Journal of Approximation Theory **94**, 203–221 (1998)
Article No. AT983201

Proximal Mappings

Jean-Paul Penot

*Laboratoire de Mathématiques Appliquées, URA 1204 CNRS,
Université de Pau et des Pays de l'Adour,
Av. de l'Université, 64000 Pau, France*

Communicated by Aldric L. Brown

Received June 11, 1996; accepted in revised form August 18, 1997

We study differentiability properties and subdifferentiability properties of the Baire approximate and of the Moreau–Yosida approximate of a nonconvex function on a Banach space. These properties are intimately linked with exactness of the infimal convolution defining the approximation. When applied to indicator functions of possibly nonconvex subsets, our results yield existence of best approximations under subdifferentiability assumptions on the distance function and suitable smoothness assumptions on the space. © 1998 Academic Press

Two kinds of approximations of functions are widely used. The first one is the integral convolution by mollifiers. It is particularly used in the study of partial differential equations. The second one is the infimal convolution method of Baire and Moreau and Yosida (see [1, 26]). It is not limited to the finite dimensional case. It is usually used for convex functions, but as shown in [13] and some of its references, its domain of application can be extended beyond this case (see also [2]).

It has been shown in these two references that a key assumption on the function f to be regularized is its growth property. For this reason, it is advisable to consider a regularization using a general kernel as in [13]. In such a way, a kernel adapted to the growth of the function can be chosen. Here, for $f : X \to \mathbb{R}^{\bullet} := \mathbb{R} \cup \{\infty\}$, lower semicontinuous (l.s.c.) on the Banach space X, and $r > 0$, we set

$$f_r(w) := \inf_{x \in X} (f(x) + r^{-1}k(w-x)),$$

where $k = h \circ j$, with $j(x) = \|x\|$, $h : \mathbb{R}_+ \to \mathbb{R}$ of class C^1, nonnegative and nondecreasing on $\mathbb{P} := (0, \infty)$ and such that $h(0) = 0$, $h(t) \to \infty$ when $t \to \infty$. The usual cases are $h(r) = r$ (Baire regularization), $h(r) = (1/2)\, r^2$ (Moreau–Yosida regularization), and more generally $h = h_p$ with $h_p(r) = (1/p)\, r^p$ for $p \geqslant 1$ and $h = h_e$ with $h_e(r) = \exp r - 1$. Combinations of these

0021-9045/98 $25.00

cases can also be considered; for instance one can take $h(r) := br^p$ for $r \in [0, a]$, $h(r) := cr^q + d$ for $r > a$, with appropriate positive numbers a, b, c, d.

The *proximal* (or prox) multimapping associated with f is given by

$$P_r^f(w) := \{x \in X : f(x) + r^{-1}k(w - x) = f_r(w)\} = P_1^{rf}(w).$$

When f is *the indicator function* i_C of a closed subset C of X (given by $i_C(x) = 0$ for $x \in C$, $i_C(x) = \infty$ for $x \in X \setminus C$), one has $f_r(w) = r^{-1}h(d_C(w))$ where $d_C(w) = \inf_{x \in C} d(w, x)$ and $P_r^f(w)$ is nothing but the set of best approximations $P_C(w)$ of w in C. In such a case the multimapping has been widely studied and used. In general it is a multimapping (or correspondence) which also has a considerable interest. In particular it is an essential tool for the study of the proximal algorithm (see [23] for a recent account and references).

Conditions ensuring that the values of $P_r^f(\cdot)$ are at most singletons are easy to find. As in [24] we are especially interested in conditions ensuring it has nonempty values. In the special case just mentioned, that means we are looking for conditions ensuring existence of best approximations. We show that subdifferentiability in the Fréchet sense of f_r is such a condition in a natural class of Banach spaces including the Hilbert space (Section 3). We also consider the case of Hadamard (or contingent) subdifferentiability. Our results generalize previous results of [11, 21, 27] related to the case of distance functions instead of general regularized functions and of usual differentiability instead of subdifferentiability as here. While preparing the list of references of the present paper we became aware of the paper [16] in which related results are proved in a Hilbert space framework for the distance function and with different subdifferentiability notions (see also the thesis [34] prepared under the guidance of the author).

Let us observe that since f_r is devised in order to regularize f, differentiability or subdifferentiability assumptions on it are quite natural; on the other hand, the knowledge of derivatives or subderivatives of f_r requires an a priori study of the proximal multimapping $P_r^f(\cdot)$. We also observe that our object of study bears some similarity with the process known as Tychonov regularization; this fact is reflected by our use of the notion of radius of essential minimization of a function. In the case of the indicator function of a subset, this notion could not be distinguished, as the distance to the set is part of the picture in such a case.

1. PRELIMINARIES

Proximal multimappings and distance functions have properties which deserve interest, even in the nonconvex case. Using convex analysis techniques

Moreau [26, Proposition 7d] has shown that the proximal mapping associated to any function f is monotone. On the other hand, E. Asplund [4, 5] has proved that if C is a closed subset of a Hilbert space X then $\frac{1}{2}d_C^2 - \frac{1}{2}\|\cdot\|^2$ is a concave function and that the projection mapping is monotone. Synthesizing both works, let us note here that the proximal multimapping associated to any (nonconvex) function f is monotone when the kernel is the Moreau–Yosida kernel.

LEMMA 1.1. *The Moreau–Yosida proximal multimapping P_r^f associated with any function $f: X \to \mathbb{R}^\bullet$ on a Hilbert space X is a monotone relation.*

Proof. Let $x \in P_r^f(w)$, $x' \in P_r^f(w')$ with $w, w' \in X$. By definition we have

$$f(x) + \frac{1}{2r}\|w - x\|^2 \leqslant f(x') + \frac{1}{2r}\|w - x'\|^2$$

$$f(x') + \frac{1}{2r}\|w' - x'\|^2 \leqslant f(x) + \frac{1}{2r}\|w' - x\|^2$$

hence, by addition, $f(x)$ and $f(x')$ being finite (unless $f \equiv +\infty$),

$$\|w - x\|^2 - \|w' - x\|^2 \leqslant \|w - x'\|^2 - \|w' - x'\|^2$$

or

$$0 \leqslant (w - w' \mid x - x'). \quad \blacksquare$$

It would be interesting to know whether the preceding argument can be extended to other kernels and to non-Hilbertian spaces. But this is not our aim here. Let us just observe that in this special case P_r^f is known to be cyclically monotone as contained in the subdifferential of the Asplund function

$$\alpha_{rf}(w) := (rf + k)^*(w)$$

(where $(rf + k)^*$ is the Fenchel conjugate of $(rf + k)$) while $-f_r$ is paraconvex (see [2] for instance):

$$f_r(w) = r^{-1}k(w) - r^{-1}\alpha_{rf}(w).$$

Throughout we suppose f is l.s.c. non-improper (i.e., assumes at least one finite value) and that its regularized function f_r takes a finite value at some given point w of X. This assumption is fulfilled for $w=0$ when f satisfies the growth condition (which is in fact a non-decay condition)

$$\text{there exist } a>0,\ b\in\mathbb{R}_+ \text{ such that } f(x)\geqslant b-ak(x) \text{ for each } x\in X, \tag{G}$$

and we take $r\in(0,a^{-1})$. Such a condition is obviously satisfied for any $r>0$ when f is an indicator function.

Furthermore, under a mild condition on h, the function f_r takes only finite values. Thus, in the sequel we assume h satisfies the hypothesis

$$\begin{aligned}&\text{For any } c\in(0,1),\ d\in\mathbb{R}_+ \text{ there exists } m\in\mathbb{R}_+ \text{ such that}\\ &h(t)\geqslant ch(t+d)-m \qquad \forall t\in\mathbb{R}_+.\end{aligned} \tag{H}$$

This condition is satisfied when $h(t)=h_p(t)=(1/p)\,t^p$. Appropriate modifications of what follows (taking restrictions on balls) would enable one to replace (H) by the following weaker assumption:

$$\begin{aligned}&\text{For any } d\in\mathbb{R}_+ \text{ there exist } c\in(0,1) \text{ and } m\in\mathbb{R}_+ \text{ such that}\\ &h(t)\geqslant ch(t+d)-m \qquad \forall t\in\mathbb{R}_+.\end{aligned} \tag{H_0}$$

This condition is satisfied by the function h_e given by $h_e(t)=\exp t-1$.

It can be shown that any function f satisfying the growth condition (G) is such that $f_r(w)>-\infty$ for each $w\in X$ whenever $r\in(0,a^{-1})$. In fact we will establish a more precise result in Lemma 1.3 below.

It will be convenient to introduce the following terminology. A function $f: X\to\mathbb{R}\cup\{\infty\}$ is said to be *finitely minimizable* if there exists $r\geqslant 0$ such that for each $t>m_f:=\inf f(X)$ one has

$$rB_X\cap[f\leqslant t]\neq\varnothing,$$

where $[f\leqslant m]=\{x\in X: f(x)\leqslant m\}$ and B_X is the closed unit ball of X. The infimum of such r's is called in [29] the radius of essential minimization of f and is denoted by $\rho(f)$. It is shown in [29] that

$$\rho(f)=\inf\{\operatorname{rad}(x_n):(x_n)\in\mathcal{M}(f)\},$$

where $\mathcal{M}(f)$ denotes the set of minimizing sequences of f and $\operatorname{rad}(x_n)=\liminf_n\|x_n\|$. A minimizing sequence (x_n) such that $(\|x_n\|)\to\rho(f)$ will be called an *essential minimizing sequence*.

It is easy to show that any coercive function is finitely minimizable.

LEMMA 1.2. *Suppose that f is coercive* (*i.e.* $\lim_{\|x\|\to\infty} f(x)=\infty$) *or, more generally, semicoercive* (*i.e.* $f(\infty):=\liminf_{\|x\|\to\infty} f(x)>\inf f(X)=: m_f$). *Then f is finitely minimizable.*

Proof. For any $s \in (m_f, f(\infty))$ there exists $r \geqslant 0$ such that $f(x) > s$ for $x \in X \backslash rB_X$, so that for each $t > m_f$, setting $s = \min(t, \frac{1}{2}(m_f + f(\infty)))$ one gets $rB_X \cap [f \leqslant t] \supset [f \leqslant s] \neq \emptyset$. ▮

The example of $X = \mathbb{R}$, $f(x) = x^2 \exp(-x^2)$ shows that a finitely minimizable function may be non-semicoercive. When the set S_f of minimizers of f is nonempty one has $\rho(f) \leqslant \inf\{\|x\| : x \in S_f\}$; it is easy to find examples showing that this inequality may be strict. When X is reflexive and f is finitely minimizable and weakly lower semicontinuous (l.s.c.), the set S_f is nonempty; in that case equality holds.

LEMMA 1.3. *Suppose f satisfies the growth condition* (G). *Then for each $w \in X$ and for each $r \in (0, a^{-1})$ the function $f_{r,w}$ given by*

$$f_{r,w}(y) := f(w-y) + r^{-1}k(y)$$

is coercive, hence finitely minimizable and $f_r(w)$ is finite.

Proof. Let $w \in X$, $d \geqslant \|w\|$, $r \in (0, a^{-1})$, $x = w - y$. Then, taking $c \in (ar, 1)$, we have for some $m \in \mathbb{R}$

$$\begin{aligned} f_{r,w}(y) &\geqslant b - ah(\|x\|) + r^{-1}h(\|x\| - d) \\ &\geqslant b - ah(\|x\|) + r^{-1}ch(\|x\|) - mr^{-1} \\ &\geqslant b - r^{-1}m + r^{-1}(c - ar)\, h(\|x\|) \end{aligned}$$

by assumption (H). Since $h(\|w-y\|) \to \infty$ as $\|y\| \to \infty$, $f_{r,w}$ is coercive. ▮

The preceding proof shows that for each $d \in \mathbb{R}_+$ the function f_r is bounded below on the ball dB_X. It also shows that given $d \in \mathbb{R}_+$ there exists $d' \in \mathbb{R}_+$ such that for each $w \in dB_X$

$$f_r(w) = \inf\{f(x) + r^{-1}h(\|w-x\|) : x \in d'B_X\}$$

so that f_r is Lipschitzian on dB_X (see also [13]). Then, if the space X has a Fréchet differentiable norm (resp. is an Asplund space and f is convex) f_r is densely (resp. generically) Fréchet subdifferentiable [33]. Let us recall basic facts about this notion and related properties.

Given $f : X \to \mathbb{R} \cup \{\infty\}$ finite at $w \in X$, the *firm* (or Fréchet) *subdifferential* of f at w is the set $\partial^- f(w)$ of $w^* \in X^*$ such that for each $\varepsilon > 0$ there exists $\delta > 0$ such that

$$f(x) \geqslant f(w) + \langle w^*, x - w \rangle - \varepsilon \|x - w\|$$

for $x \in B(w, \delta)$. This set is contained in the *contingent* (*or Hadamard*) *subdifferential* $\partial f(w)$ of f at w which is the set of $w^* \in X^*$ such that $\langle w^*, u \rangle \leqslant f'(w, u)$ for each $u \in X$, where

$$f'(w, u) := \liminf_{(t, v) \to (0_+, u)} t^{-1}(f(w + tv) - f(w))$$

is the lower (or contingent) derivative of f at w. Both $\partial^- f(w)$ and $\partial f(w)$ are reduced to the singleton $\{f'(w)\}$ when f is Fréchet differentiable at w. Moreover, when f is convex, they coincide with the subdifferential of convex analysis.

Let us present an elementary estimate on the elements of $\partial f(w)$ which we will use later. To this aim, recall that $f : X \to \mathbb{R}^{\bullet}$ is said to be *calm* at x_0 if it is finite at x_0 and if there exist $r > 0$, $c \geqslant 0$ such that for $x \in B(x_0, r)$ one has

$$f(x) \geqslant f(x_0) - c \, \|x - x_0\|,$$

or, equivalently, if $\liminf_{x \to (\neq) x_0} \|x - x_0\|^{-1} (f(x) - f(x_0)) > -\infty$. The function f is said to be *quiet* at x_0 if $-f$ is calm at x_0. The infimum of the family of constants c as above (i.e., $-\liminf_{x \to x_0} \|x - x_0\|^{-1} (f(x) - f(x_0))$ is called the *constant of calmness* of f at x_0; the constant of quietness of f at x_0 is the constant of calmness of $-f$ at x_0.

Lemma 1.4. *Suppose f is quiet at x_0 with constant of quietness c. Then for any $x_0^* \in \partial f(x_0)$ one has $\|x_0^*\| \leqslant c$.*

Proof. This follows immediately from the inequalities

$$\langle x_0^*, u \rangle \leqslant f'(x_0, u) \leqslant c \, \|u\| \qquad \forall u \in X. \blacksquare$$

The following lemma could be given in the general framework of performance functions but we prefer to keep close to the specific case we study.

Lemma 1.5. *Suppose the function $f_{r,w}$ given by $f_{r,w}(y) := f(w - y) + r^{-1}k(y)$ is finitely minimizable. Then f_r is quiet at w and its rate of quietness is majorized by $r^{-1}h'(\rho)$, where $\rho = \rho(f_{r,w})$ is the radius of essential minimization of $f_{r,w}$. In particular, for any $w^* \in \partial f_r(w)$ one has $\|w^*\| \leqslant r^{-1}h'(\rho)$.*

Proof. Let (y_n) be a minimizing sequence of $f_{r,w}$ such that $(\|y_n\|) \to \rho$ (such a sequence does exist). Then, for some sequence (ε_n) of $\mathbb{R}_+$ with limit 0 one has for any $v \in X$

$$f_r(w) \geqslant f(w - y_n) + r^{-1}k(y_n) - \varepsilon_n,$$

$$f_r(w + v) \leqslant f(w - y_n) + r^{-1}k(v + y_n),$$

so that, h being of class C^1 around ρ and nondecreasing

$$\begin{aligned}\limsup_{\|v\|\to 0_+} \frac{1}{\|v\|}\,(f_r(w+v)-f_r(w)) &\leqslant r^{-1} \limsup_{\|v\|\to 0_+} \frac{1}{\|v\|}\,(h(\|y_n\|+\|v\|)-h(\|y_n\|))\\ &\leqslant r^{-1}h'(\rho).\end{aligned}$$

The last assertion then follows from Lemma 1.4. ∎

For elements of the Fréchet subdifferential one can be more precise.

LEMMA 1.6. *For any* $w^* \in \partial^- f_r(w)$ (*in particular for any* $w^* \in \partial f_r(w)$ *when* f_r *is convex*) *with* $\rho := \rho(f_{r,w}) > 0$ *one has* $\|w^*\| = r^{-1}h'(\rho)$.

Proof. The proof is inspired by the corresponding one with the distance function in [11]: when one takes for f an indicator function the result reduces to [11] Theorem 11. Let (y_n) be a minimizing sequence of $f_{r,w}$ such that $(\|y_n\|) \to \rho(f_{r,w})$. We may suppose $y_n \neq 0$ for each n, otherwise $w \in P_r^f(w)$ and $\rho = 0$. Let $(t_n) \to 0_+$ be such that

$$f(w-y_n) + r^{-1}k(y_n) \leqslant f_r(w) + t_n^2.$$

Since for each n we have

$$f_r(w - t_n y_n) \leqslant f(w-y_n) + r^{-1}k(y_n - t_n y_n)$$

and since for each bounded subset M of X (in particular for $M = \{-y_n : n \in \mathbb{N}\}$) we have

$$\liminf_{t\to 0_+} \inf_{z\in M} t^{-1}(f_r(w+tz) - f_r(w) - \langle w^*, tz\rangle) \geqslant 0$$

we get

$$\liminf_n t_n^{-1}(r^{-1}h((1-t_n)\,\|y_n\|) - r^{-1}h(\|y_n\|) - \langle w^*, -t_n y_n\rangle) \geqslant 0$$

hence

$$\liminf_n \left\langle w^*, \frac{y_n}{\|y_n\|}\right\rangle \geqslant r^{-1}h'(\rho). \tag{1}$$

As $\|w^*\| \leqslant r^{-1}h'(\rho)$ by the preceding lemma, we can conclude that $\|w^*\| = r^{-1}h'(\rho)$. ∎

2. CONSEQUENCES OF EXACT PROXIMATION

In this section we suppose the proximal regularized function f_r is *exact* at w for some $r>0$, $w\in X$, i.e., that the set

$$P(w) := P_r^f(w) := \{x\in X : f(x)+r^{-1}k(w-x)=f_r(w)\}$$

is nonempty and we draw some easy consequences pertaining to the subdifferentials and the super-differentials of f and f_r.

This exactness property is satisfied when X is a dual space and $f_{r,w}$ is weak* lower semicontinuous (l.s.c.) and finitely minimizable, in particular when X is reflexive and f satisfies the growth condition (G) and is quasiconvex (the sum of two l.s.c. functions being l.s.c.).

Our first observation is an easy generalization to the nonconvex case of a well-known result of convex analysis (see [22] for instance). It could be deduced from general results about performance functions. Here, given proper functions $f, g : X\to \mathbb{R}\cup\{\infty\}$ we set

$$(f\,\Box\, g)(w) := \inf\{f(x)+g(w-x) : x\in X\}$$

$$P(w) := P_{f,g}(w) := \{x\in X : f(x)+g(w-x) = (f\,\Box\, g)(w)\}.$$

In the sequel we will take $g=r^{-1}k$ fixed and we will use the following easy lemma.

LEMMA 2.1. *Suppose $f\,\Box\, g$ is finite at w and $P(w)$ is nonempty. Then*

$$\partial(f\,\Box\, g)(w)\subset \bigcap_{x\in P(w)} \partial f(x)\cap \partial g(w-x),$$

$$\partial^-(f\,\Box\, g)(w)\subset \bigcap_{x\in P(w)} \partial^- f(x)\cap \partial^- g(w-x).$$

Proof. Let us prove the first inclusion, the proof of the second one being similar (see also [17, Lemma 3.6]). In view of the symmetry of the roles of f and g it suffices to show that any $w^*\in\partial(f\,\Box\, g)(w)$ belongs to $\partial g(w-x)$. By definition, for each $u\in X$ there exists a function $\varepsilon : \mathbb{R}_+\times X\to \mathbb{R}_+\cup\{\infty\}$ such that $\lim_{(t,v)\to(0_+,u)}\varepsilon(t,v)=0$ such that for any $t\in\mathbb{R}_+$, $v\in X$

$$\langle w^*, tv\rangle \leqslant (f\,\Box\, g)(w+tv)-(f\,\Box\, g)(w)+\varepsilon(t,v)\,t.$$

Using the relations

$$(f\,\Box\, g)(w+tv)\leqslant f(x)+g(w+tv-x)$$

$$(f\,\Box\, g)(w) = f(x)+g(w-x)$$

we get

$$\langle w^*, tv \rangle \leqslant g(w-x+tv) - g(w-x) + \varepsilon(t, v)\, t,$$

so that $w^* \in \partial g(w-x)$. ∎

Taking $g = r^{-1}k$ we get the following consequence about regularized functions.

PROPOSITION 2.2. *Suppose that for some $r \in (0, a^{-1})$ and some $w \in X$ the set $P_r^f(w)$ is nonempty. Then, for each $x \in P_r^f(w)$, one has*

$$\partial f_r(w) \subset r^{-1}\, \partial k(w-x) \cap \partial f(x).$$

In particular, if the norm $j = \|\cdot\|$ is Gâteaux-differentiable at $w-x$ and if f_r is Hadamard-subdifferentiable at w then

$$\partial f_r(w) = r^{-1} h'(\|w-x\|)\, S(w-x),$$

where, for $u \in X$, $S(u) := \partial j(u) = \{u^* \in S_{X^*} : \langle u^*, u \rangle = \|u\|\}$. If ρ is the radius of essential minimization of the function $f_{r,w}$ and if there exists some $x \in P_r^f(w)$ with $\|x-w\| = \rho$ (in particular if X is reflexive and if f is weakly l.s.c.) then for any $w^* \in \partial f_r(w)$ one has $\|w^*\| = r^{-1}h'(\rho)$.

When X is infinite dimensional and f is not an indicator function it may happen that $P_r^f(w)$ is nonempty but that $P_r^f(w) \cap B(w, \rho) = \varnothing$.

EXAMPLE. Let (x_n) be a sequence without cluster point of the unit sphere S_X of X and let f be given by $f(x_n) = r_n$, with $(r_n) \to 0_+$, $f(x) = -1$ for $x \in 2S_X$ and $f(x) = \infty$ otherwise. Then for $w = 0$, $r = 1$, $h(t) = t$ we have $\rho = 1$, $P_r^f(w) = 2S_X$ hence $P_r^f(w) \cap B(w, \rho) = \varnothing$.

Taking for f an indicator function and setting $r = 1$, $h(t) = t$, we get a simple consequence which generalizes results of [8, 21, 36].

COROLLARY 2.3. *Let C be a nonempty closed subset of X and let $w \in X \backslash C$ be such that $P_C(w)$ and $\partial d_C(w)$ are nonempty. Let $N(C, x)$ be the normal cone to C at x. Then, for each $x \in P_C(w)$ one has*

$$\partial d_C(w) \subset S(w-x) \cap N(C, x) \subset S_{X^*}.$$

COROLLARY 2.4. *Let C be a nonempty closed subset of a strictly convex Banach space X. Then, for each $w \in W \backslash C$ such that d_C is H-subdifferentiable, the set $P_C(w)$ is at most a singleton.*

Proof. This follows from the fact $u=v$ whenever $S(u)\cap S(v)\neq\varnothing$ when X is strictly convex. ▮

This uniqueness result can be extended to proximal multimappings.

COROLLARY 2.5. *Suppose that for some $r\in(0,a^{-1})$ and some $w\in X$ the function f_r is a Hadamard subdifferentiable at w. Suppose h' is positive and increasing on $(0,\infty)$ and the norm is strictly convex. Then the set $P_r^f(w)$ is at most a singleton.*

Proof. Given $w^*\in\partial f_r(w)$, for each $x\in P_r^f(w)$ we must have

$$w^*\in r^{-1}\,\partial k(w-x).$$

Then, if $x\neq w$ one has $\partial k(w-x)=h'(\|w-x\|)\,S(w-x)$, so that the preceding inclusion yields

$$rw^*\in h'(\|w-x\|)\,S(w-x),$$

hence $r\,\|w^*\|=h'(\|w-x\|)$, a relation which determines $t:=\|w-x\|$. Then one cannot have $w\in P_r^f(w)$ which would imply $rw^*\in h'(0)\,B_{X^*}$, $r\,\|w^*\|\leqslant h'(0)<h'(t)$. Since $S(w-x)=h'(\rho)^{-1}\,rw^*$, and since S is injective on spheres centered at 0, as easily checked, the determination of x is complete. ▮

Now let us turn to superdifferentials. We recall that the Hadamard superdifferential of f at x is the set

$$\breve{\partial} f(x)=-\partial(-f)(x);$$

a similar notation can be used for the Fréchet superdifferential $\breve{\partial}^- f(x)$.

LEMMA 2.6. *Given $w\in X$ such that $f\,\Box\,g$ is finite and exact at w, given $x\in P_r^f(w)$, for any $z^*\in\breve{\partial} g(w-x)$ (resp. $z^*\in\breve{\partial}^- g(w-x)$) one has*

$$z^*\in\breve{\partial}(f\,\Box\,g)(w)\qquad (\textit{resp. } z^*\in\breve{\partial}^-(f\,\Box\,g)(w)).$$

Proof. This time we just prove the Fréchet case. Then, for $z^*\in\breve{\partial}^- g(w-x)$ we can find a modulus $\varepsilon(\cdot)$, i.e., a function $\varepsilon(\cdot)$ satisfying $\lim_{t\to 0}\varepsilon(t)=0$, such that

$$g(w+u-x)-g(w-x)\leqslant\langle z^*,u\rangle+\varepsilon(\|u\|)\,\|u\|.$$

Then, as $(f \,\square\, g)(w) = f(x) + g(w-x)$ we get

$$\begin{aligned}
&(f \,\square\, g)(w+u) - (f \,\square\, g)(w) \\
&\quad \leqslant (f(x) + g(w+u-x)) - (f(x) + g(w-x)) \\
&\quad \leqslant \langle z^*, u \rangle + \varepsilon(\|u\|)\,\|u\|. \quad \blacksquare
\end{aligned}$$

Applying this lemma to our regularization process we get the following result.

PROPOSITION 2.7. *Suppose for some* $r \in (0, a^{-1})$ *and some* $w \in X$ *the set* $P_r^f(w)$ *is nonempty. Let* $x \in P_r^f(w)$ *be such that* $k = h \circ j$ *is H-differentiable (resp. F-differentiable) at* $w - x$. *Then*

$$r^{-1} k'(w-x) \in \breve{\partial} f_r(w) \qquad (\textit{resp.} \quad r^{-1} k'(w-x) \in \breve{\partial}^- f_r(w)).$$

In particular, if j *is G-differentiable (resp. F-differentiable) at* $w - x$ *then*

$$r^{-1} h'(j(w-x))\, S(w-x) \in \breve{\partial} f_r(w) \qquad (\textit{resp.}\ \breve{\partial}^- f_r(w)).$$

COROLLARY 2.8. *Suppose with the assumptions of the preceding proposition that* f_r *is H-subdifferentiable (resp. F-subdifferentiable) at* w *(this is the case when* f_r *is convex). Then* f_r *is H-differentiable (resp. F-differentiable) at* w *and*

$$f_r'(w) = r^{-1} k'(w-x).$$

Proof. This follows from the fact that if $w^* \in \breve{\partial} f_r(w)$ and if $z^* \in \partial f_r(w)$ then $w^* = z^*$ and f_r is H-differentiable at w. $\blacksquare$

Let us show now that a continuity assumption on the proximal multimapping P_r^f entails a strong differentiability property of the regularized functions. Let us recall that a function f is said to be *strictly differentiable* at x if it is Fréchet differentiable at x and if

$$\|u - v\|^{-1}\,(f(u) - f(v) - f'(x)(u-v)) \to 0$$

as $u, v \to x$, $u \neq v$. This condition is satisfied if f is continuously differentiable at x, i.e., when f is differentiable on a neighborhood of x and its derivative f' is continuous at x.

We will also need the notion of *lower semicontinuity* (l.s.c.) of a multimapping $M : W \rightrightarrows X$ between two topological spaces W, X: M is said to be l.s.c. at $(w, x) \in M$ when for each neighborhood V of x there exists a neighborhood U of w such that $M(u) \cap V \neq \emptyset$ for each $u \in U$. When X is metrizable this condition amounts to $d(x, M(u)) \to 0$ as $u \to w$.

LEMMA 2.9. *Suppose the multimapping $P := P_{f,g}$ is l.s.c. at (w, x) and g is strictly differentiable at $w-x$. Then $f \,\Box\, g$ is strictly differentiable at w and*

$$(f \,\Box\, g)'(w) = g'(w-x).$$

Proof. By assumption, we can find a neighborhood V of w in X and a mapping $v \mapsto x_v$ from V to X such that $x_v \in P(v)$ for $v \in V$ and $x_v \to x$ as $v \to w$.

Substituting v to w and x_v to x in the estimates of Lemma 2.6 we obtain, as g is strictly differentiable at $w-x$ and $x_v \to x$ as $v \to w$

$$\begin{aligned}(f \,\Box\, g)(v+u) - (f \,\Box\, g)(v) &\leqslant g(v+u-x_v) - g(v-x_v)\\ &\leqslant g'(w-x)(u) + \varepsilon(v, u)\, \|u\|\end{aligned}$$

with $\varepsilon(v, u) \to 0$ as $v \to w$, $u \to 0$. Setting $u' = -u$, $v' = v+u$ we get an inequality in the opposite direction, and we have shown strict differentiability of $f \,\Box\, g$ at w. ∎

COROLLARY 2.10. *Suppose the norm j is Fréchet differentiable on $X_0 = X \backslash \{0\}$ and for some $r \in (0, a^{-1})$, $w \in W$, $x \in P_r^f(w)$ the multimapping P_r^f is l.s.c. at (w, x). Then the regularized function f_r is strictly differentiable at w and*

$$f_r'(w) = r^{-1} h'(\|w-x\|)\, S(w-x).$$

The lower semicontinuity assumption (which can be dropped when f_r is convex) is stringent, especially in the nonconvex case, but it is satisfied in a number of cases.

EXAMPLES. (a) Let $f = i_C$, where $C = X \backslash U$, U being the open unit ball of the Hilbert space X. Then $P_C := P_r^f$ is not l.s.c. at 0 but it is l.s.c. at $(w, P_C(w))$ for each $w \in X \backslash \{0\}$.

(b) Let $X = \mathbb{R}^2$, $f = i_C$ where $C = \mathbb{R}^2 \backslash \mathbb{P}^2$, with $\mathbb{P} = (0, \infty)$. Then P_r^f is l.s.c. at each point $X \backslash \Delta_+$ where $\Delta_+ = \{(r, r) : r \in \mathbb{P}\}$.

(c) When X is an uniformly convex Banach space, it can be shown that the projection mapping on convex subsets is uniformly continuous on bounded subsets (see [30]) so that the lower semicontinuity assumption is satisfied.

(d) This assumption is also satisfied if X is a strictly convex reflexive Kadec space and f is convex.

(e) When X is a Hilbert space and f is convex, P_r^f is nonexpansive [26].

3. EXISTENCE OF PROXIMAL POINTS

In this section we examine the consequences of subdifferentiability of the regularized function f_r. We have seen in the preceding section that super-differentiability of f_r is for free when the norm j is assumed to be differentiable and $P_r^f(w)$ is nonempty. On the contrary, subdifferentiability of f_r is a stringent assumption. We note however that this assumption is satisfied at a dense set of points when f_r is convex and X is an Asplund space or when f_r is l.s.c. and X has a Lipschitzian differentiable bump function.

In the sequel we denote by WP the set of well-posed linear forms on X, in other words, the set of $x^* \in X^*$ which firmly (or strongly) exposes B_X, i.e., such that each maximizing sequence of x^* in B_X converges. It follows from the general theory of well-posed optimization problems that any $x^* \in WP$ attains its maximum on B_X at a unique point. We also denote by GWP the set of generalized well-posed linear forms on X, i.e., the set of $x^* \in X^*$ such that any maximizing sequence (x_n) of x^* in the unit ball B_X has a converging subsequence. Then

$$S_*(x^*) := S^*(x^*) \cap X := \{x \in S_X : \langle x^*, x \rangle = \|x^*\|\}$$

is nonempty, S_X being the unit sphere of X, S^* being the subdifferential of the dual norm. The following definition will be convenient.

DEFINITION 3.1. The Banach space X (or rather $(X, \|\cdot\|)$ will be said to be metrically reflexive (in short M-reflexive) if $X^* = GWP \cup \{0\}$.

This property obviously bears on the metric structure of X; by the theorem of James, it implies that X is reflexive. On the other hand, the theorem of Trojanski ensures that any reflexive Banach space can be renormed into an M-reflexive space. In fact, any Banach space whose dual norm is Fréchet differentiable off 0 is M-reflexive in view of the Šmulyan Theorem [19, Theorem 1.4(ii), p. 3]. In particular, if X is reflexive and has the Kadec–Klee property (i.e., weak convergence and norm convergence coincide in the unit sphere) then X is M-reflexive. If moreover X is strictly convex then $WP \cup \{0\} = X^*$. The class of M-reflexive spaces is stable under quotients (with the quotient norms) and products (with the sup norm).

Without any assumption on X one can assert that $x^* \in WP$, hence $x^* \in GWP$, whenever the dual norm on X^* is Fréchet differentiable at x^* [19]. Moreover, it is known [32, 35] that for any Banach space X verifying the Radon–Nikodym property the sets GWP and WP are generic subsets of X^* (i.e., dense $\mathcal{G}_\delta$ subsets of X^*).

THEOREM 3.1. *Suppose f is l.s.c. and $f_r(w)$ is finite for some $r>0$ and some $w \in X$. Suppose $w^* \in \partial f_r(w) \cap GWP$, $\|w^*\| = r^{-1}h'(\rho)$ where ρ is the radius of essential minimization of $f_{r,w}$ and there exists $z \in S_*(w^*)$ at which the norm j of X is Fréchet differentiable. Then $P_r^f(w)$ is nonempty. Moreover, it contains some point x such that $\|x-w\| = \rho$ and any essential minimizing sequence of $f_{r,w}$ has a limit point in $P_r^f(w)$. If moreover $w^* \in WP$, in particular if X is strictly convex, then any essential minimizing sequence of $f_{r,w}$ converges to the unique element of $P_r^f(w)$.*

Proof. When $\rho = 0$ we have $w \in P_r^f(w)$, the function $f_{r,w}$ being l.s.c. Thus we suppose $\rho > 0$.

Let $z \in S_*(w^*)$ and let $(z_n) \to z$, $(t_n) \to 0_+$ be such that the contingent derivative $f'_r(w, z)$ is given by

$$f'_r(w, z) = \lim t_n^{-1}(f_r(w + t_n z_n) - f_r(w)),$$

and let $y_n \in X$ be such that $(\|y_n\|) \to \rho$, $f(w - y_n) + r^{-1}k(y_n) \leqslant f_r(w) + t_n^2$. Then,

$$\begin{aligned} &f_r(w + t_n z_n) - f_r(w) \\ &\quad \leqslant f(w - y_n) + r^{-1}k(y_n + t_n z_n) - (f(w - y_n) + r^{-1}k(y_n) - t_n^2) \\ &\quad \leqslant r^{-1}(h(\|y_n + t_n z_n\|) - h(\|y_n\|)) + t_n^2 \end{aligned}$$

and the definitions give

$$\begin{aligned} \|w^*\| &= \langle w^*, z \rangle \leqslant f'_r(w, z) \\ &\leqslant \liminf_n t_n^{-1} r^{-1}(h(\|y_n + t_n z_n\|) - h(\|y_n\|)). \end{aligned}$$

The Lebourg Mean Value Theorem and the chain rule for Clarke's subdifferentials of locally Lipschitzian functions [15, Theorems 2.3.7 and 2.3.9] yields some $s_n \in [0, t_n]$ and some

$$w_n^* \in r^{-1}\partial k(y_n + s_n z_n) \subset r^{-1}h'(\|y_n + s_n z_n\|)\, S(y_n + s_n z_n),$$

where $S(x) := \partial j(x)$, such that

$$k(y_n + t_n z_n) - k(y_n) = \langle r w_n^*, t_n z_n \rangle$$

for each $n \in \mathbb{N}$. It follows that

$$\begin{aligned} \|w^*\| &\leqslant \liminf \langle w_n^*, z_n \rangle = \liminf \langle w_n^*, z \rangle \\ &\leqslant \limsup \|w_n^*\| = \limsup r^{-1}h'(\|y_n + s_n z_n\|) = r^{-1}h'(\rho). \end{aligned}$$

Since j is Fréchet differentiable at z, the Šmulyan Theorem ensures that z strongly exposes $\|w^*\| B_{X^*}$ at w^* and (w_n^*) converges to w^*.

Now, setting $\rho_n := \|y_n\|$, $u_n := \rho_n^{-1} y_n$, we have by our assumptions

$$r^{-1}h'(\rho) = \|w^*\| \geqslant \langle w^*, u_n \rangle \geqslant \langle w_n^*, u_n \rangle - \|w^* - w_n^*\|.$$

Since $(w_n^*) \to w^*$ and since $w_n^* \in r^{-1}h'(\|y_n + s_n z_n\|)\, S(u_n + \rho_n^{-1} s_n z_n)$ we get that

$$r^{-1}h'(\rho) \geqslant \limsup \langle w_n^*, u_n \rangle \geqslant \liminf \langle w_n^*, u_n \rangle = \liminf \langle w_n^*, u_n + \rho_n^{-1} s_n z_n \rangle \geqslant \lim r^{-1}h'(\|y_n + t_n z_n\|),$$

hence $(\langle w^*, u_n \rangle) \to r^{-1}h'(\rho)$. As w^* is generalized well-posed on B_X we get that (u_n) has a converging subsequence. It follows that $x_n := w - \rho_n u_n$ converges to $x := w - \rho u$ where $u = \lim u_n$; as f is l.s.c. we get $x \in P_r^f(w)$. When $w^* \in WP$ the whole sequence (u_n) converges. ∎

We observe that the Šmulyan Theorem ensures that the assumption on w^* is satisfied whenever $w^* \in \partial f_r(w) \cap J(F) \cap F^*$, where $J = \partial \frac{1}{2} j^2$, F (resp. F^*) denoting the set of points at which the norm j (resp. the dual norm) is Fréchet differentiable.

Corollary 3.2. *Suppose $f_r(w)$ is finite for some $r > 0$ and some $w \in X$. Suppose $\partial f_r(w)$ contains an element with norm $r^{-1}h'(\rho)$ and the norms of X and X^* are Fréchet differentiable off 0. Then $P_r^f(w)$ is nonempty.*

Using the observations preceding the theorem we get the following consequence.

Corollary 3.3. *Suppose f is l.s.c. and satisfies the growth condition* (G). *Let $r \in (0, a^{-1})$, $w \in X$. Suppose X is M-reflexive and the norm of X is Fréchet differentiable off 0. If $\partial f_r(w)$ contains a non-zero element with norm $r^{-1}h'(\rho)$ then $P_r^f(w)$ is nonempty.*

Taking $h = I_{\mathbb{R}}$, $r = 1$, $f = i_C$ for a nonempty closed subset C of X we get an existence result for best approximation which extends a result in [21] in which d_C is assumed to be Gâteaux differentiable at w, the norm of X is supposed to be uniformly differentiable, and the norm of X^* is Fréchet differentiable.

Corollary 3.4. *Let C be a nonempty closed subset of a Banach space X and let $w \in X \backslash C$. Suppose $w^* \in \partial d_C(w) \cap GWP \cap S_{X^*}$ for some $w \in X \backslash C$*

and there exists some $z \in S_*(w^*)$ *at which the norm of* X *is Fréchet differentiable. Then* w *has a best approximation in* C.

When one takes an element in the Fréchet subdifferential instead of the contingent subdifferential of f_r one can drop the assumptions that $\|w^*\| = r^{-1}h'(\rho)$ and that the norm of X is Fréchet differentiable at some $z \in S_*(w^*)$.

THEOREM 3.5. *Suppose* f *is l.s.c. and satisfies the growth condition* (G). *Let* $r \in (0, a^{-1})$. *If for some* $w \in X$ *the set* $\partial^- f_r(w) \cap GWP$ *is nonempty then* $P_r^f(w)$ *is nonempty. In particular, when* X *is M-reflexive and* h' *is positive on* $(0, \infty)$, *the set* $P_r^f(w)$ *is nonempty whenever* f_r *is Fréchet subdifferentiable at* w.

Of course, the conclusion is valid when X is reflexive and satisfies the Kadec–Klee property and f is weakly l.s.c., but here we do not make this stringent assumption.

Proof. Again we may suppose the radius of essential minimization ρ of $f_{r,w}$ is positive. Let (y_n) be an essential minimizing sequence. The proof of Lemma 1.6 shows that the sequence $(u_n) := (\|y_n\|^{-1} y_n)$ is a maximizing sequence of w^* in B_X. From the fact that w^* is in GWP we get that (u_n) has a converging subsequence. As $(\|y_n\|) \to \rho$, the sequence (y_n) has a converging subsequence and if y is the limit of such a sequence, y is a minimizer of $f_{r,w}$ as $f_{r,w}$ is l.s.c. Then $x := w - y \in P_r^f(w)$. Moreover we have $\|y\| = \rho$, an observation which we will use later.

When X is M-reflexive we have $GWP \cup \{0\} = X^*$. Thus it remains to observe that if $w^* = 0$ we must have $\rho = 0$ by inequality (1) above and the positivity of h' on $\mathbb{P}$. Then $w \in P_r^f(w)$ by the lower semicontinuity of f and k. ∎

Taking for f the indicator function of a closed subset we get an extension of results of [21, 27] in which differentiability was assumed and a result in [11] in which X is supposed to be reflexive with the Kadec–Klee property. It has been pointed out by an anonymous referee that when the dual norm is Fréchet differentiable off 0 then X must be Kadec by [12, Theorem 6.6].

COROLLARY 3.6. *Let* C *be a nonempty closed subset of* X. *Suppose that for some* $w \in X \backslash C$ *and some* $w^* \in \partial^- d_C(w)$ *one has* $w^* \in GWP \cup \{0\}$ *(in particular suppose the dual norm is Fréchet differentiable at* w^**). Then* w *has a best approximation in* C.

COROLLARY 3.7. *Suppose* X *is M-reflexive, the norm of* X *is Fréchet differentiable off* 0, *and* h' *is positive on* $(0, \infty)$. *Then, for each element* w *of a dense subset* D *of* X, *the set* $P_r^f(w)$ *is nonempty.*

Proof. It is a consequence of Theorem 3.5 and of the Theorem of Preiss [33]: the function f_r being locally Lipschitzian is densely Fréchet subdifferentiable. ∎

Finally let us deal with the stabilized (or limiting) subdifferential

$$\bar{\partial}^- f_r(w) := \limsup_{v \to w} \partial^- f_r(v)$$

of the function f_r, lim sup denoting the limit superior for the strong topology.

PROPOSITION 3.8. *Suppose f is l.s.c. and satisfies the growth condition* (G), *X is M-reflexive, and h' is positive on* $(0, \infty)$. *Let $r \in (0, a^{-1})$. Suppose the radius $\rho(f_{r,v})$ of essential minimization of $f_{r,v}$ is continuous at $w \in X$ as a function of v. Then for each non-zero $w^* \in \bar{\partial}^- f_r(w)$ there exists some $x \in P_r^f(w)$ such that*

$$w^* \in r^{-1} h'(\rho(f_{r,w}))\, S(w - x).$$

In particular the set $P_r^f(w)$ is nonempty.

We observe that the continuity assumption on the radius of essential minimization is satisfied in the case of an indicator function since the distance function is Lipschitzian.

Proof. Let $w^* \in \bar{\partial}^- f_r(w) \backslash \{0\}$: there exists sequences $(w_n) \to w$, $(w_n^*) \to w^*$ such that $w_n^* \neq 0$, $w_n^* \in \partial^- f_r(w_n)$ for each n, hence $w_n^* \in \partial^- f_r(w_n) \cap GWP$. The proof of Theorem 3.5 shows that there exists $x_n \in P_r^f(w_n)$ such that $\|y_n\| = \rho_n$ for $y_n := w_n - x_n$, $\rho_n := \rho(f_{r,w_n})$ and $r w_n^* \in h'(\rho_n)\, S(y_n)$. Since $(\rho_n) \to \rho := \rho(f_{r,w})$ and since h' is continuous, we see that $(y_n^*) := (r h'(\rho_n)^{-1} w_n^*) \to y^* := r h'(\rho)^{-1} w^*$. Then, setting $u_n := \|y_n\|^{-1} y_n$ we observe that (u_n) is a maximizing sequence of y^* on the unit ball as

$$\langle y^*, u_n \rangle - \|y^*\| = \langle y^* - y_n^*, u_n \rangle + \|y_n^*\| - \|y^*\| \to 0.$$

Thus (a subsequence of) (u_n) converges and (y_n) converges too. Then $(x_n) = (w_n - y_n)$ converges to some x with $x \in P_r^f(w)$, $\|w - x\| = \rho$. Passing to the limit in the inclusion of Proposition 2.2 we get the announced formula. ∎

ACKNOWLEDGMENTS

The author is grateful to the editor and to an anonymous referee who pointed out a number of oversights, misuses, and misprints in the manuscript.

REFERENCES

1. H. Attouch, "Variational Convergence for Functions and Operators," Pitman, Boston, 1984.
2. H. Attouch and D. Azé, Approximation and regularization of arbitrary functions in Hilbert spaces by the Lasry–Lions method, *Ann. Inst. H. Poincaré Anal. Nonlinéaire* **10** (1993), 289–312.
3. J.-P. Aubin and H. Frankowska, "Set-Valued Analysis," Birkhäuser, Boston, 1990.
4. E. Asplund, Cebyshev sets in Hilbert spaces, *Trans. Amer. Math. Soc.* **144** (1969), 235–240.
5. E. Asplund, Differentiability of the metric projection in finite-dimensional Euclidean space, *Proc. Amer. Math. Soc.* **38** (1973), 218–219.
6. D. Azé and J.-P. Penot, Uniformly convex and uniformly smooth convex functions, *Ann. Fac. Sci. Toulouse Math.* (6) **4**, No. 4 (1995), 705–730.
7. B. Beauzamy, "Introduction to Banach Spaces and Their Geometry," Mathematic Studies, Vol. 68, North-Holland, Amsterdam, 1982.
8. J. Blatter, Weiteste Punkte und nächste Punkte, *Rev. Roumaine Math. Pures Appl.* **14** (1969), 615–622.
9. J. M. Borwein, S. P. Fitzpatrick, and J. R. Giles, The differentiability of real functions on normed linear space using generalized subgradients, *J. Math. Anal. Appl.* **128** (1987), 512–534.
10. J. M. Borwein and S. P. Fitzpatrick, Existence of nearest points in Banach spaces, *Canad. J. Math.* **61** (1989), 702–720.
11. J. M. Borwein and J. R. Giles, The proximal normal formula in Banach space, *Trans. Amer. Math. Soc.* **302** No. 1 (1987), 371–381.
12. J. M. Borwein and M. Fabian, A note on regularity of sets and of distance functions in Banach space, *J. Math. Anal. Appl.* **182** (1994), 566–570.
13. M. Bougeard, J.-P. Penot, and A. Pommellet, Towards minimal assumptions for the infimal convolution regularization, *J. Math. Anal. Appl.* **64** (1991), 245–270.
14. J. V. Burke, M. C. Ferris, and M. Qian, On the Clarke subdifferential of the distance function of a closed set, *J. Math. Anal. Appl.* **166** (1992), 199–213.
15. F. H. Clarke, "Optimization and Nonsmooth Analysis," Wiley, New York, 1983.
16. F. H. Clarke, R. J. Stern, and P. R. Wolenski, Proximal smoothness and the lower-C^2 property, *J. Convex Anal.* **2** (1995), 117–144.
17. R. Correa, A. Jofre, and L. Thibault, Characterization of lower semicontinuous convex functions, *Proc. Amer. Soc.* **116** (1992), 67–72.
18. K. Deimling, "Nonlinear Functional Analysis," Springer-Verlag, Berlin, 1985.
19. R. Deville, G. Godefroy, and V. Zizler, "Smoothness and Renormings in Banach Spaces," Longman, Harlow, New York, 1993.
20. J. Diestel, "Geometry of Banach Spaces: Selected Topics," Lecture Notes in Mathematics, Vol. 485, Springer-Verlag, Berlin, 1975.
21. S. Fitzpatrick, Metric projection and the differentiability of the distance function, *Bull. Austral. Math. Soc.* **22**, No. 270 (1980), 773–802.
22. P.-J. Laurent, "Approximation et Optimisation," Hermann, Paris, 1972.
23. B. Lemaire, The proximal algorithm, *in* "New Methods in Optimization and Their Industrial Uses" (J.-P. Penot, Ed.), Internat. Ser. Numer. Math., Vol. 87, pp. 73–87, Birkhäuser, Basel, 1989.
24. C. Lescarret, Application "prox" dans un espace de Banach, *C.R. Acad. Sci. Paris Sér. A* **265** (1967), 676–678.
25. P. K. Lin, Strongly unique best approximation in uniformly convex Banach spaces, *J. Approx. Theory* **56** (1989), 101–107.

26. J.-P. Moreau, Proximité et dualité dans un espace hilbertien, *Bull. Soc. Math. France* **93** (1965), 273–299.
27. J.-P. Penot, A characterization of tangential regularity, *Nonlinear Anal.* **25**, No. 6 (1981), 625–643.
28. J.-P. Penot, On regularity conditions in mathematical programming, *Math. Programming* **19** (1982), 167–199.
29. J.-P. Penot, Miscellaneous incidences of convergence theories in optimization and nonlinear analysis. I. Behavior of solutions, *Set-Valued Anal.* **2** (1994), 259–274.
30. J.-P. Penot, Conditioning convex and nonconvex problems, *J. Optim. Theory Appl.* **90** No. 3 (1996), 539–558.
31. R. R. Phelps, Metric projections and the gradient projection method in Banach spaces, *SIAM J. Control Optim.* **23**, No. 6 (1985), 973–977.
32. R. R. Phelps, "Convex Functions, Monotone Operators and Differentiability," Lecture Notes in Mathematics, Vol. 1364, Springer-Verlag, New York, Berlin, 1988.
33. D. Preiss, Differentiability of Lipschitz functions on Banach spaces, *J. Funct. Anal.* **91** (1990), 312–345.
34. R. Ratsimahalo, "Étude de la projection métrique dans les espaces de Banach," Thesis Univ. of Pau, 1992–1996.
35. C. Stegall, Optimization of functions on certain subsets of Banach spaces, *Math. Ann.* **236** (1978), 171–176.
36. L. P. Vlasov, On Cebyshev sets, *Soviet Math.* **8** (1967), 401–404.