Journal of Approximation Theory **116**, 397–415 (2002)
doi:10.1006/jath.2002.3679

Strong Subdifferentiability of Convex Functionals and Proximinality

G. Godefroy

Équipe d'Analyse, Université Paris 6, Case 186, 4, Place Jussieu, 75252 Paris Cedex 05, France
E-mail: gig@ccr.jussieu.fr

V. Indumathi

Department of Mathematics, University of Pondicherry, Kalapet, Pondicherry 605014, India
E-mail: indu@pu.pon.nic.in

and

F. Lust-Piquard

Département de Mathémathiques, Université de Cergy-Pontoise, 2, Avenue A. Chauvin, 95302 Cergy, France
E-mail: Francoise.piquard@math.u-cergy.fr

Communicated by Frank Deutsch

Received February 27, 2001; accepted in revised form January 9, 2002

Using strong subdifferentiability of convex functionals, we give a new sufficient condition for proximinality of closed subspaces of finite codimension in a Banach space. We apply this result to the Banach space $K(l_2)$ of compact operators on l_2 and we show that a finite codimensional subspace Y of $K(l_2)$ is strongly proximinal if and only if every linear form which vanishes on Y attains its norm.

1. INTRODUCTION AND NOTATION

Compactness is crucially needed for showing that distances to closed subsets are attained and that multivalued maps with closed graphs are upper-semicontinuous. Compactness arguments usually fail in infinite dimensional Banach spaces and the question arises whether substitutes can be used. The interesting answer is that it is sometimes, but not always, possible.

0021-9045/02 $35.00

In this work, we investigate proximinality of closed vector subspaces, in its usual and in its strong form, according to the following terminology.

Let X be a normed linear space and let Y be a closed linear subspace of X. For $x \in X$, define

$$d(x, Y) = \inf\{\|x-y\|; y \in Y\},$$
$$P_Y(x) = \{y \in Y; \|x-y\| = d(x; Y)\},$$

and for any $\delta > 0$

$$P_Y(x, \delta) = \{y \in Y; \|x-y\| < d(x, Y)+\delta\}.$$

The space Y is said to be *proximinal* in X, if for each $x \in X$, the set $P_Y(x)$ is nonempty. We say that Y is *strongly proximinal* (See [G-I1]) if it is proximinal and if, for any $x \in X$ and $\epsilon > 0$, there exists $\delta > 0$ such that $y \in P_Y(x, \delta)$ implies that $d(y, P_Y(x)) < \epsilon$. We note that strong proximinality is not a new notion; indeed, it has been introduced under the name *H-set* in [V1] and further developed in [V2 , V3, V4, B].

It is difficult to find usable conditions which determine if a given subspace is proximinal. Various characterizations and sufficient conditions for proximinality of closed subspaces of finite codimension are available (see [Ga, S, V5, V6, I1, I2]), and the strong proximinality case has been recently investigated in [GI1]. An easy consequence of a proximinality characterization of Garkavi [Ga] is the following necessary condition for proximinality of finite codimentional subspaces:

$$Y \text{ is proximinal in } X \Rightarrow \text{ every } f \in Y^{\perp} \text{ attains its norm on } X. \tag{1}$$

This necessary condition is easy to check, but unfortunately it is far from being sufficient for proximinality [P1]. For some spaces X it does suffice (see [D, I1, I2, GI3]). In fact, this is true under smoothness assumptions on the dual space X^* or alternatively under convexity assumptions on the space X. It turns out that the assumptions which have been used so far are too strong for being satisfied by classical nonreflexive spaces equipped with their *natural* norm. This work provides weaker assumptions which do apply to such spaces.

In the first part of this paper (Sections 2 and 3), we present a new geometric criterion, which implies strong proximinality, and thus proximinality, of closed subspaces Y of finite codimension of a Banach space X such that every $f \in Y^{\perp}$ attains its norm. Unexpectedly, equivalence between proximinality and its strong version thus holds for many spaces which are not reflexive and not strictly convex. In the second part of this paper (Sections 4 and 5), we illustrate our criterion by applying it to

$X = K(l_2)$ and we clear up proximinality questions for finite codimensional subspaces of $K(l_2)$.

Notation. In the following, X denotes a real Banach space, X^* its dual, B_X the closed unit ball of X, and S_X the unit sphere of X. The class of all functionals in X^* which attain their norm on X is denoted by $NA(X)$. All subspaces are assumed to be closed. If Y is a closed subspace of finite codimension in a general normed linear space X, let $Y^{\perp}$ denote the annihilator space of Y given by

$$\{f \in X^*; f(y) = 0 \; \forall y \in Y\}.$$

If $C \subseteq X$, then $Co(C)$ and $\mathscr{E}xt(C)$ denote the convex hull and the set of extreme points of C respectively. Also, $\bar{C}^{w*}$ denotes the weak* closure of C in the second dual space X^{**}.

Let F be a real valued convex function defined on a Banach space Z. For fixed z and y in Z, $(F(z+ty)-F(z))/t$ is an increasing function of t and therefore $\lim_{t\to 0^+} (F(z+ty)-F(z))/t$ exists. Further, if $t > 0$,

$$(F(z+ty)-F(z))/t \geqslant \lim_{t\to 0^+} (F(z+ty)-F(z))/t. \tag{2}$$

The set of subdifferentials of F at z, denoted by $\partial F(z)$, is defined by

$$\partial F(z) = \{\phi \in X^* : \phi(h) \leqslant F(z+h) - F(z) \; \forall h \in X\}.$$

For $\phi \in \partial F(z)$ and $h \in X$, we have

$$\phi(h) \leqslant \lim_{t\to 0^+} (F(z+th)-F(z))/t$$

and moreover there exists (See Proposition 2.24 of [P2]) some $\phi \in \partial F(z)$ such that

$$\phi(h) = \lim_{t\to 0^+} (F(z+th)-F(z))/t. \tag{3}$$

We say that F is *strongly subdifferentiable (SSD)* at $z \in Z$ (see [FP, DGZ]) if the one-sided limit

$$\lim_{t\to 0^+} (F(z+th)-F(z))/t$$

exists uniformly for $h \in S_Z$.

2. THE CONVEX FUNCTIONAL S_C

Let X be a Banach space and let C be a closed, convex, and bounded subset of X. We define a real valued, convex functional S_C on X^* by

$$S_C(f) = \sup_C (f).$$

This functional is usually called the support function of the set C. Our proximinality result involves strong subdifferentiability of convex functionals. The purpose of this section is to find necessary and sufficient conditions for the strong subdifferentiability of S_C. We first recall a simple fact.

FACT 2.1. $\partial S_C(f) = \{\phi \in \bar{C}^{w*}; \phi(f) = S_C(f)\}$.

Proof. If $\phi \in \partial S_C(f)$, then

$$\phi(h) \leqslant S_C(f+h) - S_C(f) \leqslant S_C(h) \qquad \forall h \in X^*$$

and hence $\phi \in \bar{C}^{w*}$ by the bipolar theorem. Since $S_C(0) = 0$,

$$\phi(-f) \leqslant -S_C(f)$$

and therefore $\phi(f) = S_C(f)$. Conversely, if $\phi \in \bar{C}^{w*}$, $\phi(g) \leqslant S_C(g)$ for all $g \in X^*$ and if $\phi(f) = S_C(f)$, it follows that

$$\phi(g-f) = \phi(g) - S_C(f) \leqslant S_C(g) - S_C(f)$$

and so $\phi \in \partial S_C(f)$. This completes the proof of Fact 2.1.

Remark 2.1. If we let

$$J_C(f) = \{x \in C : f(x) = S_C(f)\},$$

then by Fact 2.1, we have $J_C(f) = \partial S_C(f) \cap X$. When $C = B_X$, we simply denote $J_{B_X}(f) = J_X(f)$. Note that although the set $\partial S_C(f)$ is always nonempty, the set $J_C(f)$ can be empty.

We now characterize strong subdifferentiability of the convex functional S_C.

PROPOSITION 2.2. *Let X be a Banach space and C be a closed, convex, and bounded subset of X. Then the following are equivalent:*

(a) *The convex functional S_C is SSD at $f \in X^*$.*

(b) *Given $\epsilon > 0$, there exists $\delta > 0$ such that*

$$\phi \in \bar{C}^{w*}, \qquad \phi(f) > S_C(f) - \delta \Rightarrow d(\phi, \partial S_C(f)) < \epsilon.$$

(c) *The set* $J_C(f)$ *is nonempty and given* $\epsilon > 0$, *there exists* $\delta > 0$ *such that*

$$x \in C, \qquad f(x) > S_C(f) - \delta \Rightarrow d(x, J_C(f)) < \epsilon.$$

Proof.

(a) $\Rightarrow$ (b): Assume (a). If (b) fails, we can find a sequence $\phi_n \subseteq \bar{C}^{w*}$ such that

$$\phi_n(f) \to S_C(f) \qquad \text{and} \qquad d(\phi_n, \partial S_C(f)) > \epsilon_0 > 0.$$

By Fact 2.1, $\partial S_C(f)$ is a weak* compact convex subset of X^{**} and the Hahn–Banach theorem provides a sequence $h_n \subseteq S_{X^*}$ such that

$$\phi_n(h_n) - \psi(h_n) > \epsilon_0 \qquad \forall \psi \in \partial S_C(f).$$

Also by (3), for each n, there exists $\psi_n \in \partial S_C(f)$ such that

$$\psi_n(h_n) = \lim_{t \to 0^+} t^{-1}[S_C(f + th_n) - S_C(f)]. \tag{4}$$

Now,

$$\begin{aligned} S_C(f + th_n) - S_C(f) &\geqslant \phi_n(f + th_n) - \psi_n(f) \\ &= (\phi_n - \psi_n)(f) + t(\phi_n - \psi_n)(h_n) + t\psi_n(h_n). \end{aligned}$$

Hence if $t > 0$,

$$\begin{aligned} t^{-1}[S_C(f + th_n) - S_C(f)] - \psi_n(h_n) &\geqslant (\phi_n - \psi_n)(h_n) + t^{-1}(\phi_n - \psi_n)(f) \\ &\geqslant \epsilon_0 + t^{-1}(\phi_n - \psi_n)(f). \end{aligned}$$

So, if $t_n = 2(\psi_n - \phi_n)(f)/\epsilon_0$, then

$$t_n^{-1}[S_C(f + t_n h_n) - S_C(f)] - \psi_n(h_n) \geqslant \epsilon_0/2, \qquad \forall n.$$

Since $\lim_n (\phi_n - \psi_n)(f) = 0$, $\lim_{n \to \infty} t_n = 0$, and this together with (4) shows that

$$\lim_{t \to 0^+} t^{-1}[S_C(f + th_n) - S_C(f)]$$

is not uniform on $(h_n)_{n \geqslant 1} \subseteq S_{X^*}$, contradicting (a).

(b) $\Rightarrow$ (c): We will show the following fact, which extends [GI1, Lemma 1.1] with a similar proof.

FACT 2.3. *If* (b) *holds, then* $J_C(f) \neq \emptyset$ *and for each* $x \in C$, $d(x, \partial S_C(f)) = d(x, J_C(f))$.

Proof. Choose any $\alpha > d(x, \partial S_C(f))$ and $\epsilon > 0$. Then there exists $\phi \in \partial S_C(f)$ such that $\|x - \phi\| < \alpha$. Using the principle of local reflexivity, we can get a net $(x_\lambda) \subseteq C$ such that

$$\|x - x_\lambda\| < \alpha \ \forall \lambda \qquad \text{and} \qquad \lim_\lambda f(x_\lambda) = \phi(f) = S_C(f).$$

Now, (b) implies that $\lim_\lambda d(x_\lambda, \partial S_C(f)) = 0$. Hence there exists $x_1 \in C$ such that

$$\|x - x_1\| < \alpha \qquad \text{and} \qquad d(x_1, \partial S_C(f)) < \epsilon.$$

Repeating the above argument with x_1 and ϵ replacing x and α, respectively, we get $x_2 \in C$ satisfying

$$\|x_1 - x_2\| < \epsilon \qquad \text{and} \qquad d(x_2, \partial S_C(f)) < \epsilon/2.$$

Proceeding thus inductively, we obtain a norm Cauchy sequence $(x_n) \subseteq C$ such that if $x_\infty = \lim_n x_n$, then $x_\infty \in \partial S_C(f) \cap C = J_C(f)$ by Fact 2.1 and

$$\|x - x_\infty\| < \alpha + 2\epsilon.$$

Hence $J_C(f) \neq \emptyset$ and $d(x, J_C(f)) < \alpha + 2\epsilon$. Since $\epsilon > 0$ and $\alpha > d(x, \partial S_C(f))$ were chosen arbitrarily, this proves Fact 2.3 and the implication (b) $\Rightarrow$ (c).

(c) $\Rightarrow$ (b): If $\phi \in \bar{C}^{w^*}$ and $\phi(f) > S_C(f) - \delta$, there exists a net $(x_\lambda) \subseteq C$ such that

$$\phi = w^* - \lim x_\lambda \qquad \text{and} \qquad f(x_\lambda) > S_C(f) - \delta \ \forall \lambda.$$

Hence $d(x_\lambda, J_C(f)) < \epsilon$ for each λ. If $(y_\lambda) \subseteq J_C(f)$ is such that $\|x_\lambda - y_\lambda\| < \epsilon$ and ψ is a weak* cluster point of the net (y_λ), then $\psi \in \partial S_C(f)$ and $\|\phi - \psi\| \leqslant \epsilon$.

(b) $\Rightarrow$ (a): Let $f \in X^*$ and $\epsilon > 0$. Let δ be as in (b). As C is bounded, so is $\partial S_C \subseteq \bar{C}^{w^*}$. Hence there exists $\epsilon_0 > 0$ such that for all $g \in X^*$ satisfying $\|f - g\| < \epsilon_0$, and for all $\phi \in \partial S_C(g)$, we have

$$|\phi(g) - \phi(f)| < \delta/2 \qquad \text{and} \qquad |S_C(g) - S_C(f)| < \delta/2.$$

Hence $\phi(f) > S_C(f) - \delta$. This with (b) implies that there exists $\psi \in \partial S_C(f)$ such that

$$\|\phi - \psi\| < \epsilon. \tag{5}$$

Applying Fact 2.1 twice, we have

$$\phi(g-f) \geqslant S_C(g) - S_C(f) \geqslant \psi(g-f). \tag{6}$$

Also, for any $h \in X^*$ and $\phi_0 \in \partial S_C(f)$, we have

$$\lim_{t \to 0^+} t^{-1}[S_C(f+th) - S_C(f)] \geqslant \phi_0(h). \tag{7}$$

Now using (5), taking $g = f + th$ in (6) and $\phi_0 = \psi$ in (7), we get for all $t \in (0, \epsilon_0)$ and for all $h \in X^*$ with $\|h\| = 1$,

$$\begin{aligned} \epsilon &\geqslant (\phi - \psi)(h) \\ &\geqslant t^{-1}[S_C(f+th) - S_C(f)] - \lim_{t \to 0^+} t^{-1}[S_C(f+th) - S_C(f)], \end{aligned}$$

which implies that S_C is SSD at f. This concludes the proof of the proposition.

We conclude this section with the following corollary.

COROLLARY 2.4. *Let C be as in Proposition* 2.2. *If S_C is SSD at f, then $J_C(f)$ is nonempty and $\partial S_C(f) = \overline{J_C(f)}^{w^*}$.*

Proof. The fact that the set $J_C(f)$ is nonempty follows from the implication (a) ⇒ (c) given above. Clearly $\overline{J_C(f)}^{w^*} \subseteq \partial S_C(f)$. If $\phi \in \partial S_C(f)$, by Fact 2.1 there exists a net $(x_\lambda) \subseteq C$ such that $\phi = w^* - \lim x_\lambda$ and $\lim f(x_\lambda) = S_C(f)$. Now by (a) ⇒ (c) of Proposition 2.2, $\lim_\lambda d(x_\lambda, J_C(f)) = 0$ and thus $\phi \in \overline{J_C(f)}^{w^*}$.

3. STRONG PROXIMINALITY OF SUBSPACES OF FINITE CODIMENSION

The main result of this section (Theorem 3.2) is a general statement which relates the strong sub differentiability of convex functionals on the dual space with proximinality. Let X be a normed linear space and Y be a subspace of finite codimension in X. Let Q: $X \to X/Y$ be the quotient map. Then it is easy to check that

$$Y \text{ is proximinal} \Leftrightarrow Q(B_X) = B_{X/Y}.$$

Since $X/Y = (Y^\perp)^*$,

$$Y \text{ is proximinal} \Leftrightarrow (B_X)_{|Y^\perp} = B_{(Y^\perp)^*},$$

where, for $x \in X$, $x_{|Y^\perp}$ denotes the restriction of the canonical image of x in X^{**} to the subspace $Y^\perp$. Now, by the Krein–Milman theorem, $B_{(Y^\perp)^*} = Co(\mathscr{E}xt\, B_{(Y^\perp)^*})$ and therefore,

$$Y \text{ is proximinal} \Leftrightarrow \mathscr{E}xt\, B_{(Y^\perp)^*} \subseteq (B_X)_{|Y^\perp}. \tag{8}$$

We now need some notation from [I1]. If Z is an n- dimensional normed linear space and $(f_1, f_2, ..., f_n)$ is a basis of Z, let

$$D_0 = B_{Z^*}, \qquad D_i = \{t \in D_{i-1} : f_i(t) = \sup_{D_{i-1}} f_i\}, \qquad 1 \leqslant i \leqslant n.$$

Then we have

Lemma 3.1 [I1]. *Let $e \in \mathscr{E}xt\, B_{Z^*}$. Then there exists a basis $(f_1, f_2, ..., f_n)$ of Z such that $\{e\} = D_n$.*

We refer to Remark 2.1 for the notation $J_C(f)$. Let us state and prove the main result of this section.

Theorem 3.2. *Let X be a Banach space and Y be a subspace of X with $\dim X/Y = n$. For each basis $(f_1, f_2, ..., f_n)$ of $Y^\perp$, we set $C_0 = B_X$ and $C_i = J_{C_{i-1}}(f_i)$ $(1 \leqslant i \leqslant n)$ and we assume that $S_{C_{i-1}}$ is SSD at f_i for $1 \leqslant i \leqslant n$. Then Y is strongly proximinal in X.*

Proof. Let $(f_1, f_2, ..., f_n)$ be a basis of $Y^\perp = Z$. We define the sets $(D_i)_{0 \leqslant i \leqslant n}$ as above. By the Hahn–Banach theorem,

$$D_0 = B_{(Y^\perp)^*} = B_{X^{**}|Y^\perp} = \overline{C_{0|Y^\perp}}^{w^*}.$$

Assume that $C_{i-1} \neq \varnothing$ and that $D_{i-1} = \overline{C_{i-1|Y^\perp}}^{w^*}$ for some i with $1 \leqslant i \leqslant n-1$. Since $S_{C_{i-1}}$ is SSD at f_i, Corollary 2.4 implies

$$\overline{C_i}^{w^*} = \overline{J_{C_{i-1}}(f_i)}^{w^*} = \partial S_{C_{i-1}}(f_i);$$

hence $C_{i \neq \varnothing}$, and by Fact 2.1,

$$\partial S_{C_{i-1}}(f_i) = \{\phi \in \overline{C_{i-1}}^{w^*} : \phi(f_i) = S_{C_{i-1}}(f_i)\}$$

and thus

$$\overline{C_i}^{w^*} = \{\phi \in \overline{C_{i-1}}^{w^*} : \phi(f_i) = \sup_{D_{i-1}} (f_i)\}.$$

It follows that $C_i \neq \varnothing$ and that $D_i = \overline{C_{i|Y^\perp}}^{w^*}$ for every i with $1 \leqslant i \leqslant n$.

Pick now $e \in \mathscr{E}xt\, B_{(Y^\perp)^*}$. By Lemma 3.1, there is a basis $(f_1, f_2, ..., f_n)$ of $Y^\perp = Z$ such that $\{e\} = D_n$. By the above, we have $\{e\} = \overline{C_{n|Y^\perp}}^{w^*}$. Since

$C_n \neq \emptyset$, this implies that $\{e\} = C_{n|Y^\perp}$ and hence $e \in B_{X|Y^\perp}$. It follows now from (8) that Y is proximinal.

In order to show that Y is in fact strongly proximinal, it suffices, by the characterization from [I1] (see [GI1, Section 2]) and Theorem 2.5 from [GI1], to show that for every basis $(f_1, f_2, \ldots, f_n)$ of $Y^\perp$ and every $1 \leqslant j \leqslant n$, one has

$$\lim_{\epsilon \to 0} [\sup \{d(x, C_j) : x \in J(f_1, f_2, \ldots, f_j, \epsilon)\}] = 0, \tag{9}$$

where we set

$$J(f_1, f_2, \ldots, f_j, \epsilon) = \bigcap_{i=1}^{i=j} \{x \in B_X : f_i(x) > S_{C_{i-1}}(f_i) - \epsilon\}.$$

Note that with the notation of [GI2], we have

$$C_i = J(f_1, f_2, \ldots, f_i).$$

Let us now proceed to prove (9). For $j = 1$, it holds true by the implication (a) $\Rightarrow$ (c) of Proposition 2.2. This same implication gives that for $1 < i \leqslant n$

$$\lim_{\epsilon \to 0} [\sup \{d(x, C_i); x \in J_{C_{i-1}}(f_i, \epsilon)\}] = 0, \tag{10}$$

where we denote

$$J_{C_{i-1}}(f_i, \epsilon) = \{x \in C_{i-1}; f_i(x) > S_{C_{i-1}}(f_i) - \epsilon\}.$$

Now applying (10) for $i = 1, \ldots, j \leqslant n$ shows Eq. (9) for j. This concludes the proof of Theorem 3.2.

Remarks 3.3. (1) The assumptions of Theorem 3.2 can be reformulated in such a way that they make no reference to the predual X. Indeed, set $J^{**}(\emptyset) = B_{X^{**}}$, $S_0(f) = \|f\|$ and

$$J^{**}(f_1) = \{f \in J^{**}(\emptyset); F(f_1) = S_0(f_1)\}.$$

For $1 \leqslant i \leqslant n-1$, we define inductively

$$S_i(f) = \sup \{F(f); F \in J^{**}(f_1, f_2, \ldots, f_i)\}$$

and

$$J^{**}(f_1, f_2, \ldots, f_{i+1}) = \{F \in J^{**}(f_1, f_2, f_i); F(f_{i+1}) = S_i(f_{i+1})\}.$$

Corollary 2.4 shows that the assumptions of Theorem 3.2 are equivalent to saying that for every $0 \leqslant i \leqslant n-1$, the functional S_i is SSD on X^* at f_{i+1}.

(2) Using the methods of [GI2], one can show that if the subdifferential map $\partial S_C(f)$ is norm-to-weak upper-semicontinuous at f, then $J_C(f) \neq \emptyset$ and $\partial S_C(f) = \overline{J_C(f)}^{w^*}$. Hence, we may conclude proximinality under the weaker assumption of norm to weak upper semicontinuity of the set valued maps $S_{C_{i-1}}$ at f_i for $1 \leqslant i \leqslant n$, since the first part of the proof works under this assumption.

We state now an important special case of Theorem 3.2.

COROLLARY 3.4. *Let X be a Banach space such that the norm of X^* is Fréchet differentiable at every nonzero norm- attaining functional of X^*. Then a finite codimensional subspace Y of X is strongly proximinal in X if and only if $Y^\perp$ is contained in the set of norm-attaining functionals, and in this case the set $P_Y(x)$ is a singleton for every $x \in X$ and the nearest point projection is continuous from X onto Y.*

Proof. It is clear by (1) that the norm-attainment condition is necessary. Conversely, pick any basis $(f_1, f_2, ..., f_n)$ of $Y^\perp$ and assume that $Y^\perp$ consists of norm-attaining functionals. Since S_{C_0} is the dual norm, it is in particular SSD at f_1. Moreover C_1 is a singleton, which makes it obvious that S_{C_1}, and then S_{C_i} for all $i < n$, is SSD at every point and in particular at the functionals f_j. Theorem 3.2 provides the strong proximinality.

Note now that if $x \in S_X$ is such that $d(x, Y) = 1$, and $f \in Y^\perp$ is such that $\|f\| = 1 = f(x)$, then $f(x-y) = 1 = \|x-y\|$ for every $y \in P_Y(x)$. Since the dual norm is smooth at f, this implies that $P_Y(x)$ reduces to $\{0\}$, and the result follows. The continuity of the nearest point projection is clear by strong proximinality.

The assumptions of Corollary 3.4 are satisfied exactly by the spaces which have an average locally uniformly rotund norm, that is, a strictly convex norm such that the weak and norm topologies coincide on the unit sphere (see [DGZ, Theorem IV.2.2]). In particular, its conclusion holds when the norm of X is locally uniformly rotund. This is a strong version of Proposition 1 from [I2]. We shall see below that Theorem 3.2 also applies to spaces whose norm is far from being strictly convex.

4. THE SPACE $K(l_2)$ AND THE NORM-ATTAINING FUNCTIONALS OF ITS DUAL

Our goal is now to show that Theorem 3.2 applies to the space of compact operators on the Hilbert space l_2. In order to prove this, we first investigate in Section 4 which linear forms on that space attain their norm and the subsets of the sphere where they attain it.

Let $l_2 = l_2(\mathbf{N}; \mathbf{C})$, where $\mathbf{N}$ denotes the set of natural numbers and $\mathbf{C}$ the complex field, be the separable complex Hilbert space. Let $B(l_2)$ be the space of bounded linear operators on l_2 and let $K(l_2)$ be the space of compact linear operators on l_2. We denote the space of trace class linear operators on l_2 by $N(l_2)$ and by $\mathcal{U}(l_2)$ the group of unitary operators on l_2. We use the operator norms on the spaces $B(l_2)$ and $K(l_2)$, and we denote the identity operator on l_2 by Id_{l_2}. If $A \in B(l_2)$, $rk(A)$ denotes the rank of the operator A and $Ker(T)$ is the kernel of the operator T. If $A \in N(l_2)$, $tr(A)$ denotes the trace of the operator A.

The description of the dual of the space of compact operators on the Hilbert space goes back to R. Schatten (see [Sc]). Let $X = K(l_2)$. Then $X^* = N(l_2)$ is equipped with the nuclear norm, and for $A \in N(l_2)$ and $T \in K(l_2)$, the duality is given by

$$A(T) = tr(A^*T),$$

where the adjoint A^* of A is given by

$$\langle y, Ax \rangle = \langle A^*y, x \rangle, \qquad \forall x, y \text{ in } l_2.$$

If E is a closed subspace of l_2, we will denote the orthogonal complement $\{x \in l_2 : \langle x, y \rangle = 0 \ \forall y \in E\}$ of E by $E^\perp$ and by P_E the orthogonal projection onto E. If z is a complex number, Rez and Imz denote the real and imaginary part of z. If $f \in X^*$, where X is a complex Banach space, Ref denotes the real valued linear functional, over the real Banach space X, given by

$$(Ref)(x) = Ref(x) \qquad \forall x \in X.$$

Finally, if Y is a subspace of X, dim Y denotes the dimension of Y.

Remark 4.1. Let $\mathbf{R}$ denote the real field. If $Z = Z_{\mathbf{C}}$ is a complex Banach space, and $Z_{\mathbf{R}} = Z$ is equipped with the real structure, the map $f \to Re(f)$ is a $\mathbf{R}$-linear isometry between $(Z_{\mathbf{C}})^*$ and $(Z_{\mathbf{R}})^*$. Clearly, norm attainment is preserved. In what follows, $K(l_2) = X$ is equipped with its real structure. Of course complex subspaces are in particular real subspaces, and proximinality notions do not depend upon the scalar field. Hence our results apply in particular to the complex subspaces. If $T \in X$ and $A \in X^*$, the real duality bracket is given by

$$\langle A, T \rangle = Re[tr(A^*T)].$$

We start with the following observation about norm attaining trace class operators, which is needed in the remainder of the paper.

LEMMA 4.2. *Let $X = K(l_2)$. If $A \in X^*$, then A attains its norm over X if and only if $rk(A)$ is finite.*

Proof. It is clear that every finite rank operator attains its norm over X. Let us prove the converse. We can and do assume that $\|A\| = 1$. Also, to begin with, we assume that A is a self-adjoint operator. Then there exists an orthonormal basis (v_i) of l_2 and a sequence of real scalars (α_i) such that

$$A = \sum_i \alpha_i v_i \otimes v_i, \qquad \sum_i |\alpha_i| = 1.$$

That is,

$$A(y) = \sum_i \alpha_i \langle v_i, y \rangle \, v_i, \qquad \text{for} \quad y \in l_2.$$

For $T \in X$, we have

$$Re[tr(A^*T)] = \sum_i \alpha_i \; Re[\langle v_i, Tv_i \rangle].$$

If $\langle A, T \rangle = 1 = \|T\|$, then

$$|\langle v_i, Tv_i \rangle| \leqslant 1, \sum_i |\alpha_i| = 1 \qquad \text{and} \qquad \sum_i , \alpha_i \; Re[\langle v_i, Tv_i \rangle] = 1,$$

and therefore we have

$$\langle v_i, Tv_i \rangle = sign(\alpha_i) \qquad \forall i.$$

This implies

$$Tv_i = sign(\alpha_i) \, v_i \qquad \forall i. \tag{11}$$

Since T is compact, $\dim(Ker(T \pm I)) < \infty$ and this together with (11) implies $rk(A) < \infty$. We now consider any X* . Then

$$A = U \, |A|,$$

where $|A| = (A^*A)^{1/2} \geqslant 0$, and U is a partial isometry, hence $\|U\| = 1$. For $T \in X$, we have

$$\langle A, T \rangle = Re[tr(A^*T)] = Re[tr(|A| \, U^*T)] = \langle |A|, U^*T \rangle.$$

Hence for $B \in X^*$, if we let

$$J_X(B) = \{T \in X; \langle B, T \rangle = \|B\|, \|T\| = 1\},$$

then we have

$$T \in J_X(A) \Rightarrow U^*T \in J_X(|A|). \tag{12}$$

Therefore $J_X(A) \neq \emptyset$ implies $J_X(|A|) \neq \emptyset$. Since the operator $|A|$ is self-adjoint, by the first part of the proof, $rk(|A|) < \infty$ and thus $rk(A) < \infty$.

Remark 4.3. If $rk(A) < \infty$ (or equivalently $rk(|A|) < \infty$), then $A = U\,|A|$, with $UU^* = U^*U = Id_{l_2}$. That is, $U \in \mathcal{U}(l_2)$.

LEMMA 4.4. *Let $X = K(l_2)$. If $A \in X^*$ is a norm attaining functional, then*

$$J_X(A) = U \cdot J_X(|A|), \qquad \textit{where} \quad U \in \mathcal{U}(l_2).$$

Proof. By Lemma 4.1 and Remark 4.3, $A = U\,|A|$, with $U \in \mathcal{U}(l_2)$. By (12),

$$T \in J_X(A) \Rightarrow U^*T \in J_X(|A|).$$

Conversely,

$$\begin{aligned} U^*T \in J_X(|A|) &\Leftrightarrow U^*T \in J_X(U^*A) \\ &\Rightarrow \|A\| = Re[tr(A^*UU^*T)] = Re[tr(A^*T)] \\ &\Rightarrow T \in J_X(A). \end{aligned}$$

Hence the lemma.

We need the following simple observation in subsequent proofs.

FACT 4.5. *Let $T \in B(l_2)$, $\|T\| = 1$. Let $E_{\pm 1}$ be the eigenspace of T associated to the eigenvalues ± 1. Then $E_1^\perp$ and $E_{-1}^\perp$ are T-invariant subspaces.*

Proof. Assume $\|v\| = \|w\| = 1$, $Tv = \pm v$, and $\langle v, w\rangle = 0$. Then

$$\|v + tw\| = \sqrt{1 + t^2} \leqslant 1 + t^2/2, \qquad \forall t \in \mathbf{R}.$$

This implies

$$1 + t^2/2 \geqslant |\langle v, T(v + tw)\rangle| = |\pm 1 + t\langle v, v, Tw\rangle|, \qquad \forall t \in \mathbf{R}$$

which, in turn, implies $\langle v, Tw\rangle \in i\mathbf{R}$. We now replace w by iw in the above to conclude $\langle v, Tw\rangle = 0$.

We now describe the set $J_X(A)$ for a norm-attaining $A \in N(l_2)$.

LEMMA 4.6. *Let $X = K(l_2)$ and $A \in X^*$ be a norm- attaining functional. Then there exists a $T_0 \in X$, with $rk(T_0) < \infty$, $U_0 \in \mathscr{U}(l_2)$, and a finite dimensional subspace E of l_2 such that*

$$J_X(A) = \{T_0 + U_0 P_{E^\perp} V P_{E^\perp}, V \in X, \|V\| \leqslant 1\}.$$

Proof. If A is self-adjoint, by the proof of Lemma 4.1, there exists an orthonormal set $\{v_i: 1 \leqslant i \leqslant n\}$ such that

$$T \in J_X(A) \Leftrightarrow \|T\| = 1 \qquad \text{and} \qquad T v_i = \varepsilon_i v_i, \qquad \forall 1 \leqslant i \leqslant n$$

with $\varepsilon_i = sign\langle A v_i, v_i\rangle$, for $1 \leqslant i \leqslant n$. Let $E = span_{\mathbb{C}}\{v_i : i \leqslant n\}$. By Fact 4.5, if $T \in J_X(A)$ then $T(E^\perp) \subseteq E^\perp$. Hence if A is self-adjoint,

$$J_X(A) = \{\Delta + P_{E^\perp} V P_{E^\perp} : \|V\| \leqslant 1\},$$

where Δ is a diagonal $n \times n$ matrix, with each entry on the diagonal being either 1 or -1. For general A, $A = U_0\,|A|$ with $U_0 \in \mathscr{U}(l_2)$ and we apply the first part of the proof to the self-adjoint operator $|A|$ and use Lemma 4.4 to conclude that

$$J_X(A) = \{U_0\,\Delta + U_0 P_{E^\perp} V P_{E^\perp}, \|V\| \leqslant 1\}.$$

5. SSD AT NORM-ATTAINING FUNCTIONALS AND PROXIMINALITY IN $K(l_2)$

The main result of this section (Theorem 5.3) is that our general theorem applies to $X = K(l_2)$. This extends results from [GI3] on $X = c_0$, which were obtained by different techniques, to the noncommutative case. We first prove that the norm of $\mathscr{N}(l_2) = (K(l_2))^*$ is SSD at all the norm-attaining functionals. Then we proceed to show SSD for relevant functionals S_C, which we need for performing an easy induction on the codimension which provides the result.

LEMMA 5.1. *Let $X = K(l_2)$. If $A \in X^*$ is norm- attaining, then the norm of X^* is SSD at A.*

Proof. We may and do assume $\|A\| = 1$. Also, using multiplication with $U \in \mathscr{U}(l_2)$, we may and do assume that A is self-adjoint and $A \geqslant 0$. Thus

$$A = \sum_{i=1}^{n} \lambda_i v_i \otimes v_i,$$

with $0 \leqslant \lambda_i \leqslant 1$, $\sum \lambda_i = 1$, and $\{v_i: 1 \leqslant i \leqslant n\}$ an orthonormal subset of l_2. We set $E = span_{\mathbf{C}}\{v_i: 1 \leqslant i \leqslant n\}$. If $T \in B_X = \{T \in X; \|T\| \leqslant 1\}$, then

$$\langle A, T \rangle = Re[tr(A^*T)] = \sum_{i=1}^{n} \lambda_i \; Re[\langle v_i, Tv_i \rangle].$$

If $(T_k)_k \subseteq B_X$ are such that

$$\lim_k \langle A, T_k \rangle = 1$$

then, for all $1 \leqslant i \leqslant n$,

$$\lim_k Re[\langle v_i, T_k v_i \rangle] = 1$$

which implies

$$\lim_k \langle v_i, T_k v_i \rangle = 1, \qquad \forall 1 \leqslant i \leqslant n. \tag{13}$$

If E_i denotes $span_{\mathbf{C}}(v_i)$ for $1 \leqslant i \leqslant n$, it follows that

$$\lim_k P_{E_i}(T_k v_i) = v_i. \tag{14}$$

Now we have for $1 \leqslant i \leqslant n$, $1 \leqslant j \leqslant n$, and $i \neq j$,

$$\|P_{E_i} T_k v_i\|^2 + \|P_{E_j} T_k v_i\|^2 + \|P_{E^\perp} T_k v_i\|^2 \leqslant \|T_k v_i\|^2 \leqslant 1.$$

Note that (14) implies

$$\lim_k \|P_{E_i} T_k v_i\| = 1.$$

Hence

$$\lim_k \|P_{E^\perp} T_k v_i\| = 0 \qquad \text{and} \qquad \lim_k \langle T_k v_i, v_j \rangle = 0.$$

So,

$$\lim_k \|P_{E^\perp} T_k P_E\| = 0 \tag{15}$$

and

$$\lim_k \left\| P_E T_k P_E - \sum_{i=1}^{n} \langle T_k(v_i), v_i \rangle \, v_i \otimes v_i \right\| = 0. \tag{16}$$

If now $w \in E^{\perp}$ and $\|w\| = 1$, we have, since $\|T_k\| \leqslant 1$,

$$Re[\langle v_i, T_k(v_i + tw)\rangle] \leqslant \sqrt{1+t^2} \leqslant 1 + t^2/2 \qquad \forall t \in \mathbf{R}.$$

Hence, for all $\epsilon > 0$, if $Re[\langle v_i, T_k v_i\rangle] > 1 - \epsilon$,

$$1 - \epsilon + t\, Re[\langle v_i, T_k w\rangle] \leqslant 1 + t^2/2$$

which implies

$$t\, Re[\langle v_i, T_k w\rangle] - t^2/2 \leqslant \epsilon.$$

Taking $t = \sqrt{\epsilon}$, we get

$$Re[\langle v_i, T_k w\rangle] \leqslant \sqrt{\epsilon}\,/2.$$

Applying this to $w' = \alpha w$ with $|\alpha| = 1$ gives

$$|\langle v_i, T_k w\rangle| \leqslant \sqrt{\epsilon}\,/2.$$

It now follows that

$$\lim_k \|P_E T_k P_{E^{\perp}}\| = 0. \tag{17}$$

If $d_k = d(T_k, J_X(A))$, then by Lemma 4.6

$$d_k \leqslant \|P_E T_k P_{E^{\perp}}\| + \|P_{E^{\perp}} T_k P_E\| + \left\| P_E T_k P_E - \sum_{i=1}^{n} v_i \otimes v_i \right\|.$$

Hence by (13)–(17), $lim_k\, d_k = 0$. Now, by Proposition 2.2, $\|\cdot\|_{X^*}$ is SSD at A and this completes the proof of the lemma.

We now proceed to prove SSD of the convex functional S_C at A, where $C \subseteq K(l_2)$ is a suitably defined closed, convex, bounded set and A is a finite rank operator. We also provide a description of the set $J_C(A)$ in this case. Pick $T_0 \in K(l_2)$ with $rk(T_0) < \infty$, $U_0 \in \mathcal{U}(l_2)$, P_0 an orthogonal projection with $\dim(Ker(P_0)) < \infty$. Then we have the following result.

LEMMA 5.2. *Let* $X = K(l_2)$ *and* $C = \{T_0 + U_0 P_0 V P_0 : \|V\| \leqslant 1\}$. *For* $A \in X^*$, *we recall that* $S_C(A) = \sup\{\langle A, T\rangle : T \in C\}$. *If* $rk(A) < \infty$, *then*

(a) *The convex functional* S_C *is SSD at* A.

(b) *There exists a* $T_1 \in X$ *with* $rk(T_1) < \infty$, $U \in \mathcal{U}(l_2)$, *and an orthogonal projection* P_1 *with* $\dim(Ker(P_1)) < \infty$ *such that*

$$J_C(A) = \{T_1 + U_1 P_1 W P_1 : \|W\| \leqslant 1\}.$$

Proof. (a) If S and T are in $K(l_2) = X$ and $A \in X^*$, we have

$$tr(A^*ST) = tr((S^*A)^*T) = tr(TA^*S) = tr((AT^*)^*S).$$

Hence, by taking real parts, we get

$$\langle A, ST\rangle = \langle S^*A, T\rangle = \langle AT^*, S\rangle. \tag{18}$$

For all $A \in X^*$,

$$S_C(A) = \langle A, T_0\rangle + \sup\{\langle A, U_0P_0VP_0\rangle : \|V\| \leqslant 1\}.$$

Since $P_0 = P^*{}_0 = P_0^2$, we have by (18)

$$\langle A, U_0P_0VP_0\rangle = \langle P_0U^*{}_0AP_0, P_0VP_0\rangle. \tag{19}$$

Hence

$$S_C(A) = \langle A, T_0\rangle + \|P_0U^*{}_0AP_0\|_{X^*} = \langle A, T_0\rangle + \|A_1\|_{X^*}, \tag{20}$$

where $A_1 = P_0U^*{}_0AP_0$. Now assume that $rk(A) < \infty$. Then $rk(A_1) < \infty$. The maps

$$A \to \langle A, T_0\rangle \qquad \text{and} \qquad A \to A_1$$

are both linear and $\|\cdot\|_{X^*}$ is SSD at A_1 by Lemma 5.1. Hence S_C is SSD at A.

(b) Pick $T = T_0 + U_0P_0VP_0 \in C$. By (19),

$$T \in J_C(A) \Leftrightarrow P_0VP_0 \in J_X(A_1).$$

We have

$$A_1 = U' \cdot |A_1|$$

with $U' \in \mathscr{U}(l_2)$ satisfying $P_0U'P_0 = U'P_0$. By Lemma 4.6 applied to $|A_1|$ restricted to the closed subspace $P_0(l_2)$, we can get $T'_0 \in X$ with $rk(T'_0) < \infty$, satisfying $T'_0 = T'_0P_0 = P_0T'_0$, and an orthogonal projection, P'_0, with $P'_0P_0 = P_0P'_0 = P'_0$ and $\dim(Ker(P'_0)) < \infty$, such that

$$P_0J_X(|A_1|)\,P_0 = \{T'_0 + P'_0WP'_0;\ \|W\| \leqslant 1\}.$$

Hence

$$P_0J_X(A_1)\,P_0 = \{U'T'_0 + U'P'_0WP'_0;\ \|W\| \leqslant 1\}.$$

Now it follows that

$$\begin{aligned} J_C(A) &= \{T_0 + U_0 U'(T'_0 + P'_0 W P'_0); \|W\| \leqslant 1\} \\ &= \{T_1 + U_1 P_1 W P_1; \|W\| \leqslant 1\}, \end{aligned}$$

where

$$T_1 = T_0 + U_0 U' T'_0, \qquad U_1 = U_0 U', \qquad \text{and} \qquad P_1 = P'_0.$$

This completes the proof of the lemma.

It is now easy to show the main result of Sections 4 and 5, which is an application of Theorem 3.2 to the space $K(l_2)$. We refer to [GI3] for the much simpler commutative case.

THEOREM 5.3. *Let $X = K(l_2)$ and Y be a subspace of codimension n in X. Then Y is strongly proximinal in X if and only if $Y^\perp \subseteq NA(X)$, where $NA(X)$ denotes the set of all norm-attaining functionals on X.*

Proof. Since $Y^\perp \subseteq NA(X)$, by Lemma 4.1, $Y^\perp$ is contained in the space of finite rank operators on l_2. Select any basis $A_1, A_2, \ldots, A_n$ of $Y^\perp$. Each A_i, $1 \leqslant i \leqslant n$ is a finite rank operator. Set

$$C_0 = B(X) \qquad \text{and} \qquad C_k = J_{C_{k-1}}(A_k) \qquad \text{for} \quad 1 \leqslant k \leqslant n.$$

It follows from Lemma 4.6, Lemma 5.2, and a straightforward induction that the convex functionals $S_{C_{k-1}}$ are SSD at A_k for $1 \leqslant k \leqslant n$. Now Theorem 3.2 shows the strong proximinality for Y.

REFERENCES

[B] V. I. Berdyshev, Stability of a minimization problem under perturbation of the set of admissible elements, *Mat. Sb.* **103** (145), 4 (1977), 467–479. English transl. in *J. Math. USSR Sb.* **32**, 4 (1977), 401–412.

[D] F. Deutsch, Representers of linear functionals, norm- attaining functionals and best approximation by cones and linear varieties in inner product spaces, *J. Approx. Theory* **36** (1982), 226–236.

[DGZ] R. Deville, G. Godefroy, and V. Zizler, "Smoothness and Renormings in Banach Spaces," Pitmann Monographs and Surveys in Pure and Applied Mathematics, Vol. 64, Langman, Harlow, 1993.

[FP] C. Franchetti and R. Paya, Banach spaces with strongly subdifferentiable norm, *Boll. Uni. Mat. Ital.* **VII**.B (1993), 45–70.

[Ga] A. L. Garkavi, On the best approximation by elements of infinite dimensional subspaces of a certain class, *Mat. Sb.* **62** (1963), 104–120.

[GI1] G. Godefroy and V. Indumathi, Strong Proximinality and Polyhedral spaces, *Rev. Math.* **XIV**, 1 (2001), 105–125.

[GI2] G. Godefroy and V. Indumathi, Norm-weak upper semi-continuity of duality and preduality mappings, *J. Set-Valued Anal.*, in press.

[GI3] G. Godefroy and V. Indumathi, Proximinality in subspaces of c_0, *J. Approx. Theory* **101** (1999), 175–181.

[I1] V. Indumathi, Proximinal subspaces of finite codimension in general normed linear spaces, *Proc. London Math. Soc. (3)* **45** (1982), 435–455.

[I2] V. Indumathi, On transitivity of proximinality, *J. Approx. Theory* **49**, 2 (1987), 130–143.

[P1] R. R. Phelps, Chebychev subspaces of finite codimension in $C(X)$, *Pacific J. Math.* **13** (1963), 647–655.

[P2] R. R. Phelps, "Convex Functions Monotone Operators, and Differentiability," Lecture Notes in Mathematics, Vol. 1364, Springer-Verlag, Berlin/New York, 1989.

[Sc] R. Schatten, "Norm Ideals of Completely Continuous Operators," 2nd printing, Ergebnise der Mathematik, Vol. 27, Springer-Verlag, Berlin/New York, 1970.

[S] I. Singer, On best approximation in normed linear spaces by elements of subspaces of finite codimension, *Rev. Roumaine Math. Pure. Appl.* **17** (1972), 1245–1256.

[V1] L. P. Vlasov, The concept of approximative compactness and its variants, *Mat. Zametki* **16** (1974), 337–348. English transl. in *Math. Notes* **16**, No. 2 (1974), 786–792.

[V2] L. P. Vlasov, Properties of generalized elements of best approximation, *Mat. Zametki* **24**, No. 4 (1978), 513–522. English transl. in *Math. Notes* **24**, No. 4 (1978), 774–778.

[V3] L. P. Vlasov, Uniqueness of generalized elements of best approximation, *Mat. Zametki* **25**, No. 2 (1979), 161–175. English transl. in *Math. Notes* **25**, No. 2 (1979), 83–91.

[V4] L. P. Vlasov, Approximative properties of subspaces of finite codimension in $C(Q)$, *Mat. Zametki* **28**, No. 2 (1980), 205–222. English transl. in *Math. Notes* **28**, No. 2 (1980), 565–574.

[V5] L. P. Vlasov, Elements of best approximation relative to subspaces of finite codimension, *Mat. Zametki* **32**, No. 3 (1982), 325–341.

[V6] L. P. Vlasov, Subspaces of finite codimension: Existence of elements of best approximation, *Mat. Zametki* **37**, No. 1 (1985), 78–85.