MEAN VALUE PROPERTY AND SUBDIFFERENTIAL CRITERIA FOR LOWER SEMICONTINUOUS FUNCTIONS

DIDIER AUSSEL, JEAN-NOËL CORVELLEC, AND MARC LASSONDE

ABSTRACT. We define an abstract notion of subdifferential operator and an associated notion of smoothness of a norm covering all the standard situations. In particular, a norm is smooth for the Gâteaux (Fréchet, Hadamard, Lipschitzsmooth) subdifferential if it is Gâteaux (Fréchet, Hadamard, Lipschitz) smooth in the classical sense, while on the other hand any norm is smooth for the Clarke-Rockafellar subdifferential. We then show that lower semicontinuous functions on a Banach space satisfy an Approximate Mean Value Inequality with respect to any subdifferential for which the norm is smooth, thus providing a new insight on the connection between the smoothness of norms and the subdifferentiability properties of functions. The proof relies on an adaptation of the "smooth" variational principle of Borwein-Preiss. Along the same vein, we derive subdifferential criteria for coercivity, Lipschitz behavior, conemonotonicity, quasiconvexity, and convexity of lower semicontinuous functions which clarify, unify and extend many existing results for specific subdifferentials.

1. Introduction

Let X be a Banach space. Since the pioneering observation of Asplund [1] that there is a close connection between Gâteaux-differentiability of the norm of X and Gâteaux-differentiability of convex continuous functions on dense subsets of X, considerable efforts have been made to refine or extend this connection. See, e.g., Phelps [25] and the references therein for a comprehensive account on this topic. One major contribution is due to Borwein and Preiss [3] who proved that, more generally, the existence of an equivalent #-differentiable norm on X implies that lower semicontinuous functions are #-subdifferentiable on dense subsets of X, where the # symbol can be replaced by any common notion of differentiability (e.g., Gâteaux, Hadamard, Fréchet, Hölder, ...).

The purpose of this paper is to show that the connection between the differentiability of a renorm of X and the subdifferentiability properties of lower semicontinuous functions on X can be given both a more systematic and a sharper form. More systematic since we include virtually all types of subdifferentiability in the same framework, and sharper since we establish in Theo-

Received by the editors June 10, 1994 and, in revised form, December 7, 1994.

1991 Mathematics Subject Classification. Primary 49J52; Secondary 46G05, 49J50, 58C20.

Key words and phrases. Nonsmooth analysis, renorming, variational principle, subdifferential, mean value theorem, Lipschitz behavior, quasiconvexity, convexity.

rem 4.2 that, roughly speaking, if X admits an equivalent ∂ -smooth norm for some subdifferential ∂ , then lower semicontinuous functions on X actually satisfy an Approximate Mean Value Inequality with respect to ∂ . The case where ∂ is the Clarke-Rockafellar subdifferential yields a slight improvement (with a far simpler proof) of Zagrodny's result [34]. The cases where ∂ is a subdifferential of the class # (see above) all appear to be new. Our approach follows the lines of Borwein and Preiss [3]: it is based on an adaptation of their "smooth" variational principle.

In the second part of the paper, we further refine the connection. Our motivation here comes from recent development in nonsmooth analysis where the challenge is to characterize the behavior of lower semicontinuous functions in terms on their subdifferential properties. See, e.g., Treiman [33], Poliquin [26], Correa et al. [9, 10], Luc [20, 21], Clarke et al. [8], Thibault-Zagrodny [32], Aussel et al. [2]. There is no clear link between all these results. Except in [10, 32], they are usually stated for one specific subdifferential, either the Clarke-Rockafellar subdifferential [33, 26, 9, 20, 21, 2] or the proximal subdifferential [26, 8]. They also usually deal with only one or two specific properties, like Lipschitz behavior [33, 8, 32], convexity [26, 9, 10, 21, 8, 2], quasi-convexity [20, 2] or cone-monotonicity [8]. Moreover they are sometimes restricted to some specific spaces, like reflexive Banach spaces [9], Hilbert spaces [8] or finite dimensional spaces [26]. It is plain that our approach enables to establish at once a variety of such "subdifferential criteria", with a common formulation which clarifies the role of the norm of the underlying space: roughly speaking, for a given subdifferential ∂ , the criterion holds in any space with an equivalent ∂ -smooth norm. All the results mentioned above follow as special cases.

We point out that Borwein-Preiss' result has been (partially) improved by Deville $et\ al.$ [14] who showed that, in the case where # denotes the Gâteaux, Hadamard or Fréchet notion of differentiability, the existence of a Lipschitz #-differentiable bump function on a Banach space X suffices to assure the dense #-subdifferentiability of lower semicontinuous functions on X. However, we do not develop this connection here since our main objective is to cover at once a large amount of cases in a systematic and simple way, for which a key role is played by the definition of a ∂ -smooth norm.

The paper is organized as follows. In §2 we introduce a class of subdifferentials containing (among others) the *-subdifferentials considered by Borwein and Preiss and the Clarke-Rockafellar subdifferential. Its definition is axiomatic: the class is determined by a list of three elementary properties. Next, for any subdifferential ∂ of this class we define the notion of a ∂-smooth norm, extending the classical notion. In the rest of the paper, we study the ∂-subdifferential properties of lower semicontinuous functions defined on Banach spaces with an equivalent ∂-smooth norm. In §3 we show the existence of subgradients on dense subsets of the effective domain of the functions. In §4 we establish the aforementioned Approximate Mean Value Inequality. In §5 we derive subdifferential criteria for coercivity, Lipschitz behavior, cone-monotonicity, quasiconvexity, and convexity of lower semicontinuous functions. In §6 we observe that all the previous results hold without any assumption on the smoothness of a renorm of the space, provided that the functions are either Gâteaux differentiable or convex lower semicontinuous.

2. Notations and definitions

Throughout this paper, X stands for a real Banach space, X^* for its topological dual, and $\langle .,. \rangle$ for the duality pairing. For x, a, $b \in X$ and $\|.\|$ a norm on X, we let

$$d_a(x) := \|x - a\|, \quad d_{[a,b]}(x) := \min_{c \in [a,b]} \|x - c\|,$$

where $[a, b] := \{c \in X \mid c = \lambda a + (1 - \lambda)b \text{ for some } \lambda \in [0, 1]\}$ is the closed segment joining a and b. For $\lambda > 0$ and $x \in X$, we let $B_{\lambda}(x) = \{y \in X \mid ||y - x|| < \lambda\}$.

All the functions $f: X \to \mathbb{R} \cup \{+\infty\}$ considered in this paper are lower semicontinuous. As usual, we let dom $f = \{x \in X \mid f(x) < \infty\}$ and, for $x \in \text{dom } f$, we write $x' \to_f x$ to express that $x' \to x$ and $f(x') \to f(x)$.

For a set-valued operator $A: X \to X^*$, we let dom $A = \{x \in X \mid A(x) \neq \emptyset\}$.

Definition 2.1. We call *subdifferential*, denoted by ∂ , any operator which associates a subset $\partial f(x)$ of X^* to any lower semicontinuous $f: X \to \mathbb{R} \cup \{+\infty\}$ and any $x \in X$, and satisfies the following properties:

- (P1) $\partial f(x) = \{x^* \in X^* \mid \langle x^*, y x \rangle + f(x) \le f(y), \forall y \in X \}$ whenever f is convex:
- (P2) $0 \in \partial f(x)$ whenever $x \in \text{dom } f$ is a local minimum of f;
- (P3) $\partial(f+g)(x) \subset \partial f(x) + \partial g(x)$ whenever g is a real-valued convex continuous function which is ∂ -differentiable at x,

where g is ∂ -differentiable at x means that both $\partial g(x)$ and $\partial (-g)(x)$ are nonempty. We say that f is ∂ -subdifferentiable at x when $\partial f(x)$ is nonempty. The elements of $\partial f(x)$ are called the subgradients of f at x.

Abstract classes of subdifferentials were already considered by Ioffe [17] in another context, and more recently by Correa *et al.* [10] and Thibault-Zagrodny [32]. Our definition is more general mainly because of a much less restrictive additivity property (P3).

Main examples: (1) All the subdifferentials $\partial^{\#}$ considered by Borwein-Preiss [3] satisfy the above properties: this readily follows from their definition. A function f is ∂ -differentiable at x for a subdifferential $\partial = \partial^{\#}$ of this class if and only if $\partial f(x) = -\partial (-f)(x)$ contains a single element, which then coincides with the usual Gâteaux derivative $\nabla^G f(x)$. We recall that this class contains as special cases the following subdifferentials:

Hadamard

$$\partial^H f(x) := \left\{ x^* \in X^* \mid \langle x^*, v \rangle \leq \liminf_{\substack{t \geq 0 \\ v' \rightarrow v}} \frac{f(x + tv') - f(x)}{t} \quad \forall v \in X \right\};$$

Fréchet

$$\partial^{F} f(x) := \{ x^{*} \in X^{*} \mid \liminf_{\|v\| \searrow 0} \frac{f(x+v) - f(x) - \langle x^{*}, v \rangle}{\|v\|} \ge 0 \};$$

Lipschitz-smooth

$$\partial^{LS} f(x) := \{ x^* \in X^* \mid \liminf_{\|v\| \searrow 0} \frac{f(x+v) - f(x) - \langle x^*, v \rangle}{\|v\|^2} > -\infty \}.$$

It is a straightforward verification that $\partial^{LS} f \subset \partial^F f \subset \partial^H f$. Moreover, when f is locally Lipschitz the Hadamard subdifferential coincides with the Gâteaux subdifferential

$$\partial^G f(x) := \left\{ x^* \in X^* \mid \langle x^*, v \rangle \le \liminf_{t \searrow 0} \frac{f(x + tv) - f(x)}{t} \quad \forall v \in X \right\},\,$$

and when X is a Hilbert space, the Lipschitz-smooth subdifferential coincides with Rockafellar's proximal subdifferential. See Borwein-Preiss, loc. cit., Borwein-Strojwas [4], Penot [23, 24], Ioffe [18, 19], Rockafellar [30], Clarke et al. [7] for further properties of these subdifferentials. Let us finally mention that the following operators built from the ∂^* -subdifferentials:

$$\tilde{\partial}^{\#} f(x) := \{ \nabla^{G} \varphi(x) \mid \varphi : X \to \mathbb{R} \text{ is continuous, } \partial^{\#}\text{-differentiable at } x,$$
 and $f - \varphi$ has a (local) minimum at $x \}$,

are also easily seen to be subdifferentials in the sense of Definition 2.1 (this type of operator naturally arises from the study of Hamilton-Jacobi's equations: see Crandall *et al.* [12], Deville *et al.* [14]).

(2) The Clarke-Rockafellar subdifferential given by $\partial^{CR} f(x)$

$$:= \{x^* \in X^* \mid \langle x^*, v \rangle \leq \sup_{\varepsilon > 0} \limsup_{\substack{x' \to f^x \\ t \searrow 0}} \inf_{v' \in B_{\varepsilon}(v)} \frac{f(x' + tv') - f(x')}{t} \quad \forall v \in X\}$$

also satisfies the above properties: see for instance Clarke [6]. A function is ∂^{CR} -differentiable at x if it is locally Lipschitz at x. It is obvious that $\partial^H f \subset \partial^{CR} f$. Moreover, when f is locally Lipschitz the Clarke-Rockafellar subdifferential coincides with the Clarke subdifferential

$$\partial^C f(x) := \{ x^* \in X^* \mid \langle x^*, v \rangle \leq \limsup_{\stackrel{x' \to f^x}{t \searrow 0}} \frac{f(x' + tv) - f(x')}{t} \quad \forall v \in X \}.$$

See Clarke [6] and Rockafellar [28, 29] for further properties.

(3) The subdifferentials considered by Correa et al. [10, Definition 2] and Thibault-Zagrodny [32, §1] evidently satisfy the above properties. See [10, 32] for examples of such subdifferentials. We point out that the subdifferentials $\partial^{\#}$ do not satisfy the conditions in [10, Definition 2] nor in [32, §1]: they are at best presubdifferentials in the sense of [32].

Definition 2.2. A norm $\|.\|$ on X is said to be ∂ -smooth if all the (real-valued, convex, continuous) functions of the following form are ∂ -differentiable:

(i)
$$d_{[a,b]}^2(x) := \min_{c \in [a,b]} ||x-c||^2$$
, where $[a,b]$ is a closed segment in X ,

(ii) $\Delta_2(x) := \sum_n \mu_n \|x - v_n\|^2$, where $\sum_n \mu_n = 1$, $\mu_n \ge 0$, and (v_n) converges in X.

We say that a Banach space admits a ∂ -smooth renorm if it admits an equivalent norm which is ∂ -smooth.

Since the functions $d_{[a,b]}^2$ and Δ_2 play a crucial role in the first part of this paper, it is worth giving an explicit expression for their subdifferential:

Proposition 2.3. (i)
$$\partial d^2_{[a,b]}(x) = \{ 2 d_{[a,b]}(x) \zeta_x \mid \zeta_x \in \partial d_{[a,b]}(x) \};$$

(ii)
$$\partial \Delta_2(x) = \{ 2 \sum_n \mu_n \|x - v_n\| \xi_n \mid \xi_n \in \partial d_{v_n}(x) \}.$$

Proof. We include the proof for completeness. (i) The function $d_{[a,b]}^2$ being the composition of the convex continuous function $d_{[a,b]}$ with the function $t \mapsto t^2$, assertion (i) follows from Property (P1) and the classical chain rule of subdifferential calculus.

(ii) For k = 1, 2, ..., define the functions φ_k , $\varphi^k : X \to \mathbb{R}$ by letting

$$\varphi_k(x) := \sum_{n=1}^k \mu_n \|x - v_n\|^2, \quad \varphi^k(x) := \sum_{n=k+1}^\infty \mu_n \|x - v_n\|^2.$$

Since (v_n) is bounded and $\sum \mu_n$ is absolutely convergent, the sequence of convex continuous functions (φ_k) converges to Δ_2 uniformly on bounded sets, so that the functions Δ_2 and φ^k are real-valued, convex, and continuous.

An easy computation shows that

$$\partial \varphi_k(x) = \{ 2 \sum_{n=1}^k \mu_n \|x - v_n\| \xi_n \mid \xi_n \in \partial d_{v_n}(x) \}.$$

As above, for any $x \in X$ and any fixed sequence (ξ_n) with $\xi_n \in \partial d_{v_n}(x)$, the sequence $(2 \sum_{n=1}^k \mu_n \|x - v_n\|\xi_n)_k$ is seen to converge strongly in X^* , so that the set

$$A(x) := \{ 2 \sum_{n=1}^{\infty} \mu_n \|x - v_n\| \xi_n \mid \xi_n \in \partial d_{v_n}(x) \}$$

is well-defined. The inclusion $A(x)\subset\partial\Delta_2(x)$ is clear: if $x^*\in A(x)$, then for any $v\in X$ we have

$$\langle x^*, v \rangle = \lim_{k \to \infty} \langle 2 \sum_{n=1}^k \mu_n || x - v_n || \xi_n, v \rangle$$

$$\leq \lim_{k \to \infty} (\varphi_k(x+v) - \varphi_k(x))$$

$$= \Delta_2(x+v) - \Delta_2(x).$$

To prove the converse, let $x^* \in \partial \Delta_2(x)$. The formula

$$\varphi^{k}(x) = \mu_{k+1} ||x - v_{k+1}||^{2} + \varphi^{k+1}(x)$$

leads to the subdifferential equalities

$$\partial \varphi^k(x) = 2\mu_{k+1} ||x - v_{k+1}|| \partial d_{v_{k+1}}(x) + \partial \varphi^{k+1}(x), \quad k = 0, 1, \dots$$

It is therefore possible to define inductively a sequence $\xi_k \in \partial d_{v_k}(x)$, $k = 1, 2, \ldots$, such that

$$\delta^{k} := x^{*} - 2 \sum_{n=1}^{k} \mu_{n} ||x - v_{n}|| \xi_{n} \in \partial \varphi^{k}(x), \quad k = 1, 2, \dots.$$

To complete the proof, it remains to observe that $\|\delta^k\|$ converges to 0: indeed, for every $v \in X$ with $\|v\| \le 1$ we have

$$\langle \delta^{k}, v \rangle \leq \sum_{n=k+1}^{\infty} \mu_{n}(\|x+v-v_{n}\|^{2} - \|x-v_{n}\|^{2})$$

 $\leq \sum_{n=k+1}^{\infty} \mu_{n}(2\|x-v_{n}\|+1). \quad \Box$

Main examples: (1) A norm is $\partial^{\#}$ -smooth if and only if it is $^{\#}$ -smooth in the sense of Borwein-Preiss [3]. In particular, a norm is ∂^{F} -smooth (resp. ∂^{G} -smooth, ∂^{LS} -smooth) if and only if it is Fréchet-differentiable (resp. Gâteaux-differentiable, Lipschitz-smooth) off the origin. We recall that separable Banach spaces admit a ∂^{G} -smooth renorm, that reflexive Banach spaces admit a ∂^{F} -smooth renorm, and that the natural norms of Hilbert spaces and L^{p} spaces, for $2 \le p < +\infty$, are ∂^{LS} -smooth.

(2) Since for any norm the functions $d_{[a,b]}^2$ and Δ_2 are locally Lipschitz, any norm is ∂^{CR} -smooth.

3. Existence of subgradients

It is not apparent from Definition 2.1 that the set $\partial f(x)$ may be nonempty except, of course, if f is convex continuous or if x is a local minimum of f on its domain. However it turns out that the set of points where $\partial f(x)$ is nonempty is dense in dom f. This is a consequence of the following basic result which asserts that in any neighborhood of an approximate minimum of f there are points where f does have subgradients whose norm can moreover be estimated.

Theorem 3.1 (Approximate minimum). Let X be a Banach space with a ∂ -smooth norm and let $f: X \to \mathbb{R} \cup \{+\infty\}$ be lower semicontinuous. Let $A \subset X$ be a closed set and let $\varepsilon > 0$ be a given constant. Suppose that $x_0 \in A$ and $\lambda > 0$ satisfy

$$B_{\lambda}(x_0) \subset A$$
 and $f(x_0) < \varepsilon + \inf_{A} f$.

Then there exists \bar{x} in $B_{\lambda}(x_0)$ verifying

$$f(\bar{x}) < \varepsilon + \inf_{A} f$$
 and $\partial f(\bar{x}) \cap 2(\varepsilon/\lambda)B^* \neq \emptyset$,

where B^* is the dual unit ball.

Proof. Let $g := f + \Psi_A$ where Ψ_A is the indicator function of A. Then g is lower semicontinuous and verifies

$$g(x_0) < \varepsilon + \inf_{x} g$$
.

Consider Borwein-Preiss' smooth variational principle [3, Theorem 2.6] applied to g with the given ε and λ , and p := 2. According to the proof, there exists a sequence $v_n \to \bar{x} \in X$ such that:

$$(1) v_0 := x_0 \text{ and } ||v_n - \bar{x}|| < \lambda, \quad \forall n,$$

(2)
$$g(\bar{x}) < \varepsilon + \inf_{X} g,$$

(3)
$$g(x) + (\varepsilon/\lambda^2)\Delta_2(x) \ge g(\bar{x}) + (\varepsilon/\lambda^2)\Delta_2(\bar{x}), \quad \forall x \in X,$$

where $\Delta_2(x) := \sum_n \mu_n ||x - v_n||^2$, with $\sum_n \mu_n = 1$ and $\mu_n \ge 0$.

It follows from (1) that \bar{x} lies in $B_{\lambda}(x_0) \subset A$. Then (2) means that $f(\bar{x}) < \varepsilon + \inf_A f$, while (3) implies that \bar{x} is a local minimum of the function $f + (\varepsilon/\lambda^2)\Delta_2$. Hence, according to Properties (P2)-(P3) and Definition 2.2 of a ∂ -smooth norm, we have

$$0 \in \partial f(\bar{x}) + \partial ((\varepsilon/\lambda^2)\Delta_2)(\bar{x}).$$

Now, we derive from (1) and Proposition 2.3 (ii) that

$$\partial((\varepsilon/\lambda^2)\Delta_2)(\bar{x}) \subset 2(\varepsilon/\lambda)B^*$$
,

which completes the proof.

Remark 3.1.1. When $\partial = \partial^{\#}$, Theorem 3.1 essentially formulates the final part of Borwein-Preiss' smooth variational principle, *loc. cit.* The case $\partial = \partial^{CR}$ is proved in Rockafellar [28, Theorem 7] as an application of Ekeland's variational principle [15].

Corollary 3.2 (Subdifferentiability on dense subset). Let X be a Banach space with a ∂ -smooth renorm. Then any lower semicontinuous function $f: X \to \mathbb{R} \cup \{+\infty\}$ is ∂ -subdifferentiable at a dense subset of points in its graph, i. e., for each $x \in \text{dom } f$ there exists $(x_n) \subset \text{dom } \partial f$ such that $x_n \to_f x$.

Proof. Let $x_0 \in \text{dom } f$ and let $\varepsilon > 0$. Since f is lower semicontinuous, there exists a closed neighborhood A of x_0 such that $f(x_0) < \varepsilon + \inf_A f$. Hence, by Theorem 3.1, there exists $\bar{x} \in \text{dom } \partial f$ such that $\bar{x} \in A$ and $f(x_0) - \varepsilon \leq f(\bar{x}) \leq f(x_0) + \varepsilon$. \square

Remark 3.2.1. The case $\partial = \partial^{\#}$ is Theorem 3.1 in Borwein-Preiss, *loc. cit.* The case $\partial = \partial^{CR}$ is due to McLinden [22].

Corollary 3.3 (Differentiability on dense subset). Let X be a Banach space with a ∂ -smooth renorm. Then any convex continuous function $f: X \to \mathbb{R}$ is ∂ -differentiable at a dense subset of points in X.

Proof. It follows from (P1) and elementary convex analysis that $\partial f(x) \neq \emptyset$ for every x in X, and from Corollary 3.2 that $\partial (-f)(x) \neq \emptyset$ for every x in a dense subset of X. \square

Remark 3.3.1. The case $\partial = \partial^{\#}$ is Theorem 4.1(a) in Borwein-Preiss, *loc. cit.* It covers many earlier results, like the one by Asplund [1] mentioned in the introduction.

As we already pointed out in the introduction, in the case where ∂ denotes the Gâteaux, Hadamard or Fréchet subdifferential, the above corollaries remain valid in larger classes of spaces, namely the spaces admitting a Lipschitz ∂ -differentiable bump function (instead of a ∂ -smooth renorm): see Deville *et al.* [14].

4. THE APPROXIMATE MEAN VALUE INEQUALITY

Throughout this section, X denotes a Banach space with a ∂ -smooth renorm and $f:X\to\mathbb{R}\cup\{+\infty\}$ denotes a lower semicontinuous function. The Approximate Mean Value Inequality is a straightforward consequence of the following theorem which is a natural complement to Property (P2) and to Theorem 3.1. Indeed, its conclusion reads thus: in any neighborhood of a local minimum of f in a given direction v there are points where f has subgradients whose value in the direction v can be arbitrary small.

Theorem 4.1 (Directional minimum). Let $a, b \in X$ be such that $[a, b] \cap \text{dom } f \neq \emptyset$, and let c be a minimum of f on [a, b]. Then there exist $x_n \to_f c$ and $x_n^* \in \partial f(x_n)$ such that

$$\liminf_{n\to+\infty} \langle x_n^*, x-x_n\rangle \ge 0 \quad for \ every \ x\in [a,b].$$

Moreover, if $c \neq b$, then

$$\liminf_{n\to+\infty} \langle x_n^*, b-a\rangle \ge 0.$$

Consequently, if c is distinct from both end-points a and b, we have

$$\lim_{n\to+\infty}\langle x_n^*,\,b-a\rangle=0.$$

Proof. We essentially follow the lines of the first part of the proof of Corollary 3.2 in Borwein-Preiss [3]. Let $\|.\|$ be a ∂ -smooth norm on X and A be a closed neighborhood of [a, b] on which f is bounded below. For any positive integer n so large that the ball $B_{1/n}(c)$ is contained in A, let $r_n > 0$ be such that

$$f(c)-1/n^2<\inf_{B_{r_n}}f,$$

where $B_{r_n} := \{x \in X \mid d_{[a,b]}(x) < r_n\}$, and let then $K_n > 0$ be such that

$$f(c) - 1/n^2 < \inf_{A} f + K_n r_n^2$$
.

Consider now the function $f_n := f + K_n d_{[a,b]}^2$. It follows from the above inequalities that

$$f_n(c) = f(c) < \inf_A f_n + 1/n^2.$$

Apply Theorem 3.1 to f_n with the given A, $\varepsilon := 1/n^2$, $x_0 := c$ and $\lambda := 1/n$. This produces a sequence $(x_n) \subset A$ verifying

$$||c-x_n||<1/n\;,$$

$$f(x_n) \le f_n(x_n) < f(c) + 1/n^2$$
, and

$$\partial f_n(x_n) \cap (2/n)B^* \neq \emptyset$$
.

Since f is lower semicontinuous, the first two formulae show that $x_n \to_f c$. Since the norm is ∂ -smooth, the third formula and Property (P3) show that there exist $x_n^* \in \partial f(x_n)$, $\delta_n^* \in \partial d_{[a,b]}^2(x_n)$, and $\beta_n^* \in B^*$ with $x_n^* + K_n \delta_n^* = (2/n)\beta_n^*$.

From Proposition 2.3(i) we get $\delta_n^* = 2d_{[a,b]}(x_n)\zeta_n$ with $\zeta_n \in \partial d_{[a,b]}(x_n)$, hence for any $x \in [a,b]$ we have

(1)
$$\langle \delta_n^*, x - x_n \rangle = 2d_{[a,b]}(x_n) \langle \zeta_n, x - x_n \rangle \leq -2d_{[a,b]}^2(x_n) \leq 0,$$

and for any point Px_n in [a, b] such that $||x_n - Px_n|| = d_{[a,b]}(x_n)$ it holds

(2)
$$\langle \delta_n^*, b - Px_n \rangle = \langle \delta_n^*, b - x_n \rangle + \langle \delta_n^*, x_n - Px_n \rangle$$

$$\leq -2d_{[a,b]}^2(x_n) + 2\|x_n - Px_n\|^2 = 0.$$

We derive from (1) that $\langle x_n^*, x - x_n \rangle \ge (2/n) \langle \beta_n^*, x - x_n \rangle$, which proves the first assertion. If $c \ne b$, we may assume that $Px_n \ne b$ (because $Px_n \to c$), so that (2) implies:

$$\langle \delta_n^*, b-a \rangle = \frac{\|b-a\|}{\|b-Px_n\|} \langle \delta_n^*, b-Px_n \rangle \leq 0,$$

which yields $\langle x_n^*, b-a \rangle \ge (2/n) \langle \beta_n^*, b-a \rangle$ and proves the second assertion.

Symmetrically, if $c \neq a$ we have $\liminf_{n \to +\infty} \langle x_n^*, a - b \rangle \geq 0$, and the third assertion follows. \square

We are now in position to state our main result:

Theorem 4.2 (Approximate mean value inequality). Let $a, b \in X$ with $a \in \text{dom } f$ and $a \neq b$, and let $r \in \mathbb{R}$ be such that $r \leq f(b)$. Then there exist $c \in [a, b]$, $c \neq b$, and sequences $x_n \to_f c$ and $x_n^* \in \partial f(x_n)$ such that

- (i) $\liminf_{n\to+\infty} \langle x_n^*, c-x_n \rangle \ge 0$;
- (ii) $\liminf_{n\to+\infty} \langle x_n^*, b-a \rangle \ge r f(a)$.

Proof. Let $\alpha \in X^*$ be such that $\langle \alpha, a - b \rangle = r - f(a)$. Consider the function $g := f + \alpha$. Since g is lower semicontinuous and $g(a) = r + \langle \alpha, b \rangle \leq g(b)$, there exists $c \in [a, b]$, $c \neq b$, such that $g(c) = \min_{[a, b]} g$, so we may apply Theorem 4.1 to find sequences $x_n \to_g c$ and $y_n^* \in \partial g(x_n)$ verifying

$$\liminf_{n \to +\infty} \langle y_n^*, c - x_n \rangle \ge 0 \quad \text{and} \quad \liminf_{n \to +\infty} \langle y_n^*, b - a \rangle \ge 0.$$

According to Property (P3), for any n there exists $x_n^* \in \partial f(x_n)$ such that $y_n^* = x_n^* + \alpha$ and substitution in the above inequalities yields

- (i) $\liminf_{n \to +\infty} \langle x_n^*, c x_n \rangle \ge \lim_{n \to +\infty} \langle \alpha, x_n c \rangle = 0$, and
- (ii) $\liminf_{n \to +\infty} \langle x_n^*, b a \rangle \ge \langle \alpha, a b \rangle = r f(a).$

Finally, $x_n \to_f c$ since $x_n \to_g c$. \square

The following immediate corollary of Theorem 4.2 will be used several times in the sequel.

Corollary 4.3. Let $a, b \in X$ with f(a) < f(b). Then there exist $c \in [a, b]$, $c \neq b$, and sequences $x_n \to_f c$ and $x_n^* \in \partial f(x_n)$ with

$$\langle x_n^*, x - x_n \rangle > 0$$
 for every $x = c + t(b - a)$ with $t > 0$.

Proof. According to Theorem 4.2, we can find $c \in [a, b]$, $c \neq b$ and sequences $x_n \to_f c$ and $x_n^* \in \partial f(x_n)$ verifying

$$\liminf_{n \to +\infty} \langle x_n^*, c - x_n \rangle \ge 0 \quad \text{ and } \liminf_{n \to +\infty} \langle x_n^*, b - a \rangle > 0.$$

Letting x = c + t(b - a) with t > 0 it holds

$$\langle x_n^*, x - x_n \rangle = \langle x_n^*, c - x_n \rangle + t \langle x_n^*, b - a \rangle > 0$$

for n sufficiently large. \square

Remark 4.3.1. (a) Similarly as in Corollary 4.3, it follows immediately from (i) and (ii) of Theorem 4.2 that

$$\liminf_{n \to +\infty} \langle x_n^*, x - x_n \rangle \ge \frac{\|x - c\|}{\|b - a\|} (r - f(a))$$

for every x = c + t(b - a) with $t \ge 0$. Furthermore, whenever the point c is different from a, it holds

$$\lim_{n\to+\infty} \langle x_n^*, b-a\rangle = r - f(a).$$

(See Theorem 4.1.) Thus we recover in a far simpler way Zagrodny's approximate mean value theorem [34, Theorem 4.3], where $\partial = \partial^{CR}$, with the slight improvements that b need not be in dom f and that x_n is shown to graphically converge to c.

- (b) The Gâteaux, Hadamard, Fréchet, Lipschitz-smooth (or, more generally, ") cases in Theorem 4.2 appear to be new, although closely related results have been given by Penot [24] for the Hadamard subderivative and by Clarke et al. [7] for the Lipschitz-smooth subdifferential. It is mentioned in Thibault-Zagrodny [32] that some of these cases could also be derived from Zagrodny's theorem by using the presubdifferential approach and "fuzzy calculus" on the subdifferentials.
- (c) The case $\partial = \partial^{CR}$ in Corollary 4.3 slightly refines results in Zagrodny [34], Luc [21] and Aussel *et al.* [2]. For convex functions (hence with the Fenchel subdifferential), a variant of Corollary 4.3 is established in Simons [31].

5. SUBDIFFERENTIAL CRITERIA

In this section again, X is a Banach space with a ∂ -smooth renorm and $f: X \to \mathbb{R} \cup \{+\infty\}$ is a lower semicontinuous function. As our main applications of the results of the preceding sections, we shall derive various subdifferential criteria which have been the object of a number of recent publications, dealing with specific subdifferentials. Here, these criteria are seen to be easy consequences of Theorem 3.1 or Theorem 4.2 (compare with the references given in the remarks).

Theorem 5.1 (Coercivity). Let f be bounded below. If f satisfies the following Palais-Smale type condition:

if (x_n) is a sequence in X such that $\gamma(x_n) := \inf\{\|x_n^*\| \mid x_n^* \in \partial f(x_n)\}$ converges to 0 and $(f(x_n))$ is bounded, then (x_n) has a bounded subsequence,

then f is coercive, that is, $f(x) \to +\infty$ when $||x|| \to +\infty$.

Proof. Set $X_0 := X$ and $X_n := X \setminus B_n(0)$, $n \in \mathbb{N}$. Of course, f is coercive if and only if the nondecreasing sequence $(\inf_{X_n} f)$ goes to infinity. Assume, on the contrary, that the sequence $(\inf_{X_n} f)$ is convergent. For each $n \in \mathbb{N}$, let $x_0^n \in X_{n+1}$ be such that

$$f(x_0^n) < \inf_{X_{n+1}} f + 1/n = \inf_{X_n} f + \alpha_n + 1/n$$
,

where $\alpha_n := \inf_{X_{n+1}} f - \inf_{X_n} f \ge 0$. Apply Theorem 3.1 to f with $A := X_n$, $\varepsilon := \alpha_n + 1/n$, $x_0 := x_0^n$, and $\lambda := 1$, to obtain a sequence $(x_n) \subset X$ verifying

$$x_n \in X_n$$
,

$$f(x_n) < \inf_{X_{n+1}} f + 1/n, \text{ and }$$

$$\gamma(x_n) \leq 2(\alpha_n + 1/n) .$$

Observe that (x_n) has no bounded subsequence while $(f(x_n))$ is bounded and $\gamma(x_n)$ converges to 0, so the Palais-Smale condition is not satisfied. \square

Remark 5.1.1. (a) The converse is obviously true.

(b) In critical point theory, the "usual" Palais-Smale condition reads:

if (x_n) is a sequence in X such that $\gamma(x_n) := \inf\{\|x_n^*\| \mid x_n^* \in \partial f(x_n)\}$ converges to 0 and $(f(x_n))$ is bounded, then (x_n) possesses a (strongly) convergent subsequence.

The weaker condition stated in Theorem 5.1 often appears as an intermediate step when verifying the above one.

(c) In Section 6, we give a "smooth" version of Theorem 5.1.

Theorem 5.2 (Lipschitz behavior). Let $U \subset X$ be an open convex set with $U \cap \text{dom } f \neq \emptyset$ and let C > 0. If

$$\sup\{\|x^*\| \mid x^* \in \partial f(x)\} \le C \quad \text{for all } x \in U,$$

then f is Lipschitz of rank C on U.

Proof. Let $a, b \in U$ with a in dom f and $a \neq b$, let $r \in \mathbb{R}$ such that $r \leq f(b)$, and let $\varepsilon > 0$. It follows from Theorem 4.2 (ii) that there exist $x \in U$ and $x^* \in \partial f(x)$ such that

$$r - f(a) \le \langle x^*, b - a \rangle + \varepsilon \le C ||b - a|| + \varepsilon.$$

Since $r \le f(b)$ and $\varepsilon > 0$ are arbitrary, we derive that $f(b) - f(a) \le C \|b - a\|$. Hence b belongs to dom f and we can exchange the role of a and b to conclude that f is Lipschitz of rank C on U. \square

Remark 5.2.1. (a) The converse is true whenever $\partial f \subset \partial^{CR} f$.

(b) The case $\partial = \partial^{CR}$ is due to Treiman [33, Theorem 5], the case $\partial = \partial^{LS}$ and X a Hilbert space is given in Clarke *et al.* [8, Theorem 3.6].

Theorem 5.3 (Cone-monotonicity). Let K be a cone in X. If

$$\partial f(x) \subset \{x^* \in X^* \mid \langle x^*, y \rangle \le 0 \quad \forall y \in K \} \quad \text{for all } x \in X,$$

then f is K-nonincreasing, that is, $y \in x + K$ implies $f(y) \le f(x)$.

Proof. Let $x, y \in X$ such that f(x) < f(y). It follows from Theorem 4.2 (ii) that there exist $z \in \text{dom}\,\partial f$ and $z^* \in \partial f(z)$ with $\langle z^*, y - x \rangle > 0$, hence y - x does not belong to K. \square

Remark 5.3.1. (a) The converse is true whenever $\partial f \subset \partial^{CR} f$.

(b) The case where X is a Hilbert space and $\partial = \partial^{LS}$ is given in Clarke *et al.* [8, Theorem 3.2].

The next two results are stated in Aussel *et al.* [2] for $\partial = \partial^{CR}$, but the proofs work as well in the present setting: we give them for the sake of completeness. We recall that a function $f: X \to \mathbb{R} \cup \{+\infty\}$ is said to be *quasiconvex* if

$$x, y \in \text{dom } f, z \in [x, y] \implies f(z) \le \max\{f(x), f(y)\},$$

and that a multivalued operator $A: X \to X^*$ is said to be quasimonotone if

$$x^* \in A(x), y^* \in A(y), \text{ and } \langle x^*, y - x \rangle > 0 \implies \langle y^*, y - x \rangle \geq 0.$$

Theorem 5.4 (Quasiconvexity). If ∂f is quasimonotone, then f is quasiconvex. Proof. We show that if f is not quasiconvex then ∂f is not quasimonotone. Assume that there are some x, y, z in X with $z \in [x, y]$ and $f(z) > \max\{f(x), f(y)\}$. According to Corollary 4.3 applied with a := x and b := z, there exist sequences $y_n \in \text{dom } \partial f$ and $y_n^* \in \partial f(y_n)$ such that

$$y_n \to \hat{y} \in [x, z], \ \hat{y} \neq z \text{ and } \langle y_n^*, y - y_n \rangle > 0.$$

Let $0 < \lambda \le 1$ be such that $z = \hat{y} + \lambda(y - \hat{y})$ and set $z_n := y_n + \lambda(y - y_n)$, so that $z_n \to z$. Since f is lower semicontinuous, we may pick $n \in \mathbb{N}$ large enough with $f(z_n) > f(y)$. Apply Corollary 4.3 again with a := y and $b := z_n$ to find sequences $x_k \in \text{dom } \partial f$, $x_k^* \in \partial f(x_k)$ such that

$$x_k \to \hat{x} \in [y, z_n], \ \hat{x} \neq z_n \text{ and } \langle x_k^*, y_n - x_k \rangle > 0.$$

In particular, $\hat{x} \neq y_n$ and

$$\langle y_n^*, \hat{x} - y_n \rangle = \frac{\|\hat{x} - y_n\|}{\|y - y_n\|} \langle y_n^*, y - y_n \rangle > 0 ;$$

hence, $\langle y_n^*, x_k - y_n \rangle > 0$ for k sufficiently large. But $\langle x_k^*, y_n - x_k \rangle > 0$, showing that ∂f is not quasimonotone. \square

Remark 5.4.1. (a) The converse is true whenever $\partial f \subset \partial^{CR} f$, see Aussel et al. [2].

(b) The case where $\partial = \partial^{CR}$ can also be found in Luc [20].

Theorem 5.5 (Convexity). If ∂f is monotone, then f is convex.

Proof. If ∂f is monotone then for each $\alpha \in X^*$ the operator $x \mapsto \partial f(x) + \alpha = \partial (f + \alpha)(x)$ is monotone, hence quasimonotone. According to Theorem 5.4, for each $\alpha \in X^*$ the function $f + \alpha$ is then quasiconvex, which is easily shown to be equivalent to the convexity of f. \square

Remark 5.5.1. (a) The converse is true because of Property (P1).

(b) The case $\partial = \partial^{CR}$ is established in Correa *et al.* [10] and Luc [21], generalizing earlier results by Poliquin [26] (in finite dimensional spaces) and Correa *et al.* [9] (in reflexive Banach spaces). The case $\partial = \partial^{LS}$ is proved in Clarke *et al.* [8] for Hilbert spaces and in Poliquin [26] for finite dimensional spaces. The case $\partial = \partial^F$ in Hilbert spaces goes back to Degiovanni *et al.* [13].

Theorem 5.6 (Maximal monotonicity). If $dom f \neq \emptyset$ and ∂f is monotone, then ∂f is actually maximal monotone.

Proof. Let $x \in X$ and $x^* \in X^*$ be such that $x^* \notin \partial f(x)$; we show that there exist $y \in \operatorname{dom} \partial f$ and $y^* \in \partial f(y)$ with $\langle y^* - x^*, y - x \rangle < 0$, proving the result. First we claim that there is some point z in X such that $(f - x^*)(z) < (f - x^*)(x)$: this is evident if $x \notin \operatorname{dom} f$; otherwise, since 0 does not belong to $\partial f(x) - x^* = \partial (f - x^*)(x)$, we derive from Property (P2) that x cannot be a minimum point of the function $f - x^*$. Next, applying Corollary 4.3 to $f - x^*$ with a := z and b := x, we find $y \in \operatorname{dom} \partial (f - x^*)$ and $\bar{y}^* \in \partial (f - x^*)(y) = \partial f(y) - x^*$ such that $\langle \bar{y}^*, x - y \rangle > 0$. Writing $\bar{y}^* = y^* - x^*$ with $y^* \in \partial f(y)$ yields $\langle y^* - x^*, x - y \rangle > 0$.

Remark 5.6.1. The idea of the above proof is due to Simons [31], who proved Rockafellar's theorem [27] (that the Fenchel subdifferential of a proper lower semicontinuous convex function f is maximal monotone) using an analogue of Corollary 4.3 for the Fenchel subdifferential. Then Luc [21] used the same method and a variant of Corollary 4.3 to obtain Theorem 5.6 for the Clarke-Rockafellar subdifferential.

6. A CLASS OF "SMOOTH" FUNCTIONS

In this concluding section we consider functions f of the form

f = g + h with $g: X \to \mathbb{R}$ lower semicontinuous and Gâteaux-differentiable, and $h: X \to \mathbb{R} \cup \{+\infty\}$ convex and lower semicontinuous.

A function f of this type may be considered "smooth" in the sense that Theorem 3.1 remains true, replacing ∂ by ∂^G , in an arbitrary Banach space X. Indeed, for such a function f it holds $\partial^G f = \nabla^G g + \partial h$, and Theorem 3.1 can actually be proved using the variational principle of Ekeland instead of the smooth variational principle of Borwein-Preiss.

Then, all the previous results hold for this class of functions in arbitrary Banach spaces. As an example, let us mention the following coercivity criterion due to Goeleven [16]:

Theorem 6.1 (Coercivity). Let X be a Banach space and $f: X \to \mathbb{R} \cup \{+\infty\}$ be of the form f = g + h, where g is lower semicontinuous Gâteaux-differentiable and h is convex and lower semicontinuous. If f is bounded below and satisfies the following Palais-Smale type condition:

if (x_n) is a sequence in X such that $\lambda(x_n) := \inf\{\|x_n^*\| \mid x_n^* \in \partial^G f(x_n)\}$ converges to 0 and $(f(x_n))$ is bounded, then (x_n) has a bounded subsequence,

then f is coercive.

Remark 6.1.1. The case of f Gâteaux-differentiable was previously treated by Caklovic et al. [5]. In [11], Costa and Silva consider the Fréchet-differentiable case but mention that the Gâteaux-differentiable case and the locally Lipschitz case (using Clarke subdifferential) can be obtained equally. Theorem 5.1 contains the latter case.

REFERENCES

- 1. E. Asplund, Fréchet differentiability of convex functions, Acta Math. 121 (1968), 31-47.
- 2. D. Aussel, J.-N. Corvellec, and M. Lassonde, Subdifferential characterization of quasiconvexity and convexity, J. Convex Analysis 1 (1994), 1-7.
- 3. J. M. Borwein and D. Preiss, A smooth variational principle with applications to subdifferentiability and to differentiability of convex functions, Trans. Amer. Math. Soc. 303 (1987), 517-527.
- 4. J. M. Borwein and H. Strojwas, *Proximal analysis and boundaries of closed sets in Banach space, Part II: Applications*, Canad. J. Math. 39 (1987), 428-472.
- 5. L. Caklovic, S. Li, and M. Willem, A note on Palais-Smale condition and coercivity, Differential and Integral Equations 3 (1990), 799-800.
- 6. F. H. Clarke, Optimization and nonsmooth analysis, Wiley-Interscience, New York, 1983.
- 7. F. H. Clarke, Yu. S. Ledyaev, R. J. Stern, and P. R. Wolenski, *Introduction to nonsmooth analysis*, book in preparation.
- 8. F. H. Clarke, R. J. Stern, and P. R. Wolenski, Subgradient criteria for monotonicity, the Lipschitz condition, and convexity, Canad. J. Math. 45 (1993), 1167-1183.
- 9. R. Correa, A. Jofré, and L. Thibault, Characterization of lower semicontinuous convex functions, Proc. Amer. Math. Soc. 116 (1992), 67-72.
- 10. ______, Subdifferential monotonicity as characterization of convex functions, Numer. Funct. Anal. Optim. 15 (1994), 531-535.
- 11. D. G. Costa and E. A. de B. e Silva, *The Palais-Smale condition versus coercivity*, Nonlinear Anal. 16 (1991), 371-381.
- 12. M. G. Crandall, L. C. Evans, and P. L. Lions, Some properties of viscosity solutions of Hamilton-Jacobi equations, Trans. Amer. Math. Soc. 282 (1984), 487-502.
- 13. M. Degiovanni, A. Marino, and M. Tosques, Evolution equations with lack of convexity, Nonlinear Anal. 9 (1985), 1401-1443.
- 14. R. Deville, G. Godefroy, and V. Zizler, A smooth variational principle with applications to Hamilton-Jacobi equations in infinite dimensions, J. Funct. Anal. 111 (1993), 197-212.
- 15. I. Ekeland, On the variational principle, J. Math. Anal. Appl. 47 (1974), 324-353.
- 16. D. Goeleven, A note on Palais-Smale condition in the sense of Szulkin, Differential and Integral Equations 6 (1993), 1041-1043.
- 17. A. D. Ioffe, Approximate subdifferentials and applications. I: The finite dimensional theory, Trans. Amer. Math. Soc. 281 (1984), 389-416.
- 18. _____, Calculus of Dini subdifferentials of functions and contingent coderivatives of set-valued maps, Nonlinear Anal. 8 (1984), 517-539.
- 19. _____, Proximal analysis and approximate subdifferentials, J. London Math. Soc. 41 (1990), 175–192.
- D. T. Luc, Characterizations of quasiconvex functions, Bull. Austral. Math. Soc. 48 (1993), 393-405.
- 21. _____, On the maximal monotonicity of subdifferentials, Acta Math. Vietnam. 18 (1993), 99-106.
- 22. L. McLinden, An application of Ekeland's theorem to minimax problems, Nonlinear Anal. 6 (1982), 189-196.
- 23. J.-P. Penot, Sous-différentiels de fonctions numériques non convexes, C.R. Acad. Sci. Paris 278 (1974), 1553-1555.

- 24. _____, On the mean value theorem, Optimization 19 (1988), 147–156.
- 25. R. R. Phelps, Convex functions, monotone operators and differentiability, Lecture Notes in Math., vol. 1364, 2nd ed., Springer-Verlag, Berlin, 1993.
- 26. R. A. Poliquin, Subgradient monotonicity and convex functions, Nonlinear Anal. 14 (1990), 305-317.
- 27. R. T. Rockafellar, On the maximal monotonicity of subdifferential mappings, Pacific J. Math. 33 (1970), 209-216.
- 28. _____, Directionally Lipschitzian functions and subdifferential calculus, Proc. London Math. Soc. 39 (1979), 331–355.
- 29. _____, Generalized directional derivatives and subgradients of nonconvex functions, Canad. J. Math. 32 (1980), 257-280.
- 30. _____, Proximal subgradients, marginal values and augmented Lagrangians in nonconvex optimization, Math. Oper. Res. 6 (1981), 424-436.
- 31. S. Simons, The least slope of a convex function and the maximal monotonicity of its subdifferential, J. Optim. Theory Appl. 71 (1991), 127-136.
- 32. L. Thibault and D. Zagrodny, Integration of subdifferentials of lower semicontinuous functions on Banach spaces, J. Math. Anal. Appl. 189 (1995), 33-58.
- 33. J. S. Treiman, Generalized gradients, Lipschitz behavior and directional derivatives, Canad. J. Math. 37 (1985), 1074-1084.
- 34. D. Zagrodny, Approximate mean value theorem for upper subderivatives, Nonlinear Anal. 12 (1988), 1413-1428.

MATHÉMATIQUES, UNIVERSITÉ DE PERPIGNAN, 66860 PERPIGNAN CEDEX, FRANCE *E-mail address*: corvellec@syspo.univ-perp.fr

MATHÉMATIQUES APPLIQUÉES, UNIVERSITÉ BLAISE PASCAL, CLERMONT-FERRAND, 63177 AUBIÈRE CEDEX, FRANCE

E-mail address: aussel@ucfma.univ-bpclermont.fr

Mathématiques, Université des Antilles et de la Guyane, 97167 Pointe-à-Pitre Cedex, Guadeloupe, France

E-mail address: lassonde@univ-ag.fr