COMMUTATOR AND SELF-COMMUTATOR APPROXIMANTS II

P. J. Maher

Abstract

We minimize the quantities (i) \( \| T - (AX -XA) \| \), (ii) \( \| T - (X^*X - XX^*) \| \) and (iii) \( \| T - (AX -XB) \| \) where \( T \) is isometric and where in (i) \( A \) is paranormal and commutes with \( T \), in (ii) \( X^* \) (or \( X \)) is paranormal and commutes with \( T \), and in (iii) \( A \) and \( B \) are paranormal and \( AT = TB \) and \( TA = BT \). The upshot is that these quantities are minimized when \( 0 = AX -XA = X^*X - XX^* = AX -XB \). To prove these results we obtain the power norm equality for paranormal operators: if \( A \) is paranormal then \( \| A^n \| = \| A \|^n \) if \( n \in \mathbb{N} \).

1 Introduction

As in [11] and [12] we approximate an operator by a commutator \( AX -XA \) of operators, by a self-commutator \( X^*X - XX^* \) and, as in [4], by a "generalized commutator" \( AX -XB \). There, in [4], [11] and [12] the approximation is in the von Neumann-Schatten norm \( \| \cdot \|_p \), where \( 1 \leq p < \infty \), on the von Neumann-Schatten classes \( C_p \); here, the approximation is in the sup norm on \( L(H) \) (For operators on Banach space, see the recent paper by Duggal [6]).

The pertinent concept is that of paranormality which, as is well known from [9], is a strong generalization of hyponormality.

For self-commutator approximation with paranormal \( X \) we have to restrict ourselves to the sup norm. Consider the more transparent hyponormal special case of this: that is, approximation by a self-commutator \( X^*X -XX^* \) for hyponormal \( X \); that is, approximation by a positive self-commutator. This topic may be regarded as an obvious extension of [12] since there can be no question of minimizing \( \| T - (X^*X - XX^*) \|_p \) where \( X^*X -XX^* \) is compact and \( X \) is hyponormal. For if \( X^*X -XX^* \) is compact then \( X^*XP = (X^*X -XX^*)P \) is compact where \( P \) is the orthogonal projection onto \( \text{Ker}X^* \) (that is, \( I - P \) is the orthogonal projection onto \( \text{Ker}X^* \)).


Key words and Phrases. Commutator, self-commutator.

Received: December 1, 2009

Communicated by Dragan S. Djordjević
being hyponormal, is normal [10, Problem 206]. Thus, if \(X^*X - XX^*\) is compact for hyponormal \(X\) then \(X^*X - XX^* = 0\). The same result holds if, more generally, \(X\) is paranormal; for, as is proved in Theorem 2.2 below, a compact paranormal operator is normal.

Another property the paranormal operators share with the hyponormal ones is the power norm equality: if \(A\) is paranormal then \(\|A^n\| = \|A\|^n\) if \(n \in \mathbb{N}\) as is proved in Theorem 2.1.

We use the power norm equality to obtain the approximation results here (Theorems 3.1, 3.2 and 3.3). Theorem 3.1 says that if \(A\) is a paranormal operator commuting with the isometry \(T\) then \(\|T - (AX -XA)\| \geq T\). Theorem 3.2 gives a similar result about minimizing \(\|T - (X^*X - XX^*)\|\) for (a) paranormal \(X^*\) commuting with the isometry \(T\) and for (b) paranormal \(X\) commuting with \(T\); Example 3.1 shows that this commutativity assumption is necessary. From Theorem 3.1 we obtain - via operator matrices - Theorem 3.3, a result about minimizing \(\|T - (AX -XB)\|\) for paranormal \(A\) and \(B\). Those minimization results are interpreted geometrically in Corollaries 3.1 and 3.4.

# 2 Paranormality

**Definition 2.1.** An operator \(A\) in \(L(H)\) is paranormal if
\[
(A^*)^2 - 2\lambda A^*A + \lambda^2 \geq 0
\]
for all real \(\lambda \geq 0\).

With the definition of positivity in \(L(H)\) and the discriminant criterion for the quadratic in \(\lambda\), Definition 2.1 is easily proved to be equivalent to Definition 2.2 [3, Theorem 4].

**Definition 2.2.** An operator \(A\) in \(L(H)\) is paranormal if
\[
\|Af\|^2 \leq \|A^2f\|\|f\|
\]
for all \(f\) in \(H\).

The class of paranormal operators strictly contains many other classes of operators including the hyponormal operators [5, §3], [7, §1], [9]. Paranormal operators share with the hyponormal ones the following properties given in Theorems 2.1 and 2.2 below.

**Theorem 2.1** (Power norm equality). If \(A\) is paranormal, \(\|A^n\| = \|A\|^n\) where \(n \in \mathbb{N}\).

**Proof.** Equality is trivial for \(n = 1\). Proceed by induction. Now, by Definition 2.2
\[
\|A^n f\|^2 = \|A(A^{n-1}f)\|^2 = \|Ag\|^2, \text{ say}
\]
\[
\leq \|A^2 g\|\|g\|
\]
\[
= \|A^{n+1}f\|\|A^{n-1}f\|
\]
\[
\leq \|A^{n+1}\|\|A^{n-1}\|\|f\|
\]
for all $f$ in $H$. Therefore, using the induction hypothesis $\|A^k\| = \|A\|^k$ for $1 \leq k \leq n$, we get

$$\|A^n\|^2 = \|A\|^{2n} \leq \|A^{n+1}\| \|A\|^{n-1}$$

whence $\|A\|^{n+1} \leq \|A^{n+1}\|$. Since $\|A^{n+1}\| \leq \|A\|^{n+1}$ automatically the inductive step follows.

**Theorem 2.2.** A compact paranormal operator is normal.

**Proof.** Since the paranormal operator $A$, say, is compact it has a countable spectrum (cf. [13, Theorem 1.8.2]). Isolated points of the spectrum are poles, hence eigenvalues; further, the eigenspaces corresponding to these eigenvalues are mutually orthogonal. So if one generates the space corresponding to these eigenvalues one obtains a diagonal operator. What is left is at best the limit point which is the limit point of the diagonal entries. Conclusion: $A$ is normal.

## 3 Approximation results

The proofs of the approximation results below use Theorem 2.1 and hinge on the following identity: if $AT = TA$ then

$$nTA^{n-1} = A^nB - BA^n + \sum_{i=0}^{n-1} A^{n-i-1}(T - (AB - BA))A^i$$

(3.1)

for all $B$ in $L(H)$.

The next result is a variant of the well-known result of Anderson [1, Theorem 1.7] on minimizing $\|T - (AX - XA)\|$ for normal $A$.

**Theorem 3.1.** If $A$ is paranormal and $T$ is an isometry such that $AT = TA$ then

$$\|T - (AX - XA)\| \geq \|T\|$$

for all $X$ in $L(H)$.

**Proof.** Let "$B^n = X"$ in (3.1). Take norms:

$$n\|TA^{n-1}\| \leq 2\|A^n\|\|X\| + \|T - (AX - XA)\|\sum_{i=0}^{n-1} \|A^{n-i-1}\|\|A^i\|.$$ 

Since $A$ is paranormal then, by Theorem 2.1, $\|A^k\| = \|A\|^k$ for all $k$ in $\mathbb{N}$ and so the summation above equals $n\|A\|^{n-1}$; and, further, since $T$ is isometric then $n\|TA^{n-1}\| = n\|A^{n-1}\| = n\|A\|^{n-1}$. Dividing through by $n\|A\|^{n-1}$ gives

$$1 \leq \frac{2}{n} \|A\|\|X\| + \|T - (AX - XA)\|.$$ 

Since this holds for all $n$ we have $\|T - (AX - XA)\| \geq 1 = \|T\|$. 

\qed
This result may be expressed geometrically. Consider the linear map \( \Delta_A : L(H) \to L(H) \) given by
\[
\Delta_A = AX - XA
\]
for fixed \( A \) and varying \( X \). Let the linear subsets \( \text{Ran} \, \Delta_A \) and \( \text{Ker} \, \Delta_A \) be given by
\[
\text{Ran} \, \Delta_A = \{ Y \in L(H) : Y = \Delta_A(X) \text{ for varying } X \text{ in } L(H) \}, \\
\text{Ker} \, \Delta_A = \{ X \in L(H) : \Delta_A(X) = 0 \}.
\]

Let \( \mathcal{M} \) and \( \mathcal{N} \) be linear subsets of a normed space \( \mathcal{L} \), say. We say \( \mathcal{M} \) is orthogonal to \( \mathcal{N} \), denoted \( \mathcal{M} \perp \mathcal{N} \), if for all \( m \) in \( \mathcal{M} \) and \( n \) in \( \mathcal{N} \)
\[
\| m + n \| \geq \| m \|.
\]

For historical remarks on this asymmetric definition of orthogonality see [8, p. 93].

**Corollary 3.1.** If \( A \) is paranormal and \( T \) is an isometry in \( \text{Ker} \, \Delta_A \) then
\[
\text{Ker} \, \Delta_A \perp \text{Ran} \, \Delta_A
\]

The next result generalizes a well-known result for hyponormal \( A \) [10, Problem 233].

**Corollary 3.2.** If \( A \) is paranormal then
\[
\| I - (AX - XA) \| \geq \| I \|
\]
for all \( X \) in \( L(H) \).

**Theorem 3.2.** (a) If \( X^* \) is paranormal and \( T \) is an isometry such that \( X^*T = TX^* \) then
\[
\| T - (X^*X - XX^*) \| \geq \| T \|;
\]
(b) The same conclusion holds if, instead, \( X \) is paranormal and \( T \) is an isometry such that \( XT = TX \).

**Proof.** (a) In the identity (3.1) take "\( A \)" = \( X^* \) and "\( B \)" = \( X \) and proceed as in the proof of Theorem 3.1: then
\[
1 \leq \frac{2}{n} \| X^* \| \| X \| - \| T - (X^*X - XX^*) \| \]
which gives the result.

(b) In (3.1) take "\( A \)" = \( -X \) and "\( B \)" = \( X^* \). Then, because \( X \) is paranormal, \( \| A^n \| = \| (-X)^n \| = \| X^n \| = \| X \|^n = \| A \|^n \) and because \( \| T - (X^*X - X^*(-X)) \| = \| T - (X^*X - XX^*) \| \) we get, as in the proof of Theorem 3.1, the inequality 3.2 above, giving the result.

The following example shows that Theorem 3.2 (b) fails if \( XT \neq TX \).
Example 3.1. Let $H = l_2$, the space of square-summable sequences of complex numbers, and let $X = S$, the simple unilateral shift. Then $S$ is hyponormal and hence paranormal: for, with $f = (x_n)_{n=1}^\infty$, 
\[
\langle (S^*S - SS^*)f, f \rangle = |x_1|^2 \geq 0
\]

Let $T$ be given by $T(x_1, x_2, x_3,...) = (x_2, x_1, x_3,...)$. Then $\|Tf\| = \|f\|$ and $ST \neq TS$. Now,
\[
\|T - (S^*S - SS^*)\| = \sup_{\|f\| = 1} |(Tf, f) - \langle (S^*S - SS^*)f, f \rangle| = \sup_{\|f\| = 1} |((x_2, x_1, x_3,...)(\bar{x}_1, \bar{x}_2, x_3,...)) - |x_1|^2| = \sup_{\|f\| = 1} |x_2\bar{x}_1 + x_1\bar{x}_2 + \sum_{n=3}^\infty |x_n|^2 - |x_1|^2|.
\]

Without loss of generality suppose that $Rx_1 \geq Rx_2, Jx_1 \geq Jx_2$ and $x_2 \neq 0$ and $\|f\| = 1$. Then one can check that
\[
(3.3) \leq \sup_{\|f\| = 1} |2|x_1|^2 + \sum_{n=3}^\infty |x_n|^2 - |x_1|^2| = \sup_{\|f\| = 1} \|f\|^2 - |x_2|^2 < \|f\|^2 = 1 = \|T\|.
\]
The next result deals with minimizing $\|T - (AX - XB)\| \geq \|T\|$. It reduces to Theorem 3.1 if $A = B$.

Theorem 3.3. If $A$ and $B$ are paranormal and $T$ is an isometry such that $AT = TB$ and $TA = BT$ then
\[
\|T - (AX - XB)\| \geq \|T\|.
\]

Proof. On $H \oplus H$, let $T = \begin{pmatrix} 0 & T \\ T & 0 \end{pmatrix}$, $A = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$ and $X = \begin{pmatrix} 0 & X \\ 0 & 0 \end{pmatrix}$. Then $T$ is isometric on $H \oplus H$ (since $T$ is isometric on $H$) and $AT = TA$ (since $AT = TB$ and $TA = BT$). Further, $A$ is paranormal on $H \oplus H$: for, with $f = \begin{pmatrix} f_1 \\ f_2 \end{pmatrix}$, on using the paranormality of $A$ and $B$ we get
\[
\|Af\|^4 = \|A \begin{pmatrix} f_1 \\ f_2 \end{pmatrix}\|^4 = (\|A^2f_1\|^2 + \|Bf_2\|^2)^2 = (\|A^2f_1\|^2 + \|B^2f_2\|^2)^2 + 2\|A^2f_1\|\|Bf_2\|\|A^2f_1\|\|Bf_2\|
\]

\[
\leq (\|A^2f_1\|^2 + \|B^2f_2\|^2)^2 + 2\|A^2f_1\|\|B^2f_2\|\|A^2f_1\|\|Bf_2\|)
\]

(3.4)
so that \( \| A f \|^2 \leq \| A^2 f \| \| f \| \) as desired (The inequality (3.4) comes from \( 0 \leq (\| A^2 f_1 \| \| f_2 \| - \| B^2 f_2 \| \| f_1 \|)^2 \)). Therefore, \( A \) and \( T \) satisfy Theorem 3.1 and so \( \| T - (AX - XA) \| \geq \| T \| \) whence

\[
\| T - (AX - XA) \| \geq \| T \|.
\]

\( \square \)

Let the linear map \( \Delta_{A,B} : L(H) \to L(H) \) and the linear subsets \( \text{Ran} \Delta_{A,B} \) and \( \text{Ker} \Delta_{A,B} \) be given by, for fixed \( A \) and \( B \),

\[
\begin{align*}
\Delta_{A,B}(X) &= AX - XB; \\
\text{Ran} \Delta_{A,B} &= \{ Y \in L(H) : Y = \Delta_{A,B}(X) \text{ for varying } X \in L(H) \}; \\
\text{Ker} \Delta_{A,B} &= \{ X \in L(H) : \Delta_{A,B}(X) = 0 \}.
\end{align*}
\]

Theorem 3.3 can be expressed geometrically as follows.

**Corollary 3.3.** If \( A \) and \( B \) are paranormal and \( T \) is an isometry such that \( T \in \text{Ker} \Delta_{A,B} \cap \text{Ker} \Delta_{B,A} \) then \( \text{Ker} \Delta_{A,B} \cap \text{Ker} \Delta_{B,A} \perp \text{Ran} \Delta_{A,B} \).

It is proved in [2, Theorem 1.5] that if \( A \) and \( B \) are normal and \( T \) is such that \( AT = TB \) then

\[
\| T - (AX - XB) \| \geq \| T \|.
\]

Geometrically: if \( A \) and \( B \) are normal and if \( \text{Ker} \Delta_{A,B} \neq \{ 0 \} \) then

\[
\text{Ker} \Delta_{A,B} \perp \text{Ran} \Delta_{A,B}.
\]

This last result, (A-F), together with the rest of this paper, prompts the following questions.

**Question 1.** Can the condition in Theorems 3.1, 3.2 and 3.3 (and in Corollaries 3.1 and 3.3) that \( T \) is isometric be dropped?

**Question 2.** More generally, can the condition of normality in (A-F) be weakened to that of paranormality?

**Acknowledgement** I thank Professor B. P. Duggal for very helpful conversations about paranormality and for giving me his proof of Theorem 2.2.

**References**


Department of Economics and Statistics Middlesex University Hendon Campus
The Burroughs London NW4 4BT England

E-mail: p.maher@mdx.ac.uk