ISRAEL JOURNAL OF MATHEMATICS, Vol. 32, No. 4, 1979

A NOTE ON NORM IDEALS AND THE OPERATOR $X \to AX - XB$

BY

LAWRENCE A. FIALKOW[†]

ABSTRACT

For operators A and B on a Hilbert space $\mathcal{H}$, let τ denote the operator on $\mathcal{L}(\mathcal{H})$ defined by $\tau(X) = AX - XB$. Several equivalent conditions are given for τ to be surjective or bounded below. Analogues of these results are given for the restrictions of τ to norm ideals, and the norms of these restrictions are estimated.

1. Introduction

Let $\mathcal{H}$ denote an infinite dimensional complex Hilbert space and let $\mathcal{L}(\mathcal{H})$ denote the algebra of all bounded linear operators on $\mathcal{H}$. For operators A and B in $\mathcal{L}(\mathcal{H})$, let $\tau = \tau(A, B)$ denote the operator on $\mathcal{L}(\mathcal{H})$ defined by $\tau(X) = AX - XB$. In this note we give several equivalent conditions for τ to be surjective or bounded below. We give analogues of these results for the restrictions of τ to norm ideals, and we also estimate the norms of these restrictions.

By a well known theorem of M. Rosenblum [25], τ is invertible if $\sigma(A) \cap \sigma(B) = \emptyset$ (see below for notation). C. Davis and P. Rosenthal [11] proved that τ is bounded below if and only if $\sigma_\pi(A) \cap \sigma_\delta(B) = \emptyset$ and that τ is surjective if and only if $\sigma_\delta(A) \cap \sigma_\pi(B) = \emptyset$. In Theorem 2.1 it is proved that τ is surjective if and only if the range of τ contains each rank one operator. In Theorem 3.1 it is proved that τ is surjective if and only if τ has a bounded right inverse, and Theorem 3.4 contains the analogues of these results for the case when τ is bounded below.

Let $(\mathcal{J}, |||\cdot|||)$ denote a (uniform) norm ideal in $\mathcal{L}(\mathcal{H})$ in the sense of R. Schatten [26, pp. 54–55]. Then $\tau(\mathcal{J}) \subset \mathcal{J}$ and we denote the restriction of τ to $\mathcal{J}$

[†] The author gratefully acknowledges support by a grant from the National Science Foundation.
Received May 25, 1978 and in revised form October 8, 1978

by $\tau_{\mathcal{J}}$. For each X in $\mathcal{J}$ we have $|||AX - XB||| \leqq (\|A\| + \|B\|)|||X|||$, and thus $\tau_{\mathcal{J}}$ is a bounded operator on $(\mathcal{J}, |||\cdot|||)$. A. Brown and C. Pearcy [10] proved that $\sigma(\tau_{\mathcal{J}}) = \sigma(\tau)$; in particular, it is proved in [10] that if some $\tau_{\mathcal{J}}$ is surjective, then τ is surjective, and if $\tau_{\mathcal{J}}$ is bounded below, then τ is bounded below. For the Schatten ideals $\mathcal{C}_p$ $(1 \leqq p \leqq \infty)$ these results were independently obtained in [14] and their converses were proved for $\mathcal{C}_1$ (the trace class) and $\mathcal{C}_\infty$ (the ideal of all compact operators on $\mathcal{H}$). In the present note we obtain the complete converses. In Theorem 3.2 it is proved that if τ is surjective, then $\tau_{\mathcal{J}}$ has a right inverse in $\mathcal{L}(\mathcal{J})$ for each norm ideal $\mathcal{J}$, and thus $\tau_{\mathcal{J}}$ is surjective. In Theorem 3.5 it is proved that if τ is bounded below, then each $\tau_{\mathcal{J}}$ has a left inverse in $\mathcal{L}(\mathcal{J})$ and is thus bounded below. These results provide affirmative answers to Questions (i)–(iii) of [14, section 3].

In [28], J. G. Stampfli proved the identity $\|\tau(A,B)\| = \inf_{\lambda \in \mathbf{C}} \|A - \lambda\| + \|B - \lambda\|$; in particular, if $A = B$ and $\delta(A) = \tau(A, A)$, then $\|\delta(A)\| = 2\inf_\lambda \|A - \lambda\|$. In [16], C. K. Fong proved the analogous result for the induced operator, $\tilde{\delta}(A)$, acting on the Calkin algebra: $\|\tilde{\delta}(A)\| = 2\inf_\lambda \|\pi(A - \lambda)\|$. In section 4 we examine the extent to which Stampfli's identity applies to $\delta_{\mathcal{J}}(A)$, the restriction of $\delta(A)$ to the norm ideal $\mathcal{J}$. To this end, we say that an operator $A \in \mathcal{L}(\mathcal{H})$ is *S-universal* if $\|\delta_{\mathcal{J}}(A)\| = 2\inf_\lambda \|A - \lambda\|$ for each norm ideal $\mathcal{J}$. (This is equivalent to the condition that $\|\delta_{\mathcal{J}}(A)\|$ is independent of $\mathcal{J}$.) For an arbitrary operator A and each norm ideal $\mathcal{J}$, we prove that $\operatorname{diam}(W(A)) \leqq \|\delta_{\mathcal{J}}(A)\| \leqq 2\inf_\lambda \|A - \lambda\|$, where the lower bound represents the diameter of the numerical range of A. As an application, we show that when A is subnormal, the shape of $\sigma(A)$ determines whether A is S-universal; precisely, a subnormal operator is S-universal if and only if the diameter of the spectrum is equal to twice the radius of the smallest disk containing it (Theorem 4.12).

We conclude this section with some notation and a survey of pertinent results from the literature. Let $\mathcal{A}$ denote a complex Banach algebra with identity. For $A \in \mathcal{A}$, $\sigma(A)$, $\sigma_r(A)$, and $\sigma_l(A)$ denote, respectively, the spectrum, right spectrum, and left spectrum of A. For a Banach space $\mathcal{X}$, $\mathcal{L}(\mathcal{X})$ denotes the algebra of all bounded linear operators on $\mathcal{X}$. For $T \in \mathcal{L}(\mathcal{X})$, let $\sigma_\pi(T) = \{\lambda \in \mathbf{C} : T - \lambda$ is not bounded below$\}$, the approximate point spectrum of T. Following [11], we set $\sigma_\delta(T) = \{\lambda \in \mathbf{C} : T - \lambda$ is not surjective$\}$, the approximate defect spectrum of T. Of course, when $\mathcal{X} = \mathcal{H}$ (a Hilbert space), then $\sigma_\pi(T) = \sigma_l(T)$ and $\sigma_\delta(T) = \sigma_r(T)$. We record for ease of reference the following fundamental results concerning the mapping $\tau = \tau(A, B) : \mathcal{A} \to \mathcal{A}$, where $A, B \in \mathcal{A}$ and $\tau(X) = AX - XB$.

THEOREM 1.1. (Rosenblum [25]) *(i)* $\sigma(\tau) \subset \sigma(A) - \sigma(B) = \{\alpha - \beta: \alpha \in \sigma(A),\ \beta \in \sigma(B)\}$; *(ii) if* $\tau - z$ *is invertible, then* $(\tau - z)^{-1}(X) = (1/2\pi i)\int_\gamma (A - z - w)^{-1} X (w - B)^{-1} dw$ *for each* X *in* $\mathcal{A}$ *(where* γ *is a suitable contour independent of* X*).*

THEOREM 1.2. (Kleinecke [22] [25]) *If* $\mathcal{A} = \mathcal{L}(\mathcal{X})$, *where* $\mathcal{X}$ *is a Banach space, then* $\sigma(\tau) = \sigma(A) - \sigma(B)$.

THEOREM 1.3. (Davis–Rosenthal [11]) *Let* $\mathcal{A} = \mathcal{L}(\mathcal{H})$, *where* $\mathcal{H}$ *is a Hilbert space.* (i) $\sigma_\pi(\tau) = \sigma_\pi(A) - \sigma_\delta(B)$; (ii) $\sigma_\delta(\tau) = \sigma_\delta(A) - \sigma_\pi(B)$.

The main ingredient in the proofs of the results in section 3 is the use of analytic left and right inverses for certain elements of a Banach algebra. G. R. Allan [1] [2] proved that if $A \in \mathcal{A}$, then there exists an analytic function $L_A : \mathbf{C}\setminus\sigma_l(A) \to \mathcal{A}$ such that $L_A(\lambda)(A - \lambda) = 1$; similarly, there exists an analytic function $R_A : \mathbf{C}\setminus\sigma_r(A) \to \mathcal{A}$ such that $(A - \lambda)R_A(\lambda) = 1$; of course, for $\lambda \in \mathbf{C}\setminus\sigma(A)$, $L_A(\lambda) = R_A(\lambda) = (A - \lambda)^{-1}$. Refinements and extensions of these results for the case $\mathcal{A} = \mathcal{L}(\mathcal{H})$ may be found in [6] and [7]. Using Allan's results, the proofs in section 3 entail suitable modifications of Rosenblum's resolvent formula (Theorem 1.1(ii)).

Let $\mathcal{K}(\mathcal{H})$ (or $\mathcal{C}_\infty(\mathcal{H})$) denote the ideal of all compact operators in $\mathcal{L}(\mathcal{H})$, and let $\pi : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H})/\mathcal{K}(\mathcal{H})$ denote the canonical projection onto the Calkin algebra. For T in $\mathcal{L}(\mathcal{H})$ we set $\sigma_{le}(T) = \sigma_l(\pi(T))$ and $\sigma_{re}(T) = \sigma_r(\pi(T))$ (see [15]). For A and B in $\mathcal{L}(\mathcal{H})$, $\tilde{\tau}$ denotes the operator on the Calkin algebra defined by $\tilde{\tau}(\pi(X)) = \pi(A)\pi(X) - \pi(X)\pi(B)$.

Let $(\mathcal{J}, |||\cdot|||)$ denote a uniform (or symmetrically normed [17, p. 68]) norm ideal in $\mathcal{L}(\mathcal{H})$. If $X \in \mathcal{J}$ and $R, S \in \mathcal{L}(\mathcal{H})$, then $|||RXS||| \leqq \|R\|\|S\|\,|||X|||$, with equality if R and S are unitary. Let $\mathcal{C}_p$ (or $\mathcal{C}_p(\mathcal{H})$) denote the Schatten p-ideal, $1 \leqq p < \infty$. $\mathcal{C}_p$ consists of the compact operators X such that $\Sigma_i |\lambda_i|^p < \infty$, where $\{\lambda_i\}$ denotes the sequence of nonzero eigenvalues of $(X^*X)^{1/2}$, each repeated according to its multiplicity. For $X \in \mathcal{C}_p$, $\|X\|_p \equiv (\Sigma_i |\lambda_i|^p)^{1/p}$ and $(\mathcal{C}_p, \|\cdot\|_p)$ is a uniform norm ideal. If $(\mathcal{J}, |||\cdot|||)$ is an arbitrary (uniform) norm ideal and $X \in \mathcal{F}$, then $\|X\| \leqq |||X||| \leqq \|X\|_1$; in particular, if X is a rank one operator, then $|||X||| = \|X\|$. We denote the restriction of τ to $\mathcal{C}_p$ by τ_p.

The author wishes to thank E. Azoff, J. G. Stampfli, K. Clancey, J. Deddens, and the referee for helpful suggestions.

2. Rank one operators

Let $\mathcal{F}$ denote the ideal of all finite rank operators and let $\mathcal{F}_1$ denote the set of all rank one operators in $\mathcal{L}(\mathcal{H})$. In [10] and [14] it was shown that τ is bounded below if and only if the restriction of τ to $\mathcal{F}_1$ is bounded below, and in [14] it was also shown that τ is surjective if and only if $\mathcal{K}(\mathcal{H}) \subset \mathcal{R}(\tau)$. We next obtain an analogue of the former result by means of the following extension of the latter result: τ is surjective if and only if $\mathcal{F}_1 \subset \mathcal{R}(\tau)$. In contrast to this result, [14] contains an example of an invertible τ such that $\mathcal{F}_1 \not\subset \mathcal{R}(\tau \mid \mathcal{F})$; of course, since τ is invertible, Theorem 3.2 (below) implies that $\mathcal{F}_1 \subset \mathcal{R}(\tau \mid \mathcal{J})$ $(=\mathcal{J})$ for each norm ideal $\mathcal{J}$.

THEOREM 2.1. *$\tau \equiv \tau(A, B)$ is surjective if and only if the range of τ contains each rank one operator.*

PROOF. If τ is surjective, then clearly $\mathcal{F}_1 \subset \mathcal{R}(\tau)$; for the converse, we assume that τ is not surjective and we will show that $\mathcal{F}_1 \not\subset \mathcal{R}(\tau)$. Since τ is not surjective, Theorem 1.3 implies that $\sigma_r(A) \cap \sigma_l(B) \neq \emptyset$, and we may assume that $0 \in \sigma_r(A) \cap \sigma_l(B)$. For T in $\mathcal{L}(\mathcal{H})$, let $\delta_r(T) = \sigma_r(T) \setminus \sigma_{re}(T) = \{\lambda \in \mathbf{C} : T - \lambda$ has closed range and $0 < \dim(\ker((T-\lambda)^*)) < \infty\}$; $\delta_l(T) = \sigma_l(T) \setminus \sigma_{le}(T) = \{\lambda \in \mathbf{C} : T - \lambda$ has closed range and $0 < \dim(\ker(T-\lambda)) < \infty\}$. Now $0 \in (\delta_r(A) \cup \sigma_{re}(A)) \cap (\delta_l(B) \cup \sigma_{le}(B))$ and we consider the following cases for the location of 0.

(i) $0 \in \delta_r(A) \cap \delta_l(B)$. Let e and f denote vectors such that $\|e\| = \|f\| = 1$ and $Be = A^*f = 0$. Let Y denote the rank one partial isometry defined by $Yg = (g, e)f$ for $g \in \mathcal{H}$. For each operator X in $\mathcal{L}(\mathcal{H})$,

$$\|AX - XB - Y\|^2 \geqq \|AXe - XBe - Ye\|^2 = \|AXe - f\|^2 = \|AXe\|^2 + \|f\|^2 \geqq 1.$$

Thus $Y \notin \mathcal{R}(\tau)^-$ and so in this case $\mathcal{F}_1 \not\subset \mathcal{R}(\tau)^-$.

(ii) $0 \in \sigma_{re}(A) \cap \sigma_{le}(B)$. For this case we may adapt some results of J. G. Stampfli pertaining to the case when $A = B$ [29, lemma 1] [29, theorem 2]. For the sake of completeness, we sketch the argument; our proof also closes some apparent (minor) gaps in [29, theorem 2] concerning the dimension of some of the subspaces involved in the construction. Suppose $Y \in \mathcal{L}(\mathcal{H})$ is not a scalar multiple of the identity. We show there exists a unitary operator U such that $U^*YU \notin \mathcal{R}(\tau)$; if we apply this result to any rank one operator, then we may conclude that $\mathcal{F}_1 \not\subset \mathcal{R}(\tau)$.

Let $\mathcal{M}$ denote a separable, infinite dimensional reducing subspace for Y such that $Y \mid \mathcal{M}$ is not a scalar multiple of $1_{\mathcal{M}}$. (We do not assume that $\mathcal{M}$ is a proper subspace; indeed, if $\mathcal{H}$ is separable, we may set $\mathcal{M} = \mathcal{H}$.) Let $Y_1 = Y \mid \mathcal{M}$; since $\mathcal{M}$ is separable, there exists an orthonormal basis $\{h_n\}_{n=1}^{\infty}$ for $\mathcal{M}$ such that $(Y_1 h_n, h_m) \neq 0$ for $1 \leqq m, n$ [23, theorem 2]. Let $\alpha_n = (Yh_{3n}, h_{3n+1})$ for $n \geqq 1$. Since $0 \in \sigma_{re}(A) \cap \sigma_{le}(B)$, by a straightforward modification of [29, lemma 1] there exist orthonormal sequences $\{f_n\}_{n=1}^{\infty}$ and $\{g_n\}_{n=1}^{\infty}$ in $\mathcal{H}$ such that the following properties are satisfied: (i) $\|Bf_n\| \leqq |\alpha_n|/n$; (ii) $\|A^*g_n\| \leqq |\alpha_n|/n$; (iii) $(f_n, g_m) = 0$ for $1 \leqq n, m$; (iv) the subspace spanned by all of the vectors f_n and g_n $(n \geqq 1)$ has an infinite dimensional orthocomplement in $\mathcal{H}$.

We define a unitary operator U by the following relations: (a) $Uf_n = h_{3n}$; (b) $Ug_n = h_{3n+1}$; (c) extend U to a unitary operator ((iv) permits this). Now U^*YU is not in the range of τ; indeed, if $AX - XB = U^*YU$, then

$$|\alpha_n| = |(U^*YUf_n, g_n)| = |((AX - XB)f_n, g_n)|$$
$$= |(Xf_n, A^*g_n) - (Bf_n, X^*g_n)| \leqq 2\|X\| |\alpha_n|/n.$$

Thus $\|X\| \geqq n/2$ for each n, which is a contradiction.

(iii) $0 \in \delta_r(A) \cap \sigma_{le}(B)$. Let f be a vector such that $\|f\| = 1$ and $A^*f = 0$; thus $f \in \mathcal{H} \ominus \mathcal{R}(A)$. Since $0 \in \sigma_{le}(B)$, there exists an orthonormal sequence $\{e_n\}_{n=1}^{\infty}$ such that $\|Be_n\| < 1/n^2$. Let $\mathcal{M}$ denote the subspace spanned by $\{e_n\}_{n=1}^{\infty}$ and let Y denote the rank one operator defined as follows: $Ye_n = (1/n)f$, $n \geqq 1$; $Yg = 0$ for $g \in \mathcal{H} \ominus \mathcal{M}$. Suppose that X is in $\mathcal{L}(\mathcal{H})$ and that $AX - XB = Y$. Then $AXe_n - XBe_n = Ye_n = (1/n)f$, and so $XBe_n = AXe_n - (1/n)f$. Since $f \in \mathcal{H} \ominus \mathcal{R}(A)$, then

$$\|XBe_n\|^2 = \|AXe_n - (1/n)f\|^2 = \|AXe_n\|^2 + \|(1/n)f\|^2 \geqq 1/n^2,$$

and so $\|XBe_n\| \geqq 1/n > 0$. Thus $\|Be_n\| \neq 0$, and since $1/\|Be_n\| > n^2$, it follows that $\|XBe_n\|/\|Be_n\| > n^2/n = n$. Thus $\|X\| > n$ for each n, which is a contradiction, implying that the rank one operator Y is not in the range of τ.

(iv) $0 \in \sigma_{re}(A) \cap \delta_l(B)$. The equation $AX - XB = Y$ is equivalent to the equation (*) $B^*X^* - X^*A^* = -Y^*$. Now $0 \in \sigma_{le}(A^*) \cap \delta_r(B^*)$ and thus, from the preceding case, there exists a rank one operator Y such that (*) has no solution. Thus $Y \notin \mathcal{R}(\tau)$ and the proof is complete.

REMARK. D. A. Herrero [21] has posed the problem of characterizing when τ has dense range. By modifying some of the proofs of Stampfli's results in [29], Herrero has shown that if $\sigma_{re}(A) \cap \sigma_{le}(B) \neq \emptyset$, or if there exists a scalar λ such that $\dim\ker((A-\lambda)^*) > 0$ and $\dim\ker(B-\lambda) > 0$, then τ does not have dense range. Indeed, the proof of case (i) above is due to Herrero, and his idea of modifying Stampfli's proofs (in connection with the above problem) motivated the proof of case (ii). (However, the proof that in this case τ does not have dense range entails a different argument than that given above to prove that $\mathcal{F}_1 \not\subset \mathcal{R}(\tau)$.) Herrero has also given examples of the case when the range of τ is proper but dense.

3. Norm ideals

In this section we show that τ is surjective if and only if each $\tau_{\mathcal{J}}$ is surjective, and we then give the analogue for the case when τ is bounded below.

THEOREM 3.1. *Let A and B be in* $\mathcal{L}(\mathcal{H})$. *If* $\tau \equiv \tau(A, B)$ *is surjective, then* τ *is right invertible in* $\mathcal{L}(\mathcal{L}(\mathcal{H}))$. *Moreover, if* $\mathcal{J}$ *is a norm ideal in* $\mathcal{L}(\mathcal{H})$, *then* $\tau_{\mathcal{J}}$ *is right invertible in* $\mathcal{L}(\mathcal{J})$.

PROOF. Since τ is surjective, Theorem 1.3 implies that $\sigma_\delta(A) \cap \sigma_\pi(B) = \varnothing$. Since $\sigma_\delta(A)$ and $\sigma_\pi(B)$ are nonempty, compact, and disjoint, [25, theorem 2.1] implies that there exists a Cauchy domain D such that $\sigma_\delta(A) \subset D$ and $\sigma_\pi(B) \cap D^- = \varnothing$. In particular, D satisfies the following properties: (i) D is bounded and open; (ii) D has a finite number of components, C_i, $i = 1, \cdots, n$, and $C_i^- \cap C_j^- = \varnothing$ for $1 \leqq i, j \leqq n$ and $i \neq j$; (iii) bdry(D) is the disjoint union of a finite number of closed rectifiable Jordan curves $J_1, \cdots, J_p$. Let b(D) denote the positively oriented boundary of D.

Let $R_A(\lambda)$ denote an analytic right inverse for $A - \lambda$ defined on $\mathbf{C} \setminus \sigma_\delta(A)$ and let $L_B(\lambda)$ denote an analytic left inverse for $B - \lambda$ on $\mathbf{C} \setminus \sigma_\pi(B)$. Let $E = \mathbf{C} \setminus D^-$; since D^- is compact, E has a unique unbounded component $\mathcal{O}_\infty$, and (ii)–(iii) imply that E has a finite number of bounded components $\mathcal{O}_1, \cdots, \mathcal{O}_q$. For $1 \leqq i \leqq q$, $R_A(\lambda)$ is analytic in a neighborhood of $\mathcal{O}_i^-$; thus, from Cauchy's Theorem, it follows that $\int_{b(\mathcal{O}_i)} R_A(\lambda) d\lambda = 0$ (where b$(\mathcal{O}_i)$ denotes the positively oriented boundary of $\mathcal{O}_i$, which consists of certain of the curves J_k, $1 \leqq k \leqq p$).

Since bdry$(D) =$ bdry(E), the boundary of $\mathcal{O}_\infty$ is of the form bdry$(\mathcal{O}_\infty) = J_{i_1} \cup \cdots \cup J_{i_k}$, and we claim that $\sigma(A) \subset \operatorname{int dom}(J_{i_1}) \cup \cdots \cup \operatorname{int dom}(J_{i_k})$ (where int dom(J) denotes the interior domain determined by the Jordan curve J). If the claim is false, there is some point of $\sigma(A)$ in $\mathcal{O}_\infty^-$, and from this it follows that bdry$(\sigma(A)) \cap \mathcal{O}_\infty^- \neq \varnothing$. Since bdry$(\sigma(A)) \subset \sigma_\delta(A) \subset D \subset \mathbf{C} \setminus \mathcal{O}_\infty^-$, we have a contradiction.

Thus $\sigma(A) \subset \operatorname{int dom}(J_{i_1}) \cup \cdots \cup \operatorname{int dom}(J_{i_k})$, and it follows that $R_A(\lambda) = (A - \lambda)^{-1}$ in a neighborhood of $\mathcal{O}_\infty^-$. The Riesz functional calculus thus implies that $(1/2\pi i) \int_{b(\mathcal{O}_\infty)} R_A(\lambda) d\lambda = (1/2\pi i) \int_{b(\mathcal{O}_\infty)} (A - \lambda)^{-1} d\lambda = 1$ [24, p. 421]. Now (i)–(iii) imply that $\mathcal{O}_i^- \cap \mathcal{O}_j^- = \varnothing$ for $i \neq j$ and that each curve J_k is a subset of the boundary of exactly one component of E. Thus

$$\frac{1}{2\pi i} \int_{b(D)} R_A(\lambda) d\lambda = \sum_{j=1}^{q} \frac{1}{2\pi i} \int_{b(\mathcal{O}_j)} R_A(\lambda) d\lambda + \frac{1}{2\pi i} \int_{b(\mathcal{O}_\infty)} (A - \lambda)^{-1} d\lambda$$

$$= 1.$$

For $1 \leqq i \leqq n$, $L_B(\lambda)$ is analytic in a neighborhood of C_i^-, so Cauchy's Theorem implies $\int_{b(C_i)} L_B(\lambda)d\lambda = 0$, and it follows that $\int_{b(D)} L_B(\lambda)d\lambda = 0$.

Now for Y in $\mathcal{L}(\mathcal{H})$, let $\rho(Y) = X = -(1/2\pi i)\int_{b(D)} R_A(\lambda)YL_B(\lambda)d\lambda$. The map ρ is clearly linear and we will verify that ρ is a bounded right inverse for τ. Note that

$$\begin{aligned} AX - XB &= -\frac{1}{2\pi i}\int_{b(D)} AR_A(\lambda)YL_B(\lambda)d\lambda + \frac{1}{2\pi i}\int_{b(D)} R_A(\lambda)YL_B(\lambda)Bd\lambda \\ &= -\frac{1}{2\pi i}\int_{b(D)} (A - \lambda + \lambda)R_A(\lambda)YL_B(\lambda)d\lambda \\ &\quad + \frac{1}{2\pi i}\int_{b(D)} R_A(\lambda)YL_B(B - \lambda + \lambda)d\lambda \\ &= -Y\frac{1}{2\pi i}\int_{b(D)} L_B(\lambda)d\lambda + \left(\frac{1}{2\pi i}\int_{b(D)} R_A(\lambda)d\lambda\right)Y \\ &= Y. \end{aligned}$$

Let L denote the length of $b(D)$. From the continuity of R_A and L_B on $b(D)$, let $M_1 = \max_{\lambda \in b(D)}\|R_A(\lambda)\|$ and $M_2 = \max_{\lambda \in b(D)}\|L_B(\lambda)\|$. For each Y in $\mathcal{L}(\mathcal{H})$, it follows that $\|\rho(Y)\| \leqq (1/2\pi)LM_1M_2\|Y\|$; thus ρ is a right inverse for τ in $\mathcal{L}(\mathcal{L}(\mathcal{H}))$. If Y is in the norm ideal $(\mathcal{J}, |||\cdot|||)$, then it follows exactly as in [14, theorem 3.20] that $\rho(Y) \in \mathcal{J}$, and moreover, $|||\rho(Y)||| \leqq (1/2\pi)LM_1M_2|||Y|||$. Thus $\rho_{\mathcal{J}} \equiv \rho \mid \mathcal{J}$ is a right inverse for $\tau_{\mathcal{J}}$ in $\mathcal{L}(\mathcal{J})$.

The following result summarizes the various conditions equivalent to the surjectivity of τ.

THEOREM 3.2. *For $\tau \equiv \tau(A, B)$ the following are equivalent*:

(i) *τ is surjective*;

(ii) $\sigma_r(A) \cap \sigma_l(B) = \varnothing$;

(iii) $\mathcal{F}_1 \subset \mathcal{R}(\tau)$;

(iv) *$\tau_{\mathcal{J}}$ is surjective for some norm ideal $\mathcal{J}$*;

(v) *$\tau_{\mathcal{K}}$ is surjective for each norm ideal $\mathcal{K}$*;

(vi) *τ is right invertible in $\mathcal{L}(\mathcal{L}(\mathcal{H}))$*;

(vii) *$\tau_{\mathcal{J}}$ is right invertible in $\mathcal{L}(\mathcal{J})$ for some norm ideal $\mathcal{J}$*;

(viii) *$\tau_{\mathcal{K}}$ is right invertible in $\mathcal{L}(\mathcal{K})$ for each norm ideal $\mathcal{K}$.*

PROOF. The equivalence of (i) and (ii) is given by the theorem of Davis and Rosenthal (Theorem 1.3), and the equivalence of (i) and (iii) is the content of Theorem 2.1. Theorem 3.1 implies that (i) and (vi) are equivalent, and from Theorem 3.1 we also have (i) $\Rightarrow$ (viii) $\Rightarrow$ (v) $\Rightarrow$ (iv). Moreover, from [10] we have (iv) $\Rightarrow$ (i), completing the equivalence of each of (i)–(vi) and (viii). Theorem 3.1 also shows that (i) $\Rightarrow$ (vii) $\Rightarrow$ (iv) $\Rightarrow$ (i), thus completing the proof.

COROLLARY 3.3. *If τ is surjective, then* $\ker(\tau)$ *and* $\mathcal{R}(\rho)$ *are complementary closed subspaces of* $\mathcal{L}(\mathcal{H})$, *i.e.* $\mathcal{L}(\mathcal{H}) = \ker(\tau) \dotplus \mathcal{R}(\rho)$. *If $\mathcal{J}$ is a norm ideal, then* $\mathcal{J} = \ker(\tau_{\mathcal{J}}) + \mathcal{R}(\rho_{\mathcal{J}})$.

THEOREM 3.4. *Let A and B be in $\mathcal{L}(\mathcal{H})$. If $\tau \equiv \tau(A, B)$ is bounded below, then τ is left invertible in $\mathcal{L}(\mathcal{L}(\mathcal{H}))$; if $\mathcal{J}$ is a norm ideal in $\mathcal{L}(\mathcal{H})$, then $\tau_{\mathcal{J}}$ is left invertible in $\mathcal{L}(\mathcal{J})$.*

PROOF. The method of proof is analogous to that of Theorem 2.1. We include an outline, but omit certain details. Since τ is bounded below, Theorem 1.3 implies that $\sigma_\pi(A) \cap \sigma_\delta(B) = \varnothing$. Let E denote a Cauchy domain such that $\sigma_\delta(B) \subset E$ and $\sigma_\pi(A) \cap E^- = \varnothing$. Then, as in the proof of Theorem 3.1, $\int_{b(E)} L_A(\lambda) d\lambda = 0$ and $(1/2\pi i) \int_{b(E)} R_B(\lambda) d\lambda = 1$. For Y in $\mathcal{L}(\mathcal{H})$ we set $\psi(Y) = (1/2\pi i) \int_{b(E)} L_A(\lambda) Y R_B(\lambda) d\lambda$. As in Theorem 3.1, $\psi \in \mathcal{L}(\mathcal{L}(\mathcal{H}))$, and if $\mathcal{J}$ is a norm ideal, then $\psi(\mathcal{J}) \subset \mathcal{J}$ and $\psi_{\mathcal{J}} = \psi \mid \mathcal{J} \in \mathcal{L}(\mathcal{J})$. To complete the proof it suffices to verify that ψ is a left inverse for τ. If $X \in \mathcal{L}(\mathcal{H})$, then

$$\begin{aligned}
\psi(\tau(X)) &= \psi(AX - XB) \\
&= \frac{1}{2\pi i} \int_{b(E)} L_A(\lambda)(AX - XB) R_B(\lambda) d\lambda \\
&= \frac{1}{2\pi i} \int_{b(E)} L_A(\lambda)[((A-\lambda)+\lambda)X - X((B-\lambda)+\lambda)] R_B(\lambda) d\lambda \\
&= X\left(\frac{1}{2\pi i}\right) \int_{b(E)} R_B(\lambda) d\lambda - \left(\frac{1}{2\pi i} \int_{b(E)} L_A(\lambda) d\lambda\right) X \\
&= X.
\end{aligned}$$

The proof is now complete.

The following result summarizes the conditions for τ to be bounded below. We omit the proof, which is analogous to that of Theorem 3.2.

THEOREM 3.5. *For $\tau = \tau(A, B)$ the following are equivalent:*

(i) *τ is bounded below;*

(ii) $\sigma_\pi(A) \cap \sigma_\delta(B) = \varnothing$ [11];

(iii) *$\tau \mid \mathcal{F}_1$ is bounded below* [10], [14];

(iv) *τ is left invertible in $\mathcal{L}(\mathcal{L}(\mathcal{H}))$;*

(v) *$\tau_{\mathcal{J}}$ is bounded below for some norm ideal $\mathcal{J}$;*

(vi) $\tau_{\mathcal{J}}$ *is left invertible in* $\mathcal{L}(\mathcal{J})$ *for some norm ideal* $\mathcal{J}$;

(vii) $\tau_{\mathcal{K}}$ *is bounded below for each norm ideal* $\mathcal{K}$;

(viii) $\tau_{\mathcal{K}}$ *is left invertible in* $\mathcal{L}(\mathcal{K})$ *for each norm ideal* $\mathcal{K}$.

COROLLARY 3.6. *If* τ *is bounded below, then* $\ker(\psi)$ *and* $\mathcal{R}(\tau)$ *are complementary closed subspaces of* $\mathcal{L}(\mathcal{H})$, *i.e.* $\mathcal{L}(\mathcal{H}) = \ker(\psi) \dotplus \mathcal{R}(\tau)$. *If* $\mathcal{J}$ *is a norm ideal in* $\mathcal{L}(\mathcal{H})$, *then* $\mathcal{J} = \ker(\psi_{\mathcal{J}}) \dotplus \mathcal{R}(\tau_{\mathcal{J}})$.

We conclude this section with analogues of the above results for the operator $\tilde{\tau}$ acting on the Calkin algebra $\mathcal{C}$. In [14] it was proved that $\tilde{\tau}$ is bounded below if and only if $\sigma_{le}(A) \cap \sigma_{re}(B) = \varnothing$ and that $\tilde{\tau}$ is surjective if and only if $\sigma_{re}(A) \cap \sigma_{le}(B) = \varnothing$. Applying Allan's results to the Calkin algebra, and using the preceding method of contour integration, the following results may be proved exactly as the results for the operator case; for this reason we omit the details.

PROPOSITION 3.7. *The following are equivalent*:

(i) $\tilde{\tau}$ *is surjective*;

(ii) $\sigma_{re}(A) \cap \sigma_{le}(B) = \varnothing$ [14];

(iii) $\tilde{\tau}$ *is right invertible in* $\mathcal{L}(\mathcal{C})$. *In this case, a right inverse for* $\tilde{\tau}$ *is given by* $\tilde{\rho}(\pi(Y)) = -(1/2\pi i)\int_\gamma R_{\pi(A)}(\lambda)\pi(Y)L_{\pi(B)}(\lambda)d\lambda$, *where* γ *is a contour suitably separating* $\sigma_{re}(A)$ *and* $\sigma_{le}(B)$. *Moreover*, $\mathcal{C} = \ker(\tilde{\tau}) \dotplus \mathcal{R}(\tilde{\rho})$.

PROPOSITION 3.8. *The following are equivalent*:

(i) $\tilde{\tau}$ *is bounded below*;

(ii) $\sigma_{le}(A) \cap \sigma_{re}(B) = \varnothing$ [14];

(iii) $\tilde{\tau}$ *has a left inverse in* $\mathcal{L}(\mathcal{C})$. *A left inverse for* $\tilde{\tau}$ *is given by* $\tilde{\psi}(\pi(Y)) = (1/2\pi i)\int_\gamma L_{\pi(A)}(\lambda)\pi(Y)R_{\pi(B)}(\lambda)d\lambda$, *where* γ *separates* $\sigma_{le}(A)$ *and* $\sigma_{re}(B)$ *in a suitable fashion. In this case*, $\mathcal{C} = \ker(\tilde{\psi}) \dotplus \mathcal{R}(\tilde{\tau})$.

4. On the norm of $\tau_{\mathcal{J}}$

In this section we obtain some estimates for the norm of $\tau_{\mathcal{J}}$. Let $\tau = \tau(A, B)$. If $x \in \mathcal{J}$ and $\lambda \in \mathbf{C}$, then

$$|||\, \tau(X)\, ||| = |||\, (A-\lambda)X - X(B-\lambda)\, ||| \leqq (\|A-\lambda\| + \|B-\lambda\|)\, |||\, X\, |||,$$

and it follows that $\|\tau_{\mathcal{J}}\| \leqq \inf_{\lambda\in\mathbf{C}}(\|A-\lambda\| + \|B-\lambda\|)$. We consider primarily the case $A = B$. We set $\delta_{\mathcal{J}}(A) = \tau_{\mathcal{J}}(A, A)$ and in case $\mathcal{J} = \mathcal{C}_p$ $(1 \leqq p \leqq \infty)$ we denote $\delta_{\mathcal{J}}(A)$ by $\delta_p(A)$. For $A \in \mathcal{L}(\mathcal{H})$, let $d(A) = \inf_{\lambda\in\mathbf{C}}\|A-\lambda\|$; thus $\|\delta_{\mathcal{J}}(A)\| \leqq 2d(A)$, and in the sequel we obtain certain lower bounds for $\|\delta_{\mathcal{J}}(A)\|$.

Let e_1 and e_2 denote vectors in $\mathcal{H}$ such that $\|e_1\| = \|e_2\| = 1$ and $(e_1, e_2) = 0$. Let

V denote the rank two partial isometry defined by the following relations: $Ve_1 = e_1$, $Ve_2 = -e_2$, $Vh = 0$ if $(h, e_1) = (h, e_2) = 0$. (Note that V is the same as the operator "V_n" in the proof of [28, lemma 3].) For the norm ideal $(\mathcal{J}, |||\cdot|||)$, we set $\alpha_{\mathcal{J}} = |||V|||$; of course, $\alpha_{\mathcal{J}}$ depends only on the unitary equivalence class of V. Since $1 = \|V\| \leqq |||V||| \leqq \|V\|_1 = 2$, then $1 \leqq \alpha_{\mathcal{J}} \leqq 2$.

Proposition 4.1. *Let $A \in \mathcal{L}(\mathcal{H})$. If $\mathcal{J}$ is a norm ideal, then* $(2/\alpha_{\mathcal{J}})d(A) \leqq \|\delta_{\mathcal{J}}(A)\| \leqq 2d(A)$.

Proof. By a result of T. Ando [4, theorem 1], $d(A) = \sup\|(1-P)AP\|$, where the supremum is with respect to the set of all rank one orthogonal projections in $\mathcal{L}(\mathcal{H})$. Let $0 < \varepsilon < d(A)$ and let P denote a rank one projection such that $\|(1-P)AP\| \geqq d(A) - \varepsilon > 0$. Let e denote a unit vector in the range of P and let $f = (1-P)Ae$. Since $\|f\| = \|(1-P)AP\| > 0$, then $f \neq 0$ and $(e,f) = 0$. Define an operator W as follows: $We = e$, $Wf = -f$, $Wh = 0$ if $(h,e) = (h,f) = 0$. A calculation shows that $(AW - WA)e = 2(1-P)Ae = 2f$; also, since W is unitarily equivalent to V, then $|||W||| = |||V||| = \alpha_{\mathcal{J}}$, and thus

$$\begin{aligned} 2d(A) &\geqq \|\delta_{\mathcal{J}}(A)\| \geqq |||AW - WA|||/|||W||| \geqq \|AW - WA\|/\alpha_{\mathcal{J}} \\ &\geqq \|(AW - WA)e\|/\alpha_{\mathcal{J}} = \|2(1-P)Ae\|/\alpha_{\mathcal{J}} \\ &= (2/\alpha_{\mathcal{J}})\|(1-P)AP\| \geqq (2/\alpha_{\mathcal{J}})(d(A) - \varepsilon). \end{aligned}$$

The result follows directly.

Remark. The preceding argument also recaptures Stampfli's indentity $\|\delta(A)\| = 2d(A)$ (note that $\|W\| = 1$). The apparent absense of convexity arguments in the above proof should not mislead the reader; convexity results do play a role in the proof of Ando's theorem.

Corollary 4.2. *For* $1 \leqq p \leqq \infty$, $2^{1-(1/p)}d(A) \leqq \|\delta_p(A)\| \leqq 2d(A)$.

The preceding estimate is not sharp. For example, if $\mathcal{J} = \mathcal{C}_1$, then $\alpha_{\mathcal{J}} = 2$; however, as we next show, Stampfli's identity does hold for the trace class.

Proposition 4.3. *Let $\tau = \tau(A, B)$. Then*

$$\|\tau_\infty\| = \|\tau_1\| = \|\tau\| = \inf_\lambda (\|A - \lambda\| + \|B - \lambda\|).$$

Proof. Recall that if $\mathcal{X}$ is a Banach space and $T \in \mathcal{L}(\mathcal{X})$, then $\|T\| = \|T^*\|$. Now from [14, theorem 3.13] we have $(\tau_\infty)^{**} = \tau$, so $\|\tau_\infty\| = \|\tau\|$. From [14] we also have $(\tau_\infty)^* = -\tau_1(B,A)$, so that $\|\tau_1\| = \|(\tau_\infty(B,A))^*\| = \|\tau_\infty(B,A)\| = \|\tau(B,A)\| = \inf_\lambda (\|B-\lambda\| + \|A-\lambda\|) = \|\tau\|$.

We seek to sharpen the lower bound for $\|\tau_{\mathcal{J}}(A)\|$, at least in special cases, and to this end we introduce some notation. Let K be a nonempty bounded subset of the plane, and let $\mathrm{diam}(K)$ denote the diameter of K, i.e. $\mathrm{diam}(K) = \sup_{\lambda,\beta\in K} |\lambda-\beta|$. If K is compact, among all the closed disks containing K there is a (possibly degenerate) disk D_K of smallest radius, and we let $\mathrm{rad}(K)$ denote the radius of D_K; clearly $\mathrm{diam}(K) \leqq 2\mathrm{rad}(K)$. For $T \in \mathcal{L}(\mathcal{H})$, we set $\mathcal{D}(T) = \mathrm{diam}(\sigma(T))$ and $\mathcal{R}(T) = \mathrm{rad}(\sigma(T))$. Now $\mathcal{D}(T)/\mathcal{R}(T) \leqq 2$ and the next two results give some other relationships among $d(T)$, $\mathcal{D}(T)$, and $\mathcal{R}(T)$.

Proposition 4.4. *If $T \in \mathcal{L}(\mathcal{H})$, then $\mathcal{R}(T) \leqq d(T)$.*

Proof. Let $\lambda \in \mathbf{C}$ and $\beta \in \sigma_\pi(T)$. Let $\{x_n\}_{n=1}^\infty \subset \mathcal{H}$ denote a sequence such that $\|x_n\| = 1$ for $n \geqq 1$ and $\lim \|(T-\beta)x_n\| = 0$. Since $\|T-\lambda\| \geqq \|(\beta-\lambda+T-\beta)x_n\| \geqq |\lambda-\beta| - \|(T-\beta)x_n\|$, it follows that $\|T-\lambda\| \geqq |\lambda-\beta|$. Thus $\|T-\lambda\| \geqq \sup_{\beta\in\sigma_\pi(T)} |\lambda-\beta|$, and we claim that the supremum is not less than $\mathcal{R}(T)$. If the claim is false, let γ be such that $\sup_{\beta\in\sigma_\pi(T)} |\lambda-\beta| < \gamma < \mathcal{R}(T)$. Now $\sigma_\pi(T) \subset \{z : |z-\lambda| \leqq \gamma\}$ and thus $\sigma(T) \subset \{z : |z-\lambda| \leqq \gamma\}$, contradicting the definition of $\mathcal{R}(T)$. Thus $d(T) = \inf_{\lambda\in\mathbf{C}} \|T-\lambda\| \geqq \inf_\lambda \sup_{\beta\in\sigma_\pi(T)} |\lambda-\beta| \geqq \mathcal{R}(T)$.

Corollary 4.5. *If $T \in \mathcal{L}(\mathcal{H})$, then $\mathcal{D}(T) \leqq 2d(T)$.*

Proof. From Proposition 4.4 we have $\mathcal{D}(T) \leqq 2\mathcal{R}(T) \leqq 2d(T)$.

For T in $\mathcal{L}(\mathcal{H})$ let $W(T)$ denote the numerical range of T, i.e. $W(T) = \{(Tx,x) : x \in \mathcal{H}, \|x\| = 1\}$. Let $W_e(T)$ denote the essential numerical range of T given by $W_e(T) = \{\lambda \in \mathbf{C} : \lambda = \lim(Te_n, e_n)$ for some orthonormal sequence $\{e_n\}_{n=1}^\infty \subset \mathcal{H}\}$ [15, section 5]. $W(T)$ and $W_e(T)$ are bounded and convex, and unlike $W(T)$, $W_e(T)$ is necessarily compact [15]. Since $W_e(T) \subset W(T)^-$, it follows that $\mathrm{diam}(W_e(T)) \leqq \mathrm{diam}(W(T))$, and considerations with nonzero compact nilpotent operators show that the inequality may be strict.

Lemma 4.6. *Let T be in $\mathcal{L}(\mathcal{H})$ and let α and β denote distinct elements of $W(T)$. Then there exist $\alpha', \beta' \in W(T)$ and orthogonal unit vectors e and f such that $\alpha' = (Te,e)$, $\beta' = (Tf,f)$ and $|\alpha'-\beta'| \geqq |\alpha-\beta|$. Moreover, if $|\alpha-\beta| = \mathrm{diam}(W(T))$, then the choices $\alpha' = \alpha$ and $\beta' = \beta$ are permitted.*

Proof. Let x and y be unit vectors in $\mathcal{H}$ such that $\alpha = (Tx,x)$ and $\beta = (Ty,y)$. Since $\alpha \neq \beta$, x and y are independent, and we let $\mathcal{M}$ denote the two dimensional subspace spanned by x and y. Let P denote the orthogonal projection onto $\mathcal{M}$ and let S denote the compression of T to $\mathcal{M}$, i.e. $S = PT \mid \mathcal{M}$. It is easy to verify that α and β are elements of $W(S)$ and that $W(S) \subset W(T)$. Since $\mathcal{M}$ is two dimensional, relative to a suitable orthonomal basis $\{x_1, x_2\}$ for $\mathcal{M}$, the matrix of S is upper triangular. Up to a further unitary equivalence, followed by a translation and then a rotation, we may assume, without loss of generality in the following argument, that the matrix of S relative to $\{x_1, x_2\}$ is of the form $\left(\begin{smallmatrix} \lambda & 2\theta \\ 0 & -\lambda \end{smallmatrix}\right)$, where λ and θ are nonnegative real numbers.

We consider first the case $\lambda, \theta > 0$. In this case $W(S)$ consists of the closed ellipsoidal disk $\{x + iy : x, y \in \mathbf{R}$ and $x^2/\mu^2 + y^2/\theta^2 \leqq 1\}$, where $\mu = (\lambda^2 + \theta^2)^{1/2}$ (see [9] [18, p. 109]). Clearly $\operatorname{diam}(W(S)) = 2\mu$, and since $\alpha, \beta \in W(S)$ and $W(S) \subset W(T)$, it suffices to find orthogonal unit vectors e and f in $\mathcal{M}$ such that $(Se, e) = \mu$ and $(Sf, f) = -\mu$. Let e denote a unit vector in $\mathcal{M}$ such that $(Se, e) = \mu$. If $e = \gamma x_1 + \rho x_2$, then $(\lambda^2 + \theta^2)^{1/2} = \mu = \lambda(|\gamma|^2 - |\rho|^2) + 2\theta\bar{\gamma}\rho$, so $\gamma, \rho \neq 0$, and we set $f = (|\rho|\gamma/|\gamma|)x_1 - (|\gamma||\rho|/\bar{\rho})x_2$. Now $\|f\| = 1$, $(e, f) = 0$, and a calculation shows that $(Sf, f) = \lambda(|\rho|^2 - |\gamma|^2) - 2\theta\bar{\gamma}\rho = -\mu$. Moreover, if $|\alpha - \beta| = \operatorname{diam}(W(T))$, then α and β must be the endpoints of the major axis of the ellipse. Thus $\alpha = \mu$ and $\beta = -\mu$ (or $\alpha = -\mu$ and $\beta = \mu$), so the proof is complete in this case.

If $\theta = 0$, then $W(S)$ is the closed interval $[-\lambda, \lambda]$. Thus we may take $e = x_1$, $f = x_2$ and set $\alpha' = \lambda$ and $\beta' = -\lambda$. If $|\alpha - \beta| = \operatorname{diam}(W(T))$, then α and β are the interval endpoints, so (up to a change of notation) we may assume $\alpha' = \alpha$ and $\beta' = \beta$.

If $\lambda = 0$, then $W(S)$ is the closed circular disk centered at the origin with radius θ. We may set $e = (1/2^{1/2}, 1/2^{1/2})$, $\alpha' = \theta$, $f = (1/2^{1/2}, -1/2^{1/2})$, and $\beta' = -\theta$. If $|\alpha - \beta| = \operatorname{diam}(W(T))$, α and β must be the endpoints of a diameter of $W(S)$, i.e. $\alpha = \theta e^{i\gamma}$ and $\beta = -\alpha$. In this case we may set $e = (e^{-i\gamma}/2^{1/2}, 1/2^{1/2})$, $\alpha' = \alpha$, $f = (1/2^{1/2}, -e^{i\gamma}/2^{1/2})$, $\beta' = \beta$. The proof is now complete.

REMARK. We cannot in general take $\alpha' = \alpha$ and $\beta' = \beta$. To see this, consider the operator T on $\mathbf{C}^2$ whose matrix is $\left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$. The only unit vectors e such that $(Te, e) = 1/2$ are of the form $(e^{ir}/2^{1/2}, e^{is}/2^{1/2})$ where $r - s = 2\pi n$ for some integer n. The only unit vectors f such that $(Tf, f) = 0$ are those of the form $(\gamma, 0)$ or $(0, \gamma)$ where $|\gamma| = 1$; clearly e and f are not orthogonal. On the other hand, an analogue of Lemma 4.6 for the essential numerical range does allow at least an asymptotic version of the stronger conclusion. Indeed, from [3, lemma 2 (corollary)], if $\alpha, \beta \in W_e(T)$, then there is an orthogonal decompositon of $\mathcal{H}$ relative to which T has the matrix

$$\begin{pmatrix} \begin{matrix} \alpha_1 & & & & & & & \\ & \beta_1 & & & & & & \\ & & \alpha_2 & & & 0 & & \\ & & & \beta_2 & & & & \\ & 0 & & & \ddots & & & \\ & & & & & \alpha_n & & \\ & & & & & & \beta_n & \\ & & & & & & & \ddots \end{matrix} & \Big| & * \\ \hline * & \Big| & * \end{pmatrix}$$

where $\alpha_n \to \alpha$ and $\beta_n \to \beta$. Thus $(Te_n, e_n) \to \alpha$ and $(Tf_n, f_n) \to \beta$ for certain mutually orthogonal orthonormal sequences $\{e_n\}$ and $\{f_n\}$.

Proposition 4.7. *If $T \in \mathcal{L}(\mathcal{H})$, then $\|\delta_{\mathcal{J}}(T)\| \geqq \operatorname{diam}(W(T))$ for each norm ideal $\mathcal{J}$.*

Proof. We may assume $\operatorname{diam}(W(T)) > 0$. Let $0 < \varepsilon < \operatorname{diam}(W(T))$ and let $\alpha, \beta \in W(T)$ be such that $|\alpha - \beta| > \operatorname{diam}(W(T)) - \varepsilon$. From Lemma 4.6, there exist $\alpha', \beta' \in W(T)$ and orthogonal unit vectors e and f such that $(Te, e) = \alpha'$, $(Tf, f) = \beta'$, and $|\alpha' - \beta'| \geqq |\alpha - \beta|$. Relative to the orthogonal decomposition $\mathcal{H} = \langle e \rangle \oplus \langle f \rangle \oplus \mathcal{H}'$, the operator matrix of T is of the form

$$\left(\begin{array}{cc|c} \alpha' & \delta & * \\ \gamma & \beta' & \\ \hline * & & * \end{array}\right).$$

Let V denote the rank one partial isometry with the following matrix:

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Let $(\mathcal{J}, |||\cdot|||)$ be a norm ideal. Now $|||V||| = 1$ and a matrix calculation shows that

$$\begin{aligned} \|\delta_{\mathcal{J}}(T)\| &\geqq |||TV - VT||| \geqq \|TV - VT\| \\ &\geqq |\alpha' - \beta'| \geqq |\alpha - \beta| \geqq \operatorname{diam}(W(T)) - \varepsilon. \end{aligned}$$

The result follows directly.

Corollary 4.8. *If $T \in \mathcal{L}(\mathcal{H})$, then $\|\delta_{\mathcal{J}}(T)\| \geqq \mathcal{D}(T)$ for each norm ideal $\mathcal{J}$.*

PROOF. Since $\sigma(T) \subset W(T)^-$ [18, p. 111], we have $\operatorname{diam}(W(T)) \geqq \mathcal{D}(T)$, so the result follows from Proposition 4.7.

We note that if T is subnormal, then $\mathcal{D}(T) = \operatorname{diam}(W(T))$ [18, p. 321].

COROLLARY 4.9. *If $T \in \mathcal{L}(\mathcal{H})$ and $\mathcal{D}(T) = 2d(T)$ or $\operatorname{diam}(W(T)) = 2d(T)$, then T is S-universal, i.e. $\|\delta_{\mathcal{J}}(T)\| = 2d(T)$ for each norm ideal $\mathcal{J}$.*

LEMMA 4.10. *If $T \in \mathcal{L}(\mathcal{H})$ is hyponormal and $\mathcal{D}(T) = 2\mathcal{R}(T)$, then T is S-universal.*

PROOF. Since T is hyponormal, $d(T) = \mathcal{R}(T)$ [28, corollary 1, section 1] thus, from Corollary 4.8, $2\mathcal{R}(T) = \mathcal{D}(T) \leqq \|\delta_{\mathcal{J}}(T)\| \leqq 2d(T) = 2\mathcal{R}(T)$, and the proof is complete.

LEMMA 4.11. *If $T \in \mathcal{L}(\mathcal{H})$ is subnormal, then $\|\delta_2(T)\| = \mathcal{D}(T)$.*

PROOF. We assume first that T is a normal operator with finite spectrum. Let $\{\lambda_1, \cdots, \lambda_n\}$ denote the distinct elements of $\sigma(T)$. Thus, relative to the decomposition $\mathcal{H} = \ker(T - \lambda_1) \oplus \cdots \oplus \ker(T - \lambda_n)$ we have $T = \lambda_1 \oplus \cdots \oplus \lambda_n$. Let $X \in \mathcal{C}_2$ and let $(X_{ij})_{1 \leqq i, j \leqq n}$ denote the operator matrix of X with respect to the above decomposition. The matrix of $TX - XT$ assumes the form $((\lambda_i - \lambda_j)X_{ij})_{1 \leqq i, j \leqq n}$, so $\|TX - XT\|_2^2 = \sum_{1 \leqq i, j \leqq n} |\lambda_i - \lambda_j|^2 \|X_{ij}\|_2^2 \leqq \mathcal{D}(T)^2 \|X\|_2^2$. Thus $\|\delta_2(T)\| \leqq \mathcal{D}(T)$, and the reverse inequality follows from Corollary 4.8.

We next consider the case when T is a normal operator whose spectrum is infinite. Thus there exists a sequence of normal operators $\{N_k\}$ such that $\|N_k - T\| \to 0$ and such that each $\sigma(N_k)$ is finite. Note that in general the mapping $T \to \mathcal{D}(T)$ is not norm continuous. However, using properties of the spectral measure, we may assume that in this case $\mathcal{D}(N_k) \to \mathcal{D}(T)$. Since $\sigma(N_k)$ is finite it follows from above that $\|\delta_2(N_k)\| \to \mathcal{D}(T)$. Now since $N_k \to T$, we have $\|\delta_2(N_k)\| \to \|\delta_2(T)\|$, so the proof of this case is complete.

For the general case, let N denote the minimal normal extension of the subnormal operator T. We assume N acts on the Hilbert space $\mathcal{H}_E$. Relative to the decomposition $\mathcal{H}_E = \mathcal{H} \oplus (\mathcal{H}_E \ominus \mathcal{H})$, the operator matrix of N is of the form $\begin{pmatrix} T & * \\ 0 & * \end{pmatrix}$. For $X \in \mathcal{C}_2(\mathcal{H})$, let X_E denote the operator in $\mathcal{C}_2(\mathcal{H}_E)$ whose matrix is of the form $\begin{pmatrix} X & 0 \\ 0 & 0 \end{pmatrix}$; thus $\|X\|_2 = \|X_E\|_2$.

The preceding case for normal operators implies that $\mathcal{D}(N)\|X\|_2 = \mathcal{D}(N)\|X_E\|_2 \geqq \|NX_E - X_EN\|_2 \geqq \|TX - XT\|_2$. Since N is the minimal normal extension of T, it follows from [18, p. 103] that $\mathcal{D}(T) = \mathcal{D}(N)$, and thus $\mathcal{D}(T)\|X\|_2 \geqq \|\delta_2(T)(X)\|_2$. Now $\|\delta_2(T)\| \leqq \mathcal{D}(T)$ and the reverse inequality follows from Corollary 4.8, completing the proof.

THEOREM 4.12. *A subnormal operator $T \in \mathcal{L}(\mathcal{H})$ is S-universal if and only if $\mathcal{D}(T) = 2\mathcal{R}(T)$.*

PROOF. If $\mathcal{D}(T) = 2\mathcal{R}(T)$, the conclusion that T is S-universal follows from Lemma 4.10 (since T is hyponormal). If $\mathcal{D}(T) < 2\mathcal{R}(T)$, then Lemma 4.11 implies that $\|\delta_2(T)\| = \mathcal{D}(T) < 2\mathcal{R}(T) = 2d(T)$, and the proof is complete.

QUESTION 4.13. Is the identity $\|\delta_2(T)\| = \mathcal{D}(T)$ valid for each hyponormal operator T?

An affirmative answer would imply that a hyponormal operator T is S-universal if and only if $\mathcal{D}(T) = 2\mathcal{R}(T)$.

We note that there exist S-universal operators which are not subnormal. It is not difficult to verify that each hyponormal injective unilateral weighted shift satisfies the hypothesis of Lemma 4.10 and is thus S-universal; it is known that not all such shifts are subnormal [27, theorem 4]. If U is a nonunitary isometry, then it follows readily that $T = U \oplus U^*$ satisfies $\mathcal{D}(T) = 2d(T)$; thus T is an S-universal operator that is not hyponormal. We next consider examples with nilpotent operators.

EXAMPLE 4.14. Let $\{e_i\}_{i=1}^n$ denote an orthonormal basis for the n-dimensional Hilbert space $\mathcal{H}_n$. Let $q_n \in \mathcal{L}(\mathcal{H}_n)$ be the nilpotent operator defined by the relations $q_n(e_1) = 0$ and $q_n(e_i) = e_{i-1}$ for $2 \leqq i \leqq n$. Let T_n denote the countably infinite direct sum of copies of q_n ($T_n = q_n \oplus q_n \oplus \cdots$). Now $d(T_n) = 1$ and it follows from [9] that $\operatorname{diam}(W(T_n)) = 2\cos(\pi/(n+1))$; thus $2 \geqq \|\delta_{\mathcal{J}}(T_n)\| \geqq 2\cos(\pi/(n+1))$ for each norm ideal $\mathcal{J}$. In this case, since $\mathcal{D}(T_n) = 0$, the preceding estimate is clearly better than that using Corollary 4.8, and the estimate improves as $n \to \infty$.

Note that if $T = \sum_{n=1}^{\infty} \oplus q_n$, then $2d(T) = \operatorname{diam}(W(T)) = \mathcal{D}(T) = 2$, so T is an example of a direct sum of nilpotents which is S-universal. We know of no example of a non-zero quasinilpotent operator that is S-universal, and we next show that if $T \neq 0$ and $T^2 = 0$, then T is not S-universal. Indeed, from [19, theorem 1], we may assume that the operator matrix of T is of the form $\begin{pmatrix} 0 & M \\ 0 & 0 \end{pmatrix}$; thus $d(T) = \|M\|$. If $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ denotes the matrix of a Hilbert–Schmidt operator X satisfying $\|X\|_2 = 1$, then a calculation shows that

$$\begin{aligned}\|TX - XT\|_2^2 &= \|MC\|_2^2 + \|MD - AM\|_2^2 + \|CM\|_2^2 \\ &\leqq \|M\|^2(2\|C\|_2^2 + (\|D\|_2 + \|A\|_2)^2).\end{aligned}$$

Since $\|A\|_2^2 + \|D\|_2^2 + \|C\|_2^2 \leqq 1$, it follows that

$$2\|C\|_2^2 + (\|D\|_2 + \|A\|_2)^2 = \|C\|_2^2 + (\|C\|_2^2 + \|D\|_2^2 + \|A\|_2^2) + 2\|D\|_2\|A\|_2$$
$$\leqq 1 - \|A\|_2^2 - \|D\|_2^2 + 1 + 2\|D\|_2\|A\|_2$$
$$= 2 - (\|A\|_2 - \|D\|_2)^2$$
$$\leqq 2.$$

Thus $\|\delta_2(T)\| \leqq 2^{1/2}\|M\| < 2d(T) = \|\delta_1(T)\|$, so that T is not S-universal.

EXAMPLE 4.15. As a final example, we consider the Volterra operator V [18, p. 94]. It is known that $W(V)$ is determined by the boundary curves $(1-\cos(t))/t^2 \pm i(t-\sin(t))/t^2$ $(0 \leqq t \leqq 2\pi)$ [18, p. 110]. The rightmost boundary point is $(1/2, 0)$ (corresponding to $t = 0$), and the maximum height occurs at $(2/\pi, 1/\pi)$ (when $t = \pi$). It follows that $\operatorname{diam}(W(V)) < 5/6$. On the other hand, J. Deddens [12] has determined that $d(V) = (w^2 + w^4)^{1/2}$, where $\tan(1/w) = -1/w$ and $\pi/2 \leqq 1/w \leqq \pi$; a computer calculation implies that $0.5495394 \leqq d(V) \leqq 0.5495397$. Thus $\operatorname{diam}(W(V)) < 2d(V)$, so Corollary 4.9 cannot be applied; we do not know whether the Volterra operator is S-universal.

The preceding examples also show that none of Lemma 4.10, Lemma 4.11, or Theorem 4.12 can be extended to arbitrary operators, and they suggest the following questions.

QUESTION 4.16. Does there exist a nonzero quasinilpotent operator T such that $\operatorname{diam}(W(T)) = 2d(T)$?

QUESTION 4.17. Does there exist a nonzero quasinilpotent operator that is S-universal?

The results of this section show that for T in $\mathcal{L}(\mathcal{H})$, $2d(T) \geqq \|\delta_{\mathcal{J}}(T)\| \geqq \operatorname{diam}(W(T)) \geqq \mathcal{D}(T)$ and $2d(T) \geqq 2\mathcal{R}(T) \geqq \mathcal{D}(T)$. There appears to be no simple relationship between $\operatorname{diam}(W(T))$ and $2\mathcal{R}(T)$. If T is a nonzero nilpotent, then $\operatorname{diam}(W(T)) > 2\mathcal{R}(T) = 0$; however, if N is a normal operator whose spectrum is an equilateral triangle, then $2d(N) = 2\mathcal{R}(N) > \mathcal{D}(N) = \operatorname{diam}(W(N))$. Note also that $2d(N) = \|\delta_1(N)\| > \operatorname{diam}(W(N))$, while $2d(N) > \|\delta_2(N)\| = \operatorname{diam}(W(N))$. In this example $\mathcal{D}(N)/\mathcal{R}(N) = 3^{1/2}$, and it is known that $3^{1/2}$ is the smallest value for the ratio $\mathcal{D}(T)/\mathcal{R}(T)$ (if $\sigma(T)$ is not a singleton). (This fact was shown to the author by E. Azoff and also pointed out by the referee.)

REMARKS. (1) After completing this paper, we obtained the following characterization of the case when $\tau(A, B)$ has dense range (cf. section 2). The following are equivalent: (i) $\tau(A, B)$ has dense range; (ii) $\sigma_{re}(A) \cap \sigma_{le}(B) = \emptyset$ and there exists no nonzero trace class operator X such that $BX = XA$; (iii) Given $Y \in \mathcal{L}(\mathcal{H})$ and $\varepsilon > 0$, there exists $X \in \mathcal{L}(\mathcal{H})$ such that $K = AX - XB - Y$ is compact and $\|K\| < \varepsilon$. These results will appear in a forthcoming paper.

(2) The identity for the spectrum of τ_p apparently first appeared in [8]; it also appears in [13]. The inclusions $\sigma_r(\tau) \subset \sigma_r(A) - \sigma_l(B)$, $\sigma_l(\tau) \subset \sigma_l(A) - \sigma_r(B)$ (but without the resolvent formulas of section 3) are contained in [20]. Several authors have studied the norms of derivations on certain C^*-algebras (with identity) [5], [28]; these results appear to have no bearing on the estimation of $\|\tau_{\mathcal{J}}\|$. Other related results are contained in R. E. Harte's paper [Proc. Roy. Irish Acad., **73** (1973), 285–302].

References

1. G. R. Allan, *On one-sided inverses in Banach algebras of holomorphic vector-valued functions*, J. London Math. Soc. **42** (1967), 463–470, MR35#5939.

2. G. R. Allan, *Holomorphic vector-valued functions on a domain of holomorphy*, J. London Math. Soc. **42** (1967), 509–513, MR35#5940.

3. J. Anderson and J. G. Stampfli, *Commutators and compressions*, Israel J. Math. **10** (1971), 433–441.

4. T. Ando, *Distance to the set of scalars*, preprint.

5. C. Apostol and L. Zsido, *Ideals in W^*-algebras and the function η of A. Brown and C. Pearcy*, Rev. Roumaine Math. Pures Appl. **18** (1973), 1151–1170.

6. C. Apostol and K. Clancey, *Generalized inverses and spectral theory*, Trans. Amer. Math. Soc. **215** (1976), 293–300.

7. C. Apostol and K. Clancey, *On generalized resolvents*, Proc. Amer. Math. Soc. **58** (1976), 163–168.

8. A. Brown and C. Pearcy, *Spectra of tensor products of operators*, Proc. Amer. Math. Soc. **17** (1966), 162–169.

9. A. Brown and C. Pearcy, *Operators on Hilbert Space*, Springer-Verlag, to appear.

10. A. Brown and C. Pearcy, *On the spectra of derivations on norm ideals*, submitted to Acta Sci. Math. (Szeged).

11. C. Davis and P. Rosenthal, *Solving linear operator equations*, Canad. J. Math. **26** (1974), 1384–1389.

12. J. Deddens, private communication.

13. M. R. Embry and M. Rosenblum, *Spectra, tensor products, and linear operator equations*, Pacific J. Math. **53** (1974), 95–107.

14. L. A. Fialkow, *A note on the operator $X \to AX - XB$*. Trans. Amer. Math. Soc. **243** (1978), 147–168.

15. P. A. Fillmore, J. G. Stampfli, and J. P. Williams, *On the essential numerical range, the essential spectrum, and a problem of Halmos*, Acta Sci. Math. **33** (1972), 179–192, MR48#896.

16. C. K. Fong, *Some aspects of derivations on $\mathcal{B}(\mathcal{H})$*, Part I, Seminar Notes, University of Toronto, 1978, preprint.

17. I. C. Gohberg and M. G. Krein, *Introduction to the Theory of Linear Nonselfadjoint Operators*, Vol. 18, Translations of Mathematical Monographs, Amer. Math. Soc., 1969.

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

19. P. R. Halmos, *Capacity in Banach algebras*, Indiana Univ. Math. J. **20** (1970/71), 855–863, MR42 # 3569.

20. R. E. Harte, *Spectral mapping theorems on a tensor product*, Bull. Amer. Math. Soc. **79** (1973), 367–372.

21. D. A. Herrero, private communication.

22. G. Lumer and M. Rosenblum, *Linear operator equations*, Proc. Amer. Math. Soc. **10** (1959), 32–41.

23. H. Radjavi and P. Rosenthal, *Matrices for operators and generators of* $\mathcal{B}(\mathcal{H})$, J. London Math. Soc. **2** (1970), 557–560.

24. F. Riesz and B. Sz.-Nagy, *Functional Analysis*, Ungar, New York, 1955.

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

26. R. Schatten, *Norm Ideals of Completely Continuous Operators*, Ergebnisse der Math., Springer-Verlag, 1960.

27. J. G. Stampfli, *Which weighted shifts are subnormal*, Pacific J. Math. **17** (1966), 367–379, MR33#1740.

28. J. G. Stampfli, *The norm of a derivation*, Pacific J. Math. **33** (1970), 737–747.

29. J. G. Stampfli, *Derivations on* $\mathcal{B}(\mathcal{H})$: *The range*, Illinois J. Math. **17** (1973), 518–524.

DEPARTMENT OF MATHEMATICS
WESTERN MICHIGAN UNIVERSITY
KALAMAZOO, MICHIGAN 49008 USA