PACIFIC JOURNAL OF MATHEMATICS
Vol. 33, No. 3, 1970

THE NORM OF A DERIVATION

JOSEPH G. STAMPFLI

In this paper, we determine the norm of the inner derivation $\mathfrak{D}_T$: $A \to TA - AT$ acting on the Banach algebra $\mathfrak{B}(H)$ of all bounded linear operators on Hilbert space. More precisely, we show that $\|\mathfrak{D}_T\| = \inf\{2\|T - \lambda I\| : \lambda \text{ complex}\}$. If T is normal, then $\|\mathfrak{D}_T\|$ can be specified in terms of the geometry of the spectrum of T.

A derivation on a Banach algebra $\mathfrak{A}$ is a linear transformation $\mathfrak{D} : \mathfrak{A} \to \mathfrak{A}$ which satisfies $\mathfrak{D}(ab) = a\mathfrak{D}(b) + \mathfrak{D}(a)b$ for all $a, b \in \mathfrak{A}$. If for a fixed a, $\mathfrak{D} : b \to ab - ba$, then $\mathfrak{D}$ is called an inner derivation. Sakai has shown that every derivation on a von Neumann algebra [8] or on a simple C^*-algebra [9] is inner. See also [3] and [4].

In [7], Rosenblum determined the spectrum of an inner derivation. Our estimates on the norm of $\mathfrak{D}_T$ have some applications of general operator theory as a by product. Kadison, Lance, and Ringrose [5] have investigated the derivation $\mathfrak{D}_T$ acting on a general C^*-algebra, when T is self adjoint. In §2, we study $\mathfrak{D}_T$ acting on an irreducible C^*-algebra. There appears to be little common ground in the two approaches. In the last section we consider the operator which sends $X \to AX - XB$ for $A, B, X \in \mathfrak{B}(H)$.

1. From now on, all operators are bounded and act on a Hilbert space. We shall denote the complex numbers by $\boldsymbol{C}$.

DEFINITION. The maximal numerical range of T is the set

$$W_0(T) = \{\lambda : (Tx_n, x_n) \to \lambda \text{ where } \|x_n\| = 1 \text{ and } \|Tx_n\| \to \|T\|\}\,.$$

When H is finite dimensional, $W_0(T)$ corresponds to the numerical range produced by the maximal vectors (vectors x such that $\|x\| = 1$ and $\|Tx\| = \|T\|$).

LEMMA 1. *If* $\|T\| = \|x\| = 1$ *and* $\|Tx\|^2 \geqq (1 - \varepsilon)$, *then* $\|(T^*T - I)x\|^2 \leqq 2\varepsilon$.

Proof. Note that $0 \leqq \|(T^*T-I)x\|^2 = \|T^*Tx\| - 2\|Tx\|^2 + \|x\|^2 \leqq 2(1 - \|Tx\|^2) \leqq 2\varepsilon$.

LEMMA 2. *The set* $W_0(T)$ *is nonempty, closed, convex, and contained in the closure of the numerical range.*

Proof. Everything but convexity is obvious. Therefore, let $\lambda, \mu \in W_0(T)$. Assume, without loss of generality, that $\|T\| = 1$. Assume also that $\|x_n\| = \|y_n\| = 1$, $(Tx_n, x_n) \to \lambda$ and $(Ty_n, y_n) \to \mu$. Consider $T_n = P_n T P_n$, where P_n is the projection on H of $\{x_n, y_n\}$. Let η be a point on the line segment joining λ and μ. Then for each n, it is possible, by the Toeplitz-Hausdorff Theorem, to choose α_n, β_n such that $(Tu_n, u_n) = (T_n u_n, u_n) \to \eta$ and $\|u_n\| = 1$, where $u_n = \alpha_n x_n + \beta_n y_n$. Note that $|(x_n, y_n)| \leqq \theta < 1$ for n sufficiently large; that is, the angle between x_n and y_n is bounded away from 0. (It is not difficult to compute an explicit upper bound for $\limsup |(x_n, y_n)|$ in terms of λ and μ.) Thus, there exists a constant M such that $|\alpha_n| \leqq M$ and $|\beta_n| \leqq M$ for large n, where $\|\alpha_n x_n + \beta_n y_n\| = 1$. By Lemma 1, $\|Tu_n\| = (T^* Tu_n, u_n) = \|u_n\|^2 - 2M\varepsilon_n$ where $\varepsilon_n \to 0$, and thus it follows that $\|Tu_n\| \to 1$. Since $(Tu_n, u_n) \to \eta$ this completes the proof of convexity.

LEMMA 3. *Let $\mu \in W_0(T)$. Then* $\|\mathfrak{D}_T\| \geqq 2(\|T\|^2 - |\mu|^2)^{1/2}$.

Proof. Note that $\|\mathfrak{D}_T\| = \sup\{\|TA - AT\| : A \in \mathfrak{B}(H) \text{ and } \|A\| = 1\}$. Since $\mu \in W_0(T)$, there exist $x_n \in H$ such that $\|x_n\| = 1$, $\|Tx_n\| \to \|T\|$, and $(Tx_n, x_n) \to \mu$. Set $Tx_n = \alpha_n x_n + \beta_n y_n$, where $(x_n, y_n) = 0$ and $\|y_n\| = 1$. Set $V_n x_n = x_n$, $V_n y_n = -y_n$ and $V_n = 0$ on $\{x_n, y_n\}$. Then $\|(TV_n - V_n T)x_n\| = 2|\beta_n| \geqq 2(\|T\| - |\alpha_n|^2)^{1/2} - \varepsilon_n$, where $\varepsilon_n \to 0$. Since $\alpha_n \to \mu$, this completes the proof.

THEOREM 1. $\|\mathfrak{D}_T\| = 2\|T\|$ *if and only if* $0 \in W_0(T)$.

Proof. It follows from the previous lemma that $\|\mathfrak{D}_T\| \geqq 2\|T\|$ if $0 \in W_0(T)$. Since $\|\mathfrak{D}_T\| \leqq 2\|T\|$ for any T, sufficiency is proved. We now assume $\|\mathfrak{D}_T\| = 2\|T\|$, and hence there exist x_n and A_n such that $\|x_n\| = \|A_n\| = 1$ and $\|TA_n - A_n T)x_n\| \to 2\|T\|$. Clearly, $\|A_n x_n\| \to 1$, $\|Tx_n\| \to \|T\|$, and $\|TA_n x_n\| \to \|T\|$. Moreover, since $\|TA_n - A_n T)x_n\| \to 2\|T\|$, $TA_n x_n = -A_n Tx_n + \vec{\varepsilon}_n$ where $\|\vec{\varepsilon}_n\| \to 0$. Let $(Tx_n, x_n) \to \mu$ by choosing subsequence if necessary, i.e., $\mu \in W_0(T)$. Observe that

$$\begin{aligned}(TA_n x_n, A_n x_n) &= -(A_n Tx_n, A_n x_n) + \varepsilon_n \\ &= -(Tx_n, A_n^* A_n x_n) = -(Tx_n, x_n) + \varepsilon_n'\end{aligned}$$

where the last step follows from Lemma 1. Thus, $\lim_{n\to\infty}(TA_n x_n, A_n x_n) = -\mu$. Since both μ and $-\mu \in W_0(T)$, it follows that $0 \in W_0(T)$.

THEOREM 2. *If $0 \in W_0(T)$, then $\|T\|^2 + |\lambda|^2 \leqq \|T + \lambda\|^2$ for all $\lambda \in \boldsymbol{C}$. Conversely, if $\|T\| \leqq \|T + \lambda\|$ for all $\lambda \in \boldsymbol{C}$, then $0 \in W_0(T)$.*

Proof. If $0 \in W_0(T)$, then there exist $x_n \in H$, $||x_n|| = 1$, such that $||(T + \lambda)x_n||^2 = ||Tx_n||^2 + \alpha \operatorname{Re} \bar{\lambda}(Tx_n, x_n) + |\lambda|^2 \to ||T||^2 + |\lambda|^2$. Conversely, let $||T|| \leqq ||T + \lambda||$ for $\lambda \in \boldsymbol{C}$. Assume $0 \notin W_0(T)$. By rotating T, we may assume that $\operatorname{Re} W_0(T) \geqq \tau > 0$. Let $\mathfrak{S} = \{x \in H : ||x|| = 1$ and $\operatorname{Re}(Tx, x) \leqq \tau/2\}$. Let $\eta = \sup\{||Tx|| : x \in \mathfrak{S}\}$. Then $\eta < ||T||$. Let $\mu = \min\{\tau/2, (||T|| - \eta)/2\}$. Consider $(T - \mu)$. If $x \in \mathfrak{S}$, then $||(T - \mu)x|| \leqq ||Tx|| + \mu \leqq \eta + \mu < ||T||$. Let $Tx = (a + ib)x + y$ where $x \notin \mathfrak{S}$, $||x|| = 1$, and $(x, y) = 0$. Then $||(T - \mu)x||^2 = (a - \mu)^2 + b^2 + ||y||^2 = ||Tx||^2 + (\mu^2 - 2a\mu) < ||T||^2$ since $a > \mu > 0$. Thus, $||T - \mu|| < ||T||$, contrary to hypothesis.

COROLLARY. (*Pythagorean relation for operator*). *Let T be a bounded linear operator. Then there exists a unique $z_0 \in \boldsymbol{C}$, such that $||T - z_0||^2 + |\lambda|^2 \leqq ||(T - z_0) + \lambda||^2$ for all $\lambda \in \boldsymbol{C}$. Moreover, $0 \in W_0(T - \lambda)$ if and only if $\lambda = z_0$.*

Proof. Obviously, there exists a $z_0 \in \boldsymbol{C}$ such that $||T - z_0|| \leqq ||(T - z_0) + \lambda||$ for all $\lambda \in \boldsymbol{C}$. The rest of the corollary follows easily from Theorem 2.

REMARK. Given an operator T, we define the center (or center of mass) of T to be the point z_0 specified in the corollary, and designate it by c_T. Given an operator, how does one determine c_T? In general, there is no simple answer. However, if T is normal (or hyponormal) then c_T is the center of the smallest circle containing the spectrum. (See Corollary 1 of Theorem 4.) In any event, $c_T \in$ closure $W(T)$ as can be seen by a variation of the proof of Theorem 2. However, c_T need not be contained in the convex hull of $\sigma(T)$. There are nilpotents of order 3 which bear out this remark. A further example is provided by the Volterra operator $V\left(V : f(x) \to \int_0^x f(t)dt \text{ for } f \in L^2[0, 1]\right)$.

THEOREM 3. *Let $||S - T|| \leqq \delta$. Then*

$$|c_S - c_T| \leqq (\delta + [\delta^2 + 8\delta||S - c_S||]^{1/2})/2 .$$

In particular, the map $T \to c_T$ is continuous in the uniform operator topology.

Proof. We first assume that $c_S = 0$. Then,

$$\begin{aligned} ||T||^2 &\geqq |c_T|^2 + ||T - c_T||^2 \\ &\geqq |c_T|^2 + ||S - c_T||^2 - 2\delta||S - c_T|| + \delta^2 \\ &\geqq 2|c_T|^2 + ||S||^2 - 2\delta(||S|| + |c_T|) + \delta^2 \\ &\geqq ||T||^2 + (2|c_T|^2 - 2\delta|c_T| - 4\delta||S||) . \end{aligned}$$

Solving for c_T in the last expression on the right, we find that $|c_T| \leqq (\delta + [\delta^2 + 8\delta||S||]^{1/2})/2$. To handle the case when $c_S \neq 0$, we merely translate both T and S by $c_S I$.

LEMMA 4. *$W_0(T) \cap W_0(T + \alpha) = \emptyset$, for any $\alpha \in \boldsymbol{C}$, $\alpha \neq 0$.*

Proof. Let $\mu \in W_0(T) \cap W_0(T + \alpha)$. By an argument similar to one in Theorem 2, we see that $||T||^2 + |\lambda|^2 + 2 \operatorname{Re} \bar{\lambda}\mu \leqq ||T + \lambda||^2$ for $\lambda \in \boldsymbol{C}$, and $||T + \alpha||^2 + |\lambda|^2 + 2 \operatorname{Re} \bar{\lambda}\mu \leqq ||T + \alpha + \lambda||^2$ for $\lambda \in \boldsymbol{C}$. Letting $\lambda = \alpha$ in the first inequality, we obtain $||T||^2 + |\alpha|^2 + 2 \operatorname{Re} \bar{\alpha}\mu \leqq ||T + \alpha||^2$. Letting $\lambda = -\alpha$ in the second inequality, we obtain $||T + \alpha||^2 + |\alpha|^2 - 2 \operatorname{Re} \bar{\alpha}\mu \leqq ||T||^2$. Combining these yields $|\alpha|^2 \leqq 0$, which completes the proof.

Unlike the usual numerical range, $W_0(T)$ is extremely unstable under translation, as can be seen from Lemma 4. Indeed, under an ε perturbation, it may jump from a disk to a point (consider the bilateral shift). It is this property which makes it useful for our purposes.

The maximal range $W_0(T)$ does not satisfy the power inequality (as does the numerical range). More explicitly, $|W_0(T^n)| \not\leqq |W_0(T)|^n$ for $n = 1, 2, \cdots$. It is quite easy to construct counter examples using finite dimensional weighted shifts.

THEOREM 4. *Let $\mathfrak{D}_T$ be a derivation on $\mathfrak{B}(H)$. Then, $||\mathfrak{D}_T|| = \sup\{||TA - AT|| : A \in \mathfrak{B}(H)$ and $||A|| = 1\} = \inf_{\lambda \in \boldsymbol{C}} 2||T - \lambda||$.*

Proof. Since $||TA - AT|| = ||(T - \lambda)A - A(T - \lambda)|| \leqq 2||T - \lambda||\, ||A||$, it follows that $||\mathfrak{D}_T|| \leqq \inf_{\lambda \in \boldsymbol{C}} 2||T - \lambda||$. On the other hand, $||T - \lambda||$ is large for λ large, so $\inf ||T - \lambda||$ must be taken on at some point, say z_0. But $||T - z_0|| \leqq ||(T - z_0) + \lambda||$ for all $\lambda \in \boldsymbol{C}$ implies that $0 \in W_0(T - z_0)$. Hence, $||\mathfrak{D}_T|| = ||\mathfrak{D}_{(T - z_0)}|| = 2||T - z_0||$; which completes the proof.

REMARK. Rosenblum [7] proved that $\sigma(\mathfrak{D}_T) = \sigma(T) - \sigma(T) = \{(\lambda - \mu) : \lambda, \mu \in \sigma(T)\}$. There seems to be no simple relation between the norm and the spectral radius of $\mathfrak{D}_T$. For example, if $T = \begin{vmatrix} 0 & 1 \\ 0 & 0 \end{vmatrix}$, then $\sigma(\mathfrak{D}_T) = \{0\}$ but $||\mathfrak{D}_T|| = 1$. (In fact, $\mathfrak{D}_T$ is nilpotent of order 3, while T is obviously nilpotent of order 2.)

DEFINITION. Let K be a compact set in the complex plane. Then the radius of K is the radius of the smallest disk containing K. Caution: There is no obvious relation between the radius of a set and its diameter.

COROLLARY 1. *Let T be a normal (hyponormal) operator. Then* $\|\mathfrak{D}_T\| = \sup\{\|TA - AT\| : A \in \mathfrak{B}(H)$ *and* $\|A\| = 1\} = 2R_T$, *where* R_T *is the radius of the spectrum of* T.

Proof. Since $\|T - \lambda\| =$ spectral radius $(T - \lambda)$; clearly $\inf_{\lambda \in C}\|T - \lambda\| = R_T$.

COROLLARY 2. *Let* $0 \leqq A \leqq 1$, $0 \leqq B \leqq 1$. *Then* $\|AB - BA\| = 2\|\mathrm{Im}\, AB\| \leqq 1/2$.

Let A and B be self adjoint. The last corollary bounds the norm of the imaginary part of AB. In general, AB will not be self adjoint. However, the spectrum of AB is real and positive (see [10]). The obvious estimate $\|\mathrm{Re}\, AB\| \leqq \|A\|\,\|B\|$ can not be improved without additional restrictions. However, one can ask for a lower bound for $\mathrm{Re}\, AB$.

PROPOSITION 1. *Let* $0 \leqq A \leqq I$ *and* $0 \leqq B \leqq I$. *Then* $\mathrm{Re}\, AB \geqq -1/8$. *More generally,* $\mathrm{Re}\, AB \geqq k_1k_2 - (K_1 - k_1)(K_2 - k_2)/8$ *for* $0 \leqq k_1 \leqq A \leqq K_1$ *and* $0 \leqq k_2 \leqq B \leqq K_2$.

Proof. Let $Ax = \alpha x + \lambda y$, where $(x, y) = 0$ and $\|x\| = \|y\| = 1$. Let $(Ay, y) = \gamma$. Then, $|\lambda|^2 \leqq \alpha\gamma$ since $A \geqq 0$; and $|\lambda|^2 \leqq (1 - \alpha)(1 - \gamma)$ since $I - A \geqq 0$. Combining these yields $|\lambda|^2 \leqq \alpha(1 - \alpha)$. Let $Bx = \beta x + \eta v$ where $(x, v) = 0$. By a similar argument, $|\eta|^2 \leqq \beta(1 - \beta)$. Since $(ABx, x) = \alpha\beta + \eta\bar{\lambda}(v, y)$, it follows that

$$\mathrm{Re}\,(ABx, x) \geqq \alpha\beta - [\alpha\beta(1 - \alpha)(1 - \beta)]^{1/2}$$

and a standard argument shows that the last term has a minimum of $-1/8$ for $0 \leqq \alpha \leqq 1$, $0 \leqq \beta \leqq 1$; which proves the first part of the proposition. The rest is obvious.

It is not hard to see that these estimates are sharp. For example, if

$$A = \begin{vmatrix} 1 & 0 \\ 0 & 0 \end{vmatrix} \text{ and } B = \begin{vmatrix} 1/4 & \sqrt{3/4} \\ \sqrt{3/4} & 3/4 \end{vmatrix}$$

then $\mathrm{Re}\,(ABx, x) = -1/8$ for suitable chosen x.

2. Theorem 4 also holds for derivations on certain C^*-algebras. A C^*-algebra is a concretely given algebra of operators (on a Hilbert space H) which is uniformly closed and contains adjoints, as well as

an identity. A C^*-algebra $\mathfrak{A}$ is *irreducible* if the commutant of $\mathfrak{A}$ contains only the scalars.

THEOREM 5. *Let $\mathfrak{A}$ be an irreducible C^*-algebra on H. Let $T \in \mathfrak{A}(H)$. Then*

$$\|\mathfrak{D}_T|\mathfrak{A}\| = \sup\{\|TA - AT\| : A \in \mathfrak{A} \text{ and } \|A\| = 1\} = \inf_{\lambda \in \boldsymbol{C}} 2\|T - \lambda\| .$$

Proof. In the proof of Theorem 3, we used the fact that $\mathfrak{B}(H)$ contains an operator V such that $Vx = x$, $Vy = -y$ and $\|V\| = 1$ for any $x, y \in H$ where $(x, y) = 0$. This was really the only special feature of the algebra $\mathfrak{B}(H)$ which we needed. However, if $\mathfrak{A}$ is an irreducible C^*-algebra, then by the Kadison density theorem [2], there exists a unitary operator $U \in \mathfrak{A}$ such that $Ux = x$ and $Uy = -y$ whenever $(x, y) = 0$. The rest of the proof carries over with only trivial modifications which we shall omit.

COROLLARY. *Let $\mathfrak{A}_\alpha$ be an irreducible C^*-algebra on the Hilbert space H_α for α in the index set K. Let $\mathfrak{A} = \Sigma_\alpha \oplus \mathfrak{A}_\alpha$ on $H = \Sigma_\alpha \oplus H_\alpha$ where $\|A\| = \sup_\alpha \|A_\alpha\|$ for $A \in \mathfrak{A}$. Let $T \in \mathfrak{B}(H)$, and assume $\mathfrak{D}_T : \mathfrak{A} \to \mathfrak{A}$. Then, $\|\mathfrak{D}_T\| = \sup\{\|TA - AT\| : A \in \mathfrak{A}$ and $\|A\| = 1\} = \inf\{2\|T - Z\| : Z \in \mathfrak{Z}(\mathfrak{A})\}$, where $\mathfrak{Z}(\mathfrak{A})$ is the center of $\mathfrak{A}$.*

Proof. Since $\mathfrak{D}_T : \mathfrak{A} \to \mathfrak{A}$ it follows that $T = \Sigma \oplus T_\alpha$ where $T_\alpha \in \mathfrak{B}(H_\alpha)$. For each α choose λ_α such that $\|\mathfrak{D}_{T_\alpha}\| = 2\|T - \lambda_\alpha\|$. Then, $\|\mathfrak{D}_T\| = \sup_{A \in \mathfrak{A}} \|TA - AT\| = \sup_{A \in \mathfrak{A}} \sup_\alpha \|T_\alpha A_\alpha - A_\alpha T_\alpha\| = \sup_\alpha \|\mathfrak{D}_{T_\alpha}\| = \sup_\alpha 2\|T - \lambda_\alpha\| = 2\|T - Z_0\|$ where $Z_0 = \Sigma_\alpha \oplus \lambda_\alpha I_\alpha$. Since it is obvious that $\|\mathfrak{D}_T\| \leqq 2\|T - Z\|$, for $Z \in \mathfrak{Z}(\mathfrak{A})$ the proof is complete.

Note that the corollary is not true if we relax our conditions on $\mathfrak{A}$. For example, let $\mathfrak{A}$ consist of operator valued 2×2 matrices on $H \oplus H$ of the form $\begin{vmatrix} A & 0 \\ 0 & A \end{vmatrix}$ where $A \in \mathfrak{B}(H)$. Let $T = \begin{vmatrix} 0 & I \\ I & 0 \end{vmatrix}$. Then, $\mathfrak{D}_T : \mathfrak{A} \to \mathfrak{A}$. Indeed, $\mathfrak{D}_T = \mathfrak{D}_0$, so clearly, $\|\mathfrak{D}_T\| = 0$. But, $\inf\{\|T - Z\| : Z \in \mathfrak{Z}(\mathfrak{A})\} = 1$. Of course, the conclusion of the corollary would remain valid if we took the infimum over the commutant of $\mathfrak{A}$ in this example.

REMARK. Kadison, Lance and Ringrose have proved a variant of Theorem 4. Given a derivation $\mathfrak{D}_A$ on a general C^*-algebra, where A is self adjoint, and $A \in \mathfrak{A}$ they show there exists a $A' \in \mathfrak{A}^-$, the weak closure of $\mathfrak{A}$, such that $\mathfrak{D}_A = \mathfrak{D}_{A'}$, and $\|\mathfrak{D}_{A'}\| = 2\|A'\|$. (Actually, they prove more; namely that $\|\mathfrak{D}_{A'}|Q\mathfrak{A}\| = 2|QA'\|$ where Q is any central projection of $\mathfrak{A}^-$.) It is not difficult to modify their result to make it look more like Theorem 4. Clearly, $A - A' \in \mathfrak{A}'$ the commutant of $\mathfrak{A}$. Thus, $\|\mathfrak{D}_A\| = \inf\{2\|A - B\| : B \in \mathfrak{A}'\}$. This implies our result for irreducible C^*-algebras (but only for A self adjoint, of course).

3. In this section, we will study an operator from $\mathfrak{B}(H)$ to $\mathfrak{B}(H)$ which is not a derivation, but is related to $\mathfrak{D}_T$ of §1. Let $A, B \in \mathfrak{B}(H)$. Set $\mathfrak{T}_{AB}(X) = AX - XB$ for $X \in \mathfrak{B}(H)$. Clearly, $\mathfrak{T}_{AB}$ is a bounded linear operator on $\mathfrak{B}(H)$. Before estimating its norm we will need some additional information about $W_0(\cdot)$.

LEMMA 5. *Let* Re $W_0(A) \leqq a$. *Then, given* $\varepsilon > 0$, *there exists a* $\delta > 0$ *such that* Re $W_0(A + \lambda) < a + \varepsilon$ *for* $|\lambda| < \delta$.

Proof. Assume, without loss of generality, that $\|A\| = 1$. Let $\tau = \sup\{\|Ax\| : \|x\| = 1$ and $\operatorname{Re}(Ax, x) \geqq a + \varepsilon\}$. It is clear that $\|A + \lambda\| \geqq 1 - |\lambda|$. However, for $y \in H$ when $\|y\| = 1$ and $\operatorname{Re}(Ay, y) \geqq a + \varepsilon$, we see that $\|(A + \lambda)y\|^2 \leqq \tau^2 + 2|\lambda| + |\lambda|^2$. Thus, for $|\lambda| < (1 - \tau^2)/4$, it follows that Re $W_0(A + \lambda) < a + \varepsilon$.

DEFINITION. The set valued mapping $\lambda \to M(\lambda)$ from the complex plane to subsets of the complex plane is upper semi continuous at λ_0 if $\lim_{\lambda \to \lambda_0} \operatorname{dist}[M(\lambda), M(\lambda_0)] = 0$, or equivalently, the set $\{M(\lambda_0) + \varepsilon\} \supset M(\lambda)$ for λ sufficiently small. When the mapping is locally bounded; upper semi continuity is equivalent to the map having a closed graph.

THEOREM 6. *The mapping* $\lambda \to W_0(A + \lambda)$ *is upper semi continuous.*

Proof. Since $W_0(A + \lambda_0)$ is convex for fixed λ_0, we may box it with a finite number of support lines. By the previous lemma, $W_0(A + \lambda)$ will be contained in the box for λ close to λ_0.

DEFINITION. We define the normalized maximal numerical range $W_N(A)$ of the operator A to be the set $W_0(A/\|A\|)$ for $A \neq 0$. Although this definition may seem artificial, it is the relevant concept for studying the norm of $\mathfrak{T}_{AB}$.

COROLLARY. *If* $\|A + \lambda\| \neq 0$ *for any* λ, *then the map* $\lambda \to W_N(A + \lambda)$ *is upper semi continuous.*

THEOREM 7. *Let* $A, B \neq 0$. *Then* $\|\mathfrak{T}_{AB}\| = \sup\{\|AX - XB\| : X \in \mathfrak{B}(H)$ *and* $\|X\| = 1\} = \|A\| + \|B\|$ *if and only if* $W_N(A) \cap W_N(-B) \neq \varnothing$.

Proof. The proof is very similar to that of Theorem 1, and so we will only sketch a portion. Let $\lambda \in W_N(A) \cap W_N(-B)$. There exist $f, g \in H$ such that $||f|| = ||g|| = 1$ and $(Af, f) = \lambda||A|| + \varepsilon$ and $(Bg, g) = -\lambda||B|| + \varepsilon$. Since $(Af, f)/||A|| = -(Bg, g)/||B|| + \varepsilon'$ it is possible to define an operator U of norm $1 + \varepsilon''$ which sends g to f and $-Bg/||B||$ to $Af/||A||$. The rest of the proof is virtually unchanged.

Given two operators A and B, there exists a λ_0 such that

$$\inf_{\lambda \in C} \{||A - \lambda|| + ||B - \lambda||\} = ||A - \lambda_0|| + ||S - \lambda_0|| .$$

Unfortunately, λ_0 is no longer unique as simple examples demonstrate. However, the following lemma gives a criteria for deciding which λ_0's are minimal.

LEMMA 6. *Assume that neither A nor B is a scalar multiple of the identity. Then* $\inf_{\lambda \in C}\{||A - \lambda|| + ||B - \lambda||\} = ||A - \lambda_0|| + ||B - \lambda_0||$ *if and only if* $W_N(A - \lambda_0) \cap W_N(-(B - \lambda_0)) \neq \emptyset$.

Proof. Assume $W_N(A - \lambda_0) \cap W_N(-(B - \lambda_0)) \neq \emptyset$. Then $||\mathfrak{T}_{AB}|| = ||\mathfrak{T}_{(A-\lambda_0),(B-\lambda_0)}|| = ||A - \lambda_0|| + ||B - \lambda_0||$. But, since $||AX - XB|| = ||(A-\lambda)X - X(B-\lambda)|| \leqq ||A-\lambda|| + ||B-\lambda||$, it is clear that $||\mathfrak{T}_{AB}|| \leqq \inf_{\lambda \in C}\{||A - \lambda|| + ||B - \lambda||\}$ which proves the necessity.

For the sufficiency, we may assume without loss of generality that $\lambda_0 = 0$. Thus, for any pre-assigned λ, $\varepsilon > 0$, there exist $x, y \in H$ of unit norm, such that $||(A + \lambda)x|| + ||(B + \lambda)y|| \geqq ||A|| + ||B|| - \varepsilon$. After some algebra, we find that $\operatorname{Re} \bar{\lambda}\,[(Ax, x)/||A|| + (By, y)/||B||] \leqq K(|\lambda|^2 + \varepsilon)$ where K is a constant, independent of λ and ε.

Assume that $W_N(A) \cap W_N(-B) \neq \emptyset$. Then, $\operatorname{dist}[W_N(A), W_N(-B)] = \delta > 0$ and by continuity, $\operatorname{dist}[W_N(A + \lambda), W_N(-(B + \lambda))] > \delta/2$, for small λ. Thus, by convexity and continuity, any choice of x, y which satisfies the above conditions, must satisfy the inequality $|(Ax, x)/||A|| + (By, y)/||B|| \geqq \delta/4$ for λ small. But then we are lead to the inequality $|\lambda|\delta/4 \leqq K(|\lambda|^2 + \varepsilon)$ for a suitable choice of $\arg \lambda$ and small $|\lambda|$, which is impossible. Thus, λ_0 was not minimal, which completes the proof.

THEOREM 8. *Let $A, B \in \mathfrak{B}(H)$. Then,* $||\mathfrak{T}_{AB}|| = \sup\{||AB - XB||$: $X \in \mathfrak{B}(H)$ *and* $||X|| = 1\} = \inf_{\lambda \in C}\{||A - \lambda|| + ||B - \lambda||\}$.

Proof. Clearly, $||\mathfrak{T}_{AB}|| \leqq \inf\{||A - \lambda|| + ||B - \lambda||\}$. If A or B is equal to αI, the rest of the proof is trivial. (Just take $\lambda = \alpha$ and check.) Let $\inf_{\lambda \in C}\{||A - \lambda|| + ||B - \lambda||\} = ||A - \lambda_0|| + ||B - \lambda_0||$. Then it follows from Lemma 6 and Theorem 7 that $||\mathfrak{T}_{AB}|| = ||\mathfrak{T}_{A-\lambda_0, B-\lambda_0}|| = ||A - \lambda_0|| + ||B - \lambda_0||$, which completes the proof.

COROLLARY 1. *Let* $A \in \mathfrak{B}(H)$, *where* $||A|| = 1$ *and* $W_0(A) = \{|z| \leqq 1\}$. *Then, for any* $B \in \mathfrak{B}(H)$, $||\mathfrak{T}_{AB}|| = 1 + ||B||$.

COROLLARY 2. *Let* $\mathfrak{A}$ *be an irreducible* C^**-algebra. Set* $\mathfrak{T}_{AB}(X) = AX - XB$ *for* $A, B, X \in \mathfrak{A}$. *Then,* $||\mathfrak{T}_{AB}|| = \sup\{||AX - XB|| : X \in \mathfrak{A}$ *and* $||X|| = 1\} = \inf_{\lambda \in C}\{||A - \lambda|| + ||B - \lambda||\}$.

Proof. Simply use the Kadison density theorem, as in the proof of Theorem 5.

We will now present another proof of Theorem 8 which bypasses Lemma 6 and is interesting in its own right. The author would like to thank W. Gustin, who contributed a substantial portion of the proof including the following version of the next theorem:

THEOREM (*Kakutani* [6]). *Let* $\lambda \rightarrow M(\lambda)$ *be a upper semi continuous set valued mapping of the n-cube into the n-cube, where* $M(\lambda)$ *is a closed convex set for each* λ. *If the map leaves each point in the boundary fixed, then its image covers the n-cube.*

Although this theorem is not stated explicitly in [6], it is easily obtainable from the results found there.

Another proof of Theorem 8. One half the proof of Theorem 8 is trivial. For the other half, it is sufficient to show that $W_N(A + \lambda) \cap W_N(-(B + \lambda)) \neq \emptyset$, for some $\lambda \in \boldsymbol{C}$. We again assume that neither A nor B is equal to αI. We begin by defining a map ϕ of the open unit disc $\{|z| < 1\}$ onto the complex plane. Any reasonable, argument preserving, continuous map, such as $\phi(re^{i\theta}) = r(1 - r)^{-1}e^{i\theta}$, will do. Let $M(\lambda) = [W_N(A + \lambda) - W_N(-(B + \lambda))]/2$. We now define $\Phi(\lambda) = \lambda$ for $|\lambda| = 1$ and $\Phi(\lambda) = M(\phi(\lambda))$ for $|\lambda| < 1$. The $W_N(\cdot)$'s are closed and convex, and thus, $\Phi(\lambda)$ is a closed, convex set for each λ. The map Φ is upper semi continuous for points inside the disc by the corollary to Theorem 6.

It is easy to see that for θ fixed, $W_N(A + re^{i\theta}) \rightarrow e^{i\theta}$ as $r \rightarrow \infty$. Observe that $W_0(A + re^{i\theta}) \subset$ closure $W(A + re^{i\theta})$. This fact makes our map Φ upper semi continuous on the boundary. By the Kakutani fixed point theorem $0 \in M(\lambda_0) = [W_N(A + \lambda_0) - W_N(-(B + \lambda_0))]/2$ for some λ_0. But then $W_N(A + \lambda_0) \cap W_N(-(B + \lambda_0)) \neq \emptyset$, which is all we needed to prove, in light of Theorem 7.

QUESTIONS. Is Theorem 4 true for an operator T on a *Banach* space? Is Theorem 5 true for an arbitrary C^*-algebra (with the infimum taken over the commutant)? We may generalize the definition of $W_0(T)$ in the following way. For T an operator on a Hilbert space, set $W_\delta(T) = \text{closure}\,\{(Tx, x) : \|x\| = 1 \text{ and } \|Tx\| \geqq \delta\}$. Clearly, $W_\delta(T)$ is a closed subset of the closure of the usual numerical range, and $W_0(T) = \bigcap_{\delta < \|T\|} W_\delta(T)$. By a slight modification of a theorem of Dekker [1], it is not hard to see that $W_\delta(T)$ is connected. It would be interesting to know if $W_\delta(T)$ is convex. It is, if T is normal, or if the underlying Hilbert space is two-dimensional.

Added in proof: It is easy to see from the Kaplansky Density Theorem that, given an inner derivation $\mathfrak{D}_T$ on the C^*-algebra $\mathfrak{A}$, one might as well consider $\mathfrak{D}_T$ acting on $\mathfrak{A}^-$, the weak closure of $\mathfrak{A}$, if one wishes to evaluate $\|\mathfrak{D}_T\|$. Thus our second question has recently been answered by P. Gajendragadkar in her thesis (Indiana University, 1970). More precisely, she shows that if $\mathfrak{A}$ is a W^* algebra on a separable Hilbert space, and $\mathfrak{D}_T$ is an inner derivation on $\mathfrak{A}$ where $T \in \mathfrak{A}$, then

$$\|\mathfrak{D}_T\| = 2 \inf \{\|T - Z\|: \quad Z \text{ in the center of } \mathfrak{A}\} .$$

If $T \notin \mathfrak{A}$ then there is an example due to C. A. McCarthy, which shows that $\|\mathfrak{D}_T\|$ maybe be smaller than

$$2 \inf \{\|T - B\|; B \in \mathfrak{A}', \text{ the commutant of } \mathfrak{A}\} ,$$

where $\mathfrak{A}$ is a C^* or W^* algebra according to choice, and T can even be taken to be self adjoint. Finally, Proposition 1 appears in a paper by G. Strang in the Monthly, Jan. 1962.

The author would like to express his appreciation to W. B. Arveson, C. Davis, W. Gustin, G. Minty and J. P. Williams for many stimulating discussions. In particular, the latter conjectured Theorem 4 during one of these conversations.

REFERENCES

1. N. P. Dekker, *Joint numerical range and joint spectrum of Hilbert space operators*, Thesis, University of Amsterdam, 1969.

2. J. G. Glimm and R. V. Kadison, *Unitrary operators in C*-algebras*, Pacific J. Math., 547-556.

3. B. E. Johnson and J. R. Ringrose, *Derivations of operator algerbas and discrete group algebras*, Bull. London Math. Soc. **1** (1969), 70-74.

4. R. V. Kadison, *Derivations of operator algebras*, Ann. of Math. **83** (1966), 280-293.

5. R. V. Kadison, E. C. Lance, and J. R. Ringrose, *Derivations and automorphisms of operator algebras II*, J. of Functional Analysis **1** (1947), 204-221.

6. S. Kakutani, *A generalization of Brouwer's fixed point theorem*, Duke Math. J. **7** (1941), 457-459.

7. M. Rosenblum, *On the operator equation $BX - XA = Q$*, Duke Math. J. **23** (1956), 263-269.

8. S. Sakai, *Derivations of W*-algebras*, Ann. of Math. **83** (1966), 273-279.
9. ———, *Derivations of simple C*-algebras*, J. of Functional Analysis **2** (1968), 202-206.
10. J. P. Williams, *Spectra of products and numerical ranges*, J. Math. Anal. Appl. **17** (1967), 214-220.

Received September 2, 1969, and in revised form December 15, 1969. The author gratefully acknowledges the support of the National Science Foundation.

INDIANA UNIVERSITY