LEIBNIZ SEMINORMS AND BEST APPROXIMATION FROM C*-SUBALGEBRAS

MARC A. RIEFFEL

In celebration of the successful completion by Richard V. Kadison of 85 circumnavigations of the sun

Abstract. We show that if $B$ is a C*-subalgebra of a C*-algebra $A$ such that $B$ contains a bounded approximate identity for $A$, and if $L$ is the pull-back to $A$ of the quotient norm on $A/B$, then $L$ is strongly Leibniz. In connection with this situation we study certain aspects of best approximation of elements of a unital C*-algebra by elements of a unital C*-subalgebra.

1. Introduction

From my attempts to understand C*-metrics, as defined in [18], I have recently been trying to discover the mechanisms that can lead to (continuous) seminorms, $L$, defined on a C*-algebra $A$, that are Leibniz, that is, satisfy the Leibniz inequality

$$L(AC) \leq L(A)\|C\| + \|A\|L(C)$$

for all $A, C \in A$. Also of much interest to me, because of their importance in [17] and in potential non-commutative versions of the results in [17], are seminorms on a unital C*-algebra that are strongly-Leibniz, that is, satisfy, in addition to the Leibniz inequality, the property that if $A \in A$ and if $A$ is invertible in $A$, then

$$L(A^{-1}) \leq \|A^{-1}\|^2L(A),$$

and also $L(1_A) = 0$, again as defined in [18]. This last inequality seems to have received virtually no attention in the mathematics literature.

Actually, for infinite-dimensional C*-algebras, the C*-metrics as defined in [18] are discontinuous and only densely defined. But they are

2000 Mathematics Subject Classification. Primary 46L87; Secondary 53C23, 58B34, 81R15, 81R30.

Key words and phrases. C*-subalgebras, best approximation, distance formula, Leibniz inequality, strongly-Leibniz, minimal elements.

The research reported here was supported in part by National Science Foundation grant DMS-0753228.
required to be lower semi-continuous with respect to the C*-norm, and in all of the examples that I know of one proves that they are lower semi-continuous by showing that they are the supremum of an infinite family of continuous strongly-Leibniz seminorms. This provides ample reason for studying continuous strongly-Leibniz seminorms.

My investigations have led me to consider the situation in which \(B\) is a C*-subalgebra of a C*-algebra \(A\) and \(L\) is the quotient norm on \(A/B\) pulled back to \(A\). That is,

\[
L(A) = \inf\{\|A - B\| : B \in B\}
\]

for \(A \in A\), so that \(L(A)\) is the norm-distance from \(A\) to \(B\). One motivation for studying such quotient seminorms is that they arise naturally when considering matricial Lipschitz seminorms, as defined in [25, 26, 27] (where the Leibniz aspect was not used) and as discussed in the final section of [18]. These involve a unital C*-algebra \(D\), and for each natural number \(m\) the corresponding matrix C*-algebra \(A = M_m(D) = M_m(\mathbb{C}) \otimes D\) over \(D\). The Leibniz seminorm on \(A\) from a matricial Leibniz seminorm on \(D\) will take value 0 on the C*-subalgebra \(B = M_m(\mathbb{C}) \otimes 1_D\), and so can be viewed as a seminorm on \(A/B\). Furthermore, in this situation \(B\) is finite-dimensional, and so elements of \(A\) will always have a best approximation by an element of \(B\) (probably not unique). This is good motivation for studying best approximations. Although there is a very large literature dealing with best approximation of elements of a Banach space by elements of a closed subspace [22], I have found almost no literature concerning this topic for the case of C*-algebras and their C*-subalgebras. The later sections of this paper are devoted to developing some basic results about this for C*-algebras. We show that the situation for C*-algebras has some nice features in comparison to the general case, and we are able to tie together a few results scattered in the literature.

Somewhat to my surprise I have found that the seminorms from quotients of C*-algebras are usually strongly-Leibniz. To be specific, the most important theorem of this paper states:

**Theorem.** Let \(A\) be a C*-algebra and let \(B\) be a C*-subalgebra of \(A\). Assume that \(B\) contains a bounded approximate identity for \(A\). Let \(L\) be defined as above. Then \(L\) is Leibniz, that is,

\[
L(AC) \leq L(A)\|C\| + \|A\|L(C)
\]

for all \(A, C \in A\). If \(A\) is unital and if \(1_A \in B\), then \(L\) is strongly-Leibniz, that is, \(L\) is Leibniz, \(L(1) = 0\), and if \(A\) is invertible in \(A\) then

\[
L(A^{-1}) \leq \|A^{-1}\|^2L(A).
\]
This theorem is a fairly simple consequence of a C*-algebra version of the Arveson distance formula [2] that has been widely used in the study of nest algebras. This C*-algebra version and the proof of the above theorem are given in Section 3.

One can, of course, ask what happens for Banach algebras $\mathcal{A}$ and their closed subalgebras $\mathcal{B}$. In Section 2 we discuss a property, that we call the “same-norm approximation property”, that guarantees that the seminorm on $\mathcal{A}$ pulled back from the quotient norm on $\mathcal{A}/\mathcal{B}$ is Leibniz. We then show that if $\mathcal{A}$ is a C*-algebra and $\mathcal{B}$ is a C*-subalgebra of $\mathcal{A}$ that is central in $\mathcal{A}$, then $\mathcal{B}$ has the same-norm approximation property in $\mathcal{A}$.

But a main application of our later discussion of best approximation is to give an example of a finite-dimensional C*-algebra and a (non-central) unital C*-subalgebra that does not have the same-norm approximation property. The nature of the construction strongly suggests that the same-norm approximation property will usually fail unless one is dealing with central subalgebras.

It is a pleasure to thank Man-Duen Choi, Erik Christensen, and Dick Kadison for very helpful conversations concerning the subject of this paper.

2. The Leibniz inequality and Banach algebras

In this section we make some elementary observations in the setting of Banach algebras. These observations relate the Leibniz inequality for quotient seminorms with a property of approximations by elements of a subalgebra. We then apply these observations to the setting of C*-algebras.

Our experience with Hilbert spaces leads us to expect that if $B$ is a best approximation to $A$ then $\|B\| \leq \|A\|$. But this easily fails for the seminorms coming from best approximation in C*-algebras, as we will see later. In general, the most that one can say is that

$$\|B\| \leq 2\|A\|$$

because $B$ must be at least as close to $A$ as is 0. Easy examples show that 2 is the best constant that holds here in general. But the stronger inequality is desirable for our purposes, as we will show in this section, and so we make the following definition:

**Definition 2.1.** Let $\mathcal{A}$ be a normed vector space and let $\mathcal{B}$ be a closed subspace of $\mathcal{A}$. Let $M$ be the quotient norm on $\mathcal{A}/\mathcal{B}$, pulled back to $\mathcal{A}$. That is,

$$M(A) = \inf\{\|A - B\| : B \in \mathcal{B}\}.$$
Given $A \in \mathcal{A}$, we say that $A$ is *same-norm approximatable* in $\mathcal{B}$ if for every $\varepsilon > 0$ there is a $B \in \mathcal{B}$ with $\|B\| \leq \|A\|$ and $\|A - B\| < M(A) + \varepsilon$. If every element of $\mathcal{A}$ is same-norm approximatable in $\mathcal{B}$ then we say that $\mathcal{B}$ has the *same-norm approximation property* in $\mathcal{A}$.

**Proposition 2.2.** Let $\mathcal{A}$ be a normed algebra, let $\mathcal{B}$ be a closed subalgebra of $\mathcal{A}$, and let $M$ be the quotient norm on $\mathcal{A}/\mathcal{B}$, pulled back to $\mathcal{A}$. Let $A \in \mathcal{A}$. If $A$ is same-norm approximatable in $\mathcal{B}$, then for every $C \in \mathcal{A}$ the Leibniz inequality

$$M(AC) \leq M(A)\|C\| + \|A\|M(C)$$

holds for all $C \in \mathcal{A}$, and similarly for $M(CA)$.

**Proof.** Given $\varepsilon > 0$, choose $D \in \mathcal{B}$ such that $\|C - D\| < M(C) + \varepsilon$, and choose $B \in \mathcal{B}$ such that $\|B\| \leq \|A\|$ and $\|A - B\| < M(A) + \varepsilon$. Then

$$M(AC) \leq \|AC - BD\| \leq \|A - B\|\|C\| + \|B\|\|C - D\|$$

$$\leq (M(A) + \varepsilon)\|C\| + \|A\|(M(C) + \varepsilon).$$

Since $\varepsilon$ is arbitrary, we obtain the desired inequality. A similar calculation works for $M(CA)$. \qed

We remark that the above proof is somewhat parallel to the proof of proposition 5.2 of [18].

In the setting of normed algebras it would be interesting to have conditions that also ensure that $M$ is strongly-Leibniz, as defined earlier. For this to make sense we need that $\mathcal{A}$ is unital, (There may well be a suitable generalization of the notion of strongly-Leibniz to non-unital algebras, but I have not explored this possibility.)

For C*-algebras the main situation in which the same-norm approximation property seems to arise is the following.

**Proposition 2.3.** Let $\mathcal{A}$ be a C*-algebra and let $\mathcal{B}$ be a C*-subalgebra of $\mathcal{A}$ such that $\mathcal{B}$ is central in $\mathcal{A}$. Then $\mathcal{B}$ has the same-norm approximation property in $\mathcal{A}$.

More specifically, let $A \in \mathcal{A}$, let $F \in \mathcal{B}$, and let $G$ be the radial retraction of $F$ to the ball about $0$ of radius $\|A\|$ in $\mathcal{B}$ defined by the continuous-function calculus for normal elements. Then $\|A - G\| \leq \|A - F\|$, and, of course, $\|G\| \leq \|A\|$.

**Proof.** The assertion in the second paragraph implies the assertion in the first paragraph because it shows that whatever approximation we have, we can always get from it another that is as close but also satisfies the same-norm condition.
If \( A \) is not unital, then when we adjoin an identity element to \( A \) the subalgebra \( B \) is still central, and distances are not changed. So we will assume now that \( A \) is unital. Let \( \tilde{B} \) denote \( B \) with the identity element of \( A \) adjoined (if it is not already in \( B \)).

To prove the assertion of the second paragraph, notice that when we view \( \tilde{B} \) as the algebra of continuous functions on a compact space, we see that \( G = HF \) where \( H = K^{-1} \) and \( K = \max\{1, |F|/\|A\|\} \) as functions. Notice that \( K - 1 \) will be in \( B \), as will then \( H - 1 \), so that \( K \) and \( H \) are both multipliers of \( B \). To prove the assertion of the second paragraph, it suffices to show that as operators

\[
(A - F)^* (A - F) \geq (A - G)^* (A - G),
\]

that is, upon simplification, that

\[
|F|^2 (1 - H^2) \geq (1 - H) (F^* A + A^* F).
\]

Let \((\pi, \mathcal{H})\) be an irreducible representation of \( A \), so that \( \pi(F) \) is a scalar multiple, say, \( \lambda \), of \( I_{\mathcal{H}} \), since \( B \) is central. Then \( \pi(H) = \mu I_{\mathcal{H}} \) where \( \mu = 1 \) if \( |\lambda| \leq \|A\| \) while \( \mu = \|A\|/|\lambda| \) otherwise. Then for the above inequality we need to know that

\[
|\lambda|^2 (1 - |\mu|^2) \geq (1 - \mu) (\bar{\lambda} \pi(A) + \lambda \pi(A^*)).
\]

Now \( \bar{\lambda} \pi(A) + \lambda \pi(A^*) \) has norm \( \leq 2|\lambda| \|A\| \) and is Hermitian, and thus it suffices to know that

\[
|\lambda|^2 (1 - |\mu|^2) \geq (1 - \mu) 2|\lambda| \|A\|.
\]

If \( \mu = 1 \) then both sides are 0, while if \( |\lambda| > \|A\| \) so that \( \mu = \|A\|/|\lambda| < 1 \) then a simple calculation shows that this inequality holds. Since all of this works for any irreducible representation, we obtain the desired inequality. \( \square \)

**Corollary 2.4.** Let \( A \) be a \( C^* \)-algebra and let \( B \) be a \( C^* \)-subalgebra of \( A \) such that \( B \) is central in \( A \), and let \( M \) be the quotient norm on \( A/B \) pulled back to \( A \). Then \( M \) satisfies the Leibniz inequality.

This corollary is related to the main result of [10], which, however, is more in the spirit of the next section.

We remark that if \( A \) is unital and if \( 1_A \notin B \), then \( M(1_A) \neq 0 \), so that \( M \) can not be strongly-Leibniz.

**Corollary 2.5.** Let \( A \) be a unital \( C^* \)-algebra, let \( B = \mathbb{C}1_A \), and let \( M \) be the corresponding seminorm on \( A \). Then \( M \) satisfies the Leibniz inequality.
This last corollary shows that we do obtain a Leibniz seminorm in the $m = 1$ case of the situation in which $\mathcal{A} = M_m(\mathcal{D})$ and $\mathcal{B} = M_m(\mathbb{C}1_{\mathcal{D}})$ that was mentioned in introduction. But it does not show that the seminorm is strongly-Leibniz. (We will see in the next section that it is indeed strongly-Leibniz.) This corollary is also strongly related to an often-cited result of Stampfli [23], and to, for example, equation 1.6 of [3] and section 3.1 of [5].

We remark that if $\mathcal{A} = M_2(\mathbb{C})$ and if $\mathcal{B}$ is the C*-subalgebra of all diagonal matrices in $\mathcal{A}$, then $\mathcal{B}$ has the same-norm approximation property in $\mathcal{A}$. This is seen easily by direct calculation. Thus a proper C*-subalgebra does not need to be central in order to have the same-norm approximation property. But it would be interesting to know if there are any other C*-algebras that have a proper non-central C*-subalgebra that has the same-norm approximation property. This may be related to the fact shown in [16] that $M_2(\mathbb{C})$ is the only C*-algebra that has a proper Chebyshev C*-subalgebra of dimension strictly greater than 1.

**Corollary 2.6.** Let $\mathcal{A}$ be a C*-algebra, let $\mathcal{B}$ be a C*-subalgebra of $\mathcal{A}$, and let $M$ be the quotient norm on $\mathcal{A}/\mathcal{B}$ pulled back to $\mathcal{A}$. Let $A \in \mathcal{A}$. If there is a best approximation $B \in \mathcal{B}$ to $A$ which is normal and commutes with $A$, then $A$ is same-norm approximatable by $B$ (so Proposition 2.2 is applicable).

*Proof.* By Fuglede’s theorem (theorem 4.76 of [8]) $A$ also commutes with $B^*$. Thus the C*-algebra $C^*(B)$ generated by $B$ is a central C*-subalgebra of the C*-algebra $C^*(A, B)$ generated by $A$ and $B$. Thus we can apply Proposition 2.3. Since $C^*(B)$ contains a best approximation to $A$ in $\mathcal{B}$, namely $B$, the desired conclusion follows. $\square$

**Corollary 2.7.** Let $\mathcal{D}$ be a C*-algebra, and for a natural number $m$ let $\mathcal{A} = M_m(\mathcal{D})$ and let $\mathcal{B}$ be the C*-subalgebra $\mathcal{B} = M_m(\mathbb{C}) \otimes 1_{\mathcal{D}}$ as mentioned in the introduction. Then any element $A$ of $\mathcal{A}$ that is in the commutant of $\mathcal{B}$ in $\mathcal{A}$ (i.e. is in $I_m \otimes \mathcal{D}$) is same-norm approximatable in $\mathcal{B}$ (so that Proposition 2.2 is applicable).

*Proof.* It is easily seen that if $A = I_m \otimes D$ for some $D \in \mathcal{D}$, and if $\lambda \in \mathbb{C}$ is such that $\lambda 1_{\mathcal{D}}$ is a best approximation to $D$ in $\mathbb{C}1_{\mathcal{D}}$, chosen via Corollary 2.5 so that $|\lambda| \leq \|D\|$, then $B = \lambda I_m \otimes 1_{\mathcal{D}}$ is a best approximation to $A$ in $\mathcal{B}$ and $\|B\| \leq \|A\|$. $\square$

It would be interesting to have examples of Banach algebras that are not C*-algebras but that have subalgebras that satisfy the same-norm approximation property.
3. MANY QUOTIENT NORMS OF C*-ALGEBRAS ARE STRONGLY LEIBNIZ

Let $\mathcal{A}$ be a C*-algebra and let $\mathcal{B}$ be a C*-subalgebra of $\mathcal{A}$. As before we let $L$ denote the pull-back to $\mathcal{A}$ of the corresponding quotient norm on $\mathcal{A}/\mathcal{B}$, so that

$$L(A) = \inf\{\|A - B\| : B \in \mathcal{B}\}$$

for all $A \in \mathcal{A}$. For the purposes of my general investigation of Leibniz seminorms on C*-algebras, the most important theorem of this paper is the following:

**Theorem 3.1.** Let $\mathcal{A}$ be a C*-algebra and let $\mathcal{B}$ be a C*-subalgebra of $\mathcal{A}$. Assume that $\mathcal{B}$ contains a bounded approximate identity for $\mathcal{A}$. Let $L$ be defined as above. Then $L$ is Leibniz, that is,

$$L(AC) \leq L(A)\|C\| + \|A\|L(C)$$

for all $A,C \in \mathcal{A}$. If $\mathcal{A}$ is unital and if $1_A \in \mathcal{B}$, then $L$ is strongly-Leibniz, that is, $L$ is Leibniz, $L(1) = 0$, and if $A$ is invertible in $\mathcal{A}$ then

$$L(A^{-1}) \leq \|A^{-1}\|^2L(A).$$

In proposition 1.2iii of [18] it is seen that any supremum of a (possibly infinite) family of Leibniz seminorms will again be Leibniz, with a similar statement for strongly-Leibniz seminorms. In proposition 2.1 of [18] it is seen that for any derivation $d$ from $\mathcal{A}$ into a normed bimodule, if we set $L_d(A) = \|d(A)\|$ then $L_d$ is a Leibniz seminorm on $\mathcal{A}$, and if $\mathcal{A}$ is unital and the bimodule is non-degenerate, then $L$ is strongly-Leibniz. Combining these two facts gives the last part of the next theorem. Furthermore, Theorem 3.1 is an immediate consequence of this next theorem.

**Theorem 3.2.** Let $\mathcal{A}$ be a C*-algebra and let $\mathcal{B}$ be a C*-subalgebra of $\mathcal{A}$. Assume that $\mathcal{B}$ contains a bounded approximate identity for $\mathcal{A}$. Let $L$ be defined as above. For any $A \in \mathcal{A}$ there is a non-degenerate $*$-representation, $(\mathcal{H}, \pi)$, of $\mathcal{A}$, and a Hermitian unitary operator $U \in \mathcal{L}(\mathcal{H})$, such that $[U, \pi(B)] = 0$ for all $B \in \mathcal{B}$, and

$$L(A) = (1/2)\|[U, \pi(A)]\|.$$ 

Furthermore, $L$ is the supremum of all the seminorms $L_{(\mathcal{H}, \pi, U)}$ defined by $L_{(\mathcal{H}, \pi, U)}(C) = (1/2)\|[U, \pi(C)]\|$ for $C \in \mathcal{A}$, as $(\mathcal{H}, \pi, U)$ ranges over all non-degenerate $*$-representations, $(\mathcal{H}, \pi)$, of $\mathcal{A}$ and all Hermitian unitary elements $U \in \mathcal{L}(\mathcal{H})$ such that $[U, \pi(B)] = 0$ for all $B \in \mathcal{B}$. Since each $L_{(\mathcal{H}, \pi, U)}$ is Leibniz, so is $L$. If $\mathcal{A}$ is unital and if $1_A \in \mathcal{B}$, then $L$ is strongly-Leibniz.
Theorem 3.2 is in turn a C*-algebraic variation on Arveson’s distance formula for nest algebras [2], as used in the case of von Neumann algebras [6]. I thank Erik Christensen for elucidating for me this formula for the case of von Neumann algebras.

I have not seen a good way to weaken the assumption that \( \mathcal{B} \) contains a bounded approximate identity for \( \mathcal{A} \).

Proof of theorem 3.2. Notice that if \( C \in \mathcal{A} \) then for any \((\mathcal{H}, \pi, U)\) as above we have \( \| [U, \pi(C)] \| = \|[U, \pi(C - B)]\| \leq 2\|C - B\| \) for all \( B \in \mathcal{B} \), and thus

\[
L_{(\mathcal{H}, \pi, U)}(C) \leq L(C)
\]

for all \( C \in \mathcal{A} \). The fact that \( L \) is the supremum of the seminorms \( L_{(\mathcal{H}, \pi, U)} \) thus follows from the first statement in the theorem. The fact that \( L \) is strongly-Leibniz if \( \mathcal{A} \) is unital and if \( 1_{\mathcal{A}} \in \mathcal{B} \) then follows from the comments made before the statement of the theorem. Thus it remains to prove the first statement of the theorem.

So let \( A \in \mathcal{A} \) be given. By scaling, we see that it suffices to treat the case in which \( L(A) = 1 \), and so for ease of bookkeeping we assume this. We now use the first basic tool of linear approximation theory. By applying the Hahn-Banach theorem to the image of \( A \) in \( \mathcal{A}/\mathcal{B} \), we see that there is a linear functional, \( \psi \), on \( \mathcal{A} \), such that \( \| \psi \| = 1 \), \( \psi(A) = 1 \), and \( \psi(B) = 0 \) for all \( B \in \mathcal{B} \). Our proof is then based on the following key lemma, which is very closely related to the polar decomposition of linear functionals in the predual of a von Neumann algebra. See notably corollary 7.3.3 of [15]. This lemma can also be obtained by examining the universal representation of \( \mathcal{A} \), as suggested in the proof of lemma 1 of [16].

Lemma 3.3. Let \( \mathcal{A} \) be a C*-algebra, and let \( \psi \) be a linear functional on \( \mathcal{A} \) such that \( \| \psi \| = 1 \). Then there is a cyclic \( \ast \)-representation, \((\mathcal{H}, \pi, \xi)\), of \( \mathcal{A} \), and a vector \( \eta \in \mathcal{H} \), such that \( \| \xi \| = 1 = \| \eta \| \) and

\[
\psi(C) = \langle \pi(C)\xi, \eta \rangle
\]

for all \( C \in \mathcal{A} \).

Before giving the proof of this lemma, we show how to use it to complete the proof of Theorem 3.2. As in the first fragment of the proof of Theorem 3.2 given above, let \( \psi \) be the linear functional on \( \mathcal{A} \) with \( \| \psi \| = 1 \), \( \psi(A) = L(A) = 1 \), and \( \psi(B) = 0 \) for all \( B \in \mathcal{B} \). According to Lemma 3.3 there is a non-degenerate \( \ast \)-representation \((\mathcal{H}, \pi)\), of \( \mathcal{A} \), and vectors \( \xi, \eta \in \mathcal{H} \), such that \( \| \xi \| = 1 = \| \eta \| \) and

\[
\psi(C) = \langle \pi(C)\xi, \eta \rangle
\]
for all \( C \in \mathcal{A} \). Let \( P \) denote the orthogonal projection of \( \mathcal{H} \) onto the closure of \( \pi(B)\xi \). Since the closure of \( \pi(B)\xi \) is clearly \( \mathcal{B} \)-invariant, we have \([P, \pi(B)] = 0 \) for all \( B \in \mathcal{B} \). Notice that for all \( B \in \mathcal{B} \) we have \( 0 = \langle \pi(B)\xi, \eta \rangle \), and thus \( \eta \perp \pi(B)\xi \), so that \( P(\eta) = 0 \). Also, because \( \mathcal{B} \) contains a bounded approximate identity for \( \mathcal{A} \), we see that \( \xi \) is in \( \pi(B)\xi \), so that \( P(\xi) = \xi \). For notational simplicity, let us now write \( C \) for \( \pi(C) \), etc. Then

\[
\langle [C, P] \xi, \eta \rangle = \langle CP\xi, \eta \rangle - \langle PC\xi, \eta \rangle = \langle C\xi, \eta \rangle = \psi(C)
\]

for all \( C \in \mathcal{A} \). In particular, \( \|[[A, P]] \| \geq |\psi(A)| = L(A) \). This does not quite fit our needs, as we only have the general estimate that for \( C \in \mathcal{A} \) and \( B \in \mathcal{B} \)

\[
\|[C, P] \| = \|[C - B, P] \| \leq 2\|C - B\|,
\]

so that \( \|[C, P] \| \leq 2L(C) \), and we do not want the factor of 2 here. (But notice the importance of having \( P \) commute with all the \( B \)'s.) To rectify this, let \( U = 2P - I \), so that \( U \) is a Hermitian unitary, and in particular, \( \|U\| = 1 \). Then we will have

\[
\psi(C) = (1/2) \langle [C, U] \xi, \eta \rangle
\]

for all \( C \in \mathcal{A} \), so that in particular \( (1/2)\|[U, A] \| \geq L(A) \). But for any \( C \in \mathcal{A} \) and \( B \in \mathcal{B} \) we still have, much as above,

\[
\|[C, U] \| = \|[C - B, U] \| \leq 2\|C - B\|,
\]

so that \( (1/2)\|[U, C] \| \leq L(C) \). This gives the desired result. \( \square \)

For a related result in the finite-dimensional case see remark 3.1 of [3].

**Proof of Lemma 3.3.** For the reader’s convenience we now give a proof of Lemma 3.3 that uses as its biggest tool just the Jordan decomposition of Hermitian linear functionals into differences of two positive linear functionals.

Let \( \mathcal{A} \) and \( \psi \) be as in the statement of Lemma 3.3. If \( \mathcal{A} \) is not unital, adjoin an identity element in the usual way so as to obtain a \( C^* \)-algebra. By the technical step in the proof of the Hahn-Banach theorem we can extend \( \psi \) to the resulting \( C^* \)-algebra, with no increase in its norm. Until the end of the proof we will now assume that \( \mathcal{A} \) is unital.

Form the algebra \( M_2(\mathcal{A}) \) of \( 2 \times 2 \) matrices with entries in \( \mathcal{A} \), with its unique \( C^* \)-algebra structure. Define on \( M_2(\mathcal{A}) \) a linear functional, \( \psi_2 \), by

\[
\psi_2 \left( \begin{pmatrix} A & C \\ B & D \end{pmatrix} \right) = (1/2)(\psi(B) + \overline{\psi(C^*)}).
\]
Then $\psi_2$ is a Hermitian linear functional, that is, $\psi_2(T^*) = \overline{\psi(T)}$ for any $T \in M_2(\mathcal{A})$. Note also that $\|\psi_2\| = 1$.

Let $\psi_2 = (1/2)(\phi^+ - \phi^-)$ be the Jordan decomposition of $\psi_2$ (theorem 4.3.6 and remark 4.3.12 of [13]), so that $\phi^+$ and $\phi^-$ are positive linear functionals on $M_2(\mathcal{A})$ such that $\|\phi^+\| + \|\phi^-\| = 2\|\psi_2\| = 2$. Note that $\psi_2(I_2) = 0$, where $I_2$ is the identity element of $M_2(\mathcal{A})$. It follows that $\phi^+(I_2) = \phi^-(I_2)$, so that $\|\phi^+\| = \|\phi^-\| = 1$. Thus both $\phi^+$ and $\phi^-$ are states on $M_2(\mathcal{A})$.

Let $(\mathcal{H}^+, \pi^+, \xi^+)$ and $(\mathcal{H}^-, \pi^-, \xi^-)$ be the GNS representations for $\phi^+$ and $\phi^-$. Set

$$
\mathcal{H}_2 = \mathcal{H}^+ \oplus \mathcal{H}^- \quad \text{and} \quad \pi_2 = \pi^+ \oplus \pi^-,
$$

and set

$$
\xi_2 = (\xi^+ \oplus \xi^-)/\sqrt{2} \quad \text{and} \quad \eta_2 = (\xi^+ \oplus -\xi^-)/\sqrt{2}.
$$

Note that $\|\xi_2\| = 1 = \|\eta_2\|$. Then for any $T \in M_2(\mathcal{A})$ we have

$$
\psi_2(T) = (1/2)(\phi^+(T) - \phi^-(T)) = (1/2)((\pi^+(T))\xi^+, \xi^+) - (\pi^-(T))\xi^-, \xi^-)
$$

$$\quad = (\pi_2(T))\xi_2, \eta_2).
$$

Now for any $A \in \mathcal{A}$ we have

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

Thus

$$
\psi(A) = \psi_2 \left( \begin{pmatrix}
0 & 0 \\
A & 0
\end{pmatrix} \right) = \langle \pi_2 \left( \begin{pmatrix}
A & 0 \\
0 & 0
\end{pmatrix} \right) \xi_2, \pi_2 \left( \begin{pmatrix}
0 & 1 \\
1 & 0
\end{pmatrix} \right) \eta_2 \rangle.
$$

Let $P = \pi_2((0 0 \\ 1 0))$, and then set $\mathcal{H} = P\mathcal{H}_2$. Notice that $\pi_2((0 0 \\ 1 0))$ carries $\mathcal{H}$ into itself for every $A \in \mathcal{A}$. We obtain in this way a unital *-representation, $\pi$, of $\mathcal{A}$ on $\mathcal{H}$. Set $\xi = P\xi_2$ and $\eta = P\pi_2((0 0 \\ 1 0))\eta_2$. From the equation displayed just above we see that

$$
\psi(A) = \langle \pi(A)\xi, \eta \rangle
$$

for all $A \in \mathcal{A}$. From the definitions given above it is easily seen that $\|\xi\| \leq 1$ and $\|\eta\| \leq 1$. One can restrict $\pi$ to the cyclic $\mathcal{A}$-subspace of $\mathcal{H}$ generated by $\xi$. Then from the fact that $\|\psi\| = 1$ it follows that $\|\xi\| = 1$ and $\|\eta\| = 1$, as desired. If an identity element had been adjoined to the original $\mathcal{A}$, one can restrict $\pi$ further to the cyclic $\mathcal{A}$-subspace of $\mathcal{H}$ generated by $\xi$. \qed
If we now let \( \{ \mathcal{H}, \pi \} \) be the Hilbert-space direct sum over all \( A \in \mathcal{A} \) of the representations obtained in Theorem 3.2, and if we let \( U \) be the direct sum of the corresponding Hermitian unitaries obtained there, then we immediately obtain:

**Corollary 3.4.** Let \( \mathcal{A} \) be a C*-algebra, and let \( \mathcal{B} \) be a C*-subalgebra of \( \mathcal{A} \) that contains a bounded approximate identity for \( \mathcal{A} \). Let \( L \) be defined as above for \( \mathcal{A} \) and \( \mathcal{B} \). Then there is a non-degenerate unitary representation \( \{ \mathcal{H}, \pi \} \) of \( \mathcal{A} \) and a Hermitian unitary operator \( U \) on \( \mathcal{H} \) that commutes with \( \pi(B) \) for every \( B \in \mathcal{B} \), such that

\[
L(A) = (1/2)\| [U, A] \|
\]

for every \( A \in \mathcal{A} \).

Note that if \( L \) is defined as above, and if \( M \) is any seminorm on \( \mathcal{A} \) that satisfies the conditions that \( M(A) \leq \| A \| \) for all \( A \in \mathcal{A} \), and that \( M(B) = 0 \) for all \( B \in \mathcal{B} \), then we have \( M(A) \leq L(A) \) for all \( A \in \mathcal{A} \). Thus we obtain immediately the following corollary, which is of interest in particular for the situation discussed in the introduction in which we have a unital C*-algebra \( \mathcal{D} \) and \( A = M_n(\mathbb{C}) \otimes \mathcal{D} \) while \( \mathcal{B} = M_n(\mathbb{C}) \otimes 1_D \).

**Corollary 3.5.** Let \( \mathcal{A} \) be a C*-algebra, and let \( \mathcal{B} \) be a C*-subalgebra of \( \mathcal{A} \) that contains a bounded approximate identity for \( \mathcal{A} \). Let \( L \) be defined as above. Then \( L \) is the maximal Leibniz seminorm on \( \mathcal{A} \) that takes value 0 on \( \mathcal{B} \) and is dominated by the C*-norm of \( \mathcal{A} \). If \( \mathcal{A} \) is unital and if \( 1_A \in \mathcal{B} \), then \( L \) is the maximal strongly-Leibniz seminorm on \( \mathcal{A} \) that takes value 0 on \( \mathcal{B} \) and is dominated by the C*-norm of \( \mathcal{A} \).

Presumably Arveson’s distance formula provides strongly-Leibniz seminorms on suitable nest algebras, but I have not explored this matter.

We remark that from the proof given above for Theorem 3.2 one can extract a corollary that is closely related to lemma 2.1 of [12] (which deals with finite-dimensional matrix algebras).

Now let the notation be as in Corollary 3.4, and let \( V = iU \), so that \( V \) is a skew-Hermitian operator. Define \( \delta_V \) on \( \mathcal{A} \) by \( \delta_V(A) = (1/2)[V, A] \). Then \( \delta_V \) is a \(*\)-derivation of \( \mathcal{A} \) on \( \mathcal{H} \), where the general definition is:

**Definition 3.6.** Let \( \mathcal{A} \) be a C*-algebra, and let \( \mathcal{H} \) be a Hilbert space. We say that a pair \( (\pi, \delta) \) is a \(*\)-derivation of \( \mathcal{A} \) on \( \mathcal{H} \) if \( \pi \) is a non-degenerate \(*\)-representation of \( \mathcal{A} \) on \( \mathcal{H} \) and \( \delta \) is a linear function from \( \mathcal{A} \) into \( \mathcal{B}(\mathcal{H}) \) such that \( \delta(A^*) = (\delta(A))^* \) and

\[
\delta(AC) = \delta(A)C + A\delta(C)
\]
for all $A, C \in \mathcal{A}$. By the *kernel of $(\pi, \delta)$* we will mean the kernel of $\delta$.

It is known that every $*$-derivation on a C*-algebra is necessarily continuous. (See 4.6.66 of [13] and 4.6.66 of [14].) It is evident that the kernel of a $*$-derivation is a C*-subalgebra of $\mathcal{A}$. If $\mathcal{A}$ is unital then it is easily seen that $\delta(1) = 0$, so that the kernel of $(\pi, \delta)$ is a unital C*-subalgebra of $\mathcal{A}$. Then Corollary 3.4 quickly gives, by means of $\delta_V$, the following characterization of the kernels of $*$-derivations of unital C*-algebras:

**Corollary 3.7.** Let $\mathcal{A}$ be a unital C*-algebra. Then the possible kernels of $*$-derivations of $\mathcal{A}$ on various Hilbert spaces are exactly the unital C*-subalgebras of $\mathcal{A}$.

If $\mathcal{A}$ is not unital, then I do not know whether the kernel of a $*$-derivation must contain a bounded approximate identity for $\mathcal{A}$, and thus I do not know how to characterize all of the possible kernels of $*$-derivations of $\mathcal{A}$.

## 4. Chebyshev subalgebras

We now begin our study of best approximations in C*-algebras. We start by recalling some relevant well-known facts about approximation of Banach-space elements by elements of a closed subspace, and by relating them to our C*-algebra situation. Much of the literature on that subject seems to be motivated by Chebyshev’s famous theorem that a best uniform approximation of a continuous real-valued function on the unit interval by polynomials of degree no greater than a given natural number is unique. (See, e.g., section 7.6 of [7].) Accordingly, when $\mathcal{A}$ is a Banach space and $\mathcal{B}$ is a closed subspace, one says that $\mathcal{B}$ is a **Chebyshev subspace** if every element of $\mathcal{A}$ has a best approximation by an element of $\mathcal{B}$, and that best approximation is unique. For example, every closed subspace of a Hilbert space is a Chebyshev subspace.

Following partial results by A. G. Robertson [19, 20] and a few others, Gert Pedersen showed [16] that if $\mathcal{A}$ is a unital C*-algebra and $\mathcal{B}$ is a unital C*-subalgebra that is a Chebyshev subspace of $\mathcal{A}$, then either $\mathcal{B} = \mathcal{A}$, or $\mathcal{B} = C1_\mathcal{A}$, or $\mathcal{A} = M_2(\mathbb{C})$ and $\mathcal{B}$ is the subalgebra of diagonal matrices.

**Example 4.1.** Let $\mathcal{A}$ be $(M_2(\mathbb{C}))^3$, the C*-algebra of 3-tuples of elements of $M_2(\mathbb{C})$, and let $\mathcal{B}$ be its C*-subalgebra of constant 3-tuples. Let

$$A = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \right\}.$$
Since the distance between the first two entries of $A$ is 2, it is clear that any $B = \begin{pmatrix} t & 0 \\ 0 & 0 \end{pmatrix}$ (viewed as a constant 3-tuple) with $0 \leq t \leq 2$ is a best approximation to $A$ in $\mathcal{B}$. In particular, for $t = 2$ we see that $B$ is a best approximation with $\|B\| > \|A\|$. But since $B$ for $t = 0$ is also a best approximation, $A$ is still same-norm approximatable in $\mathcal{B}$. In the last section of this paper we will give examples that are not same-norm approximatable but have a unique best approximation.

We remark that anyway, in the Banach space case, same-norm approximation can easily fail for Chebyshev subspaces. This already happens for uniform approximation of continuous real-valued functions on the unit interval by polynomials, as is seen when uniformly approximating the function $f(t) = t^{1/n}$ by polynomials of degree 1 or less. As $n$ goes to $\infty$ the norm of the best approximating linear polynomial goes to 2.

However, one striking phenomenon concerning polynomial approximation of continuous real-valued functions on the interval is that the error function equi-oscillates. (See, e.g., theorem 7.6.2 of [7].) That is, if $f \in C_{\mathbb{R}}([0, 1])$ and if $p$ is the best uniform approximation to $f$ by polynomials of degree no greater than $n$, then one can find $n + 2$ distinct points in $[0, 1]$ at which $|f - p|$ takes its maximum value with the sign of $f - p$ evaluated at these points alternating at successive points. We will see an echo of this equi-oscillation later in our C*-algebraic setting.

5. Witnesses for best approximation in Banach spaces

If we want to give specific counter-examples involving best approximations, we need to prove that a candidate best approximation is indeed a best approximation. For this purpose we now discuss further the first basic tool of linear approximation theory [22] that we used already in the second paragraph of the proof of Theorem 3.2. In the next sections we will see that in our C*-algebraic setting this tool has some nice special features.

Let $\mathcal{A}$ be a Banach space, let $\mathcal{B}$ be a proper closed subspace of $\mathcal{A}$, and let $M$ be the quotient norm on $\mathcal{A}/\mathcal{B}$, often viewed as a seminorm on $\mathcal{A}$. Let $A \in \mathcal{A}$. Denote by $\mathcal{A}'$ the Banach-space dual of $\mathcal{A}$, and by $\mathcal{A}'_1$ its unit ball. Then, much as in the proof of Theorem 3.2, by the Hahn-Banach theorem applied to the image of $A$ in $\mathcal{A}/\mathcal{B}$ there is a $\psi \in \mathcal{A}'$, such that $\|\psi\| = 1$, $\psi \in B^\perp$ (in the sense that $\psi(B) = 0$ for all $B \in \mathcal{B}$) and $\psi(A) = M(A)$. It may well happen that $\psi$ is not unique.
Suppose that $A$ has a best approximation, $B$, in $\mathcal{B}$. Then $\|A - B\| = M(A) = \psi(A)$. The following proposition is a basic tool for proving that one has a best approximation.

**Proposition 5.1.** Let $A$, $B$, and $M$ be as above, and let $A \in \mathcal{A}$ and $B \in \mathcal{B}$ be given. Let $\psi \in \mathcal{A}_1'$ be given. If $\psi \in \mathcal{B}^\perp$ and $\psi(A) = \|A - B\|$, then $B$ is a best approximation to $A$ in $\mathcal{B}$.

**Proof.** Note that necessarily $\|\psi\| = 1$, as long as $A \notin \mathcal{B}$. For any $C \in \mathcal{B}$ we have

$$\psi(A - C) = \psi(A) = \|A - B\|.$$  

It follows that $\|A - C\| \geq \|A - B\|$.

Thus we can say that $\psi$ is a “witness” for the fact that $B$ is a best approximation to $A$ in $\mathcal{B}$. That is, if we can find such a $\psi$ then we have a proof that $B$ is a best approximation.

As suggested by our comments above concerning equi-oscillation, it is useful to look at the error of an approximation, that is, at $Z = A - B$. If $B$ is a best approximation to $A$ in $\mathcal{B}$, then it is clear that $0$ is a best approximation to $Z$ in $\mathcal{B}$ (thus our use of the symbol $Z$, but see also its use in [9, 1]), and if $\psi$ is a witness for $B$ as above, then we have $\psi(Z) = \psi(A) = \|Z\|$.

**Definition 5.2.** With $A$ and $\mathcal{B}$ as above, we say that an element $Z \in \mathcal{A}$ is **minimal for $\mathcal{B}$**, or **$\mathcal{B}$-minimal**, if $0$ is a best approximation to $Z$ in $\mathcal{B}$, so that $M(Z) = \|Z\|$, and $M(Z + B) = \|Z\|$ for every $B \in \mathcal{B}$.

Then a witness to the $\mathcal{B}$-minimality of $Z$ will be a $\psi \in \mathcal{A}_1'$ such that $\psi \in \mathcal{B}^\perp$ and $\psi(Z) = \|Z\|$. Note that for any $B \in \mathcal{B}$ we will have

$$M(Z + B) = \|Z\| = \psi(Z).$$

This can be of help in calculating $M(A)$ for $A \in \mathcal{A}$.

Suppose now that $\mathcal{B}$ is finite-dimensional, so that best approximations always exist. If $\mathcal{A}$ is over the complex numbers, we will forget that and view $\mathcal{A}$ as being over $\mathbb{R}$. This has no effect on best approximations, but dimensions will usually be over $\mathbb{R}$ as this simplifies a bit the bookkeeping.

Recall Carathéodory’s theorem (see, e.g., exercise 19 of chapter 3 of [21]) that if $\mathcal{C}$ is a finite-dimensional Banach space of dimension $q$ and if $K$ is a closed bounded convex subset of $\mathcal{C}$, with $K_\varepsilon$ its set of extreme points, then $K = \text{convex}(K_\varepsilon)$, with no need to take closure, and in fact that every element of $K$ is a convex combination of at most $q + 1$ elements of $K_\varepsilon$. Furthermore, one can show that if the element is in the boundary of $K$ then at most $q$ elements of $K_\varepsilon$ are needed. (See lemma 1.1 of chapter II of [22].)
Suppose now that \( A \in A \) but \( A \notin B \). Let \( F = B \oplus \mathbb{R}A \), a subspace of \( A \), equipped with the norm from \( A \). Let \( p = \text{dim}_{\mathbb{R}}(B) \), so that \( \text{dim}_{\mathbb{R}}(F) = p + 1 \). When \( A \) is viewed as an element of \( F \), its best approximations in \( B \) are the same as the best approximations when \( A \) is viewed as an element of \( A \). Suppose now that \( B \in B \) is a best approximation to \( A \) in \( B \). By the results stated above there is a \( \psi \in F' \) such that \( \psi \in B^1 \) and \( \psi(A) = \| A - B \| \). Since \( \psi \) is in the boundary of the compact convex set \( F'_1 \), there exist extreme points, \( \psi_1, \ldots, \psi_k \) of \( F'_1 \), with \( k \leq p + 1 \), and there exist positive real numbers \( t_1, \ldots, t_k \) with \( \sum t_j = 1 \) such that \( \psi = \sum t_j \psi_j \). Each \( \psi_j \) can be extended to an extreme point of \( A'_1 \). We denote these extensions again by \( \psi_j \). Then we extend \( \psi \) to \( A \) by setting \( \psi = \sum t_j \psi_j \). Note that \( \| \psi \| = 1 \), \( \psi \in B^1 \), and \( \psi(A) = \| A - B \| \). Since \( | \psi_j(A - B) | \leq \| A - B \| \) for each \( j \), we must have \( \psi_j(A - B) = \| A - B \| \) for each \( j \). (But we can not expect that \( \psi_j \in B^1 \).) In this way we obtain:

**Proposition 5.3.** Let \( A \) and \( B \) be as above, with \( p = \text{dim}(B) \). Let \( A \in A \) and \( B \in B \) be given. Then \( B \) is a best approximation to \( A \) in \( B \) if and only if there are \( k \) extreme points, \( \psi_1, \ldots, \psi_k \), of \( A'_1 \), with \( k \leq p \), and there are positive real numbers \( t_1, \ldots, t_k \) with \( \sum t_j = 1 \), such that when we set \( \psi = \sum t_j \psi_j \), we have \( \psi \in A'_1 \), \( \psi \in B^1 \), and \( \psi(A) = \| A - B \| \). Furthermore, \( \psi_j(A) = \| A - B \| \) for each \( j \).

6. **Witnesses for best approximation in \( C^* \)-algebras**

In this section we assume for simplicity that \( A \) is a unital \( C^* \)-algebra and that \( B \) is a unital \( C^* \)-subalgebra of \( A \) (with \( 1_A \in B \)). We let \( A^h \) denote the subspace of Hermitian elements of \( A \), and similarly for \( B^h \).

It is easily seen that if \( A \in A^h \) and \( A \) has a best approximation \( B \) in \( B \), then it has a best approximation in \( B \) that is actually in \( B^h \) and of no greater norm. If \( A \neq A^* \), then we can consider the Hermitian element \( \begin{pmatrix} 0 & A^* \\ A & 0 \end{pmatrix} \) in \( M_2(A) \). If this element has a best approximation in \( M_2(B) \), then it is easily seen that it has a best approximation in \( M_2(B) \) that is of the form \( \begin{pmatrix} 0 & B^* \\ B & 0 \end{pmatrix} \) for \( B \in B \) of no greater norm, so that \( B \) is a best approximation to \( A \) in \( B \) of no greater norm. In this way we can reduce the study of best approximation in \( C^* \)-algebras to that for Hermitian elements. One does not need to take this path — one can, for example, instead work directly with elements of \( A'_1 \), much as we did in Section 3. But the path via Hermitian elements involves somewhat more familiar arguments, using states, as we will see. Note that \( A^h \) is a vector space over \( \mathbb{R} \).
We now apply the results of the previous section. We seek witnesses for elements of $A \in A^h$ to be $B$-minimal, since if $A \in A^h$ and $B \in B^h$ then $B$ is a best approximation to $A$ in $B$ exactly if $A - B$ is $B$-minimal. So suppose now that $Z \in A^h$ (with $Z \neq 0$), and that $Z$ is $B$-minimal. By rescaling, we see that it is sufficient to treat the case in which $\|Z\| = 1$. We now assume this, as it slightly simplifies the bookkeeping.

Then, as seen in the previous section, there exists a $\psi \in (A^h)^\perp$ such that $\psi(Z) = 1$. Note that necessarily $\|\psi\| = 1$. We can extend $\psi$ (uniquely) to be a Hermitian element of $A'$ (of same norm).

We are now exactly at the point where we can take advantage of our C*-algebra setting. Let $\psi = (\psi^+ - \psi^-)/2$ be the Jordan decomposition of $\psi$, as discussed, for example, in theorem 4.3.6 and remark 4.3.12 of [13]. Thus $\psi^+$ and $\psi^-$ are positive linear functionals on $A$ that are orthogonal in the sense that

$$\|\psi^+\| + \|\psi^-\| = \|\psi^+ - \psi^-\| = 2\|\psi\| = 2.$$ 

Now by assumption $1_A \in B$, and so

$$0 = \psi(1_A) = (\psi^+(1_A) - \psi^-(1_A))/2,$$

so that $\psi^+(1_A) = \psi^-(1_A)$, and consequently $\|\psi^+\| = \|\psi^-\|$. It follows that each of $\psi^+$ and $\psi^-$ is a state of $A$. Now

$$2 = 2\psi(Z) = \psi^+(Z) - \psi^-(Z).$$

Since $\|Z\| = 1$, it follows that $\psi^+(Z) = 1$ while $\psi^-(Z) = -1$. This is an echo of the Chebyshev equi-oscillation phenomenon mentioned earlier, with more echo to come in the next section. Because $\psi^-(Z) = -1$ we have

$$\psi^-((Z + 1_A)^2) = \psi^-(Z^2) + 2\psi^-(Z) + 1 = \psi^-(Z^2) - 1.$$ 

Since the left-hand side is clearly non-negative and the right-hand side is clearly non-positive, we see that

$$\psi^-((Z + 1_A)^2) = 0 \quad \text{and} \quad \psi^-(Z^2) = +1 = (\psi^-(Z))^2.$$ 

A similar calculation shows that

$$\psi^+((Z - 1_A)^2) = 0 \quad \text{and} \quad \psi^+(Z^2) = +1 = (\psi^+(Z))^2.$$ 

This means that each of $\psi^+$ and $\psi^-$ is “definite” on $Z$, as defined, for example, in exercise 4.6.16 of [13] (and see also 4.6.16 of [14]). I thank Dick Kadison for pointing out to me the relevance of this definition to the present situation. Another way of expressing this definiteness is that the “mean-square deviation” of $Z$ for $\psi^+$ and for $\psi^-$, as considered
in remark 2.2.23(4) of [24], is 0. As follows from exercise 4.6.16 of [13],
this definiteness implies that
\[ \psi^+(ZC) = \psi^+(C) \quad \text{and} \quad \psi^-(ZC) = -\psi^-(C) \]
for all \( C \in A \), as can be seen by applying the Cauchy-Schwarz inequality
to \( \psi^+((Z - 1_A)C) \), and similarly for \( \psi^- \).

Let us now set \( \varphi = (\psi^+ + \psi^-)/2 \), which is a state on \( A \). In analogy
with measure theory, we could call \( \varphi \) the "total variation" of \( \psi \) and
write \( \varphi = |\psi| \). From the calculations done some lines above we see
that \( \varphi(Z^2) = 1 \) (from which one can show that \( \varphi \) is definite on \( Z^2 \)).
But we also see from these calculations that for every \( B \in B \)
we have
\[ 2\varphi(ZB) = \psi^+(ZB) + \psi^-(ZB) = \psi^+(B) - \psi^-(B) = \psi(B) = 0, \]
so that \( \varphi(ZB) = 0 \) for all \( B \in B^h \). Since \( Z \) is Hermitian, it follows
that also \( \varphi(BZ) = 0 \) for all \( B \in B^h \), and so
\[ \varphi(ZB + B^*Z) = 0 \]
for all \( B \in B \). Note that \( ZB + B^*Z \) is Hermitian. We arrive in this
way at one direction of the proof of the next theorem, which is essentially theorem 2.2 of [1], which in turn has antecedents in section 5 of [9]. The authors of these two papers have a purpose quite dif-
ferent from ours, namely to understand geodesics in the homogeneous spaces \( U_A/U_B \) where \( U_A \) denotes the unitary group of the C*-algebra \( A \), and similarly for \( U_B \). They use the Finsler metric from the operator
norm. The authors give a direct proof of the next theorem, a proof that
is somewhat more complicated than what we have done above, and they provide almost no indication that this theorem has anything to
do with best approximation, much less that it relates strongly to basic central results in the literature on best approximation along the lines
that we have discussed above. The authors do examine some inter-
esting specific finite-dimensional examples in which \( A \) is a full matrix
algebra and \( B \) is the subalgebra of diagonal matrices, or a subalgebra
of block-diagonal matrices.

**Theorem 6.1.** Let \( A \) be a unital C*-algebra and let \( B \) be a unital C*-
subalgebra of \( A \). Let \( A \in A^h \). If \( A \) is \( B \)-minimal, then there exists a
state \( \varphi \) of \( A \) such that \( \varphi(A^2) = ||A||^2 \) and \( \varphi(AB + B^*A) = 0 \) for all
\( B \in B \). Conversely, if \( A \in A^h \) and if there is a state \( \varphi \) of \( A \) such
that \( \varphi(A^2) = ||A||^2 \) and \( \varphi \) satisfies the slightly weaker condition that
\( \varphi(AB + BA) = 0 \) for all \( B \in B^h \), then \( A \) is \( B \)-minimal.

**Proof.** By scaling we can assume that \( ||A|| = 1 \). We have given above
the proof that if \( A \) is \( B \)-minimal then there exists a state \( \varphi \) satisfying
the given conditions. Conversely, suppose that there exists a state \( \varphi \)
that satisfies the slightly weaker conditions. Then for any \( B \in \mathcal{B}^h \) we have
\[
\varphi((A - B)^*(A - B)) = \varphi(A^2) - \varphi(BA + AB) + \varphi(B^2) \\
= \|A\|^2 + \varphi(B^2) \geq \|A\|^2.
\]
Consequently \( \|A - B\|^2 \geq \|A\|^2 \), so that \( A \) is \( \mathcal{B} \)-minimal. \( \square \)

Thus a state satisfying the (weaker) conditions of this theorem can serve as a witness for the \( \mathcal{B} \)-minimality of \( A \). Since \( \varphi \) is a state and \( A^2 \) is positive, this may be more convenient to use than our earlier \( \psi \).

We remark that the condition that \( \varphi(AB + BA) = 0 \) for all \( B \in \mathcal{B}^h \) is natural from the following point of view. Let \( \varphi \) be any state of \( \mathcal{A} \), let \( B \in \mathcal{B} \) be given, and define a function, \( f \), on \( \mathbb{R} \) by
\[
f(t) = \varphi((A + tB)^2).
\]
If \( A \) is \( \mathcal{B} \)-minimal, we can hope that \( f \) takes a minimum value at \( t = 0 \). But
\[
f'(t) = \varphi(AB + BA) + 2t\varphi(B^2),
\]
and so if \( f'(0) = 0 \) then \( \varphi(AB + BA) = 0 \). Notice that \( f''(0) = 2\varphi(B^2) \geq 0 \). But these calculations do not provide part of a proof for the theorem.

The proof of Theorem 6.1 can be used to give some information about the uniqueness of best approximations:

**Theorem 6.2.** Let \( \mathcal{A} \) be a unital \( C^* \)-algebra and let \( \mathcal{B} \) be a unital \( C^* \)-subalgebra of \( \mathcal{A} \). Let \( A \in \mathcal{A}^h \). Suppose that \( \varphi \) is a state of \( \mathcal{A} \) such that \( \varphi(A^2) = \|A\|^2 \), and that \( \varphi \) satisfies the condition that \( \varphi(AB + BA) = 0 \) for all \( B \in \mathcal{B}^h \), so that \( A \) is \( \mathcal{B} \)-minimal. If the restriction of \( \varphi \) to \( \mathcal{B} \) is faithful, then 0 is the unique best approximation to \( A \) in \( \mathcal{B} \).

**Proof.** The last display of the proof of Theorem 6.1 gives
\[
\varphi((A - B)^*(A - B)) = \|A\|^2 + \varphi(B^2).
\]
Consequently if the restriction of \( \varphi \) to \( \mathcal{B} \) is faithful, then for \( B \in \mathcal{B}^h \) with \( B \neq 0 \) we have \( \varphi(B^2) > 0 \), so that \( \|A - B\| > \|A\| \). \( \square \)

**Corollary 6.3.** Let \( \mathcal{A} \) be a unital \( C^* \)-algebra and let \( \mathcal{B} \) be a unital \( C^* \)-subalgebra of \( \mathcal{A} \). Let \( A \in \mathcal{A}^h \). Suppose that \( A \) has a best approximation, \( B \), in \( \mathcal{B}^h \). Let \( \varphi \) be a state of \( \mathcal{A} \) such that \( \varphi((A - B)^2) = \|A - B\|^2 \) and \( \varphi((A - B)D + D(A - B)) = 0 \) for all \( D \in \mathcal{B}^h \). (Such a state is guaranteed to exist by Theorem 6.1.) If the restriction of \( \varphi \) to \( \mathcal{B} \) is faithful, then \( B \) is the unique best approximation to \( A \) in \( \mathcal{B} \).
Thus in the case when \( \varphi \) is not unique, it can be useful to seek a \( \varphi \) that is faithful, or at least has support as large as possible.

Recall that we are interested in upper bounds on the norms of best approximations. Knowing \( \varphi(B^2) \) gives us a lower bound on \( \|B\| \), but that may not be very useful.

For the present special context we can give a considerably simpler proof of the key part of Theorem 3.2:

**Corollary 6.4.** Let \( \mathcal{A} \) be a unital C*-algebra and let \( \mathcal{B} \) be a unital C*-subalgebra of \( \mathcal{A} \). Let \( A \in \mathcal{A}^b \). Suppose that \( A \) has a best approximation, \( B \), in \( \mathcal{B}^b \). Then there is a non-degenerate \(*\)-representation, \((\mathcal{H}, \pi)\), of \( \mathcal{A} \), and a Hermitian unitary operator \( U \in \mathcal{L}(\mathcal{H}) \), such that

\[
L(A) = \left(\frac{1}{2}\right)\|[U, A]\| \quad \text{while} \quad \left(\frac{1}{2}\right)\|[C, U]\| \leq L(C)
\]

for all \( C \in \mathcal{A} \).

**Proof.** Let \( Z = A - B \), so that \( Z \) is \( \mathcal{B} \)-minimal. Let \( \varphi \) be the state whose existence is guaranteed by Theorem 6.1, and let \( (\mathcal{H}, \pi, \xi) \) be the GNS representation for \( \varphi \). Let \( P \) be the orthogonal projection of \( \mathcal{H} \) onto the closure of \( \pi(\mathcal{B})\xi \). As before, it is clear that we have \([P, \pi(B)] = 0\) for all \( B \in \mathcal{B} \). Notice that \( \pi(Z)\xi \perp \pi(\mathcal{B})\xi \) because \( Z \) is Hermitian and \( \varphi(ZB) = 0 \) for all \( B \in \mathcal{B} \). Thus \( P(Z\xi) = 0 \). Also, because \( \mathcal{B} \) is unital we see that \( \xi \) is in \( \pi(\mathcal{B})\xi \), so that \( P(\xi) = \xi \). Finally, \( \|Z\xi\| = \|Z\| \) because \( \varphi(Z^2) = \|Z\|^2 \). Set \( \eta = \pi(Z)\xi \), and let \( U = 2P - I \). Then by the same calculations as done in the proof of Theorem 3.2 we find that

\[
L(A) = \left(\frac{1}{2}\right)\|[U, A]\| \quad \text{while} \quad \left(\frac{1}{2}\right)\|[C, U]\| \leq L(C)
\]

for all \( C \in \mathcal{A} \).

\( \square \)

7. The case when \( \mathcal{B} \) is finite-dimensional

In this section we assume that \( \mathcal{A} \) is a unital C*-algebra and that \( \mathcal{B} \) is a unital C*-subalgebra that is finite-dimensional. In this case best approximations always exist. We let \( p = \dim(\mathcal{B}) \). (We remark that if, instead, \( \mathcal{B} \) is a closed two-sided ideal of \( \mathcal{A} \), then again best approximations always exist. This is an immediate consequence of proposition II.5.1.5 of [4]. But I have not seen how to make use of this fact in the present context.)

The case in which \( \mathcal{A} \) itself is finite-dimensional is already quite interesting. In particular, for any natural numbers \( m \) and \( n \) we can set \( \mathcal{A} = (M_n(\mathbb{C}))^m \), the C*-algebra of \( m \)-tuples of elements of \( M_n(\mathbb{C}) \). Equally well, \( \mathcal{A} \) can be viewed as \( M_n(\mathcal{D}) \) where \( \mathcal{D} \) is the commutative C*-algebra of functions on a set with \( m \) elements. Then, along the lines of our comments in the introduction concerning matricial norms,
it is natural to take \( B \) to be the C*-subalgebra of constant \( m \)-tuples, so that \( B \cong M_n(\mathbb{C}) \). This approximation problem can be viewed as the problem of finding a center of a smallest ball in \( M_n(\mathbb{C}) \) containing the \( m \) entries of a given element of \( A \). One might try taking the average of the \( m \) elements, or at least expect that the center will be in the convex hull of the \( m \) elements. But there is a striking theorem of Garvaki [11] that says that if \( V \) is a Banach space of dimension at least 3, and if for every three points of \( V \) there exists a center for a ball of smallest radius containing the three points, such that this center lies in a plane containing the three point, then \( V \) is a Hilbert space. Thus in our C*-algebra setting we should not expect that centers will even be in the convex hull of the points. Interested readers can try their hand at finding a center (that is, a best approximation in \( B \)) for the case in which \( m = 3 \), \( n = 2 \), and

\[
A = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right\}.
\]

This example is small enough that it is not difficult, but it does seem to illustrate fairly well some of the challenges of this kind of problem.

Let us now apply some of the results of Section 5 to refine those of Section 6. Let \( Z \) be an element of \( A^h \) that is \( B \)-minimal. We assume as before that \( \|Z\| = 1 \). Much as in the discussion leading to Proposition 5.3, we set \( F = B^h \oplus \mathbb{R}Z \). Then, as seen in that discussion, there exist extreme points \( \psi_1, \ldots, \psi_k \) of \( F_1 \) with \( k \leq p + 1 \), and there exist positive real numbers \( t_1, \ldots, t_k \) with \( \sum t_j = 1 \), such that when we set

\[
\psi = \sum t_j \psi_j
\]

we have \( \psi(Z) = \|Z\| = 1 \) and \( \psi \in (B^h)^1 \). Each \( \psi_j \) can be extended to be an extreme point of \( (A^h)_1 \) and extended further to be a Hermitian extreme point of \( A'_1 \). We denote these extensions again by \( \psi_j \). Then we extend \( \psi \) to \( A \) by setting \( \psi = \sum t_j \psi_j \). Note that \( \|\psi\| = 1 \), \( \psi(Z) = 1 \), and \( \psi \in B^\perp \).

Of course

\[
1 = \|Z\| = \psi(Z) = \sum t_j \psi_j(Z),
\]

and since \( |\psi_j(Z)| \leq 1 \) for each \( j \), it follows that \( \psi_j(Z) = 1 \) for each \( j \). But when we form the Jordan decompositions of the \( \psi_j \)'s, theorem 4.3.6 of [13], much as we did in the previous section, the fact that each \( \psi_j \) is an extreme point tells us that either \( \psi_j \) is a pure state of \( A \) or the negative of a pure state. Thus for each \( j \) there is an integer \( \varepsilon_j \) which is either +1 or −1 such that \( \varepsilon_j \psi_j \) is a pure state. We set \( \varphi_j = \varepsilon_j \psi_j \) for each \( j \). Thus each \( \varphi_j \) is a pure state, and \( \varphi_j(Z) = \varepsilon_j \) for each \( j \). This is again an echo of the Chebyshev equi-oscillation phenomenon mentioned earlier, with a bit more echo to come. Of course \( \psi = \sum t_j \varphi_j \).
Because \( \varphi_j(Z) = \varepsilon_j \) for each \( j \), we can argue exactly as in the previous section to conclude that for each \( j \)

\[
\varphi_j(Z^2) = 1 = \|\varphi_j(Z)\|^2,
\]

and that consequently \( \varphi_j \) is “definite” on \( Z \), so that \( \varphi_j(ZC) = \varepsilon_j \varphi_j(C) \) for every \( C \in \mathcal{A} \).

Let us now set \( \varphi = \sum t_j \varphi_j \), so that \( \varphi \) is a state of \( \mathcal{A} \), not usually pure. Then for every \( B \in \mathcal{B} \) we have

\[
\varphi(ZB) = \sum t_j \varphi_j(ZB) = \sum t_j \varepsilon_j \varphi_j(B) = \sum t_j \psi_j(B) = \psi(B) = 0.
\]

Since \( Z \) is Hermitian, this implies, as in the previous section, that \( \varphi(ZB + B^*Z) = 0 \) for all \( B \in \mathcal{B} \). We are now exactly in the situation for one direction of Theorem 6.1, except that we now have the additional information about the decomposition of \( \varphi \) into pure states that are definite on \( Z \).

It is natural to set

\[
\psi^+ = 2 \sum \{ t_j \psi_j : \varepsilon_j = +1 \} \quad \text{and} \quad \psi^- = 2 \sum \{ t_j \psi_j : \varepsilon_j = -1 \},
\]

for then \( \psi = (\psi^+ - \psi^-)/2 \), while \( \psi^+ \) and \( \psi^- \) are positive and \( \varphi = (\psi^+ + \psi^-)/2 \). Since \( 1_\mathcal{A} \in \mathcal{B} \) so that \( \psi(1_\mathcal{A}) = 0 \), we see that \( \psi^+(1_\mathcal{A}) = \psi^-(1_\mathcal{A}) \), so that \( \psi^+ \) and \( \psi^- \) must be states. Since \( \|\psi\| = 1 \), we see that \( \psi^+ \) and \( \psi^- \) are orthogonal, and so give the Jordan decomposition of \( 2\psi \). It follows easily that \( \psi^+(Z) = 1 \) while \( \psi^-(Z) = -1 \), so that each of \( \psi^+ \) and \( \psi^- \) is definite on \( Z \), as is to be expected.

We summarize part of the above discussion as follows:

**Theorem 7.1.** Let \( \mathcal{A} \) be a unital \( C^* \)-algebra and let \( \mathcal{B} \) be a unital \( C^* \)-subalgebra of \( \mathcal{A} \) that is finite-dimensional, of dimension \( p \). Let \( Z \) be an element of \( \mathcal{A}^b \) that is \( \mathcal{B} \)-minimal. Then there exist pure states \( \varphi_1, \ldots, \varphi_k \) of \( \mathcal{A}'_1 \) with \( k \leq p + 1 \), such that each \( \varphi_j \) is definite on \( Z \), and there are positive real numbers \( t_1, \ldots, t_k \) with \( \sum t_j = 1 \), such that when we set \( \varphi = \sum t_j \varphi_j \) then \( \varphi(Z^2) = \|Z\|^2 \) and \( \varphi(ZB + B^*Z) = 0 \) for all \( B \in \mathcal{B} \).

Thus we are assured that we can always find a witness, \( \varphi \), for the minimality of \( Z \) that is expressed in terms of pure states definite on \( Z \) in the way stated in the theorem. This suggests that one way to find a witness \( \varphi \) for the minimality of \( Z \) is to examine the pure states that are definite on \( Z \). We use this approach in the next section.
8. The failure of the same-norm property

Example 8.1. As earlier, we let \( \mathcal{A} = (M_2(\mathbb{C}))^3 \), and we let \( \mathcal{B} \) be its subalgebra of constant 3-tuples. We will actually just work with real Hermitian matrices, and it is easily seen that if \( A \in \mathcal{A}^h \) and if its 3 entries are all real matrices, then there will be a best approximation to \( A \) in \( \mathcal{B} \) that is a real symmetric matrix. We let

\[
Z = \left\{ \begin{pmatrix} 2 & 0 \\ 0 & 5 \end{pmatrix}, \begin{pmatrix} 4 & -3 \\ -3 & -4 \end{pmatrix}, \begin{pmatrix} 4 & 3 \\ 3 & -4 \end{pmatrix} \right\}.
\]

We do not require that \( \|Z\| = 1 \) so that we can work with integer entries. Then

\[
Z^2 = \left\{ \begin{pmatrix} 4 & 0 \\ 0 & 25 \end{pmatrix}, \begin{pmatrix} 25 & 0 \\ 0 & 25 \end{pmatrix}, \begin{pmatrix} 25 & 0 \\ 0 & 25 \end{pmatrix} \right\}.
\]

We now use the tools developed in the previous section to prove that \( Z \) is minimal. There are 5 pure states on \( \mathcal{A} \) that are definite on \( Z \) and take value 25 on \( Z^2 \). Each of the entries of \( Z \) is self-adjoint, and we use the pure states corresponding to the eigenvectors of the entries for the eigenvalues \( \pm 5 \) of \( Z \). We denote these states by \( \varphi^j_\varepsilon \) where \( j \) corresponds to the entry used, so \( j = 1, 2, 3 \), and \( \varepsilon = + \) or \( - \) depending on the sign of the corresponding eigenvalue. For instance, \( \varphi^1_+ \) is the state that acts on the first entry of the elements of \( \mathcal{A} \) and is determined by the eigenvector \( \begin{pmatrix} 1 \\ 1 \end{pmatrix} \) of the first entry of \( Z \), while \( \varphi^2_- \) is the state that acts on the second entry of elements of \( \mathcal{A} \) and is determined by the eigenvector for the eigenvalue \(-5\) of the second entry of \( Z \). Thus \( \varphi^2_-(Z) = -5 \).

We seek \( \psi \) and \( \varphi \) with \( \varphi = |\psi| \), satisfying the conditions of Theorem 7.1, expressed as a convex combination of the 5 pure states. Thus we seek coefficients that must be non-negative and sum to 1. The subspace of \( \mathcal{B}^h \) consisting of real matrices has dimension 3, and so we expect to need only 4 of the 5 pure states. This indicates that \( \psi \) and \( \varphi \), if they exist, may not be unique, and calculations show that this is indeed the case. In particular, calculations show that there do exist solutions, and that one of them is

\[
\psi = (1/18)(8\varphi^1_+ + \varphi^2_+ - 4\varphi^-_+ - 5\varphi^3_-),
\]

so that \( \psi^+ = (1/9)(8\varphi^1_+ + \varphi^2_+) \) and \( \psi^- = (1/9)(4\varphi^-_+ + 5\varphi^3_-) \) (for our conventions), and \( \varphi = (\psi^+ + \psi^-)/2 \). Once found, it is not hard to check that this is in fact a solution. Thus \( Z \) is minimal. Furthermore, \( \varphi \) restricted to \( \mathcal{B} \) is faithful, because here this amounts to the fact that the set of eigenvectors defining the pure states used above to express \( \psi \) spans \( \mathbb{R}^2 \). Then Theorem 6.2 tells us that 0 is the unique nearest element to \( Z \) in \( \mathcal{B} \).
Now let $B = \begin{pmatrix} -8 & 0 \\ 0 & 0 \end{pmatrix}$, viewed as a constant 3-tuple in $B^h$. Let $A = Z + B$. It is easily seen that $\|A\| = 7$. But by the uniqueness of the best approximation to $Z$, the best approximation to $A$ in $B$ is unique, and is just $B$. But $\|B\| = 8$. Thus we see that $A$ has no best approximation of norm no bigger than $\|A\|$. We see in this way the failure of the same-norm approximation property.

References

[1] Esteban Andruchow, Luis E. Mata-Lorenzo, Alberto Mendoza, Lázaro Recht, and Alejandro Varela, *Minimal matrices and the corresponding minimal curves on flag manifolds in low dimension*, Linear Algebra Appl. 430 (2009), no. 8-9, 1906–1928. MR 2503942 (2010b:46133)

[2] William Arveson, *Interpolation problems in nest algebras*, J. Functional Analysis 20 (1975), no. 3, 208–233. MR 0383098 (52 #3979)

[3] Rajendra Bhatia and Peter Šemrl, *Orthogonality of matrices and some distance problems*, Linear Algebra Appl. 287 (1999), no. 1-3, 77–85, Special issue celebrating the 60th birthday of Ludwig Elsner. MR 1662861 (99k:15042)

[4] B. Blackadar, *Operator algebras*, Encyclopaedia of Mathematical Sciences, vol. 122. Springer-Verlag, Berlin, 2006, Theory of $C^*$-algebras and von Neumann algebras, Operator Algebras and Non-commutative Geometry, III. MR 2188261 (2006k:46082)

[5] Man-Duen Choi and Chi-Kwong Li, *The ultimate estimate of the upper norm bound for the summation of operators*, J. Funct. Anal. 232 (2006), no. 2, 455–476. MR 2200742 (2006j:47010)

[6] Erik Christensen, *Perturbations of operator algebras. II*, Indiana Univ. Math. J. 26 (1977), no. 5, 891–904. MR 0512368 (58 #23628b)

[7] Philip J. Davis, *Interpolation and approximation*, Blaisdell Publishing Co. Ginn and Co. New York-Toronto-London, 1963. MR 0157156 (28 #393)

[8] Ronald G. Douglas, *Banach algebra techniques in operator theory*, second ed., Graduate Texts in Mathematics, vol. 179, Springer-Verlag, New York, 1998. MR 1634900 (99c:46040)

[9] Carlos E. Durán, Luis E. Mata-Lorenzo, and Lázaro Recht, *Metric geometry in homogeneous spaces of the unitary group of a $C^*$-algebra. II. Geodesics joining fixed endpoints*, Integral Equations Operator Theory 53 (2005), no. 1, 33–50. MR 2183595 (2007a:58006)

[10] P. Gajendragadkar, *Norm of a derivation on a von Neumann algebra*, Trans. Amer. Math. Soc. 170 (1972), 165–170. MR 0305090 (46 #4220)

[11] A. L. Garkavi, *On the Čebyšev center and convex hull of a set*, Uspehi Mat. Nauk 19 (1964), no. 6 (120), 139–145. MR 0175035 (30 #5221)

[12] D. Hadwin, D. R. Larson, and D. Timotin, *Approximation theory and matrix completions*, Linear Algebra Appl. 377 (2004), 165–179. MR 2021609 (2004j:47028)

[13] R. V. Kadison and J. R. Ringrose, *Fundamentals of the theory of operator algebras. Vol. I*, American Mathematical Society, Providence, RI, 1997, Elementary theory, Reprint of the 1983 original. MR 98f:46001a
[14] ———, Fundamentals of the theory of operator algebras. Vol. III, Birkhäuser Boston Inc., Boston, MA, 1991, Special topics, Elementary theory—an exercise approach. MR 1134132 (92m:46084)

[15] ———, Fundamentals of the theory of operator algebras. Vol. II, Graduate Studies in Mathematics, vol. 16, American Mathematical Society, Providence, RI, 1997, Advanced theory, Corrected reprint of the 1986 original. MR 1468230 (98f:46001b)

[16] Gert Kjaergård Pedersen, Čebyšev subspaces of $C^\ast$-algebras, Math. Scand. 45 (1979), no. 1, 147–156. MR 567440 (81i:46077)

[17] Marc A. Rieffel, Vector bundles and Gromov-Hausdorff distance, J. K-Theory 5 (2010), 39–103, arXiv:math.MG/0608266.

[18] ———. Leibniz seminorms for “Matrix algebras converge to the sphere”, Quanta of Maths, 543–578, Clay Mathematics Proceedings, vol. 11, Amer. Math. Soc., Providence, R.I., 2011, arXiv:0707.3229.

[19] A. Guyan Robertson, Best approximation in von Neumann algebras, Math. Proc. Cambridge Philos. Soc. 81 (1977), no. 2, 233–236. MR 0473860 (57 #13519)

[20] A. Guyan Robertson and David Yost, Chebyshev subspaces of operator algebras, J. London Math. Soc. (2) 19 (1979), no. 3, 523–531. MR 540068 (80j:46097)

[21] Walter Rudin, Functional analysis, second ed., International Series in Pure and Applied Mathematics, McGraw-Hill Inc., New York, 1991. MR 1157815 (92k:46001)

[22] Ivan Singer, Best approximation in normed linear spaces by elements of linear subspaces, Die Grundlehren der mathematischen Wissenschaften, Band 171, Publishing House of the Academy of the Socialist Republic of Romania, Bucharest, 1970. MR 0270044 (42 #4937)

[23] Joseph G. Stampfli, The norm of a derivation, Pacific J. Math. 33 (1970), 737–747. MR 0265952 (42 #861)

[24] Walter Thirring, A course in mathematical physics. Vol. 3, Springer-Verlag, New York, 1981, Quantum mechanics of atoms and molecules, Translated from the German by Evans M. Harrell, Lecture Notes in Physics, 141. MR 625662 (84m:81006)

[25] Wei Wu, Non-commutative metrics on matrix state spaces, J. Ramanujan Math. Soc. 20 (2005), no. 3, 215–254, arXiv:math.OA/0411475. MR 2181130

[26] ———, Non-commutative metric topology on matrix state space, Proc. Amer. Math. Soc. 134 (2006), no. 2, 443–453, arXiv:math.OA/0410587. MR 2176013 (2006f:46072)

[27] ———, Quantized Gromov-Hausdorff distance, J. Funct. Anal. 238 (2006), no. 1, 58–98, arXiv:math.OA/0503344. MR 2234123

Department of Mathematics, University of California, Berkeley, CA 94720-3840
E-mail address: rieffel@math.berkeley.edu