Canad. Math. Bull. Vol. **40** (1), 1997 pp. 88–102

THE MULTIDIRECTIONAL MEAN VALUE THEOREM IN BANACH SPACES

M. L. RADULESCU AND F. H. CLARKE

ABSTRACT. Recently, F. H. Clarke and Y. Ledyaev established a multidirectional mean value theorem applicable to lower semi-continuous functions on Hilbert spaces, a result which turns out to be useful in many applications. We develop a variant of the result applicable to locally Lipschitz functions on certain Banach spaces, namely those that admit a C^1-Lipschitz continuous bump function.

1. **Introduction.** The multidirectional mean-value inequality established in [6] is a generalization of the mean-value theorem, in the following sense: it gives an estimate for the differences $f(y)-f(x)$ where y is no longer the end of a fixed segment, but ranges over a set Y. For example, when f is a smooth function on $\mathbb{R}^n$, the theorem asserts the existence of a point z in the "interval" $[x,Y]$ (*i.e.* the convex hull of $\{x\}\cup Y$) such that:

$$\min_Y f - f(x) \le \langle f'(z), y-x\rangle \quad \forall y \in Y.$$

The result is developed in [6] for lower semicontinuous functions defined on a Hilbert space.

In this article we will establish a similar result for locally Lipschitz functions in the context of a Banach space that admits a C^1 Lipschitz continuous bump function. We will discuss also a straightforward generalization of the multidirectional mean-value inequality for uniformly smooth Banach spaces.

Let us establish some notation: X is a Banach space, $\|\cdot\|$ its norm, B the closed unit ball in X, and $B(x,\rho)$, the closed unit ball centered at x and of radius ρ. If x is a point and Y a set in X then $[x,Y] := \{z : z = x + t(y-x) \text{ for some } t \in [0,1] \text{ and } y \in Y\}$.

Let $f\colon X \longrightarrow \mathbb{R}\cup\{+\infty\}$ be a function and $x \in X$ be such that $f(x) < \infty$.

DEFINITION 1.1. We say that f attains its *strong minimum* on X at x if $f(x) = \inf\{f(x') : x' \in X\}$ and $\|x_n - x\| \longrightarrow 0$ whenever $x_n \in X$ are such that $f(x_n) \longrightarrow f(x)$.

DEFINITION 1.2. We say that f is *Fréchet subdifferentiable* at x with $x^* \in X^*$ belonging to the *Fréchet subdifferential* at x, denoted $\partial_F f(x)$, provided that

$$\liminf_{y\longrightarrow 0} \frac{f(x+y)-f(x)-\langle x^*, y\rangle}{\|y\|} \ge 0.$$

Let X_1 and X_2 be two Banach spaces and $(a,b) \in \operatorname{dom} f$, where f is defined on $X_1 \times X_2$. We denote by $\partial_{1F} f(a,b)$ the (partial) Fréchet subdifferential of $f(\cdot,b)$ at a and $\partial_{2F} f(a,b)$ that of $f(a,\cdot)$ at b.

Received by the editors July 25, 1995; revised January 4, 1996.
AMS subject classification: 26B05, 49J52.

If C is a closed nonempty subset of X then the distance function d_C is defined by $d_C(x) = \inf\{\|x - c\| : c \in C\}$. The distance function is Lipschitz continuous of rank 1 and in case C is convex, it is convex too. Also, for $\eta > 0$ we define $C(\eta)$ by

$$C(\eta) := \{z \in X : d_C(z) \leq \eta\}.$$

DEFINITION 1.3. A functional $x^* \in X^*$ is said to be a *Fréchet normal to C* at x, $(x \in C)$ if for any $\varepsilon > 0$ there exists $\eta_\varepsilon > 0$ such that

$$\langle x^*, x' - x\rangle \leq \varepsilon\|x' - x\| \quad \text{for all } x' \in C \cap B(x, \eta_\varepsilon).$$

We will denote by $N_C^F(x)$ the set of all Fréchet normals to C at x. If C is convex $N_C^F(x)$ coincides with the normal convex cone to C at x which will be denoted as usual by $N_C(x)$.

We will make the following hypothesis regarding the space X.

(H1) X is a Banach space that admits a Lipschitz continuous bump function which is of class C^1 on X. By a *bump function on X* we mean a function with bounded nonempty support on X.

2. **Preliminaries.** We recall the following results that we will use later on. We suppose throughout this section that X satisfies (H1).

THEOREM 2.1 (THE SMOOTH VARIATIONAL PRINCIPLE[10]). *Let $f: X \longrightarrow (-\infty, \infty]$ be a* l.s.c. *function bounded below. Then for every $\varepsilon > 0$ there exists a function g, which is Lipschitzian, Fréchet differentiable on X and with g' norm to norm continuous on X, $\|g\|_\infty \leq \varepsilon$, $\|g'\|_\infty \leq \varepsilon$ and such that $f + g$ attains its strong minimum on X.*

Theorem 2.1 is a version of the Borwein-Preiss smooth variational principle [1].

PROPOSITION 2.1 ([11]). *There exists a function $d: X \longrightarrow \mathbb{R}^+$ and $K > 1$ such that*
(i) d is bounded, Lipschitzian on X and C^1 on $X \setminus \{0\}$.
(ii) $\|x\| \leq d(x) \leq K\|x\|$ if $\|x\| \leq 1$ and $d(x) = 2$ if $\|x\| \geq 1$.

THEOREM 2.2 ([11]). *Let $f: X \longrightarrow \mathbb{R}$, $x_0 \in X$ and $p \in X^*$. Then the following are equivalent.*
(i) There exists a Fréchet differentiable function $\varphi: X \longrightarrow \mathbb{R}$ such that $f - \varphi$ attains a local minimum at x_0, $\varphi'(x_0) = p$ and φ' is norm to norm continuous at x_0.
(ii) There exists a neighbourhood U of x_0 and a Fréchet differentiable function $\varphi: U \longrightarrow \mathbb{R}$ such that $f - \varphi$ attains a local minimum at x_0, $\varphi'(x_0) = p$, and φ' is norm to norm continuous at x.
(iii) $p \in \partial_F f(x_0)$.

THEOREM 2.3 (FRÉCHET SUBDIFFERENTIAL SUM RULE[12]). *Let $x_0 \in X$ and f_1, f_2 two extended-valued functions defined on X, such that f_1 is* l.s.c. *near x_0 and f_2 is uniformly continuous near x_0. Suppose that $p \in \partial_F(f_1 + f_2)(x_0)$ is given. Then for each $\varepsilon > 0$, there exist $x_i \in X$, $p_i \in \partial_F f_i(x_i)$ $i = 1, 2$ such that*

$$\|x_i - x_0\| < \varepsilon \quad i = 1, 2,\ |f_i(x_i) - f_i(x_0)| < \varepsilon \quad i = 1, 2,\ \|p_1 + p_2 - p\| < \varepsilon.$$

The Fréchet subdifferential sum rule was initially derived by Ioffe in [14], for two functions, one l.s.c. and the other Lipschitz near x_0 defined on a Banach space with Fréchet differentiable norm.

PROPOSITION 2.2 (EXACT PENALIZATION [4, P. 52]). *Let X be any Banach space. Suppose that f is Lipschitz of rank K near x and attains a minimum over C at x. Then $f + Kd_C$ has a local minimum at x.*

Now, by definition $\|(x, y)\|^2 = \|x\|^2 + \|y\|^2$, for all $(x, y) \in X \times Y$.

PROPOSITION 2.3. *Let $f: X \times Y \longrightarrow \overline{\mathbb{R}}$ be a* l.s.c. *function and $(a, b) \in \operatorname{dom} f$. Then $\partial_F f(a, b) \subset \partial_{1F} f(a, b) \times \partial_{2F}(a, b)$.*

The proof is omitted.

PROPOSITION 2.4. *Let $A \subset X$, $B \subset Y$ be two closed sets and $a \in X$, $b \in Y$ such that $d_A(a) = \alpha$, $d_B(b) = \beta$, $\alpha \geq 0$, $\beta \geq 0$. Then $\partial_{1F} d_{A\times B}(a, b) \subset N^F_{A(\alpha)}(a)$.*

PROOF. Let $a^* \in \partial_{1F} d_{A\times B}(a, b)$ and $\varepsilon > 0$. Then there exists $\delta > 0$ such that: $\langle a^*, a' - a\rangle \leq d_{A\times B}(a', b) - d_{A\times B}(a, b) + \varepsilon\|a' - a\|$ whenever $a' \in B(a, \delta)$.

Since

$$\begin{aligned} d_{A\times B}(a', b) \leq d_{A\times B}(a, b) \quad & \text{iff } d^2_A(a') + d^2_B(b) \leq \alpha^2 + \beta^2 \\ & \text{iff } d_A(a') \leq \alpha \\ & \text{iff } a' \in A(\alpha), \end{aligned}$$

we conclude that

$$\langle a^*, a' - a\rangle \leq \varepsilon\|a' - a\| \quad \text{for all } a' \in B(a, \delta) \cap A(\alpha)$$

so

$$a^* \in N^F_{A(\alpha)}(a).$$

∎

For later purpose we will derive another piece of "fuzzy" calculus namely a Fréchet subdifferential "chain rule" for certain functions f and g.

PROPOSITION 2.5. *Let X and Y be two Banach spaces. Assume that $g: X \longrightarrow Y$ is Lipschitz continuous near x_0 and that $f: Y \longrightarrow \mathbb{R}$ is Fréchet differentiable near $g(x_0)$. Then*

$$\partial_F(f \circ g)(x_0) \subset \partial_F\Big\langle f'\big(g(x_0)\big), g(\cdot)\Big\rangle(x_0).$$

PROOF. Since g is Lipschitz continuous near x_0 there exist K and $\eta_1 > 0$ such that $\|g(x) - g(x_0)\| \leq K\|x - x_0\|$, for all $x \in B(x_0, \eta_1)$.

Let $x^* \in \partial_F(f \circ g)(x_0)$ and $\varepsilon > 0$. Then there exists $\eta_2 > 0$ such that for all $x \in B(x_0, \eta_2)$,

$$f\big(g(x_0)\big) - \langle x^*, x_0\rangle \leq f\big(g(x)\big) - \langle x^*, x\rangle + \frac{\varepsilon}{2}\|x - x_0\|.$$

Since f is Fréchet differentiable at $g(x_0)$, there exists $\delta > 0$ s.t. for all $y \in B\big(g(x_0), \delta\big)$

$$f(y) - f\big(g(x_0)\big) \leq \big\langle f'\big(g(x_0)\big), y - g(x_0)\big\rangle + \frac{\varepsilon}{2K}\|g(x_0) - y\|.$$

Set $\eta = \min(\eta_1, \eta_2, \delta/K)$. Combining the three inequalities with $y = g(x)$ in the third results in

$$\langle x^*, x - x_0\rangle \leq \big\langle f'\big(g(x_0)\big), g(x)\big\rangle - \big\langle f'\big(g(x_0)\big), g(x_0)\big\rangle$$
$$+\frac{\varepsilon}{2K}\cdot K\|x - x_0\| + \frac{\varepsilon}{2}\|x - x_0\| \quad \forall x \in B(x_0, \eta)$$

which completes the proof. ∎

COROLLARY 2.1. *Assume that* $h: X \times Y \longrightarrow \mathbb{R}$ *is Fréchet differentiable near* $\big(x_0, g(x_0)\big)$ *and*

$$x^* \in \partial_F\Big(f\big(g(\cdot)\big) + h\big(\cdot, g(\cdot)\big)\Big)(x_0).$$

Then

$$x^* - \frac{\partial h}{\partial x}\big(x_0, g(x_0)\big) \in \partial_F\Big\langle f'\big(g(x_0)\big) + \frac{\partial h}{\partial y}\big(x_0, g(x_0)\big), g(\cdot)\Big\rangle(x_0).$$

THEOREM 2.4. *Let X and Y be two Banach spaces satisfying (H1).*

Let $g: X \longrightarrow Y$ *be Lipschitz continuous near* x_0 *and let* $f: Y \longrightarrow \overline{\mathbb{R}}$ *be uniformly continuous near* $g(x_0)$.

Suppose that $x^* \in \partial_F(f \circ g)(x_0)$. *Then for any given* $\varepsilon > 0$ *there exist* $\bar{x}$, $\bar{z}$ *and* ψ *such that*

$$\|\bar{x} - x_0\| < \varepsilon, \quad \|g(x_0) - \bar{z}\| < \varepsilon, \quad \psi \in \partial_F f(\bar{z}) \quad \textit{and} \quad x^* \in \partial_F\langle\psi, g(\cdot)\rangle(\bar{x}) + \varepsilon B.$$

PROOF. Let $\varphi: X \longrightarrow \mathbb{R}$ be a C^1 function such that $f\big(g(x)\big) - \varphi(x)$ attains its minimum at x_0 and $\varphi'(x_0) = x^*$ (see Theorem 2.2). We can assume without loss of generality that $f\big(g(x)\big) - \varphi(x)$ attains its strong minimum at x_0 (see [12] for example).

Let $\varepsilon > 0$. Since φ is C^1, f is uniformly continuous near $g(x_0)$ and g is Lipschitz continuous near x_0, there exists δ_1, $0 < \delta_1 < \varepsilon$ and $K > 0$ such that:

$$(1) \qquad \|\varphi'(x) - x^*\| \leq \frac{\varepsilon}{2},$$

$$(2) \qquad \|g(x) - g(x_0)\| \leq K\|x - x_0\| \quad \text{whenever } \|x - x_0\| < \delta_1$$

and for any $\gamma > 0$, there exists $\lambda < (\delta_1)/2$ such that

$$(3) \qquad \|f(u) - f(v)\| < \gamma \quad \text{whenever } u, v \in B\big(g(x_0), \delta_1\big) \text{ and } \|u - v\| < \lambda.$$

https://doi.org/10.4153/CMB-1997-011-2 Published online by Cambridge University Press

Let

$$\delta := \min\Big(\frac{\delta_1}{2}, \frac{\delta_1}{2K}\Big). \tag{4}$$

Since $f\big(g(x)\big) - \varphi(x)$ has a strong minimum at x_0, there exists γ_1, $0 < \gamma_1 < \delta$ such that if

$$f\big(g(x)\big) - \varphi(x) \le f\big(g(x_0)\big) - \varphi(x_0) + \gamma_1 \quad \text{then } \|x - x_0\| < \frac{\delta}{2}. \tag{5}$$

Now (3) is true in particular for $\gamma = (\gamma_1)/3$, $u = g(x) + y$, $v = g(x)$; so according to (2) and (4), there exists $\lambda < \delta$ such that

$$\Big\|f\big(g(x) + y\big) - f\big(g(x)\big)\Big\| < \frac{\gamma_1}{3} \tag{6}$$

whenever $\|y\| < \lambda$ and $x \in B(x_0, \delta)$.

(Indeed: $\|g(x) + y - g(x_0)\| \le \|g(x) - g(x_0)\| + \|y\| \le \delta_1/2 + K \cdot \delta \le \delta_1$).

According to Proposition 2.1 there exists $d\colon X \longrightarrow \mathbb{R}^+$, Lipschitz continuous and C^1 on $X \setminus \{0\}$ such that $d(x) \ge \|x\|$, if $\|x\| \le 1$.

For each $n \ge 1$ we define:

$$H_n(x, y) = \begin{cases} f\big(g(x) + y\big) - \varphi(x) + nd^2(y) & \text{if } \|x - x_0\| \le \delta \text{ and } \|y\| \le \lambda \\ +\infty & \text{otherwise.} \end{cases}$$

$X \times Y$ is a Banach space such that there exists a C^1 Lipschitz continuous bump function on it and $H_n(x, y)$ is l.s.c. and bounded below on $X \times Y$. So according to the smooth variational principle, there exists a C^1 function $h\colon X \times Y \longrightarrow \mathbb{R}$ such that $H_n(x, y) + h(x, y)$ attains its strong minimum at some point $(x_n, y_n) \in B(x_0, \delta) \times B(0, \lambda)$, and such that

$$\|h\|_\infty < \frac{\gamma_1}{3} \quad \text{and} \quad \|h'\|_\infty < \frac{\varepsilon}{2}. \tag{7}$$

CLAIM. $(x_n, y_n) \in \operatorname{int} B(x_0, \delta) \times \operatorname{int} B(0, \lambda)$, *so* (x_n, y_n) *is a local minimum for* $H_n(x, y) + h(x, y)$.

PROOF OF THE CLAIM. Indeed, in particular

$$H_n(x_n, y_n) + h(x_n, y_n) \le H_n(x_0, 0) + h(x_0, 0)$$

so,

$$\begin{aligned} n\|y_n\|^2 \le nd^2(y_n) &\le f\big(g(x_0)\big) - \varphi(x_0) + h(x_0, 0) \\ &\quad - f\big(g(x_n) + y_n\big) + \varphi(x_n) - h(x_n, y_n) \\ &\le f\big(g(x_0)\big) - f\big(g(x_n) + y_n\big) + \varphi(x_n) - \varphi(x_0) + 2\frac{\gamma_1}{3}. \end{aligned} \tag{8}$$

The right-hand side is bounded, so choosing n large enough,

$$\|y_n\| < \lambda < \delta, \tag{9}$$

https://doi.org/10.4153/CMB-1997-011-2 Published online by Cambridge University Press

so $y_n \in \operatorname{int} B(0, \lambda)$. Also by (6)

$$\left|f\bigl(g(x_n)+y_n\bigr)-f\bigl(g(x_n)\bigr)\right|<\gamma_1/3.$$

Combining (8) and the above inequality we conclude that

$$f\bigl(g(x_n)\bigr)-\varphi(x_n)\le f\bigl(g(x_0)\bigr)-\varphi(x_0)+\gamma_1$$

so by (5), $\|x_n-x_0\|<\delta/2$ which completes the proof of the claim.

Now set

$$g(x)+y=z,\quad g(x_n)+y_n=z_n. \tag{10}$$

Then

$$z\longrightarrow f(z)-\varphi(x_n)+nd^2\bigl(z-g(x_n)\bigr)+h\bigl(x_n,z-g(x_n)\bigr)$$

has a local minimum at z_n, so

$$\psi:=-n(d^2)'\bigl(z_n-g(x_n)\bigr)-\frac{\partial h}{\partial y}\bigl(x_n,z_n-g(x_n)\bigr)\in\partial_F f(z_n). \tag{11}$$

Also,

$$x\longrightarrow f(z_n)-\varphi(x)+nd^2\bigl(z_n-g(x)\bigr)+h\bigl(x,z_n-g(x)\bigr)$$

has a local minimum at x_n, so

$$\varphi'(x_n)\in\partial_F\Bigl(nd^2\bigl(z_n-g(\cdot)\bigr)+h\bigl(\cdot,z_n-g(\cdot)\bigr)\Bigr)(x_n).$$

We deduce from Corollary 2.1 that

$$\begin{aligned}p:=\varphi'(x_n)-\frac{\partial h}{\partial x}\bigl(x_n,z_n-g(x_n)\bigr)&\\ \in\partial_F\Bigl\langle-n(d^2)'\bigl(z_n-g(x_n)\bigr)-\frac{\partial h}{\partial y}\bigl(x_n,z_n-g(x_n)\bigr),g(\cdot)\Bigr\rangle(x_n)&\\ =\partial_F\langle\psi,g(\cdot)\rangle(x_n).&\end{aligned}$$

Relations (1), (4), (7) and the definition of p imply

$$\|x^*-p\|\le\frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon.$$

According to (9) and (10) we also have

$$\begin{aligned}\|g(x_0)-z_n\|&\le\|g(x_0)-g(x_n)\|+\|y_n\|\le K\|x_0-x_n\|+\lambda\\ &\le K\delta+\frac{\delta_1}{2}\le K\frac{\delta_1}{2K}+\frac{\delta_1}{2}<\varepsilon.\end{aligned}$$

Consequently, the conclusion follows with

$\bar{x}=x_n,\quad \bar{z}=z_n$ and ψ given by (11). ■

https://doi.org/10.4153/CMB-1997-011-2 Published online by Cambridge University Press

COROLLARY 2.2. *Besides the hypothesis in Theorem 2.4, assume that g is Fréchet differentiable near* x_0. *Then for any given* $\varepsilon > 0$, *there exist* $\overline{x} \in X$, $\overline{z} \in Y$ *and* $\psi \in Y^*$ *such that*

$$\|\overline{x} - x_0\| < \varepsilon, \quad \|g(x_0) - \overline{z}\| < \varepsilon,$$
$$\psi \in \partial_F f(\overline{z}) \quad \textit{and} \quad \|x^* - \psi \circ g'(\overline{x})\| \leq \varepsilon,$$

or equivalently, there exists $w^* \in X^*$, $\|w^*\| \leq 1$ *such that*

$$\langle x^* - \varepsilon w^*, x\rangle = \big\langle \psi, \langle g'(\overline{x}), x\rangle\big\rangle \quad \forall x \in X.$$

3**. The main result.** Let X be a Banach space satisfying (H1), and $Y \subset X$ a closed, bounded, convex set. Suppose $f: X \longrightarrow \mathbb{R}$ is locally Lipschitz and bounded below on $[x_0, Y] + \delta_0 B$ for some $\delta_0 > 0$. Define

$$\hat{r} := \inf_Y f(y) - f(x_0)$$

and

$$V_a := \{x_0 + t(y - x_0) : t \in [a, 1), y \in Y\}.$$

THEOREM 3.1. *Let* $r < \hat{r}$. *Then there exists* a, $0 < a < 1$, *such that for any* $\delta > 0$, *there exist* $z \in V_a + \delta B$ *and* $\xi \in \partial_F f(z)$ *such that*

$$r < \langle \xi, y - x_0\rangle \quad \forall y \in Y.$$

PROOF. We may assume without loss of generality that $x_0 = 0$.

Since Y is bounded, there exists $M > 0$ such that $\|y\| < M$ for all $y \in Y$ and since f is Lipschitz near 0 there exist $K > 0$ and $\gamma > 0$ such that

$$|f(z) - f(0)| \leq K\|z\| \quad \forall z, \|z\| \leq \gamma. \tag{12}$$

Choose $\overline{r}$ satisfying $\hat{r} > \overline{r} > r$.

$$\text{If } 1 + KM - \overline{r} \leq 0 \text{ we choose } 0 < a < \min(\gamma/M, 1) \text{ and}$$
$$\text{if } 1 + KM - \overline{r} > 0 \text{ we choose } 0 < a < \min\big(\gamma/M, (\hat{r} - \overline{r})/(1 + KM - \overline{r}), 1\big).$$

Define

$$H(t, y) = \begin{cases} f(ty) - \overline{r}t & \text{if } (t, y) \in [a, 1] \times Y \\ \infty & \text{otherwise.} \end{cases}$$

(For a justification of the choice of H, we refer to the "motivating idea" behind the multidirectional mean value inequality exposed in [6].)

Let

$$0 < \varepsilon < \min\Big(\frac{a}{2}, \delta_0, \delta, \frac{\overline{r} - r}{2(1 + M/a)}, \hat{r} - r\Big). \tag{13}$$

There exists a C^1 Lipschitz bump function on $\mathbb{R} \times X$ and H is lower semicontinuous and bounded below on $\mathbb{R} \times X$, so according to the smooth variational principle, there exists a C^1 Lipschitz function $g: \mathbb{R} \times X \longrightarrow \mathbb{R}$ such that

$$H(t,y) + g(t,y) \quad \text{attains its minimum at some point } (\bar{t},\bar{y}) \in [a,1] \times Y$$

and

$$\|g\|_\infty < \varepsilon, \quad \|\nabla_t g\|_\infty \le \frac{\varepsilon}{2}, \quad \|\nabla_y g\|_\infty \le \frac{\varepsilon}{2}. \tag{14}$$

CLAIM. $\bar{t} \neq 1$.

PROOF OF THE CLAIM. Suppose $\bar{t} = 1$; then,

$$f(\bar{y}) - \bar{r} + g(1,\bar{y}) \le f(a\bar{y}) - a\bar{r} + g(a,\bar{y}),$$

so

$$f(\bar{y}) - f(a\bar{y}) \le \bar{r} - a\bar{r} + 2\varepsilon < \bar{r} - a\bar{r} + a.$$

Since a was chosen such that $a \le \gamma/M$, $a\|\bar{y}\| \le \gamma$ and according to (12)

$$-KaM + f(\bar{y}) - f(0) < \bar{r} - a\bar{r} + a,$$

so

$$\inf_Y f(y) - f(0) < \bar{r} - a\bar{r} + a + KaM$$

that is

$$\hat{r} - \bar{r} < a(1 + KM - \bar{r}).$$

If $1 + KM - \bar{r} \le 0$ we have a contradiction since $\hat{r} - \bar{r} > 0$ and if $1 + KM - \bar{r} > 0$, then $a > (\hat{r} - \bar{r})/(1 + KM - \bar{r})$ and we have a contradiction too, with the choice of a. This establishes the claim. ∎

Since f is locally Lipschitz near $\bar{t}\bar{y}$, there exists $\bar{K}$ and $\bar{\gamma}$ such that

$$B(\bar{t}\bar{y},\bar{\gamma}) \subset V_a + a\delta_0 B \quad \text{and} \tag{15}$$
$$\|f(y) - f(z)\| < \bar{K}\|y - z\| \quad \forall y,z \in B(\bar{t}\bar{y},\bar{\gamma}).$$

$(V_a + a\delta_0 B \subset \{ty : t \in [a,1], y \in Y + \delta_0 B\} \subset V_a + \delta_0 B \subset [0,Y] + \delta_0 B)$.

Denote by $\hat{K}$ the Lipschitz constant of $f(ty) - \bar{r}t + g(t,y)$ near $(\bar{t},\bar{y})$ and by A the set $[a,1]$. Then by Proposition 2.2, $f(ty) - \bar{r}t + g(t,y) + \hat{K}d_{A\times Y}(t,y)$ has a local minimum at $(\bar{t},\bar{y})$. Consequently

$$(t^*,y^*) := \big(\bar{r} - \nabla_t g(\bar{t},\bar{y}), -\nabla_y g(\bar{t},\bar{y})\big) \in \partial_F\big(F(\cdot,\cdot) + \hat{K}d_{A\times Y}(\cdot,\cdot)\big)(\bar{t},\bar{y}), \tag{16}$$

where $F(t,y) := f(ty)$. Set

$$\lambda_0 := \min\Big(\frac{1-\bar{t}}{2}, \frac{2\varepsilon}{2+M}, \frac{2\varepsilon}{3K}, \frac{\bar{\gamma}}{2+M}\Big). \tag{17}$$

https://doi.org/10.4153/CMB-1997-011-2 Published online by Cambridge University Press

According to Theorem 2.3, for any $0 < \lambda < \lambda_0$ there exist (t_i, y_i) $i = 1, 2$ such that

$$|\bar{t} - t_i| < \frac{\lambda}{2}, \quad \|\bar{y} - y_i\| < \frac{\lambda}{2} \quad i = 1, 2 \text{ and} \tag{18}$$

$$(t^*, y^*) \in \partial_F F(t_1, y_1) + \hat{K}\partial_F d_{A\times Y}(t_2, y_2) + \frac{\lambda}{4}B.$$

Furthermore, by Proposition 2.3

$$(t^*, y^*) \in \partial_F F(t_1, y_1) + \hat{K}\partial_{1F} d_{A\times Y}(t_2, y_2) \times \partial_{2F} d_{A\times Y}(t_2, y_2) + \frac{\lambda B}{4}.$$

Now, denote $\alpha := d_A(t_2)$ and $\beta := d_Y(y_2)$.

Proposition 2.4 and the fact that A and Y are convex sets, imply that there exist

$$n_t \in N_{A(\alpha)}(t_2), \quad n_y \in N_{Y(\beta)}(y_2), \quad \theta_0 \in [0, 1] \quad \text{and} \quad w \in X^* \text{ with } \|w\| \le 1 \tag{19}$$

such that

$$\Big(t^* - \hat{K}n_t - \frac{\lambda}{4}\theta_0, y^* - \hat{K}n_y - \frac{\lambda}{4}w_0\Big) \in \partial_F F(t_1, y_1).$$

According to Corollary 2.2, for any $\lambda, 0 < \lambda < \lambda_0$, there exist $(\tilde{t}, \tilde{y}) \in \mathbb{R} \times X$, $\tilde{z} \in X$, $\theta \in \mathbb{R}$, $\psi, w \in X^*$ such that

$$|\tilde{t} - t_1| < \frac{\lambda}{2}, \quad \|\tilde{y} - y_1\| < \frac{\lambda}{2} \tag{20}$$

$$\|t_1 y_1 - \tilde{z}\| < \frac{\lambda}{2}, \quad \psi \in \partial_F f(\tilde{z}), \tag{21}$$

$$\theta \in [0, 1], \quad \|w\| \le 1 \quad \text{and} \tag{22}$$

$$\alpha\Big(t^* - \hat{K}n_t - \frac{\lambda}{4}(\theta_0 + \theta)\Big) + \Big\langle y^* - \hat{K}n_y - \frac{\lambda}{4}(w_0 + w), v\Big\rangle = \langle \psi, \alpha\tilde{y} + \tilde{t}v\rangle$$

for all $(\alpha, v) \in \mathbb{R} \times X$, where (t^*, y^*) is given by (16).

For $\alpha = 0$, $v = y - y_2$, $y \in Y$, the above inequality results in

$$\Big\langle -\nabla_y g(\bar{t}, \bar{y}) - \hat{K}n_y - \frac{\lambda}{4}(w_0 + w), y - y_2\Big\rangle = \tilde{t}\langle \psi, y - y_2\rangle \quad \text{for any } y \in Y.$$

Using (14), (19) and (22) we obtain

$$\langle \psi, y - y_2\rangle \ge -\frac{\varepsilon}{2} \cdot \frac{M}{\tilde{t}} - \frac{\varepsilon}{2}\frac{M}{\tilde{t}} \ge -\frac{M}{\tilde{t}} \cdot \varepsilon \quad \text{for any } y \in Y. \tag{23}$$

Now if we set $v = 0$, $\alpha \neq 0$, and use again (14), (19) and (22),

$$\langle \psi, \tilde{y}\rangle = \bar{r} - \nabla_t g(\bar{t}, \bar{y}) - \frac{\lambda}{4}(\theta_0 + \theta) - \hat{K}n_t \ge \bar{r} - \frac{\varepsilon}{2} - \frac{\varepsilon}{2} \tag{24}$$

(because $n_t \le 0$, t_2 being $\neq 1$).

https://doi.org/10.4153/CMB-1997-011-2 Published online by Cambridge University Press

Combining (17), (18) and (21) we obtain

$$\|\tilde{z}-\bar{t}\bar{y}\| \leq \|\tilde{z}-t_1y_1\| + t_1\|y_1-\bar{y}\| + \|\bar{y}\|\,|t_1-\bar{t}| < \frac{\lambda}{2}+\frac{\lambda}{2}+\frac{M\lambda}{2} < \frac{\bar{\gamma}}{2}. \tag{25}$$

So by (15) f is locally Lipschitz near $\tilde{z}$ with the same constant $\bar{K}$ which implies that $\|\psi\| < \bar{K}$.

Starting with (24) and using once again (17), (18) and (21),

$$\bar{r}-\varepsilon \leq \langle\psi,\tilde{y}\rangle \leq \langle\psi,y_2\rangle + \bar{K}\|\tilde{y}-y_2\| \leq \langle\psi,y_2\rangle + 3\bar{K}\frac{\lambda}{2} \leq \langle\psi,y_2\rangle+\varepsilon. \tag{26}$$

Adding the inequalities (23) and (26) results in

$$\bar{r}-\varepsilon\Big(2+\frac{M}{\tilde{t}}\Big) \leq \langle\psi,y\rangle \quad \text{for any } y\in Y.$$

According to (13), (17), (18) and (20), $|\tilde{t}-\bar{t}| < \lambda < a/2$ and since $\bar{t} \geq a$ it follows that $\tilde{t} > a/2$. So from the last inequality we get

$$\bar{r}-\varepsilon\Big[2+\frac{2M}{a}\Big] \leq \langle\psi,y\rangle \quad \text{for all } y\in Y.$$

Since ε was chosen less then $(\bar{r}-r)/\big(2(1+M/a)\big)$,

$$r < \langle\psi,y\rangle \quad \text{for all } y\in Y.$$

By (17) and (25) $\|\tilde{z}-\bar{t}\bar{y}\| < \varepsilon < \delta$ where $\bar{t}\in[a,1)$ and $\bar{y}\in Y$. Consequently $\tilde{z}\in V_a+\delta B$. According to (21), $\psi\in\partial_F f(\tilde{z})$ so the proof is complete. ∎

We may reformulate the theorem in the following way:

COROLLARY 3.1. *Under the same hypotheses on f, let $\delta > 0$. Then there exist a, $0 < a < 1$, $z\in V_a+\delta B$ and $\xi\in\partial_F f(z)$ such that,*

$$\inf_Y f - f(x_0) \leq \langle\xi, y-x_0\rangle + \delta \quad \forall y\in Y.$$

Obviously, the unidirectional "fuzzy" mean-value inequality follows from Theorem 3.1. We refer to [16] for a discussion of various consequences when Y is a singleton.

In the following corollary we will denote by $\partial_C f$ the generalized gradient of f; for definition and properties see [4, chapter 2]. We mention that

$$\partial_F f(x) \subset \partial_C f(x).$$

COROLLARY 3.2. *Suppose in addition that Y is compact. Then there exist $z\in[x_0,Y]$ and $\xi\in\partial_C f(z)$ such that*

$$\min_Y f - f(x_0) \leq \langle\xi, y-x_0\rangle \quad \textit{for any } y\in Y.$$

https://doi.org/10.4153/CMB-1997-011-2 Published online by Cambridge University Press

PROOF. Let $\delta_i > 0$, $\delta_i \longrightarrow 0$. Then by Corollary 3.1, there exist (a_i), $a_i > 0$, $a_i \longrightarrow 0$, $z_i \in V_{a_i} + \delta_i B$ and $\xi_i \in \partial_F f(z_i)$ such that

$$\min_Y f - f(x_0) < \langle \xi_i, y - x_0 \rangle + \delta_i \quad \text{for any } y \in Y.$$

By the above inclusion, $\xi_i \in \partial_C f(z_i)$. Since Y is compact and $\partial_C f$ is a weak* closed multifunction we can pass to the limit in the inequality on a subsequence and the result follows. ■

We now state without proof another version of the mean-value inequality which is obtained by minor modifications of the original one [6]. Suppose X satisfies the following hypothesis:

(H2) X is a uniformly smooth Banach space.

Of course (H2) is stronger than (H1). However uniformly smooth Banach spaces include a quite large class of spaces as for example L^P-spaces with $1 < p < \infty$. On the other hand the hypothesis on f will be weaker, namely f need only be l.s.c.

Let $Y \subset X$ be a nonempty, closed, bounded subset of X (not necessarily convex).

Let $f: X \times (-\infty, +\infty]$ be a l.s.c. function, finite at x, and define

$$\hat{r} := \lim_{\delta \downarrow 0} \inf_{y \in Y + \delta B} \{f(y) - f(x)\}.$$

■

THEOREM 3.2. *Let f be bounded below on $[x, Y] + \delta B$ for some $\delta > 0$. Then for any $r < \hat{r}$ and ε between 0 and δ, there exist $z_0 \in [x, Y] + \varepsilon B$, $y_0 \in Y$ and $\xi \in \partial_F f(z_0)$ such that for any $\lambda > 0$, there exists $\gamma > 0$ such that for any $y \in Y \cap B(y_0, \gamma)$*

$$r < \langle \xi, y - x \rangle + \lambda \|y - y_0\|.$$

(If Y is convex the theorem applies with $\lambda = 0$ and $\gamma = +\infty$).

4. **Examples.** As a first application we will give an "infinitesimal version" of Theorem 3.1 (see Section 3 in [6]) which extends an important result of Subbotin [9] to the Banach space setting.

For a given subset E of X and a point x at which the function f is finite we introduce $\underline{D}f(x; E)$, the quantity given by

$$\underline{D}f(x; E) := \liminf_{t \downarrow 0} \inf_{\delta \downarrow 0} \inf_{e \in E + \delta B} \frac{f(x + te) - f(x)}{t}.$$

Suppose X is a reflexive Banach space. Then it can be given an equivalent Fréchet differentiable norm (see for example [13]) and consequently the space satisfies hypothesis (H1) (see [11]) for details and further comments).

THEOREM 4.1. *Let E be a nonempty closed, bounded, convex subset of X and let $f: X \longrightarrow \mathbb{R}$ be a locally Lipschitz function. Suppose for some scalar ρ we have*

$$\underline{D}f(x; E) > \rho. \tag{27}$$

Then for any $\varepsilon' > 0$ *there exist a point* z *and* $\xi \in \partial_F f(z)$ *such that* $\|z - x\| < \varepsilon'$, $|f(z) - f(x)| < \varepsilon'$ *and*

$$\langle \zeta, e\rangle > \rho, \quad \forall e \in E.$$

The proof is based on Theorem 3.1 as for the Hilbert space case given in [6].

We now introduce the *weak Dini derivative* (or *weak-Hadamard derivative*)

$$D^w f(x; e) := \inf_{\{e_i\}} \lim_{t\downarrow 0} \inf_{e_i \xrightarrow{w} e} \frac{f(x + te_i) - f(x)}{t}$$

and the *weak Dini* (or *weak-Hadamard*) *subgradient*

$$\partial_D f(x) = \{x^* \in X \mid D^w f(x; v) \geq \langle x^*, v\rangle, \ \forall v \in X\}.$$

In general $\partial_F f(x) \subseteq \partial_D f(x)$ and in case X is reflexive they coincide (see for example [3]).

PROPOSITION 4.1. *If Y is a nonempty closed, bounded, convex subset of X then*

$$\inf_{e\in E} D^w f(x; e) \leq \underline{D}f(x; E). \tag{28}$$

PROOF. Let (t_i), (δ_i) and (e_i) be sequences such that $t_i \downarrow 0$, $\delta_i \downarrow 0$, $e_i \in E + \delta_i B$ and

$$\frac{f(x + t_i e_i) - f(x)}{t_i} \longrightarrow \underline{D}f(x; E).$$

Since (e_i) is bounded there exists a weak convergent subsequence (without relabeling), $e_i \xrightarrow{w} e$. Since Y is closed and convex it is also weakly closed so $e \in E$. Hence there exists $e \in E$ such that

$$D^w f(x; e) \leq \lim_{i\to\infty} \frac{f(x + t_i e_i) - f(x)}{t} = \underline{D}f(x; E)$$

which completes the proof. ∎

THEOREM 4.2. *Let E be a nonempty closed, bounded, convex subset of a reflexive Banach space X, and let $f : X \longrightarrow \mathbb{R}$ be a locally Lipschitz function.*

Suppose that for some ρ we have

$$\inf_{e\in E} D^w f(x; e) > \rho. \tag{29}$$

Then for any $\varepsilon > 0$ *there exist* $z \in X$ *and* $\zeta \in \partial_D f(z)$ *such that* $\|z - x\| < \varepsilon$, $|f(z) - f(x)| < \varepsilon$ *and*

$$\langle \zeta, e\rangle > \rho \quad \forall e \in E.$$

PROOF. In view of Proposition 4.1, (29) implies (27), so Theorem 4.1 applies. ∎

REMARK 4.1. The same result holds under hypothesis (H2) for a l.s.c. function, finite at x. In this case the inequality $|f(z) - f(x)| < \varepsilon$ is no longer implicit as in the Lipschitz case. We refer again to [6] for the details.

As a second example of the use of the theorem we characterize monotonicity of a function with respect to a cone by means of its Fréchet subdifferential. We refer to [6, 7, 8] for the Hilbert version and for other applications.

DEFINITION 4.1. Let $K \subset X$ be a cone. A function $f: X \longrightarrow (-\infty, \infty]$ is said to be *K-nonincreasing* if

$$y \in x + K \Longrightarrow f(y) \leq f(x).$$

The *polar* of a nonempty set $K \subset X$ is the set

$$K^* = \{y \in X : \langle y, x \rangle \leq 0, \forall x \in K\}.$$

PROPOSITION 4.2. *Suppose X satisfies (H1) (respectively (H2)), $f: X \longrightarrow (-\infty, +\infty]$ is a locally Lipschitz (respectively lower-semicontinuous) function, and $K \subset X$ is a cone. Then f is K-nonincreasing if and only if*

$$\partial_F f(x) \subseteq K^* \quad \forall x \in X. \tag{30}$$

PROOF. In order to prove the necessity let $x \in X$ and $z \in K$. By assumption, $f(x + tz) \leq f(x)$ for all $t > 0$. Suppose $p \in \partial_F f(x)$ and let $\varepsilon > 0$. Then there exists $\eta > 0$, such that $\langle p, tz \rangle - \varepsilon t \|z\| \leq f(x + tz) - f(x) \leq 0$ for any t, $0 < t < \eta$. Dividing by t and letting $\varepsilon \longrightarrow 0$ leads to $\langle p, z \rangle \leq 0$. Since $z \in K$ is arbitrary we conclude that $p \in K^*$.

Now assume that (30) holds and suppose that f is not K-nonincreasing. Then there exist points x and y such that $y \in x + K$ but $f(y) > f(x)$. We now apply Theorem 3.1 with $Y := \{y\}$. So there exist z and $p \in \partial_F f(z)$ such that

$$0 < f(y) - f(x) < \langle p, y - x \rangle.$$

But since $y - x \in K$ this contradicts the assumption that $p \in K^*$. ∎

Now we will characterize weak-monotonicity. For a detailed discussion see [6] and [9]. Let D be a nonempty, compact, convex subset of X, where X satisfies (H1) (respectively (H2)) and let $f: X \longrightarrow \mathbb{R}$, be locally Lipschitz (respectively lower-semicontinuous).

PROPOSITION 4.3. *Suppose that*

$$u \in X, \quad p \in \partial_F f(u) \Longrightarrow \min_{d \in D} \langle p, d \rangle \leq 0.$$

Then for any $x \in X$ and for any $t > 0$, we have

$$\min_{y \in x + tD} f(y) \leq f(x).$$

https://doi.org/10.4153/CMB-1997-011-2 Published online by Cambridge University Press

PROOF. According to Theorem 3.1, for some z and some $p \in \partial_F f(z)$

$$r < \min_{d\in D} \langle p, td\rangle$$

for any r such that $r < \min_{y\in x+tD} f(y) - f(x)$. By hypothesis, $\min_{d\in D}\langle p, d\rangle \leq 0$ so $r < 0$ and the conclusion follows. ■

We remark that the strong-monotonicity characterization uses only the unidirectional mean-value theorem whereas the weak-monotonicity one requires the multidirectional mean value theorem.

Finally we want to point out the specificity of the multidirectional mean value results already known. In [6] Clarke and Ledyaev treat functions which are just l.s.c. and in the framework of a Hilbert space. The same authors derived in [7] a multidirectional mean value theorem for locally Lipschitz functions defined on a general Banach space, in the "two-set" case (*i.e.* the point x_0 is replaced by a closed, convex, bounded set). However the subgradient figuring in that theorem is the generalized gradient, $\partial_C f$. Consequently, on the one hand our result generalizes the result in [6] from the point of view of the space but under the hypothesis (H1) is more restrictive with respect to the class of functions. On the other hand it generalizes the result in [7] from the point of view of the subgradient class, but is more restrictive with respect to the space and it covers just the "point-set" case. We remark that it gives a more precise estimate of the set in which the point z lies in terms of the initial tolerance in the choice of r.

REFERENCES

1. J. M. Borwein and D. Preiss, *A smooth variational principle with applications to subdifferentiability and to differentiability of convex functions*, Trans. Amer. Math. Soc. **303**(1987), 517–527.

2. J. M. Borwein and H. M. Strojwas, *Proximal analysis and boundaries of closed sets in Banach space, I, Theory*, Canad. J. Math. **38**(1986), 431–452.

3. ———, *Proximal analysis and boundaries of closed sets in Banach space, II, Applications*, Canad. J. Math. **39**(1987), 428–472.

4. F. H. Clarke, *Optimization and nonsmooth analysis*, Classics Appl. Math. **5**, SIAM Publications, Philadelphia, 1990, originally published by Wiley Interscience, New York, 1983.

5. ———, *Methods of dynamic and nonsmooth optimization*, CBMS-NSF Regional Conf. Ser. Appl. Math. **57**, SIAM Publications, Philadelphia, 1989.

6. F. H. Clarke and Yu. S. Ledyaev, *Mean value inequalities in Hilbert space*, Trans. Amer. Math. Soc. **344**(1994), 307–324.

7. ———, *Mean value inequalities*, Proc. Amer. Math. Soc. **122**(1994), 1075–1083.

8. F. H. Clarke, Yu S. Ledyaev, R. J. Stern, and P. R. Wolenski, *Subdifferential and nonlinear analysis*, monograph, in preparation.

9. F. H. Clarke, R. J. Stern and P. R. Wolenski, *Subgradient criteria for monotonicity, the Lipschitz condition and convexity*, Canad. J. Math. **45**(1993), 1167–1183.

10. R. Deville, G. Godefroy and V. Zizler, *A smooth variational principle with applications to Hamilton-Jacobi equations*, Infinite Dimensions, J. Funct. Anal. **111**(1993), 197–212.

11. ———, *Smoothness and renormings in Banach spaces*, Pitman Monographs, Surveys Pure Appl. Math. **64**, Longman, 1993.

12. R. Deville and E. M. El Haddad, *The subdifferential of the sum of two functions in Banach spaces, I. First order case*, to appear.

13. J. Diestel, *Geometry of Banach spaces-selected topics*, Lecture Notes in Math. **485**, Springer, 1975.

14. A. D. Ioffe, *Proximal analysis and approximate subdifferentials*, J. London Math. Soc. (2) **41**(1990), 175–192.

15. A. S. Lewis, D. Ralph, *A nonlinear duality result equivalent to the Clarke-Ledyaev mean value inequality*, Nonlinear Anal., to appear.
16. P. D. Loewen, *Optimal control via nonsmooth analysis*, CRM Proceedings and Lecture Notes **2**, Amer. Math. Soc., Providence, 1993.
17. A. Subbotin, *On a property of a subdifferential*, Mat. Sb. **1982**(1991) 1315–1330; Russian.
18. ———, *Generalized solutions of first order PDEs: The dynamical optimization perspective*, Birkhaüser, to appear.

Centre de recherches mathématiques
Université de Montréal
C.P. 6128, Succ. Centre-ville
Montréal, Québec
H3C 3J7

https://doi.org/10.4153/CMB-1997-011-2 Published online by Cambridge University Press