GENERIC DIFFERENTIABILITY OF CONVEX FUNCTIONS AND MONOTONE OPERATORS

Thesis by

Maria Elena Verona

In Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy

California Institute of Technology

Pasadena, California

1989

(Submitted May 15, 1989)

ACKNOWLEDGEMENTS

I would like to express my thanks to Professor W. A. J. Luxemburg, my advisor, for his continued support, guidance and confidence in my work.

I also wish to thank Professor R. R. Phelps from Washington University (Seattle, Washington) for the interest shown in my work and the helpful comments he made on some of my previous papers.

The Department of Mathematics of Caltech had a positive influence on my mathematical development and teaching training. I acknowledge it with gratitude.

Finally, I would like to thank Andrei for his support and encouragement.

ABSTRACT

The aim of this paper is to investigate to what extent the known theory of subdifferentiability and generic differentiability of convex functions defined on open sets can be carried out in the context of convex functions defined on not necessarily open sets. Among the main results obtained I would like to mention a Kenderov type theorem (the subdifferential at a generic point is contained in a sphere), a generic Gâteaux differentiability result in Banach spaces of class **S** and a generic Fréchet differentiability result in Asplund spaces. At least two methods can be used to prove these results: first, a direct one, and second, a more general one, based on the theory of monotone operators. Since this last theory was previously developed essentially for monotone operators defined on open sets, it was necessary to extend it to the context of monotone operators defined on a larger class of sets, our "quasi open" sets. This is done in Chapter III. As a matter of fact, most of these results have an even more general nature and have roots in the theory of minimal usco maps, as shown in Chapter II.

Table of Contents

INTRODUCTION

Consider a real Banach space X and a closed, convex subset C of X with non-empty interior. Recall that a boundary point x of C is called *smooth* if there exists a unique hyperplane H of X passing through x and such that C is contained in one of the halfspaces determined by H. In 1933 S. Mazur [**10**] proved that, under the assumption that X is separable, the set of all smooth points of C is a dense G_δ subset of the boundary of C. Assuming that zero is an interior point of C, we can consider the Minkowski functional associated to C:

$$p_C(x) = \inf\{t > 0 : x \in tC\}, \quad x \in X.$$

This functional is always sublinear and continuous. There exists a close relationship between the "smoothness" of C and the Gâteaux differentiability of p_C, namely: a point x in the boundary of C is smooth if and only if p_C is Gâteaux differentiable at any point of the ray $\{tx : t > 0\}$. It follows that the result mentioned above can be restated by saying that in a separable Banach space a Minkowski functional is Gâteaux differentiable on a dense G_δ subset.

It is now natural to ask to what extent convex functions are Gâteaux differentiable on "large" subsets of their domains. As a matter of fact, by using Mazur's result, one can prove that any continuous, convex function defined on an open, convex subset of a separable Banach space is Gâteaux differentiable on a dense G_δ subset of its domain. However, this result is not true in an arbitrary Banach space. For example, in l^∞ the continuous convex function $p(x) = \limsup |x_n|$ is nowhere Gâteaux differentiable. Asplund [**2**] introduced the following classes of Banach spaces: the class of weak differentiability spaces (nowadays called weak Asplund spaces) and the class of strong differentiability spaces (nowadays called Asplund spaces). The first one consists of those Banach spaces in which any continuous, convex function defined on an open, convex set is Gâteaux differentiable on a dense G_δ subset of its domain. The second one is defined similarly, except that Gâteaux differentiability is replaced by Fréchet differentiability. In these terms Mazur's theorem can be restated by saying that separable Banach spaces are weak Asplund.

Asplund showed that a larger class of Banach spaces are weak Asplund, namely the Banach spaces that admit an equivalent norm whose dual norm is strictly convex. Besides separable spaces, this class contains the weakly compactly generated Banach spaces, in particular the reflexive ones.

More interesting and better known is the class of Asplund spaces. For example, a Banach space X is Asplund if and only if each closed and separable subspace of X has a separable dual. For more details on Asplund and weak Asplund spaces see Phelps' Lecture Notes [**14**].

In a finite dimensional space a convex set a has non-empty interior if and only if its affine hull is the whole space. This is no longer true in infinite dimensional spaces, as shown by the positive cone in $l^2(N)$ (see Remark 2.7 (1)). However, when the interior is empty, there exists sometimes another subset, the "quasi-interior," which plays a similar role (see Section 2). It was the striking similarity between interior and quasi-interior that triggered my interest to study differentiability properties of convex functions defined on quasi-open sets. As a first result in this direction, I proved that in a separable Banach space any convex and locally

Lipschitz function defined on a quasi-open convex set is Gâteaux differentiable on a dense G_δ subset of its domain (see [**23**]). Even in the case of an open domain this result is stronger than Mazur's initial theorem: the set of points where the function is not Gâteaux differentiable cannot contain a closed, convex set with non-empty quasi-interior. Soon after, my result was extended by Rainwater [**16**] who proved a generic Gâteaux (respectively Fréchet) differentiability result for Banach spaces of class **S** (respectively Asplund spaces). His proofs use a subdifferentiability result from [**23**] and techniques from the theory of multivalued maps.

In this thesis all these results are further generalized to the case of convex locally Lipschitz functions defined on arbitrary convex sets. In this more general context we have to deal with a new fact: the subdifferential map, the key to differentiability results, is no longer locally bounded, one of its main features in the case of open or quasi-open domains. Fortunately it still has a useful property: it is "locally efficient." It turns out that this property is sufficient to derive directly the main differentiability results of Chapter IV, as I did in a first version of this thesis. However, most of those results have a more general nature and can be studied in the context of monotone operators, or even more general, in the context of minimal usco maps. This is done in Chapter II and Chapter III and some of the results obtained there may be of independent interest.

Chapter I

INTRODUCTORY TOPICS

1. Tangent and Normal Cones

1.1. Let A be a subset of the Banach space X and $x \in A$. The simplest cone associated to A at x is the *cone generated by A from x*, denoted A_x and defined by

$$A_x = \{y \in X : \text{ there exists } t > 0 \text{ such that } x + ty \in A\}.$$

Let us notice that if $0 \in A$ and A is a cone, then $A_0 = A$; this shows that in general A_x is neither convex, nor closed. If the set A is convex, then A_x is a convex cone, not necessarily closed even if A is closed, as the following example shows it: let A be the closed unit ball in R^2 and $x = (0, -1)$; then A_x consists of the open upper half plane and the origin.

A point $x \in A$ is called an *absorbing point* of A if $A_x = X$. Obviously any interior point of A is an absorbing point of A. The converse is not true in general, as the following example shows it: let $A = \{x \in X : \|x\| = 1\} \cup \{0\}$; then 0 is an absorbing, but not interior, point of A. However, if A is closed and convex then any absorbing point of A is an interior point of A (this is true in any barreled space).

1.2. A smaller cone, $T_x(A)$, can be defined as follows:

$$T_x(A) = \{y \in X : \text{ there exists a sequence } (t_n) \downarrow 0 \text{ such that } x + t_n y \in A\}.$$

Of course, if A is convex the two cones coincide.

1.3. Another useful cone that we shall consider in this paper is the *contingent cone to A at x*, denoted $K_x(A)$. It is defined for any $x \in \mathrm{cl}\,(A)$ as follows: $y \in K_x(A)$ if and only if there exist a sequence of positive real numbers $(t_n) \downarrow 0$ and a sequence $(y_n) \subset X$ (norm) convergent to y such that $x + t_n y_n \in A$ for each n. Equivalently, $y \in K_x(A)$ if and only if for any neighborhood U of x and any neighborhood V of y there exist $t > 0$ and $z \in V$ such that $x + tz \in U \cap A$.

One can easily check that $K_x(A)$ is a closed cone. If A is a cone then $K_0(A)$ is the closure of A. It follows that in general $K_x(A)$ is not convex.

1.4. LEMMA. *$K_x(A) = K_x(\mathrm{cl}\,(A))$ for every $x \in \mathrm{cl}\,(A)$.*

PROOF. Trivially we have that $K_x(A) \subseteq K_x(\mathrm{cl}\,(A))$. To prove the other inclusion, let $y \in K_x(\mathrm{cl}\,(A))$. Then there exists a sequence of positive real numbers $(t_n) \downarrow 0$ and a sequence $(y_n) \subset X$ convergent to y such that $x + t_n y_n \in \mathrm{cl}\,A$. Let $\varepsilon_n = t_n/n$. Then there exists $z_n \in A$ with $\|z_n - x - t_n y_n\| < \varepsilon_n$. Let $u_n = \dfrac{1}{t_n}(z_n - x)$. Since

$$\|u_n - y\| \leq \|u_n - y_n\| + \|y_n - y\| = \frac{1}{t_n}\|z_n - x - t_n y_n\| + \|y_n - y\| < \frac{1}{n} + \|y_n - y\|$$

it follows that the sequence (u_n) converges to y. Since $z_n = x + t_n u_n \in A$, it follows that $y \in K_x(A)$, completing the proof of the lemma. □

Except the obvious inclusions $T_x(A) \subseteq A_x$ and $T_x(A) \subseteq K_x(A)$, $x \in A$, in general there are no other relations between the cones introduced above. However, in the case of convex sets, we have

1.5. LEMMA. *Let C be a convex subset of X and $x \in C$. Then $K_x(C) = \mathrm{cl}\,(C_x)$.*

PROOF. Clearly $C_x \subseteq K_x(C)$ and, since $K_x(C)$ is closed, we have $\mathrm{cl}\,(C_x) \subseteq K_x(C)$. Conversely, if $y \in K_x(C)$ there exist a sequence of positive real numbers $(t_n) \downarrow 0$ and a sequence $(y_n) \subset X$ converging to y such that $x + t_n y_n \in C$. Hence $y_n \in C_x$ for each n and therefore $y \in \mathrm{cl}\,(C_x)$. □

1.6. If C is a convex subset of X, the *normal cone to C at $x \in C$*, denoted C_x^*, is defined by

$$C_x^* = \{x^* \in X^* : \langle x^*, z\rangle \leq 0 \text{ for every } z \in C_x\}.$$

It is easy to see that C_x^* can also be described by

$$C_x^* = \{x^* \in X^* : \langle x^*, y - x\rangle \leq 0 \text{ for every } y \in C\}.$$

2. Quasi interior. Support Points

2.1. Let us notice that for an open subset A of X we have $K_x(A) = X$ for any $x \in A$. This motivates the following definitions.

DEFINITION. (1) A point $x \in A$ is called a *quasi interior point* of A if $K_x(A) = X$. The set $\mathrm{qi}(A)$ of all quasi-interior points of A is called the *quasi-interior* of A. The set A is called *quasi-open* if it is equal to its quasi-interior.

(2) Let $A \subseteq B$ be subsets of X. A is called *quasi open in B* if $K_x(A) = K_x(B)$ for all $x \in A$.

Obviously an open set is quasi-open. Since $K_x(A) = K_x(\mathrm{cl}\,(A))$, it follows that any dense subset of a quasi-open set is quasi-open. Also, a quasi-open subset of a quasi-open set is quasi-open. The quasi-interior of convex sets will play an important role in what follows.

2.2. DEFINITION. Let C be a convex subset of X. A point $x \in C$ is called a *support point of C* if there exists a non zero $x^* \in X^*$ such that

$$\langle x^*, x\rangle = \sup\{\langle x^*, y\rangle : y \in C\}.$$

The functional x^* is called a *support functional to C at x*. We shall denote by $\mathrm{S}(C)$ the set of all support points of C.

It follows immediately from the definitions that

$$x \in \mathrm{S}(C) \text{ if and only if } C_x^* \neq \{0\}$$

and, since $K_x(C) = \mathrm{cl}\,(C_x)$,

$$x \in \mathrm{S}(C) \text{ if and only if } 0 \in \mathrm{S}(C_x) \text{ if and only if } 0 \in \mathrm{S}(K_x(C)).$$

It is also clear that C_x^*, the normal cone to C at x, consists of 0 and all the support functionals to C at x.

2.3. LEMMA. *Let $C \subseteq X$ be convex and let $x \in C$. Then $x \in \mathrm{S}(C)$ if and only if $x \notin \mathrm{qi}(C)$.*

PROOF. Let $x \in \mathrm{S}(C)$. Then $0 \in \mathrm{S}(K_x(C))$ and thus $K_x(C) \neq X$, proving that $x \notin \mathrm{qi}(C)$. Conversely, if $x \notin \mathrm{qi}(C)$, then the closed convex cone $K_x(C)$ is different from X. Let $z \notin K_x(C)$; by the separation theorem there exists $x^* \in X^*$ such that $\langle x^*, u\rangle \leq \langle x^*, z\rangle$ for each $u \in K_x(C)$. It follows easily that x^* is a support functional to $K_x(C)$ at zero and this implies that $0 \in \mathrm{S}(K_x(C))$. As noticed above, this is equivalent to $x \in \mathrm{S}(C)$. □

2.4. LEMMA. *Let $C \subseteq X$ be a convex subset.*

(1) *For any $x \in \mathrm{qi}(C)$ and $y \in C$, the segment $[x, y)$ is contained in* $\mathrm{qi}(C)$.

(2) *For any $x \in \mathrm{int}(C)$ and $y \in C$, the segment $[x, y)$ is contained in* $\mathrm{int}(C)$.

PROOF. (1) Recall that $z \in [x, y)$ if and only if there exists $t \in [0, 1)$ such that $z = (1-t)x + ty$. Assume that such a z is a support point of C. Then there exists a non zero $z^* \in C_z^*$ and

$$(1-t)\langle z^*, x\rangle = \langle z^*, z\rangle - t\langle z^*, y\rangle \geq (1-t)\langle z^*, z\rangle \geq (1-t)\langle z^*, x\rangle$$

implying that $\langle z^*, x\rangle = \langle z^*, z\rangle = \sup\{\langle z^*, u\rangle : u \in C\}$, hence x is a support point of C. This contradiction proves the first assertion of the lemma.

(2) Let z be as above, but assume now that $x \in \text{int}(C)$. Then there exists $\varepsilon > 0$ such that $B(x, \varepsilon) \subseteq C$. For any $u \in B(z, (1-t)\varepsilon)$ let $v = \dfrac{1}{1-t}u - \dfrac{t}{1-t}y$. Since

$$\|v - x\| = \frac{1}{1-t}\|u - z\| < \varepsilon$$

it follows that $v \in B(x, \varepsilon) \subseteq C$ and therefore $u = (1-t)v + ty \in C$. Thus $B(z, (1-t)\varepsilon) \subseteq C$, proving that $z \in \text{int}(C)$. □

2.5. COROLLARY. *If $C \subseteq X$ is convex then* $\text{qi}(C)$ *(resp.* $\text{int}(C)$*), if non-empty, is convex and dense in C.*

It is obvious that the interior of a convex set is contained in its quasi-interior. As we shall see, it may happen that the interior is empty while the quasi-interior is non-empty. However, if the interior is non-empty then it is equal to the quasi-interior.

2.6. LEMMA. *Let $C \subseteq X$ be convex and assume that* $\text{int}(C)$ *is non-empty. Then* $\text{int}(C) = \text{qi}(C)$.

PROOF. It is enough to prove that if $x \notin \text{int}(C)$ then $x \in \mathrm{S}(C)$. Since $\text{int}(C)$ is open and convex, by the separation theorem there exists a non zero $x^* \in X^*$ such that $\langle x^*, x\rangle \geq \langle x^*, y\rangle$ for any $y \in \text{int}(C)$. Since $\text{int}(C)$ is dense in C, the preceeding inequality is true for any $y \in C$ and thus $x \in \mathrm{S}(C)$. □

2.7. REMARKS. (1) It is well known that if X is finite dimensional, then a convex subset C of X has a non-empty interior if and only if the affine hull of C is X. This result cannot be extended to infinite dimensional spaces, as shown by the following example.

Let $X = l^2 = l^2(N)$ and C be the positive cone in X, i.e., $C = \{(x_n) \in l^2 : x_n \geq 0 \text{ for every } n\}$. Clearly the affine hull of C is X. However, the interior of C is empty. To prove this let $\varepsilon > 0$ and $x \in C$. Let n_0 be such that $|x_n| < \varepsilon/2$ if $n \geq n_0$. Let $y_n = x_n$ if $n \neq n_0$ and $y_n = -\varepsilon/2$ if $n = n_0$. Clearly (y_n) is contained in $B(x, \varepsilon)$ but not in C. Thus C cannot contain any ball and therefore its interior is empty. On the other side, C has quasi-interior points, namely $\mathrm{qi}(C)$ consists of those $(x_n) \in l^2$ such that $x_n > 0$ for all n. To prove this let $\mathbf{x} = (x_n) \in C$ be such that $x_n > 0$ for any n and let $\mathbf{z} = (z_n) \in l^2$ be such that $z_n = 0$ for all but finitely many n. Then, if $t > 0$ is sufficiently small, $\mathbf{x} + t\mathbf{z} \in C$ and therefore $\mathbf{z} \in C_x$. Since the set of all $\mathbf{z}$ with the above property is dense in l^2, it follows that $K_{\mathbf{x}}(C) = \mathrm{cl}\, C_{\mathbf{x}} = l^2$, which means that $\mathbf{x} \in \mathrm{qi}(C)$. It remains to be seen that if $\mathbf{x} = (x_n) \in C$ has a zero coordinate, say $x_k = 0$, then the functional $x^* \in (l^2)^*$ defined by

$$\langle x^*, \mathbf{y} \rangle = -y_k, \qquad \text{for any } \ \mathbf{y} = (y_n) \in l^2$$

is a support functional to C at $\mathbf{x}$ and therefore $\mathbf{x} \in \mathrm{S}(C)$.

An obvious necessary condition for a convex set C in a Banach space X to have non-empty quasi-interior is that $\mathrm{cl}\,\mathrm{aff}\, C$ be equal to X. In view of the preceeding discussion we might suspect that this condition is also sufficient. In general this is not the case. As a matter of fact if in the previous example we replace $l^2(N)$ by $l^2(U)$, where U is an uncountable set and let C be again the positive cone in $l^2(U)$, then $\mathrm{cl}\,\mathrm{aff}\, C = l^2(U)$ but C has no quasi-interior points. However, as we shall see later, in the case of separable Banach spaces the above condition is also sufficient (see Theorem 2.8).

(2) If C is a closed convex set whose interior is empty, then $\mathrm{qi}(C)$, if non-empty, is neither relatively open nor relatively closed in C (since $\mathrm{S}(C)$ is dense in C by the Bishop-Phelps theorem (see for example [**14**]).

2.8. THEOREM. *Let X be a separable Banach space and C be a closed convex subset of X which is not contained in any hyperplane (i.e., $\mathrm{cl}\,\mathrm{aff}\, C = X$). Then the quasi interior of C is non-empty.*

PROOF. Without any loss of generality we can assume that $0 \in C$. Since X is separable, C is also separable and there exists a sequence $(x_n) \subset C$ which is dense in C. Define a new sequence by

$$y_n = \begin{cases} x_n, & \text{if } \|x_n\| < 1 \\ \dfrac{1}{\|x_n\|} x_n, & \text{if } \|x_n\| \geq 1. \end{cases}$$

Since $0 \in C$, it follows that $(y_n) \subset C$. The series $\sum 2^{-n} y_n$ is clearly absolutely convergent and its partial sums are all contained in C (since $0 \in C$). Using the fact that C is closed we obtain that $x = \sum 2^{-n} y_n$ is also contained in C. We claim that x is a quasi-interior point of C. Assume the contrary, i.e., assume that x is a support point of C. Then there exists a non-zero $x^* \in C_x^*$. Since $\langle x^*, y_n \rangle \leq \langle x^*, x \rangle$ for every n, we have

$$\begin{aligned} \langle x^*, x \rangle &= \lim_{k \to \infty} \sum_{n=1}^{k} \langle x^*, 2^{-n} y_n \rangle \leq \lim_{k \to \infty} \sum_{n=1}^{k} 2^{-n} \langle x^*, x \rangle \\ &= \langle x^*, x \rangle \sum_{n=1}^{\infty} 2^{-n} = \langle x^*, x \rangle \end{aligned}$$

and therefore $\langle x^*, y_n \rangle = \langle x^*, x \rangle$ for every n.

If $\|x_n\| < 1$ then $x_n = y_n$ and thus $\langle x^*, x_n \rangle = \langle x^*, x \rangle$. If $\|x_n\| \geq 1$ then $x_n = \|x_n\| y_n$ and thus $\langle x^*, x \rangle \geq \langle x^*, x_n \rangle = \|x_n\| \langle x^*, y_n \rangle \geq \langle x^*, y_n \rangle = \langle x^*, x \rangle$, showing that $\langle x^*, x_n \rangle = \langle x^*, x \rangle$ in this case too. It follows that (x_n) is contained in the hyperplane $H = \{z \in X : \langle x^*, z \rangle = \langle x^*, x \rangle\}$. Since (x_n) is dense in C, C is also contained in H, contradicting our assumption that C is not contained in any hyperplane. It follows that x must be a quasi-interior point of C and the theorem is proved. □

2.9. Remark. We have already seen (Remark 2.7 (1)) that the separability condition imposed on X in the previous theorem is essential. As the following example shows, the requirement that C be closed is also essential.

Let $X = l^2(N)$ and let C_0 consist of those (x_n) in X such that $x_n \geq 0$ for every n and only finitely many of them are non-zero. Then C_0 is a convex set whose closure is the set C consisting of those (x_n) such that $x_n \geq 0$ for every n. As a consequence, C_0 is not closed in X. Clearly $\operatorname{cl}\operatorname{aff} C_0 = X$, but C_0 has no quasi-interior points since all its points are support points (this can be proved exactly as in Remark 2.7 (1)).

We have already mentioned (Remark 2.7 (2)) that if C is a closed convex subset of the Banach space X with empty interior but with non-empty quasi-interior, then both $\mathrm{qi}(C)$ and $\mathrm{S}(C)$ are dense in C. The next theorem shows that $\mathrm{qi}(C)$ (if non-empty) is "much larger" than $\mathrm{S}(C)$.

2.10. THEOREM. *Let C be a closed convex subset of the Banach space X. Then $\mathrm{S}(C)$ is a (relative) F_σ subset of C, while $\mathrm{qi}(C)$ is a (relative) G_δ subset of C.*

PROOF. Recall first that a subset A of a topological space T is called an F_σ (resp. a G_δ) if it is the union (resp. intersection) of countably many closed (resp. open) subsets of T. Since $\mathrm{qi}(C)$ is the complement in C of $\mathrm{S}(C)$, it is sufficient to prove that the $\mathrm{S}(C)$ is an F_σ in C. If $\mathrm{S}(C) = C$, this is obvious. So we shall assume that $\mathrm{S}(C) \neq C$. Then there exists $x_0 \in \mathrm{qi}(C)$. For any positive integer n let

$$F_n = \{x \in C : \text{there exists } x^* \in C_x^*, \text{ such that } \|x^*\| \leq n \text{ and } \langle x^*, x - x_0\rangle = 1\}.$$

Next notice that $\mathrm{S}(C) = \bigcup F_n$. Indeed, let $x \in \mathrm{S}(C)$. Then there exists a non-zero $y^* \in C_x^*$. If $\langle y^*, x - x_0\rangle = 0$, then $\langle y^*, x\rangle = \langle y^*, x_0\rangle$ and this would imply that x_0 is a support point of C, contradicting our choice of x_0. Hence $\langle y^*, x - x_0\rangle = t > 0$ and we can consider $x^* = \dfrac{1}{t}y^*$. It is clear now that x is contained in any F_n such that $n \geq \|x^*\|$.

It remains to prove that each set F_m is closed in C. To this end let (x_n) be a sequence in F_m convergent to $x \in C$. For every n choose $x_n^* \in C_{x_n}^* \bigcap mB^*$ such that $\langle x_n^*, x_n - x_0\rangle = 1$. Since mB^* is w^* compact, there exists $x^* \in mB^*$, which is a w^* cluster point of (x_n^*). For any $\varepsilon > 0$ there exists k_ε such that $\|x - x_k\| < \varepsilon$ for all $k \geq k_\varepsilon$. Since x^* is a w^* cluster point of (x_n^*), there exists $k \geq k_\varepsilon$ such that

$|\langle x^* - x_k^*, x - x_0\rangle| < \varepsilon$. We have

$$\begin{aligned}|\langle x^*, x - x_0\rangle - 1| &= |\langle x^*, x - x_0\rangle - \langle x_k^*, x_k - x_0\rangle| \\ &= |\langle x^* - x_k^*, x - x_0\rangle + \langle x_k^*, x - x_k\rangle| \\ &\le |\langle x^* - x_k^*, x - x_0\rangle| + \|x_k^*\| \cdot \|x_k - x\| \le \varepsilon + m\varepsilon\end{aligned}$$

and therefore $\langle x^*, x - x_0\rangle = 1$.

We still have to check that $x^* \in C_x^*$. Let $y \in C$ and $\varepsilon > 0$. As before, we can find k such that $\|x - x_k\| < \varepsilon$ and $|\langle x^* - x_k^*, x - y\rangle| < \varepsilon$. We have

$$\begin{aligned}\langle x^*, x\rangle &= \langle x^* - x_k^*, x - y\rangle + \langle x_k^*, x - y\rangle + \langle x^*, y\rangle \\ &\ge -\varepsilon + \langle x_k^*, x - x_k\rangle + \langle x_k^*, x_k - y\rangle + \langle x^*, y\rangle \\ &\ge -\varepsilon - \|x_k^*\| \cdot \|x - x_k\| + \langle x^*, y\rangle \\ &\ge -\varepsilon - m\varepsilon + \langle x^*, y\rangle\end{aligned}$$

and therefore $\langle x^*, x\rangle \ge \langle x^*, y\rangle$ for any $y \in C$, proving that $x^* \in C_x^*$. Thus $x \in F_m$ and the proof of the theorem is complete. □

Comments

Most of the properties of non-support points (our quasi-interior) of a convex set can be found in [**9**]. Our proof of Theorem 2.10 follows [**15**], where it is proved in the more general setting of metrizable, locally convex spaces.

Chapter II

MINIMAL USCO MAPS

3. Usco and Minimal Usco Maps

3.1. Given a set Z we shall denote by 2^Z the set of all non-empty subsets of Z. If A is another set, a *multivalued map* from A to Z is a map $F\colon A \to 2^Z$. A *selection* for a multivalued map $F\colon A \to 2^Z$ is a map $\sigma\colon A \to Z$ such that $\sigma(a) \in F(a)$ for all $a \in A$.

DEFINITIONS. Assume that A and Z are topological spaces (always Hausdorff).

(1) A multivalued map $F\colon A \to 2^Z$ is called *upper semicontinuous* at $a \in A$ if for every open subset U of Z such that $F(a) \subset U$, the set $F^{-1}(U) = \{x \in A : F(x) \subset U\}$ is a neighborhood of a in A. If it is upper semicontinuous at each point of A, F is called *upper semicontinuous* (on A).

(2) A multivalued map $F\colon A \to 2^Z$ is called *usco* if it is upper semicontinuous and compact valued (i.e., $F(a)$ is a compact subset of Z for every $a \in A$).

(3) A multivalued map $F\colon A \to 2^Z$ is called *convex* if Z is a vector space and $F(a)$ is a convex subset of Z for every $a \in A$.

(4) The *graph* of a multivalued map $F\colon A \to 2^Z$ is the set

$$\mathcal{G}(F) = \{(a, z) \in A \times Z \,:\, z \in F(a)\}.$$

We shall partially order the set of all multivalued maps from A to Z as follows: $F_1 \preceq F_2$ if $\mathcal{G}(F_1) \subseteq \mathcal{G}(F_2)$. Thus a *minimal usco map* is a usco map whose graph does not properly contain the graph of any other usco map. Similarly, a *minimal convex usco map* is a convex usco map whose graph does not properly contain the graph of any other convex usco map.

3.2. LEMMA. *For every (convex) usco map $F\colon A \to 2^Z$ there exists a minimal (convex) usco map $F_0\colon A \to 2^Z$ such that $F_0 \preceq F$.*

PROOF. This follows from Zorn's Lemma. Indeed, let $(F_\alpha)_\alpha$ be a decreasing chain of (convex) usco maps contained in F. For each $a \in A$ set $G(a) = \bigcap F_\alpha(a)$. It is sufficient to prove that G is a (convex) usco map. Let $a \in A$. Since every $F_\alpha(a)$ is compact and non-empty, $G(a)$ is compact and non-empty (and convex in the convex case). To prove that G is upper semicontinuous at a, let U be an open subset of Z such that $G(a) \subseteq U$. Assume that for any α, $F_\alpha(a) \not\subseteq U$ and therefore $F_\alpha(a) \setminus U \neq \emptyset$. By a compactness argument, $G(a) \setminus U = \bigcap (F_\alpha(a) \setminus U) \neq \emptyset$; this contradiction shows that there exists α such that $F_\alpha(a) \supseteq U$. It follows that $F_\alpha^{-1}(U)$ is a neighborhood of a in A. Since, by definition, $G^{-1}(U) \supseteq F_\alpha^{-1}(U)$, $G^{-1}(U)$ is a neighborhood of a in A, proving that G is upper semicontinuous at a.

□

3.3. LEMMA. *Let $F\colon A \to 2^Z$ be a multivalued map and $a \in A$. The following assertions are equivalent:*

(1) *$F(a)$ is compact and F is upper semicontinuous at a.*

(2) *If (a_α) is a net in A converging to a and (z_α) is a net in Z such that $z_\alpha \in F(a_\alpha)$ for any α, then the set consisting of all cluster points of the net (z_α) is non-empty and is contained in $F(a)$.*

PROOF. Assume that (1) is true and that the net (z_α) has no cluster points. Then every $z \in F(a)$ has an open neighborhood V_z and there exists α_z such that $z_\alpha \notin V_z$ if $\alpha \geq \alpha_z$. Since $F(a)$ is compact, there exist $z_1, z_2, \ldots, z_n \in F(a)$ such that $F(a) \subseteq V = V_{z_1} \cup \ldots \cup V_{z_n}$. Also let β be such that $\beta \geq \alpha_{z_i}$, $i = 1, \ldots, n$. Since F is upper semicontinuous and $F(a) \subseteq V$, $D = F^{-1}(V)$ is a neighborhood of a in A and hence there exists $\gamma \geq \beta$ such that $a_\alpha \in D$ for any $\alpha \geq \gamma$. Thus $z_\alpha \in F(a_\alpha) \subseteq V$ for any $\alpha \geq \gamma$, contradicting the construction of V. It follows that the net (z_α) has cluster points. Suppose next that $z \notin F(a)$ is a cluster point of the net (z_α). Then there exist an open neighborhood U of z and an open set W containing $F(a)$ such that $U \cap W = \emptyset$. There exist α_1 and α_2 such that $z_\alpha \in U$ if $\alpha \geq \alpha_1$ (since z is a cluster point of the net (z_α)) and $F(a_\alpha) \subseteq W$ if $\alpha \geq \alpha_2$ (since F is upper semicontinuous at a). Thus, if $\alpha \geq \alpha_1$ and $\alpha \geq \alpha_2$, then $z_\alpha \in U \cap W$. This contradiction proves that (2) is true.

Assume now that (2) is true. Then any net in $F(a)$ has a cluster point in $F(a)$, proving that $F(a)$ is compact. Suppose that F is not upper semicontinuous at a. Then there exists an open set V containing $F(a)$ and such that for each neighborhood U of a there exists $a_U \in U$ and $z_U \in F(a_U)$ with $z_U \notin V$. This implies that the net (z_U) has no cluster points in $F(a)$, contradicting (2). Therefore F is upper semicontinuous at a and the lemma is completely proved. □

3.4. COROLLARY. *Let $F\colon A \to 2^Z$ be a usco map. Then $\mathcal{G}(F)$, the graph of F, is a closed subset of $A \times Z$.*

PROOF. Let (a_α, z_α) be a net in $\mathcal{G}(F)$ converging to $(a.z) \in A \times Z$. Then (a_α) converges to a, z_α converges to z and $z_\alpha \in F(a_\alpha)$ for any α. By the previous lemma $z \in F(a)$ and therefore $(a, z) \in \mathcal{G}(F)$. □

3.5. LEMMA. *Let F, $G\colon A \to 2^Z$ be multivalued maps such that $F \preceq G$, $\mathcal{G}(F)$ is closed and G is usco. The F is usco.*

PROOF. Let (a_α) be a net in A converging to a and (z_α) be a net in Z such that $z_\alpha \in F(a_\alpha)$ for every α, that is (a_α, z_α) is a net in $\mathcal{G}(F)$. Since G is usco and $\mathcal{G}(F) \subseteq \mathcal{G}(G)$, Lemma 3.3 implies that this net has at least one cluster point. Since $\mathcal{G}(F)$ is closed, all its cluster points belong to $F(a)$. The assertion follows now from Lemma 3.3. □

3.6. PROPOSITION. *Let $F\colon A \to 2^Z$ be a usco map and $p\colon A \times Z \to A$ be the projection. The following assertions are equivalent:*

(1) F is a minimal usco map.

(2) For any proper, closed subset S of $\mathcal{G}(F)$, $p(S)$ is strictly contained in A.

(3) For any $(a, z) \in \mathcal{G}(F)$ and any open neighborhoods U of a and W of z, there exists a non-empty open subset V of U such that $F(V) \subseteq W$.

PROOF. Assume that F is minimal and that $p(S) = A$ for a proper and closed subset S of $\mathcal{G}(F)$. Then S is the graph of a multivalued map $G\colon A \to 2^Z$ and $G \preceq F$. Since S is closed, the above lemma implies that G is usco and therefore, by the minimality of F, $G = F$. This implies that $S = \mathcal{G}(G) = \mathcal{G}(F)$, contradicting one of our assumptions. This shows that (1) implies (2).

Assume now that (2) is true and let (a, z), U and W be as in (3). Then $p(\mathcal{G}(F) \setminus U \times W) \neq A$; let $a_1 \in A \setminus p(\mathcal{G}(F) \setminus U \times W) \neq A$. It follows that $a_1 \in U$ and $F(a_1) \subseteq W$. Since F is usco, there exists an open neighborhood V_1 of a_1 such that $F(V_1) \subseteq W$. Then $V = V_1 \cap U$ satisfies the conclusion of (3). This shows that (2) implies (3).

Finally, assume that (3) is true and that F is not minimal. Then there exists a usco map $H\colon A \to 2^Z$ such that $H \prec F$. It follows that there exist $a \in A$ and $z \in F(a)$ such that $z \notin H(a)$. Since $H(a)$ is compact, we can find open neighborhoods W of z and W_1 of $H(a)$ such that $W \cap W_1 = \emptyset$. The fact that H is usco implies that there exists an open neighborhood U of a such that $H(U) \subseteq W_1$. By our assumption, there exists a non-empty open subset V of U such that $F(V) \subseteq W$. But then $F(V) \cap H(V) = \emptyset$, which is impossible. This proves that (3) implies (1).

□

3.7. LEMMA. *Let A be a Hausdorff space, X be a complete, locally convex, Hausdorff topological vector space and $F\colon A \to 2^X$ be a usco map. Then the multivalued map $\overline{\mathrm{co}}\, F\colon A \to 2^X$ defined by $(\overline{\mathrm{co}}\, F)(x) = \overline{\mathrm{co}}\,\big(F(x)\big)$ is a convex usco map. If in addition F is a minimal usco map, then $\overline{\mathrm{co}}\, F$ is a minimal convex usco* *map.*

PROOF. First recall that for a subset S of X, $\overline{\text{co}}(S)$ denotes the closed convex hull of S. Then $\overline{\text{co}}\, F$ is convex and compact valued. We need to show that it is upper semicontinuous. We shall use the fact that in a locally convex vector space a compact and convex set K has a neighborhood base consisting of sets of the form $K+W$, where W is a closed convex neighborhood of the origin. Let $a \in A$ and V be an open subset of X such that $\overline{\text{co}}\, F(a) \subseteq V$. Then there exists a closed and convex neighborhood W of the origin such that $\overline{\text{co}}\, F(a) \subseteq \overline{\text{co}}\, F(a) + W \subseteq V$; in particular $\overline{\text{co}}\, F(a) + W$ is a neighborhood of $F(a)$. Since F is upper semicontinuous, there exists an open neighborhood U of a in A such that $F(U) \subseteq \overline{\text{co}}\, F(a) + W$. It follows that $(\overline{\text{co}}\, F)(U) \subseteq \overline{\text{co}}\, F(a) + W \subseteq V$, proving that $\overline{\text{co}}\, F$ is upper semicontinuous at a.

To prove the second assertion, consider a convex usco map $H\colon A \to 2^X$ such that $H \preceq \overline{\text{co}}\, F$. We need to show that $H = \overline{\text{co}}\, F$. This is obvious if $F \preceq H$. Assume therefore that $F \not\preceq H$. Then there exist $a_0 \in A$ and $x_0 \in F(a_0)$ such that $x_0 \notin H(a_0)$. Since $H(a_0)$ is closed and convex, by the separation theorem there exists $z^* \in X^*$ and $t \in R$ such that

$$\langle z^*, y\rangle < t < \langle z^*, x_0\rangle, \qquad \text{for all } y \in H(a_0).$$

Let

$$W^- = \{x \in X : \langle z^*, x\rangle < t\} \quad \text{and} \quad W^+ = \{x \in X : \langle z^*, x\rangle > t\}.$$

Then W^- is an open subset of X containing $H(a_0)$. Since H is upper semicontinuous, there exists an open neighborhood U of a_0 such that $H(a) \subset W^-$ for all $a \in U$. Since W^+ is also open in X and $x_0 \in F(a_0) \cap W^+$, by Proposition 3.6 there exists a non-empty open subset V of U such that $F(V) \subset W^+$. Then $\overline{\text{co}}\, F(a) \subset X \setminus W^-$ for all $a \in V$, implying that $\overline{\text{co}}\, F(a) \cap H(a) = \emptyset$ for all $a \in V$. Since $H \preceq \overline{\text{co}}\, F$, this is impossible. This contradiction shows that $F \preceq H$ and, as noticed above, this proves the minimality of $\overline{\text{co}}\, F$. □

3.8. LEMMA. *Let $F: A \to 2^Z$ be a minimal (convex) usco map and U be an open or dense subset of A. Then the restriction of F to U is a minimal (convex) usco map.*

PROOF. Consider first the case when U is open. Let $(a, z) \in \mathcal{G}(F|U)$ and let U_1 be an open neighborhood of a in U and W be an open neighborhood of z. Since $(a, z) \in \mathcal{G}(F)$ and F is minimal, there exists a non-empty open set V of U_1 such that $F(V) \subseteq W$. Since $(F|U)(V) = F(V)$, Proposition 3.6 implies that $F|U$ is minimal.

Similar arguments can be used to prove the minimality of $F|U$ in the case of a dense U.

Assume now that $F: A \to 2^Z$ be a minimal convex usco map and let $F_1: A \to 2^Z$ be a minimal usco map such that $F_1 \preceq F$. Clearly $\overline{\mathrm{co}}\, F_1 = F$ and therefore

$$F|U = (\overline{\mathrm{co}}\, F_1)|U = \overline{\mathrm{co}}\,(F_1|U).$$

We just proved that $F_1|U$ is a minimal usco map; from Lemma 3.7 it follows that $\overline{\mathrm{co}}\,(F_1|U)$ is a minimal convex usco map and therefore so is $F|U$. □

4. Existence of Selections

There exists an interesting relation between the fact that a minimal usco map F is single valued at a point and the continuity of a selection for F at that point. Namely

4.1. LEMMA. *Let $F: A \to 2^Z$ be a minimal usco map and $\sigma: A \to Z$ be a selection for F. Then σ is continuous at $x \in A$ if and only if $F(x)$ is a singleton (as a matter of fact $F(x) = \{\sigma(x)\}$).*

PROOF. Define $G: A \to 2^Z$ as follows: $z \in G(x)$ if and only if there exists a net (x_α) in A converging to x such that z is a cluster point of the net $(\sigma(x_\alpha))$. Since

F is usco, by Lemma 3.3, $G \preceq F$. It is obvious that $\mathcal{G}(G) = \mathrm{cl}(\mathcal{G}(\sigma))$ and therefore, by Lemma 3.5, G is a usco map. The minimality of F implies that $G = F$.

Assume now that σ is continuous at x. Then, by definition, $G(x) = \{\sigma(x)\}$ and thus $F(x) = \{\sigma(x)\}$. Conversely, if $F(x) = \{\sigma(x)\}$, the fact that F is upper semicontinuous at x implies that σ is continuous at x. □

4.2. COROLLARY. *Let $F\colon A \to 2^Z$ be a minimal usco map, Z being metrizable. Then the set of points at which F is single valued is a G_δ subset of A.*

PROOF. Let $\sigma\colon A \to Z$ be a selection for F. It is well known that σ is continuous on a G_δ subset of A. The corollary follows immediately from the previous lemma. □

4.3. The conclusion of the above corollary is unfortunately not strong enough to imply the existence of at least one point at which a minimal usco map is single valued, even if Z is metrizable. In studying this type of problems Stegall (see [**21**], [**22**]) was led to introduce and develop the following classes of topological spaces, called **C** and **S** respectively.

DEFINITIONS. (1) A Hausdorff topological space Z is in **C** if any minimal usco map $F\colon A \to 2^Z$ defined on a Baire space A is single valued on a dense G_δ subset of A.

(2) A Banach space X is in the class **S** if X^* endowed with the w* topology is in **C**.

REMARKS. (1) If Z is in **C** and A is a Baire space, then any usco map $F\colon A \to 2^Z$ has a selection that is continuous on a dense G_δ subset of A. Indeed, by Lemma 3.2 there exists a minimal usco map $F_0\colon A \to 2^Z$ such that $F_0 \preceq F$; since any selection of F_0 is a selection of F too, the assertion follows from the definition of the class **C** and Lemma 4.1. As a matter of fact one can easily see that the converse assertion is also true.

(2) Let X be a Banach space in the class **S** and let $F\colon A \to 2^{X^*}$ be a minimal convex w* usco map, A being a Baire space. By Lemma 3.2 there exists a minimal w* usco map $G\colon A \to 2^{X^*}$ such that $G \preceq F$. Lemma 3.7 and the minimality of F imply that $\overline{\text{co}}(G) = F$. Since X^* is in the class **C**, G is single valued on a dense G_δ subset A_0 of A and therefore F is also single valued on A_0.

Before stating the next result we need to introduce some additional notation. Let $F\colon A \to 2^Z$ be a usco map and let τ be the topology of Z. If we need to emphasize that F is usco with respect to this topology on Z, we shall say that F is τ-usco. The same convention will be used for other topological notions.

4.4. PROPOSITION. *Let B be a Baire space, let Z be a Hausdorff topological space and let τ denote the topology of Z. Let $F\colon B \to 2^Z$ be a minimal τ-usco map. Assume that there exists a metric d on Z with the following property: every open non-empty subset U of B has an open non-empty subset V such that $F(V)$ contains non-empty, relatively τ-open subsets of arbitrarily small d-diameter. Then there exists a dense G_δ subset of B on which F is single valued and d-upper semicontinuous.*

PROOF. For any $\varepsilon > 0$ let

$$O_\varepsilon = \bigcup\{G : G \subseteq B \text{ is open and } \operatorname{diam}(F(G)) \leq \varepsilon\}.$$

Clearly O_ε is open in B. We want to show that O_ε is also dense in B.

Let U be a non-empty open subset of B. By hypothesis there exists a non-empty open subset V of U and a τ-open subset W of Z such that $F(V) \cap W \neq \emptyset$ and $\operatorname{diam}(F(V) \cap W) \leq \varepsilon$. Take $z \in F(V) \cap W$ and $x \in V$ such that $z \in F(x)$. Since F is minimal, by Proposition 3.6 there exists a non-empty open subset G contained in V such that $F(G) \subseteq W$, hence $\operatorname{diam}(F(G) \leq \varepsilon$. Then $G \subseteq O_\varepsilon$ and thus $U \cap O_\varepsilon \neq \emptyset$. This proves that O_ε is dense in B.

Now let $D = \bigcap O_{1/n}$. Since A is a Baire space, D is dense in A. From the definition of D one can easily deduce that F is single valued and d-upper semicontinuous at all points of D. □

4.5. COROLLARY. *If Z is a metric space, then $Z \in \mathbf{C}$.*

PROOF. It is enough to apply the previous proposition to the case when τ is the metric topology on Z. □

4.6. Before we can state our next result, we need to introduce two more definitions. Let X be a Banach space.

DEFINITIONS. (1) Let V be a non-empty, bounded subset of X^* and let $x \in X$ and $\alpha > 0$. A w* *slice* of V is the subset of V defined by

$$\Sigma(x, V, \alpha) = \{x^* \in V \colon \langle x^*, x\rangle > \sup\{\langle y^*, x\rangle : y^* \in V\} - \alpha\}.$$

(2) We say that X^* has the *Radon-Nikodým property* if every non-empty bounded subset of X^* has w* slices of arbitrarily small diameter.

4.7. COROLLARY. *Let X be a Banach space such that X^* has the Radon-Nikodým property and let $F\colon A \to 2^{X^*}$ be minimal (convex) w* usco map, A being a Baire space. Then there exists a dense G_δ subset A_0 of A such that F is single valued and norm upper semicontinuous at each point of A_0. If F is not minimal, then there exists a dense G_δ subset A_0 of A and a selection σ for F which is norm continuous at every point of A_0. In particular, $X \in \mathbf{S}$.*

PROOF. We can use the lemma immediately following this corollary to find a dense, relatively open subset D of A such that $F|D\colon D \to 2^{X^*}$ is a locally bounded, minimal w* usco map. Let U be any relatively open subset of A. Since $F|D$ is locally bounded, there exists an open subset V of U such that $F(V)$ is bounded in X^*. Let $\varepsilon > 0$. By assumption, there exists a slice Σ of $F(V)$ whose norm diameter is less than ε. By definition, Σ is relatively w* open in $F(V)$. Thus $F(V)$ contains non-empty, relatively w* open subsets of arbitrarily small norm diameter. By Proposition 4.4, there exists a dense G_δ subset A_0 of A such that $F(x)$ is single valued and norm usco at each $x \in A_0$.

The case of a minimal convex w* usco can be handled as in the remark in section 4.3.

If F is not minimal, apply the first part of the corollary to a minimal w* usco map contained in F and then use Lemma 4.1. The last assertion is obvious. □

4.8. LEMMA. *Let X be a Banach space, A be a Baire space and $F\colon A \to 2^{X^*}$ be a minimal (convex) w* usco map. Then there exists an open, dense subset D of A such that $F|D\colon D \to 2^{X^*}$ is a locally (norm) bounded, minimal (convex) w* usco map.*

PROOF. For every positive integer n define

$$A_n = \{x \in A : F(x) \cap nB^* \neq \emptyset\}.$$

Using the w* upper semicontinuity of F and the lower w* semicontinuity of the norm, one can easily prove that A_n is closed in A. Clearly $A = \bigcup A_n$ and, since A is Baire, the open subset $D = \bigcup \mathrm{int}(A_n)$ is dense in A. Let $n > 0$ be such that $\mathrm{int}(A_n) \neq \emptyset$ and set $F_n = F|\mathrm{int}(A_n)$. By Lemma 3.8, F_n is a minimal (convex) w* usco map. Since $\mathcal{G}(F_n) \cap (A \times nB^*)$ is a closed subset of $\mathcal{G}(F_n)$ whose projection on $\mathrm{int}(A_n)$ is $\mathrm{int}(A_n)$, Proposition 3.6 (or the remark following it) implies that

$$\mathcal{G}(F_n) \cap (A \times nB^*) = \mathcal{G}(F_n)$$

and therefore

$$\mathcal{G}(F|\mathrm{int}(A_n)) \subseteq A \times nB^*.$$

This proves that $F|D$ is a locally (norm) bounded, multivalued map. From Lemma 3.8 we get that $F|D$ is a minimal w* usco map. □

4.9. THEOREM. *Let A be a Baire space, X be a Banach space and $F\colon A \to 2^{X^*}$ be a minimal (convex) w* usco map. Then there exists a dense G_δ subset A_0 of A such that $F(x)$ is contained in a sphere of X^* for every $x \in A_0$.*

PROOF. Step I. Define $\psi\colon A \to R$ as follows:

$$\psi(a) = \min\{\|x^*\| : x^* \in F(a)\}$$

(the definition is correct since $F(a)$ is w* compact and $\| \ \|$ is w* lower semicontinuous). We shall prove that ψ is lower semicontinuous on A. Let (a_α) be a net in A converging to a. For every α choose $x^*_\alpha \in F(a_\alpha)$ such that $\psi(a_\alpha) = \|x^*_\alpha\|$. Since F is a w* usco map, the set of w* cluster points of the net (x^*_α) is non-empty and contained in $F(A)$. Since the $\| \ \|$ is w* lower semicontinuous, we can find a w* cluster point $x^* \in F(a)$ of the net (x^*_α) such that the first inequality below is satisfied

$$\liminf \psi(a_\alpha) = \liminf \|x^*_\alpha\| \geq \|x^*\| \geq \psi(a).$$

This proves our assertion.

Step II. Define a (convex) multivalued map $F_0 \colon A \to 2^{X^*}$ by

$$F_0(a) = \{x^* \in F(a) : \|x^*\| = \psi(a)\} = F(a) \cap \{x^* \in X^* : \|x^*\| \leq \psi(a)\}.$$

Since $F(a)$ is w* compact and $\| \ \|$ is w* lower semicontinuous, $F_0(a)$ is also w* compact. As a matter of fact, F_0 is w* upper semicontinuous at each point were ψ is continuous. Indeed, let $a \in A$ be a point at which ψ is continuous. Consider a net (a_α) in A converging to a and a net (x^*_α) in X^* such that $x^*_\alpha \in F_0(a_\alpha)$. Since $F_0(a_\alpha) \subseteq F(a_\alpha)$ and F is a w* usco map, by Lemma 3.3 the set of w* cluster points of the net (x^*_α) is non-empty and contained in $F(a)$. As above, we can find a w* cluster point $x^* \in F(a)$ of this net such that $\liminf \|x^*_\alpha\| \geq \|x^*\|$. Using the continuity of ψ at a we get

$$\psi(a) = \lim \psi(a_\alpha) = \liminf \|x^*_\alpha\| \geq \|x^*\| \geq \psi(a).$$

It follows that $\psi(a) = \|x^*\|$ and therefore $x^* \in F_0(a)$. By Lemma 3.3 again, we conclude that F_0 is w* upper semicontinuous at a.

Step III. The function ψ being defined on a Baire space and being lower semicontinuous, is continuous at each point of a dense G_δ subset A_0 of A. It follows that F_0 is w* upper semicontinuous at each point of A_0.

Let $\mathcal{G}$ be the closure of $\mathcal{G}(F_0)$ in $A \times X^*$; clearly $\mathcal{G} \subset \mathcal{G}(F)$. Let G be the unique multivalued map whose graph is $\mathcal{G}$ (in the convex case G may not be convex valued and we'll have to replace it in the following discussion by $\overline{\text{co}}\, G$). By Lemma 3.5, G is a usco map and since F is a minimal usco map, it follows that $G = F$. Since F_0 is upper semicontinuous at each point of A_0, Lemma 3.3 implies that $F_0 \,|\, A_0 = G \,|\, A_0$ and therefore $F \,|\, A_0 = F_0 \,|\, A_0$. This shows that for every $a \in A_0$ and $x \in F(a)$ we have $\|x^*\| = \psi(a)$, which proves our assertion. □

Before giving another example of Banach spaces contained in the class $\mathbf{S}$, let us recall that a norm $\| \ \|$ on a vector space X is called *strictly convex* if $\|x + y\| < \|x\| + \|y\|$ for any two distinct points x, $y \in X$ such that $\|x\| = \|y\|$.

4.10. COROLLARY. *Let X be a Banach space that admits an equivalent norm whose dual norm is strictly convex. Then $X \in \mathbf{S}$.*

PROOF. Let A be a Baire space and $F\colon A \to 2^{X^*}$ be a minimal w* usco map. By Theorem 4.9 there exists a G_δ subset A_0 of A such that $\overline{\text{co}}\, F(a)$ is contained in sphere of X^* for each $a \in A_0$. Since the norm of X^* is strictly convex, any convex subset of a sphere in X^* must be a singleton. Thus $\overline{\text{co}}\, F$ is single valued on A_0, implying that F is single valued on A_0. This shows that $X^* \in \mathbf{C}$ and therefore $X \in \mathbf{S}$. □

4.11. COROLLARY. *Any separable, or any reflexive, or, more generally, any weakly compactly generated Banach space is in the class* $\mathbf{S}$.

PROOF. Any separable (respectively reflexive) Banach space is weakly compactly generated. By Theorem 3 in [**1**], any weakly compactly generated Banach space admits an equivalent norm whose dual norm is strictly convex. Now we can apply Corollary 4.10. □

Comments

In general we have followed [**14**] (see also [**5**] and [**6**]). We were able to simplify some of the proofs by making consistent use of Lemma 3.3 (the fact that (2) implies (1) seems to be new). Proposition 4.4 is due to J. Rainwater [**16**]. The fact that a

Banach space whose dual has the Radon-Nikodým property is in class **S** (Corollary 4.4) is due to Stegall [**22**]. Theorem 4.9 is an extension of a result of Kendorov [**7**], [**8**] who proved it for monotone operators. I could not find any reference for Corollary 4.10; the proof given here is an improvement of one sketched to the author by Kenderov in August 1988. As already mentioned, the classes **C** and **S** were introduced by Stegall ([**21**] and [**22**]) who studied them intensively.

Chapter III

MONOTONE OPERATORS

5. Locally Efficient Monotone Operators

5.1. From now on we shall denote by X a Banach space, by X^* its topological dual, by B^* the unit ball in X^* and by $B(x, \varepsilon)$ the ball in X centered at x and with radius $\varepsilon > 0$. Let A be a subset of X.

DEFINITIONS. (1) A subset $G \subseteq A \times X^*$ is called *monotone* if

$$\langle x^* - y^*, x - y \rangle \geq 0 \quad \text{for any } (x, x^*),\ (y, y^*) \in G\,.$$

If, in addition, G is maximal (under set inclusion) in the family of all monotone sets contained in $A \times X^*$, then G is called *maximal monotone in* $A \times X^*$.

(2) A multivalued map $T\colon A \to 2^{X^*}$, with $T(x) \neq \emptyset$ for all $x \in A$, is called a *(maximal) monotone operator on* A if its graph $\mathcal{G}(T)$ is a (maximal) monotone set in $A \times X^*$.

(3) Let $F\colon A \to 2^{X^*}$ be a multivalued map and $r > 0$. F is called r*-efficient* at $a \in A$ if there exists a relative neighborhood U of a in A such that $F(x) \cap rB^* \neq \emptyset$ for all $x \in U$. If F is r-efficient at a but we don't need to specify r, we shall say simply that F is *efficient* at a. If for every $a \in A$ there exists a real number $r_a > 0$ such that F is r_a-efficient at a, we call F *locally efficient*.

REMARKS. (1) Let $F\colon A \to 2^{X^*}$ be a multivalued map and $r > 0$. Set

$$A_r = \{x \in A : F(x) \cap rB^* \neq \emptyset\}$$

and define $F_r\colon A_r \to 2^{X^*}$ by

$$F_r(x) = F(x) \cap rB^*.$$

It is obvious that F is r-efficient at $a \in A$ if and only if $F_r(x) \neq \emptyset$ for all x in a relative neighborhood of a in A. Let us also notice that if F is r-efficient at a, then F_r is bounded in a relative neighborhood of a in A.

(2) An easy application of Zorn's lemma shows that for every monotone operator $T\colon A \to 2^{X^*}$ there exists a maximal monotone operator $S\colon A \to 2^{X^*}$ such that $T \preceq S$.

5.2. EXAMPLES. (1) If $A \subseteq R$, then a function $f\colon A \to R$ is non decreasing if and only if it is monotone in the above sense (we identify R^* with R as usually).

(2) If H is a real Hilbert space and $T : H \to H$ is a linear map, then T is monotone if and only if it is a positive operator, i.e., $\langle Tx, x\rangle \geq 0$ for all $x \in H$ (we identify H^* with H as usual).

(3) Let C be a bounded, closed, convex subset of the Hilbert space H and let $U\colon C \to C$ be a (generally non linear) non expansive map, i.e., $\|U(x) - U(y)\| \leq \|x - y\|$ for any $x, y \in C$. Then $T = Id - U\colon C \to H$ is monotone. Indeed

$$\begin{aligned}
\langle T(x) - T(y), x - y\rangle &= \langle x - y - U(x) + U(y), x - y\rangle \\
&= \|x - y\|^2 - \langle U(x) - U(y), x - y\rangle \\
&\geq \|x - y\|^2 - \|U(x) - U(y)\| \cdot \|x - y\| \geq 0.
\end{aligned}$$

Note that $T^{-1}(0)$ is the fixed point set of U.

(4) The most important example for us is the following one. Let C be a convex subset of the Banach space X, A be a relatively open subset of C and $f: C \to R$ be convex and locally Lipschitz on A. Then the subdifferential map $\partial f: A \to 2^{X^*}$ is a monotone operator that is efficient at each $x \in A$. The proof of this assertion and of other important properties of the subdifferential map will be given in the next chapter.

5.3. LEMMA. *Let $A \subseteq X$ and $T: A \to 2^{X^*}$ be a monotone operator.*

(1) T is maximal if and only if $(y, y^*) \in A \times X^*$ and $\langle y^* - x^*, y - x\rangle \geq 0$ for all $(x, x^*) \in \mathcal{G}(T)$ imply $y^* \in T(y)$.

(2) If T is maximal, then $T(x)$ is a convex w^* closed subset of X^* for all $x \in A$.

PROOF. (1) Assume first that T is a maximal monotone operator and let $(y, y^*) \in A \times X^*$ satisfy $\langle y^* - x^*, y - x\rangle \geq 0$ whenever $x \in A$ and $x^* \in T(x)$. Define $S: A \to 2^{X^*}$ by $S(y) = T(y) \cup \{y^*\}$ and $S(x) = T(x)$ if $x \neq y$. Clearly S is a monotone operator on A and $\mathcal{G}(T) \subseteq \mathcal{G}(S)$. Since T is maximal on A, it follows that $\mathcal{G}(T) = \mathcal{G}(S)$ and thus $y^* \in T(y)$.

Conversely, assume that T satisfies the stated condition and let $S: A \to 2^{X^*}$ be a monotone operator such that $\mathcal{G}(T) \subseteq \mathcal{G}(S)$. Let $(y, y^*) \in \mathcal{G}(S)$ and $(x, x^*) \in \mathcal{G}(T) \subseteq \mathcal{G}(S)$; since S is monotone, $\langle y^* - x^*, y - x\rangle \geq 0$. In view of the hypothesis, $y^* \in T(y)$. This shows that $\mathcal{G}(T) = \mathcal{G}(S)$ and therefore T is maximal.

(2) Let $x \in A$, x^*, $y^* \in T(x)$, $\alpha \in [0, 1]$ and $u^* = (1 - \alpha)x^* + \alpha y^*$. For every $(z, z^*) \in \mathcal{G}(T)$ we have

$$\begin{aligned}\langle u^* - z^*, x - z\rangle &= \langle (1 - \alpha)x^* + \alpha y^* - z^*, x - z\rangle \\ &= (1 - \alpha)\langle x^* - z^*, x - z\rangle + \alpha\langle y^* - z^*, x - z\rangle \geq 0\end{aligned}$$

(the inequality is due to the monotonicity of T). Since T is maximal, (1) implies that $u^* \in T(x)$, proving that $T(x)$ is convex.

To prove that $T(x)$ is w* closed in X^*, let (x_i^*) be a net in $T(x)$, w* convergent to x^*, and let $(y, y^*) \in \mathcal{G}(T)$. Then, since T is monotone,

$$\langle x_i^* - y^*, x - y \rangle \geq 0$$

and therefore

$$\langle x^* - y^*, x - y \rangle \geq 0.$$

Using again the maximality of T and (1) we obtain that $x^* \in T(x)$. □

5.4. LEMMA. *Let A be a quasi-open subset of X, $T\colon A \to 2^{X^*}$ be a monotone operator and let $a \in A$. Then T is locally bounded at a if and only if it is efficient at a.*

PROOF. If T is locally bounded at a, it is obviously efficient at a. Assume it is r-efficient at $a \in A$. Then there exists $\varepsilon > 0$ such that $T(x) \cap rB^* \neq \emptyset$ for every $x \in B(a, \varepsilon) \cap A$. Let $x \in B(a, \varepsilon) \cap A$ and $y \in X$; since $x \in \mathrm{qi}(A)$ there exist a sequence (t_n) of positive real numbers convergent to 0 and a sequence $(y_n) \subset X$ convergent to y such that $x + t_n y_n \in A \cap B(a, \varepsilon)$. It follows that for each n there exists $y_n^* \in T(x + t_n y_n) \cap rB^*$. Let $x^* \in T(x)$. Since T is monotone

$$\langle y_n^* - x^*, t_n y_n \rangle \geq 0,$$

which implies that

$$x^*(y_n) \leq y_n^*(y_n) \leq r\|y_n\|$$

and thus $x^*(y) \leq r\|y\|$. This shows that $\|x^*\| \leq r$ and therefore the lemma is proved. □

5.5. LEMMA. *Let $T\colon A \to 2^{X^*}$ be a monotone operator and a be an absorbing point of A. Then T is locally bounded at a, i.e., there exist a neighborhood U of a in A and $r > 0$ such that $T(U) \subseteq rB^*$. In particular, T is locally bounded on $\mathrm{int}(A)$.*

Proof. Let $a^* \in T(a)$, $A_0 = A - a$ and define $T_0 \colon A_0 \to 2^{X^*}$ by $T_0(x) = T(x+a) - a^*$, $x \in A_0$. Clearly $0 \in A_0$, $0 \in T_0(0)$ and T is locally bounded at a if and only if T_0 is locally bounded at 0. So, without loss of generality, we can assume that $a = 0$ and $0 \in T(0)$.

Define $f \colon X \to (-\infty, +\infty]$ by

$$f(x) = \sup\{\langle y^*, x - y\rangle \colon y \in A,\ \|y\| \le 1,\ y^* \in T(y)\}.$$

Observe that f is always non negative (since $0 \in T(0)$) and that $f(0) = 0$ (since T is monotone). Being the supremum of a family of continuous affine functionals, f is convex and lower semicontinuous. We want to prove first that 0 is an absorbing point of $\operatorname{dom} f = \{x \in X \colon f(x) < \infty\}$, the effective domain of f.

Let $x \in X$. Since 0 is an absorbing point of A, there exists $t > 0$ such that $tx \in A$. Fix $u^* \in T(tx)$. Then, since T is monotone, for any $y \in A$ such that $\|y\| \le 1$ and any $y^* \in T(y)$ we have

$$\langle y^*, tx - y\rangle \le \langle u^*, tx - y\rangle \le \|u^*\|(t\|x\| + 1)$$

showing that $f(tx) < \infty$, i.e., $tx \in \operatorname{dom}(f)$. Thus 0 is an absorbing point of $\operatorname{dom}(f)$.

Let $C = \{x \in X \colon f(x) \le 1\}$. Since f is convex, C is convex. Since f is lower semicontinuous, C is closed. Clearly $0 \in C$. Since 0 is an absorbing point of $\operatorname{dom}(f)$, for any $x \in X$ there exists $t > 0$ such that $f(tx) < \infty$. Using the convexity of f, it follows that for a sufficiently small s we have

$$f(stx) = f((1-s)0 + stx) \le (1-s)f(0) + sf(tx) = sf(tx) \le 1,$$

which proves that 0 is an absorbing point of C. Then $C_0 = C \cap (-C)$ is a barrel and, since X is Banach, there exists $0 < \delta < 1/2$ such that $B(0, 2\delta) \subseteq C_0$. In particular

$$f(x) \le 1 \quad \text{if } \|x\| < 2\delta.$$

For any $y \in A \cap B(0,\delta)$, $y^* \in T(y)$ and $x \in B(0, 2\delta)$ we have

$$\langle y^*, x - y\rangle \le f(x) \le 1$$

and thus

$$\langle y^*, x\rangle \leq 1 + \langle y^*, y\rangle \leq 1 + \delta\|y^*\|.$$

Therefore

$$2\delta\|y^*\| \leq 1 + \delta\|y^*\|$$

or

$$\|y^*\| \leq \frac{1}{\delta},$$

proving that T is locally bounded at 0. □

5.6. Lemma. *Let $T\colon A \to 2^{X^*}$ be a maximal monotone operator and $(x_n) \subset A$ be a sequence norm convergent to $x \in A$. Let also (x_n^*) be a sequence in X^* with $x_n^* \in rB^* \cap T(x_n)$ (for some $r > 0$). Then the set consisting of all w* cluster points of (x_n^*) is non-empty and is contained in $T(x) \cap rB^*$.*

Proof. Since rB^* is w* compact, it follows that (x_n^*) has w* cluster points and all are contained in rB^*. Let x^* be such a point. We need to show that $x^* \in T(x)$. Let $y \in A$ and $y^* \in T(y)$. We have

$$(*)\qquad \begin{aligned}\langle x^* - y^*, y - x\rangle &= \langle x^* - x_n^*, y - x\rangle + \langle x_n^* - y^*, y - x\rangle \\ &= \langle x^* - x_n^*, y - x\rangle + \langle x_n^* - y^*, y - x_n\rangle + \langle x_n^* - y^*, x_n - x\rangle.\end{aligned}$$

Since x^* is a w* cluster point of (x_n^*), by passing to a subsequence if necessary, we can assume that

$$\lim\langle x^* - x_n^*, y - x\rangle = 0.$$

Since T is monotone,

$$\langle x_n^* - y^*, y - x_n\rangle \leq 0.$$

Finally, since $(x_n^* - y^*)$ is bounded and (x_n) is norm convergent to x, it follows that

$$\lim\langle x_n^* - y^*, x_n - x\rangle = 0.$$

Returning to $(*)$, we get

$$\langle x^* - y^*, y - x\rangle \leq 0$$

and Lemma 5.3 implies that $x^* \in T(x)$. □

5.7. COROLLARY. *If $T\colon A \to 2^{X^*}$ is a maximal monotone operator and $r > 0$ is such that $A_r \neq \emptyset$, then T_r is w* usco. In particular, T is w* usco at each point where it is locally bounded.*

PROOF. This follows immediately from the above lemma and Lemma 3.3. □

We shall soon see (Corollary 6.3) that locally efficient maximal monotone operators defined on quasi-open subsets are minimal convex w* usco maps and therefore they inherit many of the properties of the minimal convex w* usco maps. However, some of these properties are true for maximal monotone operators defined on not necessarily quasi-open sets, the proofs being more or less similar. For the sake of completeness we shall include all the proofs.

5.8. THEOREM. *Let A be a subset of the Banach space X and let $T\colon A \to 2^{X^*}$ be a locally efficient maximal monotone operator. Then*

(1) there exists a continuous function $\psi\colon A \to (0,\infty)$ such that, for every $x \in A$, T is $\psi(x)$-efficient at x; for any such function ψ, the multivalued map $T_\psi\colon A \to 2^{X^*}$ given by $T_\psi(x) = T(x) \cap \psi(x)B^*$ is a locally bounded, w usco monotone operator.*

(2) the function $\psi_T\colon A \to R$, defined by $\psi_T(x) = \inf\{\|x^*\|;\, x^* \in T(x)\}$, is locally bounded and lower semicontinuous on A;

(3) the multivalued map $T_0\colon A \to 2^{X^*}$ defined by

$$T_0(x) = \{x^* \in T(x) : \|x^*\| = \psi_T(x)\}$$

is a locally bounded monotone operator and $T_0(x)$ is a non-empty, convex, w compact subset of X^* for all $x \in A$;*

(4) T_0 is w usco at each point where ψ_T is continuous.*

PROOF. (1) For each $x \in A$ let ε_x and r_x be positive real numbers such that $T(y) \cap r_x B^* \neq \emptyset$ for every $y \in B(x, \varepsilon_x) \cap A$. Since A is a metric space, it is paracompact (Stone's theorem; see for example [**11**], Theorem 4.3 in Chapter 6). Then there exists an open locally finite refinement $(U_i)_{i\in I}$ of the open cover

$\left(B(x,\varepsilon_x)\cap A\right)_{x\in A}$ of A. Since A is also normal, there exists another open cover $(W_i)_{i\in I}$ of A such that $\mathrm{cl}_A(W_i)\subseteq U_i$ for each $i\in I$. We can now find continuous functions $\varphi_i\colon A\to[0,1]$ such that $\varphi|W_i=1$ and $\varphi|A\setminus U_i=0$ for $i\in I$. For each i pick $x_i\in A$ such that $U_i\subseteq B(x_i,\varepsilon_{x_i})\cap A$ and let $r_i=r_{x_i}$. Define $\psi\colon A\to(0,\infty)$ by $\psi(x)=\sum r_i\varphi_i(x)$. It is easy to see that ψ is continuous on A. If $x\in W_i\subseteq U_i\subseteq B(x_i,\varepsilon_{x_i})\cap A$, then $\psi(x)\geq r_i$ and therefore T is $\psi(x)$-efficient at x. Clearly T_ψ is a locally bounded, monotone operator. We are left to prove that it is w* upper semicontinuous, but this follows immediately from Lemma 3.3 and an argument similar to that used in proving Lemma 5.6.

(2) Since T is locally efficient, ψ_T is locally bounded. To prove the lower semicontinuity of ψ_T, let (x_n) be a sequence in A, norm convergent to x. Let U be a neighborhood of x in A and $r>0$ be such that $T(y)\cap rB^*\neq\emptyset$ for all $y\in U$. Then there exists $n_0>0$ such that $x_n\in U$ if $n>n_0$. For each $n>n_0$ there exists $x_n^*\in T(x_n)$ such that $\psi_T(x_n)\geq\|x_n^*\|-1/n$ and $\|x_n^*\|\leq r$. By Lemma 5.6 the set of w* cluster points of the sequence $(x_n^*)_{n>n_0}$ is non-empty and contained in $x^*\in T(x)\cap rB^*$. Since the norm is w* lower semicontinuous we can find a w* cluster point $x^*\in T(x)\cap rB^*$ of this sequence such that $\liminf\|x_n^*\|\geq\|x^*\|$. Then

$$\liminf\psi_T(x_n)\geq\liminf\left(\|x_n^*\|-1/n\right)\geq\|x^*\|\geq\psi_T(x),$$

showing that ψ_T is lower semicontinuous at x.

(3) Let $x\in A$. The definition of ψ_T implies that for every n there exists $x_n^*\in T(x)$ such that $\psi_T(x)+1/n\geq\|x_n^*\|$. This implies that $x_n^*\in(\psi_T(x)+1)B^*$ and therefore, arguing as above, we can find a w* cluster point $x^*\in T(x)$ of the sequence (x_n^*) such that $\liminf\|x_n^*\|\geq\|x^*\|$. Since

$$\psi_T(x)\geq\liminf\|x_n^*\|\geq\|x^*\|\geq\psi_T(x),$$

it follows that $x^*\in T_0(x)$, showing that $T_0(x)\neq\emptyset$. Clearly $T_0(x)$ is convex and w* compact. Finally, T_0 is locally bounded because ψ_T is.

(4) The arguments we used to prove the upper semicontinuity of T_ψ in (1), can also be used to prove the upper semicontinuity of T_0 at the points where ψ_T is continuous. □

Since a lower semicontinuous function on a Baire space is continuous on a dense G_δ subset of its domain, we obtain

5.9. COROLLARY. *If A is a Baire subset of X and $T\colon A \to 2^{X^*}$ is a locally efficient maximal monotone operator then there exists a dense G_δ subset D of A such that T_0 is w* usco at each point of D.*

5.10. PROPOSITION. *Let A be a Baire subspace of X and $T\colon A \to 2^{X^*}$ be a maximal monotone operator. Then there exists a dense, relatively open subset D of A such that T is efficient at each point of D.*

PROOF. For each positive integer n let

$$F_n = \{x \in A; T(x) \cap nB^* \neq \emptyset\}.$$

Clearly $A = \bigcup F_n$. Observe that each F_n is closed in A. Indeed, let (x_k) be a sequence in F_n norm convergent to $x \in A$. For each k choose $x_k^* \in T(x_k) \cap nB^*$. By Lemma 5.6, the sequence (x_k^*) has a w* cluster point $x^* \in T(x) \cap nB^*$, showing that $x \in F_n$.

Let now denote by G_n the relative interior of F_n in A. Then, since A is Baire, $D = \bigcup G_n$ is dense in A and of course relatively open. Obviously T is efficient at each point of D. □

6. Maximality, Minimality and w* Upper Semicontinuity

6.1. LEMMA. *Let $C \subseteq X$ be convex, $A \subseteq C$ be such that $K_x(A) = K_x(C)$ for every $x \in A$ and $T\colon A \to 2^{X^*}$ be a locally efficient, convex, w* closed valued monotone operator. For every $x \in A$ let $r_x > 0$ be such that T is r_x-efficient at x and assume that T_{r_x} is w* usco at x. Assume also that $T(x) + C_x^* \subseteq T(x)$ for every $x \in A$. Then T is a maximal monotone operator.*

PROOF. Assume T is not a maximal monotone operator. Then there exists $(y, y_0^*) \in A \times X^*$ such that $\langle x^* - y_0^*, x - y \rangle \geq 0$ for all $(x, x^*) \in G(T)$, but $y_0^* \notin T(y)$. Since $T(y)$ is convex and w* closed, by the separation theorem there exist $u \in X$ and $\alpha \in R$ such that

$$\langle y^*, u \rangle < \alpha < \langle y_0^*, u \rangle, \text{ for all } y^* \in T(y).$$

First let us prove that $u \in K_y(A) = K_y(C) = \mathrm{cl}(C_y)$. Assume not. Then there exists $z^* \in X^*$ such that

$$\langle z^*, x \rangle \leq 0 < \langle z^*, u \rangle, \text{ for all } x \in C_y.$$

The left inequality implies that $z^* \in C_y^*$ and therefore $nz^* \in C_y^*$ for any positive n. Choose $y^* \in T(y)$. Then $y^* + nz^* \in T(y)$ and therefore $\langle y^* + nz^*, u \rangle < \alpha$ for any positive n, which implies that $\langle z^*, u \rangle \leq 0$, contradicting the choice of z^*. Thus $u \in K_y(A)$.

Next let $r = r_y$ and choose $\beta \in R$ such that $\alpha < \beta < \langle y_0^*, u \rangle$. Let $W = \{z^* \in X^*;\ \langle z^*, u \rangle < \alpha\}$; since W is w* open and contains $T(y)$ and T_r is w* usco at y, there exists a neighborhood U of y in X such that $T_r(z) \subseteq W$ for any $z \in U \cap A$. We can also assume that $T_r(z) \neq \emptyset$ for all $z \in U \cap A$ (T is r-efficient at y). Since $u \in K_y(A)$, there exist $t > 0$ and $v \in V = \{z \in X; \langle y_0^*, z \rangle > \beta\}$ such that $\|u - v\| < (\beta - \alpha)/r$ and $y + tv \in A \cap U$. Then $T_r(y + tv) \subseteq W$. For any $u^* \in T_r(y + tv) \subseteq W$ we have

$$0 \leq \langle y_0^* - u^*, y - y - tv \rangle = -t\langle y_0^* - u^*, v \rangle$$

and thus $\langle y_0^*, v\rangle \leq \langle u^*, v\rangle$. This implies that

$$\langle u^*, u\rangle = \langle u^*, v\rangle + \langle u^*, u - v\rangle \geq \langle y_0^*, v\rangle - \|u^*\| \cdot \|u - v\| \geq \beta - r(\beta - \alpha)/r = \alpha$$

and thus $u^* \notin W$, contradicting one of our previous choices. The lemma is therefore proved. □

Our first main result, stated below, provides a characterization of the maximal monotone operators among the locally efficient monotone operators on convex sets: a locally efficient monotone operator on a convex set C is maximal monotone if and only if it is convex and w* closed valued, w* upper semicontinuous in a certain sense and, for every $x \in C$, $T(x)$ is invariant under translations by elements in the normal cone to C at x. More precisely

6.2. THEOREM. *Let $C \subseteq X$ be convex, $A \subseteq C$ be quasi-open in C and let $T\colon A \to 2^{X^*}$ be a locally efficient monotone operator. For every $x \in A$ choose $r_x > 0$ such that T is r_x-efficient at x. The following assertions are equivalent:*

(1) *T is a maximal monotone operator on A;*

(2) *T is convex and w* closed valued, T_r is w* usco for every $r > 0$ for which $A_r \neq \emptyset$ and $T(x) + C_x^* \subseteq T(x)$ for every $x \in A$;*

(3) *T is convex and w* closed valued, T_{r_x} is w* usco at x and $T(x) + C_x^* \subseteq T(x)$ for every $x \in A$.*

PROOF. Assume that (1) is true. Then, by Lemma 5.3 (2), T is convex and w* closed valued. By Corollary 5.7, T_r is w* usco for every $r > 0$ for which $A_r \neq \emptyset$. The last assertion follows from the definitions. Obviously (2) implies (3). Finally, Lemma 6.1 shows that (3) implies (1). □

6.3. COROLLARY. *Let $A \subseteq X$ be quasi-open and $T\colon A \to 2^{X^*}$ be a locally efficient monotone operator. The following assertions are equivalent:*

(a) *T is maximal monotone;*

(b) *T is a minimal convex* w* *usco map;*

(c) *T is a convex* w* *usco map.*

PROOF. Clearly (b) implies (c). Assume (c). Since $X_x^* = \{0\}$ for each $x \in X$, condition (3) in the above theorem is satisfied (with $C = X$). It follows that T is maximal monotone. Thus (c) implies (a). Finally assume that (a) is true. By Lemma 5.4 and Corollary 5.7, T is a convex w* usco map. Let $S: A \to 2^{X^*}$ be a convex w* usco map such that $S \preceq T$. Since (c) implies (a), S must be maximal. Since T itself is maximal, $S = T$ and therefore T is a minimal convex w* usco map. □

The following result shows that monotone operators on certain domains have unique maximal monotone extensions (over the same domain).

6.4. COROLLARY. *Let $C \subseteq X$ be convex, $A \subseteq C$ be quasi-open in C and let $T: A \to 2^{X^*}$ be a locally efficient monotone operator. Then there exists a unique maximal monotone operator $M: A \to 2^{X^*}$ such that $T(x) \subseteq M(x)$ for every $x \in A$. If A is quasi-open on X then M can be described as follows: let $\overline{T}$ be the monotone operator on A whose graph is the closure (in $A \times X^*$) of the graph of T; then $M(x)$ is the* w* *closed convex hull of $\overline{T}(x)$.*

PROOF. By standard arguments there exists a maximal monotone extension M of T. To prove that M is unique, assume that there exists another maximal monotone operator S on A which contains T. Clearly $M \cap S: A \to 2^{X^*}$, $(M \cap S)(x) = M(x) \cap S(x)$, is a convex, w* closed valued monotone operator on A. It also satisfies $(M \cap S)(x) + C_x^* \subseteq (M \cap S)(x)$ for every $x \in A$ (since we can apply Theorem 2.2 to M and S). Let $r > 0$ be such that $(M \cap S)_r$ is defined. Since $(M \cap S)_r$ has a closed graph and is contained in M_r, which is w* usco (by Theorem 2.2), we can apply Lemma 3.5 to deduce that $(M \cap S)_r$ is w* usco. Finally, Theorem 2.2 implies that $(M \cap S)$ is a maximal monotone operator on A and therefore $M = M \cap S = S$. This proves the uniqueness of M. The description of M given in the last assertion of the corollary can be proved as in [14], Theorem 7.13. □

The next result may also be useful.

6.5. Corollary. *Let $T\colon A \to 2^{X^*}$ be a locally efficient maximal monotone operator and $A_0 \subseteq A$ be such that $K_x(A_0) = K_x(A)$ for every $x \in A_0$. Assume that A is either convex or quasi-open. Then the restriction of T to A_0 is maximal monotone (on A_0).*

6.6. Theorem. *Let A be a quasi-open, Baire subset of the Banach space X and let $T\colon A \to 2^{X^*}$ be a maximal monotone operator. Then there exist a dense, relatively open subset D_1 of A and a dense G_δ subset D of A such that $D \subseteq D_1$ and*

(1) *$T|D_1$ is a locally bounded, minimal convex w* usco map;*

(2) *for each $x \in D$, $T(x)$ is contained in a sphere of X^* and T is norm-to-w* upper semicontinuous at x.*

Proof. From Proposition 5.10, there exists a dense, relatively open subset D_1 of A such that T is efficient at each point of D_1. By Lemma 5.4, T is locally bounded at each point of D_1. By Corollary 6.5 and Corollary 6.3, $T|D_1$ is a minimal convex w* usco map. This proves (1). The second assertion follows immediately from Theorem 4.9. $\square$

6.7. Remark. Let A be as above and $T\colon A \to 2^{X^*}$ be a monotone operator, not necessarily maximal. Then there exists a dense G_δ subset D of A such that $T(x)$ is contained in a sphere of X^* for each $x \in D$. This follows immediately from the previous theorem and the fact that a monotone operator is always contained in a maximal one.

6.8. Corollary. *Let A be a quasi-open Baire subset of the Banach space X and let $T\colon A \to 2^{X^*}$ be a monotone operator. Assume that X^* has the Radon-Nikodým property (resp. that X is in Stegall's class $\mathbf{C}$). Then there exists a dense G_δ subset A_0 of A such that T is single valued and norm-to-norm (resp. norm-to-w*) upper semicontinuous at each point of A_0.*

PROOF. There is no loss of generality if we assume that T is maximal monotone. Since A is quasi-open, Theorem 6.6 (1) implies that there exists a dense, relatively open subset D of A such that $T|D$ is a locally bounded, minimal convex w* usco map. Our assertions follow from Lemma 4.7 and the remark in Section 4.3. □

Comments

Most of the results in this chapter were known for monotone operators on open sets (see the exposition in [**14**]). We were able to extend them to the case of quasi-open sets, the essential tool being the notion of local efficiency. Theorem 6.2 is new even in the context of open sets. The fundamental result expressed in Lemma 5.5 is due to Rockafellar [**18**] (who proved more in a less general setting); the proof presented here is due to Borwein and Fitzpatrick [**3**].

Chapter IV

SUBDIFFERENTIABILITY AND DIFFERENTIABILITY

7. Convex Functions

7.1. In this chapter we shall denote by X a real Banach space. Let A be a subset of X and let $f\colon A \to R$ be a real function. The *epigraph* of f, denoted $\operatorname{epi}(f)$, is defined as follows

$$\operatorname{epi}(f) = \{(a,r) \in X \times R : a \in A,\ r \geq f(a)\}.$$

It is well known that f is lower semicontinuous on A if and only if $\operatorname{epi}(f)$ is a relatively closed subset of $A \times R$.

Let C be a convex subset of X. A function $f\colon C \to R$ is called *convex* if

$$f((1-t)x + ty) \leq (1-t)f(x) + tf(y), \text{ for all } x,\ y \in C,\ t \in [0,1].$$

Let us observe that f is convex if and only if its epigraph is a convex subset of $X \times R$.

Here is a list of well known properties of convex functions:

(1) Any linear (or, slightly more general, any affine) function $f\colon X \to R$ is convex.

(2) Any subadditive and positively homogeneous function $f\colon X \to R$ is convex (in particular the norm is a convex function on X).

(3) The set of all convex functions defined on a convex set is a convex cone (i.e., the sum of two convex functions is a convex function and any positive multiple of a convex function is a convex function).

(4) Let $f_i\colon C \to R$ be a family of convex functions and let $C_0 = \{x \in C : \sup_i f_i(x) < \infty\}$. Then C_0, if not empty, is a convex subset of X and the pointwise supremum $\sup_i f_i$ is a convex function on C_0.

(5) Assume that C is an open convex subset of X and that $f\colon C \to R$ is continuously differentiable and has a second derivative f'' throughout C. Then f is convex if and only if $f''(x)$ is nonnegative definite for each $x \in C$.

Like linear functions, convex functions defined on open convex sets need very little to be locally Lipschitz. Namely, it is sufficient for them to be bounded from above in an open subset of their domain.

7.2. PROPOSITION. *Let D be an open convex subset of X and let $f\colon D \to R$ be convex. Then the following assertions are equivalent*

(1) *there exists $x_0 \in D$ such that f is bounded from above in a neighborhood of x_0;*

(2) *f is locally Lipschitz on D.*

PROOF. Clearly (2) implies (1). To prove that (1) implies (2), let M, $\delta > 0$ be such that $B(x_0, 2\delta) \subseteq D$ and $f(x) \leq M$ for all $x \in B(x_0, 2\delta)$. Let $x \in B(x_0, 2\delta)$. Then $2x_0 - x \in B(x_0, 2\delta)$ and (since $x_0 = \frac{1}{2}x + \frac{1}{2}(2x_0 - x)$)

$$f(x_0) \leq \frac{1}{2}f(x) + \frac{1}{2}f(2x_0 - x) \leq \frac{1}{2}f(x) + \frac{1}{2}M,$$

hence

$$2f(x_0) - M \leq f(x).$$

If we set $N = \max\{M, |2f(x_0) - M|\}$ we obtain

$$|f(x)| \leq N \text{ for all } x \in B(x_0, 2\delta).$$

Let $x, y \in B(x_0, \delta)$, $x \neq y$, $\alpha = \|x - y\|$, $\gamma = (\alpha + \delta)/\alpha$ and $z = (1-\gamma)x + \gamma y$. Notice that

$$\|z - y\| = \|(1-\gamma)x - (1-\gamma)y\| = (\gamma - 1)\|x - y\| = \delta$$

and therefore $z \in B(x_0, 2\delta)$. Since $y = (1 - \frac{1}{\gamma})x + \frac{1}{\gamma}z$ and since f is convex we have

$$f(y) \leq (1 - \frac{1}{\gamma})f(x) + \frac{1}{\gamma}f(z)$$

or

$$f(y) - f(x) \leq \frac{1}{\gamma}(f(z) - f(x)) \leq \frac{\alpha}{\alpha + \delta}2N \leq \frac{2N}{\delta}\|y - x\|.$$

Interchanging x with y we obtain that f is Lipschitz on $B(x_0, \delta)$ with Lipschitz constant $2N/\delta$.

It remains to show that f is bounded from above in the neighborhood of each point of D. So let $x \in D$ and choose $\lambda > 1$ such that $z = x_0 + \lambda(x - x_0) \in D$ (this is possible since D is open). Let $\mu = 1/\lambda$. Clearly

$$B(x, (1-\mu)\delta) = \mu z + (1-\mu)B(x_0, \delta) \subseteq D$$

and for any $y = (1-\mu)v + \mu z \in B(x, (1-\mu)\delta)$ with $v \in B(x_0, \delta)$ we have

$$f(y) \leq (1-\mu)f(v) + \mu f(z) \leq M + f(z).$$

This completes the proof of the proposition. □

7.3. It is natural to ask if the above result remains true for a convex function defined on a not open, but quasi-open set. In general the answer is negative.

EXAMPLE. Let $X = l^2(N)$ and $C = \{\mathbf{x} = (x_n) \in X : |x_n| < 2^{-n} \text{ for all } n\}$. Notice that if $\mathbf{x} \in C$ then $\mathbf{x} \pm t\mathbf{e}_n \in C$ for any n and a sufficiently small $t > 0$. It follows that $\mathbf{e}_n$ and $-\mathbf{e}_n \in C_{\mathbf{x}}$ and therefore $\mathrm{cl}(C_{\mathbf{x}}) = X$. This shows that C is quasi-open.

Let $f\colon C \to R$ be defined by

$$f(\mathbf{x}) = \sum_{n=1}^{\infty} -(2^{-n} + x_n)^{1/2}.$$

Since each summand is continuous, convex, and bounded in absolute value by $2^{(1-n)/2}$, the series is uniformly convergent and thus f is continuous and convex. On the other hand, it is not difficult to check that f is not locally Lipschitz at any $\mathbf{x}$ in C (in view of the previous proposition this shows that the interior of C is empty). As a matter of fact, we shall prove later a stronger result, namely that f is not subdifferentiable at any point of C (see Example 8.6).

8. Subdifferentiability of Convex Functions

8.1. DEFINITIONS. Let C be a convex subset of X, $f\colon C \to R$ be a convex function and $x \in C$.

(1) A functional $x^* \in X^*$ is called a *subgradient* of f at x if

$$\langle x^*, y - x\rangle \leq f(y) - f(x) \qquad \text{for all } y \in C.$$

We shall denote by $\partial f(x)$ the (possibly empty) set of all subgradients of f at x. We shall call $\partial f(x)$ the *subdifferential* of f at x.

(2) The function f is called *subdifferentiable at* x if its subdifferential at x is not empty, i.e., there exists at least one subgradient of f at x.

(3) The function f is called *subdifferentiable on* C if it is subdifferentiable at each point of C.

8.2. LEMMA. *Let $f\colon C \to R$ be a convex function, let $A = \{x \in C : \partial f(x) \neq \emptyset\}$ and let $x \in A$. Then*

(1) *f is lower semicontinuous at x.*

(2) *$\partial f(x)$ is a convex, w* closed subset of X^* and $\partial f(x) + C_x^* \subseteq \partial f(x)$.*

(3) *$\partial f\colon A \to 2^{X^*}$ is a monotone operator.*

PROOF. (1) Let (x_n) be a sequence in C converging to x and let $x^* \in \partial f(x)$. Then $\langle x^*, x_n - x\rangle \leq f(x_n) - f(x)$, which implies that $\liminf f(x_n) \geq f(x)$, proving the lower semicontinuity of f at x.

(2) All three assertions follow immediately from the definitions.

(3) Let x, $y \in A$, $x^* \in \partial f(x)$ and $y^* \in \partial f(y)$. Then

$$\langle y^* - x^*, x - y\rangle = \langle y^*, x - y\rangle + \langle x^*, y - x\rangle \leq f(x) - f(y) + f(y) - f(x) = 0,$$

which proves that ∂f is a monotone operator on A. □

Exactly as in the case of differentiable functions, there exists a close relation between the subdifferentiability of f at x and "tangential" properties of the graph of f at $(x, f(x))$. Before explaining this, we need to introduce some more notation.

Given $\Psi \in (X \times R)^*$, we shall denote by r_Ψ the real number $\Psi(0,1)$ and by $\psi \in X^*$ the functional $\psi(x) = \Psi(x, 0)$. We shall say that Ψ is *non-vertical* if $r_\Psi \neq 0$; otherwise we shall say that Ψ is *vertical*.

8.3. LEMMA. *Let C be a convex subset of X, $x \in C$ and $f\colon C \to R$ be a convex function.*

(1) *The assignment $\Psi \mapsto \psi$ determines a one-to-one correspondence between the set of all (non-vertical) functionals $\Psi \in (X \times R)^*$ with $r_\Psi = -1$ and which support* epi(f) *at $(x, f(x))$, and the subset $\partial f(x)$ of X^*.*

(2) *If Ψ is a non-zero vertical support functional to* epi(f) *at $(x, f(x))$, then $x \in S(C)$.*

PROOF. (1) Let $\Psi \in (X \times R)^*$ be such that $r_\Psi = -1$ and let $(y, t) \in \mathrm{epi}(f)$. We have $\langle \Psi, (y,t)\rangle = \langle \psi, y\rangle - t$ and therefore

$$\langle \Psi, (y,t)\rangle \leq \langle \Psi, (x, f(x))\rangle \iff \langle \psi, y - x\rangle \leq t - f(x),$$

which shows that Ψ is a support functional to $\mathrm{epi}(f)$ at $(x, f(x))$ if and only if ψ is a subgradient of f at x.

(2) Let Ψ be a non-zero vertical support functional to $\mathrm{epi}(f)$ at $(x, f(x))$ and let $y \in C$. Then $\langle \Psi, (y, f(y))\rangle \leq \langle \Psi, (x, f(x))\rangle$, hence $\langle \psi, y - x\rangle \leq 0$, showing that $x \in \mathrm{S}(C)$. □

8.4. Let $f\colon C \to R$ be convex and let $x \in C$. For every $v \in C_x$ there exists $\delta > 0$ such that $x + tv \in C$ for any $t \in [0, \delta]$. Define a function $d_{x,v}\colon (0,\delta) \to R$ by

$$d_{x,v}(t) = \frac{f(x+tv) - f(x)}{t}.$$

Since f is convex, $d_{x,v}$ is an increasing function of t and therefore

$$-\infty \le \lim_{t \downarrow 0} d_{x,v}(t) = \inf_{t>0} d_{x,v}(t) < \infty.$$

We shall denote the above limit by $f'_x(v)$. It is easy to see that $f'_x\colon C_x \to [-\infty, +\infty)$ is positively homogeneous. Moreover, since f is convex, f'_x is also subadditive (this follows from the inequality $d_{x,v+w}(t) \le d_{x,v}(2t) + d_{x,w}(2t)$).

8.5. Proposition. *Let $f\colon C \to R$ be convex, let $x \in C$ and let $x^* \in X^*$. The following assertions are equivalent:*

(1) $x^* \in \partial f(x)$.

(2) $\langle x^*, v\rangle \le f'_x(v)$ *for all* $v \in C_x$.

(3) f'_x *is finite valued and* $x^* \in \partial f'_x(0)$.

Proof. In order to prove that (1) implies (2), let $v \in C_x$. Then there exists $\delta > 0$ such that $x + tv \in C$ for every $t \in (0,\delta)$. Since $x^* \in \partial f(x)$, we have

$$\langle x^*, tv\rangle = \langle x^*, x + tv - x\rangle \le f(x+tv) - f(x), \quad \text{for } 0 < t < \delta.$$

This implies that $\langle x^*, v\rangle \le d_{x,v}(t)$, which proves (2).

Assume now that (2) is true. Then f'_x is finite valued. As noted above, f'_x is positively homogeneous and subadditive, hence convex. It makes sense therefore to consider $\partial f'_x(0)$. Since $f'_x(0) = 0$, the condition in (2) implies that $x^* \in \partial f'_x(0)$.

Finally assume that (3) is true and let $y \in C$. Then $y - x \in C_x$ and

$$\langle x^*, y - x\rangle \le f'_x(y-x).$$

On the other hand, using first the definition of f'_x and then the convexity of f, we have

$$f'_x(y-x) \le \frac{f(x + t(y-x)) - f(x)}{t} \le f(y) - f(x)$$

(the middle term exists for some $t > 0$). Combining the last two formulas, we get $\langle x^*, y - x\rangle \le f(y) - f(x)$, which shows that $x^* \in \partial f(x)$. □

8.6. Let $f: C \to R$ be convex and let $x \in C$. We have seen (Lemma 8.2) that a necessary condition for the subdifferentiability of f at x is that f be lower semi-continuous at x. If x is an interior point of C, this last condition is equivalent to f being locally Lipschitz on the interior of C (see Proposition 7.2). On the other hand, as the next example will show, if x is not an interior point of C it may happen that even a continuous f is nowhere subdifferentiable on C. It is therefore natural to investigate the subdifferentiability of locally Lipschitz functions. First, however, the example mentioned above.

EXAMPLE. Let $f: C \to R$ be as in Example 7.3. Let $\mathbf{x} \in C$ and $n \geq 1$. Since C is quasi-open, there exists $s > 0$ such that the segment $[\mathbf{x}, \mathbf{x} + s\mathbf{e}_n]$ is contained in C. Then

$$df_{\mathbf{x},\mathbf{e}_n}(t) = \frac{-\sqrt{2^{-n} + x_n + t} + \sqrt{2^{-n} + x_n}}{t} = \frac{-1}{\sqrt{2^{-n} + x_n + t} + \sqrt{2^{-n} + x_n}}$$

for any $t \in (0, s)$ and therefore

$$f'_{\mathbf{x}}(\mathbf{e}_n) = \frac{-1}{2\sqrt{2^{-n} + x_n}}.$$

Assume now that there exists $x^* \in \partial f(\mathbf{x})$. From Proposition 8.5 (2) and the above equality we get that

$$-\|x^*\| \leq \frac{-1}{2\sqrt{2^{-n} + x_n}}, \qquad \text{for any } n \geq 1,$$

which is impossible. It follows that f is nowhere subdifferentiable.

8.7. THEOREM. *Let C be a convex subset of X, A be a relatively open subset of C and $f: C \to R$ be a convex function whose restriction to A is locally Lipschitz. We have*

(1) *For every $x \in A$, $\partial f(x)$ is a non-empty, convex w* closed subset of X^*; if in addition x is a quasi-interior point of A, then $\partial f(x)$ is also w* compact.*

(2) *The restriction of ∂f to A is a locally efficient monotone operator, locally bounded at each quasi-interior point of A.*

(3) *Let $r > 0$ be such that $A_r = \{x \in A : \partial f(x) \cap rB^* \neq \emptyset\}$ is non-empty and let $(\partial f)_r \colon A_r \to 2^{X^*}$ be given by $(\partial f)_r(x) = \partial f(x) \cap rB^*$. Then A_r is closed in A and $(\partial f)_r$ is w* usco.*

PROOF. Let $x_0 \in A$, let ε, $M > 0$ be such that f is Lipschitz on $B(x_0, \varepsilon) \cap C$ with Lipschitz constant M and let $x \in B(x_0, \varepsilon) \cap C$. With the notation introduced before, we have

$$|d_{x,v}| \leq M\|v\|, \qquad v \in C_x$$

hence

$$(*) \qquad |f'_x(v)| \leq M\|v\| \qquad \text{for every } v \in C_x$$

and therefore f'_x is finite valued. Next, we need to show that f'_x is Lipschitz on C_x. To this end, let $v, w \in C_x$ and $t > 0$ be such that $x + tv$, $x + tw \in B(x_0, \varepsilon) \cap C$. Then

$$\left| \frac{f(x+tv) - f(x)}{t} - \frac{f(x+tw) - f(x)}{t} \right| = \left| \frac{f(x+tv) - f(x+tw)}{t} \right| \leq M\|v - w\|$$

and, after passing to limits,

$$|f'_x(v) - f'_x(w)| \leq M\|v - w\|.$$

Being Lipschitz, f'_x can be extended to a Lipschitz function $F_x \colon K_x(C) = \mathrm{cl}(C_x) \to R$, with the same Lipschitz constant M. Since f'_x is subadditive and positively homogeneous, it is easy to see that F_x has these properties too. It follows that $\mathrm{epi}(F_x)$ is a closed convex cone in $X \times R$. Consider the open convex cone $D = \{(v, r) \in X \times R \colon -M\|v\| > r\}$. Since $\mathrm{epi}(F_x)$ and D are disjoint, by the separation theorem there exists $\Phi \in (X \times R)^*$ such that

$$\Phi(y, s) \leq 0 \leq \Phi(z, t), \quad (y, s) \in \mathrm{epi}(F_x), \ (z, t) \in D.$$

Since $(0, -1) \in D$, $r = \Phi(0, -1) > 0$. Let $\Psi = (1/r)\Phi$. Clearly Φ supports $\mathrm{epi}(F_x)$ at $(0, 0)$ and the same is true for Ψ. Since $\Psi(0, 1) = -1$, it follows that we can apply Lemma 8.3 (1) to Ψ and, with the notation introduced there, $\psi \in \partial F_x(0)$ (recall that $\psi(y) = \Psi(y, 0)$). Since clearly $\partial F_x(0) = \partial f'_x(0)$, from Proposition 8.5 we obtain that $\psi \in \partial f(x)$. Thus $\partial f(x)$ is non-empty and, by Lemma 8.2, it is convex and w* closed.

Let us also notice that $\|\psi\| \leq M$. Indeed, for any $v \in X$ and any $\varepsilon > 0$, $(v, -M\|v\| - \varepsilon)$ and $(-v, -M\|v\| - \varepsilon)$ are contained in D. Then $\Psi(v, -M\|v\| - \varepsilon) \leq 0$ and $\Psi(-v, -M\|v\| - \varepsilon) \leq 0$, which imply that $|\psi(v)| \leq M\|v\| + \varepsilon$. Thus, $\|\psi\| \leq M$. This proves that ∂f is efficient at x_0.

Assume now that $x \in \mathrm{qi}(C)$. Then for any $x^* \in \partial f(x)$ and $v \in C_x$, from Proposition 8.5 and $(*)$ we get

$$\langle x^*, v \rangle \leq f'_x(v) \leq M\|v\|.$$

Since $\mathrm{cl}(C_x) = X$, the part of the above inequality involving the left and right sides is true for any $v \in X$ and therefore $\partial f(x) \subseteq MB^*$. Assertions (1) and (2) are now completely proved.

To prove (3), let (x_n) be a sequence in A_r norm convergent to $x \in A$ and for every n let $x_n^* \in \partial f(x_n) \cap rB^*$. Since rB^* is w* compact, the sequence (x_n^*) has at least one w* cluster point and all its w* cluster points are contained in rB^*. Let x^* be such a cluster point and let $y \in C$. We have

$$\begin{aligned} \langle x^*, y - x \rangle &= \langle x^* - x_n^*, y - x \rangle + \langle x_n^*, y - x_n \rangle + \langle x_n^*, x_n - x \rangle \\ &\leq \langle x^* - x_n^*, y - x \rangle + f(y) - f(x_n) + \|x_n^*\| \cdot \|x_n - x\| \\ &\leq \langle x^* - x_n^*, y - x \rangle + f(y) - f(x_n) + r\|x_n - x\|. \end{aligned}$$

By passing to a subsequence, we can assume that $\lim \langle x^* - x_n^*, y - x \rangle = 0$ and therefore the above inequalities imply that $\langle x^*, y - x \rangle \leq f(y) - f(x)$, which means that $x^* \in \partial f(x)$. We have already seen that $x^* \in rB^*$. Thus $x \in A_r$ (proving that A_r is closed in A) and we can use Lemma 3.3 to conclude that $(\partial f)_r$ is w* usco. □

REMARKS. (1) Assertion (3) is true for any subset A of C, without any assumption on f.

(2) In the above theorem the condition imposed on A to be relatively open is essential. The following example (due to Borwein and Fitzpatrick) will show that even if A is a dense G_δ subset of C the theorem does not hold. In particular, it will show that not all assumptions of the Proposition in [**13**] are true.

Let $X = l^2(N)$ and C be the positive cone in X (see Remark 2.7 (1)). Define $f\colon C \to R$ by $f(x) = \sup\{e^{-nx_n} : n \in N\}$. It is easy to see that f is a convex lower semicontinuous function on C and that $0 < f(x) \leq 1$ for any $x \in C$. Clearly

$$A = f^{-1}(1) = \bigcap_n \{x \in C : f(x) > 1 - \frac{1}{n}\}$$

is a G_δ subset of C (since f is lower semicontinuous). Since $S(C)$ is dense in C and $S(C)$ is contained in A, it follows that A is dense in C. Thus A is a dense G_δ subset of C and $f\,|\,A$ is Lipschitz (being constant on A!). However, $\partial f(x) = \emptyset$ for all points x in the dense G_δ subset $A \cap \mathrm{qi}(C)$, i.e., if $x \in A$ and $\partial f(x) \neq \emptyset$ then $x \in \mathrm{S}(C)$. Indeed, if $f(x) = 1$ and $x^* \in \partial f(x)$, then

$$\langle x^*, y - x\rangle \leq f(y) - f(x) \leq 0 \qquad \text{for any } y \in C,$$

hence $x^* \in C_x^*$. Since there are points $y \in C$ such that $f(y) < 1$, x^* cannot be zero and thus $x \in \mathrm{S}(C)$.

8.8. COROLLARY. *Under the same assumptions as in the previous theorem, the restriction of ∂f to A is a maximal monotone operator on A.*

PROOF. The previous theorem and Lemma 8.2 show that condition (2) in Theorem 6.2 is satisfied, hence $\partial f\,|\,A$ is a maximal monotone operator on A. □

8.9. COROLLARY. *Let C be a quasi-open convex subset of X, and A be a relatively open, Baire subset of C (e.g., $A = C$ is the quasi-interior of a closed convex subset of X). Let $f\colon C \to R$ be a convex function whose restriction to A is locally Lipschitz. Then $\partial f\,|\,A$ is a minimal convex w^* usco map and there exists a dense G_δ subset A_0 of A such that $\partial f(x)$ is contained in a sphere of X^* for each $x \in A_0$.*

PROOF. By the previous corollary, $\partial f\,|\,A$ is a maximal monotone operator on A. The assertion follows now from Corollary 6.3 and Theorem 4.9. □

We have proved so far that locally Lipschitz convex functions are subdifferentiable. We may now ask the following question: is a subdifferentiable convex function locally Lipschitz? If its domain is open, the answer is yes: being subdifferentiable, f is lower semicontinuous (see Lemma 8.2), hence it is locally Lipschitz (see Proposition 7.2; as a matter of fact we need only the subdifferentiabilty of f at one point!). We cannot prove such a strong result in the general case, but we shall come close enough.

8.10. PROPOSITION. *Let C be a convex subset of X, A be a subset of C and $f\colon C \to R$ be convex. Assume that $\partial f(x) \neq \emptyset$ for every $x \in A$ and that $\partial f\,|\,A$ is locally efficient. Then $f\,|\,A$ is locally Lipschitz. If in addition f is lower semicontinuous on $\mathrm{cl}_C(A)$, the closure of A in C, then $f\,|\,\mathrm{cl}_C(A)$ is locally Lipschitz at each $x \in A$.*

PROOF. Let $x \in A$ and let ε, $r > 0$ be such that $\partial(f)$ is r-efficient on $B(x,\varepsilon) \cap A$. Let $y, z \in B(x,\varepsilon) \cap A$ and choose $y^* \in \partial f(y) \cap rB^*$ and $z^* \in \partial f(z) \cap rB^*$. We have

$$-r\|y-z\| \leq \langle y^*, z-y\rangle \leq f(z)-f(y) \leq \langle z^*, z-y\rangle \leq r\|y-z\|$$

proving that f is locally Lipschitz on A.

Assume next that f is lower semicontinuous on $\mathrm{cl}_C(A)$ and let $x \in A$. By the first part, there exist δ, $r > 0$ such that f is Lipschitz on $B(x,\delta) \cap A$, with Lipschitz constant r. Let y, $z \in B(x,\delta) \cap \mathrm{cl}_C(A)$ and let $\varepsilon > 0$. Since f is lower semicontinuous at z, there exists ρ, $0 < \rho < \varepsilon$, such that

$$f(z') > f(z) - \varepsilon, \qquad \text{for all } z' \in B(z,\rho) \cap \mathrm{cl}_C(A).$$

Pick $y' \in B(y,\rho) \cap B(x,\delta) \cap A$, $z' \in B(z,\rho) \cap B(x,\delta) \cap A$ and $y^* \in \partial f(y') \cap rB^*$. Then

$$\begin{aligned} f(z)-f(y) &= f(z)-f(z')+f(z')-f(y')+f(y')-f(y) \\ &\leq \varepsilon + r\|z'-y'\| + \langle y^*, y'-y\rangle \\ &\leq \varepsilon + r(2\rho + \|z-y\|) + r\rho \leq r\|z-y\| + \varepsilon(1+3r) \end{aligned}$$

Since ε is arbitrary, we obtain that $f(z) - f(y) \leq r\|z - y\|$. Interchanging y and z, we finally obtain that

$$|f(z) - f(y)| \leq r\|z - y\|, \qquad \text{for any} \;\; y, z \in B(x, \delta) \cap \mathrm{cl}_C(A),$$

and the second assertion is also proved. □

8.11. THEOREM. *Let C be a convex subset of X and let $f\colon C \to R$ be convex. Assume that $\partial f(x) \neq \emptyset$ for every point x of a Baire subset A of C. Then there exists a dense, relatively open subset D of A such that the restriction of f to A is locally Lipschitz at each point of D. If in addition f is lower semicontinuous on $\mathrm{cl}_C(A)$, then the restriction of f to $\mathrm{cl}_C(A)$ is locally Lipschitz at each point of D.*

PROOF. For every positive integer n define

$$A_n = \{x \in A : \partial f(x) \cap nB^* \neq \emptyset\}.$$

From Theorem 8.7(3) (see also the remark following it), we deduce that every A_n is closed in A. Clearly $A = \bigcup A_n$ and, since A is Baire, $D = \bigcup \mathrm{int}_A(A_n)$ is dense and relatively open in A. Since ∂f is n-efficient on A_n, it is locally efficient on D and the theorem follows from the preceding proposition. □

9. Gâteaux and Fréchet Differentiability

9.1. DEFINITIONS. (1) Let X and Y be Banach spaces and A be a subset of X. A map $f\colon A \to Y$ is called *Gâteaux differentiable* at $x \in A$ if there exists a continuous linear map $df_x\colon X \to Y$, called a *Gâteaux differential* of f at x, such that $df_x(v) = f'_x(v)$ for all $v \in T_x(A)$ (see Section 1.2 for the definition of $T_x(A)$).

(2) The map $f\colon A \to Y$ is called *Fréchet differentiable* at $x \in A$ if there exists a continuous linear map $Df_x\colon X \to Y$, called a *Fréchet differential* of f at x, such that the function $\mathcal{O}_{f,x}\colon A \to Y$ defined by

$$\mathcal{O}_{f,x}(y) = \begin{cases} \dfrac{f(y) - f(x) - Df_x(y - x)}{\|y - x\|}, & y \neq x \\ 0, & y = x \end{cases}$$

is norm-to-norm continuous at x.

REMARKS. (1) The definition of Fréchet differentiability given above is equivalent to the usual one, which requires that $\lim_{y\to x} \dfrac{\|f(y)-f(x)-Df_x(y-x)\|}{\|y-x\|} = 0$.

(2) It is also clear that if f is Fréchet differentiable at x, then it is Gâteaux differentiable at x and any Fréchet differential of f at x is also a Gâteaux differential of f at x.

(3) The Gâteaux (respectively Fréchet) differential, if it exists, is not unique in general. However, if $\operatorname{cl}\operatorname{aff}(T_x(A)) = X$ then it is unique. This happens for example when A is convex and $\operatorname{cl}\operatorname{aff}(A) = X$.

9.2. Let $f\colon C \to R$ be convex and assume that f is Gâteaux differentiable at $x \in A$. From Proposition 8.5 it follows immediately that $df_x \in \partial f(x)$. The same proposition implies that $x^* \in \partial f(x)$ if and only if $\langle x^* - df_x, v\rangle \le 0$ for any $v \in C_x$, i.e., $x^* \in \partial f(x)$ if and only if $x^* - df_x \in C_x^*$. It follows that $\partial f(x) = \{df_x\}$ if $x \in \operatorname{qi}(C)$.

Conversely, if $x \in \operatorname{qi}(C)$, f is Lipschitz in a neighborhood of x and $\partial f(x)$ is a singleton, then f is Gâteaux differentiable at x. Indeed, let $\partial f(x) = \{x^*\}$. It is enough to prove that $\langle x^*, v\rangle = F_x(v)$ for all $v \in X$, where $F_x\colon X = \operatorname{cl}(C_x) \to R$ is the continuous sublinear map which extends f'_x (see the proof of Theorem 8.7). Let $v \in X$; by the Hahn-Banach theorem there exists $y^* \in X^*$ such that $\langle y^*, v\rangle = F_x(v)$ and $\langle y^*, u\rangle \le F_x(u)$ for any $u \in X$. From Proposition 8.5 we deduce that $y^* \in \partial f(x)$ and therefore $y^* = x^*$, hence $\langle x^*, v\rangle = F_x(v)$, which proves that x^* is the Gâteaux differential of f at x. To summarize, we have

9.3. LEMMA. *Let $f\colon C \to R$ be convex and locally Lipschitz on a relatively open subset A of the quasi-interior of C. The f is Gâteaux differentiable at $x \in A$ if and only if $\partial f(x)$ is a singleton.*

9.4. LEMMA. *Let $f\colon C \to R$ be convex and locally Lipschitz on a relatively open subset A of C and let $x \in A$. Then*

(1) *If A is contained in the quasi-interior of C and f is Gâteaux differentiable at x then any selection $\sigma\colon A \to X^*$ for the subdifferential map is norm-to-w* continuous at x.*

(2) *If there exists a selection $\sigma\colon A \to X^*$ for the subdifferential map which is norm-to-w* continuous at x, then f is Gâteaux differentiable at x.*

PROOF. (1) Corollary 8.9 implies that the restriction of ∂f to A is w* usco and this clearly implies our assertion.

(2) The proof is analoguous to that of assertion (2) in the next lemma and we omit it. □

A similar result can be proved for the Fréchet differential. Namely

9.5. LEMMA. *Let $f\colon C \to R$ be a convex function.*

(1) *Assume that C is open and that f is locally Lipschitz on C and Fréchet differentiable at some $x \in C$. Then any selection $\sigma\colon C \to X^*$ of the subdifferential map is norm-to-norm continuous at x.*

(2) *Assume that there exists a relatively open subset A of C such that $f\,|\,A$ is locally Lipschitz. Assume also that there exists a selection $\sigma\colon A \to X^*$ of the subdifferential map which is norm-to-norm continuous at some $x \in A$. Then f is Fréchet differentiable at x.*

PROOF. (1) There is no loss of generality in assuming that $x = 0$ and $f(0) = 0$. Let r, $k > 0$ be such that $B(0,r) \subseteq C$, k is a Lipschitz constant for f on $B(0,r)$ and $\partial f(y) \subseteq kB^*$ for all $y \in B(0,r)$.

Let $\varepsilon > 0$ and let $\alpha = \varepsilon/(1+2k)$. Since f is Fréchet differentiable at 0, there exists $\delta > 0$, $\delta < \min\{\alpha,\, r,\, 1\}$, such that

$$f(z) - \langle Df_0, z\rangle \le \alpha\|z\|, \quad \text{for all } z \in B(0,\delta).$$

Let $y \in B(0,\delta^2)$, $y^* \in \partial f(y)$ and $z \in X$ with $\|z\| = \delta$. Using the above inequality and the fact that y^* is a subgradient of f at y (see Definition 8.1 (1)), we have

$$\begin{aligned}\langle y^* - Df_0, z\rangle &\le f(z) - f(y) + \langle y^*, y\rangle - \langle Df_0, z\rangle \\ &\le \alpha\|z\| - f(y) + \langle y^*, y\rangle \le \alpha\delta + k\delta^2 + k\delta^2 < \delta\alpha(1+2k) = \delta\varepsilon\end{aligned}$$

hence $\|y^* - Df_0\| < \varepsilon$, which proves (1).

(2) Let $\sigma\colon A \to X^*$ be any selection of ∂f. Then for x, $y \in A$ we have

$$\langle \sigma(x), y - x\rangle \le f(y) - f(x) \quad \text{and} \quad \langle \sigma(y), x - y\rangle \le f(x) - f(y)$$

hence

$$0 \le f(y) - f(x) - \langle \sigma(x), y - x\rangle \le \langle \sigma(x) - \sigma(y), y - x\rangle \le \|\sigma(x) - \sigma(y)\| \cdot \|x - y\|.$$

If σ is continuous at x this shows immediately that f is Fréchet differentiable at x (and that $\sigma(x)$ is a Fréchet differential of f at x). This proves (2). □

9.6. Corollary. *Let $C \subseteq X$ be an open convex subset and let $f\colon C \to R$ be convex and locally Lipschitz. Then the set C_0 consisting of all points of C at which f is Fréchet differentiable is a relative G_δ subset of C.*

Proof. Let $\sigma\colon C \to X^*$ be any selection for ∂f. From the preceding lemma, C_0 consists of the points of C at which σ is norm-to-norm continuous, which is known to be a G_δ subset of C. □

9.7. Definitions. (1) A Banach space X is called an *Asplund space* if every convex and continuous function defined on an open convex subset of X is Fréchet differentiable on a dense G_δ subset of its domain.

(2) A Banach space X is called a *weak Asplund space* if every convex and continuous function defined on an open convex subset of X is Gâteaux differentiable on a dense G_δ subset of its domain.

REMARK. One can alternatively define Asplund spaces as being those Banach spaces with the property that every convex and continuous function defined on an open convex subset of X is Fréchet differentiable at at least one point of its domain.

Indeed, if a Banach space X satisfies this last condition, then any continuous convex function $f: C \to R$, C being open and convex, is Fréchet differentiable on a dense subset C_0 of C and, by Corollary 9.6, C_0 is also a G_δ subset of C. Thus X is an Asplund space. The other implication being obviously true, the definitions are equivalent.

9.8. THEOREM. *A Banach space X is an Asplund space if and only if X^* has the Radon-Nikodým property.*

PROOF. Assume first that X is an Asplund space. Let A be a bounded subset of X^* and let $\varepsilon > 0$. Let $p: X \to R$ be given by

$$p(x) = \sup\{\langle x^*, x\rangle : x^* \in A\}.$$

Clearly p is a sublinear functional on X. Since A is bounded, p is bounded and therefore continuous. By hypothesis, there exists $x \in X$ such that p is Fréchet differentiable at x. It follows that there exists $\delta > 0$ such that

$$p(x+v) + p(x-v) - 2p(x) \le \varepsilon\|v\|/2, \quad \text{for all } v \in B(0, 2\delta).$$

Let x^*, $y^* \in \Sigma = \Sigma(x, A, \varepsilon\delta/4)$ and $v \in X$ with $\|v\| = \delta$. From the definition of p, the definition of a slice (see 4.6) and the above inequality we have

$$\begin{aligned}\langle x^* - y^*, v\rangle &= \langle x^*, x+v\rangle + \langle y^*, x-v\rangle - \langle x^* + y^*, x\rangle \\ &\le p(x+v) + p(x-v) - 2p(x) + \varepsilon\delta/2 \\ &\le \varepsilon\|v\|/2 + \varepsilon\|v\|/2 = \varepsilon\|v\|\end{aligned}$$

and therefore $\|x^* - y^*\| \le \varepsilon$. Thus Σ has diameter less than ε, which proves that X^* has the Radon-Nikodým property.

The converse implication can be proved easily from some of the previous results. It is also a consequence of the next, more general fact. □

9.9. THEOREM. *Let X be an Asplund space, C be a convex subset of X and $f: C \to R$ be a convex function. Let A be a Baire, relatively open subset of C such that $f\,|A$ is locally Lipschitz. Then there exists a dense G_δ subset A_0 of A such that f is Fréchet differentiable at each point of A_0.*

PROOF. By Theorem 8.7(2) and Corollary 8.8, $\partial f: A \to 2^{X^*}$ is a locally efficient, maximal monotone operator. We can apply Theorem 5.8(1) to obtain a locally bounded w* usco map $(\partial f)_\psi: A \to 2^{X^*}$ for some continuous $\psi: A \to R$. By Corollary 4.7, there exists a dense G_δ subset A_0 of A and a selection σ for $(\partial f)_\psi$ which is continuous at all the points of A_0. Notice that σ is a selection for ∂f too and therefore by Lemma 9.5(2) f is Fréchet differentiable at each point of A_0. □

We shall continue by giving some characterizations for Asplund spaces.

9.10. PROPOSITION. *A Banach space X is an Asplund space if and only if for any Baire space A and any w* usco map $F: A \to 2^{X^*}$ there exists a dense G_δ subset A_0 of A and a selection σ for F which is norm continuous at each point of A_0.*

PROOF. If X is Asplund then X^* has the Radon-Nikodým property (see Theorem 9.8) and Corollary 4.7 implies that X has the stated property.

Conversely, assume that X has the stated property and let $f: C \to R$ be a continuous convex function, with C open. Then f is locally Lipschitz and therefore $\partial f: C \to 2^{X^*}$ is w* usco. By hypothesis there exists a selection σ for ∂f and a dense G_δ subset C_0 of C such that σ is norm-to-norm continuous at each point of C_0. Lemma 9.5(2) implies now that f is Fréchet differentiable at each point of C_0. Thus X is an Asplund space. □

9.11. PROPOSITION. *Let X be a weak Asplund space. Then X is an Asplund space if and only if for any Baire space A and any w* continuous map $f: A \to X^*$ there exists a dense G_δ subset $A_0 \subset A$ such that f is norm continuous at each point of A_0.*

PROOF. Let X be an Asplund space and let f be as above. Then f is a minimal w* usco map, hence, by Corollary 4.7, it is norm continuous at each point of a dense G_δ subset of A.

Conversely, let $f\colon C \to R$ be a continuous convex function, with C open. Since X is weak Asplund there exists a dense G_δ subset C_1 of C such that f is Gâteaux differentiable at each point of C_1. Let $\sigma\colon C \to X^*$ be any selection for ∂f. By Lemma 9.4, σ is w* continuous at each point of C_1. By assumption there exists a dense G_δ subset C_0 of C_1 such that σ is norm continuous at each point of C_0. By Lemma 9.5 (2), f is Fréchet differentiable at each point of C_0. Since C_0 is obviously a dense G_δ subset of C, the proposition is proved. □

REMARK. It is not my aim to discuss here all the known characterizations of Asplund spaces, however I would like to mention one more. Namely:

A Banach space X is an Asplund space if and only if every closed separable subspace of X has a separable dual. In particular, a Banach space with separable dual is an Asplund space (see, for example, [**14**]).

We now turn our attention to the class of weak Asplund spaces, which in contrast to the class of Asplund spaces has no characterization yet. However, a lot of Banach spaces have been proven to be weak Asplund. Among them, are the Banach spaces of class **S**, which in view of Corollary 4.10 and Corollary 4.11 is fairly large. Indeed we have:

9.12. THEOREM. *Let X be a Banach space in the class* **S**, *let C be a convex subset of X and let $f\colon C \to R$ be a convex function. Assume that there exists a Baire, relatively open subset A of C such that $f\,|\,A$ is locally Lipschitz. Then there exists a dense G_δ subset A_0 of A such that f is Gâteaux differentiable at each point of A_0. In particular, X is a weak Asplund space.*

PROOF. We begin as in the proof of Theorem 9.9 to find a continuous $\psi\colon A \to R$ such that $(\partial f)_\psi\colon A \to 2^{X^*}$ is a w* usco map. By Remark (1) in Section 4.3

there exists a selection $\sigma\colon A \to X^*$ for $(\partial f)_\psi$ which is w* continuous at each point of a dense G_δ subset A_0 of A. Lemma 9.4 (2) implies now that f is Gâteaux differentiable at each point of A_0. □

9.13. COROLLARY. *Let X be a Banach space in the class* **S**, *let C be a convex subset of X and let $f\colon C \to R$ be a convex and locally Lipschitz function. Then the set of points of C where f is not Gâteaux differentiable cannot contain a Baire, quasi-open convex subset of C.*

At this point we have proved our main results about generic differentiability of convex functions. We shall conclude by proving a chain rule type result.

9.14. THEOREM. *Let X and Y be Banach spaces, let B be a Baire subset of Y, C be a convex subset of X and A be a relatively open, Baire subset of C. Let $h\colon B \to A$ be continuous and Fréchet differentiable (resp. Gâteaux differentiable) and $f\colon C \to R$ be convex and such that $f\,|\,A$ is locally Lipschitz. Assume that X is Asplund (resp. in the class* **S**). *Then there exists a dense G_δ subset B_0 of B such that $f \circ h$ is Fréchet (resp. Gâteaux) differentiable at each point of B_0.*

PROOF. We shall prove only the Fréchet part, the proof of the Gâteaux one being similar (in fact simpler). Exactly as in the first part of the proof of Theorem 9.9 we can find a continuous $\psi\colon A \to R$ such that $(\partial f)_\psi\colon A \to 2^{X^*}$ is a locally bounded, w* usco map. Then the set valued map $F\colon B \to 2^{X^*}$ defined by $F(x) = (\partial f)_\psi(h(x))$ is locally bounded and w* usco. By Corollary 4.7 there exists a selection σ for F which is norm continuous at each point of a dense G_δ subset B_0 of B. For each $x \in B_0$ define

$$DF_x\colon Y \to R \qquad \text{by} \qquad DF_x = \sigma(x) \circ Dh_x,$$

where Dh_x is a Fréchet differential of h at x. Clearly DF_x is linear and continuous. We shall show that it is a Fréchet differential of $f \circ h$ at x.

To this end let $y \in B$. Since $\sigma(x) \in \partial f(h(x))$ and $\sigma(y) \in \partial f(h(y)y)$ we have

$$\langle \sigma(x), h(y) - h(x) \rangle \leq f(h(y)) - f(h(x))$$

and

$$\langle \sigma(y), h(x) - h(y) \rangle \leq f(h(x)) - f(h(y))$$

hence

$$0 \leq f \circ h(y) - f \circ h(x) - \langle \sigma(x), h(y) - h(x) \rangle \leq \langle \sigma(y) - \sigma(x), h(y) - h(x) \rangle.$$

Using the Fréchet differentiability of h at x we get

$$\begin{aligned} 0 &\leq f \circ h(y) - f \circ h(x) - \langle \sigma(x), Dh_x(y-x) + \|y-x\| \mathcal{O}_{h,x}(y) \rangle \\ &\leq \langle \sigma(y) - \sigma(x), Dh_x(y-x) + \|y-x\| \mathcal{O}_{h,x}(y) \rangle \end{aligned}$$

or

$$\begin{aligned} \langle \sigma(x), \mathcal{O}_{h,x}(y) \rangle &\leq \frac{f \circ h(y) - f \circ h(x) - \langle \sigma(x), Dh_x(y-x) \rangle}{\|y-x\|} \\ &\leq \langle \sigma(y), \mathcal{O}_{h,x}(y) \rangle + \frac{\langle \sigma(y) - \sigma(x), Dh_x(y-x) \rangle}{\|y-x\|} \\ &\leq \langle \sigma(y), \mathcal{O}_{h,x}(y) \rangle + \|\sigma(x) - \sigma(y)\| \cdot \|Dh_x\|. \end{aligned}$$

Since $\langle \sigma(x), Dh_x(y-x) \rangle = \langle DF_x, y-x \rangle$, we obtain

$$\langle \sigma(x), \mathcal{O}_{h,x}(y) \rangle \leq \mathcal{O}_{f\circ h,x}(y) \leq \langle \sigma(y), \mathcal{O}_{h,x}(y) \rangle + \|\sigma(x) - \sigma(y)\| \cdot \|Dh_x\|.$$

Since σ is norm-to-norm continuous at x and $\lim\limits_{y \to x} \mathcal{O}_{h,x}(y) = 0$ it follows that $\lim\limits_{y \to x} \mathcal{O}_{f\circ h,x}(y) = 0$, which proves our assertion. □

REMARK. If in the above theorem we take $X = Y$ and $B = A$ we reobtain Theorem 9.9 and Theorem 9.12.

Comments

Theorem 8.7 is well known for open convex sets (with the only apparently weaker assumption that f is lower semicontinuous). Without any assumption on the interior of C, it is known that the lower semicontinuity of f implies the subdifferentiability of f on a dense subset of C (see [4]). In our case C is arbitrary and by imposing stronger conditions on f we obtain its subdifferentiability on any relatively open subset of C on which it is locally Lipschitz. A first version of Theorem 8.7 was proved by the author in **[23]** for the case of quasi-open convex sets. Corollary 8.8 is due to Rockafellar **[19]** who proved it for any lower semicontinuous function; our context is less general, but the proof is simpler. Theorem 8.11 is one of the results in **[24]**. One implication in Theorem 9.8 (the fact that the dual of an Asplund space has the Radon-Nikodým property) is due to Namioka and Phelps **[12]**; the other one is due to Stegall **[20]**. Theorem 9.12 was first proved by the author in the case of a separable Banach space and of a quasi-open convex set in **[23]**; Rainwater **[16]** extended it to the case of Banach spaces in the class **S** and also proved Theorem 9.9, both in the case of quasi open convex sets. Here we extended Rainwater's results to the case of arbitrary convex sets. Propositions 9.10 and 9.11 seem to be known to the specialists in the field, but have never appeared in published form. Theorem 9.14 is an extension of a result due to Stegall **[22]** for open sets; a less general version (for quasi-open sets) is contained in **[24]**. Some of the results in this chapter (parts of Theorems 8.7 and 8.11 and Theorem 9.9) were obtained independently and by completely different methods by Noll **[13]**.

REFERENCES

1. D. Amir and J. Lindenstrauss, *The structure of weakly compact subsets of Banach spaces*, Ann. Math. **88** (1968), 35–46.
2. E. Asplund, *Fréchet differentiability of convex functions*, Acta Math. **121** (1968), 31–47.
3. J. M. Borwein and S. P. Fitzpatrick, *Local boundedness of monotone operators under minimal hypotheses*, Bull. Australian Math. Soc. (to appear).
4. A. Brønsted and R. T. Rockafellar, *On the subdifferentiability of the convex functions*, Proc. Amer. Math. Soc. **16** (1965), 605–611.
5. L. Drewnowski and I. Labuda, *On minimal convex usco and maximal monotone maps*, Preprint.
6. L. Jokl, *Minimal convex-valued weak* usco correspondences and the Radon-Nikodým property*, Comm. Math. Univ. Carolinae **28** (1987), 353-375.
7. P. S. Kenderov, *The set-valued monotone mappings are almost everywhere single-valued*, C. R. Acad. Bulgare Sci. **27** (1974), 1173–1175.
8. P. S. Kenderov, *Monotone operators in Asplund spaces*, C. R. Acad. Bulgare Sci. **30** (1977), 963–964.
9. V. L. Klee, *Convex sets in linear spaces*, Duke Math. J. **18** (1951), 443–446.
10. S. Mazur, *Über konvexe Mengen in linearen normierten Räumen*, Studia Math. **4** (1933), 70–84.
11. J. R. Munkres, *Topology: A First Course*, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1975.
12. I. Namioka and R. R. Phelps, *Banach spaces which are Asplund spaces*, Duke Math. J. **42** (1975), 735–750.
13. D. Noll, *Generic Fréchet differentiability of convex functions on small sets*, Preprint, 1987.
14. R. R. Phelps, *Convex Functions, Monotone Operators and Differentiability*, Lecture Notes in Mathematics **1364**, Springer-Verlag, Berlin, 1989.
15. R. R. Phelps, *Some topological properties of support points of convex sets*, Israel J. Math. **13** (1972), 327–336.
16. J. Rainwater, *Yet more on the differentiability of convex functions*, Proc. Amer. Math. Soc. **103** (1988), 773–778.
17. J. Rainwater, *A generic converse to Verona's subdifferentiability theorem*, Preprint, 1987.
18. R. T. Rockafellar, *Local boundedness of nonlinear monotone operators*, Mich. Math. J. **16** (1969), 397–407.
19. R. T. Rockafellar, *On the maximal monotonicity of subdifferential mappings*, Pacific J. Math. **33** (1970), 209–216.

20. C. Stegall, *The duality between Asplund spaces and spaces with the Radon-Nikodým property*, Israel J. Math. **29** (1978), 408–412.

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

22. C. Stegall, *Gâteaux differentiation of functions on a certain class of Banach spaces*, Functional Analysis: Surveys and Recent Results III, North-Holland, 1984, 35–46.

23. M. E. Verona, *More on the differentiability of convex functions*, Proc. Amer. Math. Soc. **103** (1988), 137–140.

24. M. E. Verona, *On the differentiability of convex functions*, in Proceedings of the Centre of Mathematical Analysis, Australian National University, vol. **20** (1988), edited by S. Fitzpatrick and J. Giles, 195–202.

25. A. Verona and M. E. Verona, *Locally efficient monotone operators*. Preprint, 1989.