Subdifferential characterization of approximate convexity: the lower semicontinuous case*

Aris Daniilidis†, **Florence Jules**‡ **and Marc Lassonde**§

Dedicated to Professor A. Auslender for the occasion of his Honoris Causa

Abstract It is known that a locally Lipschitz function f is approximately convex if, and only if, its Clarke subdifferential $\partial_C f$ is a submonotone operator. The main object of this work is to extend the above characterization to the class of lower semicontinuous functions. To this end, we establish a new approximate mean value inequality involving three points. We also show that an analogue of the Rockafellar maximal monotonicity theorem holds for this class of functions and we discuss the case of arbitrary subdifferentials.

Key words Approximate convexity, submonotone operator, subdifferential, mean value inequality.

1991 AMS Subject Classification *Primary* 26B25; 47H05; *Secondary* 49J52.

1 Introduction

The class of convex functions is a cornerstone of variational analysis and the origin of the so-called subdifferential theory. Convex functions have far-reaching consequences in the study of optimization problems, and have been used in both theoretical and practical purposes. Generalizations of the concept of convexity have been proposed in the literature to serve various objectives. A natural generalization of convexity, based on a first order relaxation in a local sense, has been studied in a recent work of Ngai-Luc-Théra [14], under the name of approximate convexity (see exact definition in Section 3). In [14] it is shown that, as is the case for convex functions, lower semicontinuous approximate convex functions are locally Lipschitz in the interior of their domain. The interest for this class has been grown up when it was established (see [7, Theorem 2]) that a locally Lipschitz

*The research of the first author has been supported by the Grant MTM2005-08572-C03-03 (Spain).

†Departament de Matemàtiques, Universitat Autònoma de Barcelona, E-08193 Bellaterra (Cerdanyola del Vallès), Spain, e-mail: arisd@mat.uab.es; http://mat.uab.es/~arisd

‡Université des Antilles et de la Guyane, Campus de Fouillole, 97159 Pointe-à-Pitre, Guadeloupe, France, e-mail: florence.jules@univ-ag.fr

§Université des Antilles et de la Guyane, Campus de Fouillole, 97159 Pointe-à-Pitre, Guadeloupe, France, e-mail: marc.lassonde@univ-ag.fr

function is approximately convex if, and only if, its Clarke subdifferential is submonotone (notion corresponding to a first order relaxation of monotonicity, see exact definition in Section 4). An interesting consequence of this result is that, in finite dimensional spaces, a locally Lipschitz function is approximately convex if, and only if, it is lower-C^1 ([19]), that is, locally representable as a maximum of a compactly indexed family of C^1 functions (see [7, Corollary 3]).

Even if the aforementioned characterization is expected to hold in the general lower semicontinuous case, the techniques employed in [7] are heavily based on the Lebourg mean value theorem (valid only for locally Lipschitz functions) and are not directly extendable to the use of an approximate mean value theorem in the sense of [21] or [1]. In this work, we overcome this difficulty by establishing a new approximate mean value theorem for lower semicontinuous functions, involving simultaneously three points of the segment (see Theorem 1). This turns out to be a powerful tool: not only it allows us to extend the known characterizations of approximate convexity ([7, Theorem 2 and Theorem 13]) from the locally Lipschitz to the lower semicontinuous case (see Theorem 5), but it also provides for free the analogous and well-known characterization of convexity by means of the monotonicity of its Clarke subdifferential, previously established in [5] (for reflexive Banach spaces) and [6] (for the general case). The case of arbitrary subdifferentials is briefly discussed in the last section.

2 A three-points approximate mean value inequality

Throughout, X stands for a real Banach space with norm $\|\cdot\|$, S_X for its unit sphere, X^* for its topological dual, and $\langle\cdot,\cdot\rangle$ for the duality pairing. For $x, y, z \in X$ we set

$$d_{[x,y]}(z) := \min_{\bar{z}\in[x,y]} \|\bar{z} - z\|,$$

where $[x, y] := \{\, \bar{z} \in X \mid \bar{z} = \lambda x + (1-\lambda)y \;\text{ for some }\; \lambda \in [0, 1] \,\}$ is the closed segment joining x and y. The semi-open segments $[x, y)$, $(x, y]$ and the open segment (x, y) are defined analogously. For $\lambda > 0$ and $x, y \in X$, we denote the closed uniform λ-neighborhood of $[x, y]$ by

$$B([x, y], \lambda) := \{\, z \in X \mid d_{[x,y]}(z) \le \lambda \,\}.$$

All the functions $f : X \to \mathbb{R}\cup\{+\infty\}$ considered in this paper are lower semicontinuous. As usual, we set $\operatorname{dom} f := \{x \in X \mid f(x) < \infty\}$ and, for $x \in \operatorname{dom} f$, we write $x' \to_f x$ to express that $x' \to x$ and $f(x') \to f(x)$. The *Clarke subdifferential* of f is the set-valued operator $\partial_C f : X \rightrightarrows X^*$ defined for $x \notin \operatorname{dom} f$ by $\partial_C f(x) := \emptyset$ and for $x \in \operatorname{dom} f$ by

$$\partial_C f(x) := \{x^* \in X^* \mid \langle x^*, h\rangle \le f'_{CR}(x; h),\ \forall h \in X\}, \tag{1}$$

where $f'_{CR}(x; h)$ denotes the *Clarke-Rockafellar subderivative* (see [17]):

$$f'_{CR}(x; h) := \sup_{\lambda>0} \limsup_{\substack{t\searrow 0\\ x'\to_f x}} \inf_{h'\in B(h,\lambda)} \frac{f(x' + th') - f(x')}{t}. \tag{2}$$

Given a set-valued operator $T : X \rightrightarrows X^*$, we write $\operatorname{Dom} T := \{x \in X \mid T(x) \neq \emptyset\}$ for its *domain* and, for the sake of simplicity, we also write T for its *graph*, that is,

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

The following property is a special case of the exact sum rule for the Clarke subdifferential (see, e.g., [17, Theorem 2] or [3, Corollary 1 of Theorem 2.9.8]):

Optimality Property: *If $\bar{x} \in X$ is a finite local minimum of $f + \varphi$, where $f : X \to \mathbb{R} \cup \{+\infty\}$ is lower semicontinuous and $\varphi : X \to \mathbb{R}$ is convex Lipschitz, then $0 \in \partial_C f(\bar{x}) + \partial_C \varphi(\bar{x})$.*

Combining this property with the Ekeland variational principle [9], one easily obtains (see, e.g., [17, Theorem 7]):

Subdifferential Variational Principle: *Let $f : X \to \mathbb{R} \cup \{+\infty\}$ be lower semicontinuous, $\varphi : X \to \mathbb{R}$ be convex Lipschitz, $\lambda > 0$ and $\sigma > 0$. If $\bar{x} \in X$ verifies*

$$(f + \varphi)(\bar{x}) < \inf_{B(\bar{x}, \lambda)} (f + \varphi) + \lambda\sigma,$$

then there exist $x \in X$, $x^* \in \partial_C f(x)$ and $y^* \in \partial_C \varphi(x)$ such that

$$\|x - \bar{x}\| < \lambda, \quad (f + \varphi)(x) \leq (f + \varphi)(\bar{x}), \quad \|x^* + y^*\| < \sigma.$$

These tools are crucial for proving the Zagrodny inequality [21], which can be formulated as follows (see, e.g., [1]):

Two Points Approximate Mean Value Inequality: *Let $f : X \to \mathbb{R} \cup \{+\infty\}$ be lower semicontinuous. Let $x, z \in X$ with $x \in \operatorname{dom} f$ and $x \neq z$, and let $r \in \mathbb{R}$ such that $r \leq f(z)$. Then, there exist $\bar{x} \in [x, z)$, and sequences $\{(x_n, x_n^*)\}_n \subset \partial_C f$ with $x_n \to_f \bar{x}$ such that*

$$\liminf_n \langle x_n^*, \frac{z - x_n}{\|z - x_n\|} \rangle \geq \frac{r - f(x)}{\|z - x\|}.$$

This inequality has proved to be a key tool in nonsmooth analysis and has been used, e.g., for establishing subdifferential criteria for various properties of functions such as Lipschitzness, monotonicity, convexity, integration, etc. (see, e.g., [1, 2, 4, 5, 6, 20] and the references therein). Here, we propose an approximate mean value inequality involving three points on a segment, which turns out to be well suited for establishing subdifferential criteria for approximate convexity (see Section 4).

Theorem 1 (A three points approximate mean value inequality) *Let $f : X \to \mathbb{R} \cup \{+\infty\}$ be lower semicontinuous. Let $x, y \in \operatorname{dom} f$ with $x \neq y$, and let $z \in (x, y)$ and $r \in \mathbb{R}$ such that $r \leq f(z)$. Then, there exist $\bar{x} \in [x, z)$, $\bar{y} \in (z, y]$ and sequences $\{(x_n, x_n^*)\}_n \subset \partial_C f$ with $x_n \to_f \bar{x}$, and $\{(y_n, y_n^*)\}_n \subset \partial_C f$ with $y_n \to_f \bar{y}$, such that*

$$\liminf_n \langle x_n^* - y_n^*, \frac{y_n - x_n}{\|y_n - x_n\|} \rangle \geq \frac{r - f(x)}{\|z - x\|} + \frac{r - f(y)}{\|z - y\|}. \tag{3}$$

Proof. We follow the lines of the proof of the Zagrodny approximate mean value inequality as given in [1, Theorem 4.2]. Let $p^* \in X^*$ be such that $\langle p^*, z - x\rangle = r - f(x)$. Consider the function $h := f - p^*$. Since h is lower semicontinuous and $h(x) = f(x) - p^*(x) = r - p^*(z) \leq h(z)$, there exists $\bar{x} \in [x, z)$ such that $h(\bar{x}) = \min_{[x,z]} h$. For any $n \in \mathbb{N}$ so large that h is bounded from below on $B(\bar{x}, 1/n)$ and $2/n \leq \|z - \bar{x}\|$, let $\rho_n > 0$ be such that

$$h(\bar{x}) - 1/n^2 < \inf_{B([x,z],\rho_n)} h,$$

and let then $K_n > 0$ be such that

$$h(\bar{x}) - 1/n^2 < \inf_{B(\bar{x},1/n)} h + K_n\rho_n.$$

It readily follows from the above inequalities that

$$h(\bar{x}) = (h + K_n d_{[x,z]})(\bar{x}) < \inf_{B(\bar{x},1/n)} (h + K_n d_{[x,z]}) + 1/n^2.$$

Applying the above Subdifferential Variational Principle, we get sequences $\{x_n\} \subset X$, $x_n^* \in \partial_C f(x_n)$ and $\delta_n^* \in \partial_C(K_n d_{[x,z]})(x_n)$ such that

$$\begin{cases} \|\bar{x} - x_n\| < 1/n, \\ h(x_n) \leq h(x_n) + K_n d_{[x,z]}(x_n) \leq h(\bar{x}), \\ \|x_n^* - p^* + \delta_n^*\| < 1/n. \end{cases} \tag{4}$$

Note that since $2/n \leq \|z - \bar{x}\|$, the functions $d_{[x,z]}$ and $d_{[x,y]}$ have the same restrictions on $B(\bar{x}, 1/n)$; thus we have $\delta_n^* \in \partial_C(K_n d_{[x,y]})(x_n)$.

Similarly, taking $q^* \in X^*$ so that $\langle q^*, z - y\rangle = r - f(y)$ and working as above, we obtain a point $\bar{y} \in (z, y]$ and sequences $\{y_n\} \subset X$, $y_n^* \in \partial_C f(y_n)$ and $\xi_n^* \in \partial_C(K_n' d_{[z,y]})(y_n)$, with $K_n' > 0$, such that

$$\begin{cases} \|\bar{y} - y_n\| < 1/n, \\ h(y_n) \leq h(y_n) + K_n' d_{[z,y]}(y_n) \leq h(\bar{y}), \\ \|y_n^* - q^* + \xi_n^*\| < 1/n. \end{cases} \tag{5}$$

Clearly, we can manage so that $K_n' = K_n$ and $\xi_n^* \in \partial_C(K_n d_{[x,y]})(y_n)$.

It follows from the first two inequalities in (4) and (5) and the lower semicontinuity of f that $x_n \to_f \bar{x}$ and $y_n \to_f \bar{y}$. Now, put $u_n := (y_n - x_n)/\|x_n - y_n\|$. From the third inequality in (4) and (5) we derive that

$$\begin{cases} \langle x_n^*, u_n\rangle - \langle p^*, u_n\rangle + \langle \delta_n^*, u_n\rangle > -1/n, \\ \langle y_n^*, -u_n\rangle - \langle q^*, -u_n\rangle + \langle \xi_n^*, -u_n\rangle > -1/n, \end{cases}$$

hence, adding these inequalities, we get

$$\langle x_n^* - y_n^*, u_n\rangle + \langle \delta_n^* - \xi_n^*, u_n\rangle > \langle p^* - q^*, u_n\rangle - 2/n. \tag{6}$$

Since the pairs (x_n, δ_n^*) and (y_n, ξ_n^*) belong to the monotone operator $\partial_C(K_n d_{[x,y]})$ (recall that the function $u \mapsto K_n d_{[x,y]}(u)$ is convex), we have $\langle \delta_n^* - \xi_n^*, u_n \rangle \leq 0$, so it follows from (6) that

$$\langle x_n^* - y_n^*, u_n \rangle > \langle p^* - q^*, u_n \rangle - 2/n. \tag{7}$$

It remains to observe that the sequence $\{u_n\}$ converges to

$$\frac{\bar{y} - \bar{x}}{\|\bar{x} - \bar{y}\|} = \frac{z - x}{\|z - x\|} = -\frac{z - y}{\|z - y\|},$$

hence passing to the limit in (7) we obtain

$$\begin{aligned} \liminf_n \langle x_n^* - y_n^*, u_n \rangle &\geq \lim_n \langle p^* - q^*, u_n \rangle \\ &= \langle p^*, \frac{z - x}{\|z - x\|} \rangle + \langle q^*, \frac{z - y}{\|z - y\|} \rangle \\ &= \frac{r - f(x)}{\|z - x\|} + \frac{r - f(y)}{\|z - y\|}. \end{aligned}$$

This completes the proof. ■

3 Subdifferential regularity of approximately convex functions

We recall from [14] that a function $f : X \to \mathbb{R} \cup \{+\infty\}$ is said to be *approximately convex at* $x_0 \in X$, if for every $\varepsilon > 0$ there exists $\delta > 0$ (depending on x_0 and ε) such that for all $x, y \in B(x_0, \delta)$ and all $t \in (0, 1)$

$$f(tx + (1 - t)y) \leq tf(x) + (1 - t)f(y) + \varepsilon t(1 - t)\|x - y\|. \tag{8}$$

A slightly more general notion is considered in [7]. This notion, which appears to be a natural extension in infinite dimensions for the notion of lower-C^1 functions, is obtained by imposing an additional directional constraint for (8):

Definition 2 (Directional approximate convexity) A function $f : X \to \mathbb{R} \cup \{+\infty\}$ is called *directionally approximately convex at* $x_0 \in X$, if for every $d \in S_X$ and $\varepsilon > 0$ there exists $\delta > 0$ (depending on d, x_0 and ε) such that for all $x, y \in B(x_0, \delta)$, with $x \neq y$ and $(x - y)/\|x - y\| \in B(d, \delta)$, and all $t \in (0, 1)$

$$f(tx + (1 - t)y) \leq tf(x) + (1 - t)f(y) + \varepsilon t(1 - t)\|x - y\|. \tag{9}$$

Equivalently, $f : X \to \mathbb{R} \cup \{+\infty\}$ is *directionally approximately convex at* $x_0 \in X$, if for every $d \in S_X$ and $\varepsilon > 0$ there exists $\delta > 0$ such that for all $x, y \in B(x_0, \delta) \cap \operatorname{dom} f$, with $x \neq y$ and $(x - y)/\|x - y\| \in B(d, \delta)$, and all $z \in (x, y)$ one has

$$\frac{f(z) - f(x)}{\|z - x\|} + \frac{f(z) - f(y)}{\|z - y\|} \leq \varepsilon. \tag{10}$$

We recall that the *Hadamard subdifferential* and the *Fréchet subdifferential* of a function $f : X \to \mathbb{R} \cup \{+\infty\}$ at point $x \in \operatorname{dom} f$ are respectively defined by

$$\partial_H f(x) := \{x^* \in X^* \mid \langle x^*, h\rangle \leq \liminf_{\substack{t \searrow 0 \\ h' \to h}} \frac{f(x+th') - f(x)}{t}, \forall h \in X\}, \tag{11}$$

$$\partial_F f(x) := \{x^* \in X^* \mid \liminf_{\|h\| \to 0} \frac{f(x+h) - f(x) - \langle x^*, h\rangle}{\|h\|} \geq 0\}. \tag{12}$$

One has $\partial_F f(x) \subset \partial_H f(x) \subset \partial_C f(x)$, and the inclusions are strict in general. A function f is called *subdifferentially regular* at x if $\partial_H f(x) = \partial_C f(x)$ (see [17, 3]). Convex functions are subdifferentially regular at every point: indeed $\partial_F f(x) = \partial_C f(x)$ whenever f is convex. This fact is refined below, as a consequence of the following proposition:

Proposition 3 (Clarke-Rockafellar subderivative) *Let $f : X \to \mathbb{R} \cup \{+\infty\}$ be (directionally) approximately convex at $x_0 \in X$. Then, for every $\epsilon > 0$ (and $d \in S_X$) there exists $\delta > 0$ such that for all $x \in B(x_0, \delta) \cap \operatorname{dom} f$ and all $h \neq 0$ so that $x + h \in B(x_0, \delta)$ (and $h/\|h\| \in B(d, \delta)$), one has*

$$f'_{CR}(x; h) \leq f(x+h) - f(x) + \varepsilon\|h\|. \tag{13}$$

Proof. We establish the result for directionally approximately convex f, the case of approximately convex f being analogous. Let $\varepsilon > 0$ and $d \in S_X$. By Definition 2, there exists $\delta' > 0$ such that for all $x', y' \in B(x_0, \delta')$, with $x' \neq y'$ and $(y'-x')/\|y'-x'\| \in B(d, \delta')$, and all $t \in (0, 1)$,

$$f(ty' + (1-t)x') \leq tf(y') + (1-t)f(x') + \varepsilon t(1-t)\|y' - x'\|. \tag{14}$$

We show that (13) is verified with $\delta := \delta'/2$.

Let $x \in B(x_0, \delta) \cap \operatorname{dom} f$ and $h \neq 0$ so that $x + h \in B(x_0, \delta)$ and $h/\|h\| \in B(d, \delta)$. Fix $t > 0$ and take $\lambda \in (0, \delta)$ so that $y \notin B(x, \lambda)$ and

$$\|u\| < \lambda \quad \Rightarrow \quad \left\| \frac{h+u}{\|h+u\|} - \frac{h}{\|h\|} \right\| < \delta.$$

Now, let $x' \in B(x, \lambda)$. Then,

$$\begin{aligned} \inf_{h' \in B(h,\lambda)} \frac{f(x'+th') - f(x')}{t} &\leq \frac{f(x' + t(h + x - x')) - f(x')}{t} \\ &= \frac{f(t(h+x) + (1-t)x') - f(x')}{t}. \end{aligned}$$

Since $h+x$ and x' are in $B(x_0, \delta')$, with $h+x \neq x'$ and $(h+x-x')/\|h+x-x'\| \in B(d, \delta')$, we may apply (14) to the right-hand side of the above inequality to get

$$\begin{aligned} \inf_{h' \in B(h,\lambda)} \frac{f(x'+th') - f(x')}{t} &\leq \frac{tf(h+x) + (1-t)f(x') + \varepsilon t(1-t)\|h+x-x'\| - f(x')}{t} \\ &= f(h+x) - f(x') + \varepsilon(1-t)\|h+x-x'\|. \end{aligned}$$

It follows that, for any $\lambda > 0$,

$$\limsup_{\substack{t \searrow 0 \\ x' \to_f x}} \inf_{h' \in B(h,\lambda)} \frac{f(x'+th') - f(x')}{t} \le f(x+h) - f(x) + \varepsilon \|h\|,$$

hence, $f'_{CR}(x;h) \le f(x+h) - f(x) + \varepsilon\|h\|$. ■

Corollary 3.1 (Subdifferential regularity) *If $f : X \to \mathbb{R} \cup \{+\infty\}$ is approximately convex at $x_0 \in X$, then $\partial_C f(x_0) = \partial_F f(x_0)$. If f is directionally approximately convex at $x_0 \in X$, then $\partial_C f(x_0) = \partial_H f(x_0)$.*

Proof. The first assertion is immediate from Proposition 3 and the definitions of $\partial_C f(x_0)$ and $\partial_F f(x_0)$. To prove the second assertion, let $x^* \in \partial_C f(x_0)$ and $d \in S_X$. Fix $\epsilon > 0$ and take $\delta > 0$ so that (13) is satisfied. Then, for d' sufficiently close to d so that $d'/\|d'\| \in B(d,\delta)$ and for $t > 0$ sufficiently small so that $x_0 + td' \in B(x_0,\delta)$ we have

$$\langle x^*, td' \rangle \le f'_{CR}(x_0; td') \le f(x_0 + td') - f(x_0) + \varepsilon \|td'\|.$$

Dividing by t and letting $t \searrow 0$ and $d' \to d$ give

$$\langle x^*, d \rangle \le \liminf_{\substack{t \searrow 0 \\ d' \to d}} \frac{f(x_0 + td') - f(x_0)}{t} + \varepsilon \|d\|.$$

Since $\varepsilon > 0$ is arbitrary, we conclude that

$$\langle x^*, d \rangle \le \liminf_{\substack{t \searrow 0 \\ d' \to d}} \frac{f(x_0 + td') - f(x_0)}{t},$$

proving that $x^* \in \partial_H f(x_0)$. ■

Remark. The first assertion in Corollary 3.1 is also given in [14, Theorem 3.6]; the second assertion is proved in [10, Theorem 4.1] for locally Lipschitz functions.

4 Subdifferential submonotonicity of approximately convex functions

We recall from [19] (see also [11, 7]) that a set-valued operator $T : X \rightrightarrows X^*$ is said to be *submonotone at* $x_0 \in X$ if for every $\varepsilon > 0$ there exists $\delta > 0$ such that for all $x, y \in B(x_0, \delta)$ and all $x^* \in T(x)$ and $y^* \in T(y)$, one has

$$\langle x^* - y^*, x - y \rangle \ge -\varepsilon \|x - y\|. \tag{15}$$

The directional version of this notion, introduced in [10] (see also [8, 7]), reads as follows:

Definition 4 (Directional submonotonicity) A set-valued operator $T : X \rightrightarrows X^*$ is called *directionally submonotone at* $x_0 \in X$ if for every $d \in S_X$ and $\varepsilon > 0$ there exists $\delta > 0$ such that for all $x, y \in B(x_0, \delta)$, with $x \neq y$ and $(x-y)/\|x-y\| \in B(d, \delta)$, and all $x^* \in T(x)$ and $y^* \in T(y)$, one has

$$\langle x^* - y^*, x - y \rangle \geq -\varepsilon \|x - y\|. \tag{16}$$

It is immediate from Theorem 1 that *if* $\partial_C f$ *is monotone, then* f *is convex* (this result had been previously established in [5, 6] via less direct arguments). In fact, Theorem 1 enables us to establish a more precise statement[1]:

Theorem 5 (Subdifferential characterization of approximate convexity) *Let* $f : X \to \mathbb{R} \cup \{+\infty\}$ *be lower semicontinuous and let* $x_0 \in X$. *The following are equivalent:*

(i) f *is (directionally) approximately convex at* x_0*;*

(ii) $\partial_C f$ *is (directionally) submonotone at* x_0.

Proof. We shall establish the equivalence between directional approximate convexity of f and directional submonotonicity of $\partial_C f$. The equivalence between approximate convexity of f and submonotonicity of $\partial_C f$ can be proved analogously.

(i) $\Rightarrow$ (ii). Let $d \in S_X$ and $\varepsilon > 0$. By Proposition 3 there is $\delta > 0$ such that for all $x, y \in B(x_0, \delta) \cap \operatorname{dom} f$, with $x \neq y$ and $(x-y)/\|x-y\| \in B(d, \delta) \cup B(-d, \delta)$, one has

$$f'_{CR}(x; y - x) \leq f(y) - f(x) + (\varepsilon/2)\|x - y\|,$$

and

$$f'_{CR}(y; x - y) \leq f(x) - f(y) + (\varepsilon/2)\|x - y\|.$$

Adding these inequalities, we obtain that, for all $x^* \in \partial_C f(x)$ and all $y^* \in \partial_C f(y)$,

$$\langle x^* - y^*, y - x \rangle \leq f'_{CR}(x; y - x) + f'_{CR}(y; x - y) \leq \varepsilon \|x - y\|,$$

proving that is $\partial_C f$ directionally submonotone at x_0.

(ii) $\Rightarrow$ (i). Let $d \in S_X$ and $\varepsilon > 0$. By Definition 4 of directional submonotonicity, there exists $\delta' > 0$ such that for all $x, y \in B(x_0, \delta')$, with $x \neq y$ and $(x-y)/\|x-y\| \in B(d, \delta')$, and all $x^* \in T(x)$ and $y^* \in T(y)$, it holds

$$\left\langle x^* - y^*, \frac{y - x}{\|x - y\|} \right\rangle \leq \varepsilon. \tag{17}$$

To prove (i), we show that (10) is verified with $\delta := \delta'/2$. Let $x, y \in B(x_0, \delta) \cap \operatorname{dom} f$, with $x \neq y$ and $(x-y)/\|x-y\| \in B(d, \delta)$, and let $z \in (x, y)$ and $r \in \mathbb{R}$ such that $r \leq f(z)$.

[1] One of the referees has pointed out that part of the subdifferential characterization of Theorem 5 (namely, the case of approximately convex functions) has been recently established in [15]. We thank the referee for this remark.

According to Theorem 1, there exist $\bar{x} \in [x, z)$, $\bar{y} \in (z, y]$ and sequences $\{(x_n, x_n^*)\}_n \subset \partial_C f$ with $x_n \to_f \bar{x}$, and $\{(y_n, y_n^*)\}_n \subset \partial_C f$ with $y_n \to_f \bar{y}$, such that

$$\liminf_n \langle x_n^* - y_n^*, \frac{y_n - x_n}{\|y_n - x_n\|} \rangle \geq \frac{r - f(x)}{\|z - x\|} + \frac{r - f(y)}{\|z - y\|}. \tag{18}$$

Since $\bar{x}, \bar{y} \in B(x_0, \delta)$ and $(\bar{x} - \bar{y})/\|\bar{x} - \bar{y}\| = (x - y)/\|x - y\| \in B(d, \delta)$, for n sufficiently large we have $x_n, y_n \in B(x_0, \delta')$ and $(x_n - y_n)/\|x_n - y_n\| \in B(d, \delta')$, so by (17) and (18)

$$\frac{r - f(x)}{\|z - x\|} + \frac{r - f(y)}{\|z - y\|} \leq \varepsilon.$$

Since $r \leq f(z)$ is arbitrary, we conclude that (10) is verified. ∎

Remark. For locally Lipschitz functions, Theorem 5 is proved in [7, Theorem 2 and Theorem 13].

A fundamental theorem of Rockafellar [16] asserts that *if f is convex and lower semicontinuous with* $\operatorname{dom} f \neq \emptyset$, *then $\partial_C f$ is maximal monotone.* An analogous result holds for the class of directionally approximately convex functions. A set-valued operator $T : X \rightrightarrows X^*$ is said to be *maximal directionally submonotone at $x_0 \in X$* if it is directionally submonotone at x_0 and there is no operator $S : X \rightrightarrows X^*$ directionally submonotone at x_0 such that $T(x) \subset S(x)$ for every x in some neighborhood of x_0 and $T(x_0) \neq S(x_0)$. (This definition is a pointwise version of the definition given in [8] for locally bounded directionally submonotone operators defined on a nonempty open subset U of X, see also [8, Proposition 9].)

Theorem 6 (Maximal submonotonicity of $\partial_C f$) *Let $f : X \to \mathbb{R} \cup \{+\infty\}$ be lower semicontinuous and let $x_0 \in \operatorname{dom} f$. If $\partial_C f$ is directionally submonotone at x_0, then $\partial_C f$ is actually maximal directionally submonotone at x_0.*

Proof. The proof is based on the Zagrodny approximate mean value inequality, as is the proof of the Rockafellar maximal monotonicity theorem proposed in [1, Theorem 5.6], following an idea of Simons [18]. Let $S : X \rightrightarrows X^*$ be directionally submonotone at x_0 such that $\partial_C f(x) \subset S(x)$ for every x in some neighborhood of x_0. Let $x^* \notin \partial_C f(x_0)$. We have to show that $x^* \notin S(x_0)$. Without loss of generality (replacing f by $f - x^*$ and S by $S - x^*$), we may assume that $x^* = 0$. Since $0 \notin \partial_C f(x_0)$, we have $0 \notin \partial_H f(x_0)$, hence there exist $d \in S_X$, $\varepsilon > 0$, and sequences $t_n \searrow 0$ and $d_n \to d$ such that

$$f(x_0) - f(x_0 + t_n d_n) > \varepsilon t_n. \tag{19}$$

Combining this inequality with the Two Points Approximate Mean Value Inequality on the segment $[x_0 + t_n d_n, x_0]$ (see Section 2), we obtain $\bar{y}_n \in [x_0 + t_n d_n, x_0)$, $y_n \in B(\bar{y}_n, \|\bar{y}_n - x_0\|/n)$ and $y_n^* \in \partial_C f(y_n) \subset S(y_n)$ such that

$$\langle y_n^*, \frac{x_0 - y_n}{\|x_0 - y_n\|} \rangle > \frac{\varepsilon t_n}{\|t_n d_n\|} = \frac{\varepsilon}{\|d_n\|}. \tag{20}$$

Observing that $\bar{y}_n \to x_0$, $(\bar{y}_n - x_0)/\|\bar{y}_n - x_0\| = d_n/\|d_n\| \to d/\|d\| = d$ and $\|d_n\| \to 1$, we derive that the sequence $\{(y_n, y_n^*)\}_n \subset S$ verifies $y_n \to x_0$, $(y_n - x_0)/\|y_n - x_0\| \to d$, and, for large n,

$$\langle 0 - y_n^*, \frac{x_0 - y_n}{\|x_0 - y_n\|} \rangle < -\frac{\varepsilon}{2}.$$

This shows that 0 cannot be in $S(x_0)$, since S is directionally submonotone at x_0. ∎

5 The case of arbitrary subdifferentials

In this section, following the abstract approach of subdifferential calculus initiated in [1] and developed further, e.g., in [13, 12], we show that the three-points approximate mean value inequality (see Theorem 1) is actually valid for a large class of subdifferentials in appropriate spaces.

More precisely, 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) If f is convex, then $\partial f(x) = \{ x^* \in X^* \mid \langle x^*, y - x\rangle + f(x) \le f(y), \quad \forall y \in X \}$;

(P2) If $x \in X$ is a finite local minimum of f, then $0 \in \partial f(x)$;

(P3) If φ is convex Lipschitz and ∂-differentiable at x, then $\partial(f + \varphi)(x) \subset \partial f(x) + \partial\varphi(x)$,

where φ *is ∂-differentiable at* x means that both $\partial\varphi(x)$ and $\partial(-\varphi)(x)$ are nonempty.

Further, given a subdifferential ∂, call ∂-*appropriate* any Banach space X in which the following Approximate Optimality Property holds (compare with the exact Optimality Property quoted in Section 2):

Approximate Optimality Property: *If $\bar{x} \in X$ is a finite local minimum of $f + \varphi$, where $f : X \to \mathbb{R} \cup \{+\infty\}$ is lower semicontinuous and $\varphi : X \to \mathbb{R}$ is convex Lipschitz, then there exist sequences $\{(x_n, x_n^*)\}_n \subset \partial f$ with $x_n \to_f \bar{x}$, and $\{(y_n, y_n^*)\}_n \subset \partial\varphi$ with $y_n \to \bar{x}$, such that $\|x_n^* + y_n^*\| \to 0$.*

A large class of subdifferentials (Clarke, Frechet, Hadamard, proximal etc) satisfy the above properties (P1)–(P3). Any Banach space with an equivalent ∂-smooth norm is ∂-appropriate: in particular, any Banach space is ∂_C-appropriate, any separable Banach space is ∂_H-appropriate and any Hilbert space is ∂_P-appropriate, where ∂_P denotes the proximal subdifferential (see, e.g., [2, 4]). Further, any Asplund space is ∂_F-appropriate. See [1, 13, 12] for the details.

Similarly as in Section 2, by combining the Approximate Optimality Property with the Ekeland variational principle, we obtain (see, e.g., [12]):

Approximate Subdifferential Variational Principle: *Let X be a ∂-appropriate space for an arbitrary subdifferential ∂. Let $f : X \to \mathbb{R} \cup \{+\infty\}$ be lower semicontinuous, $\varphi : X \to \mathbb{R}$ be convex Lipschitz, $\lambda > 0$ and $\sigma > 0$. If $\bar{x} \in X$ verifies*

$$(f + \varphi)(\bar{x}) < \inf_{B(\bar{x},\lambda)} (f + \varphi) + \lambda\sigma,$$

then, for any $\varepsilon > 0$, there exist (x, x^*) in ∂f and (y, y^*) in $\partial\varphi$ such that

(i) $\|x - \bar{x}\| < \lambda, \quad \|y - \bar{x}\| < \lambda, \quad \|x - y\| < \varepsilon$,

(ii) $f(x) + \varphi(y) \leq (f + \varphi)(\bar{x}) + \varepsilon$;

(iii) $\|x^* + y^*\| < \sigma$.

Using this Approximate Subdifferential Variational Principle in place of the exact one given in Section 2, it is not difficult to show that the three-points approximate mean value inequality (Theorem 1) remains valid for arbitrary subdifferentials ∂ in ∂-appropriate spaces X (details are left to the reader), as is the case for Zagrodny's two-points approximate mean value inequality (see, e.g., [1, 13]). In particular, Theorem 1 is true for the Fréchet subdifferential in Asplund spaces, for the Hadamard subdifferential in separable Banach spaces, and for the proximal subdifferential in Hilbert spaces. As a consequence, the subdifferential characterization of approximate convexity (Theorem 5) also holds in ∂-appropriate spaces X for any subdifferential ∂ contained in the Clarke subdifferential ∂_C (this generalizes [8, Proposition 14]). Finally, we sketch the proof of the following result, which is analogous to Theorem 6:

Theorem 7 (Maximal submonotonicity of $\partial_F f$ and $\partial_H f$) *Let $f : X \to \mathbb{R} \cup \{+\infty\}$ be lower semicontinuous and let $x_0 \in \operatorname{dom} f$.*

(1) *If X is Asplund and $\partial_F f$ is submonotone at x_0, then $\partial_F f$ is maximal directionally submonotone at x_0.*

(2) *If X admits an equivalent Gâteaux differentiable norm and $\partial_H f$ is directionally submonotone at x_0, then $\partial_H f$ is maximal directionally submonotone at x_0.*

Proof. (1) According to the previous discussion, Theorem 5 holds for ∂_F in Asplund spaces, so we derive that f is approximately convex at x_0, hence $\partial_C f(x_0) = \partial_H f(x_0) = \partial_F f(x_0)$ by Corollary 3.1. To show that $\partial_F f$ is maximal directionally submonotone at x_0, we argue as in Theorem 6, assuming $0 \notin \partial_F f(x_0)$, hence $0 \notin \partial_H f(x_0)$, and using the Two Points Approximate Mean Value Inequality with ∂_F in place of ∂_C.

(2) The argument is similar and therefore is not repeated. ∎

Acknowledgment. A part of this research has been made during a research visit of the first author to the Antilles-Guyane University (February 2005). The first author wishes to thank his hosts for their hospitality.

References

[1] Aussel, D., Corvellec, J.-N., & Lassonde, M., Mean value property and subdifferential criteria for lower semicontinuous functions, *Trans. Amer. Math. Soc.* **347** (1995), 4147–4161.

[2] Borwein, J., & Zhu, Q.J., *Techniques of Variational Analysis*, CMS Books in Mathematics, X, 2005.

[3] Clarke, F.H., *Optimization and Nonsmooth Analysis*, Wiley Interscience, New York, 1983. Reprinted as Vol. 5 of the series Classics in Applied Mathematics, SIAM, Philadelphia, 1990.

[4] Clarke, F.H., Ledyaev, Yu.S., Stern, R.J., & Wolenski, P.R., *Nonsmooth analysis and control theory*, Graduate Texts in Mathematics 178, Springer, New York, 1998.

[5] Correa, R., Jofré, A. & Thibault, L., Characterization of lower semicontinuous convex functions, *Proc. Amer. Math. Soc.* **116** (1992), 67–72.

[6] Correa, R., Jofré, A. & Thibault, L., Subdifferential monotonicity as characterization of convex functions, *Numer. Funct. Anal. Optim.* **15** (1994), 531–535.

[7] Daniilidis, A., & Georgiev, P., Approximate convexity and submonotonicity, *J. Math. Anal. Appl.* **291** (2004), 292–301.

[8] Daniilidis, A., Georgiev, P. & Penot, J.-P., Integration of multivalued operators and cyclic submonotonicity, *Trans. Amer. Math. Soc.* **355** (2003), 177–195.

[9] Ekeland, I., On the variational principle, *J. Math. Anal. Appl.* **47** (1974), 324–353.

[10] Georgiev, P., Submonotone Mappings in Banach Spaces and Applications, *Set-Valued Analysis* **5** (1997), 1-35.

[11] Janin, R., Sur des multiapplications qui sont des gradients généralisés, *C.R. Acad. Sc. Paris* **294** (1982), 117–119.

[12] Jules, F., Sur la somme de sous-différentiels de fonctions semi-continues inférieurement, *Diss. Math.* **423** (2003), 1–62.

[13] Lassonde, M., First-order rules for nonsmooth constrained optimization, *Nonlinear Anal., Theory Methods Appl.* **44** (2001), 1031–1056.

[14] Ngai, H.V., Luc, D.T. & Thera, M., Approximate convex functions, *Journal of Nonlinear and Convex Analysis* **1** (2000), 155–176.

[15] Ngai, H.V. & Penot, J.-P., Approximately convex functions and approximately monotone operators, *Nonlinear Anal.* (to appear).

[16] Rockafellar, R. T., On the maximal monotonicity of subdifferential mappings, *Pacific J. Math.* **33** (1970), 209–216.

[17] Rockafellar, R.T., Directionally Lipschitzian functions and subdifferential calculus, *Proc. London Math. Soc.* **39** (1979), 331–355.

[18] Simons, S., The least slope of a convex function and the maximal monotonicity of its subdifferential, *J. Optim. Theory Appl.* **71** (1991), 127–136.

[19] Spingarn, J.E., Submonotone subdifferentials of Lipschitz functions, *Trans. Amer. Math. Soc.* **264** (1981), 77-89.

[20] Thibault, L. & Zagrodny, D., Integration of subdifferentials of lower semicontinuous functions, *J. Math. Anal. Appl.* **189** (1995), 33–58.

[21] Zagrodny, D., Approximate mean value theorem for upper subderivatives, *Nonlinear Anal.* **12** (1988), 1413–1428.