

A SMOOTH VARIATIONAL PRINCIPLE WITH APPLICATIONS TO SUBDIFFERENTIABILITY AND TO DIFFERENTIABILITY OF CONVEX FUNCTIONS

J. M. BORWEIN AND D. PREISS

ABSTRACT. We show that, typically, lower semicontinuous functions on a Banach space densely inherit lower subderivatives of the same degree of smoothness as the norm. In particular every continuous convex function on a space with a Gâteaux (weak Hadamard, Fréchet) smooth renorm is densely Gâteaux (weak Hadamard, Fréchet) differentiable. Our technique relies on a more powerful analogue of Ekeland's variational principle in which the function is perturbed by a quadratic-like function. This "smooth" variational principle has very broad applicability in problems of nonsmooth analysis.

1. Introduction. Ekeland's variational principle [11–13] has proved, along with its variants, to be a potent and flexible tool in analysis and in optimization theory [2, 7, 12, 13, 17]. One notable limitation on its application is that even when the original function is differentiable the perturbed function is not. A reasonable smooth variant has long been sought.

In §2 we provide such a "smooth" variational principle. The geometric idea behind this proof is as follows. Given a fixed penalty function one cannot, in general, use penalization techniques to densely obtain minima for lower semicontinuous functions on a given Banach space (but see Theorem 5.2). One can however adaptively adjust the penalty as one moves around the epigraph of the function, and the final cumulative penalty can be well enough controlled so as to inherit the differentiability properties of the underlying norm on the space.

In §3 we deduce the existence of appropriate subderivatives. These subderivatives enable one to considerably extend and strengthen many existing nonsmooth optimization results such as those in Treiman [25], as will be illustrated elsewhere [6].

In §4 we obtain consequences for convex functions. One result particularly merits comment. We show that every Banach space with a smooth renorm is a GDS, in the language of Larman and Phelps [18]. This largely answers a long-standing question of Day's [9, p. 167]. In §5 implications for distance functions are considered. These include a strengthening of our variational principle in reflexive Banach space.

2. The smooth variational principle. Let X be a Banach space with given norm $\|\cdot\|$. For $p \geq 1$ we consider the class of convex functions, Γ_p , which consists

Received by the editors October 3, 1986. Presented at the AMS Winter Meeting in San Antonio by the first author.

1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 49A52; Secondary 58C20, 49A51.

Key words and phrases. Weak Asplund spaces, subderivatives, renorms, nonsmooth analysis, Ekeland's principle, proximal normals.

of all functions of the form

$$(2.1) \quad \Delta_p(x) := \sum_{n \in \mathbb{N}} \mu_n \|x - v_n\|^p, \quad \sum_{n \in \mathbb{N}} \mu_n = 1, \quad \mu_n \geq 0,$$

where $\{v_n\}$ converges in norm to some v in X . Thus Δ_p is a potentially infinite convex combination of translates of p th powers of norms, with the translates themselves converging.

DEFINITION 2.1. A *bornology* on X , denoted $\#$, is for us any nonempty family of bounded sets. We identify five particular bornologies. (a) G denotes the Gâteaux bornology consisting of all singletons; (b) H denotes the Hadamard bornology consisting of all norm compact sets; (c) WH denotes the weak Hadamard bornology consisting of all weakly compact sets; (d) F denotes the Fréchet bornology consisting of all bounded sets; (e) H, C denotes the bornology which adds all translates of a given weakly compact set C to H .

DEFINITION 2.2. A norm is $\#$ -smooth if it is Gâteaux differentiable away from the origin uniformly on members of $\#$. The Gâteaux derivative of the norm at x is denoted f_x .

DEFINITION 2.3. (a) Let X be a Banach space and let $f: X \rightarrow [-\infty, \infty]$ be lower semicontinuous. Suppose $f(x)$ is finite. Then f is $\#$ -subdifferentiable at x with subderivative ϕ in X^* if, for each $\varepsilon > 0$ and each set S in the $\#$ bornology, there exists $\delta > 0$ such that for $0 < t < \delta$

$$(2.2) \quad [f(x + th) - f(x)]t^{-1} - \phi(h) \geq -\varepsilon$$

uniformly for h in S . We write $\phi \in \partial^\# f(x)$.

(b) Superderivatives are similarly defined and denoted $\phi \in \partial_\# f(x)$. Thus $\partial_\# -f(x) = -\partial^\# f(x)$.

(c) Finally, f is $\#$ -differentiable if it is both $\#$ -subdifferentiable and $\#$ -superdifferentiable. The necessarily unique $\#$ -derivative is denoted $\nabla^\# f(x)$, and must coincide with $\nabla^G f(x)$. \square

Further information on subderivatives can be found in [5, 22] and the references therein.

PROPOSITION 2.4. (a) Suppose that f is concave and continuous in a neighborhood of x . If $\partial^\# f(x)$ is nonempty then f is $\#$ -differentiable at x and $\partial^\# f(x) = \{\nabla^\# f(x)\}$.

(b) If $p > 1$ and $\|\cdot\|$ is $\#$ -smooth then each function Δ_p is $\#$ -differentiable throughout X .

PROOF. (a) Since f is concave and continuous, the Hahn-Banach theorem provides a superderivative φ in the sense of convex analysis [7]. Then $\varphi \in \partial_\# f(x)$. If $\phi \in \partial^\# f(x)$, then as observed above $\varphi = \phi$ is the $\#$ -derivative of f at x .

(b) It follows directly from the Weierstrass M -test that Δ_p is Gâteaux differentiable everywhere with derivative

$$(2.3) \quad \nabla^G \Delta_p(x) = p \sum_{n \in \mathbb{N}} \mu_n \|x - v_n\|^{p-1} f_{x-v_n}.$$

Since $\{v_n\}$ is bounded and $\sum \mu_n$ is absolutely convergent, it is easily verified that Δ_p is actually $\#$ -differentiable at each x . \square

There is another class of differentiability properties inherited by Γ_p . If the norm is uniformly Fréchet differentiable on the unit sphere, then, for $p > 1$, Δ_p will be uniformly Fréchet differentiable at least on bounded sets. When X has a *power modulus of smoothness* [20] one can be more specific. By extending the argument in [15, Lemma 2.4], one may show that X has a power modulus of smoothness t^s ($1 < s \leq 2$) if and only if the function $g_s := \|\cdot\|^s/s$ satisfies

$$(2.4) \quad \|\nabla g_s(x) - \nabla g_s(y)\| \leq C\|x - y\|^{s-1}$$

for x and y in X , where C is independent of x and y . It follows that Δ_s will also satisfy (2.4) which is to say that Δ_s has a Hölder-continuous derivative. This holds in reflexive L_p spaces, in which case $s = \min\{2, p\}$. Indeed Pisier's deep renorming theorem [21] shows that every superreflexive space has a renorm with a power modulus of smoothness. This motivates the next definition.

DEFINITION 2.5. Let X be a Banach space, let $f: X \rightarrow [-\infty, \infty]$ be lower semicontinuous, and suppose $f(x)$ is finite. Then f is *s-Hölder-subdifferentiable* at x with subderivative ϕ in X^* if there exist positive constants δ_x and C_x such that

$$(2.5) \quad [f(x+h) - f(x)] - \phi(h) \geq -C_x\|h\|^{1+s}$$

whenever $\|h\| < \delta_x$. We write $\phi \in \partial^{\text{HS}} f(x)$ (or $\partial^{s\text{-HS}} f(x)$). When $s = 1$ such subderivatives are called *Lipschitz smooth*, written $\partial_{\text{LS}} f(x)$, and in Hilbert space they coincide with Rockafellar's *proximal subderivatives* [23].

THEOREM 2.6 (SMOOTH VARIATIONAL PRINCIPLE). *Let X be a Banach space, let $g: X \rightarrow (-\infty, \infty]$ be lower semicontinuous, and let constants $\varepsilon > 0$, $\lambda > 0$, and $p \geq 1$ be given. Suppose that x_0 satisfies*

$$(a) \quad g(x_0) < \varepsilon + \inf_X g.$$

Then there exist Δ_p in Γ_p and v in X such that for all x in X

$$(b) \quad g(x) + (\varepsilon/\lambda^p)\Delta_p(x) \geq g(v) + (\varepsilon/\lambda^p)\Delta_p(v),$$

while

$$(c) \quad \|x_0 - v\| < \lambda, \quad \text{and}$$

$$(d) \quad g(v) < \varepsilon + \inf_X g.$$

Moreover, if X has a $\#$ -smooth norm and $p > 1$ then

$$(e) \quad 0 \in \partial^\# g(v) + p(\varepsilon/\lambda)B^*,$$

where B^ is the dual unit ball. Finally, if X has a $(s+1)$ -power modulus of smoothness, (e) holds with s -HS replacing $\#$.*

PROOF. We iteratively construct Δ_p as follows. Use (a) to fix constants ε_1 and ε_2 with

$$(2.6) \quad g(x_0) - \inf_X g < \varepsilon_2 < \varepsilon_1 < \varepsilon$$

and now select μ with

$$(2.7) \quad 0 < \mu < 1 - (\varepsilon_1/\varepsilon).$$

Next select θ with

$$(2.8) \quad 0 < \theta/\mu < [1 - (\varepsilon_2/\varepsilon_1)^{1/p}]^p.$$

Finally, let $\delta := (1 - \mu)\varepsilon/\lambda^p$. Let $v_0 := x_0$ and $g_0 := g$ and recursively define sequences of functions $\{g_n\}$ and points $\{v_n\}$ by

$$(2.9) \quad g_{n+1}(x) := g_n(x) + \delta\mu^n\|x - v_n\|^p$$

where v_{n+1} is chosen so that

$$(2.10) \quad g_{n+1}(v_{n+1}) \leq \theta g_n(v_n) + (1 - \theta)\inf_X g_{n+1}.$$

Set $s_n := \inf_X g_n$ and $a_n := g_n(v_n)$. To see that (2.10) is possible observe that a_n also equals $g_{n+1}(v_n)$. Now θ is strictly between 0 and 1, and either $s_{n+1} < a_n$ or s_{n+1} is attained at v_n . Thus our recursion is possible. Use (2.10) to write

$$(2.11) \quad s_n \leq s_{n+1} \leq a_{n+1} \leq \theta a_n + (1 - \theta)s_{n+1} \leq a_n$$

so that

$$(2.12) \quad a_{n+1} - s_{n+1} \leq \theta(a_n - s_n) \leq \theta^{n+1}(a_0 - s_0).$$

Substitute $x := v_{n+1}$ in (2.9) to produce

$$a_n \geq a_{n+1} = g_n(v_{n+1}) + \delta\mu^n\|v_{n+1} - v_n\|^p \geq s_n + \delta\mu^n\|v_{n+1} - v_n\|^p$$

which with (2.6) and (2.12) shows that

$$(2.13) \quad \delta\mu^n\|v_{n+1} - v_n\|^p \leq \theta^n(a_0 - s_0) \leq \theta^n\varepsilon_2.$$

Thence (2.8) shows that $\{v_n\}$ is a Cauchy sequence and that for n and m in \mathbf{N}

$$\|v_m - v_n\| \leq (\varepsilon_2/\delta)^{1/p}/[1 - (\theta/\mu)^{1/p}] \leq (\varepsilon_1/\delta)^{1/p}.$$

Now (2.7) and the definition of δ imply that for n and m in \mathbf{N}

$$(2.14) \quad \|v_m - v_n\| \leq (\varepsilon_1/\delta)^{1/p} < (\varepsilon/\delta)^{1/p}(1 - \mu)^{1/p} = \lambda.$$

Letting v denote the limit of the sequence establishes (c). The desired member of Γ_p is defined by

$$(2.15) \quad \Delta_p(x) := \sum_{n \in \mathbf{N}} \mu_n\|x - v_n\|^p, \quad \mu_n := \mu^n(1 - \mu).$$

We now establish (b). For x in the domain of g we have

$$g(x) + (\varepsilon/\lambda^p)\Delta_p(x) = \sup_n g_n(x) \geq \lim_{n \rightarrow \infty} s_n.$$

Using (2.12) again gives

$$g(x) + (\varepsilon/\lambda^p)\Delta_p(x) \geq \lim_{n \rightarrow \infty} g_n(v_n) \geq \sup_m \liminf_{n \rightarrow \infty} g_m(v_n),$$

because $\{g_m\}$ increases with m . In addition, each g_m is lower semicontinuous. Thus

$$(2.16) \quad g(x) + (\varepsilon/\lambda^p)\Delta_p(x) \geq \sup_m g_m(v) = g(v) + (\varepsilon/\lambda^p)\Delta_p(v)$$

which is (b).

To establish (d) we first estimate $\Delta_p(x_0)$. Since $x_0 = v_0$, (2.14) and (2.15) imply that $\Delta_p(x_0) < \mu\lambda^p$. We then use (2.16) to derive

$$g(x_0) + \varepsilon\mu > g(x_0) + (\varepsilon/\lambda^p)\Delta_p(x_0) \geq g(v).$$

Then (2.6) and (2.7) show that

$$g(v) \leq \varepsilon_1 + \mu\varepsilon + \inf_x g < \varepsilon + \inf_x g.$$

We now suppose that $p > 1$ and that the norm is $\#$ -smooth. Then Proposition 2.4(b) shows Δ_p to be $\#$ -differentiable at v . Let $\phi := -(\varepsilon/\lambda^p)\nabla^\# \Delta_p(v)$. It follows from (b) and Definition 2.3 that ϕ is a $\#$ -subderivative of g at v ; and it is a consequence of (2.3) and (2.14) that $\|\phi\| \leq p\varepsilon/\lambda$. The Hölder-smooth case follows similarly from Definition 2.5 and the discussion preceding it. \square

REMARK 2.7. (a) One can formulate the theorem in any complete metric space, but the applications would all appear to lie in the normed setting. Various adaptations are apparent. For example one can force the sequence $\{\mu_n\}$ to zero as rapidly as wished. One may with some extra work often arrange for $v_n \neq v$, so that (e) remains valid with $p = 1$.

(b) When $p = 1$, Theorem 2.6 essentially recaptures Ekeland's principle [11] since $\Delta_1(x) - \Delta_1(v) \leq \|x - v\|$. The strict inequalities in (a) and (d) are needed for our general result.

(c) In Hilbert space we observe that $\Delta_2(x) = \|x - w\|^2$ for some w usually not equal to v . Thus in this setting we may directly establish (e) with $\#$ replaced by LS. \square

We finish this section with an easy application of the smooth variational principle.

COROLLARY 2.8. *Let X admit a $\#$ -smooth renorm and suppose that $f: X \rightarrow (-\infty, \infty]$ is somewhere finite and is lower semicontinuous. Suppose f is coercive in the sense that $f(x) \geq c(\|x\|)$ where c is continuous and $c(\|x\|)/\|x\|$ tends to infinity with $\|x\|$. Then $\partial^\# f$ has dense range.*

PROOF. Let φ be arbitrary in X^* and let $\varepsilon > 0$ be given. Let $g := f - \varphi$. As f is coercive g has a finite infimum. Apply Theorem 2.6 to g in the $\#$ -smooth equivalent renorm, with $p := 2$ and $\lambda := 2$. Since $\partial^\# f(v) = \partial^\# g(v) + \varphi$, part (e) completes the proof. \square

3. The existence of subderivatives. In light of the results of the previous section it is meaningful to consider $\#$ -subderivatives for any derivative property which Γ_p inherits from an appropriate norm. Thus where convenient we no longer notationally distinguish the two previous classes of Definitions 2.3 and 2.5 and talk about $\#$ -derivatives and subderivatives in both settings.

THEOREM 3.1. *Let X be a Banach space with a $\#$ -smooth renorm and let $f: X \rightarrow [-\infty, \infty]$ be lower semicontinuous. Then f is $\#$ -subdifferentiable at a dense subset of points in its graph.*

PROOF. Let $\varepsilon > 0$ and x_0 with $f(x_0)$ finite be given. Since f is lower semicontinuous one can choose $\varepsilon > \lambda > 0$ such that

$$(3.1) \quad \inf\{f(x): \|x - x_0\| \leq \lambda\} > f(x_0) - \varepsilon.$$

We apply Theorem 2.6 to $f_\lambda := f + \delta_\lambda$ where δ_λ is the indicator function of $\{x: \|x - x_0\| \leq \lambda\}$. Then f_λ is lower semicontinuous and (3.1) shows that (a) of Theorem 2.6 holds for f_λ . We apply the theorem with this ε and λ and $p := 2$.

This produces a point v with $\|v - x_0\| < \lambda$ such that $\partial^\# f_\lambda(v)$ is nonempty. Since f_λ and f agree in a neighborhood of v (of positive radius $\lambda - \|v - x_0\|$) and since $\#$ -subderivatives are local objects, $\partial^\# f(v)$ is nonempty. Moreover, conclusion (d) of the principle and (3.1) yield

$$f(x_0) - \varepsilon \leq f(v) = f_\lambda(v) \leq \inf_X f_\lambda + \varepsilon \leq f(x_0) + \varepsilon.$$

Thus we have produced a point $(v, f(v))$ in the graph of f , which is arbitrarily close to the original point, where $\partial^\# f(v)$ is nonempty. \square

One can use this theorem to derive very precise subgradient results for lower semicontinuous functions, and associated tangent cone formulae [6]. Here we develop only a mean value estimate for locally Lipschitz functions. We define a *generalized $\#$ -derivative set* for f at v by

$$(3.2) \quad D^\# f(v) := \{\varphi: \varphi_n \xrightarrow{w^*} \varphi, x_n \rightarrow v, \varphi_n \in \partial^\# f(x_n)\}.$$

Note that subsequential limits exist since f is locally Lipschitz and since the dual ball is w^* sequentially compact (because X admits a G -smooth renorm, see Larman and Phelps [18]).

THEOREM 3.2. *Let X be a Banach space with a $\#$ -smooth renorm and let $f: U \subset X \rightarrow \mathbf{R}$ be locally Lipschitz on the open set U . Let the closed interval $[a, b]$ lie in U . Then there exists c in the open interval (a, b) and some φ in $D^\# f(c)$ with*

$$(3.3) \quad f(b) - f(a) \leq \varphi(b - a).$$

Also, for any v in U

$$(3.4) \quad \partial f(v) = w^* \text{cl conv } D^\# f(v)$$

where $\partial f(v)$ is the Clarke subgradient of f at v .

PROOF. By subtracting an affine functional we may suppose that $f(b) = f(a) = 0$. Let $c \neq b$ attain the infimum, s , of f over the closed interval $[a, b]$. Let δ_N be the indicator function of any neighborhood of $[a, b]$ on which f is bounded below. Let n be any positive integer and select K_n sufficiently large so that for $g_n := f + \delta_N + K_n d_{[a,b]}^2$

$$f(c) - 1/n^2 = g_n(c) - 1/n^2 < \inf_x g_n$$

where $d_{[a,b]}(x) := \min\{\|x - y\|: y \in [a, b]\}$. We apply the smooth variational principle to g_n with $x_0 := c$, with $p := 2$ and with $\varepsilon/\lambda := \lambda := 1/n$. Then we obtain a sequence $\{v_n\}$ converging to c such that $0 \in \partial^\# g_n(v_n) + (2/n)B^*$. The function $d_{[a,b]}^2$ is $\#$ -differentiable since the norm is. If v_n lies in $[a, b]$ infinitely often then $0 \in \partial^\# f(v_n) + (2/n)B^*$, which implies that $0 \in D^\# f(c)$. Otherwise, we may assume that v_n does not lie in $[a, b]$ and so for some positive constant M_n

$$(3.5) \quad 0 \in \partial^\# f(v_n) + (2/n)B^* + M_n \nabla^\# d_{[a,b]}(v_n).$$

In this case, elementary convex analysis shows that $\zeta_n := \nabla^\# d_{[a,b]}(v_n)$ satisfies $\zeta_n(x - c_n) \leq 0$ for all x in $[a, b]$, where c_n is any nearest point to v_n in $[a, b]$ and

so converges to c . Since f is locally Lipschitz on U , (3.5) now allows us to extract a w^* convergent subsequence with limit φ in $D^\# f(c)$ such that $\varphi(x - c) \geq 0$ for all x in $[a, b]$. Since $c \neq b$, $\varphi(b - a) \geq 0$. Thus in both cases we establish (3.3) except that c may equal a or b . We next establish (3.4). Fix h in X and, using the definition of the Clarke derivative, select v_n converging to v and t_n decreasing to 0 such that

$$[f(v_n + t_n h) - f(v_n)]t_n^{-1} \rightarrow f^0(v; h).$$

The previously established weak form of (3.3) applied on each interval $[v_n, v_n + t_n h]$ ultimately produces a subsequence converging w^* to some φ in $D^\# f(v)$ with $\varphi(h) \geq f^0(v; h)$. Since every Gâteaux-subderivative is a Clarke subderivative, and since ∂f is norm- w^* closed we have established that

$$(3.6) \quad f^0(v; h) = \max\{\varphi(h) : \varphi \in D^\# f(v)\}.$$

This establishes (3.4). Moreover, Lebourg's mean value theorem [7] implies that (3.3) holds for some φ in $\partial f(c)$ and c in (a, b) . Finally, (3.6) allows us to replace $\partial f(c)$ by $D^\# f(c)$. \square

Recent more difficult results of Preiss [22] allows one to replace $\partial^\# f(x)$ by $\nabla^\# f(x)$ in the definition of $D^\# f(c)$.

4. Convex differentiability spaces. We now collect up the implications for convex functions.

THEOREM 4.1. *Let X be a Banach space with a $\#$ -smooth renorm and let $f: U \subset X \rightarrow \mathbf{R}$ be convex and continuous on the open set U . (a) Then f is densely $\#$ -differentiable in U . (b) In addition for v in U*

$$(4.1) \quad \partial f(v) = w^* \text{ cl conv } D^\# f(v)$$

where now

$$D^\# f(v) := \{\varphi : \varphi_n \xrightarrow{w^*} \varphi, x_n \rightarrow v, \varphi_n \in \nabla^\# f(x_n)\}.$$

PROOF. (a) follows directly from Theorem 3.1 and Proposition 2.4(a) applied to $-f$. Part (b) follows similarly from Theorem 3.2. \square

REMARK 4.2. (a) Even on the real line Theorem 3.1 guarantees only a dense set of subderivatives; not a set residual in category or measure. Otherwise, every continuous function would be somewhere differentiable. Similarly Theorem 4.1 only asserts the existence of a dense set; an example in [14] exhibits a convex function on the real line for which the Lipschitz smooth points contain no residual set.

(b) It is elementary that the Fréchet differentiable points of a continuous convex function form a (dense) G_δ set, and we recover Ekeland and Lebourg's celebrated result [13] that a space with a Fréchet renorm is an Asplund space.

(c) The Lipschitz-smooth case of Theorem 4.1 may be found in Fabian [14]. The Hölder-smooth, weak Hadamard, and Gâteaux (or Hadamard) results appear new.

(d) The discussion before Definition 2.5 shows that every superreflexive space may be renormed so that (2.4) holds for some $\rho(t) := t^{s-1}$. Thus in any superreflexive space every continuous convex function is densely Fréchet differentiable at a Hölder rate ($K\rho, K$ varying with the point). Equally in any superreflexive space every lower semicontinuous function is densely Hölder-subdifferentiable.

(e) A similar result holds for biconvex functions. \square

It is instructive to recast Theorem 4.1 as a negative result on the existence of $\#$ -smooth renorms.

COROLLARY 4.3. *Suppose that X admits a continuous convex function which is nowhere $\#$ -differentiable. Then X has no equivalent $\#$ -smooth renorm.*

The Gâteaux case of the corollary may be viewed as completing Leach and Whitfield's result that a space with a smooth norm admits no strongly rough norm [19] and Sullivan's extensions [24].

EXAMPLE 4.4. (a) The fact that $f(x) := \limsup_{n \rightarrow \infty} |x_n|$ viewed as a convex function on $l_\infty(\mathbf{N})$ has no smooth points now becomes an easy proof that $l_\infty(\mathbf{N})$ (or l_∞/c_0) has no Gâteaux renorm. Similarly the standard facts that the usual norm on $l_1(\mathbf{R})$ (as any non- σ -finite L_1) or $L_\infty[0, 1]$ has no smooth points shows that the space has no Gâteaux renorm (see Larman and Phelps [18]).

(b) $C[0, 1]$ has no weak Hadamard renorm. Indeed it was shown in [5] that the supremum norm is nowhere weak Hadamard smooth. This holds for $[0, 1]$ replaced by any perfect compact metric space.

(c) Coban and Kenderov [8] show that the supremum norm on $C[\beta(\mathbf{Q})]$ is densely but not generically Gâteaux differentiable. Here $\beta(\mathbf{Q})$ is the Čech compactification of the rationals. \square

Note also that (4.1), applied at 0 to the support function of a weak-star closed bounded convex set in X^* , shows that such a set is the weak-star closed convex hull of $\#$ -exposed points (appropriately defined).

5. Subderivatives of distance functions. Let C be a closed subset of a Banach space X and let $d_C(x) := \inf\{\|x - c\| : c \in C\}$ denote the metric distance function. Recall that a norm has the *Kadec* property if the weak and norm topologies agree on the unit sphere.

THEOREM 5.1. (a) *Suppose that the norm on X is strictly convex and X admits a smooth renorm. Then there is a dense set of points D in X such that each x in D has at most one nearest point in C .*

(b) *Suppose that X is weakly compactly generated, that C is boundedly relatively weakly compact and that the norm on X is Kadec. Then there is a dense set of points D in X such that each x in D has at least one nearest point in C .*

PROOF. (a) Since X admits a smooth renorm and d_C is Lipschitz, Theorem 3.1 applies. Let D be a dense subset of $X \setminus C$ where d_C has Gâteaux subderivatives. Let x lie in D and let $\varphi \in \partial^G d_C(x)$. Suppose that x has two nearest points p_1 and p_2 in C . Then for $i = 1$ or 2

$$\begin{aligned} \varphi(p_i - x) &\leq \liminf_{t \rightarrow 0} [d_C(x + t(p_i - x)) - d_C(x)]/t^{-1} \\ &= \liminf_{t \rightarrow 0} [d_C(x + t(p_i - x)) - \|p_i - x\|]/t^{-1} \\ &\leq \liminf_{t \rightarrow 0} [\|x + t(p_i - x) - p_i\| - \|p_i - x\|]/t^{-1} \\ &= -\|p_i - x\|. \end{aligned}$$

But $\|\varphi\| \leq 1$ as d_C is nonexpansive. Thus

$$\varphi(x - p_i) = \|p_i - x\| = d_C(x)$$

and strict convexity of the norm implies that $p_1 = p_2$.

(b) Observe first that we may assume that C is bounded. Thus X , being WCG, admits a Gâteaux renorm which is uniform on C [1]. It is easy to verify that the

norm is H, C smooth in the sense of Definition 2.3(e). Let D be a dense subset of $X \setminus C$ where d_C has H, C subderivatives. Let x lie in D and let $\varphi \in \partial^{H, C} d_C(x)$. Let $\{c_n\}$ lie in C with $\|c_n - x\| \leq d_C(x) + 1/n^2$. Much as in (a) for n large, since each $c_n - x$ is in $C - x$,

$$\begin{aligned} \varphi(c_n - x) &\leq n[d_C(x + (c_n - x)/n) - d_C(x)] + \varepsilon/2 \\ &\leq n[d_C(x + (c_n - x)/n) - \|c_n - x\|] + \varepsilon \\ &\leq n[\|x - c_n + (c_n - x)/n\| - \|c_n - x\|] + \varepsilon \\ &= \varepsilon - \|c_n - x\| \end{aligned}$$

and

$$\liminf_{n \rightarrow \infty} \varphi(x - c_n) \geq d_C(x).$$

Thus, as $\|\varphi\| \leq 1$

$$(5.1) \quad \lim_{n \rightarrow \infty} \varphi(x - c_n) = d_C(x).$$

Since C is bounded and relatively weakly compact we may assume that c_n converges weakly to some point p . Then (5.1) shows that

$$\|x - c_n\| \rightarrow \|x - p\| = d_C(x).$$

Finally the Kadec property ensures that $x - c_n$ converges in norm to $x - p$. Thus p lies in C and x has a nearest point in C . \square

Part (a) partially extends a result in [16] and (b) recaptures much more simply most of Theorem 4.1 in [1]. More on the properties of derivatives and subderivatives of distance functions may be found in [3] where the reflexive case of Theorem 5.1 is given. Note that the span of a boundedly relatively weakly compact set is always a WCG space.

The ideas behind Theorem 5.1 can be used to show that in any reflexive space Theorem 2.6 holds for very simple Δ_p . This was observed above in Hilbert space. Precisely we have

THEOREM 5.2. *Let X be a reflexive space and let $\|\cdot\|$ be any Kadec renorm on X . Then Theorem 2.6 holds with*

$$(5.2) \quad \Delta_p(x) := \|x - w\|^p.$$

PROOF. Consider the function h defined by

$$(5.3) \quad h(w) := \inf\{g(x) + (\varepsilon/\lambda^p)\|x - w\|^p : \|x - x_0\| \leq 2\lambda\}.$$

Theorem 11 in Borwein and Giles [3] shows that since x is restricted to a bounded set there is a dense set of w such that $h(w)$ is attained by some v (because h admits a Fréchet-subderivative at w). With ε_1 and ε_2 as in (2.6), select such a w with

- (i) $\|x_0 - w\| \leq \lambda[1 - (\varepsilon_1/\varepsilon)^{1/p}]$.
- (ii) $\|x_0 - w\| \leq \lambda[(\varepsilon_1 - \varepsilon_2)/\varepsilon]^{1/p}$.

Now (i) implies that

$$(5.4) \quad \|x_0 - w\|^p + \lambda^p \left\{ g(x_0) - \inf_x g \right\} / \varepsilon \leq \lambda^p.$$

Then clearly $\|x_0 - w\| \leq \lambda$. Now if $\|x_1 - x_0\| \geq 2\lambda$ we see that $\|x_1 - w\| \geq \lambda$ and

$$g(x_1) + (\varepsilon/\lambda^p)\|x_1 - w\|^p \geq \varepsilon + \inf_x g$$

while (5.4) shows that

$$(5.5) \quad h(w) \leq g(x_0) + (\varepsilon/\lambda^p)\|x_0 - w\|^p < \varepsilon + \inf_x g.$$

Thus

$$(5.6) \quad h(w) = \inf\{g(x) + (\varepsilon/\lambda^p)\|x - w\|^p\}.$$

This and the existence of a minimizing v establish (b) of Theorem 2.6. Also (5.5) proves (d) and it remains to prove (c). We observe that $\|v - w\| < \lambda(\varepsilon_1/\varepsilon)^{1/p}$. Otherwise

$$g(v) + (\varepsilon/\lambda^p)\|v - w\|^p \geq \inf_x g + \varepsilon_1 \geq g(x_0) + \varepsilon_1 - \varepsilon_2$$

while

$$g(x_0) + \varepsilon_1 - \varepsilon_2 > g(x_0) + (\varepsilon/\lambda^p)\|x_0 - w\|^p$$

by the second condition in the choice of w . This and the first condition in the choice of w yield

$$\|v - x_0\| < \|w - x_0\| + \|v - w\| \leq \lambda[1 - (\varepsilon_1/\varepsilon)^{1/p}] + \lambda(\varepsilon_1/\varepsilon)^{1/p} = \lambda. \quad \square$$

Recall that every reflexive space admits a Fréchet and Kadec renorm, and every superreflexive space admits a uniformly Fréchet and Kadec renorm [10]. In particular Theorem 5.2 holds in each reflexive L_p norm and we may rederive the Hölder nature of subderivatives theorem. Finally, Theorem 5.2 only holds in reflexive space.

PROPOSITION 5.3. *If Theorem 2.6 holds with Δ_p given by (5.2) and with $p > 1$ then X is reflexive.*

PROOF. Let f be any norm one continuous linear functional and apply the result to $\inf_B f$ where B is the unit ball in X , with $x_0 := 0$, $\lambda := 1$, and $\varepsilon := 2$. This ensures the existence of w and v with $\|v\| < 1$ such that

$$f(x) + 2\|x - w\|^p \geq f(v) + 2\|v - w\|^p$$

for all x in B . Since $\|v\| < 1$ we must have $0 \in f + 2\partial\|v - w\|^p$ and as f is not the zero functional, $v \neq w$ and f attains its norm in the direction of $v - w$. By James' theorem [10] X is reflexive. \square

In c_0 any finite sum of norm-attaining functions is norm-attaining, and the same argument shows that Theorem 2.6 generally does not hold if only finite sums are admitted.

REFERENCES

1. J. M. Borwein, *Weak local supportability and applications to approximation*, Pacific J. Math. **82** (1979), 323-338.
2. —, *Stability and regular points of inequality systems*, J. Optim. Theory Appl. **48** (1986), 9-52.
3. J. M. Borwein and J. R. Giles, *The proximal normal formula in Banach space*, Trans. Amer. Math. Soc. **302** (1987), 371-381.
4. J. M. Borwein and H. M. Strojwas, *Proximal analysis and boundaries of closed sets in Banach space. Part 1: Theory*, Canad. Math. J. **38** (1986), 431-452.
5. —, *Proximal analysis and boundaries of closed sets in Banach space, Part 2: Applications*, Canad. Math. J. (to appear).
6. —, *Subderivatives and nonsmooth analysis* (to appear).

7. F. H. Clarke, *Optimization and nonsmooth analysis*, Canadian Math. Soc. Series, Wiley, 1983.
8. M. Coban and P. S. Kenderov, *Dense Gateaux differentiability of the sup-norm in $C(T)$ and the topological properties of T* , C.R. Acad. Bulgare Sci. **38** (1985), 1603–1604.
9. M. M. Day, *Normed linear spaces*, 3rd ed., Springer-Verlag, 1973.
10. J. Diestel, *Geometry of Banach spaces—Selected topics*, Lecture Notes in Math., vol. 485, Springer-Verlag, 1975.
11. I. Ekeland, *On the variational principle*, J. Math. Anal. Appl. **47** (1974), 324–353.
12. ———, *Nonconvex minimization problems*, Bull. Amer. Math. Soc. (N.S.) **1** (1979), 443–474.
13. I. Ekeland and G. Lebourg, *Generic Fréchet differentiability and perturbed optimization problems in Banach spaces*, Trans. Amer. Math. Soc. **224** (1976), 193–216.
14. M. Fabian, *Lipschitz smooth points of convex functions and isomorphic characterizations of Hilbert spaces*, Proc. London Math. Soc. **51** (1985), 113–126.
15. M. Fabian, J. H. M. Whitfield and V. Zizler, *Norms with locally Lipschitz derivatives*, Israel J. Math. **44** (1983), 262–276.
16. M. Fabian and N. V. Zhivkov, *A characterization of Asplund spaces with the help of local ε -supports of Ekeland and Lebourg*, C.R. Acad. Bulgare Sci. **38** (1985), 671–674.
17. J. R. Giles, *Convex analysis with application in differentiation of convex functions*, Pitman Research Notes in Math., 58, Pitman, 1982.
18. D. G. Larman and R. R. Phelps, *Gateaux differentiability of convex functions on Banach spaces*, J. London Math. Soc. **20** (1979), 115–127.
19. E. B. Leach and J. H. M. Whitfield, *Differentiable functions and rough norms on Banach spaces*, Proc. Amer. Math. Soc. **33** (1972), 120–126.
20. J. Lindenstrauss and L. Tzafriri, *Classical Banach spaces. II. Function spaces*, Springer-Verlag, 1978.
21. G. Pisier, *Martingales with values in uniformly convex spaces*, Israel J. Math. **20** (1975), 326–350.
22. D. Preiss, *Fréchet derivatives of Lipschitz functions* (to appear).
23. R. T. Rockafellar, *Proximal subgradients, marginal values and augmented Lagrangians in non-convex optimization*, Math. Oper. Res. **6** (1981), 424–436.
24. F. Sullivan, *Nearly smooth norms on Banach spaces*, Rev. Roumaine Math. Pures Appl. **21** (1981), 1053–1057.
25. J. S. Treiman, *Clarke's gradients and epsilon-subgradients in Banach spaces*, Trans. Amer. Math. Soc. **294** (1986), 66–78.

DEPARTMENT OF MATHEMATICS, DALHOUSIE UNIVERSITY, HALIFAX, NOVA SCOTIA,
B3H 3J5, CANADA

DEPARTMENT OF MATHEMATICS, CHARLES UNIVERSITY, PRAGUE, CZECHOSLOVAKIA