COMMUTATOR APPROXIMANTS

P. J. MAHER

Abstract. This paper deals with minimizing \( \| B - (AX - XA) \|_p \), where \( A \) and \( B \) are fixed, \( B \in \mathcal{C}_p \), and \( X \) varies such that \( AX - XA \in \mathcal{C}_p \). (Here, \( \mathcal{C}_p \) denotes the von Neumann–Schatten class and \( \| \cdot \|_p \) denotes its norm.) The main result (Theorem 3.2) says that if \( A \) is normal and \( AB = BA \) then \( \| B - (AX - XA) \|_p \), \( 1 < p < \infty \), is minimized if and only if, \( AX - XA = 0 \); and that the map \( X \to \| B - (AX - XA) \|_p \), \( 1 < p < \infty \), has a critical point at \( X = V \) if and only if \( AV - VA = 0 \).

1. INTRODUCTION

A well-known result of Halmos [6, Problem 233; 4] says that if \( A \) (or \( B \)) commutes with \( AB - BA \) then
\[
\| \alpha I - (AB - BA) \| \geq \| \alpha I \|.
\]
The related inequality (1.2) was obtained by Anderson [2, Theorem 1.7] who showed that if \( A \) is normal and commutes with \( B \) then, for all \( X \) in \( \mathcal{L}(H) \),
\[
\| B - (AX - XA) \| \geq \| B \|.
\]

In this paper we obtain an inequality similar to (1.2) where the operator norm is replaced by the \( \| \cdot \|_p \) norm on the von Neumann-Schatten classes \( \mathcal{C}_p \), \( 1 \leq p < \infty \). This inequality, contained in Theorem 3.2(a), says that if the normal operator \( A \) commutes with \( B \), where \( B \in \mathcal{C}_p \), and if \( X \) varies such that \( AX - XA \in \mathcal{C}_p \) then, for \( 1 \leq p < \infty \),
\[
\| B - (AX - XA) \|_p \geq \| B \|_p
\]
with equality occurring, and for \( 1 < p < \infty \) only if \( AX - XA = 0 \). Thus in Halmos' terminology [5] the zero commutator is the commutator approximant in \( \mathcal{C}_p \) of \( B \).

Additionally, we classify the critical points of the map \( F_p \), on \( \mathcal{S} = \{ X : AX - XA \in \mathcal{C}_p \} \), defined by
\[
F_p : X \to \| B - (AX - XA) \|_p
\]

Received by the editors March 21, 1990 and, in revised form, January 7, 1991.

1980 Mathematics Subject Classification (1985 Revision). Primary 47B47, 47A30; Secondary 47B10.

Key words and phrases. Commutator, von Neumann–Schatten class, Fuglede's theorem, functional calculus.

©1992 American Mathematical Society
0002-9939/92 $1.00 + $.25 per page
(that is, we classify \( \{ V: \text{the Fréchet derivative } D_V F_p = 0 \} \)). The local result, Theorem 3.2(b), says that under the same hypotheses \((A \text{ normal and } AB = BA)\), \(V\) is a critical point of \(F_p\), \(1 < p < \infty\), if and only if \(AV - VA = 0\).

Note that \(\alpha I - (AX - XA)\) cannot be compact with \(A\) and \(X\) bounded. (This follows from the Wielandt/Wintner result \([9, 10]\) that a commutator of bounded operators cannot equal the identity. If \(\alpha I - (AX - XA)\) were compact, then in the Calkin algebra, the identity would be the commutator of the images of \(A\) and \(X\) contradicting Wielandt/Wintner in the context of normed algebras with unit.) Hence, in infinite dimensions there is no question of minimizing \(\|\alpha I - (AX - XA)\|_p\).

2. Preliminaries

Let \(H\) denote a separable complex Hilbert space and \(\mathcal{L}(H)\) denote the space of all bounded linear operators mapping \(H\) into itself. For details of the von Neumann-Schatten classes \(\mathcal{C}_p\) and norms \(\| \cdot \|_p\), see \([3, \text{Chapter XI}; 8, \text{Chapter 2}]\). We state below the Aiken, Erdos, and Goldstein differentiation result. The real part of a complex number \(z\) is denoted by \(\Re z\).

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

(a) for \(1 < p < \infty\), the map \(\Phi\) is 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 \(\tau\) 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 \(X\).

3. On minimizing \(\|B - (AX - XA)\|_p\)

The proof of the local result, Theorem 3.2(b), depends on Lemma 3.1, which is a variation of the well-known Kleinecke-Shirokov theorem \([6, \text{Problem 232}]\).

**Lemma 3.1.** Let \(A\) be normal and commute with \(AV - VA\). Then \(AV - VA = 0\).

**Proof.** The proof hinges on the spectral resolution of \(A\). This says that there exists a spectral measure \(E(\cdot)\) such that, for each \(\varepsilon > 0\), there exist disjoint Borel sets \(\Delta_i, 1 \leq i \leq N\), with the property that

if \(\lambda_i \in \Delta_i\) and \(S = \sum_{i=1}^{N} \lambda_i E(\Delta_i)\) then \(\|A - S\| < \varepsilon\).

The operator \(AV - VA\) commutes with \(A\) and hence with each of the spectral projections \(E(\Delta_i)\), and so, since \(\sum_{i=1}^{N} E(\Delta_i) = E(U\Delta_i) = I\),

\[
AV - VA = (AV - VA) \sum_{i=1}^{N} E(\Delta_i) = \sum_{i=1}^{N} E(\Delta_i)(AV - VA)E(\Delta_i).
\]

Since the Borel sets are pairwise disjoint, \(E(\Delta_i)E(\Delta_j)\) equals \(E(\Delta_i)\) if \(i = j\) and is zero if \(i \neq j\). Hence on substituting for \(S\), we find that for each (fixed)
We have $E(\Delta_i)(SV - VS)E(\Delta_i) = 0$. So,

$$\|E(\Delta_i)(AV - VA)E(\Delta_i)\| = \|E(\Delta_i)((SV - VS) + (A - S)V - V(A - S))E(\Delta_i)\| \leq 2\|A - S\| \|V\| \|E(\Delta_i)\| < 2\|V\|.$$ 

Since, from (1), $\|AV - VA\| = \sup \|E(\Delta_i)(AV - VA)E(\Delta_i)\|$, then $AV - VA = 0$. □

**Theorem 3.2.** Let $A$ be normal, $AB = BA$, and $B$ be in $\mathcal{C}_p$. Let $\mathcal{S} = \{X : AX - XA \in \mathcal{C}_p\}$ and $F_p : \mathcal{S} \to \mathbb{R}^+$ be given by

$$F_p : X \to \|B - (AX - XA)\|^p.$$ 

Then

(a) for $1 \leq p < \infty$, the map $F_p$ has a global minimizer at $V$ if and for $1 < p < \infty$ only if, $AV - VA = 0$;

(b) for $1 < p < \infty$, the map $F_p$ has a critical point at $V$ if and only if $AV - VA = 0$;

(c) for $0 < p \leq 1$, the map $F_p$ has a critical point at $V$ if $AV - VA = 0$ provided $\dim H < \infty$ and $B - (AV - VA)$ is invertible.

**Proof.** (a) The idea is to replace $B$ by the compact, normal operator $|B|$. Let $B = U|B|$ be the polar decomposition of $B$ so that $\text{Ker}|B| = \text{Ker}|B|$ and $|B| = U^*B \in \mathcal{C}_p$. Since $U$ is a partial isometry so is $U^*$ (so that $\|U^*\| = 1$).

As $\|U^*T\|_p \leq \|U^*\| \|T\|_p = \|T\|_p$ for arbitrary $T$ in $\mathcal{C}_p$, then

$$\|B - (AX - XA)\|^p \geq \|B - U^*(AX - XA)\|^p$$

(1)

$$\geq \sum_n |\langle [B] - U^*(AX - XA)\phi_n, \phi_n \rangle|^p = \sum_n,$$

say, for an arbitrary orthonormal basis $\{\phi_n\}$ of $H$ (the last inequality following from [8, Lemma 2.3.4]).

Because $AB = BA$ and $A$ is normal, by Fuglede's theorem, we have $AB^* = B^*A$, and hence $A|B|^2 = |B|^2A$. Moreover, by the functional calculus [8, Theorem 1.7.7(vi)], $A|B| = |B|A$ (and, indeed, $A|B|^{p-1} = |B|^{p-1}A$). Therefore, there exists an orthonormal basis $\{\xi_k\} \cup \{\psi_m\}$ of $H$ such that $\{\psi_m\}$ is an orthonormal basis of $\text{Ker}|B|$ and $\{\xi_k\}$ consists of common eigenvectors of $A$ and $|B|$. (I thank the referee for suggesting this basis.) Hence, $\sum_k \langle |B|\xi_k, \xi_k \rangle^p = \|B\|^p$. As $B^*A = AB^*$ and $A|B| = |B|A$, $B|B|^{p-1} = |B|^{p-1}A$. In (1), take $\{\phi_n\} = \{\xi_k\} \cup \{\psi_m\}$. If $\phi_n = \xi_k$, then $\langle U^*AX\xi_k, \xi_k \rangle = \langle AU^*X\xi_k, \xi_k \rangle$ and hence, as $\xi_k$ is also an eigenvector of the normal operator $A$, $\langle U^*(AX - XA)\xi_k, \xi_k \rangle = \langle AU^*X - U^*XA\xi_k, \xi_k \rangle = 0$. Thus (1) becomes

$$\sum_k \langle |B|\xi_k, \xi_k \rangle^p + \sum_m |\langle U^*(AX - XA)\psi_m, \psi_m \rangle|^p \geq \sum_k \langle |B|\xi_k, \xi_k \rangle^p = \|B\|^p$$

as desired.

For $1 < p < \infty$ the uniqueness assertion follows from the convexity of the set $\mathcal{S} = \{X : AX - XA \in \mathcal{C}_p\}$.

(b) Let $V$ be in $\mathcal{S}$ so that $B - (AV - VA) \in \mathcal{C}_p$. Let $S$ be arbitrary subject to the condition that $B - (A(V + S) - (V + S)A) \in \mathcal{C}_p$, that is, $SA - AS \in \mathcal{C}_p$. 

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use
Let $\Psi \colon X \to B - (AX - XA)$ and $\Phi \colon X \to \|X\|_p^p$. Then $F_p = \Phi \circ \Psi$. As $F_p$ is real-valued, the Fréchet derivative of $F_p$ at $V$, denoted by $D_V F_p$, is given by

$$
(D_V F_p)(S) = \lim_{h \to 0} \frac{F_p(V + hS) - F_p(V)}{h}.
$$

From this it follows that

$$
(2) \quad (D_V F_p)(S) = (D_{B - (AV - VA)} \Phi)(SA - AS).
$$

Let $V$ be a critical point of $F_p$ so that $(D_V F_p)(S) = 0$ for all $S$ in $S$. Let $B - (AV - VA) = U_1 B - (AV - VA)$ be the polar decomposition of $B - (AV - VA)$. Then from Theorem 2.1 and (2),

$$
(3) \quad 0 = pR \tau[|B - (AV - VA)|^{p-1} U_1^* (SA - AS)] = pR \tau[Y(SA - AS)],
$$

where $Y = |B - (AV - VA)|^{-1} U_1^*$. Take $S = f \otimes g$, where $f$ and $g$ are arbitrary vectors in $H$. (The rank one operator $x \mapsto (x, f)g$, where $x \in H$, is denoted $f \otimes g$. Note that $\tau[T(f \otimes g)] = (Tg, f)$, cf. [8, pp. 73, 90].) Then, as $S \in \mathcal{E}$, (whence $YS \in \mathcal{E}$), from the invariance of trace [8, Theorem 2.2.4(iv)] we have $\tau[YS] = \tau[AYS]$. Thus, from (3)

$$
(4) \quad 0 = \mathcal{R} \tau[(AY - YA)S] = \mathcal{R} \tau((AY - YA)g, f).
$$

Because $f$ and $g$ are arbitrary, $AY - YA = 0$, that is,

$$
(5) \quad A|B - (AV - VA)|^{p-1} U_1^* = |B - (AV - VA)|^{p-1} U_1^* A.
$$

We claim that if

$$
(6) \quad A|B - (AV - VA)|^{p-1} U_1^* = |B - (AV - VA)|^{p-1} U_1^* A,
$$

then the assertion that $AV - VA = 0$ will be proved. For suppose (6) holds. Then taking adjoints and using the polar decomposition of $B - (AV - VA)$ and Fuglede's theorem (which gives $BA^* = A^* B$), we get $(AV - VA)A^* = A^*(AV - VA)$. By Fuglede again, we get $(AV - VA)A = A(AV - VA)$. Hence, by Lemma 3.1, $AV - VA = 0$.

**Proof of (6).** Write $Z = |B - (AV - VA)|^{p-1}$. Then (5) says that

$$
(7) \quad AZ U_1^* = ZU_1^* A
$$

and (6) is the same as $AZ^{1/(p-1)} U_1^* = Z^{1/(p-1)} U_1^* A$. This will follow by the functional calculus (cf. [7, Theorem 4.1]) from

$$
(8) \quad AZ^n U_1^* = Z^n U_1^* A;
$$

for the function $f : t \mapsto t^{1/(p-1)}$, $1 < p < \infty$, where $t \in \mathbb{R}^+ \supseteq \sigma(Z)$ can be approximated uniformly by a sequence $(q_i)$ of polynomials without constant term (for $f(0) = 0$). Thus, (8) will imply that $Aq_i(Z) U_1^* = q_i(Z) U_1^*$ and hence that $AZ^{1/(p-1)} U_1^* = Z^{1/(p-1)} U_1^* A$.

To prove (8) we use induction. We need the following assertion: $AZ = Z A$. To prove this assertion, note that in the polar decomposition of $B - (AV - VA)$ we have $\ker U_1 = \ker |B - (AV - VA)| = \ker Z$ (by the spectral theorem) so that $(\ker U_1)^\perp = \text{Ran } Z = U_1^*$. Thus, $U_1^* U_1$, the projection onto $(\ker U_1)^\perp$, satisfies $U_1^* U_1 Z = Z$ and hence $Z U_1^* U_1 Z = Z^2$. Now take adjoints of (7): then $U_1 Z A^* = A^* U_1 Z$ and hence, by Fuglede, $U_1 Z A = A U_1 Z$. Then by (7)

$$
A Z^2 = A Z U_1^* U_1 Z = Z U_1^* A U_1 Z = Z U_1^* U_1 Z A = Z^2 A.
$$
Taking positive square roots of $Z^2$ [8, Theorem 1.7.7(vi)] we get $AZ = ZA$. Returning now to (8): for $n = 1$, (8) is just (7); whilst the inductive step is now immediate from $AZ = ZA$.

Conclusion so far: $V$ is a critical point of $F_p \Rightarrow AV - VA = 0$.

Conversely, let $V$ satisfy $AV - VA = 0$. Then $B - (AV - VA) = 0$ and so the partial isometries $U_1$ and $U$ in the polar decompositions of $B - (AV - VA)$ and $B$ coincide. Thus, $Y = (B - (AV - VA)|^{p-1}U_*^* = |B|^{p-1}U_*^* \in \mathcal{C}_1$. As in part (a), using Fuglede, we have $|B|U_*^* = |B|U_*A$ and $A|B|^{p-1} = |B|^{p-1}A$. Hence, $\text{Ran}(AU_* - UA) \subseteq \text{Ker} |B| = \text{Ker} |B|^{p-1}$. So, $AY - YA = A|B|^{p-1}U_* - |B|^{p-1}U_*A = |B|^{p-1}(AU_* - UA) = 0$.

So, as $YS \in \mathcal{C}_1$, then (cf. (4), (3), and (2)) $DF_p(S) = 0$ for all $S$ in $\mathcal{L}(H)$.

(c) For $0 < p \leq 1$, the finite-dimensionality and invertibility conditions ensure, by Theorem 2.1(b), that $F_p$ is differentiable at $V$. If $AV - VA = 0$ then $B$, and hence $|B|$, is invertible and so $|B|^{p-1}$ exist for $0 < p < 1$. The proof of the implication, $AV - VA = 0 \Rightarrow V$ is a critical point of $F_p$, is now the same as in part (a).

We make some comments.

(i) In Theorem 3.2(a) if $B = AX_1 - X_1A$ for some operator $X_1$ then the minimum of $\|B - (AX - XA)\|_p$ is 0. This does not conflict with Theorem 3.2(a) (i.e. (1.3)) because in this case $B = 0$; for since the normal operator $A$ commutes with $B(= AX_1 - X_1A)$ then, by Lemma 3.1, $AX_1 - X_1A = 0$.

(ii) The following counterexample shows that Theorem 3.2(a) does not hold if $p < 1$. Take $p = \frac{1}{2}$ and $A = \begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$, and $X = \begin{pmatrix} 0 & x \\ -x & 0 \end{pmatrix}$ where $a$ and $x$ are reals such that $0 < |ax| < 1$. Then $\|B - (AX - XA)\|_{1/2} < \|B\|_{1/2}$.

(iii) The set $\mathcal{S} (= \{X : AX - XA \in \mathcal{C}_p\})$ properly contains $\mathcal{C}_p$, for if $X \in \mathcal{C}_p$ then $X \in \mathcal{S}$ and, e.g., $I \in \mathcal{S}$ but $I \notin \mathcal{C}_p$. If $A \in \mathcal{C}_p$, the conclusions of Theorem 3.2 hold for all $X$ in $\mathcal{S}(H)$.

(iv) The converse in Theorem 3.2(b) can be proved on the basis of the global result: if $AV - VA = 0$ then, by Theorem 3.2(a), $V$ is a global minimizer, and hence a critical point, of $F_p$.

(v) The proof in Theorem 3.2(b) of the implication, $V$ is a critical point of $F_p \Rightarrow AV - VA = 0$, does not work in the $0 < p < 1$ case because the functional calculus argument involving the function $f : t \mapsto t^{1/(p-1)}$, where $0 \leq t < \infty$, is only valid for $1 < p < \infty$.

(vi) Finally, Anderson’s original result (1.2), in the special case where $B$ is compact, can be obtained similarly to Theorem 3.2(a). (Proof: using the fact that $\|S\| \geq \sup_{|\phi| = 1} |\langle S\phi, \phi \rangle|$, where $S \in \mathcal{L}(H)$, with the basis $\{\phi_n\} = \{\xi_k\} \cup \{\psi_m\}$ as defined in Theorem 3.2(a), we get

$$\|B - (AX - XA)\| \geq \sup_n \|B - U^*(AX - XA)\| \|\phi_n\| \|\psi_n\|$$

$$= \sup_k \|B\| \|\xi_k\| + \|U^*(AX - XA)\| \|\psi_m\|$$

$$\geq \sup_k \|B\| \|\xi_k\| = \|B\| \|\|B\|\|.$$
like to thank the referee for drawing my attention to Anderson's paper and for suggesting a strengthening of my original version of Theorem 3.2(a).

REFERENCES


SCHOOL OF MATHEMATICS, MIDDLESEX POLYTECHNIC, TRENT PARK, BRAMLEY ROAD, LONDON N14 4XS, ENGLAND