A one perturbation variational principle and applications

Jonathan Borwein[*], **Marián Fabian**[†], **Julian P. Revalski**[‡]

March 7, 2003

Abstract

We study a variational principle in which there is one common perturbation function φ for every proper lower semicontinuous extended real-valued function f defined on a metric space X. Necessary and sufficient conditions are given in order for the perturbed function $f + \varphi$ to attain its minimum. In the case of a separable Banach space we obtain a specific principle in which the common perturbation function is, in addition, also convex and Hadamard-like differentiable. This allows us to provide applications of the principle to differentiability of convex functions on separable and more general Banach spaces.

Mathematics Subject Classifications (2000): 49J53, 49B20, 49J50, 49K40

Key words: variational principle, well-posed optimization problem, perturbed optimization problem, separable Banach space, weak Asplund space, Gâteaux differentiability space

1 Introduction

In a typical variational principle one is given a fixed real-valued function f (usually at least lower semicontinuous) defined on a Banach (or more general, on a complete metric) space X, bounded from below, along with a class $\mathcal{P}$ of functions in X which

[*]The research of J. Borwein has been supported by NATO Collaborative Linkage Grant 978488, by NSERC and by the Canada Research Chair Program.

[†]The research of M. Fabian has been partially supported by grants GA ČR 201/01/1198 and A 101 9301 .

[‡]The research of J.P. Revalski has been supported by a Marie Curie Fellowship of the European Community program IHP under contract HPMF-CT-2002-01874, and by NATO Collaborative Linkage Grant 978488.

serves as a source of possible perturbations for f. The aim is to find conditions under which there is at least one perturbation g from $\mathcal{P}$ such that the perturbed function $f+g$ attains its minimum (or even stronger, is *well-posed*, see the definition below).

Several well-known variational principles fit in this scheme. For example: Ekeland's variational principle [9, 10] with $\mathcal{P}$ a family of suitable multiplications-translations of the norm (see also the predecessor of this result, the Bishop-Phelps lemma [1]); Stegall's variational principle, where $\mathcal{P}$ is the elements of the dual X^* [15]; the Borwein-Preiss smooth variational principle, with $\mathcal{P}$ a family of smooth combinations of the norm in X [2]; the Deville-Godefroy-Zizler smooth variational principle, with $\mathcal{P}$ a family of Lipschitz or C^1-functions (containing at least one bump function of the same type) [6, 7] (cf. also an earlier generic variational principle in [4] and a further development in [8]). In all these principles the existence of a "good" (for f) perturbation in $\mathcal{P}$ has been shown to hold densely or generically in $\mathcal{P}$ (i.e. at least in a dense set of $\mathcal{P}$, equipped with a suitable metric).

The utility of the above principles has been demonstrated in many directions — from the already vast selection of applications of Ekeland's variational principle in optimization, nonlinear and variational analysis, critical point theory,partial differential equations etc., through the applications of the (smooth) variational principles of Stegall, Borwein-Preiss and Deville-Godefroy-Zizler to optimization and control, nonsmooth analysis, geometry of Banach spaces and existence of solutions to differential equations ([2, 6, 7, 15, 14]), to the applications of generic variational principles to optimization, differentiability of convex functions and topological games (see e.g. [5, 12]).

Our goal in this note, is in some sense to reverse the above scheme and to investigate a situation in which the perturbation function is unique That is, the class $\mathcal{P}$ reduces to a single element, say φ, common for every function f from some large enough class; for example for any lower semicontinuous function, bounded from below in X. This idea has already been used by Tykhonov [16] for the case of convex functions f in the setting of a Hilbert (or a reflexive suitably renormed) space by using as φ the square of the norm. This yields the so-called *Tykhonov regularization* (so that the perturbation $f+\varphi$ always has a unique minimum). Our aim is to consider the same situation for the larger class of lower semicontinuous functions. As we shall see, the freedom to perturb, simultaneously, by one fixed function an arbitrarily large family of functions, can be of great advantage over the principles cited above.

Of course, without any further requirements on the unique perturbation φ, the proof of the existence of a single φ making the perturbations $f+\varphi$ achieve its minimum for any lower semicontinuous function f is not at all difficult. For example, such a principle is evidently true whenever Y is a finite-dimensional subspace of X

(including $Y = \{0\}$) and

$$\varphi(x) = \begin{cases} \|x\|^2, & \text{if } x \in Y, \\ +\infty, & \text{if } x \in X \setminus Y. \end{cases}$$

Evidently, one can replace finite-dimensionality by norm compactness of the level sets of the function φ (as discussed below our simple variational principle of Theorem 2.1). What is not immediate, and deserves study, is how to see that the latter compactness condition, together with lower semicontinuity of the function φ, is not only sufficient to obtain such a variational principle but is also necessary for it. This is established in our Theorem 2.4.

The second question which is worth investigating is whether outside finite dimensional spaces one can find a function φ with compact level sets which is in addition not only convex but possesses some differentiability properties. This is obtained in Proposition 3.1, which proves the existence of such a φ in any separable Banach space. In consequence, we obtain a concrete variational principle in the setting of separable Banach spaces — Theorem 3.2. Traces of this last result are implicitly present — but quite well hidden — in [3].

We conclude our paper by showing how this variational principle can be applied to re-establish the well-known result that every separable Banach space X is a *weak Asplund space* (i.e., every continuous convex function in X is Gâteaux differentiable on a dense G_δ-subset of X). Moreover, we shall see that the proof of this result, based on our new variational principle, can be extended to prove a recent result from [3], to the effect that the Cartesian product of a Gâteaux differentiability space with a separable space is again a Gâteaux differentiability space. Recall that a *Gâteaux differentiability space* is one in which every continuous convex function is densely Gâteaux differentiable. The class of Gâteaux differentiability space is strictly larger than the class of weak Asplund spaces, as has recently been shown by Moors and Somasundaram [13]. For more detail about weak Asplund spaces and Gâteaux differentiability spaces the reader is referred to [7, 11, 14].

2 A variational principle with one perturbation function

Our main goal in this section is to investigate necessary and sufficient conditions ensuring the perturbation of every lower semicontinuous function by a fixed one attains its minimum.

Let us recall that given a function $g : X \to \mathbb{R} \cup \{+\infty\}$, defined in a Hausdorff topological space, its *level set of height* $r \in \mathbb{R}$ is $g^r := \{x \in X; g(x) \leq r\}$, the *domain* of g is $\operatorname{dom} g := \{x \in X; g(x) < \infty\}$ and its *epigraph* is the set:

$$\operatorname{epi} g := \{(x,r) \in X \times \mathbb{R}; g(x) \leq r\}.$$

The function g is *proper* if $\operatorname{dom} g$ is non-empty; g is *lower semicontinuous* (lsc) if its epigraph is a closed subset of $X \times \mathbb{R}$ considered with the product topology. Equivalent definitions of lower semicontinuity are: every level set g^r is closed; or, for each $x \in X$ one has $\liminf_\alpha g(x_\alpha) \geq g(x)$ for any net $\{x_\alpha\}$ converging to x. Finally, $\inf_X g$ (or simply $\inf g$) denotes the infimum of g over X.

It is well-known that if a proper lsc function $g : X \to \mathbb{R} \cup \{+\infty\}$ defined on a Hausdorff topological space X, has at least one non-empty compact level set then it is bounded from below and attains its minimum. This is Weierstrass's theorem. Moreover, something stronger is true: the problem of minimizing g is *well-posed* in the sense that every minimizing net (or sequence in the metric case) has a convergent to a minimum subnet (subsequence).

Thus the following result holds:

Theorem 2.1 *Let X be a Hausdorff topological space which admits a proper lsc function $\varphi : X \to \mathbb{R} \cup \{+\infty\}$ whose level sets are all compact. Then for any proper lsc and bounded from below function $f : X \to \mathbb{R} \cup \{+\infty\}$ the function $f + \varphi$ attains its minimum. In particular, if $\operatorname{dom} \varphi$ is relatively compact, the conclusion is true for any proper lsc function f (not necessarily bounded from below).*

Proof. Suppose $\operatorname{dom} f \cap \operatorname{dom} \varphi \neq \emptyset$ (if the latter set is empty there is nothing to prove). Then the function $f + \varphi$ is a proper lsc function. In particular, $(f + \varphi)^r$ is non-empty (and closed) for some r. Now, if f is bounded from below then $f + \varphi$ is bounded below as well. But, $(f + \varphi)^r \subset \varphi^{r - \inf f}$ and since the latter set is compact, the same is true for $(f + \varphi)^r$. On the other hand, if the domain of φ is relatively compact, since $\operatorname{dom}(f + \varphi) \subset \operatorname{dom} \varphi$, then the domain of $f + \varphi$ is also relatively compact. Thus, again, $(f + \varphi)^r$ is compact. The result then follows by the remark before the theorem. ∎

Remark 2.2 It is clear that, if $(X, \|\cdot\|)$ is a real normed space and φ is in addition convex, then the result above remains true for every proper lsc convex f, provided only that either the level sets of φ are only weakly compact, or that $\operatorname{dom} \varphi$ is.

Remark 2.3 Again in the case of a normed space $(X, \|\cdot\|)$, allowing translations of the argument of the fixed perturbation φ, we can establsh a localization of the minimum of the perturbation as in the classical variational principles (Bishop-Phelps [1], Ekeland [9, 10] and Borwein-Preiss [2]). Namely, essentially with the same simple proof, one can see that if X admits a function φ as above, then for any proper lsc (bounded from below) function $f : X \to \mathbb{R} \cup \{+\infty\}$, for any $\bar{x} \in \operatorname{dom} f$ and each

$\lambda > 0$, the function $f(x) + \varphi((\cdot - \bar{x})/\mu)$ (for a suitable $\mu > 0$), attains its minimum at some u with $\|u - \bar{x}\| \leq \lambda$. Observe that in this case, formally, the perturbation function is already varying.

Let us now investigate whether the conditions imposed on the unique perturbation φ in Theorem 2.1 are also necessary. In fact, in the case of a metric space X, as it is seen from the following result, the property that each lsc, or even each continuous bounded, perturbation of a fixed function φ attains its minimum characterizes the lower semicontinuity of φ as well as the fact that all its level sets are compact. Namely, we have:

Theorem 2.4 *Let $\varphi : X \to \mathbb{R} \cup \{+\infty\}$ be a proper function on a metric space X with the following property: for every bounded continuous function $f : X \to \mathbb{R}$, the function $f + \varphi$ attains its minimum. Then φ is lower semicontinuous function, bounded from below, whose level sets are all compact.*

Proof. Obviously φ is bounded from below. Suppose it is not lower semicontinuous. Then its epigraph $\operatorname{epi}\varphi$ is not closed in the product topology of $X \times \mathbb{R}$, i.e. there exists a sequence $\{(x_n, r_n)\}_{n=1}^{\infty} \subset \operatorname{epi}\varphi$ such that $(x_n, r_n) \to (x_0, r) \notin \operatorname{epi}\varphi$. Fix such an x_0 and observe that, since φ is bounded from below, the sequence may be taken to have the form $\{(x_n, \varphi(x_n))\}_{n=1}^{\infty}$. This shows that for the fixed x_0 the following set is non-empty:

$$R := \{r \in \mathbb{R};\ \text{there is } \{x_n\}_{n=1}^{\infty} \subset X : (x_n, \varphi(x_n)) \to (x_0, r) \notin \operatorname{epi}\varphi\}.$$

Put $r_0 := \inf R$. Since R is bounded from below by $\inf\varphi$, the number r_0 is well-defined and obviously $r_0 < \varphi(x_0)$. Moreover, using a diagonal process, one easily sees that $r_0 \in R$.

Further, for any $n = 0, 1, 2, \ldots$, let V_n be a neighborhood of x_0 with the property:

$$\varphi(x) > r_0 - \frac{1}{2^{n+1}} \quad \text{for any } x \in V_n.$$

The existence of such a V_n is guaranteed by the definition of r_0. We may arrange, without loss of generality, that for any $n = 0, 1, 2, \ldots$, we have $\overline{V_{n+1}} \subset V_n$ and that the family $\{V_n\}_{n \in \mathbb{N}}$ provides a local base of neighborhoods for the point x_0.

Now, for any $n = 0, 1, 2, \ldots$, we define the following continuous functions $h_n : X \to [0, 1]$:

$$h_n | X \setminus V_n \equiv 1 \text{ and}$$
$$h_n(x_0) = 0.$$

The functions exist because X is a completely regular topological space. Choose, also $\alpha \in \mathbb{R}_+$ so that $\alpha > r_0 - \inf\varphi$, and let

$$f(x) := \alpha h_0(x) + \sum_{i=1}^{\infty} \frac{1}{2^n} h_n(x), \quad x \in X.$$

The function f is well-defined, continuous, bounded and such that $f(x_0) = 0$. We show that for every $x \in X$ we have $f(x) + \varphi(x) > r_0$. Indeed, if $x = x_0$ we have

$$f(x_0) + \varphi(x_0) = \varphi(x_0) > r_0.$$

If $x \neq x_0$ we have two possibilities:

Case 1: $x \notin V_n$ for any $n = 0, 1, \ldots$. In this case $h_n(x) = 1$ for every n and consequently

$$f(x) + \varphi(x) = \alpha + 1 + \varphi(x) \geq \alpha + 1 + \inf \varphi > r_0.$$

Case 2: $x \in V_n$ for some n. Since $x \neq x_0$ and the family $\{V_n\}$ is a local base for x_0, we have $x \in V_k \setminus V_{k+1}$ for some $k \in \mathbb{N}$. Then, by the definition of f and the h_n's, we have

$$f(x) + \varphi(x) \geq \frac{1}{2^{k+1}} + \varphi(x) > r_0,$$

according to the definition of V_k. Thus

$$f(x) + \varphi(x) > r_0, \quad \text{for every } x \in X. \tag{2.1}$$

On the other hand, since $r_0 \in R$, there is $x_n \to x_0$ so that $\varphi(x_n) \to r_0$. Hence, since f is continuous, $f(x_n) + \varphi(x_n) \to f(x_0) + r_0 = r_0$, which together with (2.1) shows that $\inf(f + \varphi) = r_0$. But this infimum of the perturbation $f + \varphi$ is obviously not attained, according to the same inequality (2.1), which contradicts to the assumptions of the theorem. Therefore, the function φ must be lower semicontinuous.

Secondly, let us prove that every level set of φ is compact. We may suppose, without loss of generality, that $\inf \varphi = 0$. Take some $r > 0$ and consider the level set $X_r := \{x \in X; \varphi(x) \leq r\}$. Observe that, because φ is lsc, the set X_r is closed.

Suppose X_r is not compact. Then there is a sequence $\{x_n\}_{n=1}^{\infty} \subset X_r$ which has no cluster point in X_r. We will, for the moment, restrict ourselves to the space X_r with the inherited metric topology. Using the above, one can see that there are sets $\{W_n\}_{n=1}^{\infty}$, open in X_r, such that $x_n \in W_n$ for every n and $W_i \cap W_j = \emptyset$ for every $i \neq j$. Since φ is lsc we may assume that we have in addition:

$$\varphi(x) > \varphi(x_n) - \frac{1}{2n} \quad \forall x \in W_n. \tag{2.2}$$

Further, for any $x \in X_r \setminus \bigcup_{n=1}^{\infty} W_n$ there is a relatively open neighborhood W_x so that $x_n \notin W_x$ for every n (indeed, otherwise x would be an accumulation point of the sequence $\{x_n\}$). Consider now the following open covering of X_r

$$\Gamma := \{W_x : x \notin X_r \setminus \bigcup_{n=1}^{\infty} W_n\} \cup \{W_n : n = 1, 2, \ldots\}.$$

Since X_r is a metric space this covering has an open refinement $\gamma = \{V_\alpha : \alpha \in \Lambda\}$ which is neighborhood-finite, i.e. for any $x \in X_r$ there is a neighborhood U of x so that U intersects at most finitely many members of γ. Let $\alpha_n \in \Lambda$ be such that $x_n \in V_{\alpha_n}$ for every $n = 1, 2, \ldots$ Because of the definition of the covering Γ ($x_n \notin W_x$ for every n and $x \notin \cup W_n$) and because of the fact that the W_n are pair-wise disjoint, it follows that $V_{\alpha_n} \subset W_n$ for every n. In particular, V_{α_n} are pair-wise disjoint as well. Now, for every n, there are continuous functions $h_n : X_r \to [-\varphi(x_n) - 1 + 1/n, 0]$:

$$h_n(x_n) = -\varphi(x_n) - 1 + \frac{1}{n} \text{ and}$$
$$h|X_r \setminus V_{\alpha_n} \equiv 0.$$

Finally, let

$$h(x) := \sum_{n=1}^{\infty} h_n(x), \quad x \in X_r.$$

Because the open sets V_{α_n} are pair-wise disjoint, the function h is well-defined in X_r. Obviously h takes its values in the interval $[-r-1, 0]$. Moreover, it is continuous when restricted X_r. Indeed, if $x \in V_{\alpha_n}$ for some n, then it coincides with h_n on V_{α_n}, hence is continuous at x. If $x \notin V_{\alpha_n}$ for any n then due to the properties of the covering γ, there is an open neighborhood U of the point x so that U intersects at most finitely many V_{α_n}. Consequently in U the function h is a finite sum of continuous functions. Thus again, it h. continuous at x.

Further, since X_r was closed in X, according to Dugundji's extension theorem, there is a continuous function $f : X \to [-r-1, 0]$ which is an extension of h. Let us consider the perturbation $f + \varphi$. We first show that

$$f(x) + \varphi(x) > -1 \text{ for all } x \in X. \tag{2.3}$$

Indeed, if $x \notin X_r$ we have $\varphi(x) > r$ and the above inequality is obvious having in mind the range of f. So, let $x \in X_r$. Then we again have two cases:

Case 1: $x \notin V_{\alpha_n}$ for any n. Then $f(x) = h(x) = 0$. Thus, $f(x) + \varphi(x) = \varphi(x) \geq 0$ and the inequality is obviously true;

Case 2: $x \in V_{\alpha_n}$ for some (unique!) n. Then $f(x) = h(x) = h_n(x)$ and using (2.2) we have

$$f(x)+\varphi(x) > \varphi(x_n) - \frac{1}{2n} + h_n(x) \geq \varphi(x_n) - \frac{1}{2n} - \varphi(x_n) - 1 + \frac{1}{n} = -1 + \frac{1}{2n} > -1.$$

Hence, (2.3) is also true in this case.

Finally, let us observe that $f(x_n) + \varphi(x_n) = \varphi(x_n) + h_n(x_n) = -1 + 1/n \to -1$ which, together with (2.3) yield that $\inf(f + \varphi) = -1$ and this infimum is not attained according to (2.3). This contradiction shows that X_r must be compact. The proof of the theorem is completed. ∎

3 The case of a separable Banach space

The barrier to making significant applications of Theorem 2.1 above is clear: we need a supply of functions φ that satisfy the hypotheses stated there and have enough additional regularity properties, so that combined with the existence of a minimizer of the perturbation they have interesting consequences. The following result answers this need when X is separable.

Proposition 3.1 *Every separable Banach space X admits a convex lower semicontinuous function $\varphi\colon X \to [0, +\infty]$ whose domain is relatively norm-compact and linearly dense in X. In addition φ possesses the following smoothness property:*

$$\lim_{t \searrow 0} \sup_{h \in \operatorname{dom} \varphi} \frac{\varphi(x+th) + \varphi(x-th) - 2\varphi(x)}{t} = 0, \qquad x \in \operatorname{dom} \varphi. \tag{3.1}$$

Proof. Let $\{x_i;\ i \in \mathbb{N}\}$ be a dense subset of B_X. Define $T\colon \ell_2 \to X$ (ℓ_2 being the usual space of square summable sequences) by

$$T(\alpha) = \sum_{i=1}^{\infty} \alpha_i 2^{-i} x_i, \qquad \alpha = (\alpha_i) \in \ell_2.$$

Clearly, T is a well-defined, linear, and bounded mapping. T is also compact, that is, $T(B_{\ell_2})$ is a norm-compact set. Indeed, take any sequence $\alpha^1, \alpha^2, \ldots$, of elements in B_{ℓ_2}. Since B_{ℓ_2} is weakly compact, there is a subsequence (we do not relabel) along which α^n converges weakly to $\alpha = (\alpha_i)$: in particular, $\alpha \in B_{\ell_2}$ and for each $i \in \mathbb{N}$, $\alpha_i^n \to \alpha_i$ as $n \to \infty$. Let us show that along this same subsequence, $T(\alpha^n)$ norm-converges to $T(\alpha)$. Indeed, all $n, N \in \mathbb{N}$ satisfy

$$\begin{aligned}
\left\|T(\alpha^n) - T(\alpha)\right\| &\le \Big\| \sum_{i=1}^{N} \frac{(\alpha_i^n - \alpha_i) x_i}{2^i} \Big\| + \Big\| \sum_{i=N+1}^{\infty} \frac{(\alpha_i^n - \alpha_i) x_i}{2^i} \Big\| \\
&\le \Big\| \sum_{i=1}^{N} \frac{(\alpha_i^n - \alpha_i) x_i}{2^i} \Big\| + \Big[\sum_{i=N+1}^{\infty} (\alpha_i^n - \alpha_i)^2 \Big]^{\frac{1}{2}} \cdot \Big[\sum_{i=N+1}^{\infty} 2^{-2i} \Big]^{\frac{1}{2}} \\
&\le \Big\| \sum_{i=1}^{N} \frac{(\alpha_i^n - \alpha_i) x_i}{2^i} \Big\| + \Big[2\sum_{i=1}^{\infty} (\alpha_i^n)^2 + 2\sum_{i=1}^{\infty} \alpha_i^2 \Big]^{\frac{1}{2}} \cdot \Big(\frac{4}{3 \cdot 2^{2(N+1)}} \Big)^{\frac{1}{2}},
\end{aligned}$$

and so

$$\limsup_{n \to \infty} \left\|T(\alpha^n) - T(\alpha)\right\| \le 0 + 2^{-N+1}/\sqrt{3}.$$

This holds for every $N \in \mathbb{N}$, so $\limsup_n \|T(\alpha^n) - T(\alpha)\| = 0$, as required.

We have no guarantee that T is injective. Thus we introduce the quotient space $H = \ell_2 / T^{-1}(0)$ of ℓ_2, and factorize $T = S \circ Q$, where $Q\colon \ell_2 \to H$ is the canonical map. It is not difficult to see, by using the parallelogram law, that H, with its quotient norm $\|\cdot\|_H$, is again a Hilbert space. The mapping $S\colon H\to X$ is then linear, bounded, compact, and injective. Also, the range of S, $\mathcal{R}(S)$, is dense in X, since $\mathcal{R}(S)=\mathcal{R}(T)\supset\{x_i;\ i\in\mathbb{N}\}$.

Now we define a function with the desired properties as follows: for $x\in X$, let

$$\varphi(x)=\begin{cases}\tan\left(\|S^{-1}x\|_H^2\right), & \text{if } \|S^{-1}x\|_H^2<\dfrac{\pi}{2},\\ +\infty, & \text{otherwise.}\end{cases}$$

Then $x\in\operatorname{dom}\varphi$ implies that $\|S^{-1}x\|_H^2<\pi/2$, and hence $x=S(S^{-1}x)\in S\big(\sqrt{\frac{\pi}{2}}B_H\big)$, the last set being relatively norm-compact. Also, for every $n\in\mathbb{N}$, the standard basis vector $\mathbf{e}^n$ in ℓ_2 obeys $T(2^n t\mathbf{e}^n)=tx_n$ for all real t. This implies that $S^{-1}(tx_n)$ has small norm in H whenever t is sufficiently close to 0, and hence that $tx_n\in\operatorname{dom}\varphi$ whenever $|t|$ is sufficiently small. Thus $\operatorname{dom}\varphi$ is linearly dense in X.

The function φ is convex because $\tan(\cdot)$ is increasing and convex on $[0,\infty)$ and $\|\cdot\|_H^2$ is convex on H.

Let us show that φ is lower semicontinuous. Fix any $p\in X$ and any sequence $\{p_n\}_{n\in\mathbb{N}}$ in X, with $\|p_n-p\|\to 0$ as $n\to\infty$. We will show that

$$\varphi(p)\le\liminf_{n\to\infty}\varphi(p_n).$$

The inequality is obvious if the right side equals $+\infty$. Assume therefore that the right side is finite. Choose a subsequence $\{p_{n_i}\}_{i\in\mathbb{N}}$ along which

$$\liminf_{n\to\infty}\varphi(p_n)=\lim_{i\to\infty}\varphi(p_{n_i}).$$

Then $\|S^{-1}(p_{n_i})\|_H^2<\pi/2$ for all large i, and indeed

$$\lim_{i\to\infty}\|S^{-1}(p_{n_i})\|_H^2<\frac{\pi}{2}.$$

By passing to a further subsequence if necessary we may assume that the bounded sequence $\{S^{-1}(p_{n_i})\}_{i\in\mathbb{N}}$ converges weakly to some $z\in\ell_2$ and that $\lim_{i\to\infty}\|S^{-1}(p_{n_i})\|_H$ exists: the weak lower semicontinuity of $\|\cdot\|_H$ then gives

$$\|z\|_H^2\le\lim_{i\to\infty}\|S^{-1}(p_{n_i})\|_H^2<\frac{\pi}{2}. \tag{3.2}$$

Now $p_{n_i}=S(S^{-1}p_{n_i})$ converges weakly to Sz, giving $Sz=p$, $z=S^{-1}p$. Using the continuity and monotonicity of $\tan(\cdot)$ on $[0,\pi/2)$ together with (3.2), we obtain the desired inequality:

$$\begin{aligned}\varphi(p)=\tan\big(\|S^{-1}p\|_H^2\big)&=\tan\big(\|z\|_H^2\big)\\&\le\lim_{i\to\infty}\tan\big(\|S^{-1}(p_{n_i})\|_H^2\big)\\&=\lim_{i\to\infty}\varphi(p_{n_i})=\liminf_{n\to\infty}\varphi(p_n).\end{aligned}$$

Finally, take any $x, h \in \operatorname{dom}\varphi$. Then $x \pm th \in \operatorname{dom}\varphi$ for all $t > 0$ sufficiently small, so the parallelogram law gives

$$\|S^{-1}(x+th)\|_H^2 + \|S^{-1}(x-th)\|_H^2 - 2\|S^{-1}(x)\|_H^2 = 2t^2\|S^{-1}(h)\|_H^2. \qquad (3.3)$$

(Yes, S^{-1} is linear.) Hence the convex function $t \mapsto \|S^{-1}(x+th)\|_H^2$ is differentiable at $t = 0$, and the same is true for $t \mapsto \varphi(x+th)$. This establishes the existence of all the directional derivatives

$$\varphi'(x;h) = \lim_{t\to 0} \frac{\varphi(x+th) - \varphi(x)}{t}, \qquad x, h \in \operatorname{dom}\varphi.$$

In fact, since $h \in \operatorname{dom}\varphi$ if and only if $\|S^{-1}(h)\|_H^2 < \pi/2$, then (3.3) supports a type of uniform differentiability:

$$0 = \lim_{t\searrow 0} \sup_{h\in\operatorname{dom}\varphi} \frac{\varphi(x+th) + \varphi(x-th) - 2\varphi(x)}{t}, \qquad x \in \operatorname{dom}\varphi.$$

This completes the proof. ∎

Combining Theorem 2.1 and the above proposition leads to the main result of this section:

Theorem 3.2 *Let X be a separable Banach space. Then there exists a (proper) convex lower semicontinuous function $\varphi\colon X \to [0, +\infty]$ whose domain $\operatorname{dom}\varphi$ is relatively norm-compact and linearly dense in X and which satisfies the smoothness property (3.1). In particular, for every proper lower semicontinuous function $f\colon X \to (-\infty, +\infty]$, the perturbed function $f + \varphi$ attains its infimum.*

4 Applications to differentiability

In this section we see how the variational principle in the case of a separable space can be used to obtain some known differentiability results. First, we show the following well-known result:

Theorem 4.1 *Suppose X is a separable Banach space. Then every continuous function $f : X \to \mathbb{R}$ is Gâteaux differentiable at the points of a generic (that is a dense G_δ) subset of X.*

Proof. First, we will show that f is Gâteaux differentiable at the points of a dense subset of X. After translation, it suffices to show that every non empty open set Ω of X with $0 \in \Omega$, contains a point at which f is Gâteaux differentiable. Fix such an Ω and let φ be the function given by Theorem 3.2. We may suppose $\operatorname{dom}\varphi \subset \Omega$. Then there is some $x \in \operatorname{dom}\varphi \subset \Omega$ at which the function $-f+\varphi$ attains its minimum. In particular, for any $h \in \operatorname{dom}\varphi$ and $t>0$ we have

$$-f(x \pm th) + \varphi(x \pm th) \geq -f(x) + \varphi(x).$$

Using this and the convexity of f we obtain

$$0 \leq f(x+th) + f(x-th) - 2f(x) \leq \varphi(x+th) + \varphi(x-th) - 2\varphi(x)$$

which together with the differentiability property (3.1) of φ shows that

$$\lim_{t \searrow 0} \frac{f(x+th) + f(x-th) - 2f(x)}{t} = 0,$$

for every $h \in \operatorname{dom}\varphi$. Since f is locally Lipschitz and $\operatorname{dom}\varphi$ is linearly dense, in fact, the latter limit is 0 for any $h \in X$. Finally, the fact that f is convex yields its (linear) Gâteaux differentiability at x.

In order to show that the set of points of Gâteaux differentiability of f is *exactly* a G_δ-subset of X, let us observe that property (3.1) yields a stronger conclusion in the argument above: in fact, we obtain that X possesses a dense subset in which every x obeys the following stronger condition

$$\sup\left\{f(x+th) + f(x-th) - 2f(x);\ h \in \operatorname{dom}\varphi\right\} = o(t),\ t \searrow 0. \tag{3.4}$$

On the other hand, the set of all $x \in X$ satisfying (3.4) is always G_δ (possibly empty), see the proof of [11], Proposition 1.25. Therefore, f is Hadamard-like, as well as Gâteaux, differentiable on a dense G_δ-subset of X. ∎

The theorem above can be fairly extended, thus getting a recent result from [3].

Theorem 4.2 ([3]) . *Let Y be a Gâteaux differentiability space and X a separable Banach space. Then $Y \times X$ is a Gâteaux differentiability space.*

Proof. Let $f: Y \times X \to \mathbb{R}$ be a convex continuous function, and $\Omega \subset Y \times X$ be a non empty open set. Assume, for simplicity, that $2B_Y \times 2B_X \subset \Omega$ and that f is bounded on Ω. Let $\varphi: X \to [0, +\infty]$ be the function provided by Theorem 3.2 with domain in B_X. Define $g: Y \to (-\infty, +\infty]$ by

$$g(y) = \begin{cases} \inf\{-f(y,x) + \varphi(x);\ x \in X\}, & \text{if } y \in 2B_Y, \\ +\infty, & \text{if } y \in Y \backslash 2B_Y. \end{cases}$$

Then g is concave and continuous on $2B_Y$. As Y is a Gâteaux differentiability space, the function g is Gâteaux differentiable at some $y \in B_Y$. By Theorem 3.2, there is

$x \in B_X$ so that $g(y) = -f(y,x) + \varphi(x)$. Thus, for every $k \in Y$ and every $h \in \operatorname{dom}\varphi$ we have, for all $t > 0$ sufficiently small,

$$\begin{aligned} f(y+tk, x+th) + \; & f(y-tk, x-th) - 2f(y,x) \\ & \leq -g(y+tk) + \varphi(x+th) \\ & \qquad -g(y-tk) + \varphi(x-th) + 2g(y) - 2\varphi(x) \\ & = o(t) + o(t). \end{aligned}$$

Finally, the local Lipschitzian property of f and the linear density of $\operatorname{dom}\varphi$ in X imply

$$f(y+tk, x+th) + f(y-tk, x-th) - 2f(y,x) = o(t), \quad t \searrow 0$$

for every $k \in Y$ and every $h \in X$. Therefore f is Gâteaux differentiable at the point (y,x). ∎

In light of Theorem 4.2 and [13] it is natural to ask whether the product of a weak Asplund space and a separable space can fail to be a weak Asplund space.

Acknowledgements. The authors wish to thank Philip Loewen for his significant assistance with the development of this paper.

References

[1] E. Bishop and R.R. Phelps, A proof that every Banach space is subreflexive, *Bull. Amer. Math. Soc.* **67**(1961), 97–98.

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

[3] L. Cheng, M. Fabian, The product of a Gâteaux differentiability space and a separable space is a Gâteaux differentiability space, *Proc. Amer. Math. Soc.* **129**(2001), 3539–3541.

[4] M.M. Čoban, P.S. Kenderov and J.P. Revalski, Generic well-posedness of optimization problems in topological spaces, *Mathematika* **36**(1989), 301–324.

[5] M.M. Čoban, P.S. Kenderov and J.P. Revalski, Densely defined selections of multivalued mappings, *Trans. Amer. Math. Soc.*, **344**(1994), 533-552.

[6] R. Deville, G. Godefroy and V. Zizler, A smooth variational principle with applications to Hamilton–Jacobi equations in infinite dimensions, *J. Functional Analysis* **111** (1993), 197–212.

[7] R. Deville, G. Godefroy and V. Zizler, Smoothness and renormings in Banach spaces, Pitman monographs and Surveys in Pure and Appl. Math., Longman Scientific & Technical, 1993.

[8] R. Deville and J.P. Revalski, Porosity of ill-posed problems, *Proc. Amer. Math. Soc.*, **128**(2000), 1117–1124.

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

[10] I.Ekeland, Nonconvex minimization problems, *Bull. Amer. Math. Soc.*, **1**(1979), 443–474.

[11] M. Fabian, Gâteaux differentiability of convex functions and topology – Weak Asplund spaces, J. Wiley & Sons, Interscience, New York, 1997.

[12] P.S. Kenderov and J.P. Revalski, The Banach-Mazur game and generic existence of solutions to optimization problems, *Proc. Amer. Math. Soc.* **118**(1993), 911–917.

[13] W. B. Moors and S. Somasundaram, A Gâteaux differentiability space that is not weak Asplund, *Proc. Amer. Math. Soc.*, in press.

[14] R.R. Phelps, Convex Functions, Monotone Operators and Differentiability, *Lect. Notes in Math.* **#1364**, Springer Verlag, Berlin, 1989.

[15] C. Stegall, Optimization of functions on certain subsets of Banach spaces, *Math. Ann.* **236**(1978), 171–176.

[16] A.N. Tykhonov, Solution of incorectly formulated problems and the regularization method, *Soviet Math. Dokl.*, **4**(1963), 1035–1038.

J. Borwein, CECM, Department of Mathematics, Simon Fraser University, BC, V5A 1S6, Burnaby, Canada; e-mail: jborwein@cecm.math.sfu.ca

M. Fabian, Mathematical Institute, Czech Academy of Sciences, Žitná 25, 11567 Prague, Czech Republic; e-mail: fabian@math.cas.cz

J.P. Revalski, Institute of Mathematics and Informatics, Bulgarian Academy of Sciences, Acad. G. Bonchev Street, block 8, 1113 Sofia, Bulgaria, e-mail: revalski@math.bas.bg