An Abstract Variational Theorem

Warren B. Moors[1] and Neşet Özkan Tan

Abstract. Let $(X, \|\cdot\|)$ be a Banach space and $f : X \to \mathbb{R} \cup \{\infty\}$ be a proper function. Then the **Fenchel conjugate of f** is the function $f^* : X^* \to \mathbb{R} \cup \{\infty\}$ defined by,

$$f^*(x^*) := \sup\{(x^* - f)(x) : x \in X\}.$$

In this article we will prove a theorem more general than the following.

Theorem: Let $f : X \to \mathbb{R} \cup \{\infty\}$ be a proper function on a Banach space $(X, \|\cdot\|)$. If there is a nonempty open subset A of $\mathrm{Dom}(f^*)$ such that $\mathrm{argmax}(x^* - f) \neq \varnothing$ for each $x^* \in A$, then there is a dense and G_δ subset R of A such that $(x^* - f) : X \to \mathbb{R} \cup \{-\infty\}$ has a strong maximum for each $x^* \in R$. In addition, if $0 \in A$ and $0 < \varepsilon$ then there is an $x^* \in X^*$ with $\|x^*\| < \varepsilon$ such that $(x^* - f) : X \to \mathbb{R} \cup \{-\infty\}$ has a strong maximum.

AMS (2010) subject classification: Primary 46B20; Secondary 46B10, 46B50.

Keywords: Variational theorem, James' weak compactness theorem, convex analysis.

1 Introduction

The purpose of this paper is to prove a general variational theorem (see Theorem 3.4). The problem with proving a general theorem of this type is that one cannot expect that the conclusion is very strong. One can only get out, what one puts in. So if you do not assume much, i.e., "do not put much in", then you cannot expect "much out". Having said that, the striking feature of the variational theorem stated in the abstract, is that it says anything at all. The basis for this somewhat curious variational theorem is a rather technical preliminary result (Theorem 2.13) whose proof occupies most of the next section. As a consequence, Section 2 contains a long sequence of preliminary results, which do not make for enjoyable reading. We hope that the discomfort caused by this long sequence of results is compensated by the fact that from Theorem 2.13 one may deduce several non-trivial facts, including James' weak compactness theorem, [7].

2 Preliminary Results

We start with some preliminary definitions.

A set-valued mapping Φ from a topological space (A, τ') into subsets of a topological space (X, τ) is τ*-upper semicontinuous* at a point $x_0 \in A$ if for each τ-open set W in X, containing

[1]Corresponding author

$\Phi(x_0)$, there exists an open neighbourhood U of x_0 such that $\Phi(U) \subseteq W$. In the case when Φ also has a nonempty compact image at x_0, then we say that Φ is a *τ-usco at x_0*. Furthermore, if $(X, +, \cdot, \tau)$ is a linear topological space then we call Φ a *τ-cusco at x_0*, if Φ is a τ-usco at x_0 and $\Phi(x_0)$ convex.

If Φ is τ-upper semicontinuous (a τ-usco) [a τ-cusco] at each point of A then we say that Φ is *τ-upper semicontinuous on A* (*a τ-usco on A*) [*a τ-cusco on A*].

An usco (cusco) from a topological space (A, τ') into subsets of a topological space (X, τ) (linear topological space $(X, +, \cdot, \tau)$) is said to be a *minimal usco* (*minimal cusco*) if its graph does not contain, as a proper subset, the graph of any other usco (cusco) on A.

The following proposition ensures that there is an ample supply of minimal usco mappings.

Proposition 2.1 ([2])**.** *Suppose that (X, τ) and (Y, τ') are topological spaces and $\Phi : X \to 2^Y$ is an usco on X. If (Y, τ') is Hausdorff then there exists a minimal usco mapping $\Psi : X \to 2^Y$ such that $\mathrm{Gr}(\Psi) \subseteq \mathrm{Gr}(\Phi)$ (i.e., every usco contains a minimal usco).*

Now that we have established the existence of minimal usco mappings, it is interesting to observe that there is in fact a characterisation of when an usco is a minimal usco.

Proposition 2.2 ([2, 4])**.** *Let $\Phi : A \to 2^X$ be a τ-usco acting from a topological space (A, τ') into subsets of a topological space (X, τ). Then Φ is a minimal τ-usco if, and only if, for every pair of open subsets U of A and W of X such that $\Phi(U) \cap W \neq \varnothing$, there exists a nonempty open subset V of U such that $\Phi(V) \subseteq W$.*

For most of our considerations in this paper we will be dealing with cuscos rather than uscos, so the next proposition, which explains the connection between these two notions, is very useful.

Proposition 2.3 ([8, 12])**.** *Suppose that $\Phi : A \to 2^X$ is a τ-usco acting from a topological space (A, τ') into subsets of a locally convex space $(X, +, \cdot, \tau)$. If for each $t \in A$, $\overline{co}^\tau \Phi(t)$ is a compact subset of X, then the mapping $\Psi : A \to 2^X$ defined by, $\Psi(t) := \overline{co}^\tau \Phi(t)$ for all $t \in A$, is a τ-cusco on A.*

The relationship between convex functions and minimal cuscos is revealed next.

Let $f : C \to \mathbb{R}$ be a convex function defined on a nonempty convex subset of a normed linear space $(X, \|\cdot\|)$ and let $x \in C$. Then we define the *subdifferential* $\partial f(x)$ by,

$$\partial f(x) := \{x^* \in X^* : x^*(y) - x^*(x) \leq f(y) - f(x) \text{ for all } y \in C\}.$$

Proposition 2.4 ([10, Corollory 4.8])**.** *If $\varphi : A \to \mathbb{R}$ is a continuous convex function defined on a nonempty open convex subset A of a normed linear space $(X, \|\cdot\|)$, then the subdifferential mapping, $x \mapsto \partial\varphi(x)$, is a minimal weak*-cusco on A.*

One can show, using the Uniform Boundedness Theorem, [3, page 66], that all weak* usco mappings acting from a first countable topological space into subsets of the dual of a Banach space are locally bounded (see, [1, Lemma 3]). However, one can also directly deduce the local boundedness of the subdifferential mapping from the local Lipschitz behaviour of continuous convex functions. This is recalled next.

Proposition 2.5 ([12, Proposition 1.11]). *Let A be a nonempty open and convex subset of a normed linear space $(X, \|\cdot\|)$ and let $\varphi : A \to \mathbb{R}$ be a continuous convex function. If $x_0 \in A$, then the mapping, $x \mapsto \partial\varphi(x)$, is locally bounded at x_0. That is, there exists an $0 < L$ and a $0 < \delta$ such that $B(x_0, \delta) \subseteq A$ and $\|x^*\| \leq L$ whenever $x \in B(x_0, \delta)$ and $x^* \in \partial\varphi(x)$.*

Since we will be working extensively with sequences it makes sense to introduce the following notation.

If $(x_k : k \in \mathbb{N})$ is a sequence in a topological space (X, τ) then we define

$$cl_\tau(x_k : k \in \mathbb{N}) := \bigcap_{n=1}^{\infty} \overline{\{x_k : n \leq k\}}^{\tau}.$$

That is, $cl_\tau(x_k : k \in \mathbb{N})$ is the set of all τ-cluster points of $(x_k : k \in \mathbb{N})$. Further, if $(X, +, \cdot, \tau)$ is a linear topological space then we define $K_\tau(x_k : k \in \mathbb{N}) := \overline{\text{co}}^{\tau}(cl_\tau(x_k : k \in \mathbb{N}))$.

The following six technical results are needed in the proof of Theorem 2.13. Unfortunately, there is not much more that can be done to break-up the presentation of these results.

Lemma 2.6 ([10, Lemma 4.11]). *Let $\varphi : A \to \mathbb{R}$ be a τ-continuous convex function defined on a nonempty convex subset A of a locally convex space $(X, +, \cdot, \tau)$ and let τ' be a Hausdorff locally convex topology on X such that (i) $\tau' \subseteq \tau$ and (ii) φ is τ'-lower semicontinuous. If T is a nonempty τ'-closed and convex subset of X and S is any τ-separable subset of A such that $S - T \subseteq A$ then, for every sequence $(x_n : n \in \mathbb{N})$ in T, either there exists a subsequence without any τ'-cluster points, or else, there exists a subsequence, $(x_{n_k} : k \in \mathbb{N})$ of $(x_n : n \in \mathbb{N})$ such that φ is constant on, $y - aK_{\tau'}(x_{n_k} : k \in \mathbb{N})$ for all $y \in S$ and all $a \in [0, 1]$.*

In Theorem 2.13 we shall apply this lemma in the case when τ is the norm-topology on a Banach space $(Y, \|\cdot\|)$ and τ' is the $\sigma(Y, Z)$-topology on Y, for some subset Z of Y^*. Here, $\sigma(Y, Z)$ denotes the weakest topology on Y that makes each member of Z continuous.

Proposition 2.7 ([10, Proposition 3.14]). *Let $(X, +, \cdot, \tau)$ be a locally convex space and let $(y_n : n \in \mathbb{N})$ be a sequence in a τ-compact convex subset K of X. If $(x_n : n \in \mathbb{N})$ is any sequence such that $x_n \in \text{co}\{y_k : n \leq k\}$ for all $n \in \mathbb{N}$, then $cl_\tau(x_n : n \in \mathbb{N}) \subseteq K_\tau(y_n : n \in \mathbb{N})$.*

For a nonempty subset Y of a normed linear space $(X, \|\cdot\|)$ and a point $x \in X$ we define $\text{dist}(x, Y) := \inf\{\|x - y\| : y \in Y\}$.

Lemma 2.8 ([10, Lemma 3.8]). *Let $(X, \|\cdot\|)$ be a normed linear space, let Y be a finite-dimensional subspace of X^* and let $0 < \varepsilon$. If $x^* \in X^*$ and $\varepsilon < dist(x^*, Y)$, then there exists an $x \in S_X$ such that $\varepsilon < x^*(x)$ and $0 = y^*(x)$ for all $y^* \in Y$.*

Lemma 2.9 ([11]). *Let $(V, +, \cdot)$ be a vector space (over $\mathbb{R}$) and let $\varphi : A \to \mathbb{R}$ be a convex function defined on a convex set A with $0 \in A$. If $(A_n : n \in \mathbb{N})$ is a decreasing sequence of nonempty convex subsets of V, $(\beta_n : n \in \mathbb{N})$ is a sequence of strictly positive numbers such that $(\sum_{i=1}^{\infty} \beta_i)A_1 \subseteq A$ and*

$$\beta_1 r + \varphi(0) < \inf_{a \in A_1} \varphi(\beta_1 a) \quad \textit{for some } 0 < r < \infty,$$

then there exists a sequence $(a_n : n \in \mathbb{N})$ in V such that, for all $n \in \mathbb{N}$:

(i) $a_n \in A_n$ *and*

(ii) $\varphi(\sum_{i=1}^{n} \beta_i a_i) + \beta_{n+1} r < \varphi(\sum_{i=1}^{n+1} \beta_i a_i)$.

Note that if, in Lemma 2.9, φ is sublinear then the assumption: $\beta_1 r + \varphi(0) < \inf_{a \in A_1} \varphi(\beta_1 a)$ simplifies to: $r < \inf_{a \in A_1} \varphi(a)$. However, in Theorem 2.13, we need the full version of Lemma 2.9.

The following proposition is a version of Grothendieck's compactness theorem, [6]. As it involves several potentially different types of compactness we will make precise exactly what the various notions of compactness mean, at least in the setting considered in this paper.

If $(Y, \|\cdot\|)$ is a Banach space and Z is a subset of Y^* then we shall say that a subset K of Z is a *relatively weak*-countably compact subset of* Z, if every sequence in K has a weak* cluster-point in Z. Furthermore, we shall say that a subset C of Z is *relatively weakly compact*, if $\overline{C}^{\text{weak}}$ is a weakly compact subset of Y^*.

The following proposition also involves the notion of pointwise convergence. So we shall introduce the relevant notation here. If X is a nonempty set and $Y \subseteq \mathbb{R}^X$ (i.e., Y is a family of real-valued functions defined on X) then we shall denote the topology on Y of, *pointwise convergence on* X, by $\tau_p(X)$.

Proposition 2.10. *Let* $(Y, \|\cdot\|)$ *be a Banach space and let* Z *be a subset of* Y^*. *If the* $\sigma(Y, Z)$*-topology on* B_Y *is compact and Hausdorff, then every bounded, relatively weak*-countably compact subset of* $\overline{\text{span}}(Z)$*, is relatively weakly compact. In particular, if* C *is a weakly closed and bounded, relatively weak*-countably compact subset of* $\overline{\text{span}}(Z)$*, then* C *is weakly compact.*

Proof. Let K be a bounded relatively weak* countably compact subset of $\overline{\text{span}}(Z)$. Let $C(B_Y)$ denote the set of all real-valued, $\sigma(Y, Z)$-continuous functions defined on B_Y and let $\pi : \overline{\text{span}}(Z) \to C(B_Y)$ be defined by, $\pi(z^*) := z^*|_{B_Y}$ for all $z^* \in \overline{\text{span}}(Z)$. Note that this function is well-defined, i.e., each $\pi(z^*)$ is indeed $\sigma(Y, Z)$-continuous on B_Y. To see this, note that there exists a sequence $(z_n^* : n \in \mathbb{N})$ in $\text{span}(Z)$ such that $z^* := \lim_{n\to\infty} z_n^*$ (the convergence being with respect to the dual norm on Y^*). As each z_n^* is a linear combination of elements of Z, each z_n^* is $\sigma(Y, Z)$-continuous on Y. Furthermore, as $\|\pi(z^*) - \pi(z_n^*)\|_\infty = \|z^* - z_n^*\|$ for each $n \in \mathbb{N}$, $\pi(z^*) = \lim_{n\to\infty} \pi(z_n^*)$, with respect to the $\|\cdot\|_\infty$-norm on B_Y. Hence, as the uniform limit of $\sigma(Y, Z)$-continuous functions on B_Y, $\pi(z^*)$ is $\sigma(Y, Z)$-continuous on B_Y. If we consider $\overline{\text{span}}(Z)$ with the (restriction of the) dual norm and $C(B_Y)$ equipped with the supremum norm $\|\cdot\|_\infty$, then π is a linear isometric embedding. Therefore, $\pi(\overline{\text{span}}(Z))$ is a closed linear subspace of $C(B_Y)$, and hence weakly closed [3, page 422]. Furthermore, $\pi : (\overline{\text{span}}(Z), \text{weak}^*) \to (C(B_Y), \tau_p(B_Y))$ is continuous and $\pi^{-1} : (\pi(\overline{\text{span}}(Z)), \text{weak}) \to (\overline{\text{span}}(Z), \text{weak})$ is also continuous, as π^{-1} is a bounded linear operator from $\pi(\overline{\text{span}}(Z))$ onto $\overline{\text{span}}(Z)$, [3, page 422]. Now, since $\pi : (\overline{\text{span}}(Z), \text{weak}^*) \to (C(B_Y), \tau_p(B_Y))$ is continuous, $\pi(K)$ is relatively $\tau_p(B_Y)$-countably compact. Therefore, by [6, Theorem 2], $\overline{\pi(K)}^{\tau_p(B_Y)}$ is $\tau_p(B_Y)$-compact. Now, since $\pi : (\overline{\text{span}}(Z), \|\cdot\|) \to (C(B_Y), \|\cdot\|_\infty)$ is an isometric embedding, $\pi(K)$ is bounded in $(C(B_Y), \|\cdot\|_\infty)$. Thus, by [6, Theorem 5], $\overline{\pi(K)}^{\tau_p(B_Y)}$ is weakly compact and so $\overline{\pi(K)}^{\text{weak}}$, which is a subset of $\pi(\overline{\text{span}}(Z))$, is also weakly compact. Hence,

$$K = \pi^{-1}(\pi(K)) \subseteq \pi^{-1}(\overline{\pi(K)}^{\text{weak}})$$

and the latter set is weakly compact as $\pi^{-1} : (\pi(\overline{\text{span}}(Z)), \text{weak}) \to (\overline{\text{span}}(Z), \text{weak})$ is continuous.

Note that if K is also weakly closed then K is weakly compact. $\square$

Remark 2.11. *By examining the proof of Proposition 2.10, or otherwise, one can see that for any subset Z in the dual of a normed linear space $(Y, \|\cdot\|)$ and any bounded subset A of Y, the relative $\sigma(Y, Z)$-topology on A coincides with the relative $\sigma(Y, \overline{span}(Z))$-topology on A.*

The next lemma requires the notion of lower semicontinuity. Let (X, τ) be a topological space. We say a function $f : X \to \mathbb{R} \cup \{\infty\}$ is *lower semicontinuous* if for every $\alpha \in \mathbb{R}$, $\{x \in X : f(x) \leq \alpha\}$ is a closed set.

Lemma 2.12. *Let A be a nonempty open bounded convex subset of a Banach space $(Y, \|\cdot\|)$ and let Z be a subset of Y^*. If $\varphi : A \to \mathbb{R}$ is a continuous convex function and $\partial\varphi(y) \cap \overline{span}(Z) \neq \varnothing$ for all $y \in A$, then φ is $\sigma(Y, Z)$-lower semicontinuous.*

Proof. Let $y_0 \in A$ and let $0 < \varepsilon$. Then, there exists an $z^* \in \partial\varphi(y_0) \cap \overline{\text{span}}(Z)$. Define $h : A \to \mathbb{R}$ by

$$h(y) := z^*(y - y_0) + \varphi(y_0) \quad \text{for all } y \in A.$$

Then observe that, since $z^* \in \partial\varphi(y_0)$, we have $h(y) \leq \varphi(y)$ for all $y \in A$. Now, by Remark 2.11, the set

$$U := \{y \in A : |z^*(y - y_0)| < \varepsilon\}$$

is a $\sigma(Y, Z)$-open neighbourhood of y_0 and for all $y \in U$, we have that

$$\varphi(y_0) - \varepsilon < z^*(y - y_0) + \varphi(y_0) = h(y) \leq \varphi(y).$$

Therefore, φ is $\sigma(Y, Z)$-lower-semicontinuous at y_0. Since y_0 was arbitrary, we conclude that φ is $\sigma(Y, Z)$-lower-semicontinuous on A. $\square$

Theorem 2.13. *Let A be a nonempty open convex subset of a Banach space $(Y, \|\cdot\|)$ and let Z be a subset of Y^*. If $\varphi : A \to \mathbb{R}$ is a continuous convex function and $\partial\varphi(y) \cap \overline{span}(Z) \neq \varnothing$ for all $y \in A$, and the $\sigma(Y, Z)$-topology on B_Y is compact and Hausdorff, then $\partial\varphi(y) \subseteq \overline{span}(Z)$ for all $y \in A$.*

Proof. Let $x_0 \in A$. Without loss of generality, we may assume that $x_0 = 0$. Indeed, if not, we consider the function $\psi : (A - x_0) \to \mathbb{R}$ defined by, $\psi(y) := \varphi(y + x_0)$. Note that ψ is continuous and convex and that $\partial\varphi(y + x_0) = \partial\psi(y)$ for all $y \in A - x_0$. In particular, $\partial\psi(y) \cap \overline{\text{span}}(Z) \neq \varnothing$ for all $y \in (A - x_0)$ and $\partial\varphi(x_0) = \partial\psi(0)$. So, if $x_0 \neq 0$, we can simply replace φ by ψ and use the argument at 0. Let us also note that we may assume that the set A is bounded, because if it is not, then we can replace the set A with the set $A' := A \cap \{y \in Y : \|y\| < 1\}$ and the function φ with the function $\varphi|_{A'}$, since $\partial\varphi(y) = \partial(\varphi|_{A'})(y)$ for all $y \in A'$ and, in particular, $\partial\varphi(0) = \partial(\varphi|_{A'})(0)$.

Now, since A is open and $y \mapsto \partial\varphi(y)$ is locally bounded (Proposition 2.5), there exist positive numbers m and L such that $mB_Y \subseteq A$ and $\|y^*\| \leq L$ for all $y^* \in \partial\varphi(mB_Y)$. Next, we let

$(\beta_n : n \in \mathbb{N})$ denote any sequence of strictly positive numbers such that $\sum_{n=1}^{\infty} \beta_n < m/2$ and $\lim_{n\to\infty} \frac{1}{\beta_n} \sum_{i=n+1}^{\infty} \beta_i = 0$.

Then, since $y \mapsto \partial\varphi(y)$ is a weak*-usco, (see, Corollary 2.4) there exists, by Proposition 2.1, a minimal weak*-usco, $M : A \to 2^{Y^*}$ such that $M(y) \subseteq \partial\varphi(y)$ for all $y \in A$. Furthermore, by Proposition 2.3 and the fact that $y \mapsto \partial\varphi(y)$ is a minimal weak*-cusco, (see, Corollary 2.4), $\partial\varphi(y) = \overline{\text{co}}^{w^*}[M(y)]$ for all $y \in A$.

Therefore, to show that $\partial\varphi(0) \subseteq \overline{\text{span}}(Z)$, it suffices to show that $M(0) \subseteq \overline{\text{span}}(Z)$. This is because, if $M(0) \subseteq \overline{\text{span}}(Z)$ then by Proposition 2.10, $M(0)$ is weakly compact and so, by the Krein-Phillips Theorem [9, 13], $\overline{\text{co}}[M(0)]$ is also weakly compact. Since the weak* topology is weaker than the weak-topology, $\overline{\text{co}}[M(0)]$ is then weak*-compact and hence weak*-closed. Therefore,

$$\partial\varphi(0) = \overline{\text{co}}^{w^*}[M(0)] = \overline{\text{co}}[M(0)] \subseteq \overline{\text{span}}(Z).$$

So suppose, for the purpose of obtaining a contradiction, that $M(0) \not\subseteq \overline{\text{span}}(Z)$. Then there exists an $y^* \in M(0) \setminus \overline{\text{span}}(Z)$. Since $\overline{\text{span}}(Z)$ is a closed subspace of Y^*, this means that there exists an $0 < \varepsilon$ such that $0 < \varepsilon < \text{dist}(y^*, \overline{\text{span}}(Z))$.

Part I: Let $y_0 := 0$. We inductively define sequences $(y_n : n \in \mathbb{N})$ in S_Y, $(v_n : n \in \mathbb{N})$ in $mB_Y \subseteq A$ and $(z_n^* : n \in \mathbb{N})$ in $\overline{\text{span}}(Z)$, such that the statements:

(A_n) $\quad z_n^* \in \partial\varphi(v_n) \cap \overline{\text{span}}(Z)$ and $\|v_n\| < m/n$;

(B_n) $\quad |(y^* - z_n^*)(y_j)| \leq \varepsilon/2$ for all $0 \leq j < n$;

(C_n) $\quad \varepsilon < y^*(y_n)$ and $z_j^*(y_n) = 0$ for all $1 \leq j \leq n$;

are true, for all $n \in \mathbb{N}$. [Note that for the remainder of this proof $B(0, r)$ will denote the open ball in Y, centred at 0, of radius $0 < r$.] For the first step, choose any $v_1 \in B(0, m) \subseteq A$. Then $\partial\varphi(v_1) \cap \overline{\text{span}}(Z) \neq \varnothing$ and so we may choose $z_1^* \in \partial\varphi(v_1) \cap \overline{\text{span}}(Z)$ which clearly satisfies $|(y^* - z_1^*)(y_0)| = 0 \leq \varepsilon/2$. Now note that

$$\varepsilon < \text{dist}(y^*, \overline{\text{span}}(Z)) < \text{dist}(y^*, \text{span}\{z_1^*\}),$$

and so, by Lemma 2.8, there exists $y_1 \in S_Y$ such that $\varepsilon < y^*(y_1)$ and $z_1^*(y_1) = 0$. So the statements (A_1), (B_1) and (C_1) hold.

Now fix $k \in \mathbb{N}$. Suppose that we have defined $\{v_1, \ldots, v_k\}$, $\{z_1^*, \ldots, z_k^*\}$ and $\{y_1, \ldots, y_k\}$ such that the statements (A_k), (B_k) and (C_k) hold true. Then consider the weak* open set

$$W := \bigcap_{j=0}^{k} \{x^* \in Y^* : |(y^* - x^*)(y_j)| < \varepsilon/2\}.$$

Note that $y^* \in M(0) \cap W$ and so $M(B(0, \frac{m}{k+1})) \cap W \neq \varnothing$. Therefore, by the minimality of M and Proposition 2.2, there exists a nonempty open set $V \subseteq B(0, \frac{m}{k+1})$ such that $M(V) \subseteq W$.

Choose $v_{k+1} \in V$. Then $\|v_{k+1}\| < m/(k+1)$. By hypothesis, since $v_{k+1} \in A$, we have that $\partial\varphi(v_{k+1}) \cap \overline{\text{span}}(Z) \neq \varnothing$, and so we may choose $z_{k+1}^* \in \overline{\text{span}}(Z)$ such that

$$z_{k+1}^* \in \partial\varphi(v_{k+1}) = \overline{\text{co}}^{w^*}[M(v_{k+1})] \subseteq \overline{\text{co}}^{w^*}(W) \subseteq \bigcap_{j=0}^{k} \{x^* \in Y^* : |(y^* - x^*)(y_j)| \leq \varepsilon/2\}.$$

Thus the statements (A_{k+1}) and (B_{k+1}) hold. Finally, observe that

$$\varepsilon < \operatorname{dist}(y^*, \overline{\operatorname{span}}(Z)) \leq \operatorname{dist}(y^*, \operatorname{span}\{z_1^*, \ldots, z_{k+1}^*\}),$$

and so, by Lemma 2.8, there exists $y_{k+1} \in S_Y$ such that $\varepsilon < y^*(y_{k+1})$ and $z_j^*(y_{k+1}) = 0$ for all $1 \leq j \leq k+1$. Therefore the statement (C_{k+1}) also holds. This completes the induction.

Part II: Now let $(n_k : k \in \mathbb{N})$ be a strictly increasing sequence of natural numbers. Then for all $k \in \mathbb{N}$, define $v_k' := v_{n_k}$ and $z_k^{*\prime} := z_{n_k}^*$ and $y_k' := y_{n_k}$. Also define $y_0' := 0$. Then the sequences $(v_n' : n \in \mathbb{N})$, $(z_n^{*\prime} : n \in \mathbb{N})$ and $(y_n' : n \in \mathbb{N})$ still satisfy (A_n), (B_n) and (C_n) for all $n \in \mathbb{N}$. Therefore, the properties (A_n), (B_n) and (C_n) are stable under passing to subsequences.

Now, since $\partial\varphi(y) \cap \overline{\operatorname{span}}(Z) \neq \varnothing$ for all $y \in A$, we have by Lemma 2.12 that φ is $\sigma(Y,Z)$-lower semicontinuous on A. Let $S := \frac{m}{2}B_Y \cap \operatorname{span}\{y_n : n \in \mathbb{N}\}$ and $T := \frac{m}{2}B_Y$ and note that $S - T \subseteq mB_Y \subseteq A$. Then, by passing to a subsequence and relabelling if necessary, we may assume, by applying Lemma 2.6 to the sequence $(y_n' : n \in \mathbb{N})$ defined, $y_n' := (m/2)y_n$ for all $n \in \mathbb{N}$, that

$$\varphi \text{ is constant on, } y - aK_{\sigma(Y,Z)}(y_n' : n \in \mathbb{N}), \text{ for all } 0 \leq a \leq 1 \text{ and all } y \in S.$$

Since $K_{\sigma(Y,Z)}(y_n' : n \in \mathbb{N}) = (m/2)K_{\sigma(Y,Z)}(y_n : n \in \mathbb{N})$, we have that

$$\varphi \text{ is constant on, } y - a[(m/2)]K_{\sigma(Y,Z)}(y_n : n \in \mathbb{N}), \text{ for all } 0 \leq a \leq 1 \text{ and all } y \in S. \quad (\dagger)$$

Finally, as $(y_n : n \in \mathbb{N})$ is a sequence in B_Y (which is $\sigma(Y,Z)$-compact), it must have a $\sigma(Y,Z)$-cluster point, which we will call y_∞.

Part III: Let $k \in \mathbb{N}$. For any $k \leq n$, we have that $z_k^*(y_n) = 0$ by the statement (B_n). Therefore, it follows that $z_k^*(y_\infty) = 0$. Since k was arbitrary, this is true for all $k \in \mathbb{N}$.

On the other hand, let $k \in \mathbb{N}$ and let $k < n$. Then, by the statement (A_n), we have that $|(y^* - z_n^*)(y_k)| < \varepsilon/2$. Moreover, from (B_k), we know that $\varepsilon < y^*(y_k)$. Combining these, we get that

$$\varepsilon/2 < y^*(y_k) + (z_n^* - y^*)(y_k) = z_n^*(y_k)$$

for all $k < n$. Therefore $\varepsilon/2 < z_n^*(y_k - y_\infty)$, for all $k < n$. We also note that, from the statement (A_n), we have that $(v_n : n \in \mathbb{N})$ converges to 0 in the norm topology. Therefore, since φ is norm-continuous, there exists $N_0 \in \mathbb{N}$ such that

$$|\varphi(v_n) - \varphi(0)| < \beta_1\varepsilon/8 \quad \text{for all} \quad N_0 < n.$$

Lastly observe that for all $n \in \mathbb{N}$, $v_n \in B(0,m)$ and $z_n^* \in \partial\varphi(v_n)$ and thus $\|z_n^*\| \leq L$ by the local boundedness of $\partial\varphi$. Therefore, if $\frac{8Lm}{\beta_1\varepsilon} < n$, then we have that

$$|z_n^*(v_n)| \leq \|z_n^*\|\|v_n\| \leq \frac{Lm}{n} < \frac{\beta_1\varepsilon}{8}.$$

Part IV: For each $n \in \mathbb{N}$, let $K_n := \mathrm{co}\{y_k : n \le k\} - y_\infty$ and note that $(K_n : n \in \mathbb{N})$ is a decreasing sequence of nonempty convex subsets of Y. Set, $0 < \varepsilon/8 =: r$. Then we have that

$$\beta_1 r + \varphi(0) < \inf_{x \in K_1} \varphi(\beta_1 x).$$

To see this, let $x \in K_1$. Then $x = \sum_{i=1}^k \lambda_i y_{n_i} - y_\infty$ where $0 \le \lambda_i$ for all $1 \le i \le k$ and $\sum_{i=1}^k \lambda_i = 1$. Set $\max\{n_1, \ldots, n_k, N_0, \frac{8Lm}{\beta_1 \varepsilon}\} < N$. Then we have,

$$\begin{aligned}
\beta_1\varepsilon/4 &< \beta_1 \textstyle\sum_{i=1}^k \lambda_i z_N^*(y_{n_i} - y_\infty) - \beta_1\varepsilon/4 \\
&= \beta_1(z_N^*(\textstyle\sum_{i=1}^k \lambda_i y_{n_i} - y_\infty)) - \beta_1\varepsilon/4 \\
&= \beta_1 z_N^*(x) - \beta_1\varepsilon/8 - \beta_1\varepsilon/8 \\
&< \beta_1 z_N^*(x) - z_N^*(v_N) - \beta_1\varepsilon/8 \\
&\le z_N^*(\beta_1 x) - z_N^*(v_N) + [\varphi(v_N) - \varphi(0)] \\
&\le [\varphi(\beta_1 x) - \varphi(v_N)] + [\varphi(v_N) - \varphi(0)] \qquad \text{since } z_N^* \in \partial\varphi(v_N) \\
&= \varphi(\beta_1 x) - \varphi(0).
\end{aligned}$$

Therefore, since $x \in K_1$ was arbitrary, we have that $\beta_1 r + \varphi(0) < \inf_{x \in K_1} \varphi(\beta_1 x)$ as claimed. So, by Lemma 2.9, there exists a sequence $(x_n : n \in \mathbb{N})$ such that for all $n \in \mathbb{N}$:

(i) $x_n \in \mathrm{co}\{y_k : n \le k\}$ and

(ii) $\varphi(\sum_{i=1}^n \beta_i(x_i - y_\infty)) + \beta_{n+1} r < \varphi(\sum_{i=1}^{n+1} \beta_i(x_i - y_\infty))$. $\qquad (*)$

Part V: Since $(x_n : n \in \mathbb{N})$ is a sequence in B_Y, it must have a $\sigma(Y,Z)$-cluster point. So, let x_∞ be a $\sigma(Y,Z)$-cluster point of $(x_n : n \in \mathbb{N})$. Then, by Proposition 2.7, we have that $x_\infty \in K_{\sigma(Y,Z)}(y_n : n \in \mathbb{N})$. Let $n \in \mathbb{N}$, then $0 < \sum_{i=1}^n \beta_i < m/2$ and so we have that

$$\textstyle\sum_{i=1}^n \beta_i x_i \in S = \frac{m}{2} B_Y \cap \mathrm{span}\{y_n : n \in \mathbb{N}\}.$$

Furthermore,

$$\begin{aligned}
\varphi(\textstyle\sum_{i=1}^n \beta_i(x_i - x_\infty)) &= \varphi(\textstyle\sum_{i=1}^n \beta_i x_i - (\sum_{i=1}^n \beta_i) \cdot x_\infty) \\
&= \varphi(\textstyle\sum_{i=1}^n \beta_i x_i - (\sum_{i=1}^n \beta_i) \cdot y_\infty) \quad \text{by } (\dagger) \\
&= \varphi(\textstyle\sum_{i=1}^n \beta_i(x_i - y_\infty)). \qquad (**)
\end{aligned}$$

Set $x := \sum_{i=1}^\infty \beta_i(x_i - x_\infty)$. Since $\|x_i - x_\infty\| \le 2$ for all $i \in \mathbb{N}$, we have that

$$\sum_{i=1}^\infty \|\beta_i(x_i - x_\infty)\| = \sum_{i=1}^\infty \beta_i \|x_i - x_\infty\| \le 2 \sum_{i=1}^\infty \beta_i < \infty.$$

Therefore, since Y is a Banach space, $x \in Y$. Furthermore, as φ is continuous, it is clear that

$$(\varphi(\textstyle\sum_{i=1}^n \beta_i(x_i - x_\infty)) : n \in \mathbb{N})$$

is a convergent, and hence bounded sequence in $\mathbb{R}$. Moreover, the statement $(*)$ gives that this is also an increasing sequence. Therefore, by the Monotone Convergence Theorem, $(\varphi(\sum_{i=1}^{n}\beta_i(x_i - x_\infty)) : n \in \mathbb{N})$ converges to its supremum. That is,

$$\sup_{n\in\mathbb{N}} \varphi(\textstyle\sum_{i=1}^{n}\beta_i(x_i - x_\infty)) = \lim_{n\to\infty} \varphi(\sum_{i=1}^{n}\beta_i(x_i - x_\infty)) = \varphi(\lim_{n\to\infty} \sum_{i=1}^{n}\beta_i(x_i - x_\infty)) = \varphi(x).$$

Part VI: Let us first note that $\|x\| \leq 2\sum_{i=1}^{\infty}\beta_i \leq m$ and so $x \in mB_Y \subseteq A$. Therefore, there exists $z^* \in \overline{\text{span}}(Z) \cap \partial\varphi(x)$. Then, if $1 < n$,

$$\begin{aligned}
\beta_n r &< \varphi(\textstyle\sum_{i=1}^{n}\beta_i(x_i - y_\infty)) - \varphi(\sum_{i=1}^{n-1}\beta_i(x_i - y_\infty)) && \text{by } (*)\\
&= \varphi(\textstyle\sum_{i=1}^{n}\beta_i(x_i - x_\infty)) - \varphi(\sum_{i=1}^{n-1}\beta_i(x_i - x_\infty)) && \text{by } (**)\\
&\leq \varphi(x) - \varphi(\textstyle\sum_{i=1}^{n-1}\beta_i(x_i - x_\infty)) && \text{since } \varphi(x) = \sup_{n\in\mathbb{N}}\varphi(\sum_{i=1}^{n}\beta_i(x_i - x_\infty))\\
&\leq z^*(x) - z^*(\textstyle\sum_{i=1}^{n-1}\beta_i(x_i - x_\infty)) && \text{since } z^* \in \partial\varphi(x)\\
&= z^*(\textstyle\sum_{i=n}^{\infty}\beta_i(x_i - x_\infty)) = \beta_n z^*(x_n - x_\infty) + \sum_{i=n+1}^{\infty}\beta_i z^*(x_i - x_\infty).
\end{aligned}$$

Rearranging gives us that

$$r < z^*(x_n - x_\infty) + \frac{1}{\beta_n}\sum_{i=n+1}^{\infty}\beta_i z^*(x_i - x_\infty) \leq z^*(x_n - x_\infty) + \frac{2\|z^*\|}{\beta_n}\sum_{i=n+1}^{\infty}\beta_i.$$

Since $\lim_{n\to\infty} \frac{1}{\beta_n}\sum_{i=n+1}^{\infty}\beta_i = 0$ there exists an $N \in \mathbb{N}$ such that

$$z^*(x_\infty) < z^*(x_\infty) + (r/2) \leq z^*(x_n) \quad \text{for all } N < n.$$

Therefore, as z^* is $\sigma(Y, Z)$-continuous on A, (see Remark 2.11) we have that

$$z^*(x_\infty) < z^*(x_\infty) + (r/2) \leq z^*(x_\infty);$$

which is impossible. Hence, our original assumption that $M(0) \not\subseteq \overline{\text{span}}(Z)$ must have been false. This completes the proof. $\square$

To show that Theorem 2.13 has some real "teeth" let us demonstrate how James' weak compactness theorem can be deduced from this theorem.

Corollary 2.14 ([7])**.** *Suppose that C is a nonempty closed and bounded convex subset of a Banach space $(X, \|\cdot\|)$ and suppose that every continuous linear functional on X attains its supremum over C. Then C is weakly compact.*

Proof. Suppose that C is a nonempty closed and bounded convex subset of a Banach space $(X, \|\cdot\|)$ and suppose that every continuous linear functional on X attains its supremum over C. Note that, without loss of generality, we may assume that $0 \in C$, because if x is any element of C then $C' := C - \{x\}$, is a closed and bounded convex set with the property that every continuous linear functional on X attains its supremum over C'. Furthermore, if C' is weakly compact then so is C.

We shall apply Theorem 2.13. To this end, let $(Y, \|\cdot\|)$ denote the dual space of $(X, \|\cdot\|)$ and let $Z := \widehat{X}$ - the natural embedding of X into X^{**}. Furthermore, let $A := Y$ and note that B_Y is $\sigma(Y,Z)$ compact, by the Banach-Alaoglu theorem (since $B_Y = B_{X^*}$ and $\sigma(Y,Z)$ is the weak* topology).

Define $\varphi : Y \to \mathbb{R}$ by,

$$\varphi(y) := \sup_{c\in C} y(c) \quad \text{for all } y \in Y.$$

Then φ is a continuous convex function on Y and $\widehat{C} \subseteq \partial\varphi(0)$. Furthermore, if $y \in Y$ attains its supremum over C then $\partial\varphi(y) \cap Z \neq \varnothing$. This last fact follows because, if $y \in Y$, $c \in C$ and $y(c) = \varphi(y)$ then $\widehat{c} \in \partial\varphi(y)$ since for any $x \in Y$,

$$\widehat{c}(x) - \widehat{c}(y) = x(c) - y(c) \leq \varphi(x) - \varphi(y).$$

Thus, by Theorem 2.13,

$$\overline{\widehat{C}}^{w^*} \subseteq \partial\varphi(0) \subseteq \widehat{X} \quad \text{since, } \partial\varphi(0) \text{ is weak}^*\text{-closed.}$$

Hence, C is weakly compact. $\square$

In the last part of this section we will show that under the assumptions of Theorem 2.13 the function φ is in fact Fréchet differentiable at the points of a dense and G_δ subset of A. This is the needed ingredient for the proof of the "Abstract Variational Theorem".

We shall proceed from here by investigating usco mappings that map from first countable topological spaces into spaces where countably compact sets are compact.

Lemma 2.15. *Let $\Psi : X \to 2^Y$ be set-valued mapping acting from a first countable space (X, τ_X) into nonempty subsets of a topological space (Y, τ_Y), and let $x_0 \in X$. Suppose that every countably compact subset of (Y, τ_Y) is compact. Then the following are equivalent:*

(i) *Ψ is a τ_Y-usco at $x_0 \in X$,*

(ii) *every sequence $(y_n : n \in \mathbb{N})$ in Y such that $y_n \in \Psi(x_n)$ for all $n \in \mathbb{N}$, for some sequence $(x_n : n \in \mathbb{N})$ in X, converging to x_0, has a τ_Y-cluster point in $\Psi(x_0)$.*

Proof. (i) $\Rightarrow$ (ii). Suppose that Ψ is a τ_Y-usco at $x_0 \in X$ and suppose that $(y_n : n \in \mathbb{N})$ is a sequence in Y and $(x_n : n \in \mathbb{N})$ is a sequence in X such that $y_n \in \Psi(x_n)$ for all $n \in \mathbb{N}$ and $x_0 = \lim_{n\to\infty} x_n$. Fix $n \in \mathbb{N}$. Then $\overline{\{y_k : n \leq k\}}^{\tau_Y} \cap \Psi(x_0) \neq \varnothing$ because otherwise there would exist an open neighbourhood N of x_0 such that $\Psi(N) \cap \overline{\{y_k : n \leq k\}}^{\tau_Y} = \varnothing$. However, this is impossible, since for k large enough (and certainly larger than n), $x_k \in N$. Thus, $\overline{\{y_k : n \leq k\}}^{\tau_Y} \cap \Psi(x_0) \neq \varnothing$. Since $\Psi(x_0)$ is τ_Y-compact, we have, by the finite intersection property, that $\bigcap_{n=1}^{\infty} \overline{\{y_k : n \leq k\}}^{\tau_Y} \cap \Psi(x_0) \neq \varnothing$. Therefore, $(y_n : n \in \mathbb{N})$ has a τ_Y-cluster point in $\Psi(x_0)$.

(ii) $\Rightarrow$ (i). By assumption $\Psi(x_0) \neq \varnothing$. We will show that $\Psi(x_0)$ is τ_Y-compact. To this end, let $(y_n : n \in \mathbb{N})$ be a sequence in $\Psi(x_0)$. Let $(x_n : n \in \mathbb{N})$ be defined by, $x_n := x_0$ for all $n \in \mathbb{N}$. Then $y_n \in \Psi(x_n)$ for all $n \in \mathbb{N}$ and $x_0 = \lim_{n\to\infty} x_n$. Therefore, $(y_n : n \in \mathbb{N})$ has a τ_Y-cluster point in $\Psi(x_0)$. Thus, by assumption, $\Psi(x_0)$ is τ-compact. Next, let W be a τ_Y-open set in Y, containing $\Psi(x_0)$, and let $(U_k : k \in \mathbb{N})$ be a decreasing sequence of τ_X-open neighbourhoods of x_0 that form a local base for the τ_X-topology at x_0. We claim that there exists an $n \in \mathbb{N}$ such that $\Psi(U_n) \subseteq W$. Indeed, if this is not the case then there exist sequences $(y_k : k \in \mathbb{N})$ in Y and $(x_k : k \in \mathbb{N})$ in X such that $x_k \in U_k$ and $y_k \in \Psi(x_k) \setminus W$ for all $k \in \mathbb{N}$. Then $x_0 = \lim_{k\to\infty} x_k$ and so $(y_k : k \in \mathbb{N})$ has a τ_Y-cluster point $y_\infty \in \Psi(x_0)$. However, this is impossible since $\{y_k : k \in \mathbb{N}\} \subseteq Y \setminus W$; which is τ_Y-closed. Therefore, $y_\infty \notin W$, which in turn implies that $y_\infty \notin \Psi(x_0)$. Hence there must exist an $n \in \mathbb{N}$ such that $\Psi(U_n) \subseteq W$. Thus, Φ is a τ_Y-usco at x_0. □

Theorem 2.16. *Let A be a nonempty open convex subset of a Banach space $(Y, \|\cdot\|)$ and let Z be a subset of Y^*. If $\varphi : A \to \mathbb{R}$ is a continuous convex function and $\partial\varphi(y) \cap \overline{span}(Z) \neq \varnothing$ for all $y \in A$, and the $\sigma(Y,Z)$-topology on B_Y is compact and Hausdorff, then $y \mapsto \partial\varphi(y)$ is a minimal weak-cusco on A.*

Proof. Let $y_0 \in A$. Since, $y \mapsto \partial\varphi(y)$, is a weak*-cusco (see Proposition 2.4) we know that $\partial\varphi$ has nonempty convex images. To show that $\partial\varphi$ is a weak-usco at y_0 we shall appeal to Lemma 2.15. So let us first note that A is metrizable and hence first countable and secondly, by the Eberlein-Šmulian Theorem, [3, page 430], every weakly countably compact subset of Y^* is weakly compact. Therefore we may apply Lemma 2.15. To this end, let $(z_n^* : n \in \mathbb{N})$ be a sequence in Y^* and let $(y_n : n \in \mathbb{N})$ be a sequence in Y such that $z_n^* \in \partial\varphi(y_n)$ for all $n \in \mathbb{N}$ and $y_0 = \lim_{n\to\infty} y_n$. Let $C := \partial\varphi(y_0) \cup \{z_n^* : n \in \mathbb{N}\}$. Then C is weak* compact and hence, by the Uniform Boundedness Theorem [3, page 66], bounded in Y^*. Moreover, by Lemma 2.15, $(z_n : n \in \mathbb{N})$ has a weak*-cluster point $z^* \in \partial\varphi(y_0) \subseteq C$. Now, by Theorem 2.13, $C \subseteq \overline{\text{span}}(Z)$ and so, by Proposition 2.10, C is weakly compact. In particular, the weak and weak* topologies must coincide on C. Therefore $z^* \in \partial\varphi(y_0)$ is also a weak-cluster point of $(z_n^* : n \in \mathbb{N})$. This shows that $\partial\varphi(y)$ is a weak-cusco at y_0. The fact that, $y \mapsto \partial\varphi(y)$, is a "minimal" weak-cusco simply follows from the fact that every weak-cusco is a weak*-cusco. Therefore, if $\Phi : A \to 2^{Y^*}$ is any weak-cusco such that $\text{Gr}(\Phi) \subseteq \text{Gr}(\partial\varphi)$. Then Φ is a weak*-cusco such that $\text{Gr}(\Phi) \subseteq \text{Gr}(\partial\varphi)$ and thus, from the minimality of $\partial\varphi$, $\Phi = \partial\varphi$. □

Corollary 2.17. *Let A be a nonempty open convex subset of a Banach space $(Y, \|\cdot\|)$ and let Z be a subset of Y^*. If $\varphi : A \to \mathbb{R}$ is a continuous convex function and $\partial\varphi(y) \cap \overline{span}(Z) \neq \varnothing$ for all $y \in A$, and the $\sigma(Y,Z)$-topology on B_Y is compact and Hausdorff, then the following are equivalent:*

(i) *$\partial\varphi(y) \cap \overline{span}(Z) \neq \varnothing$ for all $y \in A$;*

(ii) *$\partial\varphi(y) \subseteq \overline{span}(Z)$ for all $y \in A$;*

(iii) *$\partial\varphi(y) \cap \overline{span}(Z) \neq \varnothing$ for all $y \in A$ and $y \mapsto \partial\varphi(y)$ is a minimal weak-cusco on A;*

(iv) *$\partial\varphi(y) \subseteq \overline{span}(Z)$ for all $y \in A$ and $y \mapsto \partial\varphi(y)$ is a minimal weak-cusco on A.*

Proof. (i) ⇒ (ii) follows from Theorem 2.13. (ii) ⇒ (iii) follows from Corollary 2.16. (iii) ⇒ (iv) follows from Theorem 2.13. (iv) ⇒ (i) is trivial. □

For the last two results of this section we will need the following definition.

Let $\Phi : X \to 2^Y$ be a set-valued mapping acting between a topological space (X, τ) and a normed linear space $(Y, \|\cdot\|)$. The set-valued mapping Φ is said to be *single-valued and* *norm-upper semicontinuous* at $x_0 \in X$ if, $\Phi(x_0) = \{y_0\}$ is a singleton and for each $0 < \varepsilon$ there exists an open neighbourhood U of x_0, such that $\Phi(U) \subseteq B[y_0, \varepsilon]$.

Theorem 2.18 ([2, 5])**.** *Suppose that $\Phi : M \to 2^X$ is a minimal weak-cusco acting from a complete metric space (M, d) into subsets of a Banach space $(X, \|\cdot\|)$. Then Φ is single-valued and norm-upper semicontinuous at the points of a dense and G_δ subset of M.*

Corollary 2.19. *Let A be a nonempty open convex subset of a Banach space $(Y, \|\cdot\|)$ and let Z be a subset of Y^*. If $\varphi : A \to \mathbb{R}$ is a continuous convex function and $\partial\varphi(y) \cap \overline{span}(Z) \neq \varnothing$ for all $y \in A$, and the $\sigma(Y, Z)$-topology on B_Y is compact and Hausdorff, then $y \mapsto \partial\varphi(y)$ is single-valued and norm-upper semicontinuous at the points of a dense and G_δ subset of A. Equivalently, φ is Fréchet differentiable at the points of a dense and G_δ subset of A.*

Proof. This follows directly from Theorem 2.16 and Theorem 2.18. The fact that the conclusion of this result is equivalent to φ being Fréchet differentiable at the points of a dense and G_δ subset of A follows from [12, Proposition 2.8]. □

3 The Variational Theorem

Let X be any nonempty set and let $Y \subseteq \mathbb{R}^X$. Throughout this section the family Y of real-valued functions (i.e., the set of all perturbations) will be considered with pointwise addition, pointwise scalar multiplication and with a complete norm $\|\cdot\|$, whose topology on Y is at least as strong as the $\tau_p(X)$ topology on Y. We emphasise here, right at the outset, that there is at most one norm $\|\cdot\|$, up to equivalence of norm, that can be placed on Y with this property. In this way, we see that the norm on the set Y, of all perturbations, is completely determined by the set Y itself.

Proposition 3.1. *Let X be a nonempty set and let $Y \subseteq \mathbb{R}^X$. If $(Y, \|\cdot\|_1)$ and $(Y, \|\cdot\|_2)$ are both Banach spaces, under pointwise addition and pointwise scalar multiplication and the topologies of both $\|\cdot\|_1$ and $\|\cdot\|_2$ are at least as strong as the $\tau_p(X)$-topology on Y, then $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent norms. (i.e., there is at most one complete norm on Y (up to equivalence of norm), whose topology is at least as strong as $\tau_p(X)$).*

Proof. Suppose that $(Y, \|\cdot\|_1)$ and $(Y, \|\cdot\|_2)$ are Banach spaces and that the topologies of both norms $\|\cdot\|_1$ and $\|\cdot\|_2$ are at least as strong as $\tau_p(X)$. Then consider the identity mapping $I : Y \to Y$. Note that since the $\tau_p(X)$-topology on Y is Hausdorff,

$$\mathrm{Gr}(I) = \Delta_Y = \{(x, y) \in Y \times Y : x = y\}$$

is closed in $Y \times Y$, when $Y \times Y$ is equipped with the product topology generated by $\tau_p(X)$. Hence, $\mathrm{Gr}(I)$ is closed in $Y \times Y$, when $Y \times Y$ is endowed with the product topology generated by $(Y, \|\cdot\|_1)$ and $(Y, \|\cdot\|_2)$. Therefore, since $I : Y \to Y$ is linear, we have by the Closed Graph Theorem [3, page 57] that $I : (Y, \|\cdot\|_1) \to (Y, \|\cdot\|_2)$ is bounded. Of course, the same argument applies to $I : (Y, \|\cdot\|_2) \to (Y, \|\cdot\|_1)$ and so the norms $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent. This completes the proof. □

In the setting described above, it is possible to consider the canonical mapping $\widehat{\cdot} : X \to Y^*$ defined by, $\widehat{x}(y) := y(x)$ for all $y \in Y$. Note that this mapping is well-defined since for each $x \in X$, the mapping $\widehat{x} : Y \to \mathbb{R}$ is linear and $\tau_p(X)$-continuous, and hence norm continuous, since we are assuming that the norm topology on Y, is at least as strong as the $\tau_p(X)$-topology on Y.

Next, we consider the function $\rho_Y : X \times X \to \mathbb{R}$ defined by

$$\rho_Y(x, z) = \|\widehat{x} - \widehat{z}\| = \sup\{|y(x) - y(z)| : y \in B_Y\}.$$

It is not hard to see that ρ_Y is pseudo-metric on X and if Y separates the points on X then it is a metric on X. Furthermore, the topology induced by the pseudo-metric ρ_Y on X is always at least as strong as the $\sigma(X, Y)$-topology on X.

Now, we shall say that a function $f : X \to [-\infty, \infty)$ attains (or has) a *strong maximum at* $x_0 \in X$*, with respect to* ρ_Y, if

$$f(x_0) = \sup_{x \in X} f(x) \quad \text{and} \quad \lim_{n\to\infty} \rho_Y(x_n, x_0) = 0$$

whenever $(x_n : n \in \mathbb{N})$ is a sequence in X such that

$$\lim_{n\to\infty} f(x_n) = \sup_{x \in X} f(x) = f(x_0).$$

In order to expedite our results on maxima we shall utilise the theory of conjugates, but first let us recall that: (i) for any function $f : X \to \mathbb{R} \cup \{\infty\}$ defined on a nonempty set X

$$\text{Dom}(f) := \{x \in X : f(x) < \infty\}$$

and (ii) a function $f : X \to \mathbb{R} \cup \{\infty\}$ is said to be a *proper function* if $\text{Dom}(f) \neq \varnothing$.

Let X be an arbitrary set and $Y \subseteq \mathbb{R}^X$. For a proper function $f : X \to \mathbb{R} \cup \{\infty\}$, *the conjugate*, $f_Y^* : Y \to \mathbb{R} \cup \{\infty\}$ is defined by,

$$f_Y^*(y) := \sup_{x \in X} (y - f)(x) = \sup_{x \in \text{Dom}(f)} (y - f)(x).$$

Our results on conjugate mappings depend upon the notion of ε-derivatives. So next we give the definition of an ε-subderivative.

Suppose that $f : X \to \mathbb{R} \cup \{\infty\}$ is a proper function on a normed linear space $(X, \|\cdot\|)$ and $x \in \text{Dom}(f)$. Then, for any $0 < \varepsilon$, we define the *ε-subdifferential* $\partial_\varepsilon f(x)$ by,

$$\partial_\varepsilon f(x) := \{x^* \in X^* : x^*(y) - x^*(x) \leq f(y) - f(x) + \varepsilon \text{ for all } y \in \text{Dom}(f)\}.$$

Before we prove the main proposition for this section we first need to recall that bounded above convex functions, are in fact, locally Lipschitz.

Proposition 3.2 ([12, Proposition 1.6])**.** *Let A be a nonempty open convex subset of a Banach space $(X, \|\cdot\|)$ and let $\varphi : A \to \mathbb{R}$ be a convex function. If φ is locally bounded above on A, that is, for every $x_0 \in A$ there exists an $0 < M$ and a $0 < \delta$ such that $B(x_0, \delta) \subseteq A$ and $\varphi(x) \leq M$ for all $x \in B(x_0, \delta)$, then it is locally Lipschitz on A; that is, for every $x_0 \in A$, there exists an $0 < L$ and a $0 < \delta$ such that $B(x_0, \delta) \subseteq A$ and $|\varphi(x) - \varphi(y)| \leq L\|x - y\|$ for all $x, y \in B(x_0, \delta)$.*

For any function $f : X \to \mathbb{R} \cup \{-\infty\}$ defined on a nonempty set X we define

$$\operatorname{argmax}(f) := \{x \in X : f(y) \leq f(x) \text{ for all } y \in X\}.$$

Proposition 3.3. *Let $f : X \to \mathbb{R} \cup \{\infty\}$ be a proper function defined on a nonempty set X and let $(Y, \|\cdot\|)$ be a Banach space of real-valued functions defined on X. Suppose that the norm-topology on Y is at least as strong as the $\tau_p(X)$-topology on Y. Then the conjugate f_Y^* satisfies the following properties:*

- (i) *f_Y^* is a convex and $\tau_p(X)$-lower semicontinuous function on $Dom(f_Y^*)$;*
- (ii) *f_Y^* is continuous on $int(Dom(f_Y^*))$;*
- (iii) *If $y \in Dom(f_Y^*)$ and $x \in \operatorname{argmax}(y - f)$ then $\widehat{x} \in \partial f_Y^*(y)$;*
- (iv) *If $0 < \varepsilon$, $y \in Dom(f_Y^*)$, $x \in X$ and $f_Y^*(y) - \varepsilon < y(x) - f(x)$ then $\widehat{x} \in \partial_\varepsilon f_Y^*(y)$;*
- (v) *If $y_0 \in int(Dom(f_Y^*))$, $x_0 \in \operatorname{argmax}(y_0 - f)$ and $y \mapsto \partial f_Y^*(y)$ is single-valued and norm-upper semicontinuous at y_0, then $y_0 - f$ has a strong maximum at x_0, with respect to ρ_Y.*

Proof. For those people familiar with the Fenchel conjugate, they may wish to skip the proofs of (i)-(iv).

(i) For each $x \in \mathrm{Dom}(f)$ define $g_x : Y \to \mathbb{R}$ by, $g_x(y) := \widehat{x}(y) - f(x)$. Then each function g_x is $\tau_p(X)$-continuous and affine. Now for each $y \in Y$,

$$f_Y^*(y) = \sup_{x \in \mathrm{Dom}(f)} g_x(y).$$

Thus, as the pointwise supremum of a family of $\tau_p(X)$-continuous affine mappings, the conjugate of f, is itself convex and $\tau_p(X)$-lower semicontinuous. [Recall the general fact that the pointwise supremum of a family of convex functions is convex and the pointwise supremum of a family of lower semicontinuous mappings is again lower semicontinuous].

(ii) Since this statement is vacuously true when $\mathrm{int}(\mathrm{Dom}(f_Y^*)) = \varnothing$, we will assume that $\mathrm{int}(\mathrm{Dom}(f_Y^*))$ is nonempty. Let us first recall that by Proposition 3.2, the fact that f_Y^* is convex and the fact that $\mathrm{int}(\mathrm{Dom}(f_Y^*))$ is also convex, it is sufficient to show that f_Y^* is locally bounded above on $\mathrm{int}(\mathrm{Dom}(f_Y^*))$. In fact, as we shall now show, it is sufficient to show that f_Y^* is locally bounded above at a single point $y_0 \in \mathrm{int}(\mathrm{Dom}(f_Y^*))$. To this end, suppose that f_Y^* is locally bounded above at $y_0 \in \mathrm{int}(\mathrm{Dom}(f_Y^*))$. Then there exist an $0 < M$ and a $0 < \delta$ such that $f_Y^*(y) \leq M$ for all $y \in B[y_0, \delta]$. Let x be any point in $\mathrm{int}(\mathrm{Dom}(f_Y^*))$. Since $\mathrm{int}(\mathrm{Dom}(f_Y^*))$ is an open convex set, there exists a point $y \in \mathrm{int}(\mathrm{Dom}(f_Y^*))$ and a $0 < \lambda < 1$ such that $x = \lambda y + (1-\lambda)y_0$. Let $M^* := \max\{M, f_Y^*(y)\}$ and note that

$$x \in B[x, (1-\lambda)\delta] = \lambda y + (1-\lambda)B[y_0, \delta] \subseteq \mathrm{int}(\mathrm{Dom}(f_Y^*)), \quad \text{since } \mathrm{int}(\mathrm{Dom}(f_Y^*)) \text{ is convex.}$$

We claim that f_Y^* is bounded above by M^* on $B[x, (1-\lambda)\delta]$. To see this, let z be any element of $B[x, (1-\lambda)\delta]$. Then $z = \lambda y + (1-\lambda)w$ for some $w \in B[y_0, \delta]$ since,

$$B[x, (1-\lambda)\delta] = x + (1-\lambda)B[0, \delta] = \lambda y + (1-\lambda)y_0 + (1-\lambda)B[0, \delta] = \lambda y + (1-\lambda)B[y_0, \delta].$$

Therefore,

$$f_Y^*(z) = f_Y^*(\lambda y + (1-\lambda)w) \leq \lambda f_Y^*(y) + (1-\lambda) f_Y^*(w) \leq \lambda M^* + (1-\lambda)M^* = M^*.$$

Next, we will use the fact that since $\mathrm{int}(\mathrm{Dom}(f_Y^*))$ is a nonempty open subset of a complete metric space, it is itself a Baire space with the relative topology. Now, for each $n \in \mathbb{N}$, let

$$F_n := \{x \in \mathrm{int}(\mathrm{Dom}(f_Y^*)) : f_Y^*(x) \leq n\}.$$

Since f_Y^* is $\tau_p(X)$-lower semicontinuous, it is lower semicontinuous with respect to the norm topology too. Therefore, each set F_n is closed with respect to the relative norm topology on $\mathrm{int}(\mathrm{Dom}(f_Y^*))$. Since $\mathrm{int}(\mathrm{Dom}(f_Y^*)) = \bigcup_{n\in\mathbb{N}} F_n$, there exists an $n_0 \in \mathbb{N}$ such that $\mathrm{int}(F_{n_0}) \neq \varnothing$. Hence, f_Y^* is locally bounded above at each point of $\mathrm{int}(F_{n_0})$. This completes the proof of part (ii).

(iii) Let z be any element of $\mathrm{Dom}(f_Y^*)$. Then,

$$\begin{aligned} \widehat{x}(z) - \widehat{x}(y) &= z(x) - y(x) \\ &= [z(x) - f(x)] - [y(x) - f(x)] \\ &= [z(x) - f(x)] - f_Y^*(y) \leq f_Y^*(z) - f_Y^*(y). \end{aligned}$$

Therefore, $\widehat{x} \in \partial f_Y^*(y)$.

(iv) Let z be any element of $\mathrm{Dom}(f_Y^*)$. Then,

$$\begin{aligned} \widehat{x}(z) - \widehat{x}(y) &= z(x) - y(x) = [z(x) - f(x)] - [y(x) - f(x)] \\ &\leq [z(x) - f(x)] - [f_Y^*(y) - \varepsilon] \leq f_Y^*(z) - f_Y^*(y) + \varepsilon. \end{aligned}$$

Therefore, $\widehat{x} \in \partial_\varepsilon f_Y^*(y)$.

(v) Let $(x_n : n \in \mathbb{N})$ be a sequence in X such that

$$\lim_{n\to\infty} (y_0 - f)(x_n) = \sup_{x\in X} (y_0 - f)(x) = f_Y^*(y_0).$$

We will show that $(x_n : n \in \mathbb{N})$ converges to x_0, with respect to the metric ρ_Y. Let $0 < \varepsilon$ and let $0 < \varepsilon' < \varepsilon/2$. By (iii) and the assumption that $\partial f_Y^*(y_0)$ is a singleton we have that $\partial f_Y^*(y_0) = \{\widehat{x_0}\}$. Since, $y \mapsto \partial f_Y^*(y)$, is norm-upper semicontinuous at y_0 there exists a $0 < \delta$ such that if $\|y - y_0\| \leq \delta$ then $\|F - \widehat{x_0}\| < \varepsilon'$ for all $F \in \partial f_Y^*(y)$. Choose $N \in \mathbb{N}$ such that $f_Y^*(y_0) - \varepsilon'\delta < (y_0 - f)(x_n)$ for all $N < n$. Then, by (iv), $\widehat{x_n} \in \partial_{\varepsilon'\delta} f_Y^*(y_0)$ for all $N < n$. Fix $N < n$ and let $v \in S_Y$. Then

$$\widehat{x_n}(\delta v) - \widehat{x_0}(\delta v) \leq [f_Y^*(y_0 + \delta v) - f_Y^*(y_0) + \varepsilon'\delta] - \widehat{x_0}(\delta v). \qquad (*)$$

Let $F \in \partial f_Y^*(y_0 + \delta v)$ then,

$$F(-\delta v) \leq f_Y^*(y_0) - f_Y^*(y_0 + \delta v)$$

or, equivalently,

$$f_Y^*(y_0 + \delta v) - f_Y^*(y_0) \leq F(\delta v). \qquad (**)$$

Substituting the inequality $(**)$ into the inequality $(*)$ we get that

$$\widehat{x_n}(\delta v) - \widehat{x_0}(\delta v) \leq (F - \widehat{x_0})(\delta v) + \varepsilon'\delta \leq \|F - \widehat{x_0}\| \cdot \delta + \varepsilon'\delta < \varepsilon'\delta + \varepsilon'\delta.$$

Hence, $(\widehat{x_n} - \widehat{x_0})(v) = \widehat{x_n}(v) - \widehat{x_0}(v) \leq 2\varepsilon' < \varepsilon$. Since $v \in S_Y$ was arbitrary, $\|\widehat{x_n} - \widehat{x_0}\| < \varepsilon$ and so $\rho_Y(x_n, x_0) = \|\widehat{x_n} - \widehat{x_0}\| < \varepsilon$.

This completes the proof. □

Theorem 3.4 (Abstract Variation Theorem)**.** *Let $f : X \to \mathbb{R} \cup \{\infty\}$ be a proper function defined on a nonempty set X and let $(Y, \|\cdot\|)$ be a Banach space of real-valued functions defined on X. Suppose that the norm-topology on Y is at least as strong as the $\tau_p(X)$-topology on Y and that B_Y is $\tau_p(X)$-compact. If there exists a nonempty open subset of A of $Dom(f_Y^*)$ such that $argmax(y - f) \neq \varnothing$ for each $y \in A$, then there exists a dense and G_δ subset of R of A such that $(y - f) : X \to \mathbb{R} \cup \{-\infty\}$ has a strong maximum, with respect to ρ_Y, for each $y \in R$.*

Proof. Let $Z := \widehat{X} \subseteq Y^*$. Then $\sigma(Y, Z) = \tau_p(X)$ and so the $\sigma(Y, Z)$-topology on B_Y is compact and Hausdorff. Furthermore, by parts (i) and (ii) of Proposition 3.3, it follows that $f_Y^*|_A : A \to \mathbb{R}$ is a continuous convex function on A. Let us also note that by Proposition 3.3 part (iii) that $\partial(f_Y^*|_A)(y) \cap Z \neq \varnothing$ for all $y \in A$. Hence, by Corollary 2.19, $y \mapsto \partial(f_Y^*|_A)(y)$ is single-valued and norm-upper semicontinuous at the points of a dense and G_δ subset R of A. The result now follows from Proposition 3.3 part (v). □

In order to see the utility of Theorem 3.4 we need to see that the hypotheses of this theorem are indeed satisfied by some Banach spaces. So below we give some obvious examples of Banach spaces where all the assumptions of Theorem 3.4, (i.e., the norm-topology on Y is at least as strong as the $\tau_p(X)$-topology on Y and the ball B_Y is compact with respect to the $\tau_p(X)$-topology) are satisfied. We emphasise that this is not an exhaustive list, but rather, only a small collection of examples.

Example 3.5.

(i) *Let $(X, \|\cdot\|)$ be a normed space and let $(Y, \|\cdot\|)$ be the dual of X endowed with with the dual norm. Note that in this example, by the Hahn-Banach theorem, $\rho_Y(x, y) = \|x - y\|$.*

(ii) *Let (X, d) be a metric space, with a distinguished point x_0. Let $Y := \mathrm{Lip}_0(X)$ be the set of all (possibly unbounded) Lipschitz functions from X to $\mathbb{R}$ that vanish at x_0, endowed with the Lipschitz norm*

$$\|f\|_L := \sup\left\{\frac{|f(x) - f(y)|}{d(x, y)} : x \neq y\right\}.$$

Note that in this example, by a simple calculation, $\rho_Y(x, y) = d(x, y)$.

(iii) *Let X be a nonempty set and let $Y := B(X)$ be the set of all bounded real-valued functions endowed with supremum norm $\|\cdot\|_\infty$.*

(iv) *Let $X := [a, b]$ and $Y := BV[a.b]$ be the Banach space of real-valued functions with bounded variation, endowed with the bounded variation norm*

$$\|f\|_{BV} := \sup\{\textstyle\sum_{i=1}^{n} |f(x_i) - f(x_{i-1})| : \{x_i : i = 0, \ldots, n\} \text{ is a partition on } [a, b]\} + |f(a)|.$$

(v) *Let (X, d) be a metric space and let $Y := BL(X)$ be the set of all real-valued bounded and Lipschitz functions on X, endowed with the norm $\|\cdot\|_{BL} := \|\cdot\|_L + \|\cdot\|_\infty$.*

From Theorem 3.4 and Example 3.5 part(i) we obtain the following notable special case of the Abstract Variational Theorem.

Corollary 3.6. *Let $f : X \to \mathbb{R}\cup\{\infty\}$ be a proper function on a normed linear space $(X, \|\cdot\|)$. If there exists a nonempty open subset A of $Dom(f^*_{X^*})$ such that $argmax(x^* - f) \neq \varnothing$ for each $x^* \in A$, then there exists a dense and G_δ subset R of A such that $(x^* - f) : X \to \mathbb{R}\cup\{-\infty\}$ has a strong maximum, with respect to the norm-topology, for each $x^* \in R$. In addition, if $0 \in A$ and $0 < \varepsilon$ then there exists an $x^*_0 \in X^*$ with $\|x^*_0\| < \varepsilon$ such that $(x^*_0 - f) : X \to \mathbb{R}\cup\{-\infty\}$ has a strong maximum, with respect to the norm-topology.*

If we impose some stronger assumptions upon the function $f : X \to \mathbb{R}\cup\{\infty\}$ then we may extract more information about the mapping, $x^* \mapsto \text{argmax}(x^* - f)$.

For example, if $f : X \to \mathbb{R}\cup\{\infty\}$ is a proper convex lower semicontinuous function defined on a Banach space $(X, \|\cdot\|)$, $x^* \in \text{Dom}(f^*_{X^*})$ and $(x_n : n \in \mathbb{N})$ is any sequence in X such that

$$\lim_{n\to\infty}(x^* - f)(x_n) = \sup_{x\in X}(x^* - f)$$

then $cl_{\text{weak}}(x_n : n \in \mathbb{N}) \subseteq \text{argmax}(x^* - f)$, i.e., all the weak-cluster points of $(x_n : n \in \mathbb{N})$ lie in $\text{argmax}(x^* - f)$.

Corollary 3.7. *Let $f : X \to \mathbb{R}\cup\{\infty\}$ be a proper convex lower semicontinuous function defined on a Banach space $(X, \|\cdot\|)$. If there exists a nonempty open subset A of $Dom(f^*_{X^*})$ such that $argmax(x^* - f) \neq \varnothing$ for each $x^* \in A$, then the mapping, $x^* \mapsto argmax(x^* - f)$, is a minimal weak-cusco on A.*

Proof. Let $x^*_0 \in A$. By assumption we know that, $x^* \mapsto \text{argmax}(x^* - f)$, has nonempty convex images. To show that $x^* \mapsto \text{argmax}(x^* - f)$ is a weak-usco at x^*_0 we shall appeal to Lemma 2.15. So let us first note that A is metrizable and hence first countable and secondly, by the Eberlein-Šmulian Theorem [3, page 430], every weakly countably compact subset of X is weakly compact. Therefore, we may apply Lemma 2.15. To this end, let $(x_n : n \in \mathbb{N})$ be a sequence in X and let $(x^*_n : n \in \mathbb{N})$ be a sequence in A such that $x_n \in \text{argmax}(x^*_n - f)$ for all $n \in \mathbb{N}$ and $x^*_0 = \lim_{n\to\infty} x^*_n$. By Proposition 3.3 part (iii), $\widehat{x_n} \in \partial f^*_{X^*}(x^*_n)$ for all $n \in \mathbb{N}$.

Let $Y := X^*$ and let $Z := \widehat{X} \subseteq X^{**}$. Then $\sigma(Y, Z) = \text{weak}^*$ and so the $\sigma(Y, Z)$-topology on B_Y is compact and Hausdorff. Furthermore, by parts (i) and (ii) of Proposition 3.3, it follows that $f^*_Y|_A : A \to \mathbb{R}$ is a continuous convex function on A. Let us also note that by Proposition 3.3 part (iii) that $\partial(f^*_Y|_A)(x^*) \cap Z \neq \varnothing$ for all $x^* \in A$. Hence, by Corollary 2.17, $x^* \mapsto \partial f^*_{X^*}(x^*)$ is a minimal weak-cusco from A into $2^{\widehat{X}}$. Therefore, by Lemma 2.15, $(\widehat{x_n} : n \in \mathbb{N})$ has a weak-cluster point $\widehat{x}_\infty$ in $\partial f^*_{X^*}(x^*_0) \subseteq \widehat{X}$. Thus, it follows, that x_∞ is a weak-cluster point of $(x_n : n \in \mathbb{N})$. We now need to show that $x_\infty \in \text{argmax}(x^*_0 - f)$. Fix $n \in \mathbb{N}$. Then,

$$\begin{aligned} f^*_{X^*}(x^*_n) - \|x^*_0 - x^*_n\|\|x_n\| &\leq (x^*_n - f)(x_n) + (x^*_0 - x^*_n)(x_n) \\ &= (x^*_0 - f)(x_n) \\ &\leq f^*_{X^*}(x^*_0). \qquad (*) \end{aligned}$$

Now, $(x_n : n \in \mathbb{N})$ is bounded, since (i) $\|x_n\| = \|\widehat{x_n}\|$ and $\widehat{x_n} \in \partial f^*_{X^*}(x^*_n)$ for all $n \in \mathbb{N}$ and (ii) $x^* \mapsto \partial f^*_{X^*}(x^*)$ is locally bounded, by Proposition 2.5. Furthermore, by Proposition 3.3 part (ii), $x^* \mapsto f^*_{X^*}(x^*)$, is continuous at x^*_0. Therefore, by the Squeeze Theorem, applied to $(*)$, we get that

$$\lim_{n\to\infty} (x^*_0 - f)(x_n) = f^*_{X^*}(x^*_0).$$

Thus, by the observation made before this corollary, we have that $x_\infty \in \operatorname{argmax}(x^*_0 - f)$. This shows that, $x^* \mapsto \operatorname{argmax}(x^* - f)$, is a weak-usco at x^*_0. Since $x_0 \in A$ was arbitrary, we conclude that, $x^* \mapsto \operatorname{argmax}(x^* - f)$, is a weak-cusco on A. The fact that, $x^* \mapsto \operatorname{argmax}(x^* - f)$, is a "minimal" weak-cusco follows from the following argument. Suppose that $\Phi : A \to 2^X$ is a weak-cusco such that $\Phi(x^*) \subseteq \operatorname{argmax}(x^* - f)$ for all $x^* \in A$. Let $\widehat{\Phi} : A \to 2^{\widehat{X}}$ be defined by, $\widehat{\Phi}(x^*) := \widehat{\Phi(x^*)}$ for all $x^* \in A$ and $\Psi : A \to 2^{\widehat{X}}$ be defined by, $\Psi(x^*) := \{\widehat{x} : x \in \operatorname{argmax}(x^* - f)\}$ for all $x^* \in A$. Then, by Proposition 3.3 part (iii), $\widehat{\Phi}(x^*) \subseteq \Psi(x^*) \subseteq \partial f^*_{X^*}(x^*)$ for all $x^* \in A$. Further, $\widehat{\Phi}$ is a weak-cusco (and hence a weak*-cusco). Since, $x^* \mapsto \partial f^*_{X^*}(x^*)$, is a minimal weak*-cusco on A (see, Proposition 2.4) we must have that $\widehat{\Phi}(x^*) = \Psi(x^*) = \partial f^*_{X^*}(x^*)$ for all $x^* \in A$. This then implies that $\Phi(x^*) = \operatorname{argmax}(x^* - f)$ for all $x^* \in A$. That is, $x^* \mapsto \operatorname{argmax}(x^* - f)$, is a minimal weak-cusco on A. $\square$

References

[1] J. M. Borwein, W. B. Moors, and X. Wang, *Generalized subdifferentials: a Baire categorical approach*, Trans. Amer. Math. Soc. **353** (2001), no. 10, 3875–3893. MR 1837212

[2] J. P. R. Christensen, *Theorems of Namioka and R. E. Johnson type for upper semicontinuous and compact valued set-valued mappings*, Proc. Amer. Math. Soc. **86** (1982), no. 4, 649–655.

[3] N. Dunford and J. T. Schwartz, *Linear Operators. I. General Theory*, Interscience Publishers, Inc., New York; Interscience Publishers, Ltd., London, 1958.

[4] M. J. Fabian, *Gâteaux differentiability of convex functions and topology*, Canadian Mathematical Society Series of Monographs and Advanced Texts, John Wiley & Sons, Inc., New York, 1997, Weak Asplund spaces, A Wiley-Interscience Publication. MR 1461271

[5] J. R. Giles, P. S. Kenderov, W. B. Moors, and S. D. Sciffer, *Generic differentiability of convex functions on the dual of a Banach space*, Pacific J. Math. **172** (1996), no. 2, 413–431.

[6] A. Grothendieck, *Critères de compacité dans les espaces fonctionnels généraux*, Amer. J. Math. **74** (1952), 168–186. MR 0047313

[7] R. C. James, *Weakly compact sets*, Trans. Amer. Math. Soc. **113** (1964), 129–140.

[8] L. Jokl, *Minimal convex-valued weak* USCO correspondences and the Radon-Nikodým property*, Comment. Math. Univ. Carolin. **28** (1987), no. 2, 353–376.

[9] M. Krein, *Sur quelques questions de la géométrie des ensembles convexes situes dans un espace linéaire norme et complet.*, C. R. (Dokl.) Acad. Sci. URSS, n. Ser. **14** (1937), 5–7.

[10] W. B. Moors, *A gentle introduction to James' weak compactness theorem and beyond*, Methods Funct. Anal. Topology. - to appear.

[11] W. B. Moors and S. J. White, *An elementary proof of James' characterisation of weak compactness. II*, Bull. Aust. Math. Soc. **95** (2017), no. 1, 133–137.

[12] R. R. Phelps, *Convex functions, monotone operators and differentiability*, Lecture Notes in Mathematics, vol. 1364, Springer-Verlag, Berlin, 1993.

[13] R. S. Phillips, *On weakly compact subsets of a Banach space*, Amer. J. Math. **65** (1943), 108–136. MR 0007938

Corresponding author: Warren B. Moors.

Email: w.moors@auckland.ac.nz

Full mailing address: Warren B. Moors, Department of Mathematics
The University of Auckland, Private Bag 92019, Auckland Mail Centre
Auckland 1142, NEW ZEALAND.

Second author: Neşet Özkan Tan

Email: neset.tan@auckland.ac.nz

Full mailing address: Neşet Özkan Tan, Department of Mathematics
The University of Auckland, Private Bag 92019, Auckland Mail Centre
Auckland 1142, NEW ZEALAND.
