Canad. J. Math. Vol. **53** (6), 2001 pp. 1174–1193

A Generalized Variational Principle

Philip D. Loewen and Xianfu Wang

Abstract. We prove a strong variant of the Borwein-Preiss variational principle, and show that on Asplund spaces, Stegall's variational principle follows from it via a generalized Smulyan test. Applications are discussed.

1 Introduction

The Borwein-Preiss smooth variational principle [1] is an important tool in infinite dimensional nonsmooth analysis. Its statement, given a Banach space X, reads as follows.

Theorem 1.1 *Given $f\colon X \to (-\infty, +\infty]$ lower semicontinuous, $x_0 \in X$, $\varepsilon > 0$, $\lambda > 0$, and $p \geq 1$, suppose*

$$f(x_0) < \varepsilon + \inf_X f.$$

Then there exist a sequence $\mu_n \geq 0$, with $\sum_{n=1}^{\infty} \mu_n = 1$, and a point v in X, expressible as the (norm-) limit of some sequence (v_n), such that

$$f(x) + \frac{\varepsilon}{\lambda^p} \triangle_p(x) \geq f(v) + \frac{\varepsilon}{\lambda^p} \triangle_p(v) \quad \forall x \in X, \tag{1.1}$$

where $\triangle_p(x) := \sum_{n=1}^{\infty} \mu_n \|x - v_n\|^p$. Moreover, $\|x_0 - v\| < \lambda$, and $f(v) \leq \varepsilon + \inf_X f$. If, in addition, X has a β-smooth norm and $p > 1$, then $\partial^\beta f(v) \cap (\varepsilon/\lambda) p \mathbb{B}^* \neq \varnothing$.

In a subsequent development [4], Deville, Godefroy and Zizler improved Theorem 1.1 by considering a certain Banach space Y of bounded continuous functions $g\colon X \to \mathbb{R}$, and showing that the following subset of Y is residual:

$$\{g \in Y : f + g \text{ attains a strong minimum somewhere in } X\}.$$

(See Definition 2.1.) However, their result gives no information about the location of the strong minimizer, and offers no way to identify explicitly a perturbation g with the desired property. In [12], [2], Li and Shi give a simpler proof of Theorem 1.1.

Received by the editors March 3, 2001; revised July 10, 2001.
The first author's research was supported by NSERC. The second author's research was supported by an NSERC post-doctoral fellowship.
AMS subject classification: 49J52.
Keywords: variational principle, strong minimizer, generalized Smulyan test, Asplund space, dimple point, porosity.

https://doi.org/10.4153/CJM-2001-044-8 Published online by Cambridge University Press

In this paper, we adapt the Borwein-Preiss approach by adjusting the penalty terms to control the diameter of the resulting sublevel sets, and thus obtain a variational principle for which strong minimality holds in the analogue of (1.1). This allows us to present unified proofs of the variational principles of Ekeland [6], Borwein-Preiss [1], and Deville-Godefroy-Zizler [4]. By combining these methods with a generalized Smulyan test, we also give a simple proof of Stegall's variational principle [15], [7] on Asplund spaces. We then apply these variational principles to characterize Banach spaces and study subdifferentiability.

Notation Throughout X denotes a Banach space with closed unit ball $\mathbb{B}$, dual X^*, and closed dual ball $\mathbb{B}^*$. We write $\mathbb{B}_r(x) := \{y \in X : \|y - x\| < r\}$ and $\mathbb{B}_r[x]$ for the closure of $\mathbb{B}_r(x)$. For a set $S \subset X$ and $f \colon X \to (-\infty, +\infty]$ we let

$$I_S(x) := \begin{cases} 0, & \text{if } x \in S, \\ +\infty, & \text{if } x \notin S, \end{cases} \qquad f_S(x) := \begin{cases} f(x), & \text{if } x \in S, \\ +\infty, & \text{if } x \notin S. \end{cases}$$

We abbreviate lower semicontinuous by lsc. Our symbol for a typical bornology on X is β.

Subgradients When $f \colon X \to \mathbb{R} \cup \{+\infty\}$ is lsc and $f(x)$ is finite, the set $\partial^\beta f(x)$ consists of all those $x^* \in X^*$ for which

$$\liminf_{t \downarrow 0} \left[\inf_{h \in S} \frac{f(x + th) - f(x) - \langle x^*, th \rangle}{t} \right] \geq 0, \quad \forall S \in \beta.$$

We say f *is* β*-subdifferentiable at* x exactly when $\partial^\beta f(x) \neq \varnothing$; each $x^* \in \partial^\beta f(x)$ is called a β-subderivative of f at x. To say f *is* β*-differentiable at* x, with β-derivative $\nabla f(x)$, means that $x^* = \nabla f(x)$ is a point of X^* such that

$$\lim_{t \downarrow 0} \sup_{h \in S} \left| \frac{f(x + th) - f(x) - \langle x^*, h \rangle}{t} \right| = 0, \quad \forall S \in \beta. \tag{1.2}$$

The *Dini subdifferential*, corresponding to the bornology of all compact subsets of X, can be characterized as

$$\partial^- f(x) := \{x^* : \langle x^*, v \rangle \leq f^-(x; v) \text{ for every } v \in X\},$$

where $f^-(x; v) := \liminf_{t \downarrow 0, h \to v} t^{-1}[f(x + th) - f(x)]$. The *approximate subdifferential* of f at x is defined by

$$\partial_a f(x) := \bigcap_{L \in F} \limsup_{u \to_f x} \partial^- f_{u+L}(u),$$

where F denotes the collection of all finite dimensional subspaces of X and lim sup denotes the collection of weak* limits of converging subnets. It follows from the definition that $\partial_a f(x)$ is weak*-closed and that $\partial_a f(x) = \limsup_{u \to_f x} \partial_a f(u)$ [11, Prop. 2.3].

2 A Strong Variant of the Borwein-Preiss Variational Principle

Definition 2.1 For the function $f\colon X \to \mathbb{R} \cup \{+\infty\}$, a point $x_0 \in X$ is a *strong minimizer* if $f(x_0) = \inf_X f$ and every sequence (x_n) along which $f(x_n) \to \inf_X f$ obeys $\|x_n - x_0\| \to 0$.

A strong minimizer is a strict minimizer, that is, $f(x) > f(x_0)$ for every $x \neq x_0$, but the converse is false. (Take $X = \mathbb{R}$, $x_0 = 0$, $f(x) = x^2 e^x$.) For general f, we write

$$(2.1)\quad \Sigma_\varepsilon(f) = \Big\{x \in X : f(x) \leq \varepsilon + \inf_X f\Big\}, \quad \{f \leq \alpha\} = \{x \in X : f(x) \leq \alpha\}.$$

Evidently $\Sigma_\varepsilon(f) \neq \varnothing$ whenever $\varepsilon > 0$. It is easy to show that f attains a strong minimum on X if and only if

$$(2.2)\qquad \inf\Big\{\operatorname{diam}\big(\Sigma_\varepsilon(f)\big) : \varepsilon > 0\Big\} = 0,$$

where 'diam' denotes the norm-diameter of the indicated set. In particular, a necessary and sufficient condition for f to have *no strong minima* on X is

$$(2.3)\qquad \exists\alpha > 0 : \forall\varepsilon > 0, \quad \operatorname{diam}\big(\Sigma_\varepsilon(f)\big) > \alpha.$$

Our main result builds a perturbation by shifting and scaling a given continuous function $\rho\colon X \to [0, +\infty)$, on which our only hypotheses are

$$(2.4)\qquad \rho(0) = 0 \quad \text{and} \quad \eta := \sup\{\|x\| : \rho(x) < 1\} < +\infty.$$

(The choice $\rho(x) = \|x\|^p$ has the required properties, with $\eta = 1$, for any $p > 0$.) The penalty functions built from ρ have the form

$$(2.5)\qquad \rho_\infty(x) := \sum_{n=0}^{\infty} \rho_n(x - v_n), \quad \text{where} \quad \rho_n(x) := \mu_n \rho\big((n+1)x\big),$$

for scalars $\mu_n \in (0, 1)$ and vectors $v_n \in X$ to be specified below.

Theorem 2.2 *Given $f\colon X \to (-\infty, +\infty]$ lsc, $x_0 \in X$, and $\varepsilon > 0$, suppose*

$$f(x_0) < \varepsilon + \inf_X f.$$

Then for any continuous ρ obeying (2.4), and any decreasing sequence (μ_n) in $(0, 1)$ with $\sum \mu_n < +\infty$, there exist a (norm-) convergent sequence v_n in X and a function ρ_∞ of the form (2.5) such that $v = \lim_n v_n$ satisfies

(i) $\rho(x_0 - v) < 1$,
(ii) $f(v) + \varepsilon\rho_\infty(v) \leq f(x_0)$, *and*
(iii) *v is a strong minimizer for $f + \varepsilon\rho_\infty$.*

In particular,

$$f(v) + \varepsilon\rho_\infty(v) < f(x) + \varepsilon\rho_\infty(x) \quad \forall x \in X \setminus \{v\}\,. \tag{2.6}$$

Proof Let $v_0 = x_0$, $f_0 = f$. For integers $n \geq 0$, use induction to choose v_{n+1} and define f_{n+1}, D_n as follows:

$$f_{n+1}(x) := f_n(x) + \varepsilon\rho_n(x - v_n), \tag{2.7}$$

$$f_{n+1}(v_{n+1}) \leq \frac{\mu_{n+1}}{2} f_n(v_n) + \Big(1 - \frac{\mu_{n+1}}{2}\Big) \inf_X f_{n+1} \leq f_n(v_n), \tag{2.8}$$

$$D_n := \Big\{x : f_{n+1}(x) \leq f_{n+1}(v_{n+1}) + \frac{\mu_n\varepsilon}{2}\Big\}\,. \tag{2.9}$$

To justify (2.8), note that $\inf_X f_{n+1} \leq f_{n+1}(v_n) = f_n(v_n)$. If this inequality is strict, the existence of v_{n+1} follows from the definition of the infimum; if equality holds instead, the choice $v_{n+1} = v_n$ will serve.

Notice that $f_{n+1} \geq f_n$ and that f_{n+1} is lsc. Hence D_n is closed; also $D_n \neq \varnothing$, because $v_{n+1} \in D_n$. Since $0 < \mu_{n+1} < 1$, (2.8) implies

$$f_{n+1}(v_{n+1}) - \inf_X f_{n+1} \leq \frac{\mu_{n+1}}{2}[f_n(v_n) - \inf_X f_{n+1}] \leq f_n(v_n) - \inf_X f_n. \tag{2.10}$$

Note that $f_0(v_0) - \inf_X f_0 = f(x_0) - \inf_X f < \varepsilon$ by hypothesis.

Claim 1 $D_n \subset D_{n-1}$ for all $n \geq 1$. Indeed, pick $x \in D_n$. Then $\mu_{n-1} > \mu_n$ and (2.8) imply

$$f_n(x) \leq f_{n+1}(x) \leq f_{n+1}(v_{n+1}) + \frac{\mu_n\varepsilon}{2} \leq f_n(v_n) + \frac{\mu_{n-1}\varepsilon}{2},$$

so $x \in D_{n-1}$.

Claim 2 $\operatorname{diam}(D_n) \to 0$ as $n \to \infty$. Indeed, since $f_{n-1} \leq f_n$, case $n-1$ of (2.8) implies

$$f_n(v_n) - \inf_X f_n \leq \frac{\mu_n}{2}\big(f_{n-1}(v_{n-1}) - \inf_X f_n\big) \leq \frac{\mu_n}{2}\big(f_{n-1}(v_{n-1}) - \inf_X f_{n-1}\big) < \frac{\mu_n\varepsilon}{2}. \tag{2.11}$$

(The last inequality follows from (2.10).) For every $x \in D_n$, the definitions of D_n and f_{n+1} give

$$\varepsilon\mu_n\rho\big((n+1)(x - v_n)\big) \leq f_{n+1}(v_{n+1}) - f_n(x) + \frac{\mu_n\varepsilon}{2} \leq f_{n+1}(v_{n+1}) - \inf_X f_n + \frac{\mu_n\varepsilon}{2}. \tag{2.12}$$

Now $f_{n+1}(v_{n+1}) \leq f_n(v_n)$ from (2.8), so (2.11) gives an upper bound of $\mu_n\varepsilon$ for the right side of (2.12). This implies $\rho\big((n+1)(x - v_n)\big) < 1$ for all $n \geq 0$, so by (2.4),

$$(n+1)\|x - v_n\| \leq \eta. \tag{2.13}$$

https://doi.org/10.4153/CJM-2001-044-8 Published online by Cambridge University Press

It follows that $\operatorname{diam}(D_n) \le 2\eta/(n+1)$.

Claim 2 implies that $\bigcap_{n=0}^{\infty} D_n$ contains exactly one point. Call this point v, and note that $v_n \to v$ as $n \to \infty$. Taking $n = 0$ in the phrase before (2.13) gives $\rho(v - x_0) < 1$.

Claim 3 $f(v) + \varepsilon\rho_\infty(v) \le f_n(v_n)$ for every n. Indeed, since $f_{n+1} \ge f_n$ while $f_{n+1}(v_{n+1}) \le f_n(v_n)$ for all n, the sequence of nonempty closed sets

$$\widetilde{D}_n := \{x : f_{n+1}(x) \le f_{n+1}(v_{n+1})\}$$

is nested and obeys $\widetilde{D}_n \subset D_n$ for all n. Thus $\bigcap_{n=1}^{\infty} \widetilde{D}_n = \{v\}$. In particular, by (2.8),

$$f_k(v) \le f_k(v_k) \le f_n(v_n) \le f(x_0) \quad \forall k > n.$$

As $k \to \infty$ we get $f(v) + \varepsilon\rho_\infty(v) \le f_n(v_n) \le f(x_0)$.

Claim 4 v is a strong minimizer for $\widetilde{f} := f + \varepsilon\rho_\infty$. Indeed, if $\widetilde{f}(x) \le \inf_X \widetilde{f} + \mu_n\varepsilon/2$, then by the last line of Claim 3,

$$f_{n+1}(x) \le \widetilde{f}(x) \le \widetilde{f}(v) + \frac{\mu_n\varepsilon}{2} \le f_{n+1}(v_{n+1}) + \frac{\mu_n\varepsilon}{2}.$$

Hence $\Sigma_{\mu_n\varepsilon/2}(\widetilde{f}) \subset D_n$. Claim 2 implies

$$\lim_{n\to\infty} \operatorname{diam}\left(\Sigma_{\mu_n\varepsilon/2}(\widetilde{f})\right) = 0,$$

whence v is a strong minimizer of $\widetilde{f} = f + \varepsilon\rho_\infty$. ■

A Strong Banach Space Variant of Ekeland's Theorem Applying Theorem 2.2 with

$$\rho(x) = \frac{\|x\|}{\lambda}, \quad \mu_n = \frac{1}{2^{n+1}(n+1)},$$

produces a sequence (v_n) for which (2.6) can be rearranged as follows: for any $x \ne v$,

$$\begin{aligned} f(v) &< f(x) + \varepsilon[\rho_\infty(x) - \rho_\infty(v)] \\ &= f(x) + \frac{\varepsilon}{\lambda}\sum_{n=0}^{\infty} \frac{1}{2^{n+1}}(\|x - v_n\| - \|v - v_n\|) \qquad (2.14) \\ &\le f(x) + \frac{\varepsilon}{\lambda}\|x - v\|. \end{aligned}$$

Thus (2.6) reduces to the central inequality of Ekeland's Theorem [6]. Moreover, for any sequence (x_n) satisfying

$$f(x_n) + \frac{\varepsilon}{\lambda}\|x_n - v\| \to f(v),$$

inequality (2.14) implies that $f(x_n) + \varepsilon\rho_\infty(x_n) \to f(v) + \varepsilon\rho_\infty(v)$. Theorem 2.2 then shows that $x_n \to v$, so v is actually a *strong minimizer* for $f + (\varepsilon/\lambda)\|\cdot - v\|$. This leads to a corollary that (in the Banach space context) extends a result of Georgiev [10, Thm 1.6]:

Corollary 2.3 *Given $f\colon X \to \mathbb{R} \cup \{+\infty\}$ lsc, $x_0 \in X$, and $\varepsilon > 0$, suppose*

$$f(x_0) < \varepsilon + \inf_X f.$$

Then for any $\lambda > 0$, some point $v \in \mathbb{B}_\lambda(x_0)$ is a strong minimizer for the function $x \mapsto f(x) + (\varepsilon/\lambda)\|x - v\|$.

The Smooth Variational Principle of Borwein and Preiss To recover Theorem 1.1 from Theorem 2.2, take $p > 1$,

$$\rho(x) = \frac{\|x\|^p}{\lambda^p}, \quad \text{and} \quad \mu_n = \frac{1}{2^{n+1}(n+1)}.$$

Then there is a sequence (v_n) converging to some v where $\rho(v - x_0) < 1$, *i.e.*, $\|v - x_0\| < \lambda$, and such that the function $f + \varepsilon\rho_\infty$ has a strong minimum at v. Here the penalty has the form $\rho_\infty(x) = \lambda^{-p}\Delta_p(x)$, where

$$\Delta_p(x) := \sum_{n=0}^{\infty} \frac{(n+1)^{p-1}}{2^{n+1}} \|x - v_n\|^p,$$

so (2.6) implies (1.1). Thus Theorem 2.2 generalizes Theorem 1.1 by adding strong minimality to the properties asserted for v.

Theorem 2.2 has differentiability consequences similar to those in Theorem 1.1. Indeed, if X has a β-smooth norm, then the penalty function ρ_∞ is β-smooth also (see [1, Prop. 2.4 (a)]), and (2.13) implies

$$\nabla\rho_\infty(v) = \sum_{k=0}^{\infty} \frac{(k+1)^{p-1}}{2^{k+1}} \frac{p}{\lambda} \left\| \frac{v - v_k}{\lambda} \right\|^{p-1} \nabla \|v - v_k\|,$$

$$\|\nabla\rho_\infty(v)\| \le \sum_{k=0}^{\infty} \frac{(k+1)^{p-1}}{2^{k+1}} \frac{p}{\lambda} \frac{1}{(k+1)^{p-1}} = \frac{p}{\lambda}.$$

Thus (2.6) implies that $\partial^\beta f(v)$ contains an element of norm not exceeding $\varepsilon p/\lambda$.

Deploying Smooth Bumps A bump on a Banach space X is, by definition, a bounded real-valued function b such that $\operatorname{supp}(b) = \{x \in X : b(x) \neq 0\}$ is a bounded non-empty set. Some Banach spaces admit smooth bumps but not smooth renorms (see [13, Section 4]), so the existence of a smooth bump is a milder hypothesis than the smoothness condition in Theorem 1.1. Given a continuous bump b on X, a perturbation kernel ρ suitable for use in Theorem 2.2 can be built as follows:

$$\rho := \phi \circ \widetilde{b}, \quad \text{where} \quad \widetilde{b}(x) = \alpha b\big(\beta(x - x_0)\big), \quad x \in X. \tag{2.15}$$

Here α, β, and x_0 are chosen to arrange $\widetilde{b}(0) = 1$ and $\widetilde{b}(x) = 0$ when $\|x\| \geq 1$, and $\phi\colon \mathbb{R} \to [0,1]$ is a C^∞ function with bounded derivative such that $\phi(t) = 0$ if $t \geq 1$ and $\phi(t) = 1$ if $t \leq 0$. Notice that ρ obeys (2.4) with $\eta = 1$, and indeed, that $\rho(w) < 1$ implies $\|w\| < 1$. Based on this construction, we can derive a local version of the Deville-Godefroy-Zizler variational principle [13, Thm. 4.10] from Theorem 2.2:

Theorem 2.4 *Given $f\colon X \to \mathbb{R} \cup \{+\infty\}$ lsc, $x_0 \in X$, and $\varepsilon \in (0,1)$, suppose*

$$f(x_0) < \varepsilon + \inf_X f.$$

If the Banach space X admits a bump which is globally Lipschitz and β-differentiable, then there exist $v \in X$ and a β-differentiable g such that

(1) $f + g$ has a strong minimum at v,
(2) $\|g\|_\infty < \varepsilon$ and $\|\nabla g\|_\infty < \varepsilon$,
(3) $\|x_0 - v\| < \varepsilon$ and $f(v) \leq f(x_0)$.

Proof Using the bump b described in the statement, build ρ as in (2.15); then let $\widetilde{\rho}(x) := \rho(x/\varepsilon)$. The β-smoothness of b implies that $\widetilde{\rho}$ is β-smooth also, so we can apply Theorem 2.2 using $\widetilde{\rho}$ and

$$\mu_k = \frac{\varepsilon}{2^{k+1}(k+1)\max\left\{\|\rho\|_\infty, \|\nabla\rho\|_\infty, 1\right\}}.$$

Then conclusion (1) holds for $g = \varepsilon\widetilde{\rho}_\infty$, given in detail by

$$g(x) = \varepsilon \sum_{k=0}^{\infty} \frac{\varepsilon}{2^{k+1}(k+1)\max\left\{\|\rho\|_\infty, \|\nabla\rho\|_\infty, 1\right\}} \rho\left(\frac{(k+1)(x - v_k)}{\varepsilon}\right).$$

As b is globally Lipschitz, g is β-differentiable. The norm and gradient estimates in (2) follow immediately, and since $g \geq 0$ everywhere, we have $f(x_0) \geq f(v) + g(v) \geq f(v)$. The remaining assertion in (3) follows from $\widetilde{\rho}(x_0 - v) < 1$. ■

3 The Nonlocal DGZ Variational Principle

In [4, Lemma I.2.5], Deville, Godefroy, and Zizler consider a Banach space Y of bounded continuous real-valued functions on X with the following properties:

(Y.1) $\|g\|_Y \geq \|g\|_\infty$ for each $g \in Y$.
(Y.2) Whenever $g \in Y$ and $x \in X$, one has $\tau_x g \in Y$ and $\|\tau_x g\|_Y = \|g\|_Y$, where $\tau_x g\colon X \to \mathbb{R}$ is given by $\tau_x g(v) = g(x + v)$.

(Y.3) Whenever $g \in Y$ and $\alpha \in \mathbb{R}$, the function $x \mapsto g(\alpha x)$ lies in Y.
(Y.4) Y contains a bump function b.

Motivated by the proof of Stegall's variational principle given by Bourgain [13, pp. 89–90], our purpose in this section is to give an iterative proof of the nonlocal Deville-Godefroy-Zizler variational principle without reference Baire category. The required perturbation function is essentially an infinite sum of negative bumps. Note that an appeal to Theorem 2.2 requires us to modify the given bump to produce a nonnegative-valued ρ. Here we use the bump directly in the iteration.

Lemma 3.1 *If $f\colon X \to \mathbb{R} \cup \{+\infty\}$ is bounded below and $g\colon X \to \mathbb{R}$, then for any $\alpha > 0$ we have $\Sigma_\beta(f+g) \subset \Sigma_\alpha(f)$ whenever $\|g\|_\infty < \alpha/2$ and $0 < \beta < \alpha - 2\|g\|_\infty$.*

Proof Under the stated conditions, every $y \in \Sigma_\beta(f+g)$ obeys

$$\begin{aligned} f(y) = (f+g)(y) - g(y) &\le \inf_X (f+g) + \beta + \|g\|_\infty \\ &\le \inf_X f + \beta + 2\|g\|_\infty < \inf_X f + \alpha. \end{aligned}$$ ■

The above observation shows we can get decreasing families of sublevel sets provided we perturb the function f slightly.

Theorem 3.2 *Let X be a Banach space. Let Y be a Banach space of bounded continuous real-valued functions on X satisfying conditions (Y.1)–(Y.4) above. Let $f\colon X \to \mathbb{R} \cup \{+\infty\}$ be lsc and bounded below, with $f(x) < +\infty$ for some $x \in X$. Then for every $\varepsilon > 0$ there exists $g \in Y$ such that $\|g\|_Y < \varepsilon$ and $f+g$ attains a strong minimum on X.*

Proof We claim that for every $\delta > 0$, there exist arbitrarily small $\alpha > 0$ for which some $\rho \in Y$ obeys both

$$\|\rho\|_Y < \delta \quad \text{and} \quad \operatorname{diam}\big(\Sigma_\alpha(f+\rho)\big) < 2\delta.$$

To show this, assume without loss of generality that the given bump b satisfies $b(0) > 0$, $b(y) = 0$ for $\|y\| \ge \delta$, and $\|b\|_Y < \delta$. (Use the scaling in (2.15) if necessary.) Choose $x_\delta \in X$ such that $f(x_\delta) < b(0) + \inf_X f$, and define $\rho(y) := -b(y - x_\delta)$. Then

$$\begin{aligned} (f+\rho)(x_\delta) = f(x_\delta) - b(0) &< \inf_X f, \\ \|y - x_\delta\| \ge \delta \Longrightarrow (f+\rho)(y) = f(y) &\ge \inf_X f. \end{aligned}$$

Thus, every $\alpha < \inf_X f + b(0) - f(x_\delta)$ obeys $\Sigma_\alpha(f+\rho) \subset \mathbb{B}_\delta(x_\delta)$, as required.

To construct g, we assume without loss of generality that $0 < \varepsilon < 1$. By the claim above, with $\delta_1 = \varepsilon/2$, there exist $\alpha_1 \in (0,1)$ and $\rho_1 \in Y$ such that

$$\|\rho_1\|_Y < \delta_1, \quad \operatorname{diam}\big(\Sigma_{\alpha_1}(f+\rho_1)\big) < 2\delta_1.$$

https://doi.org/10.4153/CJM-2001-044-8 Published online by Cambridge University Press

The claim also applies to $f + \rho_1$: taking $\delta_2 = \alpha_1\varepsilon/2^2$, we obtain $\alpha_2 \in (0, \alpha_1)$ and $\rho_2 \in Y$ such that

$$\|\rho_2\|_Y < \delta_2, \quad \operatorname{diam}\bigl(\Sigma_{\alpha_2}(f + \rho_1 + \rho_2)\bigr) < 2\delta_2.$$

Continuing by induction (with $\alpha_0 = 1$), we obtain sequences $\delta_n > 0$, $\alpha_n \in (0,1)$, and $\rho_n \in Y$ such that

$$\delta_n = \frac{\alpha_{n-1}\varepsilon}{2^n}, \quad \alpha_n < \alpha_{n-1} \quad \|\rho_n\|_Y < \delta_n, \quad \operatorname{diam}\bigl(\Sigma_{\alpha_n}(f + \rho_1 + \cdots + \rho_n)\bigr) < 2\delta_n.$$

Then we define $g = \sum_{n=1}^{\infty} \rho_n$. This function lies in Y because

$$\|g_\infty\|_Y \le \sum_{n=1}^{\infty} \|\rho_n\|_Y < \sum_{n=1}^{\infty} \frac{\alpha_{n-1}}{2^n}\varepsilon < \varepsilon.$$

To show that $f + g$ attains a strong minimum on X, it suffices to show

$$\operatorname{diam}\bigl(\Sigma_\beta(f + g)\bigr) \to 0 \quad \text{as } \beta \to 0^+.$$

To do this, fix any n and split $f + g = f + \sum_{i=1}^{n} \rho_i + \sum_{i=n+1}^{\infty} \rho_i$. Since

$$\Bigl\| \sum_{i=n+1}^{\infty} \rho_i \Bigr\|_Y \le \sum_{i=n+1}^{\infty} \frac{\alpha_{i-1}}{2^i}\varepsilon \le \frac{\alpha_n}{2^n}\varepsilon < \frac{\alpha_n}{2},$$

Lemma 3.1 implies that $\Sigma_\beta(f+g) \subset \Sigma_{\alpha_n}(f + \rho_1 + \cdots + \rho_n)$ holds whenever $0 < \beta < \alpha_n - 2\| \sum_{i=n+1}^{\infty} \rho_i \|_\infty$, that is,

$$\operatorname{diam}\bigl(\Sigma_\beta(f+g)\bigr) \le \operatorname{diam}\bigl(\Sigma_{\alpha_n}(f + \rho_1 + \cdots + \rho_n)\bigr) \le 2\delta_n.$$

Since $\delta_n \to 0$ as $n \to \infty$, this completes the proof. ∎

The proof does not require the sequence of sets $S_n = \Sigma_{\alpha_n}(f + \rho_1 + \cdots + \rho_n)$ to be nested, but this additional property can be arranged by choosing smaller α_n in the construction.

4 Generalized Smulyan Test and Stegall Minimization Principle

In this section, we consider Stegall's variational principle with linear perturbations. Our proof of Stegall's variational principle (Theorem 4.2), cast in the dual of an Asplund space, simplifies the one given by Fabian and Zizler [7]. It relies on a characteristic property of the dual of any Banach space with the Radon-Nikodym Property (RNP), in constrast to Stegall's original proof of Theorem 4.4 using properties of RNP sets [15].

The following 'generalized Smulyan test' describes the relationship between Fréchet differentiability and the strong extremum.

Lemma 4.1

(i) Let $f\colon X \to \mathbb{R} \cup \{+\infty\}$ be lsc and bounded below. Define $g\colon X^* \to \mathbb{R}$ via

$$g(x^*) := \inf_{y \in S}\{f(y) + \langle x^*, y\rangle\},$$

where $S \subset X$ is nonempty, closed, bounded, and convex. Then g is Fréchet differentiable at x^*, with $g'(x^*) = x$, if and only if $f + x^* + I_S$ attains a strong minimum at x.

(ii) Let $f\colon X^* \to \mathbb{R} \cup \{+\infty\}$ be (norm-) lsc and bounded below. Define $g\colon X \to \mathbb{R}$ by

$$g(x) := \inf_{y^* \in S}\{f(y^*) + \langle y^*, x\rangle\},$$

where $S \subset X^*$ is nonempty, norm-closed, bounded and convex. Then g is Fréchet differentiable at x, with $g'(x) = x^*$, if and only if $f + x + I_S$ attains a strong minimum at x^*.

Proof Both statements are proved similarly; we provide details only for (ii). Since g is concave, its Fréchet derivative $g'(x)$ exists if and only if

$$\liminf_{\|h\| \to 0} \frac{g(x+h) + g(x-h) - 2g(x)}{\|h\|} = 0.$$

(See, *e.g.*, [13, Prop. 1.23].)

($\Rightarrow$) Suppose g is Fréchet differentiable at x. Extending the notation in (2.1), let us write

$$\Sigma^*_\alpha(f + x + I_S) = \{y^* \in S : f(y^*) + \langle y^*, x\rangle \le \alpha + g(x)\}.$$

(For each $\alpha > 0$, this set is norm-closed, and nonempty.) To prove existence of a strong minimizer, it suffices to show that $\delta_n = \operatorname{diam}(S_n) \to 0$, where $S_n = \Sigma^*_{1/n^2}(f + x + I_S)$. Since $S_n \supseteq S_{n+1}$ in general, the conclusion is automatic if some $\delta_n = 0$, so we may assume $\delta_n > 0$ for all n. Thus we may choose $x_n^*, y_n^* \in S_n$, and then a corresponding $z_n \in X$, $\|z_n\| = 1$, such that

$$\|y_n^* - x_n^*\| > \frac{\delta_n}{2}, \quad \langle y_n^* - x_n^*, z_n\rangle > \frac{\delta_n}{2}.$$

Now by the definition of g and our choice of x_n^*, y_n^*,

$$\begin{aligned} g(x + z_n/n) + g(x - z_n/n) - 2g(x) &\le [f(x_n^*) + \langle x_n^*, x + z_n/n\rangle - g(x)] \\ &\quad + [f(y_n^*) + \langle y_n^*, x - z_n/n\rangle - g(x)] \\ &\le \frac{1}{n^2} + \langle x_n^* - y_n^*, z_n/n\rangle + \frac{1}{n^2}. \end{aligned}$$

Rearranging this gives

$$\delta_n < 2\,\langle y_n^* - x_n^*, z_n\rangle \le -2\frac{g(x + z_n/n) + g(x - z_n/n) - 2g(x)}{1/n} + \frac{4}{n}.$$

The Fréchet differentiability of g implies that the right side tends to 0, as required.

To identify $g'(x)$, write x^* for the (unique) strong minimizer of $f + x + I_S$. Note that $g(x) = f(x^*) + \langle x^*, x\rangle$, so

$$
\begin{aligned}
0 &= \lim_{\|h\|\to 0} \frac{g(x+h) - g(x) - \langle g'(x), h\rangle}{\|h\|} \\
&\le \liminf_{\|h\|\to 0} \frac{f(x^*) + \langle x^*, x+h\rangle - f(x^*) - \langle x^*, x\rangle - \langle g'(x), h\rangle}{\|h\|} \qquad (4.1)\\
&= \liminf_{\|h\|\to 0} \frac{\langle x^* - g'(x), h\rangle}{\|h\|} = -\|x^* - g'(x)\|.
\end{aligned}
$$

This shows $x^* = g'(x)$.

($\Leftarrow$) Conversely, assume $f + x + I_S$ attains a strong minimum somewhere. To show that g is Fréchet differentiable at x, fix an arbitrary sequence $h_n \in X \setminus \{0\}$ with $\|h_n\| \to 0$. Choose $x_n^*, y_n^* \in S$ such that

$$
\begin{aligned}
f(x_n^*) + \langle x_n^*, x + h_n\rangle &< g(x + h_n) + \|h_n\|/n, \\
f(y_n^*) + \langle y_n^*, x - h_n\rangle &< g(x - h_n) + \|h_n\|/n.
\end{aligned}
$$

Rearrangement gives

$$
\begin{aligned}
g(x) \le f(x_n^*) + \langle x, x_n^*\rangle &< g(x + h_n) - \langle h_n, x_n^*\rangle + \|h_n\|/n, \quad \text{and} \\
g(x) \le f(y_n^*) + \langle x, y_n^*\rangle &< g(x - h_n) + \langle h_n, y_n^*\rangle + \|h_n\|/n.
\end{aligned}
$$

Hence both $f(x_n^*) + \langle x_n^*, x_n\rangle \to g(x)$, and $f(y_n^*) + \langle y_n^*, x\rangle \to g(x)$, *i.e.*, both (x_n^*) and (y_n^*) are minimizing sequences for $f + x + I_S$. Since this function has a strong minimum, we must have $\|y_n^* - x_n^*\| \to 0$. On the other hand, the inequalities above imply

$$
\begin{aligned}
g(x + h_n) + g(x - h_n) - 2g(x) &> [f(x_n^*) + \langle x_n^*, x + h_n\rangle - \|h_n\|/n - g(x)] \\
&\quad + [f(y_n^*) + \langle y_n^*, x - h_n\rangle - \|h_n\|/n - g(x)] \\
&\ge \langle x_n^* - y_n^*, h_n\rangle - 2\|h_n\|/n \\
&\ge -\|h_n\|(\|x_n^* - y_n^*\| + 2/n).
\end{aligned}
$$

Dividing by h_n and sending $n \to \infty$ shows that

$$
\liminf_{n\to\infty} \frac{g(x + h_n) + g(x - h_n) - 2g(x)}{\|h_n\|} \ge 0.
$$

Since the sequence (h_n) is arbitrary, g is Fréchet differentiable at x. The identification of $g'(x)$ proceeds as in (4.1). ■

https://doi.org/10.4153/CJM-2001-044-8 Published online by Cambridge University Press

Smulyan's classical test for the Fréchet differentiability of norms and Phelps's test for the Fréchet differentiability of Minkowski gauges [13, p. 85] both follow from the case $f = 0$ of Lemma 4.1. One may express Lemma 4.1 in terms of Fenchel conjugates, but for our applications the stated form is more convenient.

Theorem 4.2 *Let X be an Asplund space, and let $f\colon X^* \to \mathbb{R} \cup \{+\infty\}$ be norm-lsc on X^*. Let $S \subseteq X^*$ be nonempty, norm-closed, bounded, and convex. If f is bounded below on S, then the set*

$$G := \{x \in X : f + x \text{ attains a strong minimum on } S\},$$

is residual in $(X, \|\cdot\|)$.

Proof Consider the concave function $g\colon X \to \mathbb{R}$ given by

$$g(x) := \inf_{x^* \in S}\{f(x^*) + \langle x, x^*\rangle\}.$$

Since S is bounded and nonempty, g is concave and Lipschitz on X. Since X is Asplund, the set G of all Fréchet differentiability points for g is residual in $(X, \|\cdot\|)$. Thus the conclusion follows from Lemma 4.1(ii). ■

Preiss and Zajicek [14] showed every continuous convex function on a separable Asplund space is Fréchet differentiable everywhere except for a σ-porous set. Thus on separable Asplund space the set $X \setminus G$ is σ-porous. When X is finite dimensional, Rademacher's theorem ensures that the set $X \setminus G$ is Lebesgue null.

***Corollary 4.3* (Fabian [7])** *Let X be an Asplund space. Let $f\colon X^* \to \mathbb{R} \cup \{\infty\}$ be lsc, with*

$$\inf_{X^*} f > -\infty \quad \text{and} \quad \liminf_{\|x^*\|\to\infty} \frac{f(x^*)}{\|x^*\|} > 0. \tag{4.2}$$

Then for every $\varepsilon > 0$ there exists $x_0 \in X$ with $\|x_0\| < \varepsilon$ such that $f + x_0$ attains a strong minimum on X^*.

Proof We follow the idea of Phelps [13, p. 92]. From (4.2), for some $a, b > 0$ we have $f(x^*) > a\|x^*\|$ whenever $\|x^*\| > b$, so

$$f(x^*) > a\|x^*\| + \min\{0, \inf_{X^*} f - ab\} \quad \forall x^* \in X^*.$$

Thus by adding a constant to f if necessary, we may assume $f(x^*) > a\|x^*\|$ for every $x^* \in X^*$. When $\|x\| < a/2$, we have

$$f(x^*) + \langle x, x^*\rangle > a\|x^*\| - \frac{a}{2}\|x^*\| = \frac{a}{2}\|x^*\|. \tag{4.3}$$

https://doi.org/10.4153/CJM-2001-044-8 Published online by Cambridge University Press

Let $r = 2(f(0)+1)/a$, $S = r\mathbb{B}^*$, and $f_S = f + I_S$. Then Theorem 4.2 provides a point $x_0 \in X$ such that $\|x_0\| < \min\{\varepsilon, a/2\}$ and $f_S + x_0$ attains a strong minimum at some $x_0^* \in X^*$. We only need to show x_0^* is a strong minimizer of $f + x_0$ on X^*. Indeed, if some $x^* \in X^*$ obeys

$$f(x^*) + \langle x_0, x^* \rangle \le f(x_0^*) + \langle x_0, x_0^* \rangle \le f(0),$$

then our choice of $\|x_0\| < a/2$ implies via (4.3) that $\|x^*\| < 2f(0)/a < r$, so $x^* \in S$. The strong minimality of x_0^* for $f_S + x_0$ forces $x^* = x_0^*$. Similarly, if $x_n^* \in X^*$ and $(f + x_0)(x_n^*) \to (f + x_0)(x_0^*)$, then eventually $(f + x_0)(x_n^*) \le f(0) + 1$. By (4.3), this will force $\|x_n^*\| \le \big(f(0)+1\big)2/a$ for all n sufficiently large. Hence $x_n^* \in S$ for all such n, and $\|x_n^* - x_0^*\| \to 0$. ■

Note that Corollary 4.3 retains the full force of Theorem 4.2. Indeed, if the hypotheses of Theorem 4.2 are in place, choose $r > 0$ so large that $r\mathbb{B}^* \supseteq S$ and let $m = \inf_{X^*} f$. Then apply Corollary 4.3 to $g := f + I_S$, noting that

$$g(x^*) \ge \|x^*\| - r + m \quad \text{for every } x^* \in X^*.$$

This provides a single strong minimizer; to see how this conclusion extends to provide generic information, see the proof of Corollary 4.6 below.

A Banach space X is said to have *Radon-Nikodym property* (RNP) if every nonempty bounded subset A of X is dentable, that is, for every $\varepsilon > 0$ there exists $x^* \in X^*$ and $\alpha > 0$ such that the "slice"

$$S(x^*, A, \alpha) := \{x \in A : \langle x^*, x \rangle > \sup_A \langle x^*, x \rangle - \alpha\}$$

has diameter less than ε.

Theorem 4.4 *Suppose that X has the RNP, $f \colon X \to \mathbb{R} \cup \{+\infty\}$ is lsc on X, and $S \subset X$ is nonempty, closed, bounded and convex. If $\inf_S f$ is finite, then the set*

$$\{x^* \in X^* : f + x^* \text{ has a strong minimum on } S\}$$

is residual in $(X^*, \|\cdot\|)$.

Proof Consider the concave function $g \colon X^* \to \mathbb{R}$ given by

$$g(x^*) := \inf_{x \in S} \{f(x) + \langle x^*, x \rangle\}.$$

Since S is bounded, $-g$ is Lipschitz, convex, and weak* lsc on X^*. Collier [3] has shown that every weak* lsc convex function on X^* is Fréchet differentiable on a dense G_δ subset of its domain. (That is, "X^* is weak* Asplund.") Hence the conclusion follows from Lemma 4.1(i). ■

Countable intersections of residual sets are residual, so the following consequence is immediate.

Corollary 4.5 *Let X and S be as in Theorem 4.4. Suppose (f_n) is a sequence of extended-valued lsc functions on X such that $\inf_S f_n$ is finite for each n. Then for every $\varepsilon > 0$ there exists $x^* \in X^*$ such that $\|x^*\| < \varepsilon$ and each $f_n + x^*$ attains a strong minimum at some x_n in S.*

Corollary 4.6 **(Fabian)** *Suppose X has the RNP. Let $f \colon X \to \mathbb{R} \cup \{+\infty\}$ be lsc and satisfy*

$$f(x) > c\|x\| + b \quad \forall x \in X \tag{4.4}$$

for some $c > 0$ and $b \in \mathbb{R}$. Then the set

$$U = \{x^* \in \mathbb{B}^*_c(0) : f + x^* \text{ attains a strong minimum on } X\},$$

is residual in the Baire space $\left(\mathbb{B}^*_c(0), \|\cdot\|\right)$.

Proof Just as Corollary 4.3 follows from Theorem 4.2, Theorem 4.4 implies that

$$\left[\begin{array}{l}\text{For any } \varepsilon > 0 \text{ there exists } x^* \in X^* \text{ such that } \|x^*\| < \varepsilon \\ \text{and } f + x^* \text{ attains a strong minumum on } X.\end{array}\right. \tag{4.5}$$

Now let

$$U_n = \left\{y^* \in \mathbb{B}^*_c(0) : \operatorname{diam}\left(\Sigma_\alpha(f + y^*)\right) < \frac{1}{n} \text{ for some } \alpha > 0\right\}.$$

As $U = \bigcap_{n=1}^\infty U_n$, it suffices to show that:

(i) Each U_n is open. Assume $y^* \in U_n$. Then for some $\alpha > 0$, we have

$$\operatorname{diam}\left(\Sigma_\alpha(f + y^*)\right) < \frac{1}{n}.$$

Choose $0 < \beta < (c - \|y^*\|)/2$. Whenever $\|x^* - y^*\| < \beta$, by (4.4) for every $y \in X$ we have

$$(f + x^*)(y) > \frac{c - \|y^*\|}{2}\|y\| + b.$$

We may choose $r > 0$ such that the latter is larger than $f(0) + 1$, for all $\|y\| \geq r$. Thus,

$$\inf_X(f + x^*) = \inf_{r\mathbb{B}}(f + x^*), \quad \text{for all } \|x^* - y^*\| < \beta.$$

We show that $x^* \in U_n$ if $\|x^* - y^*\| < \beta_1$ with $\beta_1 < \min\{\beta, \alpha/(2r)\}$. Indeed, for $0 < \nu < \min\{1, \alpha - 2r\beta_1\}$ we have $\Sigma_\nu(f + x^*) \subset r\mathbb{B}$. If $x \in \Sigma_\nu(f + x^*)$, then

$$\begin{aligned}(f + y^*)(x) &= (f + x^*)(x) + (y^* - x^*)(x)\\ &\leq \inf_X(f + x^*) + \nu + r\|(y^* - x^*)\| = \inf_{r\mathbb{B}}(f + x^*) + \nu + r\beta_1\\ &\leq \inf_{r\mathbb{B}}(f + y^*) + 2r\beta_1 + \nu = \inf_X(f + y^*) + 2r\beta_1 + \nu\\ &< \inf_X(f + y^*) + \alpha.\end{aligned}$$

Thus, $\Sigma_\nu(f + x^*) \subset \Sigma_\alpha(f + y^*)$, and so $\operatorname{diam}\big(\Sigma_\nu(f + x^*)\big) < 1/n$.

(ii) Each U_n is dense. For any $y^* \in \mathbb{B}_c^*(0)$, we have

$$(f + y^*)(x) \geq c\|x\| + b - \|y^*\|\,\|x\| = (c - \|y^*\|)\|x\| + b.$$

By (4.5), for every $\varepsilon > 0$ there exists $x^* \in X^*$ such that $\|x^*\| < \min\{\varepsilon, c - \|y^*\|\}$ and $f + y^* + x^*$ attains a strong minimum on X. This implies $y^* + x^* \in U_n$, as required. ■

5 The Coercive Case

A function $f : X \to \mathbb{R} \cup \{+\infty\}$ is called coercive when

$$\liminf_{\|x\| \to +\infty} \big(f(x)/\|x\|\big) = +\infty.$$

Proposition 5.1 *Let $f : X \to \mathbb{R} \cup \{+\infty\}$ be lsc and coercive. Then*

(i) When X is an arbitrary Banach space, the set $\bigcup_{x \in X} \partial_c f(x)$ is norm-dense in X^*.
(ii) When X has a β-smooth renorm, the set $\bigcup_{x \in X} \partial^\beta f(x)$ is norm-dense in X^*.
(iii) When X has the RNP, the set $\{x^* \in X^* : f + x^*$ attains a strong minimum on $X\}$ is residual in $(X^*, \|\cdot\|)$.

Proof (i) Fix $y^* \in X^*$ and $\varepsilon > 0$. As f is coercive, the function $f - y^*$ has a finite infimum. Ekeland's variational principle produces $x \in X$ such that $f - y^* + \varepsilon\|\cdot - x\|$ attains a global minimum at x. Thus there exists $x^* \in \partial_c f(x)$ such that $\|x^* - y^*\| \leq \varepsilon$. Since ε is arbitrary, this proves (i).

(ii) The proof is similar, but we apply the Borwein-Preiss variational principle.

(iii) Since f is coercive, for every $c > 0$ there exists b such that $f(x) > c\|x\| + b$ for every $x \in X$. By Fabian's variational principle (Corollary 4.6),

$$U := \{x^* \in X^* : f + x^* \text{ attains a strong minimum on } X\},$$

is residual in the Baire space $\big(\mathbb{B}_c^*(0), \|\cdot\|\big)$ for every $c > 0$. Hence U is residual in $(X^*, \|\cdot\|)$. ■

When S is a bounded set, $f = I_S$ is coercive, and Proposition 5.1 unifies several well-known results.

Theorem 5.2 **(Bishop-Phelps [13])** *Let $S \subset X$ be nonempty, closed, bounded, and convex. Then the support functionals of C are norm dense in X^*.*

Theorem 5.3 **(Phelps's Lemma [13])** *Suppose X has the RNP, and let $S \subset X$ be nonempty, closed, bounded, and convex. Then the strongly supporting functionals of S are residual in $(X^*, \|\cdot\|)$.*

Together with the Generalized Smulyan Test (Lemma 4.1), the results of this section support the following (known) characterization of reflexivity [9]. The statement and proof use the standard notation

$$\sigma_C(x^*) = \sup\{\langle x^*, x\rangle : x \in C\}, \quad x^* \in E^*,$$
$$C^\circ = \{x^* \in E^* : \sigma_C(x^*) \le 1\}$$

for various sets C and Banach spaces E.

Corollary 5.4 *If X contains a subset S that is closed, bounded, and convex, with $0 \in \operatorname{int} S$, and such that σ_S is Fréchet differentiable on $X^* \setminus \{0\}$, then X is reflexive.*

Proof According to the Bishop-Phelps theorem, the set

$$D := \{x^{**} \in X^{**} : \sigma_{S^\circ}(x^{**}) = \langle x^{**}, x^*\rangle \text{ for some } x^* \in S^\circ\},$$

is norm dense in X^{**}. The hypothesis that $0 \in \operatorname{int} S$ implies that 0 lies in the norm interior of S°, so $\sigma_{S^\circ}(x^{**}) > 0$ whenever $x^{**} \neq 0$. Thus for every $x^{**} \in D \setminus \{0\}$, there exists a nonzero $x^* \in X^*$ such that

$$1 = \left\langle \frac{x^{**}}{\sigma_{S^\circ}(x^{**})}, x^* \right\rangle = \sigma_{S^{\circ\circ}}(x^*) = \sigma_S(x^*).$$

The rightmost equation holds because $S^{\circ\circ} = \overline{S}^{w^*}$ in X^{**}, so the restriction of $\sigma_{S^{\circ\circ}}$ to X^* coincides with σ_S. Now σ_S is Fréchet differentiable at x^*, by hypothesis, so Lemma 4.1(i) (with $f = 0$) gives $x^{**}/\sigma_{S^0}(x^{**}) \in S \subseteq X$. Hence $D \subseteq X$, so X is norm-dense in X^{**}, and the result follows. ■

6 Variational Characterizations of Banach Spaces

Sullivan [16] has shown that a metric space Y is complete if and only if for every bounded below, lsc and somewhere finite function f on Y there exists a Lipschitz function $\phi\colon Y \to R$ with Lipschitz constant less than 1 such that $f+\phi$ attains a global minimum on Y. Fabian and Mordukhovich [8] showed that the Banach space X has a β-smooth renorm if and only if for every $f\colon X \to \mathbb{R} \cup \{+\infty\}$ that is lsc, bounded

https://doi.org/10.4153/CJM-2001-044-8 Published online by Cambridge University Press

below, and somewhere finite, there exists a β-smooth convex function $\phi\colon X \to \mathbb{R}$ such that $f + \phi$ attains a global minimum somewhere on X.

Let $\mathcal{S}$ be a class of real-valued functions on a Banach space X. To say that X admits an $\mathcal{S}$-variational principle means that every function $f\colon X \to \mathbb{R} \cup \{+\infty\}$ that is lsc, bounded below, and somewhere finite satisfies the conditions

$$f \geq s \text{ on } X \quad \text{and} \quad f(v) = s(v)$$

for some $s \in \mathcal{S}$ and $v \in \operatorname{dom} f$. For example, we will consider the class $\mathcal{S}$ of β-smooth functions $\phi\colon X \to \mathbb{R}$ satisfying

$$\phi(x_1) < \phi(x_2) \Rightarrow \langle \nabla\phi(x_1), x_2 - x_1 \rangle > 0.$$

Such a ϕ is called *pseudoconcave*. When the statement above holds for this $\mathcal{S}$, we say X admits a β-smooth pseudoconcave variational principle. The proof in [8] makes it clear that the cited result of Fabian and Mordukhovich holds even if $-\phi$ is only required to be pseudoconcave.

***Theorem 6.1* (Fabian and Mordukhovich [8])** *The following assertions concerning X are equivalent:*

(i) The concave β-smooth variational principle holds in X.
(ii) The pseudoconcave β-smooth variational principle holds in X.
(iii) X has an equivalent β-smooth norm.

We note that the validity of Stegall's Variational Principle (Theorem 4.2) actually characterizes Asplund spaces.

Theorem 6.2 *The following assertions about a Banach space X are equivalent:*

(i) X is Asplund.
(ii) Whenever $f\colon X^* \to \mathbb{R} \cup \{+\infty\}$ is norm-lsc and has a finite infimum over some nonempty, closed, bounded, convex $S \subset X^*$, the set

$$\{x \in X : f + x \text{ attains a strong minimum on } S\}$$

is residual in $(X, \|\cdot\|)$.

Proof (i $\Rightarrow$ ii) This restates Theorem 4.2.

(ii $\Rightarrow$ i) Take $f = 0$. Then (ii) implies that every norm-closed, bounded convex set S has weak* slices of arbitrary small diameter. This implies that X is Asplund, by [13, Theorem 2.32]. ■

7 Maximality of Approximate Subdifferentials

Throughout this section, we assume that the Banach space X is separable. For any subset A of X, we write $C(A)$ for the collection of bounded continuous real-valued functions defined on A, with the uniform norm. Our purpose is to show that (when A is open) the collection of functions f in $C(A)$ for which the approximate subdifferential equals X^* at every point of A is large.

Recall that every separable Banach space has an equivalent Gâteaux-smooth norm, so the smoothness hypothesis above ensures the applicability of the smooth variational principle.

Given $\phi\colon X \to \mathbb{R} \cup \{+\infty\}$, we call $x_0 \in X$ a *dimple point* for ϕ if there exists $r > 0$ such that $\phi(x_0) < \inf\{\phi(x) : \|x - x_0\| = r\}$. In this case, given any $\varepsilon > 0$, applying the smooth variational principle to the function $\phi + I_{\mathbb{B}_r[x_0]}$, produces a point $v \in \mathbb{B}_r(x_0)$ such that some $v^* \in \partial^- \phi(v)$ obeys $\|v^*\| < \varepsilon$.

A subset $Z \subset X$ is said to be *porous* in X if there exists $\lambda \in (0, 1]$ such that for any $x \in X$ and $r > 0$ there exists $y \in X$ such that $\mathbb{B}_{\lambda r}(y) \subset \mathbb{B}_r(x) \setminus Z$. The set Z is called σ-porous in X if it can be represented as a countable union of porous sets in X.

Lemma 7.1 *For any open set $A \subseteq X$ and any continuous $f\colon A \to \mathbb{R}$, the set $C(A) \setminus T$ is σ-porous, where*

$$T := \{g \in C(A) : \textit{the dimple points of } f + g \textit{ are dense in } A\}.$$

Proof Fix $x \in A$, $n \in N$, and consider the set

$$G_{n,x} := \left\{g : f + g \text{ has a dimple point in } \mathbb{B}_{1/n}(x)\right\}.$$

We will show that $C(A) \setminus G_{n,x}$ is porous, with $\lambda = 1/6$ in the definition. Indeed, fix any $g \in C(A)$ and $r > 0$. As $f + g$ is continuous we may choose $x_n \in \mathbb{B}_{1/n}(x)$ such that

$$(f + g)(x_n) < r/2 + \inf_{\mathbb{B}_{1/n}[x]} (f + g).$$

Then choose k large enough that the support of $h_n(y) := \max\{0, 1 - 2k\|x_n - y\|\}$ lies in $\mathbb{B}_{1/n}(x)$. Let $\alpha = 5r/6$ and define $g_n(y) := -\alpha h_n(y)$.

Claim 1 $\mathbb{B}_{r/6}(g + g_n) \subset \mathbb{B}_r(g)$.

This follows from the triangle inequality, since $\|g_n\| = \alpha = 5r/6$.

Claim 2 $\mathbb{B}_{r/6}(g + g_n) \subset G_{n,x}$.

In fact, x_n is a dimple point for every function of the form $g + g_n + \phi$, where $\phi \in C(A)$ obeys $\|\phi\| < r/6$. For we have

$$\begin{aligned}(f + g + g_n + \phi)(x_n) &= f(x_n) + g(x_n) - \alpha + \phi(x_n)\\ &< \inf_{\mathbb{B}_{1/n}[x]} (f + g) + r/2 - 5r/6 + r/6 = \inf_{\mathbb{B}_{1/n}[x]} (f + g) - r/6,\end{aligned}$$

https://doi.org/10.4153/CJM-2001-044-8 Published online by Cambridge University Press

while every y with $\|y - x_n\| = (2k)^{-1}$ lies in $\mathbb{B}_{1/n}(x)$ and satisfies $g_n(y) = 0$, so

$$(f + g + g_n + \phi)(y) = f(y) + g(y) + \phi(y) \geq \inf_{\mathbb{B}_{1/n}[x]}(f + g) - r/6.$$

Taken together, these claims show that $\mathbb{B}_{r/6}(g + g_n) \subseteq \mathbb{B}_r(g) \cap G_{n,x}$. Since this holds for arbitrary $r > 0$ and $g \in C(A)$, the set $C(A) \setminus G_{n,x}$ is porous.

To complete the proof, let $\{x_i : i \in \mathbb{N}\}$ be a countable dense subset of A and define

$$G := \bigcap_{i\in\mathbb{N}} \bigcap_{n\in\mathbb{N}} G_{n,x_i}.$$

For every $g \in G$, the function $f + g$ has a dimple point in every open subset of A, so $T \supseteq G$. Hence $C(A) \setminus T \subseteq C(A) \setminus G$. The latter set is σ-porous, by the arguments above. ■

Theorem 7.2 *For any open set $A \subset X$, the set $C(A) \setminus T$ is σ-porous, where*

$$T = \{g \in C(A) : \partial_a g(x) = X^* \quad \forall x \in A\}.$$

Proof Suppose $h \in C(A)$ has a dimple point at 0, so $h(0) < \inf_{\|x\|=r} h(x)$. Then, as noted above, the smooth variational principle implies that the open ball $\mathbb{B}_r(0)$ contains points v where $\partial^- h(v)$ contains elements of arbitrarily small norm. By translating this observation, we deduce that if $z \in X$ is a limit of points where h has dimples, then $0 \in \partial_a h(z)$. We apply this observation below.

Now for every $x^* \in X^*$, by Lemma 7.1, the set

$$\{g \in C(A) : \text{the dimple points of } -x^* + g \text{ are dense in } A\},$$

is large; it follows that the superset

$$G^{x^*} := \{g \in C(A) : x^* \in \partial_a g(x) \quad \forall x \in A\},$$

has a σ-porous complement. Now $\mathbb{B}^*$ is weak*-compact and metrizable, so there exists a countable set $\{x_i^* : i \in \mathbb{N}\}$ whose weak* closure is X^*. It follows that $C(A) \setminus G$ has σ-porous complement, where

$$G = \bigcap_{i\in\mathbb{N}} G^{x_i^*}.$$

But $G = T$, since for every $g \in G$ and $x \in A$, the subdifferential $\partial_a g(x)$ must be weak* closed and contain $\{x_i^* : i \in \mathbb{N}\}$. ■

Under the hypotheses above, Deville and Revalski [5] showed that the set

$$\{g \in C(A) : g \text{ has a strong minimum in } A\}$$

has a σ-porous complement. In conjunction with Lemma 7.1, this implies that the set

$$\{g \in C(A) : g \text{ attains a unique global minimum,} \\ \text{but has a dense set of dimple points in } A\},$$

has a σ-porous complement in $C(A)$. In finite dimensional spaces, every such function has a local minimizer in every open subset of A, and exactly one global minimizer.

https://doi.org/10.4153/CJM-2001-044-8 Published online by Cambridge University Press

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] F. H. Clarke, Yu. S. Ledyaev, R. J. Stern and P. R. Wolenski, *Nonsmooth Analysis and Control Theory.* Graduate Texts in Math. **178**, Springer-Verlag, New York, 1998.
[3] J. B. Collier, *The dual of a space with the Radon-Nikodym property.* Pacific J. Math. **64**(1976), 103–106.
[4] R. Deville, G. Godefroy and V. Zizler, *Smoothness and Renorming in Banach Spaces.* Pitman Monographs Surveys Pure Appl. Math. Vol. **64**, Longman, 1993.
[5] R. Deville and J. P. Revalski, *Porosity of ill-posed problems.* Proc. Amer. Math. Soc. **128**(1999), 1117–1124.
[6] I. Ekeland, *On the variational principle.* J. Math. Anal. Appl. **47**(1974), 324–353.
[7] M. Fabian and V. Zizler, *An elementary approach to some questions in higher order smoothness in Banach spaces.* Extracta Math. **14**(1999), 295–327.
[8] M. Fabian and B. S. Mordukhovich, *Nonsmooth characterizations of Asplund spaces and smooth variational principles.* Set-Valued Anal. **6**(1998), 381–406.
[9] J. R. Giles, *Convex analysis with application in differentiation of convex functions.* Research Notes in Mathematics **58**, Pitman, 1982.
[10] P. G. Georgiev, *The strong Ekeland variational principle, the strong drop theorem and applications.* J. Math. Anal. Appl. **131**(1988), 1–21.
[11] A. D. Ioffe, *Approximate subdifferentials and applications II.* Mathematika **33**(1986), 111–128.
[12] Yongxin Li and Shuzhong Shi, *A generalization of Ekeland's ε-variational principle and its Borwein-Preiss smooth variant.* J. Math. Anal. Appl. **246**(2000), 308–319.
[13] R. R. Phelps, *Convex functions, monotone operators and differentiability.* 2nd edition, Lecture Notes in Mathematics **1364**, Springer-Verlag, Berlin, 1993.
[14] D. Preiss and L. Zajicek, *Fréchet differentiation of convex functions in a Banach space with a separable dual.* Proc. Amer. Math. Soc. **91**(1984), 202–204.
[15] C. Stegall, *Optimization of functions on certain subsets of Banach spaces.* Math. Ann. **236**(1978), 171–176.
[16] F. Sullivan, *A characterization of complete metric spaces.* Proc. Amer. Math. Soc. **83**(1981), 345–346.

Department of Mathematics
University of British Columbia
Vancouver, BC
V6T 1Z2
email: loew@math.ubc.ca

Department of Mathematics & Statistics
Okanagan University College
3333 College Way
Kelowna, BC
V1V 1V7
email: xwang@cecm.sfu.ca

https://doi.org/10.4153/CJM-2001-044-8 Published online by Cambridge University Press