = \left[\begin{array}{cccc|c} Q & & & & \\ & Q & & & \\ & & Q & & 0 \\ & & & \cdot & \\ & & & \cdot & \\ & & & \cdot & \\ \hline & & 0 & & 1 \end{array}\right] \text{ on } \Sigma \oplus \mathfrak{M}_n \oplus \mathfrak{H}_0 = \mathfrak{H}.$$

Thus, for each n there exists a unit vector $h_n \in \mathfrak{M}_n$ such that $(RTR^{-1}h_n, h_n) = (QTQ^{-1}h_n, h_n) = 0$. Let $\mathfrak{H}_1$ be the span of the orthonormal set $\{h_n\}$ and let $\mathfrak{H}_2$ be the orthogonal complement of $\mathfrak{H}_1$. Clearly, RTR^{-1} has the desired form on $\mathfrak{H}_1 \oplus \mathfrak{H}_2$.

THEOREM 2. *Let K be compact in $\mathfrak{B}(\mathfrak{H})$. Then there is an invertible operator R such that*

$$RTR^{-1} = \begin{pmatrix} 0 & * \\ * & * \end{pmatrix}$$

PROOF. By Lemma 2

$$K = \left[\begin{array}{cccc|c} \varepsilon_1 & & & & \\ & \varepsilon_2 & 0 & & * \\ 0 & & \cdot & & \\ & & & \cdot & \\ & & & & \cdot \\ \hline & & * & & * \end{array}\right]$$

where $\varepsilon_n \to 0$. Note that if infinitely many of the $\varepsilon_n = 0$ we are done. If this is not the case, by passing to a subsequence if necessary, we may assume that $|\varepsilon_{2n}| < |\varepsilon_{2n-1}|/2$. Then if

$$A = \begin{pmatrix} \varepsilon_{2n-1} & 0 \\ 0 & \varepsilon_{2n} \end{pmatrix} \quad \text{and} \quad Q = \begin{pmatrix} 1 & 7 \\ 0 & -1 \end{pmatrix},$$

it follows from Lemma 4 that $0 \in W(QAQ^{-1})$. We may now argue exactly as in the proof of the previous theorem to establish the result.

The next lemma must be considered a folk theorem. It also appears in the guise of equivalence of operators (see [9] §1). It was brought to our attention by A. Brown.

LEMMA 5. *Let T be a bounded linear transformation of $\mathfrak{H}$ into $\mathfrak{K}$ (both separable Hilbert spaces). Then for suitable decompositions of $\mathfrak{H}$ and $\mathfrak{K}$, T has the form $T = \begin{pmatrix} * & 0 \\ 0 & * \end{pmatrix}$ (where all factors are infinite dimensional).*

PROOF. If T has finite rank, there is no difficulty. If rank T is infinite let $T = VB$ be the polar decomposition of T where B is a positive operator on $\mathfrak{H}$ and V is a partial isometry from $\mathfrak{H}$ to $\mathfrak{K}$. Let $\mathfrak{M}$ be a "half" of range B which reduces B. Then $\mathfrak{H} = \mathfrak{M} \oplus \mathfrak{M}^{\perp}$ and $\mathfrak{K} = V\mathfrak{M} \oplus (V\mathfrak{M})^{\perp}$ are the desired decompositions of $\mathfrak{H}$ and $\mathfrak{K}$.

The following is due to David [4].

LEMMA 6. *Let $T \in \mathfrak{B}(\mathfrak{H})$ have the form*

$$T = \begin{pmatrix} A_1 & B_1 \\ B_2 & A_2 \end{pmatrix}$$

where A_i and B_i are operators for $i = 1,2$. If A_1 and A_2 are commutators, then so is T.

PROOF. Let $A_i = C_iX_i - X_iC_i$ for $i = 1,2$. We claim that

$$T = \begin{pmatrix} (C_1 + t) & 0 \\ 0 & C_2 \end{pmatrix} \begin{pmatrix} X_1 & Y_1 \\ Y_2 & X_2 \end{pmatrix} - \begin{pmatrix} X_1 & Y_1 \\ Y_2 & X_2 \end{pmatrix} \begin{pmatrix} (C_1 + t) & 0 \\ 0 & C_2 \end{pmatrix}$$

where Y_1, Y_2 and t are still to be chosen. By computation we find the right hand side is

$$\begin{pmatrix} A_1 & (C_1 + t)Y_1 - Y_1C_2 \\ C_2Y_2 - Y_2(C_2 + t) & A_2 \end{pmatrix}.$$

Thus, we need only satisfy the equations

$$B_1 = (C_1Y_1 - Y_1C_2) + tY_1$$

$$B_2 = (C_2Y_2 - Y_2C_1) - tY_2$$

to complete the proof. But the map $\alpha: \mathfrak{B}(\mathfrak{H}) \to \mathfrak{B}(\mathfrak{H})$ defined by $\alpha: T \to C_1T - TC_2$ s bounded and hence for $t > \| \alpha \|$, $(\alpha - t)$ maps $\mathfrak{B}(\mathfrak{H})$ onto $\mathfrak{B}(\mathfrak{H})$. Thus the above equations always have a solution and we are done.

Remark. Note that if C_i is self-adjoint for $i = 1,2$ then $T = DZ - ZD$ where D is self-adjoint. The same proof shows that an $n \times n$ operator valued matrix is a commutator if each of the diagonal terms is a commutator. Furthermore if each of the diagonal terms is a commutator with a self-adjoint factor, then T is also a commutator with a self-adjoint factor.

Theorem 3. *Let $T \in \mathfrak{B}(\mathfrak{H})$. If $T \notin \mathfrak{G}$ then T is a commutator.*

Proof. Clearly any operator similar to a commutator is a commutator. Thus, in view of Theorems 1 and 2 it suffices to consider operators of the form

$$T = \begin{pmatrix} 0 & * \\ * & * \end{pmatrix}.$$

By applying Lemma 5 to the lower left hand corner (and splitting each entry into a 2 by 2 operator matrix) we find

$$T = \left[\begin{array}{c|c} \begin{matrix} 00 \\ 00 \end{matrix} & * \\ \hline \begin{matrix} *0 \\ 0* \end{matrix} & * \end{array}\right].$$

By a change of basis $(2 \to 3 \to 4 \to 2)$ we obtain

$$T = \left[\begin{array}{cc|cc} 0 & * & \multicolumn{2}{c}{\multirow{2}{*}{$*$}} \\ 0 & * & & \\ \hline \multicolumn{2}{c|}{\multirow{2}{*}{$*$}} & 0 & * \\ & & 0 & * \end{array}\right].$$

Now by [10], any operator of the form $D = \begin{pmatrix} 0 & * \\ 0 & * \end{pmatrix}$ is a commutator. Therefore T is a 2×2 operator matrix whose diagonal entries are commutators and T itself is a commutator by Lemma 6. We will now present another proof of the fact that $\begin{pmatrix} 0 & * \\ 0 & * \end{pmatrix}$ is a commutator; more in keeping with the spirit of this paper.

Apply Lemma 5 to the upper right hand corner of D to obtain

$$D = \left[\begin{array}{cc|cc} 0 & 0 & * & 0 \\ 0 & 0 & 0 & * \\ \hline 0 & 0 & \multicolumn{2}{c}{\multirow{2}{*}{$*$}} \\ 0 & 0 & & \end{array}\right].$$

Change the basis as before so that

$$D = \left[\begin{array}{cc|cc} 0 & 0 & \multicolumn{2}{c}{\multirow{2}{*}{$*$}} \\ 0 & * & & \\ \hline \multicolumn{2}{c|}{\multirow{2}{*}{$*$}} & 0 & 0 \\ & & 0 & * \end{array}\right].$$

Again by Lemma 6, to show that D is a commutator, it is enough to show that $S = \begin{pmatrix} 0 & 0 \\ 0 & * \end{pmatrix}$ is a commutator. Halmos [7] had done this. We recall his extremely brief proof. First write $S = \begin{bmatrix} R & & & 0 \\ & 0 & & \\ & & 0 & \\ 0 & & & \ddots \end{bmatrix}$. Then $S = V\hat{R} - \hat{R}V$ where $V = \begin{bmatrix} 0 & 1 & & \\ & 0 & 1 & \\ 0 & & \ddots & \ddots \\ & & & \end{bmatrix}$ and $\hat{R} = \begin{bmatrix} 0 & & & \\ R & 0 & & \\ & R & 0 & \\ & & \ddots & \ddots \end{bmatrix}$. (Thus V is the backward shift of infinite multiplicity and $\hat{R}$ is the forward weighted shift with each weight equal to the operator R.)

Remarks. Note that $V^*\hat{R} = \hat{R}V^*$ so that $S = (V + V^*)\hat{R} - \hat{R}(V + V^*)$. Thus, the last part of the proof of Theorem 3 shows that T has a 4×4 operator matrix such that each element on the diagonal is a commutator with a self-adjoint factor. Hence, by the remarks after Lemma 6, we have actually shown that if T has an infinite dimensional 0 compression then $T = AX - XA$ where A is self-adjoint. Thus, every commutator is similar to a commutator with a self-adjoint factor. It is shown in [1] that T is a commutator with a self-adjoint factor if and only if $0 \in W_e(T)$. Finally, note that theorems 1, 2 and 3 show that the following conditions are equivalent.

a) T is a commutator.

b) $T \notin \mathfrak{G}$.

c) T is similar to an operator with an infinite dimensional zero compression.

Zero compressions played a crucial role in our investigation of commutators. The next result characterizes those operators which possess a λ-compression.

THEOREM 4. *Let $T \in \mathfrak{B}(\mathfrak{H})$. Then T has a λ-compression* i.e., $T = \begin{pmatrix} \lambda & * \\ * & * \end{pmatrix}$ *if and only if $\lambda \in W(T + F)$ for every finite dimensional operator F. (Note that we do not take the closure of $W(T + F)$.)*

PROOF. The proof of the "if" part of the theorem is implicit in the proof of Lemma 2. The proof of the reverse implication is easy and we omit it.

COROLLARY. *If $\lambda \in \operatorname{int} W_e(T)$ then T has a λ-compression.*

PROOF. Clear.

Of course T cannot have a λ-compression if $\lambda \notin W_e(T)$. Moreover, it is easy to see that an operator may not possess a λ-compression for any λ. (Just consider a positive compact operator with dense range). However, we will now exhibit a compact operator K such that $(Ke_n, e_n) = 0$ for each n where $\{e_n\}$ is an orthonormal basis for $\mathfrak{H}$ and yet K does not admit a 0-compression. Let $K_1 e_n = 2^{-n} e_n$ and let $f = \sum_{n=0}^{\infty} 2^{-n} e_n$. Then define F by $Ff = f$ and $F = 0$ on the orthogonal complement of $\{f\}$. Put $K = K_1 - F$. Clearly, $(Ke_n, e_n) = 0$ for each n and $0 \notin W(K + F) = W(K_1)$. Hence K cannot have a 0-compression by Theorem 4.

Lastly, we present a proof of an unpublished result of C. Pearcy.

THEOREM 5. *Let $T \notin \mathfrak{G}$. Let A be an arbitrary operator in $\mathfrak{B}(\mathfrak{H})$. Then there exists an invertible operator Q such that*

$$QTQ^{-1} = \begin{pmatrix} A & * \\ * & * \end{pmatrix}$$

PROOF. By a slight modification of the proofs of the corollary to Lemma 3 and Theorem 4 it is easy to see that T is similar to an operator of the form

$$\left[\begin{array}{cc|c} I & 0 & \\ & & * \\ 0 & 0 & \\ \hline * & & * \end{array}\right].$$

Now consider the operator

$$F = \begin{pmatrix} A & A(I-A) \\ I & I-A \end{pmatrix}$$

Clearly F is idempotent and thus F is similar to a projection. Since F has infinite dimensional range and null space it is similar to $\begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}$. The theorem follows.

In closing we mention that commutators in a von Neumann algebra were studied recently in [8].

REFERENCES

1. J. H. Anderson, *Derivations, commutators, and the essential numerical range*, Thesis, Indiana University, 1971.

2. A. Brown, P. Halmos and C. Pearcy, *Commutators of operators on Hilbert Space*, Canad. J. Math. **17** (1965), 695–708.

3. A. Brown and C. Pearcy, *Structure of commutators of operators*, Ann. Math., **82** (1965) 112–127.

4. M. David, *On a certain type of commutator*, J. Math. Mech., **19** (1970), 665–680.

5. P. Fillmore, J. G. Stampfli and J. P. Williams, *The essential spectrum, the essential numerical range and a problem of Halmos*, (to appear).

6. I. C. Gohberg and M. G. Krein, *Introduction to the Theory of Linear Nonselfadjoint Operators*, Amer. Math. Soc. Providence, R.I. 1969.

7. P. R. Halmos, A *Hilbert Space Problem Book*, Van Nostrand, Princeton, 1967.

8. H. Halpern, *Commutators modulo the center in a properly infinite von Neumann algebra*, Trans. Amer. Math. Soc. **150** (1970), 55–68.

9. G. Köthe, *Die Gleichunstheorie in Hilbertschen Raum*, Math. Z. **41** (1936), 153–162.

10. C. Pearcy, *On commutators of operators on Hilbert space*, Proc. Amer. Math. Soc. **16** (1965), 53–59.

INDIANA, UNIVERSITY
BLOOMINGTON, INDIANA