PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 134, Number 1, Pages 157–165
S 0002-9939(05)07871-8
Article electronically published on August 15, 2005

SELF-COMMUTATOR APPROXIMANTS

P. J. MAHER

(Communicated by Joseph A. Ball)

ABSTRACT. This paper deals with minimizing $\|B - (X^*X - XX^*)\|_p$, where B is fixed, self-adjoint and $B \in \mathcal{C}_p$, and where X varies such that $BX = XB$ and $X^*X - XX^* \in \mathcal{C}_p$, $1 \le p < \infty$. (Here, $\mathcal{C}_p$, $1 \le p < \infty$, denotes the von Neumann-Schatten class and $\|\cdot\|_p$ its norm.) The upshot of this paper is that $\|B-(X^*X-XX^*)\|_p$, $1 \le p < \infty$, is minimized if, and for $1 < p < \infty$ only if, $X^*X - XX^* = 0$, and that the map $X \to \|B-(X^*X-XX^*)\|_p^p$, $1 < p < \infty$, has a critical point at $X = V$ if and only if $V^*V - VV^* = 0$ (with related results for normal B if $p = 1$ or 2).

1. Introduction

This paper is concerned with approximating an operator by a self-commutator $X^*X - XX^*$ of operators. We study minimizing the quantity

$$\|B - (X^*X - XX^*)\|_p, \quad 1 \le p < \infty,$$

for fixed B in $\mathcal{C}_p$ and varying X such that $X^*X - XX^* \in \mathcal{C}_p$. (Here $\mathcal{C}_p$ denotes the von Neumann-Schatten class with norm $\|\cdot\|_p$, where $1 \le p < \infty$.)

The related topic of approximation by commutators $AX - XA$, which has attracted much interest, has its roots in quantum theory. The Heisenberg Uncertainty Principle may be mathematically formulated as saying that there exists a pair A, X of linear transformations and a (non-zero) scalar α for which

$$AX - XA = \alpha I. \tag{1.1}$$

Clearly, (1.1) cannot hold for square matrices A and X. (To see this, just take the trace of both sides of (1.1).) Nor can (1.1) hold for bounded linear operators A and X: two beautiful proofs of this are due to Wielandt [19] and Wintner [20]. This prompts the question: how close can $AX - XA$ be to the identity?

Halmos [9], [11, Problem 233] proved that if A is normal, or if A commutes with $AX - XA$, then, for all X in $\mathcal{L}(H)$,

$$\|I - (AX - XA)\| \ge \|I\|. \tag{1.2}$$

Anderson [2, Theorem 1.7] generalized Halmos' inequality (1.2): he proved that if A is normal and commutes with B, then, for all X in $\mathcal{L}(H)$,

$$\|B - (AX - XA)\| \ge \|B\|. \tag{1.3}$$

Received by the editors March 5, 2003 and, in revised form, March 25, 2004.
2000 *Mathematics Subject Classification.* Primary 47B20, 47A30; Secondary 47B10.
Key words and phrases. Self-commutator, von Neumann-Schatten class.

Maher [15, Theorem 3.2] obtained the $\mathcal{C}_p$ variant of Anderson's result: he proved that if the normal operator A commutes with B and if $B \in \mathcal{C}_p$, then, for all X such that $AX - XA \in \mathcal{C}_p$,

$$\|B - (AX - XA)\|_p \geq \|B\|_p, \quad 1 \leq p < \infty, \tag{1.4}$$

with equality in (1.4) if, and for $1 < p < \infty$ only if, $AX - XA = 0$.

Using the technique of [15], Bouli and Cherki [4] and Mercheri [16] studied approximation by "generalized commutators" $AX - XC$. They proved [4, Theorem 2.2], [16, Theorem 3.7] that if $B \in \mathcal{C}_p$, if $AB = BC$ and if the pair (A, C) has the Putnam-Fuglede $\mathcal{C}_p$ property (meaning that $AX = XC \Rightarrow A^*X = XC^*$ if $X \in \mathcal{C}_p$), then, for all X such that $AX - XC \in \mathcal{C}_p$,

$$\|B - (AX - XC)\|_p \geq \|B\|_p, \quad 1 \leq p < \infty, \tag{1.5}$$

with equality if, and for $1 < p < \infty$ only if, $AX - XC = 0$. (Other related work includes that of Berens and Finzel [3], Duggal [5], [6], Kittaneh [12], [13] and Mercheri [17].)

In the above results (1.2), (1.3) and (1.4), the zero commutator is, to use Halmos' terminology [10], a commutator approximant in $\mathcal{C}_p$ of B (and similarly, cf. (1.5) for generalised commutators). Here, we frame hypotheses so that the zero self-commutator, likewise, minimizes the quantity $\|B - (X^*X - XX^*)\|_p$. Since a global minimizer is a critical point, we study the local behaviour of the map

$$F_p \colon X \to \|B - (X^*X - XX^*)\|_p^p.$$

We consider the critical points of F_p (that is, $\{V :$ the Fréchet derivative $D_V F_p = 0\}$). We show in Theorem 3.1 that if a critical point (hence, in particular, a global minimizer) V of F_p satisfies $V^*V - VV^* = 0$ for self-adjoint B, then $BV = VB$. In Theorem 3.2 we classify the critical points of F_p. Theorem 3.2 says that for self-adjoint B such that $BX = XB$, the point V is a critical point of F_p if and only if $V^*V - VV^* = 0$.

It follows from Theorem 3.2 that **every** global minimizer X of F_p satisfies $X^*X - XX^* = 0$ for self-adjoint B commuting with X. The global result guarantees the **existence** of global minima. Thus, it says that under the same hypotheses ($BX = XB$, $B^* = B \in \mathcal{C}_p$) for all X such that $X^*X - XX^* \in \mathcal{C}_p$, where $1 < p < \infty$,

$$\|B - (X^*X - XX^*)\|_p \geq \|B\|_p \tag{1.6}$$

with equality in (1.6) if and only if $X^*X - XX^* = 0$.

There is a similar inequality for the trace norm ($p = 1$) proved by different arguments in Theorem 4.2. Examples 4.1, 4.2 and 4.3 illustrate, and reinforce, the results. Examples 4.1 and 4.2 show that if the (seemingly restrictive) condition $BX = XB$ is dropped, the conclusions of Theorems 4.1 and 4.2 do not hold. Finally, Example 4.3 shows that for $0 < p < 1$ the inequalities may be reversed.

2. Preliminaries

Let H denote a separable, complex Hilbert space. For details concerning the von Neumann-Schatten classes $\mathcal{C}_p$ and norms $\|\cdot\|_p$ see [7, Chapter XI], [18, Chapter 2]. The spaces $\mathcal{C}_p$ are examples of 2-sided, self-adjoint ideals. Note $\mathcal{C}_p \subseteq \mathcal{C}_q$ and $\|\cdot\|_p \geq \|\cdot\|_q$ if $1 \leq p \leq q < \infty$. The Fréchet derivative of some real-valued function

F at V is denoted by $D_V F$ and given by

$$(D_V F)(S) = \lim_{h\to 0} \frac{F(V + hS) - F(V)}{h} \qquad (2.1)$$

(provided the R.H. limit exists). We state below the Aiken, Erdos and Goldstein differentiation result.

Theorem 2.1 ([1, Theorem 2.1])**.** *Let the map $\Phi\colon \mathcal{C}_p \to \mathbb{R}^+$ be given by $\Phi\colon X \to \|X\|_p^p$. Then:*

(a) *for $1 < p < \infty$, the map Φ is Fréchet differentiable at every X in $\mathcal{C}_p$ with derivative $D_X\Phi$ given by*

$$(D_X\Phi)(S) = p\,\Re\,\tau[|X|^{p-1}U^*S],$$

where τ denotes trace, $X = U|X|$ is the polar decomposition of X and $S \in \mathcal{C}_p$;

(b) *for $0 < p \leq 1$, provided $\dim H < \infty$, the same result holds at every invertible element X.*

Observe that the formula for the R.H.S. of $(D_X\Phi)(S)$ makes sense: for $|X|^{p-1}U^* \in \mathcal{C}_1$ since, by [18, Lemma 2.3.1], $X \in \mathcal{C}_p \Leftrightarrow |X| \in \mathcal{C}_p \Leftrightarrow |X|^p \in \mathcal{C}_1$, that is, $|X|^p = (|X|^{p-1}U^*)(U|X|) \in \mathcal{C}_1$, whence $|X|^{p-1}U^* \in \mathcal{C}_1$ as $X = U|X| \in \mathcal{C}_p \supseteq \mathcal{C}_1$.

3. Local theory

The local theory of self-commutator approximation is more complicated than the local theory of commutator approximation [15, Theorem 3.2 (b)]: inevitably so, since differentiating a (non-commutative) product is more complicated than differentiating a sum. The local theory centres on Theorem 3.1.

Theorem 3.1. *Let B be self-adjoint in $\mathcal{C}_p$, where $1 < p < \infty$. Let $\mathfrak{S} = \{X : X^*X - XX^* \in \mathcal{C}_p\}$ and let $F_p\colon \mathfrak{S} \to \mathbb{R}^+$ be given by*

$$F_p\colon X \to \|B - (X^*X - XX^*)\|_p^p.$$

Then if V is a critical point of F_p such that $V^*V - VV^* = 0$ it follows that $BV = VB$.

Proof. As we shall use this proof in that of Theorem 3.2 we adopt the hypothesis that $V^*V - VV^* = 0$ only at the last step.

Step 1. Let V be in $\mathfrak{S}$ so that $B - (V^*V - VV^*) \in \mathcal{C}_p$. (Observe that the set $\mathfrak{S} = \{X : X^*X - XX^* \in \mathcal{C}_p\}$ properly contains $\mathcal{C}_p$, for if $X \in \mathcal{C}_p$ then $X \in \mathfrak{S}$ and, e.g., $I \in \mathfrak{S}$ but $I \notin \mathcal{C}_p$.) Let $\mathcal{S}$ consist of all operators S for which $B - [(V + S)^*(V + S) - (V + S)(V + S)^*] \in \mathcal{C}_p$. (Thus, $\mathcal{S}$ also properly contains $\mathcal{C}_p$.) Let $\Phi\colon X \to \|X\|_p^p$ and $\Psi\colon X \to B - (X^*X - XX^*)$. Then $F_p = \Phi \circ \Psi$. Let S be arbitrary in $\mathcal{S}$. By considering $F_p(V + S) - F_p(V)$ it follows from the definition (2.1) of the derivative that the Fréchet derivative of F_p at V is given by

$$(D_V F_p)(S) = (D_{B-(V^*V-VV^*)}\Phi)(VS^* + SV^* - V^*S - S^*V).$$

Let $B - (V^*V - VV^*) = U_1|B - (V^*V - VV^*)|$ be the polar decomposition of $B-(V^*V-VV^*)$ (so that $\operatorname{Ker} U_1 = \operatorname{Ker}|B-(V^*V-VV^*)|$). Then by Theorem 2.1, on writing

$$Y = U_1|B - (V^*V - VV^*)|^{p-1},$$

we have

$$(D_V F_p)(S) = p\,\Re\,\tau[Y^*(VS^* + SV^* - V^*S - S^*V)]$$

for all operators S in $\mathcal{S}$. Note that Y^* $(= |B - (V^*V - VV^*)|^{p-1}U_1^*) \in \mathcal{C}_1$ (cf. comments after the statement of Theorem 2.1). Therefore, as $\Re\tau(T) = \Re\tau(T^*)$ for all T in $\mathcal{C}_1$, we have $\Re\tau[Y^*VS^* - Y^*S^*V] = \Re\tau[SV^*Y - V^*SY]$. Hence, by the invariance of trace [18, Theorem 2.2.4],

$$(D_V)(F_p)(S) = p\,\Re\,\tau[(V^*Y - YV^* + V^*Y^* - Y^*V^*)S]. \tag{3.1}$$

Step 2. Let V be a critical point of F_p so that $(D_V F_p)(S) = 0$ for all operators S in $\mathcal{S}$. Take $S = f \otimes g$, where f and g are arbitrary vectors in the underlying Hilbert space H. (The rank one operator $x \to \langle x, f\rangle g$, where $x \in H$, is denoted $f \otimes g$. Note that $\tau[T(f \otimes g)] = \langle Tg, f\rangle$ for T in $\mathcal{L}(H)$; cf. [18, pp. 73, 90].) Then by (3.1)

$$\Re\langle (V^*Y - YV^* + V^*Y^* - Y^*V^*)g, f\rangle = 0$$

which, since f and g are arbitrary, means that $V^*Y - YV^* + V^*Y^* - Y^*V^* = 0$, that is,

$$(\Re Y)V = V(\Re Y). \tag{3.2}$$

Step 3. Suppose now that B is self-adjoint. Then $B - (V^*V - VV^*)$ $(= U_1|B - (V^*V - VV^*)|)$ is self-adjoint . Hence U_1 is self-adjoint and commutes with $|B - (V^*V - VV^*)|$, and hence Y $(= U_1|B - (V^*V - VV^*)|^{p-1})$ is self-adjoint. Therefore, (3.2) says that

$$YV = VY, \tag{3.3}$$

that is,

$$U_1|B - (V^*V - VV^*)|^{p-1}V = VU_1|B - (V^*V - VV^*)|^{p-1}. \tag{3.4}$$

Step 4. Assertion: V satisfies

$$BV - (V^*V - VV^*)V = VB - V(V^*V - VV^*). \tag{3.5}$$

To prove this assertion, note that equality (3.5) is equivalent to

$$U_1|B - (V^*V - VV^*)|V = VU_1|B - (V^*V - VV^*)|. \tag{3.6}$$

Write $Z = |B - (V^*V - VV^*)|^{p-1}$. Then (3.6) says that

$$U_1 Z^{\frac{1}{p-1}} V = VU_1 Z^{\frac{1}{p-1}}. \tag{3.7}$$

To prove (3.7) we approximate both sides of (3.7) by polynomials in Z. The function $f\colon t \to t^{1/(p-1)}$, where $t \in \sigma(Z) \subseteq \mathbb{R}^+$, can be approximated uniformly by a sequence (p_i) of polynomials without constant term (for $f(0) = 0$). Therefore, (3.7) will follow by the functional calculus (cf. [15, p. 998]) from $U_1 p_i(Z)V = VU_1 p_i(Z)$ and this, in turn, will follow from

$$U_1 Z^n V = VU_1 Z^n. \tag{3.8}$$

To prove (3.8) note that in the polar decomposition of $B - (V^*V - VV^*)$, $\operatorname{Ker} U_1 = \operatorname{Ker}|B - (V^*V - VV^*)| = \operatorname{Ker} Z$ (by the spectral theorem). Hence, $(\operatorname{Ker} U_1)^\perp = \overline{\operatorname{Ran} Z}$ and $U_1^*U_1$, the orthogonal projection onto $(\operatorname{Ker} U_1)^\perp$, satisfies $ZU_1^*U_1Z = Z^2$. (It is simplest to write, where necessary, U_1^*, even though U_1 is self-adjoint.) Since $Y = U_1 Z$, then (3.3) says that $U_1ZV = VU_1Z$ and, as $Y = Y^*$, $ZU_1^*V = VZU_1^*$. Thus,

$$Z^2V = ZU_1^*(U_1ZV) = (ZU_1^*V)U_1Z = VZU_1^*U_1Z = VZ^2.$$

Taking positive square roots of Z^2 [18, Theorem 1.7.7 (vi)] we get $VZ = ZV$. Returning now to (3.8): for $n = 1$, (3.8) is the equality $U_1ZV = VU_1Z$ (which is (3.3)), and the inductive step follows from $ZV = VZ$. This proves the assertion.

Step 5. If, finally, $V^*V - VV^* = 0$ then (3.5) forces $BV = VB$. $\square$

Note. The proof in Theorem 3.1 of the implication, V is a critical point of F_p $\Rightarrow BV - (V^*V - VV^*)V = VB - V(V^*V - VV^*)$, only holds for $1 < p < \infty$ since the argument involving the function $f\colon t \to t^{1/(p-1)}$, where $0 \leq t < \infty$, only holds for $1 < p < \infty$.

Observe also that for non-self-adjoint B, in the case $p = 2$, it follows from equality (3.2) of the proof of Theorem 3.1 that if V is a critical point of F_2 such that $V^*V - VV^* = 0$, then $(\Re B)V = V(\Re B)$. But this latter equality does not force $BV = VB$ even if B is normal: witness: $B = \left[\begin{smallmatrix} 1 & i \\ i & 1 \end{smallmatrix}\right]$.

Theorem 3.2. *Let B be self-adjoint, let $BX = XB$ and let B be in $\mathcal{C}_p$. Let $\mathfrak{S} = \{X : X^*X - XX^* \in \mathcal{C}_p\}$ and let $F_p\colon \mathfrak{S} \to \mathbb{R}^+$ be given by*

$$F_p\colon X \to \|B - (X^*X - XX^*)\|_p^p.$$

Then:

(a) *for $1 < p < \infty$, the map F_p has a critical point at V if and only if $V^*V - VV^* = 0$;*
(b) *for $0 < p \leq 1$, the map F_p has a critical point at V if $V^*V - VV^* = 0$ provided $\dim H < \infty$ and $B - (V^*V - VV^*)$ is invertible;*
(c) *for $p = 2$, the same result as in* (a) *holds if the condition on B of self-adjointness is replaced by normality.*

Proof. (a) Let V be a critical point of F_p. Then equality (3.5) of the proof of Theorem 3.1 holds. As $BV = VB$ then $(V^*V - VV^*)V = V(V^*V - VV^*)$, whence, by Kleinecke-Shirokov [11, Problem 232], $V^*V - VV^*$ is quasinilpotent and, hence, being self-adjoint, zero.

Conversely, let V satisfy $V^*V - VV^* = 0$. Then the partial isometries U_1 and, say, U, occurring in the polar decompositions of $B - (V^*V - VV^*)$ and of B, coincide. Thus, $Y = U|B|^{p-1} \in \mathcal{C}_1$ so that $Y^* = |B|^{p-1}U^* \in \mathcal{C}_1$.

We first prove that $Y^*V - VY^* = 0$. Since V and V^* commute with B they commute with $|B|$ (and hence with $|B|^{p-1}$). So, $|B|U^*V = |B|VU^*$. It follows that

$$\operatorname{Ran}(U^*V - VU^*) \subseteq \operatorname{Ker}|B| = \operatorname{Ker}|B|^{p-1}.$$

Hence, since $|B|^{p-1}V = V|B|^{p-1}$, therefore $Y^*V - VY^* = 0$.

Similarly, from the equality $|B|U^*V^* = |B|V^*U^*$ it follows that $Y^*V^* - V^*Y^* = 0$. Hence, $V^*Y - YV^* + V^*Y^* - Y^*V^* = 0$. Substitute into equality (3.1) of the proof of Theorem 3.1 (the expression for D_VF_p, the Fréchet derivative of F_p at V). As $YS \in \mathcal{C}_1$ and $Y^*S \in \mathcal{C}_1$, it follows by (3.1) that $(D_VF_p)(S) = 0$ for all S in $\mathcal{L}(H)$.

(b) follows immediately from (a) as in [15, Theorem 3.2 (c)].

(c) Let B be normal. If V is a critical point of F_2, then equality (3.2) of the proof of Theorem 3.1 says that

$$[\Re B - (V^*V - VV^*)]V = V[\Re B - (V^*V - VV^*)].$$

Since V commutes with B then (by Fuglede) V commutes with B^* and hence with $\Re B$. The result now follows as in (a).

The proof of the converse implication (V satisfies $V^*V - VV^* = 0 \Rightarrow V$ is a critical point of F_2) depends only on V and V^* commuting with B and is therefore the same as in (a). □

Indeed, the proof in (a) of the implication, $V^*V - VV^* = 0 \Rightarrow V$ is a critical point of F_p, for $1 < p < \infty$, holds (via Fuglede's Theorem) for normal B.

4. Global theory

Theorem 4.1. *Let B be self-adjoint, let $BX = XB$ and let B be in $\mathcal{C}_p$. Let $\mathfrak{S} = \{X : X^*X - XX^* \in \mathcal{C}_p\}$. Then, if $X \in \mathfrak{S}$,*

(a) *for $1 < p < \infty$,*

$$\|B - (X^*X - XX^*)\|_p \geq \|B\|_p \tag{4.1}$$

with equality holding in (4.1) *if and only if $X^*X - XX^* = 0$;*

(b) *for $p = 2$, the same result as in* (a) *holds if B is assumed normal rather than self-adjoint.*

Proof. (a) First, suppose the operators X in $\mathfrak{S}$ are contractions, i.e., such that $\|X\| \leq 1$. Suppose also that the underlying space H is finite dimensional. (The argument here is analogous to [14, Theorem 5.7].) The set of contractions is bounded and closed (for the condition $X^*X - I \leq 0$ characterises the contractions, and the map $X \to X^*X$ is continuous; cf. [11, Problem 129]). Hence, $\mathfrak{S}$ is compact since H is finite dimensional. Therefore, the continuous map $F_p\colon X \to \|B-(X^*X-XX^*)\|_p^p$ is bounded, attains its bounds and thus has a global minimizer, and hence a critical point, at V, say. Since, by Theorem 3.2(a), $V^*V - VV^* = 0$, therefore

$$\|B - (X^*X - XX^*)\|_p \geq \|B\|_p. \tag{4.2}$$

Conversely, if equality holds in (4.2) for some point X, then that X is a global minimizer, hence a critical point of F_p, whence, by Theorem 3.2(a), $X^*X - XX^* = 0$.

The extension to infinite-dimensional H is similar to [1, Theorem 3.5]. As the operator B is compact and normal there exists a basis $\{\phi_i\}$ of H consisting of eigenvectors of B which may be ordered such that $|\lambda_1| \geq |\lambda_2| \geq \ldots$ where $B\phi_i = \lambda_i\phi_i$ (and where the eigenvalues are repeated according to multiplicity). Let

$$H_k = \mathrm{Span}\{\phi_i : B\phi_i = \lambda_i\phi_i, i = 1, \ldots, k\}.$$

H_k is invariant under X and X^*; for if ϕ_i is an eigenvector of B, then so are $X\phi_i$ and $X^*\phi_i$ with the same eigenvalues (since B commutes with X and X^*). Therefore, if E_k denotes the orthogonal projection onto H_k, then $E_kX = XE_k$. Hence E_kBE_k commutes with E_kXE_k (and with $E_kX^*E_k$) and hence, by the finite-dimensional inequality (4.2) applied to the contraction E_kXE_k,

$$\|(E_kBE_k) - [(E_kXE_k)^*(E_kXE_k) - (E_kXE_k)(E_kXE_k)^*]\|_p \geq \|E_kBE_k\|_p,$$

that is, $\|E_k[B-(X^*X-XX^*)]E_k\|_p \geq \|E_kBE_k\|_p$. Now let $k \to \infty$. Then $E_k \to I$ and from [8, Lemma 2] (cf. [1, Theorem 3.5]), it follows that inequality (4.2) holds for infinite-dimensional H.

The condition that the operator X in $\mathfrak{S}$ is a contraction may now be lifted. Let X be arbitrary in $\mathcal{S}$; then by applying the inequality (4.2) to the contraction $X/\|X\|$, the result immediately follows.

(b) follows similarly to (a) from the corresponding local result, Theorem 3.2(c). □

Theorem 4.2. *Let B be normal, let $BX = XB$ and let B be in $\mathcal{C}_1$. Then, if $X^*X - XX^* \in \mathcal{C}_1$,*

$$\|B - (X^*X - XX^*)\|_1 \geq \|B\|_1.$$

Proof. Let $B = U|B|$ be the polar decomposition of B. As U is a partial isometry, so is U^*, and so $\|U^*\| = 1$. Since, by [18, Theorem 2.3.10], $\|U^*T\|_1 \leq \|U^*\|\|T\|_1 = \|T\|_1$ for arbitrary T in $\mathcal{C}_1$. Then, by [18, Lemma 2.3.3],

$$\begin{aligned} \|B - (X^*X - XX^*)\|_1 &\geq \||B| - U^*(X^*X - XX^*)\|_1 \\ &\geq |\tau[|B| - U^*(X^*X - XX^*)]|, \end{aligned} \tag{4.3}$$

where

$$\tau[|B| - U^*(X^*X - XX^*)] = \sum_n \langle [|B| - U^*(X^*X - XX^*)]\phi_n, \phi_n\rangle$$

for an arbitrary orthonormal basis $\{\phi_i\}$ of H.

Take $\{\phi_i\}$ as the orthonormal basis of H consisting of eigenvectors of the compact normal operator $|B|$. Let $\{\psi_m\}$ be an orthonormal basis of $\operatorname{Ker}|B|$ and let $\{\xi_k\}$ be an orthonormal basis of $(\operatorname{Ker}|B|)^\perp$ consisting of eigenvectors of $|B|$. Thus, $\{\phi_n\} = \{\psi_m\} \cup \{\xi_k\}$. Then $\sum_m \langle [|B| - U^*(X^*X - XX^*)]\psi_m, \psi_m\rangle = 0$ because $\psi_m \in \operatorname{Ker}|B| = \operatorname{Ker} U$; and $\sum_k \langle |B|\xi_k, \xi_k\rangle = \|B\|_1$. Further, since $BX = XB$ and $BX^* = X^*B$, it can be checked that $\langle U^*X^*X\xi_k, \xi_k\rangle = \langle X^*U^*X\xi_k, \xi_k\rangle$. Hence, by the invariance of trace [18, Theorem 2.2.4 (v)], $\tau[U^*(X^*X - XX^*)] = 0$. Therefore, by (4.3),

$$\|B - (X^*X - XX^*)\|_1 \geq \tau(|B|) = \|B\|_1.$$

□

In the special case when B is positive, the proof of the trace norm result is simple and does not require the commutativity condition.

Theorem 4.3. *Let B be positive and be in $\mathcal{C}_1$. Then, if $X^*X - XX^* \in \mathcal{C}_1$,*

$$\|B - (X^*X - XX^*)\|_1 \geq \|B\|_1.$$

Proof. Let B be positive so that $B = |B|$. Then, by [18, Lemma 2.3.3] and the linearity of trace [18, Theorem 2.2.4],

$$\begin{aligned} \|B - (X^*X - XX^*)\|_1 &\geq |\tau[B - (X^*X - XX^*)]| \\ &= |\tau[B]| = \tau[|B|] = \|B\|_1. \end{aligned}$$

□

If B and X do not commute, then either $\|B - (X^*X - XX^*)\|_p \geq \|B\|_p$, for $1 \leq p < \infty$, is reversed (Example 4.1) or $\|B - (X^*X - XX^*)\|_p = \|B\|_p$ without $X^*X - XX^* = 0$ (Example 4.2).

Example 4.1. Take $B = \left[\begin{smallmatrix} 3 & 0 \\ 0 & -3 \end{smallmatrix}\right]$ and $X = \left[\begin{smallmatrix} 1 & 1 \\ 2 & 2 \end{smallmatrix}\right]$ so that $B = B^*$ and $BX \neq XB$. For $1 \leq p < \infty$, as $\|T\|_p^p = \sum_i s_i^p(T)$, where $s_i(T)$ denotes the ith eigenvalue of $|T|$, we get, for $1 \leq p < \infty$,

$$\|B - (X^*X - XX^*)\|_p^p = 1^p + 1^p < 3^p + 3^p = \|B\|_p^p.$$

Example 4.2. Take $X = f \otimes g$ and $B = f \otimes f$, where $f \neq g$ and $\|f\| = \|g\| = 1$, so that $B = B^*(\geq 0)$ and $BX \neq XB$. Then $X^*X - XX^* = f \otimes f - g \otimes g \neq 0$ and, as $\|f \otimes g\|_p = \|f\|\|g\|$ for $1 \leq p < \infty$ [18, p. 90], we have

$$\|B - (X^*X - XX^*)\|_p = \|g \otimes g\|_p (= 1) = \|f \otimes f\|_p = \|B\|_p.$$

Finally, the inequality $\|B - (X^*X - XX^*)\|_p \geq \|B\|_p$ may be reversed for $0 < p < 1$.

Example 4.3. Take $B = \left[\begin{smallmatrix} 3 & 0 \\ 0 & 3 \end{smallmatrix}\right] (\geq 0)$ and $X = \left[\begin{smallmatrix} 3 & \sqrt{6} \\ -3 & 3 \end{smallmatrix}\right]$ so that $B = B^*$ and $BX = XB$. Then for $0 < p < 1$ we have the strict inequality

$$\|B - (X^*X - XX^*)\|_p^p = 6^p < 2 \cdot 3^p = \|B\|_p^p.$$

(This example also shows that even if the conditions of Theorem 4.2 are met, a minimizer of $\|B - (X^*X - XX^*)\|_1$ need not be normal: for here

$$\|B - (X^*X - XX^*)\|_p (= 6) = \|B\|_1,$$

yet $X^*X - XX^* \neq 0$.)

References

[1] J. G. Aiken, J. A. Erdos, and J. A. Goldstein, *Unitary approximation of positive operators*, Illinois J. Math **24** (1980), 61–72. MR0550652 (81a:47026)

[2] J. Anderson, *On normal derivatives*, Proc. Amer. Math. Soc. **38** (1973), 135–140. MR0312313 (47:875)

[3] H. Berens and M. Finzel, *A problem in linear matrix approximation*, Math. Nachr. **175** (1995), 33–46. MR1355011 (96i:47030)

[4] S. Bouali and S. Cherki, *Approximation by generalized commutators*, Acta Sci. Math. (Szeged) **63** (1997), 273–278. MR1459791 (98j:47077)

[5] B. P Duggal, *A remark on normal derivations*, Proc. Amer. Math. Soc. **126** (1998), 2047–2052. MR1451795 (98h:47050)

[6] B. P. Duggal, *On the range, kernel orthogonality of derivations*, Lin. Alg. Appl. **304** (2000), 1–3, 103–108. MR1734207 (2000k:47042)

[7] N. Dunford and J. T. Schwarz, *Linear operators, part II*, Interscience, New York, 1964.

[8] J. A. Erdos, *On the trace of a trace class operator*, Bull. Lond. Math. Soc. **6** (1974), 47–50. MR0370246 (51:6473)

[9] P. R. Halmos, *Commutators of operators, II*, Amer. J. Math. **76** (1954), 191–198. MR0059484 (15:538d)

[10] ———, *Positive approximants of operators*, Indiana Univ. Math. J. **21** (1972), 951–960. MR0291829 (45:919)

[11] ———, *A Hilbert space problem book*, second ed., Springer-Verlag, New York, 1982. MR0675952 (84e:47001)

[12] F. Kittaneh, *Normal derivations in norm ideals*, Proc. Amer. Math. Soc. **12** (1995), 1779–1785. MR1242091 (95g:47054)

[13] ———, *Operators that are orthogonal to the range of a derivation*, J. Math. Anal. Appl. **203** (1997), 868–873. MR1417136 (97f:47033)

[14] P. J. Maher, *Partially isometric approximation of positive operators*, Illinois J. Math. **33** (1989), 227–243. MR0987820 (90c:47038)

[15] ———, *Commutator approximants*, Proc. Amer. Math. Soc. **115** (1992), 995–1000. MR1086335 (92j:47059)

[16] S. Mercheri, *On minimizing $\|S - (AX - XB)\|_p^p$*, Serdica Math. J. **26** (2000), 119–126. MR1794930 (2001j:47033)

[17] ———, *Another version of Maher's inequality*, submitted, 2005.

[18] J. R. Ringrose, *Compact non-self-adjoint operators*, Van Nostrand Rheinhold, London, 1971.

[19] M. Wielandt, *Über die ünbeschränktheit der Operatoren des Quantenmechanik*, Math. Ann. **121** (1949), 21. MR0030701 (11:38g)
[20] A. Wintner, *The unboundedness of quantum-mechanical matrices*, Phys. Rev. **71** (1947), 738–739. MR0020724 (8:589c)

Department of Mathematics, Middlesex University, The Burroughs, London NW4 4BT, United Kingdom
E-mail address: p.maher@mdx.ac.uk