Pergamon

Nonlinear Analysis, Theory, Methods & Applications, Vol. 33, No. 1, pp. 71–90, 1998
© 1998 Elsevier Science Ltd. All rights reserved
Printed in Great Britain
0362-546X/98 \$19.00+0.00

PII: S0362-546X(97)00511-7

ε-SUBDIFFERENTIAL AND ε-MONOTONICITY†‡

ALEJANDRO JOFRÉ,§ DINH THE LUC¶ and MICHEL THÉRA||

§ Departamento Ingenieria Matematica, Casilla 170/3, Correo 3, Universidad de Chile, Santiago, Chile; ¶ Institute of Mathematics, P.O. Box 10000, Hanoï, Vietnam; and || LACO, ESA 6090, Université de Limoges, 123, avenue Albert Thomas, 87060 Limoges Cedex, France

(*Received 1 December 1995; received for publication 14 May 1997*)

Key words and phrases: Convex function, mean-value theorem, nonsmooth analysis, generalized differentiation, extended-real-valued functions, nonconvex calculus, Asplund spaces, sequential limit, monotonicity.

1. INTRODUCTION

For convex functions, the concept of approximate or ε-subgradient has become a useful tool in optimization. This concept has attracted many mathematicians who have contributed numerous and important results. ε-subgradients play a central role in the theory of extremal problems and have been successfully used to construct numerical methods in order to minimize convex functions. The reader is referred to the book of Hiriart-Urruty and Lemaréchal ([1] comments on Chap. XI, Vol. II, p. 333).

For nonconvex functions, numerous approaches have also been developed to extend the concept of ε-gradient. For instance Ekeland and Lebourg [2], defined the concept of local ε-subdifferentiability, while Goldstein [3] introduced, for locally Lipschitz functions f on $\mathbb{R}^n$, a concept of ε-subgradient whose range is the convex hull of existing values of ∇f and limits of ∇f on a closed ε-ball. As an application, the author of [3] built a descent algorithm for the optimization of locally Lipschitz continuous functions on compact subsets of $\mathbb{R}^n$. Dixon [4] introduced the concept of gradient-ball for a class of almost everywhere differentiable functions on $\mathbb{R}^n$. For locally Lipschitz functions on $\mathbb{R}^n$, Polak *et al.* [5] defined a concept of generalized ε-delocalise gradient, while Bihain [6] defined the notion of ε–η-gradient and Bustos [7] the concepts of ε-locally generalized gradient and ε-globally generalized gradient.

For general extended-real-valued functions and on the base of sequential limits of the so-called Fréchet ε-normals, Kruger and Mordukhovich [8–9] introduced a concept of subdifferential which enjoys a rich calculus and permits to derive necessary and sufficient conditions for openness, metric regularity, Lipschitzian behavior of multifunctions with applications to optimization, sensitivity analysis, $\cdots$. For more details, the reader is, for instance, referred to the work of Mordukhovich and Shao [10, 11].

While in the previous construction, $\varepsilon \downarrow 0$ is involved in the sequential limit together with points, we prefer in this paper to use instead the sequential Painlevé–Kuratowski upper limit of the ε-Fréchet subdifferential with respect to the norm topology on X and the weak* topology on X^*, $\varepsilon > 0$ fixed. By doing this, we define for an extended-real-valued function an

† The first author was partially supported by Fondo Nacional de Ciencia y Tecnología, FONDECYT and EC under Contract CI1CT940115, while the last two authors are partially supported by a project AUPELF-UREF "Coopération en Mathématiques appliquées entre la France, la Belgique et le Vietnam".

‡ The first part of the paper has been presented during the "Journées ECOS", Limoges, 26–27 February, 1995, while the second part was presented at the *15th Mathematical Programming Symposium* in Ann Harbor, 14–19 August 1994.

ε-subdifferential. We call it the limiting Fréchet ε-subdifferential, and it is denoted by $\hat{\partial}_\varepsilon$. In the second section of the paper special attention is paid to the relation of this limiting Fréchet ε-differential with different generalized subdifferentials such as the Clarke, the Kruger–Mordukhovich and the Ioffe approximate subdifferential. Finally, we show that this subdifferential enjoys a quite satisfying calculus, including calculus rules for the sum, the composition and the marginal function.

Over the last few years, it appears, since the work of Poliquin [12], that one of the interesting topics of convex analysis is the study of the link between the convexity of a function and the maximal monotonicity of its subdifferential (see [13–16]). Poliquin proved that the monotonicity of the Clarke subdifferential of a lower semicontinuous function defined on a finite dimensional space is equivalent to its convexity. Recently, Correa *et al.* [13] by using the Moreau–Yosida approximation and the mean value theorem of Zagrodny (Z.M.V.T.) [17] extended this characterization of convexity to reflexive Banach spaces. Later, an extension to any Banach space and any subdifferential (belonging to a certain class) was proved by the same authors [14] using only Z.M.V.T. In parallel, another proof was established by Luc [15] based also on Z.M.V.T. The case of the Lipschitz-smooth subdifferential is proved in Clarke *et al.* [18] for Hilbert spaces. In a recent paper Aussel *et al.* [19], using the smooth variational principle of Borwein and Preiss established an approximate mean value inequality which permitted them to unify some previous results. Finally, Correa *et al.* [20] have proved that convexity is equivalent to the monotonicity of a more general notion of subdifferential called *Presubdifferential*. In the general context of operators, using a general and yet very transparent property of certain bifunctions and a concept of maximal ε-monotonicity (introduced in an earlier version of the present paper), Oettli and Théra [21] established a characterization of the maximal monotonicity of the ε-enlargement of a given multifunction in terms of the monotonicity of the multifunction. Since, whenever f is convex the limiting Fréchet ε-subdifferential is exactly the ε-enlargement of the Frenchel subdifferential, it is natural, in the last section of the paper the use this notion of ε-monotonicity for multifunctions and to characterize the ε-monotonicity of the limiting Fréchet ε-subdifferential in terms of generalized convexity.

2. THE LIMITING FRÉCHET ε-SUBDIFFERENTIAL: DEFINITION AND BASIC PROPERTIES

Throughout this paper, let us denote by X a Banach space, X^* its topological dual, B the closed unit ball in X and B^* the closed unit ball in X^*. We adopt the following notation: "$\xrightarrow{s}$", (respectively "$\xrightarrow{w^*}$") denotes the convergence with respect to the strong (respectively the weak* topology), while $x_n \xrightarrow{f} x$ means that the sequence $\{x_n\}_{n\in\mathbb{N}}$ converges to x while the sequence $\{f(x_n)\}_{n\in\mathbb{N}}$ converges to $f(x)$. We use the abbreviations cl* and cl*conv to denote the weak* closure and the weak* closed convex hull, respectively. Let f be a lower semicontinuous function from X to $\mathbb{R} \cup \{+\infty\}$. We recall that

$$\mathrm{Dom} f := \{x \in X \mid f(x) < +\infty\}$$

stands for the *effective domain* of f and that f is *proper* if it has a nonempty domain. For a convex function the *Fenchel subdifferential* of f at $x \in \mathrm{Dom} f$ is the set

$$\partial^{\mathrm{Fen}} f(x) := \{x^* \in X^* \mid \langle x^*, y - x\rangle \le f(y) - f(x) \text{ for each } y \in X\}, \tag{2.1}$$

while the *directional derivative* of f at x in the direction d is given by

$$f'(x; d) := \lim_{t \downarrow 0} \frac{f(x + td) - f(x)}{t}. \tag{2.2}$$

Let $f: X \to \mathbb{R} \cup \{+\infty\}$ be given and let ε be a fixed nonnegative number. We recall that the *Fréchet ε-subdifferential* of f at $x \in \operatorname{Dom} f$ is defined by

$$\partial_\varepsilon^F f(x) := \left\{ x^* \in X^* \;\middle|\; \liminf_{\|h\| \to 0} \frac{f(x+h) - f(x) - \langle x^*, h\rangle}{\|h\|} \geq -\varepsilon \right\}. \tag{2.3}$$

Remark 2.1. When $\varepsilon = 0$, (2.3) reduces to the well known *Fréchet subdifferential* which is denoted by $\partial^F f(x)$. More precisely,

$x^* \in \partial^F(x)$, if and only if for each $\eta > 0$, there exists $\delta > 0$ such that $\langle x^*, y - x\rangle \leq f(y) - f(x) + \eta\|y - x\|$, for all $y \in x + \delta B$.

As noted by Treiman [22]:

$$x^* \in \partial_\varepsilon^F f(x) \Leftrightarrow x^* \in \partial^F(f + \varepsilon\|\cdot - x\|)(x). \tag{2.4}$$

Equivalently, this amounts to saying that $x^* \in \partial_\varepsilon^F f(x)$, if and only if for each $\eta > 0$, there exists $\delta > 0$ such that

$$\langle x^*, y - x\rangle \leq f(y) - f(x) + (\varepsilon + \eta)\|y - x\|, \qquad \text{for all } y \in x + \delta B. \tag{2.5}$$

Note that if f is convex, then

$$\partial_\varepsilon^F f(x) = \{x^* \in X^* \mid f(x+h) - f(x) - \langle x^*, h\rangle \geq -\varepsilon\|h\| \text{ for each } h \in X\}. \tag{2.6}$$

When f is the *indicator function* $I(\cdot; \Omega)$ of a closed set Ω defined by

$$I(x; \Omega) := \begin{cases} 0 & \text{if } x \in \Omega \\ +\infty, & \text{otherwise,} \end{cases}$$

then for each $\bar{x} \in \Omega$ and each $\varepsilon \geq 0$, one obviously has

$$\partial_\varepsilon^F I(\cdot; \Omega)(\bar{x}) = N_\varepsilon^F(\Omega; \bar{x}),$$

where

$$N_\varepsilon^F(\Omega; \bar{x}) := \left\{ x^* \in X^* \;\middle|\; \limsup_{\substack{x \to \bar{x} \\ x \in \Omega}} \frac{\langle x^*, x - \bar{x}\rangle}{\|x - \bar{x}\|} \leq \varepsilon \right\}$$

is the set of *ε-Fréchet normals* to Ω at $\bar{x}$. Note that $N_\varepsilon^F(\Omega; \bar{x})$ is convex but may be not weak* closed. The reader is referred to the papers [2, 8, 10, 11, 22–26] etc. . . . for precise details of Fréchet ε-subdifferentiation and its applications.

For our purpose, let us briefly list some properties of the Fréchet ε-subdifferential. Let $f, g: X \to \mathbb{R} \cup \{+\infty\}$ be given, then,

(i) $\partial_\varepsilon^F f(x)$ is empty if $x \notin \operatorname{Dom} f$;
(ii) $\partial_\varepsilon^F f(x)$ is a convex subset of X^*;
(iii) $\partial_{\varepsilon_1}^F f(x) + \partial_{\varepsilon_2}^F f(x) \subseteq \partial_{\varepsilon_1 + \varepsilon_2}^F (f + g)(x)$;
(iv) $\partial_{\varepsilon_1}^F f(x) \subseteq \partial_{\varepsilon_2}^F f(x)$, if $\varepsilon_1 \leq \varepsilon_2$;
(v) if $x^* \in \bigcap_{\varepsilon > 0} (\partial_\varepsilon^F f(x) \cap (-\partial_\varepsilon^F(-f)(x)))$, then x^* is the Fréchet derivative of f at x;
(vi) if $\partial_\varepsilon^F f(x)$ is nonempty, then f is lower semicontinuous at x.

We now defined a new kind of approximate subdifferential. We call *limiting Fréchet ε-subdifferential* of f at $x \in \operatorname{Dom} f$ the set given by

$$\hat{\partial}_{\varepsilon} f(x) := \limsup_{y \xrightarrow{f} x} \partial_{\varepsilon}^{F} f(y), \tag{2.7}$$

where "lim sup" stands for the sequential Painlevé–Kuratowski upper limit, i.e.

$$\limsup_{y \xrightarrow{f} z} \partial_{\varepsilon}^{F} f(y) := \{x^* \in X^* \mid \exists \text{ sequences } x_n^* \xrightarrow{f} x,\ x_n^* \xrightarrow{w^*} x^* \text{ with } x_n^* \in \partial_{\varepsilon}^{F} f(x_n)\ \forall\, n \in \mathbb{N}\}. \tag{2.8}$$

Note that in the infinite dimensional setting, the weak* topology on X^* is not sequential and therefore the sequential Painlevé–Kuratowski upper limit as defined in (2.8) does not ensure either the weak* closedness or the weak* sequential closedness of the limiting Fréchet ε-subdifferential of f at $x \in \operatorname{Dom} f$.

It appears that a nice framework to develop calculus rules for the limiting Fréchet ε-subdifferential is the broad subclass of Banach spaces called Asplund spaces. *Asplund spaces* are those spaces for which every convex lower semicontinuous functional $f : X \to \mathbb{R} \cup \{+\infty\}$ is Fréchet differentiable on a dense G_δ subset of $\operatorname{Int}(\operatorname{Dom} f)$. In particular, this class includes reflexive Banach spaces as well as $c_0(I)$ for any index set I, or spaces for which there exists an equivalent Fréchet norm. Conversely, if X is Asplund and separable, then it has an equivalent Fréchet norm. On the other hand, the class of Asplund spaces does not include $\mathscr{C}([0, 1])$, $\ell^1(\mathbb{N})$ and $\ell^\infty(\mathbb{N})$.

Let us also recall that X is declared *weakly compactly generated* (WCG for short) if there exists a weakly compact subset K whose closed span is the entire space. These include separable spaces, reflexive spaces, σ-finite L^1 spaces, $\mathscr{C}(T)$ with T an Eberlein compactum and others (see Deville *et al.* [27] for details).

In the sequel, we will also, when necessary, make use of Banach spaces whose closed balls of their dual are weak* sequentially compact. This class of Banach spaces includes the class of Asplund spaces (Stegall [28], Theorem 3.5) , as well as, subspaces of WCG spaces (Amir and Lindenstrauss's Theorem) and Banach spaces with a smooth renorm (Hagler and Sullivan's Theorem).

Let us now establish a link between the limiting Fréchet ε-subdifferential, the Fréchet ε-subdifferential and the Fenchel subdifferential, respectively.

PROPOSITION 2.2. Let $\varepsilon \geq 0$ be fixed and let $f : X \to \mathbb{R} \cup \{+\infty\}$ be a lower semicontinuous function. Suppose X is an Asplund space. Then the domain of $\hat{\partial}_{\varepsilon} f$, that is, $\{x \in X \mid \hat{\partial}_{\varepsilon} f(x) \neq \emptyset\}$ is graphically dense in $\operatorname{Dom} f$.

Proof. Combine Fabian's characterization of Asplundness with the definition of the limiting Fréchet ε-subdifferential. ■

PROPOSITION 2.3. Let $\varepsilon \geq 0$ be fixed and let $f : X \to \mathbb{R} \cup \{+\infty\}$ be convex. Suppose X is a Banach space with a weak* sequential unit dual ball. Then

$$\partial_{\varepsilon}^{F} f(x) = \partial^{\mathrm{Fen}} f(x) + \varepsilon B^* \tag{2.9}$$

and

$$\hat{\partial}_{\varepsilon} f(x) = \partial_{\varepsilon}^{F} f(x). \tag{2.10}$$

In particular,

$$\partial^{\mathrm{Fen}} f(x) = \bigcap_{\varepsilon>0} \partial_\varepsilon^F f(x) \qquad \text{for all } x \in \operatorname{Dom} f.$$

Proof. Let $x \in \operatorname{Dom} f$, and suppose $x^* \in \partial_\varepsilon^F f(x)$. Then using the characterization (2.6) we have

$$\langle x^*, y - x\rangle \le f(y) + \varepsilon\|x - y\| - f(x), \qquad \text{for all } y \in X.$$

Since the latter equation is equivalent to say that

$$x^* \in \partial^{\mathrm{Fen}}(f + \varepsilon\|x - \cdot\|)(x),$$

by virtue of the Moreau–Rockafellar Theorem on the sum, we derive that

$$x^* \in \partial^{\mathrm{Fen}} f(x) + \varepsilon B^*,$$

and therefore $\partial_\varepsilon^F f(x) \subseteq \partial^{\mathrm{Fen}} f(x) + \varepsilon B^*$. The reverse inclusion being trivial through (2.6), thus equality (2.9) is established.

Let $x^* \in \hat{\partial}_\varepsilon f(x)$. Then, there exist sequences $\{x_n\}_{n\in\mathbb{N}}$, $\{x_n^*\}_{n\in\mathbb{N}}$ such that $x_n \xrightarrow{f} x$, $x_n^* \in \partial_\varepsilon^F f(x_n)$ and $x_n^* \xrightarrow{w^*} x^*$. Using (2.9), there exist sequences $\{u_n^*\}_{n\in\mathbb{N}}$ and $\{b_n^*\}_{n\in\mathbb{N}}$ such that $b_n^* \in B^*$, $u_n^* \in \partial^{\mathrm{Fen}} f(x)$ and

$$x_n^* = u_n^* + \varepsilon b_n^*.$$

Since by hypothesis the closed unit ball B^* of X^* is weak* sequentially compact, on relabeling if necessary, we may suppose that $b_n^* \xrightarrow{w^*} b^* \in B^*$. By a standard argument and using the fact that the Fenchel subdifferential is graph-closed, we deduce that $\hat{\partial}_\varepsilon f(x) \subseteq \partial^{\mathrm{Fen}} f(x) + \varepsilon B^*$ and equality $\partial_\varepsilon^F f(x) = \hat{\partial}_\varepsilon f(x)$ follows. ■

Remark 2.4. It is important to note that given a convex extended-real-valued function f on X, thanks to (2.9), the limiting Fréchet ε-subdifferential differs from the usual Fenchel ε-subdifferential, which is always nonempty for $\varepsilon > 0$, and is defined by

$$\partial_\varepsilon^{\mathrm{Fen}} f(x) := \{x^* \in X^* \mid f(y) - f(x) \ge \langle x^*, y - x\rangle - \varepsilon \ \forall y \in X\}.$$

Consider the real function given by $f(x) = |x|$. Then,

$$\hat{\partial}_\varepsilon f(x) = \begin{cases} [1 - \varepsilon, 1 + \varepsilon] & \text{if } x > 0 \\ [-1 - \varepsilon, 1 + \varepsilon] & \text{if } x = 0 \\ [-1 - \varepsilon, -1 + \varepsilon] & \text{if } x < 0, \end{cases}$$

while

$$\partial_\varepsilon^{\mathrm{Fen}} f(x) = \begin{cases} \left[-1, -1 + \dfrac{\varepsilon}{|x|}\right] & \text{if } x > 0 \text{ and } \varepsilon < 2|x| \\ [-1, 1] & \text{if } x = 0 \text{ or } \varepsilon \ge 2|x| \\ \left[1 - \dfrac{\varepsilon}{|x|}, 1\right] & \text{if } x < 0 \text{ and } \varepsilon < 2x. \end{cases}$$

The following observation deserves being displayed.

LEMMA 2.5. Let f be a convex extended-real-valued function on X. For each $x \in \operatorname{Dom} f$ and each $\varepsilon \geq 0$ the following inclusion holds:

$$\partial_\varepsilon^{\mathrm{Fen}} F(x) \subseteq \bigcup_{\|y-x\| \leq \sqrt{\varepsilon}} \hat{\partial}_{\sqrt{\varepsilon}} f(y).$$

Proof. Let $x^* \in \partial_\varepsilon^{\mathrm{Fen}} f(x)$. By virtue of the Brøndsted–Rockafellar Theorem, there exist $x_\varepsilon \in X$ and $x_\varepsilon^* \in X^*$ such that

$$\|x_\varepsilon - x\| \leq \sqrt{\varepsilon}$$
$$\|x_\varepsilon^* - x^*\| \leq \sqrt{\varepsilon}$$
$$x_\varepsilon^* \in \partial^{\mathrm{Fen}} f(x_\varepsilon).$$

This infers, $x^* \in \bigcup_{\|y-x\| \leq \sqrt{\varepsilon}} \hat{\partial}_{\sqrt{\varepsilon}} f(y)$. ■

Equality fails in general, as for $f(x) := |x|$, $\bigcup_{\|y-x\| \leq \sqrt{\varepsilon}} \hat{\partial}_{\sqrt{\varepsilon}} f(y) = [-1 - \sqrt{\varepsilon}, 1 + \sqrt{\varepsilon}]$, while $\partial_\varepsilon^{\mathrm{Fen}} f(0) = [-1, 1]$.

COROLLARY 2.6. Let X be a Banach space with a weak* sequential unit dual ball. Suppose f and g are convex and finite at $x \in \operatorname{Dom} f \cap \operatorname{Dom} g$. Then,

$$\bigcup_{\substack{\alpha \geq 0, \beta \geq 0 \\ \alpha + \beta = \varepsilon}} \hat{\partial}_\alpha f(x) + \hat{\partial}_\beta g(x) \subset \hat{\partial}_\varepsilon (f + g)(x). \tag{2.11}$$

If furthermore a qualification condition holds true† then equality holds.

Proof. By virtue of Proposition 2.3 formula (2.10), we obtain

$$\hat{\partial}_\alpha f(x) + \hat{\partial}_\beta g(x) = \hat{\partial}_\alpha^F f(x) + \hat{\partial}_\beta^F g(x)$$

and therefore by formula (2.9) we obtain

$$\begin{aligned} \hat{\partial}_\alpha f(x) + \hat{\partial}_\beta g(x) &= \partial^{\mathrm{Fen}} f(x) + \alpha B^* + \partial^{\mathrm{Fen}} g(x) + \beta B^* \\ &\subseteq \partial^{\mathrm{Fen}} f(x) + \partial^{\mathrm{Fen}} g(x) + (\alpha + \beta) B^* \\ &\subseteq \partial^{\mathrm{Fen}} (f + g)(x) + \varepsilon B^* \\ &= \hat{\partial}_\varepsilon (f + g)(x). \end{aligned}$$

Conversely, under a qualification condition

$$\partial^{\mathrm{Fen}} (f + g)(x) = \partial^{\mathrm{Fen}} f(x) + \partial^{\mathrm{Fen}} g(x).$$

Therefore, for each $\alpha \geq 0$, $\beta \geq 0$ such that $\alpha + \beta = \varepsilon$ we have

$$\begin{aligned} \hat{\partial}_\varepsilon (f + g)(x) &= \partial^{\mathrm{Fen}} (f + g)(x) + \varepsilon B^* \\ &= (\partial^{\mathrm{Fen}} f(x) + \alpha B^*) + (\partial^{\mathrm{Fen}} g(x) + \beta B^*) \\ &= \hat{\partial}_\alpha f(x) + \hat{\partial}_\beta g(x) \\ &= \hat{\partial}_\alpha f(x) + \hat{\partial}_\beta g(x). \quad ■ \end{aligned}$$

† We may, for instance, assume that the Brézis–Attouch condition is satisfied: $\bigcup_{\lambda > 0} \lambda(\operatorname{Dom} f - \operatorname{Dom} g)$ is a closed subspace of X.

Remark 2.7. (1) It should be observed that, although the objects are different, the inclusion (2.11) is identical to the classical formula where, instead of the limiting Fréchet ε-subdifferential we use the Fenchel ε-subdifferential.

(2) When Ω is a closed convex set, $N_\varepsilon^F(\Omega; \bar{x})$ reduces to the ε-enlargement of the classical normal cone $N(\Omega; \bar{x})$, namely,

$$N_\varepsilon^F(\Omega; \bar{x}) = N(\Omega; \bar{x}) + \varepsilon B^*.$$

In the case where f is continuously differentiable, the limiting Fréchet ε-subdifferential takes a very simple form.

PROPOSITION 2.8. Let X be a Banach space and suppose that $f: X \to \mathbb{R}$ is continuously Fréchet differentiable at x with Fréchet derivative $Df(x)$. Then

$$\hat{\partial}_\varepsilon f(x) = Df(x) + \varepsilon B^*.$$

Proof. Let $x^* \in B^*$ be given. From the Fréchet differentiability of f, we have

$$\begin{aligned} f(x+h) - f(x) - \langle Df(x) + \varepsilon x^*, h\rangle &= \langle Df(x), h\rangle + r(h)\|h\| - \langle Df(x), h\rangle - \varepsilon\langle x^*, h\rangle \\ &\geq (r(h) - \varepsilon)\|h\|, \end{aligned}$$

where $\lim_{h\to 0} r(h) = 0$. Hence, for every $\eta > 0$ there exists $\delta > 0$ such that

$$f(x+h) - f(x) - \langle Df(x) + \varepsilon x^*, h\rangle \geq -(\eta + \varepsilon)\|h\| \qquad \text{for all } h \in \delta B.$$

According to the Treiman characterization (2.5), this yields:

$$Df(x) + \varepsilon x^* \in \partial_\varepsilon^F f(x) \subset \hat{\partial}_\varepsilon f(x). \tag{2.12}$$

Conversely, suppose given $x^* \in \hat{\partial}_\varepsilon f(x)$. This amounts to saying that there exist sequences $\{x_n\}_{n\in\mathbb{N}}$, $\{x_n^*\}_{n\in\mathbb{N}}$ such that $x_n \xrightarrow{f} x$, $x_n^* \in \partial_\varepsilon^F f(x_n)$ and $x_n^* \xrightarrow{w^*} x^*$. By virtue of the characterization (2.5), for each $n \in \mathbb{N}$, and each $\eta > 0$, there exists $\delta_{n,\eta} > 0$ such that

$$f(x_n + h) - f(x_n) - \langle x_n^*, h\rangle \geq -(\eta + \varepsilon)\|h\| \qquad \text{for all } h \in \delta_{n,\eta} B.$$

Equivalently,

$$\langle Df(x_n), h\rangle + r_n(h)\|h\| \geq \langle x_n^*, h\rangle \qquad \text{for all } h \in \delta_{n,\eta} B. \tag{2.13}$$

Now, using the fact that $\lim_{h\to 0} r_n(h) = 0$, (2.13) yields:

$$\langle Df(x_n), h\rangle + (2\eta + \varepsilon)\|h\| \geq \langle x_n^*, h\rangle \qquad \text{for all } h \in \delta_{n,\eta} B\ \forall n \text{ large enough.} \tag{2.14}$$

Since relation (2.14) is positively homogeneous with respect to h, it remains true for all $h \in X$, and $\delta_{n,\eta} < \delta$ by using the continuity of $Df(\cdot)$ we obtain:

$$\langle Df(x), h\rangle + (2\eta + \varepsilon)\|h\| \geq \langle x^*, h\rangle \qquad \text{for every } h \in X.$$

As η is arbitrary, (2.15) yields

$$\langle Df(x), h\rangle + \varepsilon\|h\| \geq \langle x^*, h\rangle \qquad \text{for every } h \in X.$$

Finally, since x^* is arbitrary in $\hat{\partial}_\varepsilon f(x)$, the latter relation therefore yields

$$\hat{\partial}_\varepsilon f(x) \subset Df(x) + \varepsilon B^*. \tag{2.16}$$

The proof is established by combining (2.12) and (2.16). ■

The next proposition has been pointed out to the authors by Mordukhovich and Shao, after they received a first draft of the present paper.

As we mentioned in the introduction, if we involve $\varepsilon \downarrow 0$ in the definition of the limiting Fréchet ε-subdifferential, we obtain the Kruger–Mordukhovich subdifferential introduced in [8] and [9], which is an infinite dimensional extension of the nonconvex construction first proposed by Mordukhovich [29] and developed in several papers summarized (for the period) in [30]. This subdifferential will be denoted by $\partial^{\text{K-M}}$, that is,

$$\partial^{\text{K-M}} f(x) := \limsup_{u \xrightarrow{f} x,\, \varepsilon \downarrow 0} \partial_\varepsilon^F f(u).$$

PROPOSITION 2.9. Let f be an extended-real-valued functional on an Asplund space X, lower semicontinuous around $x \in \operatorname{Dom} f$. Then,

$$\hat{\partial}_\varepsilon f(x) = \partial^{\text{K-M}} f(x) + \varepsilon B^*. \tag{2.17}$$

Remark 2.10. Approximate Lipschitz selections: we have proved that if f is a convex function or a continuously differentiable function, then, respectively,

$$\hat{\partial}_\varepsilon f(x) = \partial^{\text{Fen}} f(x) + \varepsilon B^*$$

$$\hat{\partial}_\varepsilon f(x) = Df(x) + \varepsilon B^*.$$

Let us define the set-valued mapping $T_\varepsilon : X \rightrightarrows X \times X^*$ by

$$T_\varepsilon(x) := (x + \varepsilon B, \hat{\partial}_\varepsilon f(x)).$$

According to the approximate selection theorem (see [31], p. 84) to the Fenchel subdifferential $\partial^{\text{Fen}} f(x)$, there exists a locally Lipschitz function $s_\varepsilon : X \to X^*$ such that $s_\varepsilon(X) \subset \operatorname{conv}\{\partial^{\text{Fen}} f(x) \mid x \in \operatorname{Dom} \partial^{\text{Fen}} f\}$ and $\operatorname{Graph} s_\varepsilon \subset \{T_\varepsilon(x) \mid x \in X\}$. Hence, s_ε is a Lipschitz selection of T_ε. Analogously, we may obtain an approximate selection of $Df(\cdot) + \varepsilon B^*$.

We denote by $\partial^\dagger f(x)$ the *Clarke–Rockafellar subdifferential*. More precisely,

$$\partial^\dagger f(x) := \{x^* \in X^* \mid \langle x^*, d \rangle \leq f^\dagger(x, d)\ \forall d \in X\}$$

where $f^\dagger(x, d)$ is the *generalized Clarke–Rockafellar directional derivative* of f at x in direction d and is defined by

$$f^\dagger(x, d) := \limsup_{\substack{y \xrightarrow{f} x \\ t \downarrow d}} \inf_{u \to d} \frac{f(y + tu) - f(y)}{t}.$$

When f is Lipschitz around $x \in \operatorname{Dom} f$ then $f^\dagger(x, d)$ reduces to

$$\limsup_{y \to x,\, t \downarrow 0} \frac{f(y + td) - f(y)}{t}.$$

Ioffe, in a serie of papers, [23–25, 32] developed another line of infinite dimensional extensions of the finite dimensional Mordukhovich construction by considering instead of sequential Painlevé–Kuratowski limits, topological limits of ε-Dini subdifferentials with respect to the norm topology on X and the weak* topology on X^*. More precisely, he introduced the set $\partial^A f(x)$ which he called the *Approximate subdifferential* of f at x. If we denote by $\mathcal{F}(X)$ the collection of all finite dimensional subsets of X, then

$$\partial^A f(x) := \bigcap_{L \in \mathcal{F}(X)} \operatorname*{Limsup}_{y \xrightarrow{f} x\, \varepsilon \downarrow 0} \partial_\varepsilon^{\text{Dini}} f_{y+L}(y)$$

where

$$f_{y+L}(x) := \begin{cases} f(x) & \text{if } x \in y + L \\ +\infty & \text{otherwise.} \end{cases}$$

Recall that in the last formula,

$$\partial_\varepsilon^{\mathrm{Dini}} f(y) := \left\{ x^* \in X^* \mid \langle x^*, v \rangle \leq \liminf_{u \to v,\, t \downarrow 0} \frac{f(y + tu) - f(u)}{t} + \varepsilon \|v\| \ \forall v \in X \right\}$$

stands for the *ε-Dini subdifferential* of f at y. The abbreviation "Limsup" is used to express the topological Painlevé–Kuratowski limit, namely, if F denotes a multifunction from X into the subsets of X^*, then $x^* \in \operatorname{Limsup}_{y \to x} F(y)$, if for each weak* neigbourhood W of the origin in X^* and every neighbourhood V of x in X, there exists $y \in V$ such that $(W + x^*) \cap F(y) \neq \emptyset$.

For a lower semicontinuous function on the general subclass of Banach spaces called *weakly trustworthy spaces* (which includes in particular Asplund spaces) Ioffe [23] proved that $\partial^A f(x)$ is given by:

$$\partial^A f(x) := \operatorname*{Limsup}_{y \xrightarrow{f} x,\, \varepsilon \downarrow 0} \partial_\varepsilon^{\mathrm{Dini}} f(y).$$

The right comparison between Kruger–Mordukhovich's and Ioffe's approximate subdifferential is given in Mordukhovich and Shao ([11], Theorem 9.2).

THEOREM 2.11. ([11], Theorem 9.2) Assume X is an Asplund space. Then, for any function $f: X \to \mathbb{R} \cup \{-\infty, \infty\}$, lower semicontinuous around $x \in \operatorname{Dom} f$ one has the inclusion:

$$\partial^{\mathrm{K\text{-}M}} f(x) \subset \partial^A f(x). \tag{2.18}$$

Moreover if we suppose f Lipschitz around $x \in \operatorname{Dom} f$, then

$$\mathrm{cl}^* \partial^{\mathrm{K\text{-}M}} f(x) = \partial^A f(x).$$

In the latter case, if X is WCG, the set $\limsup_{u \to x,\, \varepsilon \downarrow 0} \partial_\varepsilon^{\mathrm{F}} f(u)$ is weak* closed and

$$\partial^{\mathrm{K\text{-}M}} f(x) = \partial^A f(x).$$

As noticed by Borwein and Fitzpatrick [33], the inclusion (2.18) may be strict, even for the case of Lipschitz functions on spaces with Fréchet differentiable norms.

The link between Kruger-Mordukhovich's subdifferential, Clarke's subdifferential and the limiting Fréchet ε-subdifferential is given in the next result:

PROPOSITION 2.12. For any extended-real-valued functional f, lower semicontinuous at $x \in \operatorname{Dom} f$, one has†

$$\bigcap_{\varepsilon > 0} \hat{\partial}_\varepsilon f(x) = \partial^{\mathrm{K\text{-}M}} f(x). \tag{2.19}$$

If moreover f is Lipschitz around $x \in \operatorname{Dom} f$ and X is an Asplund space, then

$$\mathrm{cl}^* \operatorname{conv}\left(\bigcap_{\varepsilon > 0} \hat{\partial}_\varepsilon f(x) \right) = \partial^{\uparrow} f(x). \tag{2.20}$$

† In fact it is sufficient to suppose that the unit ball of X^* is weak* sequentially compact.

Proof. Let $x^* \in \partial^{\text{K-M}} f(x$. Then by definition, there exist a sequence $\{\varepsilon_n\} n \in \mathbb{N}$ converging to zero with $\varepsilon_n > 0$ and sequences $\{x_n\}_{n\in\mathbb{N}}$, $\{x_n^*\}_{n\in\mathbb{N}}$ with $x_n \xrightarrow{f} x$, $x_n^* \in \partial^F_{\varepsilon_n}(x_n)$ and $x_n^* \xrightarrow{w^*} x^*$. For an arbitrary $\varepsilon > 0$, one can find $n_0 > 0$ such that $\varepsilon_n < \varepsilon$ for n large enough. Thus, $x_n^* \in \partial^F_\varepsilon f(x_n)$ for $n \geq n_0$. It follows that $x^* \in \hat{\partial}_\varepsilon f(x)$ and the inclusion $\partial^{\text{K-M}} f(x) \subset \bigcap_{\varepsilon>0} \hat{\partial}_\varepsilon f(x)$ is established.

For the converse inclusion, let $x^* \in \bigcap_{\varepsilon>0} \hat{\partial}_\varepsilon f(x)$. For each $\varepsilon > 0$ there exist sequences $\{x_{n,\varepsilon}\}_{n\in\mathbb{N}}$ and $\{x^*_{n,\varepsilon}\}_{n\in\mathbb{N}}$ such that $x_{n,\varepsilon} \xrightarrow{f} x$, $x^*_{n,\varepsilon} \in \hat{\partial}_\varepsilon f(x_{n,\varepsilon})$ and $x^*_{n,\varepsilon} \xrightarrow{w^*} x^*$. Given $\varepsilon_i \to 0$, $\varepsilon_i > 0$, using a diagonalization procedure, we may suppose that $x_{n_i,\varepsilon_i} \to x$, $x^*_{n_i,\varepsilon_i} \xrightarrow{w^*} x^*$. Then, $x^* \in \partial^{\text{K-M}} f(x)$. In this way, formula (2.19) is true.

The second formula is obtained from (2.19) and from the fact that $\partial^\dagger f(x) = \text{cl}^* \operatorname{conv} \partial^{\text{K-M}} f(x)$ ([11] Theorem 8.11 formula (8.21)). ■

Remark 2.13. The preceding result still holds true if we only suppose that X is a separable Banach space. Let us notice that if k stands for the Lipschitz constant of f in a neighbourhood of x, we have

$$\partial^F_\varepsilon f(y) \subset (k + \varepsilon + 1)B^* \qquad \text{for } y \text{ near } x. \tag{2.21}$$

Since X is separable, it is metrizable and each bounded subset K of X^* is metrizable for the weak* topology. Therefore, by virtue of (2.21) the diagonalization procedure used above still works.

In order to have the Fermat rule by means of the limiting Fréchet ε-subdifferential we need the following definition.

Definition 2.14. Let $x_0 \in \operatorname{Dom} f$. We say that x_0 is an ε-minimizer of f if

$$f(x_0 + h) \geq f(x_0) - \varepsilon \qquad \text{for all } h \in X,$$

and we say that x_0 is an $\varepsilon\|\cdot\|$-minimizer of f if

$$f(x_0 + h) \geq f(x_0) - \varepsilon\|\eta\| \qquad \text{for all } h \in X.$$

PROPOSITION 2.15. If $\bar{x}$ is an $\varepsilon\|\cdot\|$-minimum of f, then $0 \in \hat{\partial}_\varepsilon f(\bar{x})$. If x_0 is an εminimum of f, then for every $\delta > 0$ there exists $\bar{x}$ such that $\|\bar{x} - x_0\| < \delta$ and $\bar{x}$ is an $\varepsilon/\delta\|\cdot\|$-minimum of f.

Proof. The first conclusion is a direct consequence of the definition of the Fréchet ε-subdifferential operator. The second one is deduced from the Ekeland variational principle. ■

2.1. *The sum rule for the limiting Fréchet ε-subdifferential*

In this section we obtain a satisfactory formula for the limiting Fréchet ε-subdifferential of the sum of two extended-real-valued functionals. The main tool in proving this result will be the following "fuzzy sum rule" of Fabian ([34] Theorem 3).

PROPOSITION 2.16. Let X be an Asplund space and suppose given two lower semicontinuous functions $f_1, f_2 : X \to \mathbb{R} \cup \{+\infty\}$, one of them being locally Lipschitzian. Then, for each $x \in \operatorname{Dom} f_1 \cap \operatorname{Dom} f_2$, for each $\varepsilon \geq 0$, $\delta > 0$ and $\gamma > 0$ we have:

$$\partial^F_\varepsilon (f_1 + f_2)(x) \subset \bigcup_{\substack{x_i \in X,\ \|x_i - x\| < \delta \\ |f_i(x_i) - f_i(x)| < \delta \\ i=1,2}} \{\partial^F_\varepsilon f_1(x_1) + \partial^F_\varepsilon f_2(x_2) + (\varepsilon + \gamma)B^*\}. \tag{2.22}$$

Exploiting this "fuzzy sum rule" we are now in a position to state a sum rule formula for the limiting ε-subdifferential.

THEOREM 2.17. Let X be an Asplund space. Let $f, g: X \to \mathbb{R} \cup \{+\infty\}$ be two lower semicontinuous proper functions one of them being locally Lipschitzian. Then, for each $x \in \operatorname{Dom} f \cap \operatorname{Dom} g$ and each $\varepsilon > 0$ the following inclusion holds:

$$\hat{\partial}_\varepsilon(f+g)(x) \subseteq \hat{\partial}_\varepsilon f(x) + \hat{\partial}_\varepsilon g(x).$$

Proof. Let $x^* \in \hat{\partial}_\varepsilon(f+g)(x)$. Due to formula (2.8), there exist sequences $\{x_n\}_{n\in\mathbb{N}}$, $\{x_n^*\}_{n\in\mathbb{N}}$ such that $x_n \xrightarrow{f+g} x$, $x_n^* \xrightarrow{w^*} x^*$ and $x_n^* \in \partial_\varepsilon^F(f+g)(x_n)$. By virtue of the "fuzzy sum rule" (2.22) applied for $\delta = (1/n)$ and $\gamma = \varepsilon$ there exist sequences $\{y_n\}_{n\in\mathbb{N}}$, $\{z_n\}_{n\in\mathbb{N}}$, $\{y_n^*\}_{n\in\mathbb{N}}$ and $\{z_n^*\}_{n\in\mathbb{N}}$ such that:

(a) $$\|y_n - x_n\| \le \frac{1}{n}, \qquad \|z_n - x_n\| \le \frac{1}{n},$$

(b) $$|f(y_n) - f(x_n)| \le \frac{1}{n}, \qquad |g(z_n) - g(x_n)| \le \frac{1}{n},$$

(c) $$y_n^* \in \partial_\varepsilon^F f(y_n), \qquad z_n^* \in \partial_\varepsilon^F g(z_n),$$

and

$$x_n^* \in y_n^* + z_n^* + 2\varepsilon B^*. \tag{2.23}$$

As $x_n \xrightarrow{f+g} x$, assertion (a) together with assertion (b) insure that

$$y_n \xrightarrow{f} x \qquad \text{and} \qquad z_n \xrightarrow{g} x.$$

As $x_n^* \xrightarrow{w^*} x^*$ we derive from (2.23) that $y_n^* + z_n^* \xrightarrow{w^*} x^*$.

Since $y_n^* \in \partial_\varepsilon^F f(y_n)$, for each $\eta > 0$, there exists $\delta > 0$ such that

$$\langle y_n^*, h\rangle \le f(x+h) - f(x) + (\varepsilon + \eta)\|h\| \qquad \text{if } \|h\| > \delta.$$

If we suppose, let us say, that f is locally Lipschitz, we deduce that there exists some constant $M_x > 0$ such that

$$|\langle y_n^*, h\rangle \le (M_x + \varepsilon + \eta)\|h\| \qquad \text{fore all } h \in X \qquad \text{and } \eta > 0.$$

Therefore, the sequence $\{y_n^*\}_{n\in\mathbb{N}}$ is bounded and by Asplundness, the closed unit ball of the dual being weak* sequentially compact, it (or eventually a subsequence) has a limit point which by definition belongs to $\hat{\partial}_\varepsilon f(x)$. Hence, the sequence $\{z_n^*\}_{n\in\mathbb{N}}$ has a limit point z^*. Since $y_n^* + z_n^*$ tends to x^*, necessarily $z^* = x^* - y^* \in \hat{\partial}_\varepsilon g(x)$ and the proof is complete. ■

2.2. The limiting Fréchet ε-subdifferential of a marginal function

An important concept in optimization theory is that of a minimization problem which depends on parameters. Indeed, let us consider the general minimization problem:

$$p(u) := \inf f_0(x)$$

subject to

$$\begin{cases} f_i(x) \le u_i & i = 1 \cdots s, \\ f_i(x) = u_i & i = s+1, \cdots m \\ x \in C \end{cases}$$

where f_i, $i = 0, \cdots m$ are functional defined on $\mathbb{R}^n$, C is a closed set in $\mathbb{R}^n$ and the variable u is a perturbation near $\bar{u} = 0$. By setting

$$D := \{w = (w_1, \cdots w_m) \in \mathbb{R}^m \mid w_i \le 0 \text{ for } i = 1, \cdots s, w_i = 0 \text{ for } i = s+1, \cdots m\},$$

$$F(x) := (f_1(x) \cdots f_m(x))$$

and

$$\varphi(u, x) := \begin{cases} f_0(x) & \text{if } F(x) + u \in D, x \in C \\ +\infty & \text{otherwise,} \end{cases}$$

we obtain p as a marginal function, namely,

$$p(u) = \inf_{x \in M(u)} \varphi(u, x)$$

where $M : \mathbb{R}^m \rightrightarrows \mathbb{R}^m$ is a set-valued mapping defined by

$$M(u) := \{x \in \mathbb{R}^n \mid F(x) + u \in D, x \in C\}.$$

This general example leads us to consider the special case of a general parametrized constrained optimization problem:

$$(\mathcal{P}_u) \qquad p(u) = \min_{x \in C} \varphi(u, x)$$

where $\varphi : U \times X \to \mathbb{R}$ is a lower semicontinuous function defined on the product of two Asplund Banach spaces U and X, and C is a closed subset of X. It is an easy observation to see that even in a nice situation where φ is continuously differentiable and C is the whole space, the marginal function p fails to be smooth. It is the objective of what follows to establish a formula to calculate ε-subgradients of the marginal function p at u

THEOREM 2.18. Suppose U and X are Asplund spaces and C a closed subset of X. Let $\bar{u} \in U$ and $\bar{x} \in C$ such that $p(\bar{u}) = \varphi(\bar{u}, \bar{x})$. Suppose that φ is Lipschitz around $(\bar{u}, \bar{x})$ and that the following condition is satisfied.

($\mathcal{H}$) For every sequence $\{u_n\}$ such that $u_n \xrightarrow{p} \bar{u}$, there exists a sequence $\{x_n\}_{n \in \mathbb{N}}$ with limit $\bar{x}$ such that $p(u_n) = \varphi(u_n, x_n)$.

Then

$$\hat{\partial}_\varepsilon p(\bar{u}) \times \{0\} \subset \hat{\partial}_\varepsilon \varphi(\bar{u}, \bar{x}) + \varepsilon B^* \times \hat{\partial} I(\cdot; C)(\bar{x}).$$

Proof. First, we note that $(\mathcal{P}_{\bar{u}})$ is equivalent to the unconstrained minimization problem:

$$p(\bar{u}) = \min_{x \in X} (\varphi(\bar{u}, x) + I(\cdot; C)(x)).$$

Let $f(\bar{u}, x) := \varphi(\bar{u}, x) + I(x; C)$ and take $u^* \in \hat{\partial}_\varepsilon p(\bar{u})$. By definition, this amounts to saying that, there exist sequences $\{u_n\}_{n \in \mathbb{N}}$ and $\{x_n^*\}_{n \in \mathbb{N}}$ such that $u_n^* \in \partial_\varepsilon^F p(u_n)$, $u_n \xrightarrow{p} \bar{u}$, and $u_n^* \xrightarrow{w^*} u^*$. Then, by virtue of (2.5), for each $\eta > 0$ and $n \in \mathbb{N}$, there exists some $\delta_{n,\eta} > 0$ such that

$$p(u_n + h) - p(u_n) - \langle u_n^*, h \rangle \ge -(\eta + \varepsilon)\|h\| \qquad \text{for all } h \in \delta_{n,\eta} B. \tag{2.24}$$

According to assumption ($\mathcal{H}$), it is possible to select x_n such that $p(u_n) = \varphi(u_n, x_n)$ and $x_n \to \bar{x}$. Furthermore, since φ is lower semicontinuous, we can also suppose that $\varphi(u_n, x_n) \to \varphi(\bar{u}, \bar{x})$. Thus, using (2.24) we get

$$f(u_n + h, x_n + k) - f(u_n, x_n) - \langle u_n^*, h\rangle \geq -(\eta + \varepsilon)\|h\| \qquad \text{for all } h \in \delta_{n,\eta}B. \tag{2.25}$$

Hence, for all $(h, k) \in \delta_{n,\eta}B_U \times \delta_{n,\eta}B_X$ we have

$$f(u_n + h, x_n + k) - f(u_n, x_n) \geq \langle (u_n^*, 0), (h, k)\rangle - (\eta + \varepsilon)(\|h\| + \|k\|). \tag{2.26}$$

This amounts to saying that $(u_n^*, 0) \in \partial_\varepsilon^F f(u_n, x_n)$. Since $\lim_{n\to\infty} f(u_n, x_n) = f(\bar{u}, \bar{x})$, we deduce that $(x^*, 0) \in \hat{\partial}_\varepsilon f(\bar{u}, \bar{x})$. On the other hand, by Theorem 2.17 one has

$$\hat{\partial}_\varepsilon f(\bar{u}, \bar{x}) \subset \hat{\partial}_\varepsilon \varphi(\bar{u}, \bar{x}) + \varepsilon B^* \times \hat{\partial} I(\cdot; C)(\bar{x})$$

and the proof is complete. ■

Remark 2.19. By introducing the *limiting ε-normal cone* $\hat{N}_\varepsilon(C; \bar{x})$ to the set C at $\bar{x}$ as

$$\hat{N}_\varepsilon(C; \bar{x}) := \hat{\partial}_\varepsilon I(\cdot; C)(\bar{x}),$$

one may develop calculus rules for this concept as a by-product. For example, the above formula for a value function may be rewritten as:

for each $u^* \in \hat{\partial}_\varepsilon p(u)$ there exist $(x^*, z^*) \in \hat{\partial}_\varepsilon f(u, x)$ and $b^* \in B^*$ such that

$$u^* = x^* + \varepsilon b^* \qquad \text{and} \qquad -z^* \in \hat{N}_\varepsilon(C; x).$$

Let us also notice that it is possible to introduce the ε-normal cone to a set C at a point $x \in C$ as the upper limit of ε-Fréchet normal vectors near x.

2.3. More calculus

We now establish two rules for the composition of a locally Lipschitz mapping with a Fréchet differentiable function.

THEOREM 2.20. Let $F: X \to Y$ be a continuously Fréchet-differentiable mapping and let $g: Y \to \mathbb{R}$ be a locally Lipschitz function. Then, the following formula holds true:

$$\hat{\partial}_\varepsilon(g \circ F)(x) \subset \bigcup_{y^* \in \partial^\uparrow g(F(x))} [D^*F(x) + \varepsilon B^*].$$

If moreover, X and Y are Asplund then,

$$\hat{\partial}_\varepsilon(g \circ F)(x) \subset \bigcup_{y^* \in \hat{\partial}_\varepsilon g(F(x))} [D^*F(x) \circ y^* + 2\varepsilon B^*].$$

Proof. The proof of the first formula based on Rockafellar's chain rule formula ([35]), while the second formula follows the line of (Ioffe [23]) and is derived from Theorem 2.18.

Proof of the first Formula. We know from (2.4) that

$$x^* \in \partial_\varepsilon^F(g \circ F)(x) \Leftrightarrow x^* \in \partial^F(g \circ F + \varepsilon\|\cdot - x\|)(x).$$

Since,

$$\partial^F((g \circ F)(x) + \varepsilon\|\cdot - x\|)(x) \subset \partial^F(g \circ F)(x) + \varepsilon B^*$$

and

$$\partial^F(g \circ F)(x) \subset \partial^\uparrow(g \circ F)(x),$$

we obtain:

$$\partial^F_\varepsilon(g \circ F)(x) \subset \partial^\uparrow(g \circ F)(x) + \varepsilon B^*$$

and therefore

$$\hat{\partial}_\varepsilon(g \circ F)(x) \subset \partial^\uparrow(g \circ F)(x) + \varepsilon B^*.$$

It suffices to apply the well-known chain rule formula [36]:

$$\partial^\uparrow(g \circ F)(x) \circ \partial^\uparrow g(F(x))$$

to derive the proof.

Proof of the second formula. Let us define $h(u, v) := g(v) + I(u, v)$, where $I(u, v) := I(\cdot;\, \mathrm{gph}F)$ is the indicator function of the set $\mathrm{gph}F := \{(x, y) \in X \times \mathbb{R} | y = F(x)\}$ (the graph of F) and $f(u) := \inf_{v \in X} h(u, F(u)) = g(F(u))$.

We claim that the following implication holds:

$$(x^*, -y^*) \in \hat{\partial}_\varepsilon I((x, F(x)) \Rightarrow x^* \in D^*F(x)y^* + \varepsilon B^*. \tag{2.29}$$

Indeed, let define on $X \times Y$ the equivalent norm $\|\cdot\| := k \max(\|\cdot\|_X, \|\cdot\|_Y)$ where k is a Lipschitz constant of F in some neighbourhood $x + \delta B$ of x. Let $(x^*, -y^*) \in \hat{\partial}_\varepsilon I((x, F(x))$. Then, there exist sequences $\{h_n\}_{n \in \mathbb{N}}$, $\{(x_n^*, -y_n^*)\}_{n \in \mathbb{N}}$ such that $(x_n^*, -y_n^*) \in \partial^F_\varepsilon I((x + h_n, F(x + h_n))$, $\lim_{n \to +\infty} h_n = 0$ and $\lim_{n \to +\infty}(x_n^*, -y_n^*) = (x^*, -y^*)$. Thus for each $\eta > 0$ and n there exists $\delta_{n,\eta} > 0$ such that for all $h \in \delta_{n,\eta} B$, we have

$$\begin{aligned} I((x + h_n + h, F(x + h_n + h)) - I((x + h_n, F(x + h_n)) - \langle x_n^*, h\rangle \\ + \langle y_n^*, F(x + h_n + h) - F(x + h_n)\rangle \\ \geq -(\eta + \varepsilon)(k\|h\| + \|F(x + h_n + h) - F(x + h_n)\|) \end{aligned}$$

By using the Fréchet differentiability of F, we derive

$$\langle D^*F(x + h_n)y_n^*, h\rangle + \langle y_n^*, r_n(h)\|h\|\rangle + (\eta + \varepsilon)\|h\| \geq \langle x_n^*, h\rangle \qquad \text{for all } h \in \delta_{n,\eta} B \tag{2.30}$$

where $r_n(h) \to 0$ when $h \to 0$. Hence, for every $\eta > 0$, there exists $\tilde{\delta}_{\eta,n} > 0$, $\tilde{\delta}_{n,\eta} < \delta_{n,\eta}$ such that

$$\langle D^*F(x + h_n)y_n^*, h\rangle + (\eta + \varepsilon)\|h\| \geq \langle x_n^*, h\rangle \qquad \text{for all } h \in \delta_n B. \tag{2.31}$$

As in the proof of Proposition 2.8 we can deduce:

$$\langle D^*F(x)y^*, h\rangle + (\eta + \varepsilon)\|h\| \geq \langle x^*, h\rangle \qquad \forall\, h \in X$$

and therefore $x^* \in D^*F(x)y^* + \varepsilon B^*$.

Let $x^* \in \hat{\partial}_\varepsilon f(x)$. Then according to Theorem 2.18 we have

$$(x^*, 0) \in \varepsilon B^* \times \hat{\partial}_\varepsilon g(F(x)) + \hat{\partial}_\varepsilon I((x, F(x);\, \mathrm{gph}F).$$

Equivalently, there exist $(w^*, z^*) \in \hat{\partial}_\varepsilon I((x, F(x))$, $b^* \in B^*$ and $y^* \in \hat{\partial}_\varepsilon g(F(x))$ such that

$$x^* = w^* + \varepsilon b^*, \qquad 0 = y^* + z^*.$$

Then, $(x^* - \varepsilon b^*, -y^*) \in \hat{\partial}_\varepsilon I((x, F(x))$ and therefore according to (2.29) we obtain $x^* \in D^*F(x)y^* + 2\varepsilon B^*$, and the proof of the second formula is complete. ■

Remark 2.21. The two formulae in the preceding theorem are in general quite different. For example, for $g(x) := -|x|$, $x \in \mathbb{R}$, then

$$\hat{\partial}_\varepsilon g(0) = [-1 - \varepsilon, -1 + \varepsilon] \cup [1 - \varepsilon, 1 + \varepsilon] \qquad \text{and } \partial^\uparrow g(0) = [-1, 1].$$

However, in the case where g is convex continuous or continuously differentiable we have $\partial^\uparrow g(x) \subset \hat{\partial}_\varepsilon g(x)$ (see Propositions 2.3 and 2.8).

3. ε-MONOTONICITY AND ε-CONVEXITY

Let $A : X \rightrightarrows X^*$ be a set-valued map from X to x^*. as usual, we denote by $D(A)$ the *domain* of A:

$$D(A) := \{x \in X \mid A(x) \neq \varnothing\}$$

and by $G(A)$ the *graph* of A:

$$G(A) := \{(x, y) \in X \times X^* \mid y \in A(x)\}.$$

Definition 3.1. As is said to be ε-monotone if for each $x, y \in D(A)$

$$\begin{cases} x^* \in A(x) \\ y^* \in A(y) \end{cases} \Rightarrow \langle x^* - y^*, x - y\rangle \geq -2\varepsilon\|x - y\|.$$

It is said to be maximal ε-monotone, if there does not exist another ε-monotone multifunction whose graph strictly includes the graph of A.

These definitions subsume, when $\varepsilon = 0$, the conventional well-known monotonicity (respectively, maximal monotonicity) for operators.

Remark 3.2. Given a set-valued mapping $A : X \rightrightarrows X^*$, Robert [37] introduced in his thesis a related concept "*opérateur ε-augmenté*". The graph of this operator A_ε being the set of $(x^*, y^*) \in X \times X^*$ such that

$$\langle x^* - y^*, x - y\rangle \geq -\varepsilon(1 + (\|x\| + \|y\|)^2) \qquad \forall\, (y, y^*) \in G(A).$$

Definition 3.3. We say that $f : X \to \mathbb{R} \cup \{+\infty\}$ is ε-convex if it satisfies:

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y) + \varepsilon\lambda(1 - \lambda)\|x - y\|$$

for each $x, y \in X$ and $\lambda \in [0, 1]$.

Remark 3.4. 1. Convexity subsumes ε-convexity for each $\varepsilon \geq 0$.

2. Note that when f is differentiable, we can give equivalent definitions of ε-convexity:

(a) for all $x, y \in X$,

$$f(y) \geq f(x) + \langle y - x, \nabla f(x)\rangle + \varepsilon\|x - y\|$$

or for all $x, y \in X$,

$$\langle y - x, \nabla f(y) - \nabla f(x)\rangle \geq -2\varepsilon\|x - y\|.$$

The last inequality amounts to saying that for a differentiable function f, the ε-convexity of f coincides with the ε-monotonicity of its gradient.

LEMMA 3.5. Let f be ε-convex. If $x^* \in \partial_\varepsilon^F f(x)$ then

$$f(x + h) \geq f(x) + \langle x^*, h\rangle - 2\varepsilon\|h\| \qquad \text{for all } h \in X.$$

Proof. Fix $x, x^* \in \partial_\varepsilon^F f(x)$ and $y \in X$. From the ε-convexity and the fact that $x^* \in \partial_\varepsilon^F f(x)$ we have respectively,

$$f(x + \lambda(y - x)) \leq (1 - \lambda)f(x) + \lambda f(y) + \lambda(1 - \lambda)\varepsilon\|x - y\|, \qquad \text{for all } \lambda \in [0, 1],$$

and for some $\delta > 0$

$$f(x + \lambda(y - x)) \geq f(x) + \lambda\langle x^*, y - x\rangle - \varepsilon\lambda\|x - y\| \qquad \text{for all } \lambda \in [0, \delta].$$

Hence,

$$f(y) \geq f(x) + \langle x^*, y - x\rangle - 2\varepsilon\|x - y\| - \lambda\varepsilon\|x - y\| \qquad \text{for all } \lambda \in [0, \delta].$$

This yields, $f(y) \geq f(x) + \langle x^*, y - x\rangle - 2\varepsilon\|x - y\|$. ∎

PROPOSITION 3.6. Let $\varepsilon \geq 0$ be given and $f: X \to \mathbb{R} \cup \{+\infty\}$ be ε-convex. Then $\hat{\partial}_\varepsilon f$ is 2ε-monotone.

Proof. Let $x, y \in X, x^* \in \hat{\partial}_\varepsilon f(x)$ and $y^* \in \hat{\partial}_\varepsilon f(y)$ be given. Then there exist sequences $\{x_n\}_{n\in\mathbb{N}}$, $\{y_n\}_{n\in\mathbb{N}}$, $\{x_n^*\}_{n\in\mathbb{N}}$ and $\{y_n^*\}_{n\in\mathbb{N}}$ such that

1. $\lim_{n\to\infty} x_n = x$, $\lim_{n\to\infty} y_n = y$, $w^* - \lim_{n\to\infty} x_n^* = x^*$, $w^* - \lim_{n\to\infty} y_n^* = y^*$;
2. $x_n^* \in \partial_\varepsilon^F f(x_n)$ and $y_n^* \in \partial_\varepsilon^F f(y_n)$.

From Lemma 3.5 we have

$$f(x_n + h) \geq f(x_n) + \langle x_n^*, h\rangle - 2\varepsilon\|h\|, \qquad \text{for all } h \in X \tag{3.32}$$

and

$$f(y_n + h) \geq f(y_n) + \langle y_n^*, h\rangle - 2\varepsilon\|h\|, \qquad \text{for all } h \in X \tag{3.33}$$

If we put $h := y_n - x_n$ in (3.32) and $h := x_n - y_n$ in (3.33), we deduce from these inequalities:

$$\langle x_n^* - y_n^*, x_n - y_n\rangle \geq -4\varepsilon\|x_n - y_n\|.$$

Hence, $\langle x^* - y^*, x - y\rangle \geq -4\varepsilon\|x - y\|$. ∎

PROPOSITION 3.7. Let X be a Banach space with a Fréchet smooth renorm. Let $f: X \to \mathbb{R}$ be locally Lipschitz and $\varepsilon \geq 0$ be given. If $\partial_\varepsilon f$ is ε-monotone, then f is 2ε-convex.

Proof. Suppose by contradiction that $\hat{\partial}_\varepsilon f$ is ε-monotone and f is not 2ε-convex. Then, there exist $a, b, c := \lambda a + (1 - \lambda)b$ in X with $\lambda \in]0, 1[$ such that

$$f(c) > \lambda f(a) + (1 - \lambda)f(b) + 2\varepsilon\lambda(1 - \lambda)\|a - b\|. \tag{3.34}$$

Noticing that

$$\tilde{\partial} f(c) := \{x^* \in X^* \mid \exists \text{ sequences } x_n \to x, x_n^* \xrightarrow{w^*} x^*, \text{ with } x_n^* \in \partial^F f(x_n)\ \forall\, n \in \mathbb{N}\}$$

is a subset of $\hat{\partial}_\varepsilon f(x)$, the Borwein and Preiss Mean Value Theorem [38], gives the existence of

$$x \in [a, c], y \in [c, b], x^* \in \hat{\partial}_\varepsilon f(x), y^* \in \hat{\partial}_\varepsilon f(y)$$

such that,

$$\langle x^*, c - a\rangle \geq f(c) - f(a) \tag{3.35}$$

$$\langle y^*, c - b\rangle \geq f(c) - f(b) \tag{3.36}$$

From the formula

$$\langle x^* - y^*, x - y\rangle = \|x - y\| \left\{ \frac{\langle x^*, a - c\rangle}{\|a - c\|} - \frac{\langle y^*, c - b\rangle}{\|c - b\|} \right\},$$

using (3.35) and (3.36) we obtain:

$$\langle x^* - y^*, x - y\rangle \leq \|x - y\| \left\{ \frac{f(a) - f(c)}{\|a - c\|} - \frac{f(c) - f(b)}{\|b - c\|} \right\}. \tag{3.37}$$

Noticing that $\|a - c\| = (1 - \lambda)\|a - b\|$ and $\|b - c\| = \lambda\|a - b\|$ and using (3.37), we obtain

$$\langle x^* - y^*, x - y\rangle \leq \frac{\|x - y\|}{\lambda(1 - \lambda)\|a - b\|}(\lambda f(a) + (1 - \lambda)f(b) - f(c)).$$

Using (3.34), the last inequality yields

$$\langle x^* - y^*, x - y\rangle < -2\varepsilon\|x - y\|. \tag{3.38}$$

which is a contradiction. ■

An immediate consequence of Proposition 3.7 is that if $f: X \to \mathbb{R} \cup \{+\infty\}$ is convex, then $\hat{\partial}_\varepsilon \varphi$ is 2ε-monotone for every $\varepsilon > 0$. Actually a stronger result is true.

PROPOSITION 3.8. Let $f: X \to \mathbb{R} \cup \{+\infty\}$ be convex and lower semicontinuous. Then $\hat{\partial}_\varepsilon f$ is ε-monotone for each $\varepsilon \geq 0$.

Proof. Using the technique of Lemma 3.1 and the convexity, one can show that for $x^* \in \partial_\varepsilon^F f(x)$ one has

$$f(x + h) \geq f(x) + \langle x^*, h\rangle - \varepsilon\|h\| \qquad \text{for all } h \in X.$$

Now observe that formulae (3.10) and (3.11) are still valid by using ε instead of 2ε. Hence, $\langle x^* - y^*, x - y\rangle \geq -2\varepsilon\|x - y\|$ for every $x, y \in X, x^* \in \hat{\partial}_\varepsilon f(y)$. The ε-monotonicity is therefore established. ■

Note that under the hypothesis of Proposition 2.3, if f is convex, then one has $\hat{\partial}_\varepsilon f(x) = \partial^{\mathrm{Fen}} f(x) + \varepsilon B^*$. According to this remark, the preceding proposition can be also derived from the coming up Proposition 3.11 and from the classical result on monotone subdifferential of convex functions.

COROLLARY 3.9. If f is locally Lipschitz, the set-valued mappings $\partial_\varepsilon^F f$ and $\hat{\partial}_\varepsilon f$ are never monotone.

Proof. If $\hat{\partial}_\varepsilon f$ is monotone, it is ε'-monotone for each $\varepsilon' \geq 0$. By Proposition 3.7, the function is $2\varepsilon'$-convex. Since ε' is arbitrary, f must be convex. Thus from Proposition 2.3 we deduce that $\partial^{\mathrm{Fen}} f$ is graphically strictly contained in $\hat{\partial}_\varepsilon f$. This is impossible since $\partial^{\mathrm{Fen}} f$ is a maximal monotone operator. Now, if $\partial_\varepsilon^F f$ is monotone, then $\hat{\partial}_\varepsilon f$ is monotone too, which again is impossible. ■

The fact that the Fenchel subdifferential of a convex lower semicontinuous proper function is a maximal monotone operator is due to Rockafellar ([16]). It is a natural question to ask whether the Rockafellar Theorem extends to ε-subdifferentials, if instead of maximal mononicity we use the concept of ε-monotonicity. The next result is in this direction.

PROPOSITION 3.10. Let $f: X \to \mathbb{R} \cup \{+\infty\}$ be convex lower semicontinuous functional. Then $\hat{\partial}_\varepsilon f$ is maximal ε-monotone for every $\varepsilon \geq 0$.

Proof. Let $x^* \notin \hat{\partial}_\varepsilon f(x)$. The aim is to show that there exist $y \in X$, $Y^* \in \hat{\partial}_\varepsilon f(y)$ such that

$$\langle x^* - y^*, x - y \rangle < 2\varepsilon \|x - y\|. \tag{3.39}$$

We prove first that x cannot be a local $\varepsilon\|\cdot\|$-minimum of the function $f - x^*$. Indeed, if this is not true, by Proposition 2.15, $0 \in \partial_\varepsilon^F (f - x^*)(x)$. Since f is convex, one can express $\partial_\varepsilon^F (f - x^*)(x)$ as

$$\{y^* \in X^* \mid (f - x^*)(x + h) - (f - x^*)(x) - \langle y^*, h \rangle > -\varepsilon\|h\| \text{ for every } h \in X\},$$

which is identical with the set

$$\{y^* \in X^* \mid f(x + h) - f(x) - \langle y^* + x^*, h \rangle > -\varepsilon\|h\| \text{ for every } h \in X\}.$$

Consequently,

$$y^* \in \partial_\varepsilon^F (f - x^*)(x) \qquad \text{if and only if } y^* + x^* \in \partial_\varepsilon^F f(x).$$

By this $0 \in \partial_\varepsilon^F (f - x^*)(x)$ implies that $x^* \in \partial_\varepsilon^F f(x) \subseteq \hat{\partial}_\varepsilon f(x)$, a contradiction.

We have shown that x is not a $\varepsilon\|\cdot\|$-minimum of the function $f - x^*$, hence there exists $h \neq 0$ such that

$$(f - x^*)(x + h) - (f - x^*)(x) < -\varepsilon\|h\|.$$

Define a function g on X by

$$g(z) := (f - x^*)(z) + \varepsilon\|z - x\|.$$

This function is convex and has the property that $g(x) > g(x + h)$. In view of Lemma 3.5 [15], there is a point $y \neq x$ in a neighbourhood of the interval $[x + h, x]$ and $y_0^* \in \partial^{\text{Fen}} g(y)$ such that

$$\langle y_0^*, x - y \rangle > 0. \tag{3.40}$$

Since f and g are convex, one can find $y_1^* \in \partial^{\text{Fen}} f(y)$ such that

$$\langle y_0^*, x - y \rangle \in \langle y_1^* - x^*, x - y \rangle + \varepsilon\{\langle v^*, x - y \rangle \mid v^* \in \partial^{\text{Fen}} d_x(y)\} \tag{3.41}$$

where, d_x is the distance function to x. Since $x \neq y$, one has

$$\langle v^*, x - y \rangle = -\|x - y\|.$$

Consequently,

$$\langle y_0^*, x - y \rangle = \langle y_1^* - x^*, x - y \rangle - \varepsilon\|x - y\|. \tag{3.42}$$

Furthermore, we know that $\partial^{\text{Fen}} f(y) + \varepsilon B^* \subseteq \hat{\partial}_\varepsilon f(y)$. Hence, by choosing $b^* \in B^*$ with the property that $\langle b^*, x - y \rangle = -\|x - y\|$ and setting $y^* = y_1^* + \varepsilon b^* \in \hat{\partial}_\varepsilon f(y)$ we obtain

$$\langle y^* - x^*, x - y \rangle - 2\varepsilon\|x - y\| > 0,$$

according to (3.40) and (3.42). This proves (3.39), and establishes the maximal ε-monotonicity. ■

PROPOSITION 3.11. If $A: X \rightrightarrows X^*$ is monotone, then $A + \varepsilon B^*$ is ε-monotone.

Proof. Let $x^* \in A(x) + \varepsilon B^*$ and $y^* \in A(y) + \varepsilon B^*$. Hence $x^* = \bar{x} + \varepsilon b^*$ and $y^* = \bar{y} + \varepsilon c^*$ with $\bar{x} \in A(x)$, $\bar{y} \in A(y)$, $b^* \in B^*$ and $c^* \in B^*$. Then by hypothesis,

$$\begin{aligned}\langle x^* - y^*, x - y\rangle &= \langle \bar{x} - \bar{y}, x - y\rangle + \varepsilon\langle b^* - c^*, x - y\rangle \\ &\geq \varepsilon\langle b^* - c^*, x - y\rangle \\ &\geq -2\varepsilon\|x - y\|,\end{aligned}$$

which completes the proof. ■

We recall that $A: X \rightrightarrows X^*$ is declared *cyclically monotone* if for any subset $\{x_1, x_2, \dots, x_n\} \subseteq D(A)$ such that $u_i \in A(x_i)$, $i \in \{0 \dots n\}$ we have

$$\sum_{i=0}^{i=n} \langle u_i, x_i - x_{i+1}\rangle \leq 0,$$

where $x_{n+1} = x_0$.

A is called *maximal cyclically monotone* if it has no cyclically monotone extension.

This class of operators includes for instance, monotone graphs in $\mathbb{R}^2$, self-adjoint linear unbounded operators with a dense domain.

COROLLARY 3.12. Let $A: X \rightrightarrows X^*$ be a maximal cyclically monotone operator. Then, $T_\varepsilon: X \rightrightarrows X^*$ given by

$$T_\varepsilon(x) := A(x) + \varepsilon B^*$$

is a maximal ε-monotone operator.

Proof. It results from the Rockafellar Theorem [16] that A is the subdifferential of a convex proper lower semicontinuous function $f: X \to \mathbb{R} \cup \{+\infty\}$. Apply Proposition 3.10 and Proposition 2.3 to conclude. ■

For sake of completness, we conclude the paper by recalling a general result obtained by Oettli and Théra [21].

THEOREM 3.13. [Oettli and Théra] Let $T: X \rightrightarrows X^*$ be monotone, upper semicontinuous along lines, and have compact, convex, nonempty values. Then $T + \varepsilon B^*$ is maximal 2ε-monotone.

Acknowledgements—The authors are grateful to Professors Boris Mordukhovich, Jean-Paul Penot and Lionel Thibault for fruitful conversations on this material.

REFERENCES

1. Hiriart-Urruty J.-B. and Lemaréchal C., *Convex Analysis and Minimization Algorithms II*. Springer Verlag, 1993.
2. Ekeland I. and Lebourg G., Sous-gradients approchés et applications. *C. R. Acad. Sci., Paris*, 1975, **281**, 219–222.
3. Goldstein A., Optimization of Lipschitz continuous functions. *Math. Progr.*, 1977, **13**, 14–22.
4. Dixon L. C. W., Reflections on nondifferentiable optimization, Part 1. Ball gradient. *J.O.T.A.*, 1980, **32**(2), 165–174.
5. Polak E., Mayne D. Q. and Wardi Y., On the extension of constrained optimization algorithms from differentiable to nondifferentiable problems. *Siam J. Control Opt.*, 1983, **21**, 179–203.

6. Bihain A., Optimization of upper semidifferentiable functions. *J.O.T.A.*, 1984, **44**, 545–568.
7. Bustos, ε-gradients pour les fonctions localement lipschitziennes et applications. *Numerical Functional Anal. and Opt.*, 1994, **15**, 435-453.
8. Kruger A. Ya. and Mordukhovich B., Extremal points and the Euler equation in nonsmooth optimization. *Dokl. Akad. Nauk BSSR*, 1980, **24**, 684–687.
9. Kruger A. Ya. and Mordukhovich B., Generalized normals and derivatives, and necessary optimality conditions in nondifferentiable programming, part I Depon. VINITI, No 408–80; part II: Depon VINITI, No 494–80, Moscow, 1980 (in Russian)
10. Mordukhovich B. S. and Shao Y., Extremal characterization of Asplund spaces. *Proc. Amer. Math. Soc.*, 1996, **124**(1), 197–205.
11. Mordukhovich B. S. and Shao Y., Nonsmooth sequential analysis in Asplund spaces. *Trans. Amer. Math. Soc.*, 1996, **348**, 1235–1280.
12. Poliquin R. A., Subgradient monotonicity and convex functions. *Nonlinear Anal.*, 1990, **14**, 306–317.
13. Correa R., Jofré A. and Thibault L., Characterization of lower semicontinuous convex functions. *Proc. Amer. Math. Soc.*, 1992, **16**(1), 67–72.
14. Correa R., Jofré A. and Thibault L., Subdifferential monotonicity as characterization of convex functions. *Numerical Funct. Anal. and Optimization*, 1994, **15**, 531–535.
15. Luc D. T., On the monotonicity of subdifferentials. *Acta Math. Vietnamica*, 1993, **18**, 99-106.
16. Rockafellar R. T., On the maximal monotonicity of the subdifferential mapping. *Pacific J. Math.*, 1970, **33**, 209–216.
17. Zagrodny D., Approximate mean value theorem for upper subderivatives. *Nonlinear Anal.*, 1988, **12**, 1413–1428.
18. Clarke F. H., Stern R. J. and Wolenski P. R., *Subgradient criteria for monotonicity, the Lipschitz condition and convexity. Canad. J. Math.*, 1993, **45**(6), 1167–1183.
19. Aussel D., Corvellec J.-N. and Lassonde M., Mean Value property and subdifferential criteria for lower semicontinuous functions. *Trans. Amer. Math. Soc.*, 1995, **344**(10), 4147–4161.
20. Correa R., Jofré A. and Thibault L., Subdifferential characterization of convexity. *Recent advances in nonsmooth optimization*, ed. D.-Z. Du, L. Qi and R. S. Wormersley, World Scientific, 1995, pp. 18–23.
21. Oettli W. and Théra M., On maximal monotonicity of perturbed mappings. *Bolletino U.M.I.*, 1995, **7**(9-A), 47–55.
22. Treiman J. S., Clarke's gradients and ε-subgradients in Banach spaces. *Trans. Amer. Math. Soc.*, 1984, **294**, 65–78.
23. Ioffe A., Approximate subdifferentials and applications, Part 2. *Mathematica*, 1986, **33**, 111–128.
24. Ioffe A., Approximate subdifferentials and applications, Part 3. *Mathematica*, 1988, **36**, 1–38.
25. Ioffe A., Subdifferentiability spaces and nonsmooth analysis. *Bull. Am. Math. Soc.*, 1984, **10**(1), 87–90.
26. Kruger A. Ya., Properties of generalized differentials. *Siberian Math. Journal*, 1983, 822–832.
27. Deville R., Godefroy G. and Zisler V., Smoothness and renorming in Banach spaces. *Pitman Monographs and Surveys in Pure and Applied Math.*, Vol. 64. Longman, London, 1993.
28. Stegall C., The Radon-Nikodyn property in conjugate Banach spaces II. *Trans. Amer. Math. Soc.*, 1984, **264**, 507–519.
29. Mordukhovich B. S., Maximum principle in problems of time optimal control with nonsmooth constraints. *J. Appl. Math. Mech.*, 1976, **40**. 960–969.
30. Mordukhovich B. S., Approximation Methods in Problems of Optimization and Control. Nauka, Moscow, 1988.
31. Aubin J.-P. and Cellina A. *Differential Inclusions*. Springer Verlag, 1984.
32. Ioffe A., Approximate subdifferential and applications. Part 1. *T.A.M.S.*, 1984, **281**(1), 389–416.
33. Borwein J.-M. and Fitzpatrick S., Weak* sequential compactness and bornological limit derivatives. *J. Convex Analysis: Special Issue in Celebration of R. T. Rockafellar's 60th Birthday*, 1995, Part I, 259–268.
34. Fabian M., On classes of subdifferentiability spaces of Ioffe. *Nonlinear Anal.*, 1988, **12**, 63–74.
35. Rockafellar R. T., Generalized directional derivatives and subgradients of nonconvex functions. *Can. J. Math.*, 1980, **32**, 257–280.
36. Clarke F. H., *Optimization and Nonsmooth Analysis*. Wiley, New York, 1983.
37. Robert R., Contributions à l'analyse non linéaire. Thèse d'Etat, Grenoble, 1976.
38. Borwein J.-M. and Preiss D., A smooth variational principle. *Trans. Amer. Math. Soc.*, 1987, **303**(2), 513–527.