ISRAEL JOURNAL OF MATHEMATICS, Vol. 72, No. 3, 1990

SMOOTH BANACH SPACES, WEAK ASPLUND SPACES AND MONOTONE OR USCO MAPPINGS

BY

DAVID PREISS,[a] R. R. PHELPS[b],† AND I. NAMIOKA[b]

[a]*Department of Mathematics, University College London, Gower St., London WC1E 6BT, UK;*
and [b]*Mathematics Department GN-50, University of Washington, Seattle, WA 98195, USA*

ABSTRACT

It is shown that if a real Banach space E admits an equivalent Gateaux differentiable norm, then for every continuous convex function f on E there exists a dense G_δ subset of E at every point of which f is Gateaux differentiable. More generally, for any maximal monotone operator T on such a space, there exists a dense G_δ subset (in the interior of its essential domain) at every point of which T is single-valued. The same techniques yield results about stronger forms of differentiability and about generically continuous selections for certain upper-semicontinuous compact-set-valued maps.

0. Introduction

Let E be a real Banach space, D a nonempty open convex subset of E and f a real-valued continuous convex function on D. Recall that f is said to be *Gateaux differentiable* at the point x in D if the limit

$$(*) \qquad df(x)(y) = \lim_{t\to 0} t^{-1}[f(x+ty) - f(x)]$$

exists for all $y \in E$. When this is the case, the limit is a continuous linear function of y, denoted by $df(x)$. If the difference quotient in $(*)$ converges to $df(x)(y)$ uniformly for y in the unit ball, then f is said to be *Fréchet* differentiable at x. The norm in E is said to be *smooth* if it is Gateaux differentiable at each $x \neq 0$; one then also says that the space E itself is smooth. Finally, E is called a *weak Asplund space* [*Asplund space*] or said to have the weak Asplund property if for every f and D as above, f is "generically" Gateaux [Féchet] differentiable, that is, there

†Work on this paper by the second-named author was supported in part by NSF Grant DMS 8700284.
Received February 15, 1990

exists a dense G_δ subset G of D such that f is Gateaux [Fréchet] differentiable at each point of G. In 1933 Mazur [Ma] proved (in this terminology) that separable Banach spaces have the weak Asplund property. In 1968 Asplund showed that there is a close connection between smooth norms and differentiability of convex functions when he proved [As] that E is a weak Asplund space [he used the term "weak differentiability space"] if E^* admits an equivalent strictly convex dual norm (hence the norm induced in E is smooth). Since the dual of a smooth norm need not be strictly convex (in fact, there even exist smooth Banach spaces which cannot be renormed to have strictly convex dual [Ta]), it became a natural question whether the existence of an equivalent smooth norm implies the weak Asplund property. (See, for instance, [Da].) A partial answer was obtained in 1987 by Borwein and Preiss [Bo–Pr] who showed that it implied that E is a *Gateaux Differentiability Space* (GDS); that is, every continuous convex function f is Gateaux differentiable at the points of a dense subset of its domain. In the present paper we show that equivalent smooth norms *do* imply the weak Asplund property. Somewhat more generally, we show that smoothness of the norm guarantees that each maximal monotone operator on E has a dense G_δ subset of its effective domain where it is single-valued. We recall the relevant definitions.

Definition. A set-valued mapping $T: E \to 2^{E^*}$ from E into subsets of E^* is said to be *monotone* provided

$$0 \le \langle x^* - y^*, x - y \rangle \qquad \text{whenever } x, y \in E, \quad x^* \in T(x) \quad \text{and} \quad y^* \in T(y).$$

The *effective domain* of T is $D(T) = \{x \in E : T(x) \neq \varnothing\}$. Such a mapping is said to be *maximal* monotone provided its graph

$$\{(x, x^*) \in E \times E^* : x^* \in T(x)\}$$

is maximal (under inclusion) in the collection of all graphs of monotone operators.

Definition. If f is a continuous convex function on an open convex subset D of E, the set-valued *subdifferential* $\partial f(x)$ of f at $x \in D$ is

$$\{x^* \in E^* : \langle x^*, y - x \rangle \le f(y) - f(x) \text{ for all } y \in D\}.$$

It is well known (see, for instance, [Ph]) that ∂f is a maximal monotone operator and that f is Gateaux differentiable at x if and only if $\partial f(x)$ is a singleton. Thus, theorems which yield generic single-valuedness of maximal monotone operators are generalizations of theorems about generic differentiability of convex functions. An early extension of this kind was obtained in 1973 by Zarantonello [Za], who generalized Mazur's theorem by showing that a maximal monotone operator T on a separable Banach space is single-valued at the points of a dense G_δ subset of the interior of $D(T)$. In 1974, Kenderov [Ke$_1$] extended Asplund's theorem, showing generic single-valuedness of maximal monotone operators on E whenever E can be equivalently renormed to have a strictly convex dual norm.

The reader who is primarily interested in convex functions can substitute subdifferentials for maximal monotone operators wherever the latter appear.

In Section 2, the same methods are used to show that smooth Banach spaces lie in the class (S), a subclass of the weak Asplund spaces which admits more permanence properties than the latter and which is defined in terms of selections for certain upper-semicontinuous compact-set-valued (usco) mappings. In fact, we obtain certain stronger conclusions, provided E admits an equivalent norm which is differentiable in a stronger sense (say, Fréchet or Hadamard).

In a remarkable example, Haydon [Ha] has shown that the converse to our main theorem is not valid: There exists a weak Asplund space—in fact, an Asplund space—which does not admit an equivalent smooth norm. This solves the other half of a long-standing open problem. Since Kenderov [Ke$_2$] has shown that maximal monotone operators on any Asplund space are generically single-valued (even generically norm-to-norm continuous), the conclusion to our main theorem does not imply the existence of an equivalent smooth norm.

1. Monotone operators

A maximal monotone operator T behaves very much on the interior D of $D(T)$ like the subdifferential of a continuous convex function. Some of the properties they share are well-known; they are listed in the following proposition. The proofs may be found, for instance, in [Ph].

1.1. Proposition. *Let $T: E \to 2^{E^*}$ be maximal monotone and assume that*

$$D = \operatorname{int} D(T) \equiv \operatorname{int}\{x \in E : T(x) \neq \varnothing\}$$

is nonempty. Then

(i) *T is locally bounded in D. [That is, if $x \in D$, then there exists an open neighborhood U of x in D and $M > 0$ such that $\|y^*\| \leq M$ whenever $y \in U$ and $y^* \in T(y)$.]*

(ii) *$T(x)$ is weak*-compact and convex, for all $x \in D$.*

(iii) *T is norm-to-weak* upper semicontinuous in D. [That is, for all x in D and any weak* open set W containing $T(x)$ in E^*, there exists an open neighborhood U of x in D such that $T(y) \subset W$ for all $y \in U$.]*

We next list some elementary properties of maximal monotone operators which are less standard. They arose out of the need to mimic properties of the right-hand derivative d^+f of a continuous convex function f. Recall that for such a function f on D, d^+f can be characterized by the following well-known formula (see, for instance, [Ph, p. 27]) (when A is a subset of E^* and $e \in E$, we use the notation $\sup\{\langle A, e\rangle\}$ in place of $\sup\{\langle x^*, e\rangle : x^* \in A\}$)

$$d^+f(x)(e) = \sup\{\langle \partial f(x), e\rangle\}, \qquad x \in D, \quad e \in E.$$

This suggests the following substitute for the right-hand derivative. The same function (with different notation) was used by Zarantonello [Za].

Definition. With T and D as above, let

$$\sigma_T(x, e) = \sup\{\langle T(x), e\rangle\}, \qquad x \in D, \quad e \in E.$$

(We will usually write $\sigma(x, e)$ instead of $\sigma_T(x, e)$.)

We will also need a substitute notion which parallels the situation when the directional derivative $df(x)(e)$ of f exists (at $x \in D$ in the direction $e \in E$). Recall that this will be the case if and only if $d^+f(x)(-e) = -d^+f(x)(e)$. From the formula above for d^+f, it follows that this is equivalent to

$$\sup\{\langle \partial f(x), e\rangle\} = \inf\{\langle \partial f(x), e\rangle\}.$$

This motivates the following definition.

Definition. For any $e \in E$ and T as above, let eT denote the set-valued mapping from E into the real line defined by

$$(eT)(x) = \{\langle x^*, e\rangle : x^* \in T(x)\}.$$

Our substitute for saying that $df(x)(e)$ exists will be the assertion that eT *is single-valued* at x.

1.2. Proposition. *Let T and D be as above. Then*

(i) *For each $x \in D$, the real-valued function $e \to \sigma_T(x, e)$ is subadditive and positive homogeneous and, for any $\lambda > 0$, $\sigma_{\lambda T}(x, e) = \lambda\sigma_T(x, e)$.*

(ii) *For each $x \in D(T)$,*

$$\sup\{\sigma(x, e) : \|e\| = 1\} = \sup\{\sigma(x, e) : \|e\| \le 1\} = \sup\{\|x^*\| : x^* \in T(x)\}.$$

(iii) $(eT)(x)$ *is a singleton if and only if* $\sigma(x, -e) = -\sigma(x, e)$.

(iv) *If $x_0 \in D$, $e \in E$ and $(eT)(x_0)$ is a singleton, then $x \to \sigma(x, e)$ is continuous at x_0.*

(v) *Fixing* $x \in D$ *and* $e \neq 0$, *letting* $I = \{t \in \mathbf{R} : x + te \in D\}$ *and defining*

$$f(t) = \sigma(x + te, e), \qquad t \in I,$$

the function f *is monotone nondecreasing on* I (*and hence is continuous at all but countably many points of* I). *Moreover, if* f *is continuous at* $t_0 \in I$, *then* $(eT)(x + t_0 e)$ *is a singleton*).

PROOFS. (i) and (ii) are immediate from the definitions.

(iii) If $(eT)(x)$ is a singleton, then so is $(-eT)(x)$, with $(-eT)(x) = -\sigma(x, e)$; that is, $\sigma(x, -e) = -\sigma(x, e)$. On the other hand, if $(eT)(x)$ is not a singleton, there exists $x^* \in T(x)$ such that $\langle x^*, e\rangle < \sigma(x, e)$ and hence $\sigma(x, -e) \geq \langle x^*, -e\rangle > -\sigma(x, e)$.

(iv) Given $\epsilon > 0$, let $W = \{y^* \in E^* : |\langle y^*, e\rangle - \sigma(x_0, e)| < \epsilon\}$; by norm-to-weak* upper semicontinuity of T at x_0, there exists an open neighborhood of x_0 in D such that $T(x) \subset W$ for all $x \in U$. It follows easily from this that $|\sigma(x, e) - \sigma(x_0, e)| \leq \epsilon$ for all $x \in U$.

(v) Note that we are not asserting that the open set I is an interval, but this does not affect our argument. Suppose that $t_1, t_2 \in I$ with $t_1 < t_2$; then for any $x_i^* \in T(x + t_i e)$, $i = 1,2$, we have (by monotonicity)

$$0 \leq \langle x_1^* - x_2^*, (x + t_1 e) - (x + t_2 e)\rangle = (t_1 - t_2)\langle x_1^* - x_2^*, e\rangle,$$

hence $\langle x_1^*, e\rangle \leq \langle x_2^*, e\rangle$. This shows that

$$(\#) \qquad f(t_1) \equiv \sup\{\langle T(x + t_1 e), e\rangle\} \leq \inf\{\langle T(x + t_2 e), e\rangle\};$$

in particular, $f(t_1) \leq f(t_2)$, so f is monotone. Suppose, now, that eT is not single-valued at $x + t_0 e$. Then

$$\alpha \equiv \inf\{\langle T(x + t_0 e), e\rangle\} < f(t_0).$$

Hence if $t \in I$, $t < t_0$, then (by (#)), $f(t) \leq \alpha < f(t_0)$, so f is not continuous at t_0.

There are two lemmas which are central to the proof of our main theorem. The first is the Banach-Mazur game; it allows us to obtain a dense G_δ set of points of differentiability (or of single-valuedness, in the case of monotone operators). The second—very simple—lemma is singled out because it exhibits clearly the connection between differentiability of the norm and single-valuedness of T.

DEFINITIONS. Let X be a Hausdorff space, S a subset of X and let A and B denote the players of the game. A *play* is a decreasing sequence of nonempty open sets $U_1 \supset V_1 \supset U_2 \supset V_2 \supset \cdots$ which have been chosen alternately; the U_k's by A, the V_k's by B. Player B is said to have *won a play* if $\bigcap V_n \subset S$; otherwise, A is said to have won. (It is not required that the intersection be nonempty.) Player B is said to have a *winning* strategy if, using it, she wins every play, independently of player A's choices. (A more detailed description of the Banach–Mazur game and a proof of the following lemma may be found in [Ox].)

1.3. Lemma. *If player B has a winning strategy, then S is a residual set (that is, $X \backslash S$ is of first category). In particular, if X is a completely metrizable space, then S must contain a dense G_δ subset.*

Here is the other key lemma.

1.4. Lemma. *Suppose that $x_0 \in D$ and that $e_0 \in E$, $\|e_0\| = 1$, is such that $e_0 T$ is single-valued at x_0, with value α. If*

$$\sup\{\sigma_T(x_0, e) : \|e\| = 1\} \leq \alpha,$$

then

$$T(x_0) \subset \alpha \cdot \partial \| \cdot \| (e_0).$$

In particular, if the norm is Gateaux differentiable at e_0, then $T(x_0)$ is a singleton.

Proof. Suppose that $x^* \in T(x_0)$; then from Proposition 1.2(ii), we see that $\alpha \geq 0$ and $\|x^*\| \leq \alpha$. Since $(e_0 T)(x_0)$ is a singleton, it follows that $\langle x^*, e_0 \rangle = \alpha$ and so $x^* \in \alpha \cdot \partial \| \cdot \| (e_0)$.

In our main theorem we will assume that the interior D of the effective domain $D(T)$ of our maximal monotone operator T is nonempty and we will produce a dense G_δ subset G of D where T is single-valued. This is not a loss of generality, because G will also be a dense G_δ subset of $D(T)$; this is a consequence of the following elementary proposition.

1.5. Proposition. *Suppose that $T : E \to 2^{E^*}$ is a maximal monotone operator, with effective domain $D(T)$; then there exists a first category F_σ subset A of E such that $D(T)$ is the disjoint union of A and the interior D of $D(T)$.*

Proof. The main observation is that $D(T)$ is itself an F_σ set. Indeed, let B^* denote the closed unit ball of E^* and, for $n \geq 1$, let

$$F_n = \{x \in E : T(x) \cap n \cdot B^* \neq \emptyset\}.$$

Clearly, $D(T) = \bigcup F_n$. It is a fairly standard argument, using the weak* compactness of $n \cdot B^*$ and the maximal monotonicity of T, to show that each F_n is closed.

Finally, let $A_n = F_n \backslash D$; each of these sets is closed, has empty interior and is disjoint from D, so we let $A = \cup A_n$.

In what follows, we use the standard notation

$$B(x, r) = \{y \in E : \|y - x\| < r\} \qquad (x \in E \text{ and } r > 0),$$

while $B^*(x^*, r)$ denotes the analogous open ball in E^*.

1.6. THEOREM. *Suppose that E admits an equivalent smooth norm, that T is a maximal monotone operator on E and that*

$$D \equiv \operatorname{int}\{x \in E : T(x) \neq \varnothing\}$$

is nonempty. Then there exists a dense G_δ subset $G \subset D$ such that $T(x)$ is a singleton for every $x \in G$.

PROOF. Let $p_1 = \|\cdot\cdot\|$ denote an equivalent smooth norm on E. Choose sequences of positive numbers $1/2 > \epsilon_1 > \epsilon_2 > \cdots$ and $\beta_1 > \beta_2 > \beta_3 > \cdots$ such that

$$\epsilon_k \to 0, \quad \Sigma\beta_k^2 < 3 \quad \text{and} \quad \Sigma\sqrt{\epsilon_k}/\beta_k < \infty.$$

To use the Banach–Mazur game we let D be our Hausdorff space and S the set of points in D where T is single-valued. After player A chooses a nonempty open subset U of D, player B's first choice will always be an open subset of U in which T is bounded. Thus, we may always assume that player A's first choice was an open nonempty set U_1 in which T is bounded. We may also assume that

$$\sup\{\|x^*\| : x^* \in T[U_1]\} > 0.$$

[Indeed, if this supremum is 0, then B's strategy is obvious: Since T is single-valued (equal to 0) on the entire set U_1, she need only choose $V_k = U_k$ for each $k = 1, 2, 3, \ldots$ so that $\cap V_k = U_1 \subset S$.] Now, let S_1 be the unit sphere $\{x \in E : p_1(x) = 1\}$ defined by p_1 and define

$$s_1 = \sup\{\sigma(x, e) : (x, e) \in U_1 \times S_1\}.$$

From Proposition 1.2(ii), we can also write

$$s_1 = \sup\{\|x^*\| : x^* \in T(x) \text{ and } x \in U_1\};$$

by our earlier assumption, $s_1 > 0$. From part (v) of Proposition 1.2, for any $e \neq 0$ and $x \in U_1$, there exist points of the form $x + te \in U_1$ (with $t > 0$ arbitrarily small) such that $\sigma(x, e) \leq \sigma(x + te, e)$ and eT is single-valued at $x + te$. Thus

$$s_1 = \sup\{\sigma(x, e) : (x, e) \in U_1 \times S_1 \text{ and } eT \text{ is single-valued at } x\}.$$

It follows that there exists $(x_1, e_1) \in U_1 \times S_1$ such that e_1T is single-valued at x_1 and such that $\sigma(x_1, e_1) > (1-\epsilon_1)s_1$. By continuity of the function $x \to \sigma(x, e_1)$ at x_1, there exists r_1, with $0 < r_1 < 1$, such that $B(x_1, 2r_1) \subset U_1$ and $\sigma(x, e_1) > (1-\epsilon_1)s_1$ for $x \in B(x_1, 2r_1)$. Define $V_1 = B(x_1, r_1)$. For all $x \in E$ define

$$q_1(x) = \mathrm{dist}(x, \mathbf{R}e_1) \equiv \inf\{\|x - \lambda e_1\| : \lambda \in \mathbf{R}\}$$

and define a new norm p_2 on E by

$$p_2^2 = p_1^2 + \beta_1^2 q_1^2.$$

We now have s_1, e_1, x_1, p_2 and player B's open set V_1, so player A may choose any nonempty open subset $U_2 \subset V_1$. Using a similar strategy to respond to player A's choices $U_2, U_3, \ldots$ at every step, player B chooses $V_1, V_2, \ldots$ by constructing sequences of numbers $s_1, s_2, \ldots, s_k, \ldots$, norms $p_1, p_2, \ldots, p_k, \ldots$, spheres $S_1, S_2, \ldots, S_k = \{x \in E : p_k(x) = 1\}, \ldots$ containing the vectors $e_1, e_2, \ldots, e_k, \ldots$ respectively, points $x_1, x_2, \ldots, x_k, \ldots$ and positive radii $r_1, r_2, \ldots, r_k, \ldots$ such that $V_k = B(x_k, r_k)$,

$$p_k^2 = p_{k-1}^2 + \beta_{k-1}^2 \cdot q_{k-1}^2 \equiv p_1^2 + \sum_{j=1}^{k-1} \beta_j^2 q_j^2$$

(where $q_{k-1}(x) = \inf\{\|x - \lambda e_{k-1}\| : \lambda \in \mathbf{R}\}$, $x \in E$),

$$s_k = \sup\{\sigma(x, e) : (x, e) \in U_k \times S_k\},$$

e_kT is single-valued at x_k and

$$\sigma(x, e_k) > (1-\epsilon_k)s_k \qquad \text{for } x \in B(x_k, 2r_k) \subset U_k.$$

Since $B(x_k, 2r_k) \subset U_k \subset V_{k-1} = B(x_{k-1}, r_{k-1})$, we have

$$r_k \leq (1/2)r_{k-1} \leq (1/2^2)r_{k-2} \leq \cdots \leq (1/2^{k-1})r_1 < 1/2^{k-1}$$

as well as $\bar{V}_k \subset U_k$ for each k. Also, $p_k \geq p_{k-1}$ implies that

$$S_k \subset B_{k-1} = \{x \in E : p_{k-1}(x) \leq 1\},$$

so, using $U_k \subset U_{k-1}$ and Proposition 1.2(ii),

$$\begin{aligned} s_k &\leq \sup\{\sigma(x, e) : (x, e) \in U_{k-1} \times B_{k-1}\} \\ &= \sup\{\sigma(x, e) : (x, e) \in U_{k-1} \times S_{k-1}\} = s_{k-1}. \end{aligned}$$

Note that $p_k(e_{k-1}) = p_{k-1}(e_{k-1}) = 1$ and $x_k \in U_k \subset U_{k-1}$, so necessarily $(x_k, e_{k-1}) \in U_k \times S_k$ hence $s_k \geq \sigma(x_k, e_{k-1})$, while $x_k \in U_{k-1}$ implies that

$\sigma(x_k, e_{k-1}) > (1 - \epsilon_{k-1})s_{k-1}$. Thus, $s_k > (1 - \epsilon_{k-1})s_{k-1}$. Since $s_1 > 0$ and $\epsilon_1 < 1$, this implies that $s_2 > 0$; by induction, $s_k > 0$ for all k. It follows that the decreasing sequence $\{s_k\}$ converges to a number $s_\infty \geq 0$. Also, since the diameters of the sets V_k converge to zero, the intersections of their closures consists of a single point, denoted by x_∞; necessarily, $x_\infty \in V_k$ for all k. In order to be able to apply Lemma 1.4 and complete the proof, we will show that

(i) the sequence of norms $\{p_k\}$ converges to an equivalent smooth norm p_∞ satisfying $p_1 \leq p_\infty \leq 2 \cdot p_1$,

(ii) the sequence of vectors $\{e_k\}$ is convergent, with limit e_∞ satisfying $p_\infty(e_\infty) = 1$ and

(iii) $e_\infty T$ is single-valued at x_∞, with value s_∞.

Assuming that we have proved assertions (i), (ii) and (iii), we next prove the inequality (1) in Lemma 1.4, taking, of course, x_∞ and e_∞ in place of x_0 and e_0, s_∞ in place of α and p_∞ for the norm; that is, we want to show that $\sigma(x_\infty, e) \leq s_\infty$ whenever $p_\infty(e) = 1$. For each k, x_∞ is in U_k and $p_k(e)^{-1}e \in S_k$ and therefore

$$\sigma(x_\infty, e) = p_k(e) \cdot \sigma(x_\infty, p_k(e)^{-1}e) \leq s_k \cdot p_k(e).$$

This last term converges to $s_\infty \cdot p_\infty(e) = s_\infty$, so the inequality (1) of Lemma 1.4 is satisfied and therefore $T(x_\infty) \subset s_\infty \cdot \partial p_\infty(e_\infty)$. Since p_∞ is smooth, $T(x_\infty)$ is a singleton; that is $\{x_\infty\} = \cap V_k \subset S$.

We have thus described a winning strategy for player B, so by Lemma 1.3, S contains a dense G_δ subset (the open set D being completely metrizable).

It now remains to prove assertions (i), (ii) and (iii).

(i) Since $q_k \leq p_1$ for each k, we have

$$p_1^2 \leq p_k^2 = p_1^2 + \sum_{j=1}^{k-1} \beta_j^2 q_j^2 \leq \left(1 + \sum_{j=1}^{k-1} \beta_j^2\right) \cdot p_1^2 \leq 4 \cdot p_1^2.$$

The second inequality shows that p_k^2 is the partial sum of a series which is uniformly convergent on bounded sets. Thus, the increasing sequence $\{p_k\}$ of equivalent norms converges uniformly on bounded sets to a norm p_∞ satisfying $p_1 \leq p_\infty \leq 2p_1$. To prove that p_∞ is Gateaux smooth, note first that each of the functions q_k^2 is everywhere Gateaux differentiable. In fact, if $q_k^2(y) = 0$, then it is easily seen that $\partial q_k^2(y) = \{0\}$. If $q_k^2(y) > 0$, then for some $\lambda_0 \in \mathbf{R}$,

$$0 < q_k(y) \equiv \inf\{\| y - \lambda e_k \| : \lambda \in \mathbf{R}\} = p_1(y - \lambda_0 e_k)$$

and hence, for any $t > 0$ and $u \in E$,

$$(**)\qquad 0 \le t^{-1}[q_k(y+tu)+q_k(y-tu)-2q_k(y)]$$
$$\le t^{-1}[p_1(y-\lambda_0 e_k+tu)+p_1(y-\lambda_0 e_k-tu)-2p_1(y-\lambda_0 e_k)].$$

Since Gateaux differentiability is characterized by the fact that such difference quotients converge to zero (see, for instance, [Ph, p. 15]), Gateaux differentiability of p_1 at points where it is positive implies that the last term in (**) converges to zero, hence so does the first term. It now follows by standard arguments that the infinite series defining p_∞^2 is everywhere Gateaux differentiable, hence p_∞ is Gateaux differentiable at nonzero points.

(ii) To show that $\{e_k\}$ converges, we will show that for all sufficiently large k,

$$\|e_{k+1}-e_k\| \le 6\sqrt{\epsilon_k}/\beta_k;$$

since the series $\Sigma\sqrt{\epsilon_k}/\beta_k$ converges, this will suffice to show that $\{e_k\}$ is a Cauchy sequence, hence is convergent to e_∞, say. To obtain the estimate above, recall first that $s_{k+1}>(1-\epsilon_k)s_k$ and $\epsilon_k>\epsilon_{k+1}$ so

$$\sigma(x_{k+1},e_{k+1})>(1-\epsilon_{k+1})s_{k+1}>(1-\epsilon_{k+1})(1-\epsilon_k)s_k>(1-2\epsilon_k)s_k>0,$$

while—since $(x_{k+1},p_k(e_{k+1})^{-1}e_{k+1})\in U_k\times S_k$—

$$\sigma(x_{k+1},e_{k+1})=p_k(e_{k+1})\cdot\sigma(x_{k+1},p_k(e_{k+1})^{-1}e_{k+1})\le s_k\cdot p_k(e_{k+1}).$$

Thus $p_k(e_{k+1})>1-2\epsilon_k>0$. By the definition of p_{k+1},

$$\beta_k^2 q_k^2(e_{k+1})=p_{k+1}^2(e_{k+1})-p_k^2(e_{k+1})<1-(1-2\epsilon_k)^2<4\epsilon_k.$$

Consequently, $\operatorname{dist}(e_{k+1},\mathbf{R}e_k)\equiv q_k(e_{k+1})<2\sqrt{\epsilon_k}/\beta_k$. We can write a nearest point in $\mathbf{R}e_k$ to e_{k+1} in the form $\lambda_k e_k$ for some $\lambda_k\in\mathbf{R}$; that is

$$e_{k+1}=\lambda_k e_k+u_k,\qquad \text{where } \|u_k\|=q_k(e_{k+1})<2\sqrt{\epsilon_k}/\beta_k.$$

Now, $p_{k+1}\le p_\infty\le 2p_1$, so $p_{k+1}(u_k)<4\sqrt{\epsilon_k}/\beta_k$. Since, as noted earlier, $p_{k+1}(e_k)=p_k(e_k)=1$, we have

$$|\lambda_k|=p_{k+1}(\lambda_k e_k)=p_{k+1}(e_{k+1}-u_k)\ge 1-p_{k+1}(u_k)\ge 1-4\sqrt{\epsilon_k}/\beta_k.$$

Also,

$$|\lambda_k|\le p_{k+1}(e_{k+1})+p_{k+1}(-u_k)\le 1+4\sqrt{\epsilon_k}/\beta_k.$$

We next show that if k is large enough so that $\epsilon_{k+1}+4\sqrt{\epsilon_k}/\beta_k<1$, then the constant λ_k is positive. Indeed, since $x_{k+1}\in V_k$ for all k, we have $0<(1-\epsilon_k)s_k<$

$\sigma(x_{k+1}, e_k)$. Choose an element $x_k^* \in T(x_{k+1})$ such that $\langle x_k^*, e_{k+1}\rangle > (1-\epsilon_k)s_k$. Since $(e_{k+1}T)(x_{k+1})$ is a singleton, we must have

$$\langle x_k^*, e_{k+1}\rangle = \sigma(x_{k+1}, e_{k+1}) > (1-\epsilon_{k+1})s_{k+1}.$$

Using Proposition 1.2(ii) applied to the norm p_{k+1} and its dual norm p_{k+1}^*, we have $p_{k+1}^*(x_k^*) \le s_{k+1}$; also $p_{k+1}(u_k) \le 4\sqrt{\epsilon_k}/\beta_k$. It follows that

$$\begin{aligned}(1-\epsilon_{k+1})s_{k+1} &< \langle x_k^*, \lambda_k e_k + u_k\rangle = \lambda_k\langle x_k^*, e_k\rangle + \langle x_k^*, u_k\rangle \\ &\le \lambda_k\langle x_k^*, e_k\rangle + p_{k+1}^*(x_k^*)p_{k+1}(u_k) \le \lambda_k\langle x_k^*, e_k\rangle + s_{k+1}\cdot 4\sqrt{\epsilon_k}/\beta_k,\end{aligned}$$

therefore $\lambda_k\langle x_k^*, e_k\rangle > (1-\epsilon_{k+1} - 4\sqrt{\epsilon_k}/\beta_k)s_{k+1}$. By hypothesis, this last term is positive and since $\langle x_k^*, e_k\rangle > 0$ we conclude that $\lambda_k > 0$ for all sufficiently large k. It follows that our earlier estimates on $|\lambda_k|$ become $1 - 4\sqrt{\epsilon_k}/\beta_k \le \lambda_k \le 1 + 4\sqrt{\epsilon_k}/\beta_k$, that is, $|1-\lambda_k| \le 4\sqrt{\epsilon_k}/\beta_k$. Consequently, using the fact that $\|e_k\| \equiv p_1(e_k) \le p_k(e_k) = 1$, we finally obtain the desired estimate:

$$\begin{aligned}\|e_{k+1} - e_k\| &= \|(\lambda_k - 1)e_k + u_k\| \le |\lambda_k - 1|\cdot\|e_k\| + \|u_k\| \\ &\le |\lambda_k - 1| + \|u_k\| \le 4\sqrt{\epsilon_k}/\beta_k + 2\sqrt{\epsilon_k}/\beta_k = 6\sqrt{\epsilon_k}/\beta_k.\end{aligned}$$

The fact that $p_\infty(e_\infty) = 1$ is immediate from the fact that $p_k \to p_\infty$ uniformly on bounded sets, since $p_k(e_k) = 1$ and $e_k \to e_\infty$.

(iii) Finally, we show that $e_\infty T$ is single-valued at x_∞, with value s_∞. To that end, suppose that $x^* \in T(x_\infty)$. Since $x_\infty \in U_k$ for all k and since $1 = p_\infty(e_\infty) \ge p_k(e_\infty)$, the fact (proved earlier) that

$$s_k = \sup\{\sigma(x, e) : x \in U_k,\ p_k(e) \le 1\}$$

implies that $s_k \ge \sigma(x_\infty, e_\infty) \ge \langle x^*, e_\infty\rangle$, hence $s_\infty \ge \langle x^*, e_\infty\rangle$. On the other hand, since $x^* \in T[U_1]$, we have $\|x^*\| \le s_1$ and $x_\infty \in V_{k-1}$ for all k, so

$$\langle x^*, e_\infty\rangle \ge \langle x^*, e_k\rangle - s_1\|e_k - e_\infty\| > (1-\epsilon_k)s_k - s_1\|e_k - e_\infty\|;$$

it follows that $\langle x^*, e_\infty\rangle = s_\infty$ and the proof is complete.

2. β-smooth norms and Banach spaces of class (S)

DEFINITIONS. A map φ from a Hausdorff space Ω to the power set 2^X of another Hausdorff space X is said to be *upper-semicontinuous* if, for each open subset U of X, the set $\{t \in \Omega : \varphi(t) \subset U\}$ is open (possibly empty) in Ω. If, in addition, $\varphi(t)$ is nonempty and compact for each t in Ω, φ is said to be a *usco map*.

Assertions (ii) and (iii) of Proposition 1.1 imply that any maximal monotone operator $T: E \to 2^{E^*}$, when restricted to $\operatorname{int} D(T)$, is a usco map, so the results of this section have implications for maximal monotone operators.

DEFINITIONS. The *graph* G_φ of a map $\varphi: \Omega \to 2^X$ is the subset of the product space $\Omega \times X$ consisting of all pairs (t, x) such that $x \in \varphi(t)$. It is clear that if φ is a usco map, then its graph is closed in $\Omega \times X$ and projects onto Ω. Furthermore, if F is a closed subset of G_φ that projects onto Ω, then F is itself the graph of a usco map $\psi: \Omega \to 2^X$ [Chr]. In this case, φ is said to *contain* ψ. By Zorn's lemma, each usco map from Ω into 2^X contains a minimal usco map from Ω into 2^X.

DEFINITION. A Hausdorff space X is said to be of *type* $\mathcal{S}$ if, whenever φ is a usco map from a Baire space Ω to 2^X, there exists a function $\sigma: \Omega \to X$ such that $\sigma(t) \in \varphi(t)$ for all t and the set of points of continuity of σ is residual in Ω, that is, the selection σ is continuous at each point of a dense G_δ subset of Ω.

The following characterization of spaces of type $\mathcal{S}$ is so useful that we will use it in place of the definition: A Hausdorff space X is of type $\mathcal{S}$ if and only if, whenever $\varphi: \Omega \to 2^X$ is a *minimal* usco map and Ω is a Baire space, then the set $\{t \in \Omega : \varphi(t) \text{ is a singleton}\}$ is residual in Ω.

DEFINITION. A Banach space E is said to be of *class* (S) if (E^*, weak^*) is of type $\mathcal{S}$.

These classes originated with Stegall, who has shown [St_1, St_2] that each Banach space of class (S) is a weak Asplund space and that the class (S) spaces possess good permanence properties. (See Section 3.) On the other hand (as we discuss in Section 3), many questions concerning the permanence of weak Asplund spaces are open. In recent years, various authors have shown that Banach spaces of class (S) have somewhat stronger differentiation properties for convex functions than are needed to yield weak Asplund spaces. (See, for instance, [B-F-K], [Noll], [Rain] and [V-V].) In what follows we show, in particular, that a Banach space with an equivalent smooth norm is of class (S) (Corollary 2.6). This result is an immediate consequence of the main theorem of this section, which concerns Banach spaces with an equivalent "β-smooth norm", as defined below.

DEFINITION. Let β be a family of nonempty bounded subsets S of E satisfying (a) $\lambda S \in \beta$ whenever $\lambda \in \mathbf{R}$ and $S \in \beta$ and (b) the union of the members of β is all of E. We say that a continuous convex function f on an open convex subset D of E is *β-differentiable* at $x \in D$ if for all $S \in \beta$ the limit of the difference quotient in (*) exists *uniformly for y in S*. Natural choices for β are all finite sets (Gateaux differentiability), all weakly compact sets (Hadamard differentiability) or all bounded sets (Fréchet differentiability). If the norm is β-differentiable at every $x \neq 0$, it is said to be *β-smooth*.

There is a corresponding notion for maximal monotone operators, in fact, for general set-valued mappings into E^*.

DEFINITION. Let Ω be a Hausdorff space and φ a set-valued mapping from Ω into E^*. We say that φ is *β-continuous* at a point $t \in \Omega$ provided $\varphi(t)$ is a singleton and, for every set S in β, there exists an open neighborhood U of t in Ω such that

$$\varphi(U) \subset \varphi(t) + S^0,$$

where $S^0 = \{z^* \in E^* : \langle z^*, z\rangle \leq 1 \text{ for all } z \in S\}$.

In Corollary 2.9 we will prove a β-continuity result for monotone operators. To see that it has some relevance for convex functions, we need the following proposition.

2.1. PROPOSITION. *Suppose that f is a continuous convex function on an open convex subset D of E and that β is a family of sets as described above. Then the subdifferential mapping ∂f is β-continuous at the point $x_0 \in D$ if and only if f is β-differentiable at x_0.*

PROOF. Suppose that ∂f is β-continuous at x_0. Given $S \in \beta$, $\epsilon > 0$ and $\{x_0^*\} = \partial f(x_0)$, we want to show that there exists $\delta > 0$ such that

$$(0 \leq)t^{-1}[f(x_0 + tu) - f(x_0) - \langle x_0^*, u\rangle] \leq \epsilon$$

whenever $u \in S$ and $0 < t < \delta$. By definition of β-continuity, there exists $\delta_0 > 0$ such that $B(x_0, \delta_0) \subset D$ and $\langle y^* - x_0^*, u\rangle \leq \epsilon$ whenever $y \in B(x_0, \delta_0)$, $y^* \in \partial f(y)$ and $u \in S$. Since S is bounded, we can choose $\delta > 0$ such that $\|tu\| < \delta_0$ whenever $0 < t < \delta$ and $u \in S$. Thus, whenever $0 < t < \delta$ and $u \in S$, we have $x_0 + tu \in B(x_0, \delta_0)$. Consequently, choosing any $y^* \in \partial f(x_0 + tu)$, we have $\langle y^* - x_0^*, u\rangle \leq \epsilon$ and

$$-t\langle y^*, u\rangle = \langle y^*, x_0 - (x_0 + tu)\rangle \leq f(x_0) - f(x_0 + tu),$$

so that

$$0 \leq t^{-1}[f(x_0 + tu) - f(x_0) - \langle x_0^*, u\rangle] \leq \langle y^* - x_0^*, u\rangle \leq \epsilon.$$

To prove the converse, one can follow, with minor changes, the proof (for the case of Fréchet differentiability) of Lemma 2.6 in [Ph]. [For instance, in the proof Lemma 2.6 one replaces V by S^0 and takes $\epsilon = 1/2$.]

DEFINITION. If the nonzero vector x is bounded above on the nonempty subset $A \subset E^*$ and $\alpha > 0$, a *slice* of A defined by x is the set

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

Recall that the subdifferential $\partial\|\cdot\cdot\|$ of the norm at nonzero points is always a subset of the dual unit sphere. The basic lemma which follows shows, in a quantitative way, that an ϵ-neighborhood in E^* of the image under $\partial\|\cdot\cdot\|$ of a small set will always contain a small slice of the dual ball B^*.

2.2. LEMMA. *Let x_0 be a unit vector in E; then for $\epsilon > 0$ and $r > 0$, the set $r\partial\|\cdot\cdot\|[B(x_0, \epsilon)] + B^*(0, \epsilon)$ contains the slice $S(rB^*, x_0, \epsilon^2)$.*

PROOF. Let $x^* \in S(rB^*, x_0, \epsilon^2)$, so that $r - \epsilon^2 < \langle x^*, x_0\rangle \le r$ and $\|x^*\| \le r$. For each $y \in E$ we have

$$\begin{aligned}\langle x^*, x_0 - y\rangle = \langle x^*, x_0\rangle - \langle x^*, y\rangle &> r - \epsilon^2 - \|x^*\|\cdot\|y\| \\ &\ge r - \epsilon^2 - r\|y\| = r(\|x_0\| - \|y\|) - \epsilon^2.\end{aligned}$$

Thus, $\langle x^*, y - x_0\rangle < r(\|y\| - \|x_0\|) + \epsilon^2$ for all $y \in E$; by definition, this means that x^* is in the ϵ^2-subdifferential $\partial_{\epsilon^2} r\cdot\|\cdot\cdot\|(x_0)$ of the function $r\cdot\|\cdot\cdot\|$ at x_0. (See, for instance, [Ph, p. 48].) By the Brøndsted–Rockafellar theorem ([Ph, p. 51]), there exists $x \in E$ and $y^* \in E^*$ such that $\|x - x_0\| \le \epsilon$, $\|x^* - y^*\| \le \epsilon$ and $y^* \in \partial r\cdot\|\cdot\cdot\|(x)$. Thus

$$x^* \in y^* + B^*(0, \epsilon) \in \partial r\cdot\|\cdot\cdot\|[B(x_0, \epsilon)] + B^*(0, \epsilon).$$

2.3. COROLLARY. *Suppose that $\|\cdot\cdot\|$ is a β-smooth norm on E. Then for each $x_0 \in E$ with $\|x_0\| = 1$ and for each $R > 0$ and set S in β, there exists $\delta > 0$ such that*

$$S(rB^*, x_0, \delta) \subset r\cdot\partial\|\cdot\cdot\|(x_0) + S^0 \qquad \textit{whenever } 0 < r < R.$$

PROOF. Since, by Proposition 2.1, the subdifferential $\partial\|\cdot\cdot\|$ is β-continuous at x_0, there exists $\epsilon_1 > 0$ such that

$$\partial\|\cdot\cdot\|[B(x_0, \epsilon_1)] \subset \partial\|\cdot\cdot\|(x_0) + (R + 1)^{-1}S^0.$$

Choose $0 < \epsilon < \epsilon_1$ such that $B^*(0, \epsilon) \subset (R + 1)^{-1}S^0$ and let $\delta = \epsilon^2$. By Lemma 2.2, whenever $0 < r < R$ we have

$$\begin{aligned} S(rB^*, x_0, \epsilon^2) &\subset r\cdot\partial\|\cdot\cdot\|[B(x_0, \epsilon)] + B^*(0,\epsilon) \\ &\subset r\cdot\partial\|\cdot\cdot\|(x_0) + r\cdot(R+1)^{-1}S^0 + (R+1)^{-1}S^0 \\ &\subset r\cdot\partial\|\cdot\cdot\|(x_0) + R\cdot(R+1)^{-1}S^0 + (R+1)^{-1}S^0 \\ &\subset r\cdot\partial\|\cdot\cdot\|(x_0) + S^0. \end{aligned}$$

The proof of Theorem 2.5 (below), which is the main result of this section, is essentially the same as that of Theorem 1.6. We first need to recall three elementary properties of minimal usco maps. For the convenience of the reader, the proofs are provided.

2.4. LEMMA. *Let Ω, X and Y be Hausdorff spaces and let $\varphi:\Omega\to 2^X$ be a minimal usco map.*

(a) *If $f: X\to Y$ is a continuous map, then $f\circ\varphi:\Omega\to 2^Y$ is a minimal usco map. Here, $f\circ\varphi(t) = f(\varphi(t))$.*

(b) *Let U and V be open subsets of Ω and X, respectively. If $\varphi(U)\cap V$ is nonempty, then $\{t\in U:\varphi(t)\subset V\}$ is a nonempty open subset of U. Here, $\varphi(U) = \bigcup\{\varphi(t): t\in U\}$.*

(c) *If U is an open subset of Ω, then the restriction $\varphi|_U: U\to 2^X$ is a minimal usco map.*

PROOF. (a) Let G_φ be the graph of φ. Then the graph of $f\circ\varphi$ coincides with $(\mathrm{Id}\times f)(G_\varphi)\subset\Omega\times Y$. If G is the graph of a usco map from Ω to 2^Y properly contained in the graph of $f\circ\varphi$, then $(\mathrm{Id}\times f)^{-1}(G)$ is the graph of a usco map from Ω to 2^X properly contained in G_φ, contrary to the minimality of G_φ.

(b) Suppose that $\varphi(t)\cap(X\setminus V)$ is nonempty for each $t\in U$ and let π denote the natural projection from $\Omega\times X$ onto Ω. Then

$$F = [G_\varphi\cap\pi^{-1}(\Omega\setminus U)]\cup[G_\varphi\cap\{U\times(X\setminus V)\}]^-$$

is a proper closed subset of G_φ such that $\pi(F)=\Omega$, contrary to the minimality of G_φ. Hence there is an element $t\in U$ such that $\varphi(t)\subset V$. The rest is clear from the upper-semicontinuity of φ.

(c) If G were the graph of a usco map $\psi: U\to 2^X$ properly contained in $\varphi|_U$, then $[G_\varphi\cap\{(\Omega\setminus U)\times X\}]\cup$ [The closure of G in $\Omega\times X$] would be the graph of a usco map properly contained in G_φ.

2.5. THEOREM. *Suppose that the norm in E is β-smooth and that Ω is a Baire space. Then for each minimal w^*-usco $\varphi:\Omega\to 2^{E^*}$ there is a residual subset G of Ω such that φ is β-continuous at each point of G.*

PROOF. Let $\|\cdot\cdot\|$ be an equivalent β-smooth norm on E. For each positive integer n let

$$\Omega_n = \{t \in \Omega : \varphi(t) \cap nB^* \neq \varnothing\} \quad \text{and let} \quad D_n = \operatorname{int}\Omega_n.$$

Then each Ω_n is closed, $\Omega = \cup\Omega_n$, and $D = \cup D_n$ is a dense open subset of Ω because the latter is a Baire space. By Lemma 2.4(c), $\varphi | D_n$ is a minimal w*-usco map and $t \to \varphi(t) \cap nB^*$ defines a w*-usco map in D_n contained in $\varphi | D_n$. From the minimality of the latter, it follows that $\varphi(t) \subset nB^*$ for each $t \in D_n$. Let φ_n: $D_n \to 2^{B^*}$ be the minimal w*-usco map given by $\varphi_n(t) = (1/n)\varphi(t)$. Now, we must show that

$$G = \{t \in \Omega : \varphi \text{ is } \beta\text{-continuous at } t\}$$

is residual in Ω. Note that if $t \in D_n$, then φ is β-continuous at t if and only if φ_n is β-continuous at t. Suppose we can prove that

$$G \cap D_n \equiv \{t \in D_n : \varphi_n \text{ is } \beta\text{-continuous at } t\}$$

is residual in D_n for each n; then $D_n \backslash G$ is of first category in Ω. Since

$$\Omega \backslash G \subset \cup (D_n \backslash G) \cup (\Omega \backslash D),$$

it follows that G is residual in Ω. Since an open subset of a Baire space is a Baire space in the relative topology, we have thus reduced the proof to the case where φ is a minimal w*-usco map from a Baire space Ω into 2^{B^*}.

As before, the proof consists of exhibiting a winning strategy for the second player of the Banach–Mazur game played in Ω to end up in the set G. Specifically, the two players A and B alternately choose nonempty open sets $\{U_k : k = 1,2,\dots\}$ and $\{V_k : k = 1,2,\dots\}$ contained in Ω so that

$$U_1 \supset V_1 \supset U_2 \supset V_2 \supset \cdots.$$

We shall produce a strategy for player B so that no matter how player A responds we always end up with the inclusion

$$\cap V_n \subset G.$$

Choose positive constants ϵ_k, β_k as before and suppose that player A has chosen u_1. Let $p_1 = \|\cdot\cdot\|$ and define

$$s_1 = \sup\{p_1^*(x^*) : x^* \in \varphi(U_1)\},$$

where p_1^* is the dual norm to p_1. If $s_1 = 0$, then as before the strategy for player B is to choose $V_k = U_k$ for all k, so that $\varphi(t)$ is the singleton $\{0\}$ for all $t \in \bigcap V_k = U_1$. Clearly, φ is β-continuous at t since $\varphi(U_1) = \varphi(t) = \{0\}$.

We now assume that $s_1 > 0$, so there exists $e_1 \in E$ with $p_1(e_1) = 1$ and

$$\langle x^*, e_1 \rangle > (1 - \epsilon_1)s_1 \qquad \text{for some } x^* \in \varphi(U_1).$$

[Strictly speaking, we must first fix a choice function defined on nonempty subsets of E. The point e_1 is the result of applying this choice function to the set $\{e \in E: \|e\| = 1 \text{ and } \langle x^*, e \rangle > (1 - \epsilon_1)s_1 \text{ for some } x^* \in \varphi(U_1)\}$.]

Now, the set-valued function $t \to \langle \varphi(t), e_1 \rangle$ is a minimal usco map from Ω to $2^{\mathbf{R}}$ (see Lemma 2.4(a)). Hence, by Lemma 2.4(b),

$$V_1 = \{t \in U_1: \langle \varphi(t), e_1 \rangle > (1 - \epsilon_1)s_1\}$$

is a nonempty open subset of U_1. Here, the notation $\langle \varphi(t), e_1 \rangle > a$ means that $\langle x^*, e_1 \rangle > a$ for all $x^* \in \varphi(t)$. Next define a norm p_2 equivalent to p_1 by

$$p_2^2 = p_1^2 + \beta_1 q_1^2,$$

where

$$q_1(x) = \inf\{\|x - \lambda e_1\| : \lambda \in \mathbf{R}\}, \qquad x \in E.$$

Note that if $x^* \in \varphi(V_1)$, then $x^* \neq 0$, because $0 < (1 - \epsilon_1)s_1 < \langle x^*, e_1 \rangle$.

The construction given above of V_1 from U_1 and p_1 is repeated in the subsequent plays. Thus, suppose that player A chooses U_k ($U_k \subset V_{k-1}$) and that p_k is given. Then we let

$$s_k = \sup\{p_k^*(x^*) : x^* \in \varphi(U_k)\} > 0.$$

Choose e_k so that $p_k(e_k) = 1$ and

$$\langle x^*, e_k \rangle > (1 - \epsilon_k)s_k \qquad \text{for some } x^* \in \varphi(U_k).$$

Let $V_k = \{t \in U_k : \langle \varphi(t), e_k \rangle > (1 - \epsilon_k)s_k\}$ and $p_{k+1}^2 = p_k^2 + \beta_k^2 q_k^2$, where

$$q_k(x) = \inf\{\|x - \lambda e_k\| : \lambda \in \mathbf{R}\}, \qquad x \in E.$$

We want to show that if $t \in \bigcap V_k$, then $\varphi(t)$ is a singleton. To that end, note that it follows exactly as in the proof of Theorem 1.6 (assertion (i)) that the increasing sequence of norms $\{p_k\}$ converges uniformly on bounded sets to an equivalent smooth norm p_∞ on E. (Eventually, p_∞ will be shown to be β-smooth.) We will show that

(i) the sequence $\{s_k\}$ converges to $s_\infty \geq 0$,

(ii) the sequence $\{e_k\}$ converges to an element e_∞ satisfying $p_\infty(e_\infty) = 1$, and

(iii) if $t \in \bigcap V_k$, then for each $x^* \in \varphi(t)$, $p_\infty^*(x^*) = s_\infty = \langle x^*, e_\infty \rangle$.

Since p_∞ is smooth, (ii) and (iii) show that $\varphi(t)$ is a singleton.

(i) Since $p_k \leq p_{k+1}$, we have $p_{k+1}^* \leq p_k^*$ for all k. Hence

$$0 < s_{k+1} = \sup\{p_{k+1}^*(x^*) : x^* \in \varphi(U_{k+1})\} \leq \sup\{p_k^*(x^*) : x^* \in \varphi(U_k)\} = s_k.$$

It follows that $0 \leq s_\infty = \lim s_k$ exists.

(ii) Since $p_{k+1}(e_k) = 1$, for all $x^* \in \varphi(V_k)$ we have

$$(*) \qquad (1 - \epsilon_k)s_k < \langle x^*, e_k \rangle \leq p_{k+1}^*(x^*) \cdot p_{k+1}(e_k) = p_{k+1}^*(x^*).$$

In particular (noting $U_{k+1} \subset V_k$),

$$(**) \qquad (1 - \epsilon_k)s_k < s_{k+1} \leq s_k.$$

As in the proof of Theorem 1.6, we show that $\{e_k\}$ is convergent by showing that for all sufficiently large k,

$$(***) \qquad \|e_{k+1} - e_k\| \leq 6\sqrt{\epsilon_k}/\beta_k.$$

Pick any $x_k^* \in \varphi(V_{k+1}) \subset \varphi(V_k)$; then

$$(1) \qquad \langle x_k^*, e_k \rangle > (1 - \epsilon_k)s_k$$

and

$$(2) \qquad \langle x_k^*, e_{k+1} \rangle > (1 - \epsilon_{k+1})s_{k+1}.$$

By (2), $(1 - \epsilon_{k+1})s_{k+1} < \langle x_k^*, e_{k+1} \rangle \leq p_k^*(x_k^*) \cdot p_k(e_{k+1}) \leq s_k p_k(e_{k+1})$. On the other hand, from (**) (using $\epsilon_{k+1} < \epsilon_k$),

$$(1 - \epsilon_{k+1})s_{k+1} > (1 - \epsilon_k)(1 - \epsilon_{k+1})s_k > (1 - 2\epsilon_k)s_k.$$

Consequently $p_k(e_{k+1}) > (1 - 2\epsilon_k) > 0$. The proof of (***) now follows exactly as in the proof of Theorem 1.6(ii). (Note that (2) makes the proof that $\lambda_k > 0$ a bit simpler here.)

(iii) Suppose that $t \in \bigcap V_k$ and $x^* \in \varphi(t)$. Then by (*), for all k,

$$(1 - \epsilon_k)s_k < \langle x^*, e_k \rangle \leq p_{k+1}^*(x^*) \leq s_{k+1}.$$

Let $k \to \infty$; then since the sequence of dual norms $\{p_k^*\}$ converges pointwise to p_∞^*, we have $s_\infty \leq \langle x^*, e_\infty \rangle \leq p_\infty^*(x^*) \leq s_\infty$, which completes the proof. Note that $\bigcap V_k \neq \varnothing$ implies that $s_\infty > 0$, since if $x^* \in \varphi(\bigcap V_k) \subset \varphi(V_1)$, then $x^* \neq 0$.

Finally, we must show that φ is β-continuous at each point $t \in \bigcap V_k$. Let $0 < \epsilon < 1/2$; then for all sufficiently large k we have

$$(****) \qquad p_\infty(e_k - e_\infty) < \epsilon, \quad 0 < \epsilon_k < \epsilon \quad \text{and} \quad s_k < (1+\epsilon)\cdot s_\infty.$$

(Recall that the sequence $\{s_k\}$ decreases to s_∞ and, as noted above, we have $s_\infty > 0$.) If $x^* \in \varphi(V_k)$, then

$$p_\infty^*(x^*) \le p_k^*(x^*) \le s_k < (1+\epsilon)\cdot s_\infty,$$

and hence (using the definition of V_k) x^* satisfies

$$\begin{aligned}\langle x^*, e_\infty\rangle = \langle x^*, e_k\rangle + \langle x^*, e_\infty - e_k\rangle &> (1-\epsilon_k)\cdot s_k - p_\infty(e_\infty - e_k)p_\infty^*(x^*) \\ &\ge (1-\epsilon_k)\cdot s_k - \epsilon\cdot s_k \ge (1-2\epsilon)\cdot s_k \ge (1-2\epsilon)\cdot s_\infty.\end{aligned}$$

Let $B_\infty^* = \{x^* : p_\infty^*(x^*) \le 1\}$ and note that $\sup\langle (1+\epsilon)s_\infty\cdot B_\infty^*, e_\infty\rangle = (1+\epsilon)s_\infty$. Since $(1+\epsilon)\cdot s_\infty - 3\epsilon\cdot s_\infty = (1-2\epsilon)\cdot s_\infty$, it follows that

$$\varphi(V_k) \subset S((1+\epsilon)s_\infty\cdot B_\infty^*, e_\infty, 3\epsilon\cdot s_\infty).$$

We need to show that p_∞ is β-smooth; this is straightforward once it is established that each q_k^2 is β-smooth, since the series of Gateaux derivatives

$$2[dp_1(x) + \Sigma\beta_k^2 q_k(x)\cdot dq_k(x)]$$

converges uniformly on bounded subsets to $dp_\infty(x)$, for each $x \in E$. Now, if $q_k(y) = 0$, then an examination of the difference quotient for q_k^2 shows that the latter is always Fréchet differentiable at y, with $dq_k^2(y) = 0$. If $q_k(y) > 0$, then (as before) we can write $q_k(y) = p_1(y - \lambda_0 e_k) > 0$ for some $\lambda_0 \in \mathbf{R}$ and hence, for $t > 0$ and any $u \in E$, the inequalities in (**) in the proof of Theorem 1.6 hold. Since p_1 is assumed to be β-smooth, the second term in (**) converges to 0 uniformly for $u \in S$, for all $S \in \beta$, so the same is true for q_k. [See, for instance, [Ph, p. 14] for the (typical) case of Fréchet differentiability.]

Let $S \in \beta$; by Corollary 2.3, applied to the norm p_∞, there exists $\delta > 0$ such that

$$S(rB_\infty^*, e_\infty, \delta) \subset r\cdot\partial p_\infty(e_\infty) + (1/2)S^0, \qquad \text{provided } 0 < r < 2s_\infty.$$

(Note that δ depends on S and s_∞, but not on ϵ.) Choose $0 < \epsilon < 1/2$ such that $3\epsilon s_\infty < \delta$ and $s_\infty B_\infty^*(0, \epsilon) \subset (1/2)S^0$. Then if k is large enough to satisfy (****), we have

$$\varphi(V_k) \subset (1+\epsilon)\cdot s_\infty\cdot\partial p_\infty(e_\infty) + (1/2)S^0.$$

Now if $t \in \bigcap V_k$, then from the first part of the proof, $\varphi(t)$ is the singleton $s_\infty \partial p_\infty(e_\infty)$, so that $p_\infty^*[\varphi(t)] = s_\infty$. Consequently,

$$\varphi(V_k) \subset (1+\epsilon)\cdot\varphi(t) + (1/2)S^0 \subset \varphi(t) + s_\infty B_\infty^*(0,\epsilon) + (1/2)S^0$$
$$\subset \varphi(t) + (1/2)S^0 + (1/2)S^0 \subset \varphi(t) + S^0.$$

This shows that φ is β-continuous at t.

The next corollary follows from Theorem 2.5 by letting β be the family of all finite subsets of E and by recalling the definition of class (S).

2.6. Corollary. *If the Banach space E admits an equivalent smooth norm, then it is of class (S); that is, E^* in the weak* topology is of type $\mathcal{S}$.*

Definition. A set-valued map $\varphi:\Omega \to 2^{E^*}$ is said to be a *convex* w*-usco map if it is a w*-usco map and its values are convex. It is said to be a *minimal* convex w*-usco map if it is minimal (in the usual ordering) in the family of all convex w*-usco maps.

It is known (see, for instance, [Ph]) that a maximal monotone operator T, when restricted to the interior of its effective domain $D(T)$, is a minimal convex w*-usco map (but not necessarily a minimal usco map).

2.7. Lemma. *Let Ω be a Hausdorff space and let $\varphi:\Omega \to 2^{E^*}$ be a w*-usco map. If φ is β-continuous at $t_0 \in \Omega$, then so is the w*-usco map $\overline{\mathrm{co}}\,\varphi$ [where, for $t \in \Omega$, $\overline{\mathrm{co}}\,\varphi(t)$ is the weak* closed convex hull of $\varphi(t)$].*

Proof. Let $\{x^*\} = \varphi(t_0)$ and suppose that $S \in \beta$. By continuity, there exists a neighborhood U of t_0 such that $\varphi(U) \subset x^* + S^0$. Obviously, for any $t \in U$, we have $\varphi(t) \subset x^* + S^0$. Since the right-hand side is weak* closed and convex, $\overline{\mathrm{co}}\,\varphi(t) \subset x^* + S^0$. Thus, $\overline{\mathrm{co}}\,\varphi(U) \subset x^* + S^0$ and $\{x^*\} = \overline{\mathrm{co}}\,\varphi(t_0)$. That $\overline{\mathrm{co}}\,\varphi$ is a w*-usco map is shown [Ph, p. 101].

2.8. Theorem. *The conclusion to Theorem 2.5 holds for each minimal convex w*-usco map $\varphi:\Omega \to 2^{E^*}$.*

Proof. Let φ_0 be a minimal w*-usco map contained in φ. Then by the minimality of φ, we have $\overline{\mathrm{co}}\,\varphi_0 = \varphi$ and the conclusion follows from Theorem 2.5 and Lemma 2.7.

2.9. Corollary. *If E admits an equivalent β-smooth norm, then for any maximal monotone operator T on E there exists a dense G_δ subset G of $D = \operatorname{int} D(T)$ such that T is β-continuous at each point of G.*

When applied to the case of Fréchet differentiability (that is, when the family β consists of all bounded sets) Corollary 2.9 is well known. It was first proved for (the subdifferentials of) convex functions by Ekeland and Lebourg [Ek–Leb] and extended to monotone operators (on Asplund spaces) by Kenderov [Ke_2].

3. Remarks

The utility of "smoothability" (the existence of an equivalent Gateaux differentiable norm) in Theorem 1.6 raises anew the question of permanence properties for such spaces. It is easy to see that smoothability is inherited by subspaces and preserved under finite products. It does not satisfy the "three-space property"; this was shown by Talagrand [Ta] when he showed that the space $E_1 = C(K)$, with K the compact Hausdorff "two arrow space", does not admit an equivalent smooth norm, even though it contains the separable (hence weak Asplund) subspace $C[0, 1]$, while the quotent space $E_1/C[0, 1]$ is isometric to $c_0(0, 1]$ (which is an Asplund space). Theorem 1.6 provides another proof of this result: If E_1 admitted an equivalent smooth norm, then it would be a weak Asplund space, but Coban and Kenderov [Co–Ke] have shown that the set of points of Gateaux differentiability of the supremum norm in E_1 does not contain a dense G_δ subset. It is an open question whether a quotient of a smoothable space is necessarily smoothable. As to permanence properties of weak Asplund spaces, the example E_1 described above shows that they do not satisfy the three-space property. That the continuous linear image of a weak Asplund space is again a weak Asplund space was asserted by Asplund [As], but his proof appears to assume that a continuous linear image of a G_δ set is a G_δ set. This need not be true, but the theorem is valid. One may use, for example, the topological Theorem 2.2 of [Co–Ke] and the open mapping theorem to assert that the continuous linear image of a dense G_δ set *contains* a dense G_δ set, and hence rescue Asplund's proof. It is also possible to obtain this topological result as an easy consequence of the Banach–Mazur game, as follows.

3.1. LEMMA. *Suppose that M is a complete metric space, X a Hausdorff space and $f: M \to X$ a continuous open mapping of M onto X. If $G = \bigcap G_n$ is the intersection of a decreasing sequence of dense open subsets G_n of M, then $f(G)$ is residual in X.*

PROOF. Assume that players A and B play the Banach–Mazur game in X, with $S = f(G)$. We must show that B has a winning strategy. Given the first nonempty open subset U_1 of X chosen by A, player B uses the fact that since G_1 is open and dense there exists an open metric ball $B_1 = B(x_1, r_1)$ (with $0 < r_1 < 1$) such that $\bar{B}_1 \subset f^{-1}(U_1) \cap G_1$ and therefore chooses the nonempty open set $V_1 = f(B_1) \subset U_1$. Player A having chosen a nonempty open subset $U_2 \subset V_1$, player B observes that since $f^{-1}(U_2) \cap B_1$ is nonempty and open, it must intersect the dense open set G_2, hence there exists an open metric ball $B_2 = B(x_2, r_2)$ (with $0 < r_2 < 1/2$) such that $\bar{B}_2 \subset f^{-1}(U_2) \cap B_1 \cap G_2$, so B chooses $V_2 = f(B_2) \subset U_2$. Player B's strategy is to continue in this manner, resulting in a sequence $U_1 \supset V_1 \supset U_2 \supset V_2 \supset \cdots \supset U_n \supset V_n \supset \cdots$ of nonempty open sets such that $V_n = f(B_n)$, where, for each $n > 1$,

$$B_n = B(x_n, r_n) \subset B_{n-1}, \quad 0 < r_n < 1/n \quad \text{and} \quad \bar{B}_n \subset f^{-1}(U_n) \cap G_n.$$

By using this strategy, B wins every play; indeed, the intersection of the closed balls $\bar{B}_n$ consists of a single point $\{x_0\}$ which is necessarily in G. If $y \in \cap V_n$, then (since f is onto) for all n there exists $z_n \in B_n$ such that $y = f(z_n)$. Clearly, $x_n \to x_0$ and hence $z_n \to x_0$, so $y = f(x_0) \in f(G)$. Thus, $\cap V_n \subset f(G)$ and hence, by Lemma 1.3, the latter is residual in X.

The remaining permanence properties for weak Asplund spaces remain open; for instance, it is not clear whether the product of a weak Asplund space with the real line is again a weak Asplund space. It is easily seen (using Asplund's quotient result above) that a complemented subspace of a weak Asplund space has the same property (so the parenthetical remark in [Ph, p. 34] about hyperplanes is nonsense), but the question for arbitrary subspaces remains open. Slightly more is known about permanence properties of Gateaux Differentiability Spaces; see [Ph, Ch. 6].

Various permanence properties of Banach spaces of class (S) are a consequence of good permanence properties of topological spaces of type $\mathcal{S}$. It is known [St_1] that the class (S) is closed under taking closed subspaces and countable l_p products ($1 \leq p < \infty$). Also, if E is in the class (S) and $T: E \to F$ is a bounded linear operator having dense range, then the Banach space F is in the class (S).

A natural question about monotone operators remains open: If E is a weak Asplund space, is every maximal monotone operator on E generically single-valued? (Analogously, if E is a GDS, are such operators densely single-valued?) Note that it is still open whether every GDS is a weak Asplund space.

Acknowledgements

The first-named author wishes to thank Professor B. S. Thomson for his hospitality and the Mathematics Department of the University of British Columbia (and especially Professor N. Ghoussoub) for support during the initial work on this paper. The second-named author would similarly like to express his appreciation to the Mathematics Department of the Charles University, Prague, for its hospitality while working on this paper.

REFERENCES

[As] E. Asplund, *Fréchet differentiability of convex functions*, Acta Math. **121** (1968), 31–47.

[B-F-K] J. Borwein, S. Fitzpatrick and P. Kenderov, *Minimal convex uscos and monotone operators on small sets*, preprint.

[Bo-Pr] J. M. Borwein and D. Preiss, *A smooth variational principle with applications to subdifferentiability and to differentiability of convex functions*, Trans. Am. Math. Soc. **303** (1987), 517–527.

[Chr] J. P. R. Christensen, *Theorems of Namioka and Johnson type for usco set-valued mapping*, Proc. Am. Math. Soc. **86** (1982), 649–655.

[Co-Ke] 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.

[Da] M. M. Day, *Normed Linear Spaces* (3rd ed.), Springer-Verlag, Berlin–Heidelberg–New York, 1973.

[Ek-Leb] I. Ekeland and G. Lebourg, *Generic Fréchet differentiability and perturbed optimization problems in Banach spaces*, Trans. Am. Math. Soc. **224** (1976), 193–216.

[Ha] R. Haydon, *A counterexample to several questions about scattered compact spaces*, Bull. London Math. Soc. **27** (1990), 261–268.

[Ke_1] P. S. Kenderov, *The set-valued monotone mappings are almost everywhere single-valued*, C. R. Acad. Bulgare Sci. **27** (1974), 1173–1175.

[Ke_2] P. S. Kenderov, *Monotone operators in Asplund spaces*, C. R. Acad. Bulgare Sci. **30** (1977), 963–964.

[La-Ph] D. G. Larman and R. R. Phelps, *Gateaux differentiability of convex functions on Banach spaces*, J. London Math. Soc. **20** (1979), 115–127.

[Ma] S. Mazur, *Über konvexe Mengen in linearen normierten Räumen*, Studia Math. **4** (1933), 70–84.

[Noll] D. Noll, *Generic Gateaux-differentiability of convex functions on small sets*, J. Math. Anal. Appl. **147** (1990), 531–544.

[Ox] J. O. Oxtoby, *The Banach-Mazur game and Banach Category Theorem*, in *Contributions to the Theory of Games*, Vol. III, Annals of Math. Studies **39**, Princeton, N.J., 1957, pp. 159–163.

[Ph] R. R. Phelps, *Convex functions, monotone operators and differentiability*, Lecture Notes in Math. No. 1364, Springer-Verlag, Berlin–Heidelberg–New York, 1989.

[Pr] D. Preiss, *Fréchet derivatives of Lipschitz functions*, J. Funct. Anal. **91** (1990), 312–345.

[Rain] J. Rainwater, *Yet more on the differentiability of convex functions*, Proc. Am. Math. Soc. **103** (1988), 773–778.

[St_1] C. Stegall, *A class of topological spaces and differentiation of functions on Banach spaces*, Proc. Conf. on Vector Measures and Integral Representations of Operators, Vorlesungen aus dem Fachbereich Math., Heft 10 (W. Ruess, ed.), Univ. Essen, 1983.

[St_2] C. Stegall, *More Gateaux differentiability spaces*, Proc. Conf. Banach Spaces, Univ. Missouri, 1984 (N. Kalton and E. Saab, eds.), Lecture Notes in Math., Vol. 1166, Springer-Verlag, Berlin, 1985, pp. 158–168.

[Ta] M. Talagrand, *Renormages de quelques C(K)*, Isr. J. Math. **54** (1986), 327–334.

[V-V] A. Verona and M. E. Verona, *Locally efficient monotone operators*, Proc. Am. Math. Soc., to appear.

[Za] E. Zarantonello, *Dense single-valuedness of monotone operators*, Isr. J. Math. **15** (1973), 158–166.