A Gentle Introduction to
James' Weak Compactness Theorem
and Beyond

Warren B. Moors[1]

Abstract. The purpose of this paper is twofold: firstly, to provide an accessible proof of James' weak compactness theorem that is able to be taught in a first-year graduate class in functional analysis and secondly, to explore some of the latest and possible future extensions and applications of James' theorem.

1   Introduction

The first part of this paper is devoted to providing a proof of James' weak compactness theorem that is able to be taught in a first-year graduate class in functional analysis. Usually when one teaches a first course in functional analysis one teaches the basic finite dimensional material, Hilbert space material, the open mapping theorem, the closed graph theorem, the uniform boundedness theorem and the Hahn-Banach theorem, plus applications. Then one might consider the spectral theory of compact normal operators, or even, an introduction to C^*-algebras. However, what is often neglected is the study of linear topology, which then makes it difficult to even start to contemplate how one might prove James' theorem on weak compactness. So, what we propose in the first part of this paper is a way of presenting James' theorem on weak compactness to an audience unfamiliar with linear topology, or anything other than, the most basic facts concerning normed linear spaces. Later, in the second half of Section 3, and thereafter, we present some of the latest advances/extensions and applications of James' theorem.

For the author, James' theorem on weak compactness is one of the true delights of functional analysis. Its proof is a beautiful synthesis of linear algebra and topology. The one down-side of this theorem is that its proof has an unfortunate reputation of being very difficult. We hope, among other things, to dispel this myth.

We shall start with a brief history of this problem. Back in 1933 (see, [32]) S. Mazur conjectured that a Banach space $(X, \|\cdot\|)$, over the real numbers, is reflexive if, and only if, every continuous linear functional defined on X attains its maximum value on the closed unit ball of X. In 1957 (see, [19]), R. James confirmed this conjecture for separable Banach spaces, i.e., those spaces that contain a countable dense subset. Later, in 1963 (see, [20]), R. James completely confirmed the conjecture for arbitrary Banach spaces. One year after this, in [21], R. James extended this result to show that a closed and bounded convex subset C of a Banach space X is weakly compact if, and only if, every continuous linear functional defined on X attains its maximum value over C. The fact that this result does not extend to non-complete normed linear spaces was established in [22], again by James. Almost immediately, even in 1965 (see, [44]), there was a search for a simpler proof of James' weak compactness theorem. The proof in [44] is indeed very clear and easy to read, and is in fact the basis of a lot of the work in this paper. However, [44] still contains a series of seven technical lemmas. In 1972, R. James (see, [23]) provided a simpler proof of his own weak compactness theorem, and in [48], S. Simons, using an inequality that now bears his name, proved the weak compactness theorem for separable Banach spaces. Since these early results there have been many attempts at providing a simple proof of James' theorem. Most of these require additional assumptions on the space. One approach which is quite appealing is that of (I)-generation. This first appeared in [14] and then again in [13], but it has since been shown (see, [26]) that this approach is essentially equivalent to the approach of S. Simons

---

[1]See back page for contact details

from 1972. In addition to the already mentioned papers the interested reader may also want to see the papers [5, 17, 27, 34, 37, 38], where several "simple" proofs of James' theorem are given. The paper [5] also has some interesting applications and historical facts.

We now return to the mathematics. In this paper all vector spaces and all normed linear spaces will be over the field of real numbers. The key concept, which runs throughout this paper, is the notion of a convex set. A subset C of a vector space $(V, +, \cdot)$, over the real numbers, is called *convex* if, for every pair of points $x, y \in C$ and $0 < \lambda < 1$, we have $\lambda x + (1 - \lambda)y \in C$. We encourage the reader to follow the role that this concept plays throughout the rest of this paper.

The structure of the reminder of this paper is as follows: Section 2 contains the necessary background material. In particular, it contains Subsections; 2.1 on Weak topologies, 2.2 on Linear topology, 2.3 on the Hahn-Banach Theorem, 2.4 on the Weak* topology. Readers with a background in linear topology may wish to skip this section. Section 3 contains three proofs of James' theorem given in three subsections; 3.1 on James' theorem for separable Banach spaces; 3.2 on James' theorem for spaces with a weak* sequentially compact dual ball, 3.3 the general version of James' theorem. In Subsection 3.4 some applications of James' theorem are given. In Section 4 a generalisation of James' theorem is given. To achieve this, this section contains Subsection 4.1 which gives the necessary background in convex analysis, then in Subsection 4.2 the necessary set-valued analysis is given. In Subsection 4.3 the generalisation of James' theorem is presented. Finally, in Section 5, we give a variational principle that is based upon the generalised version of James' theorem. The paper ends with an index of notation and assumed knowledge and a bibliography.

2 Preliminaries

In this section of the paper we will present the necessary background material that is required in order to prove James' theorem on weak compactness of closed and bounded convex subsets of a given Banach space.

2.1 Weak topologies on sets

An important part of general topology concerns the generation of topologies on a given set. In this subsection we will show how to construct topologies that make a given function (set of functions) continuous.

Proposition 2.1. *Let $f : X \to Y$ be a function between sets X and Y. If τ_Y is a topology on Y then $\tau_X := \{f^{-1}(U) : U \in \tau_Y\}$ is a topology on X and $f : (X, \tau_X) \to (Y, \tau_Y)$ is continuous. Furthermore, if τ is any topology on X such that $f : (X, \tau) \to (X, \tau_Y)$ is continuous then $\tau_X \subseteq \tau$. This is, τ_X is the weakest topology on X that makes f continuous (when Y is endowed with the topology τ_Y).*

Proof. First we will show that τ_X is a topology on X. Now, $\varnothing \in \tau_X$ since $\varnothing = f^{-1}(\varnothing)$ and $\varnothing \in \tau_Y$. Similarly, $X \in \tau_X$ since $X = f^{-1}(Y)$ and $Y \in \tau_Y$. Next, suppose that $V_1 \in \tau_X$ and $V_2 \in \tau_2$. Then, by the definition of τ_X, there exists $U_1 \in \tau_Y$ and $U_2 \in \tau_Y$ such that $V_1 = f^{-1}(U_1)$ and $V_2 = f^{-1}(U_2)$. Therefore,

$$V_1 \cap V_2 = f^{-1}(U_1) \cap f^{-1}(U_2) = f^{-1}(U_1 \cap U_2).$$

Since τ_Y is a topology on Y, $U_1 \cap U_2 \in \tau_Y$. Hence, $V_1 \cap V_2 \in \tau_X$. Finally, to show that τ_X is a topology on X, suppose that $\{V_\alpha : \alpha \in A\} \subseteq \tau_X$. Then, by the definition of τ_X, there exist $\{U_\alpha : \alpha \in A\} \subseteq \tau_Y$ such that $V_\alpha = f^{-1}(U_\alpha)$ for each $\alpha \in A$. Therefore,

$$\bigcup_{\alpha \in A} V_\alpha = \bigcup_{\alpha \in A} f^{-1}(U_\alpha) = f^{-1}\left(\bigcup_{\alpha \in A} U_\alpha\right).$$

Since τ_Y is a topology on Y, $\bigcup_{\alpha \in A} U_\alpha \in \tau_Y$. Hence, $\bigcup_{\alpha \in A} V_\alpha \in \tau_X$. Thus, τ_X is a topology on X. To show that $f : (X, \tau_X) \to (Y, \tau_Y)$ is continuous we consider the following. Let $U \in \tau_Y$. Then $f^{-1}(U) \in \tau_X$, by the definition of τ_X. Therefore, by the definition of continuity, $f : (X, \tau_X) \to (Y, \tau_Y)$ is continuous. For our last step of the proof, we will show that τ_X is the weakest topology on X that makes f continuous. To this end, let τ be any topology on X such that $f : (X, \tau) \to (Y, \tau_Y)$ is continuous. Let $V \in \tau_X$. Then, by the definition of τ_X, there exists an $U \in \tau_Y$ such that $V = f^{-1}(U)$. Since we are assuming that $f : (X, \tau) \to (Y, \tau_Y)$ is continuous, $V = f^{-1}(U) \in \tau$. Thus, $\tau_X \subseteq \tau$. This completes the proof. $\square$

The topology τ_X in Proposition 2.1 is called the *weak topology on X generated by f and τ_Y*, or more briefly, when the context is clear, the *weak topology on X*.

When we have more than one function we still have the following result.

Proposition 2.2. *Let X and Y be sets and let τ_Y be a topology on Y. If $\mathcal{F}$ is a nonempty family of functions from X into Y then*

$$\mathcal{B} := \left\{ \bigcap_{1 \le k \le n} f_k^{-1}(U_k) : n \in \mathbb{N}, U_k \in \tau_Y \text{ and } f_k \in \mathcal{F} \right\}$$

is a base for a topology τ_X on X. Furthermore, the topology τ_X is the weakest topology on X that make each $f \in \mathcal{F}$ continuous, when Y is endowed with the topology τ_Y.

Proof. Firstly, it is easy to see that $\varnothing$ and X are members of $\mathcal{B}$. Indeed, since $\mathcal{F} \ne \varnothing$ we may take a function $f \in \mathcal{F}$. Then $\varnothing = f^{-1}(\varnothing)$ and so $\varnothing \in \mathcal{B}$ since $\varnothing \in \tau_Y$. Similarly, $X = f^{-1}(Y)$ and so $X \in \mathcal{B}$ since $Y \in \tau_Y$. Next, let us observe that $\mathcal{B}$ is closed under taking finite intersections. Suppose $V_1 \in \mathcal{B}$ and $V_2 \in \mathcal{B}$. Then there exists $n_1 \in \mathbb{N}$, $f_k' \in \mathcal{F}$ and $U_k' \in \tau_Y$ for each $1 \le k \le n_1$ such that $V_1 = \bigcap_{1 \le k \le n_1} (f_k')^{-1}(U_k')$. Similarly, there exists $n_2 \in \mathbb{N}$, $f_k'' \in \mathcal{F}$ and $U_k'' \in \tau_Y$ for each $1 \le k \le n_2$ such that $V_2 = \bigcap_{1 \le k \le n_2} (f_k'')^{-1}(U_k'')$. Let $n := n_1 + n_2$ and for each $1 \le k \le n_1$ let $U_k := U_k'$ and for each $1 \le k \le n_2$ let $U_{n_1+k} := U_k''$. For each $1 \le k \le n_1$ let $f_k := f_k'$ and for each $1 \le k \le n_2$ let $f_{n_1+k} := f_k''$. Then,

$$V_1 \cap V_2 = \bigcap_{1 \le k \le n_1} (f_k')^{-1}(U_k') \cap \bigcap_{1 \le k \le n_2} (f_k'')^{-1}(U_k'') = \bigcap_{1 \le k \le n} f_k^{-1}(U_k) \in \mathcal{B}.$$

We now define τ_X to be the set of all subsets of X that can be expressed as a union of members of $\mathcal{B}$. From above we see that $\varnothing$ and X are members of τ_X, and τ_X is closed under taking finite intersections. For the details of this last claim consider the following. Let $V_1 \in \tau_X$ and $V_2 \in \tau_X$. Then there exist disjoint sets I_1 and I_2 such that $V_1 = \bigcup_{i \in I_1} B_i$ for some $B_i \in \mathcal{B}$ and $V_2 = \bigcup_{i \in I_2} B_i$ for some $B_i \in \mathcal{B}$. Let $I := I_1 \times I_2$ then

$$V_1 \cap V_2 = \left(\bigcup_{i \in I_1} B_i \right) \cap \left(\bigcup_{i \in I_2} B_i \right) = \bigcup_{(i,j) \in I} B_i \cap B_j \in \tau_X \quad \text{since, } B_i \cap B_j \in \mathcal{B}$$

So it remains to show that τ_X is closed under arbitrary unions. Suppose that $\{U_i : i \in I\} \subseteq \tau_X$. Then for each $i \in I$, there exist disjoint sets J_i such that $U_i = \bigcup_{j \in J_i} B_j$, where $B_j \in \mathcal{B}$. Let $J := \bigcup_{i \in I} J_i$. Then $\bigcup_{I \in I} U_i = \bigcup_{j \in J} B_j \in \tau_X$. We now show that τ_X is the weakest topology on X the makes each function in $\mathcal{F}$ continuous. So suppose that τ is a topology on X that makes each function in $\mathcal{F}$ continuous. Then clearly $\mathcal{B} \subseteq \tau$ since $f^{-1}(U) \in \tau$ for each $f \in \mathcal{F}$ and each $U \in \tau_Y$. Since τ_X is the smallest topology on X that contains $\mathcal{B}$ we must have that $\tau_X \subseteq \tau$. $\square$

The topology τ_X in Proposition 2.2 is call the *weak topology on X generated by $\mathcal{F}$ and τ_Y*, or more briefly, when the context is clear, the *weak topology on X*. For further information on general topology see [10, 28].

2.2   Linear topologies

Let $(V, +, \cdot)$ be a vector space over the field of real numbers and let τ be a topology on V. Then $(V, +, \cdot, \tau)$ is called a *linear topological space* or a *topological vector space* if vector addition from $V \times V$ into V is continuous, when $V \times V$ is considered with the product topology and scalar multiplication from $\mathbb{R} \times V$ into V is continuous, again when we consider $\mathbb{R} \times V$ with the product topology and $\mathbb{R}$ with the usual topology.

An important feature of linear topological spaces is that they are always *regular*. That is, if $(X, +, \cdot, \tau)$ is linear topological space, C is a closed subset of X and $x \in X \setminus C$ then there exist disjoint open sets U and V such that $x \in U$ and $C \subseteq V$. To see this, suppose that $x = x + 0 \in X \setminus C$; which is open. Therefore, from the continuity of addition, there exist open neighbourhoods U of x and W of 0 such that $U + W \subseteq X \setminus C$, i.e., $(U + W) \cap C = \varnothing$. Therefore, $U \cap (C + (-W)) = \varnothing$. Let $V := C + (-W) = \bigcup_{c \in C} c - W$. Then V is an open set containing the set C and $U \cap V = \varnothing$. Thus, (X, τ) is a regular topological space.

Let $(V, +, \cdot, \tau)$ be a linear topological space over $\mathbb{R}$. We shall say that $(V, +, \cdot, \tau)$ is a *locally convex space* if for each open set U in V, containing 0, there exists an open convex set W such that $0 \in W \subseteq U$, or, equivalently, (V, τ) has a local base consisting of open convex sets.

If $(X, \|\cdot\|)$ is a normed linear space and $\varnothing \ne \mathcal{F} \subseteq X^*$ - the set of all continuous linear functionals on X, then $\sigma(X, \mathcal{F})$ denotes the weak topology on X generated by $\mathcal{F}$. We shall simply call $\sigma(X, X^*)$ the *weak topology on X* and write (X, weak) for $(X, \sigma(X, X^*))$.

Sometimes it is convenient to work with a more concrete representation of the $\sigma(X, \mathcal{F})$-topology. Fortunately such a representation exists and furthermore, the representation is very similar to the way in which open sets are defined in metric spaces. Let $(X, \|\cdot\|)$ be a normed linear space, let $x_0 \in X$, $\varepsilon > 0$ and let F be a nonempty finite subset of X^*. Then

$$N_X(x_0, F, \varepsilon) := \bigcap_{f \in F} \{x \in X : |f(x) - f(x_0)| < \varepsilon\}.$$

Note: sometimes it is also convenient to write $N_X(x_0, f_1, f_2, \ldots, f_n, \varepsilon)$ when the finite set F is enumerated as $F := \{f_1, f_2, \ldots, f_n\}$. When the context is clear we simply write, $N(x_0, F, \varepsilon)$ or $N(x_0, f_1, f_2, \ldots, f_n, \varepsilon)$.

Given a nonempty subset $\mathcal{F}$ of X^* we shall say that a subset U of X is $\mathcal{F}$-*open* if for every $x_0 \in U$ there exists a nonempty finite subset F of $\mathcal{F}$ and an $\varepsilon > 0$ such that $N(x_0, F, \varepsilon) \subseteq U$.

Proposition 2.3. *If $(X, \|\cdot\|)$ is a normed linear space and $\varnothing \neq \mathcal{F} \subseteq X^*$ then the set of all $\mathcal{F}$-open sets forms a topology on X. Furthermore, the set of all $\mathcal{F}$-open sets coincides with the $\sigma(X, \mathcal{F})$-topology on X.*

Proof. First we will show that the set of all $\mathcal{F}$-open sets forms a topology on X. It is easy to see that vacuously, $\varnothing$ is $\mathcal{F}$-open. To see that X is $\mathcal{F}$-open, consider any element $x_0 \in X$. Now, let x^* be any element in $\mathcal{F}$. Then $N(x_0, \{x^*\}, 1) \subseteq X$. Therefore, X is $\mathcal{F}$-open. Next, suppose that U and V are both $\mathcal{F}$-open subsets of X. We will show that $U \cap V$ is also $\mathcal{F}$-open. To this end, let $x_0 \in U \cap V$. Then, since $x_0 \in U$, there exists a finite subset F_U of $\mathcal{F}$ and an $\varepsilon_U > 0$ such that $N(x_0, F_U, \varepsilon_U) \subseteq U$. Similarly, there exists a finite subset F_V of $\mathcal{F}$ and an $\varepsilon_V > 0$ such that $N(x_0, F_V, \varepsilon_V) \subseteq V$. Let $F := F_U \cup F_V$ and $\varepsilon := \min\{\varepsilon_U, \varepsilon_V\}$. Then,
$$x_0 \in N(x_0, F, \varepsilon) \subseteq N(x_0, F_U, \varepsilon_U) \cap N(x_0, F_V, \varepsilon_V) \subseteq U \cap V.$$

So it remains to show that an arbitrary union of $\mathcal{F}$-open sets is again $\mathcal{F}$-open. Let $\{U_\alpha : \alpha \in A\}$ be a family of $\mathcal{F}$-open sets. Let x_0 be any element of $\bigcup_{\alpha \in A} U_\alpha$. Then there exists an $\alpha_0 \in A$ such that $x_0 \in U_{\alpha_0}$. Since U_{α_0} is $\mathcal{F}$-open there exists a finite subset F of $\mathcal{F}$ and an $\varepsilon > 0$ such that $N(x_0, F, \varepsilon) \subseteq U_{\alpha_0}$. Now, $U_{\alpha_0} \subseteq \bigcup_{\alpha \in A} U_\alpha$ and so $N(x_0, F, \varepsilon) \subseteq \bigcup_{\alpha \in A} U_\alpha$. Therefore, $\bigcup_{\alpha \in A} U_\alpha$ is $\mathcal{F}$-open. We will now show that the two topologies coincide. Suppose that U is an $\mathcal{F}$-open set. Then, for each $x \in U$, there exists a finite subset F_x of $\mathcal{F}$ and an $\varepsilon_x > 0$ such that $x \in N(x, F_x, \varepsilon_x) \subseteq U$. Therefore, $\bigcup_{x \in U} N(x, F_x, \varepsilon_x) = U$. Thus, to show that U is $\sigma(X, \mathcal{F})$-open it is sufficient to show that every set of the form: $N(x, F, \varepsilon)$ is $\sigma(X, \mathcal{F})$-open, where $x \in X$, F is a finite subset of $\mathcal{F}$ and $\varepsilon > 0$. So suppose that $x_0 \in X$, $F = \{f_1, f_2, \ldots, f_n\} \subseteq \mathcal{F}$ and $\varepsilon > 0$. Let $U_k := (f_k(x_0) - \varepsilon, f_k(x_0) + \varepsilon)$ for each $1 \leq k \leq n$. Then,

$$N(x, F, \varepsilon) = \bigcap_{f \in F} \{x \in X : |f(x) - f(x_0)| < \varepsilon\} = \bigcap_{1 \leq k \leq n} f_k^{-1}(U_k). \qquad (*)$$

Therefore, by the definition of the $\sigma(X, \mathcal{F})$-topology, $N(x, F, \varepsilon)$ is $\sigma(X, \mathcal{F})$-open. To show that every $\sigma(X, \mathcal{F})$-open set is $\mathcal{F}$-open it is sufficient to show that each member of $\mathcal{F}$ is continuous with respect to the topology generated by the $\mathcal{F}$-open sets. However, this is obvious from the definition of the $\mathcal{F}$-open sets. If you want to see the details, then let $f \in \mathcal{F}$, $x_0 \in X$ and $\varepsilon > 0$. Then

$$f(N(x_0, \{f\}, \varepsilon)) \subseteq (f(x_0) - \varepsilon, f(x_0) + \varepsilon).$$

This completes the proof. $\quad\square$

Remark 2.4. *It follows from Proposition 2.3 and equation $(*)$ that for each $x \in X$, finite set $\varnothing \neq F \subseteq \mathcal{F}$ and $\varepsilon > 0$, the set $N(x, F, \varepsilon)$ is $\sigma(X, \mathcal{F})$-open in X.*

By using Proposition 2.3 and Remark 2.4, one can easily deduce the following result.

Proposition 2.5. *Let Y be a subspace of a normed linear space $(X, \|\cdot\|)$ and let $\varnothing \neq \mathcal{F} \subseteq X^*$. Then a subset U of Y is open in the relative $\sigma(X, \mathcal{F})$-topology on Y if, and only if, for each $y \in U$ there exists a finite subset F of $\mathcal{F}$ and an $\varepsilon > 0$ such that $N_X(y, F, \varepsilon) \cap Y \subseteq U$.*

Proposition 2.6. *If $(X, \|\cdot\|)$ is a normed linear space and $\mathcal{F} \subseteq X^*$, then $(X, \sigma(X, \mathcal{F}))$ is a locally convex topological space.*

Proof. Let us first show that $(X, \sigma(X, \mathcal{F}))$ is a linear topology. Let $S : X \times X \to X$ be defined by, $S(x, y) := x + y$. We need to show that S is continuous. To this end, let W be a $\sigma(X, \mathcal{F})$-open subset of X and let $(x, y) \in S^{-1}(W)$, i.e., $S(x, y) \in W$. By Proposition 2.3 there exists a finite subset F of $\mathcal{F}$ and an $\varepsilon > 0$ such that $N(S(x, y), F, \varepsilon) \subseteq W$. We claim that $S(N(x, F, \varepsilon/2) \times N(y, F, \varepsilon/2)) \subseteq N(S(x, y), F, \varepsilon) \subseteq W$. To see this, let $(x', y') \in N(x, F, \varepsilon) \times N(y, F, \varepsilon)$ and let $f \in F$. Then, $|f(x') - f(x)| < \varepsilon/2$ and $|f(y') - f(y)| < \varepsilon/2$, and so

$$
\begin{aligned}
|f(S(x', y')) - f(S(x, y))| &= |f(x' + y') - f(x + y)| \\
&= |f(x' - x) + f(y' - y)| \\
&\leq |f(x' - x)| + |f(y' - y)| \\
&= |f(x') - f(x)| + |f(y') - f(y)| < \varepsilon/2 + \varepsilon/2 = \varepsilon.
\end{aligned}
$$

Therefore, $S(x', y') \in N(S(x, y), F, \varepsilon)$; which proves the claim. Now since both $N(x, F, \varepsilon/2)$ and $N(y, F, \varepsilon/2)$ are $\sigma(X, \mathcal{F})$-open we see that $S^{-1}(W)$ is open in $X \times X$, with the product topology and so S is continuous. Let $M : \mathbb{R} \times X \to X$ be defined by, $M(r, x) := rx$. We need to show that M is continuous. To this end, let W be a $\sigma(X, \mathcal{F})$-open subset of X and let $(r, x) \in M^{-1}(W)$, i.e., $M(r, x) \in W$. By Proposition 2.3 there exists a finite subset F of $\mathcal{F}$ and an $1 > \varepsilon > 0$ such that $N(M(r, x), F, \varepsilon) \subseteq W$. Set

$$
\varepsilon_1 := \frac{\varepsilon}{2(|f(x)| + 1)} \quad \text{and} \quad \varepsilon_2 := \frac{\varepsilon}{2(|r| + 1)}.
$$

We claim that $M((r - \varepsilon_1, r + \varepsilon_1) \times N(x, F, \varepsilon_2)) \subseteq N(M(r, x), F, \varepsilon) \subseteq W$. To see this is true, let $(r', x') \in (r - \varepsilon_1, r + \varepsilon_1) \times N(x, F, \varepsilon_2)$ and let $f \in F$. Then, $|r' - r| < \varepsilon_1$ and $|f(x') - f(x)| < \varepsilon_2$, and so

$$
\begin{aligned}
|f(M(r', x')) - f(M(r, x))| &= |f(r'x') - f(rx)| \\
&= |f(r'x') - f(r'x) - [f(rx) - f(r'x)]| \\
&\leq |f(r'x') - f(r'x)| + |f(rx) - f(r'x)| \\
&= |r'||f(x') - f(x)| + |r - r'||f(x)| \\
&\leq (|r| + 1)\varepsilon_2 + \varepsilon_1|f(x)| < \varepsilon \quad \text{since, } |r'| \leq |r| + \varepsilon_1 < |r| + 1.
\end{aligned}
$$

Therefore, $M(r', x') \in N(M(r, x), F, \varepsilon)$; which proves the claim. Now since $(r - \varepsilon_1, r + \varepsilon_1)$ is open in $\mathbb{R}$ and $N(x, F, \varepsilon_2)$ is $\sigma(X, \mathcal{F})$-open we see that $M^{-1}(W)$ is open in $\mathbb{R} \times X$, with the product topology and so M is continuous. This shows that $(X, \sigma(X, \mathcal{F}))$ is a linear topological space. To see that $(X, \sigma(X, \mathcal{F}))$ is locally convex we merely appeal to Proposition 2.3 and the fact that for each finite subset F of $\mathcal{F}$ and $\varepsilon > 0$, $N(0, F, \varepsilon)$ is a convex open neighbourhood of 0. $\quad \square$

The beauty of linear topology lies in the interplay between linear algebra and topology. This is highlighted in Proposition 2.8, which is based upon the following result from linear algebra.

Lemma 2.7 ([9, page 421]). *Let V be a vector space over $\mathbb{R}$ and suppose that $(f_i)_{i=1}^n$ are linear functionals on V. If g is a linear functional on V such that $\bigcap_{i=1}^n \ker(f_i) \subseteq \ker(g)$, then $g \in \operatorname{span}\{f_1, \ldots, f_n\}$.*

Proof. Define $T : V \to \mathbb{R}^n$ by

$$
T(x) := (f_1(x), \ldots, f_n(x)) \text{ for all } x \in V.
$$

Observe that T is clearly linear and that $\ker(T) = \bigcap_{i=1}^n \ker(f_i)$. We may assume that $(f_i)_{i=1}^n$ is a minimal (in terms of cardinality) family of functions such that $\bigcap_{i=1}^n \ker(f_i) \subseteq \ker(g)$, and from this we claim that T is also surjective.

Fix $1 \leq k \leq n$. Then, by the minimality assumption on n, we have that $\bigcap\{\ker(f_i) : 1 \leq i \leq n, \ i \neq k\} \not\subseteq \ker(g)$, and so in particular $\bigcap\{\ker(f_i) : 1 \leq i \leq n, \ i \neq k\} \not\subseteq \bigcap_{i=1}^n \ker(f_i)$. Now we may choose

$$
x_k \in \bigcap\{\ker(f_i) : 1 \leq i \leq n, \ i \neq k\} \setminus \bigcap_{i=1}^n \ker(f_i).
$$

Then, after scaling x_k if necessary, we have that $f_k(x_k) = 1$ and $f_i(x_k) = 0$ for $i \neq k$. Therefore, $T(x_k) = e_k$, where e_k is the kth standard basis vector of $\mathbb{R}^n$. Then, since $1 \leq k \leq n$ was arbitrary, we have that

$\mathbb{R}^n = \mathrm{span}\{e_1, \ldots, e_n\} \subseteq T(V)$ and so T is surjective as claimed.

Now define $g^* : \mathbb{R}^n \to \mathbb{R}$ by

$$g^*(x) := g(z) \text{ for any } z \in T^{-1}(x).$$

Then g^* is well-defined. Indeed, let $x \in \mathbb{R}^n$. Since T is onto, we have that $T^{-1}(x) \neq \varnothing$. So, suppose $z_1, z_2 \in T^{-1}(x)$. Then $T(z_1) = x = T(z_2)$ and so $z_1 - z_2 \in \ker(T) \subseteq \ker(g)$. Thus $g(z_1) = g(z_2)$ as required. Moreover, a routine calculation shows that g^* is linear, so $g^* \in (\mathbb{R}^n)^*$. Since $(\mathbb{R}^n)^* = \mathrm{span}\{e_1^*, \ldots, e_n^*\}$, there exist $(c_i)_{i=1}^n$ such that $g^* = \sum_{i=1}^n c_i e_i^*$, where here $e_i^*(e_j) = \delta_{ij}$, the ij-Kroeneker delta.

Finally, note that $x \in T^{-1}(T(x))$ and so $g^*(T(x)) = g(x)$ for all $x \in V$. Therefore,

$$g = g^* \circ T = \left(\sum_{i=1}^n c_i e_i^*\right) \circ T = \sum_{i=1}^n c_i \left(e_i^* \circ T\right) = \sum_{i=1}^n c_i f_i$$

since $e_i^* \circ T = f_i$ for all $1 \leq i \leq n$, and thus $g \in \mathrm{span}\{f_1, \ldots, f_n\}$. $\quad\square$

Proposition 2.8 ([9, page 422]). *If $(X, \|\cdot\|)$ is a normed linear space, $\mathcal{F} \subseteq X^*$ and $x^* \in X^*$ then the following are equivalent:*

(i) *x^* is $\sigma(X, \mathcal{F})$-continuous;*

(ii) *x^* is bounded on a $\sigma(X, \mathcal{F})$ neighbourhood of 0;*

(iii) *$x^* \in \mathrm{span}(\mathcal{F})$.*

Proof. $(i) \implies (ii)$. Suppose that $x^* \in X^*$ is $\sigma(X, \mathcal{F})$-continuous on X. Then, in particular, x^* is continuous at $0 \in X$. Therefore, there exists a $\sigma(X, \mathcal{F})$-open neighbourhood N of 0 such that $|x^*(x)| = |x^*(x) - x^*(0)| < 1$ for all $x \in N$. Hence x^* is bounded on N. $(ii) \implies (iii)$. Suppose that x^* is bounded on a $\sigma(X, \mathcal{F})$-open neighbourhood N of 0. Then there exists a finite subset G of $\mathcal{F}$ and an $0 < \varepsilon$ such that $N(0, G, \varepsilon) \subseteq N$. Let $S := \bigcap_{y^* \in G} \ker(y^*)$. Then S is a subspace of X and furthermore, $S \subseteq N(0, G, \varepsilon) \subseteq N$. Hence, $x^*|_S$ is bounded on S. Thus, $x^*|_S \equiv 0$ and so $\bigcap_{y^* \in G} \ker(y^*) = S \subseteq \ker(x^*)$. The result now follows from Lemma 2.7. $(iii) \implies (i)$. Suppose that $x^* = \sum_{k=1}^n \lambda_k y_k^*$, where $\lambda_k \in \mathbb{R}$ and $y_k^* \in \mathcal{F}$ for all $1 \leq k \leq n$. Define $f : \mathbb{R}^n \to \mathbb{R}$ by, $f(x_1, x_2, \ldots, x_n) := \sum_{k=1}^n \lambda_k x_k$ and $F : X \to \mathbb{R}^n$ by, $F(x) := (y_1^*(x), y_2^*(x), \ldots, y_n^*(x))$. Then f is a continuous function on $\mathbb{R}^n$ and F is a $\sigma(X, \mathcal{F})$-continuous function on X. Therefore, $x^* = f \circ F$, is also a $\sigma(X, \mathcal{F})$-continuous function on X. $\quad\square$

Remark 2.9. *It follows from Proposition 2.8 that for any normed linear space $(X, \|\cdot\|)$ and any $\mathcal{F} \subseteq X^*$, $\sigma(X, \mathcal{F}) = \sigma(X, \mathrm{span}(\mathcal{F}))$. To see this, first note the general fact that if $\mathcal{F} \subseteq \mathcal{F}'$ then $\sigma(X, \mathcal{F}) \subseteq \sigma(X, \mathcal{F}')$ (i.e., to make more functions continuous you need more open sets) and then the equivalence of (i) and (iii) above.*

Proposition 2.10. *If $T : X \to Y$ is a continuous linear operator acting between normed linear spaces $(X\|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ then $T : (X, \mathrm{weak}) \to (Y, \mathrm{weak})$ is also continuous.*

Proof. Let W be a weak open subset of Y. We will show that $T^{-1}(W)$ is open in the weak topology on X. To this end, let $x_0 \in T^{-1}(W)$. Then $T(x_0) \in W$ and so by Proposition 2.3 there exist $\{y_1^*, y_2^*, \ldots, y_n^*\} \subseteq Y^*$ and $\varepsilon > 0$ such that $N_Y(T(x_0), \{y_1^*, y_2^*, \ldots, y_n^*\}, \varepsilon) \subseteq W$. For each $1 \leq k \leq n$ let $x_k^* := y_k^* \circ T$. Then $\{x_1^*, x_2^*, \ldots, x_n^*\} \subseteq X^*$. We claim that $T(N_X(x_0, \{x_1^*, x_2^*, \ldots, x_n^*\}, \varepsilon)) \subseteq N_Y(T(x_0), \{y_1^*, y_2^*, \ldots, y_n^*\}, \varepsilon) \subseteq W$. To see this, let $y \in T(N_X(x_0, \{x_1^*, x_2^*, \ldots, x_n^*\}, \varepsilon))$. Then there exists a $x \in N_X(x_0, \{x_1^*, x_2^*, \ldots, x_n^*\}, \varepsilon)$ such that $y = T(x)$. Fix $1 \leq k \leq n$. Then,

$$\begin{aligned} |y_k^*(y) - y_k^*(T(x_0))| &= |y_k^*(T(x)) - y_k^*(T(x_0))| \\ &= |x_k^*(x) - x_k^*(x_0)| < \varepsilon, \quad \text{since } x \in N_X(x_0, \{x_1^*, x_2^*, \ldots, x_n^*\}, \varepsilon). \end{aligned}$$

Therefore, $y \in N_Y(T(x_0), \{y_1^*, y_2^*, \ldots, y_n^*\}, \varepsilon) \subseteq W$. This completes the proof of the claim. Hence

$$x_0 \in N_X(x_0, \{x_1^*, x_2^*, \ldots, x_n^*\}, \varepsilon) \subseteq T^{-1}(W).$$

Thus, by Proposition 2.3, $T^{-1}(W)$ is open in the weak topology on X. $\quad\square$

2.3 Hahn-Banach Theorem

A real-valued function p defined on a vector space V is called *sublinear* if for every $x, y \in V$ and $0 \leq \lambda < \infty$, $p(\lambda x) = \lambda p(x)$ and $p(x + y) \leq p(x) + p(y)$.

Although it is easy, using linear algebra, to construct linear functionals on a vector space, it is not so easy to construct continuous linear functions on a linear topological space. The key to constructing continuous linear functionals on locally convex spaces is given next.

Theorem 2.11 (Hahn-Banach Theorem [9, page 62]). *Let Y be a subspace of a vector space V (over $\mathbb{R}$) and let $p : V \to \mathbb{R}$ be a sublinear functional on V. If f is a linear functional on Y and $f(y) \leq p(y)$ for all $y \in Y$ then there exists a linear functional $F : V \to \mathbb{R}$ such that $F|_Y = f$ and $F(x) \leq p(x)$ for all $x \in V$.*

Proof. Let $\mathscr{P}$ be the collection of all ordered pairs (M', f'), where M' is a subspace of V containing Y and $f' : M' \to \mathbb{R}$ is a linear functional defined on M' such that $f'|_Y = f$ and satisfies $f'(x) \leq p(x)$ for all $x \in M'$. $\mathscr{P}$ is nonempty because $(Y, f) \in \mathscr{P}$. We partially order $\mathscr{P}$ by, $(M', f') \leq (M'', f'')$ if $M' \subseteq M''$ and $f''|_{M'} = f'$. If $\{(M_\alpha, f_\alpha) : \alpha \in A\}$ is a nonempty totally ordered sub-family of $\mathscr{P}$, then set $M' := \bigcup \{M_\alpha : \alpha \in A\}$ and define the linear functional $f' : M' \to \mathbb{R}$ by, $f'(x) := f_\alpha(x)$ if $x \in M_\alpha$. Then $(M', f') \in \mathscr{P}$ and $(M_\alpha, f_\alpha) \leq (M', f')$ for all $\alpha \in A$. Therefore, by Zorn's lemma, $\mathscr{P}$ has a maximal element (M, F). We must show that $M = V$. So suppose, in order to obtain a contradiction, that $M \neq V$ and pick $x_0 \in V \setminus M$ and put $M^* := \mathrm{span}\{M, x_0\}$. We will define $F^* : M^* \to \mathbb{R}$ so that $(M^*, F^*) \in \mathscr{P}$ and $(M, F) < (M^*, F^*)$; which will be our desired contradiction. For each $\alpha \in \mathbb{R}$ we define F_α on M^* by, $F_\alpha(m + \lambda x_0) := f(m) + \lambda \alpha$. It is easy to check that F_α is well defined and linear on M^*. Moreover, $F_\alpha|_M = f$. So it remains to show that $F_\alpha(x) \leq p(x)$ for all $x \in M^*$. To achieve this, we need to select the right value of $\alpha \in \mathbb{R}$.

Selection of α: For any $m_1, m_2 \in M$ and $0 < \lambda_1 < \infty$ and $0 < \lambda_2 < \infty$ we have:

$$f(\lambda_1^{-1} m_1 + \lambda_2^{-1} m_2) \leq p(\lambda_1^{-1} m_1 + \lambda_2^{-1} m_2) \leq p(\lambda_1^{-1} m_1 - x_0) + p(\lambda_2^{-1} m_2 + x_0).$$

Therefore,

$$f(\lambda_1^{-1} m_1) - p(\lambda_1^{-1} m_1 - x_0) \leq p(\lambda_2^{-1} m_2 + x_0) - f(\lambda_2^{-1} m_2)$$

for all $m_1, m_2 \in M$ and $0 < \lambda_1 < \infty$, $0 < \lambda_2 < \infty$. Hold m_2 and λ_2 fixed and take the supremum over $m_1 \in M$ and $0 < \lambda_1 < \infty$. Then for each $m_2 \in M$ and $0 < \lambda_2 < \infty$ we have that:

$$\sup_{\substack{m \in M \\ 0 < \lambda < \infty}} \left(f(\lambda^{-1} m) - p(\lambda^{-1} m - x_0) \right) \leq p(\lambda_2^{-1} m_2 + x_0) - f(\lambda_2^{-1} m_2).$$

Now we take the infimum over $m_2 \in M$ and $0 < \lambda_2 < \infty$. Then,

$$a := \sup_{\substack{m \in M \\ 0 < \lambda < \infty}} \left(f(\lambda^{-1} m) - p(\lambda^{-1} m - x_0) \right) \leq \inf_{\substack{m \in M \\ 0 < \lambda < \infty}} \left(p(\lambda^{-1} m + x_0) - f(\lambda^{-1} m) \right) =: b.$$

Choose $\alpha^* \in [a, b]$. Then from the left-hand side of the equation we get that:

$$f(m) + (-\lambda)\alpha^* \leq p(m + (-\lambda)x_0) \quad \text{for all } m \in M \text{ and } 0 < \lambda < \infty.$$

From the right-hand side of the equation we get that:

$$f(m) + \lambda \alpha^* \leq p(m + \lambda x_0) \quad \text{for all } m \in M \text{ and } 0 < \lambda < \infty.$$

From these two equations we see that:

$$F^*(x) := F_{\alpha^*}(x) \leq p(x) \quad \text{for all } x \in M^*.$$

That is, $(M, F) < (M^*, F^*) \in \mathscr{P}$. $\square$

We now give some applications of this famous theorem.

Corollary 2.12. *Let Y be a subspace of a normed linear space $(X, \|\cdot\|)$ (over $\mathbb{R}$). If $f \in Y^*$ then there exists an $F \in X^*$ such that $F|_Y = f$ and $\|F\| = \|f\|$.*

Proof. Consider the sublinear functional $p : X \to \mathbb{R}$ defined by, $p(x) := \|f\|\|x\|$. Then $f(y) \le p(y)$ for all $y \in Y$. By the Hahn-Banach Theorem, (Theorem 2.11) there exists a linear functional $F : X \to \mathbb{R}$ such that $F|_Y = f$ and $F(x) \le p(x)$ for all $x \in X$. Therefore, $-F(x) = F(-x) \le p(-x) = p(x)$ for all $x \in X$ too. Thus, $|F(x)| \le p(x)$ for all $x \in X$. This in turn implies that $\|F\| \le \|f\|$. On the other hand, since F is an extension of f, we must also have that $\|f\| \le \|F\|$. $\square$

Corollary 2.13. *Let $(X, \|\cdot\|)$ be a normed linear space. For every $x \in X \setminus \{0\}$ there exists an $f \in S_{X^*}$ such that $f(x) = \|x\|$.*

Proof. Let $Y := \operatorname{span}\{x\}$ and define $f \in Y^*$ by, $f(\lambda x) := \lambda \|x\|$. Clearly, $\|f\| = 1$ and $f(x) = \|x\|$. By Corollary 2.12 there exists an $F \in X^*$ such that $\|F\| = \|f\| = 1$ and $F|_Y = f$. Therefore, in particular we have that $F(x) = f(x) = \|x\|$. $\square$

Proposition 2.14. *Let Y be a subspace of a normed linear space $(X, \|\cdot\|)$. Then the topology $\sigma(Y, Y^*)$ on Y coincides with the relative $\sigma(X, X^*)$ topology on Y.*

Proof. Let us first show that every relatively $\sigma(X, X^*)$-open set in Y is $\sigma(Y, Y^*)$ open. To this end, let U be a relatively $\sigma(X, X^*)$-open set in Y. Let $y \in U$. By Proposition 2.5, there exists a finite set $\{x_1^*, x_2^*, \ldots, x_n^*\} \subseteq X^*$ and an $\varepsilon > 0$ such that $N_X(y, \{x_1^*, x_2^*, \ldots, x_n^*\}, \varepsilon) \cap Y \subseteq U$. For each $1 \le k \le n$ let $y_k^* := x_k^*|_Y$. Then

$$N_Y(y, \{y_1^*, y_2^*, \ldots, y_n^*\}, \varepsilon) = N_X(y, \{x_1^*, x_2^*, \ldots, x_n^*\}, \varepsilon) \cap Y \subseteq U.$$

Thus, by Proposition 2.3, U is $\sigma(Y, Y^*)$-open. Now, suppose that U is a $\sigma(Y, Y^*)$-open subset of Y. Then, by Proposition 2.3, there exists a finite set $\{y_1^*, y_2^*, \ldots, y_n^*\} \subseteq Y^*$ and an $\varepsilon > 0$ such that $N_Y(y, \{y_1^*, y_2^*, \ldots, y_n^*\}, \varepsilon) \subseteq U$. By Corollary 2.12, for each $1 \le k \le n$, there exists an $x_k^* \in X^*$ such that $x_k^*|_Y = y_k^*$. Then

$$N_X(y, \{x_1^*, x_2^*, \ldots, x_n^*\}, \varepsilon) \cap Y = N_Y(y, \{y_1^*, y_2^*, \ldots, y_n^*\}, \varepsilon) \subseteq U.$$

Therefore, by Proposition 2.5, U is open in the relative $\sigma(X, X^*)$-topology on Y. $\square$

Next we will show how to use the Hahn-Banach Theorem to obtain some geometric properties of locally convex spaces.

Let S be a nonempty subset of a vector space V. We shall say that a point $x \in S$ is a *core point* of S if for every $v \in V$ there exists a $0 < \delta < \infty$ such that $x + \lambda v \in S$ for all $0 \le \lambda < \delta$. The set of all core points of S is called the *core* of S and is denoted by $\operatorname{Cor}(S)$.

Let C be a convex set in a vector space V with $0 \in \operatorname{Cor}(C)$. Then the functional $\mu_C : V \to \mathbb{R}$ defined by,

$$\mu_C(x) := \inf\{\lambda > 0 : x \in \lambda C\}$$

is called the *Minkowski functional* generated by the set C.

Theorem 2.15. *Let C be a convex subset of a vector space V with 0 in the core of C. Then $\mu_C : V \to \mathbb{R}$ is a sublinear functional. Moreover,*

$$\{x \in V : \mu_C(x) < 1\} \subseteq C \subseteq \{x \in V : \mu_C(x) \le 1\}.$$

Proof. Given $\alpha > 0$ and $\lambda > 0$, clearly $x \in \lambda C$ if, and only if, $\alpha x \in \lambda \alpha C$. Therefore, $\mu_C(\alpha x) = \alpha \mu_C(x)$ and thus μ_C is *positively homogeneous*. We claim that μ_C is *subadditive*; that is, $\mu_C(x + y) \le \mu_C(x) + \mu_C(y)$. Fix any $s > \mu_C(x)$ and $t > \mu_C(y)$. We have that there is some $s_0 < s$ such that $x \in s_0 C$. Note that $s_0 C \subseteq sC$. Indeed, $0 \in sC$ and if $c \in C$, then by the convexity of sC,

$$s_0 c = \frac{s_0}{s}(sc) + \left(1 - \frac{s_0}{s}\right)0 \in sC.$$

We see that $x \in sC$ and similarly $y \in tC$. Then $x + y \in sC + tC$ and thus by the convexity of C,

$$x + y \in (s+t)\left(\frac{s}{s+t}C + \frac{t}{s+t}C\right) \subseteq (t+s)C.$$

Therefore, $\mu_C(x+y) \le s+t$ and so by the choice of s and t we have that $\mu_C(x+y) \le \mu_C(x) + \mu_C(y)$. If $\mu_C(x) < 1$ then $x \in \lambda C$ for some $0 < \lambda < 1$ and so $(1/\lambda)x \in C$. Since $0 \in C$ and C is convex,

$$x = \lambda\left(\frac{x}{\lambda}\right) + (1-\lambda)\,0 \in C.$$

If $x \in C$ then $\mu_C(x) \le 1$ by the definition of the Minkowski functional. $\square$

Remark 2.16. *If the set C in Theorem 2.15 is a closed and convex subset of a topological vector space (V, τ), with $0 \in \mathrm{Cor}(C)$ and $x_0 \notin C$ then it is an easy exercise to show that $1 < \mu_C(x_0)$.*

We now give the geometric version of the Hahn-Banach Theorem.

Theorem 2.17 (Separation Theorem [9, page 418]). *Suppose that (X, τ) is a locally convex space over $\mathbb{R}$ and C is a nonempty closed convex subset of X. If $x_0 \notin C$ then there exists a continuous linear functional x^* on X such that*

$$\sup\{x^*(c) : c \in C\} < x^*(x_0).$$

Proof. We may assume, without loss of generality, that $0 \in C$; because otherwise we would consider $C - x$ and $x_0 - x$ for some $x \in C$. Since vector addition is continuous and $x_0 + 0 \notin C$ there exist convex open neighbourhoods U of x_0 and V of 0 such that $(U + V) \cap C = \varnothing$. Thus, $U \cap [C + (-V)] = \varnothing$. Now, $-V$ is also a convex open neighbourhood of 0 and so $C + (-V)$ is a convex open set containing the set C and disjoint from U. Let $D := \overline{C + (-V)}$, then D is a closed and convex set with $0 \in \mathrm{int}(D)$ and $x_0 \notin D$. Let μ_D be the Minkowski functional for D. Since D is closed and $x_0 \notin D$ we have $\mu_D(x_0) > 1$ (see Remark 2.16). Define a linear functional on $\mathrm{span}\{x_0\}$ by, $f(\lambda x_0) := \lambda \mu_D(x_0)$. Then on $\mathrm{span}\{x_0\}$ we have that $f(\lambda x_0) \le \mu_D(\lambda x_0)$. Indeed, for $0 \le \lambda$ it is clear from the definition of f; whereas for $\lambda < 0$ we have $f(\lambda x_0) = \lambda \mu_D(x_0) < 0$ while $\mu_D(\lambda x_0) \ge 0$. By using the Hahn-Banach Theorem we may extend f onto X so that $f(x) \le \mu_D(x)$ for all $x \in X$. If $x \in D$ then $\mu_D(x) \le 1$ and thus, $f(x) \le \mu_D(x) \le 1$. Since D contains a neighbourhood of the origin we have that f is a bounded on a neighbourhood of 0 and so by Proposition 2.8, $f \in X^*$. Since $f(x_0) = \mu_D(x_0) > 1$ we get that $\sup\{f(x) : x \in C\} \le \sup\{f(x) : x \in D\} \le 1 < f(x_0)$. $\square$

An immediate consequence of the Separation Theorem is the following result, which is sometimes known as Mazur's Theorem.

Proposition 2.18 ([9, page 422]). *Let C be a closed convex subset of a normed linear space $(X, \|\cdot\|)$. Then C is also closed with respect to the weak topology on X.*

Proof. If C is empty or the whole space, then C is weakly closed, so let us suppose otherwise. Let $x_0 \in X \setminus C$. Since C is closed and convex, we have, by the Separation Theorem (Theorem 2.17), the existence of an $f_{x_0} \in X^*$ such that $f_{x_0}(x_0) > \sup_{x \in C} f_{x_0}(x)$. Thus, $x_0 \in f_{x_0}^{-1}\big((\sup_{x \in C} f_{x_0}(x), \infty)\big)$, which, being the inverse image of an open set, is weakly open. It is then straightforward to check that $X \setminus C = \bigcup_{x_0 \in X \setminus C} f_{x_0}^{-1}\big((\sup_{x \in C} f_{x_0}(x), \infty)\big)$. Hence, $X \setminus C$, being the union of weakly open sets, is weakly open. Thus, C is weakly closed. $\square$

2.4 Weak* topology

Let $(X, \|\cdot\|)$ be a normed linear space. For each $x \in X$ we define, $\widehat{x} \in X^{**} := (X^*)^*$ by, $\widehat{x}(x^*) := x^*(x)$ for all $x^* \in X^*$. To show that $\widehat{x}$ is really in X^{**} we must first check that it is linear and then check that it is continuous. So suppose that x^* and y^* are in X^*, then

$$\widehat{x}(x^* + y^*) = (x^* + y^*)(x) = x^*(x) + y^*(x) = \widehat{x}(x^*) + \widehat{x}(y^*).$$

Also, for any $\lambda \in \mathbb{R}$ and $x^* \in X^*$ we have that

$$\widehat{x}(\lambda x^*) = (\lambda x^*)(x) = \lambda x^*(x) = \lambda \widehat{x}(x^*).$$

Now, $|\widehat{x}(x^*)| = |(x^*)(x)| \le \|x^*\| \cdot \|x\|$. Therefore, $\|\widehat{x}\| \le \|x\|$ and so $\widehat{x} \in X^{**}$.

Proposition 2.19. *Let* $(X, \|\cdot\|)$ *be a normed linear space. Then the mapping* $x \mapsto \widehat{x}$ *is a linear isometry from* X *into* X^{**}.

Proof. The mapping $x \mapsto \widehat{x}$ from X into X^{**} is linear, since for all $x^* \in X^*$

$$\widehat{(x+y)}(x^*) = x^*(x+y) = x^*(x) + x^*(y) = \widehat{x}(x^*) + \widehat{y}(x^*).$$

Therefore, $\widehat{x+y} = \widehat{x} + \widehat{y}$. Also, for any $\lambda \in \mathbb{R}$ and $x^* \in X^*$,

$$\widehat{(\lambda x)}(x^*) = x^*(\lambda x) = \lambda x^*(x) = \lambda \widehat{x}(x^*).$$

Therefore, $\widehat{(\lambda x)} = \lambda \widehat{x}$. Next we show that $x \mapsto \widehat{x}$ is an isometry. For each $x \in X$, we have by Corollary 2.13, a linear function $x^* \in S_{X^*}$ such that $x^*(x) = \|x\|$. Therefore, $\|\widehat{x}\| \geq \frac{|\widehat{x}(x^*)|}{\|x^*\|} = |\widehat{x}(x^*)| = |x^*(x)| = \|x\|$. $\quad\square$

If $(X, \|\cdot\|)$ is a Banach space then $\widehat{X}$ is a closed subspace of X^{**} where $\widehat{X}$ is defined as $\{\widehat{x} : x \in X\}$. We call $\widehat{X}$ the *natural embedding* of X into X^{**} and we call $x \mapsto \widehat{x}$ from X into X^{**} the *natural embedding mapping*.

An important topology for our concerns is the weak* topology. Suppose that $(X, \|\cdot\|)$ is a normed linear space. Then we call the topology $\sigma(X^*, \widehat{X})$ on X^*, the *weak* topology* on X^* and we write (X^*, weak^*) for $(X^*, \sigma(X^*, \widehat{X}))$. It follows from Proposition 2.8 that $F \in X^{**}$ is weak* continuous if, and only if, $F \in \widehat{X}$.

Let $(X, \|\cdot\|)$ be a normed linear space and let $A \subseteq X$. We define the *(upper) polar of A* to be the subset A° of X^* defined by

$$A^\circ := \{x^* \in X^* : x^*(a) \leq 1 \text{ for all } a \in A\}.$$

Proposition 2.20. *Let* $(X, \|\cdot\|)$ *be a normed linear space and let* $A \subseteq X$. *Then* A° *is convex, weak*-closed, and contains* 0.

Proof. The fact that $0 \in A^\circ$ is trivial. To see that A° is weak*-closed and convex, note that $A^\circ = \bigcap_{a \in A} \widehat{a}^{-1}(-\infty, 1]$ is the intersection of weak*-closed and convex sets, and so is itself, weak*-closed and convex. $\quad\square$

There are many useful properties of polars that can be easily verified. For example, (i) if $A \subseteq B$ then $B^\circ \subseteq A^\circ$, (ii) $(B_X)^\circ = B_{X^*}$, (iii) for any $r > 0$, $(rA)^\circ = r^{-1}A^\circ$. By combining these we see that if $0 \in \text{int}(A)$ then A° is bounded and if A is bounded then $0 \in \text{int}(A^\circ)$.

There is also a dual version of (upper) polars. Let $(X, \|\cdot\|)$ be a normed linear space and let $A \subseteq X^*$. We define the *(lower) polar of A* to be the subset $A_\circ$ of X defined by

$$A_\circ := \{x \in X : a^*(x) \leq 1 \text{ for all } a^* \in A\}.$$

There are many interesting relationships between these polars. For example, for any $\varnothing \neq A \subseteq X$, $(A^\circ)_\circ = \overline{\text{co}}(A \cup \{0\})$ and for any $\varnothing \neq B \subseteq X^*$, $(B_\circ)^\circ = \overline{\text{co}}^{w^*}(B \cup \{0\})$.

Perhaps the most famous theorem concerning polars is the following theorem.

Theorem 2.21 (Bipolar Theorem)**.** *Let C be a closed, convex subset of a normed linear space* $(X, \|\cdot\|)$ *with* $0 \in C$. *Then* $C^{\circ\circ} := (C^\circ)^\circ = \overline{\widehat{C}}^{w^*}$.

Proof. It follows directly from the definition of $(C^\circ)^\circ$ that $\widehat{C} \subseteq C^{\circ\circ}$. Moreover, by Proposition 2.20 we know that $C^{\circ\circ}$ is weak*-closed and so $\overline{\widehat{C}}^{w^*} \subseteq C^{\circ\circ}$. Now suppose, in order to obtain a contradiction, that $\overline{\widehat{C}}^{w^*} \subsetneq C^{\circ\circ}$. Then there exists an $F_0 \in C^{\circ\circ} \setminus \overline{\widehat{C}}^{w^*}$. By Theorem 2.17, applied in (X^{**}, weak^*), there exists an $x^* \in X^*$ such that

$$0 \leq \sup\{F(x^*) : F \in \overline{\widehat{C}}^{w^*}\} = \sup\{\widehat{x^*}(F) : F \in \overline{\widehat{C}}^{w^*}\} < \widehat{x^*}(F_0) = F_0(x^*). \qquad (\dagger)$$

If necessary, we may replace x^* by λx^*, (for some $0 < \lambda$ and relabelling), so that

$$\sup\{x^*(c) : c \in C\} = \sup\{\widehat{c}(x^*) : \widehat{c} \in \widehat{C}\} \le \sup\{F(x^*) : F \in \overline{\widehat{C}}^{w^*}\} = 1.$$

Therefore, $x^* \in C^\circ$ and so $F_0(x^*) \le 1$, since $F_0 \in C^{\circ\circ}$. However, this contradicts the inequality:

$$1 = \sup\{F(x^*) : F \in \overline{\widehat{C}}^{w^*}\} = \sup\{\widehat{x^*}(F) : F \in \overline{\widehat{C}}^{w^*}\} < F_0(x^*), \qquad \text{by (†)}.$$

This completes the proof. □

An important application of the Bipolar Theorem is given next.

Corollary 2.22 (Goldstine's Theorem [9, page 424]). *Let $(X, \|\cdot\|)$ be a normed linear space then $B_{\widehat{X}}$ is weak* dense in $B_{X^{**}}$.*

Proof. We apply twice, the general fact (observed before) that if B_Y is the closed unit ball of a normed linear space Y then $B_{Y^*} = (B_Y)^\circ$ to obtain

$$B_{X^{**}} = (B_{X^*})^\circ = ((B_X)^\circ)^\circ = (B_X)^{\circ\circ}$$

and then apply the Bipolar Theorem. □

Perhaps the main reason for the interest in the weak* topology is contained in the next theorem. It says that, although it is too much to ask that the dual ball be compact with respect to the norm topology (unless the space is finite dimensional), it is possible that it is compact with respect to a weaker topology.

Theorem 2.23 (Banach-Alaoglu Theorem [1]). *Let $(X, \|\cdot\|)$ be a normed linear space. Then $(B_{X^*}, weak^*)$ is compact.*

Proof. For each $x \in X$, let $I_x := [-\|x\|, \|x\|]$ and let $Y := \prod_{x \in X} I_x$ be endowed with the product topology. By Tychonoff's Theorem, Y is compact. It follows from the definition of the product topology and Proposition 2.3 that $\pi : B_{X^*} \to Y$, defined by, $\pi(x^*)(x) := x^*(x)$ for all $x \in X$, is a homeomorphic embedding of B_{X^*} into Y. So to show that $(B_{X^*}, weak^*)$ is compact it is sufficient to show that $\pi(B_{X^*})$ is a closed subset of Y, that is, it is sufficient to show that $\overline{\pi(B_{X^*})} \subseteq \pi(B_{X^*})$. To this end, let $g \in \overline{\pi(B_{X^*})}$. We will show that g is "linear". Let $x, y \in X$ and $\varepsilon > 0$. Then there exists $x^* \in B_{X^*}$ such that $|g(x) - \pi(x^*)(x)| < \varepsilon/3$, $|g(y) - \pi(x^*)(y)| < \varepsilon/3$ and $|g(x + y) - \pi(x^*)(x + y)| < \varepsilon/3$. Then, since x^* is linear,

$$
\begin{aligned}
\big|g(x+y) - [g(x) + g(y)]\big| &= \big|[g(x+y) - \pi(x^*)(x+y)] + \pi(x^*)(x+y) - [g(x) + g(y)]\big| \\
&= \big|[g(x+y) - \pi(x^*)(x+y)] + [\pi(x^*)(x) - g(x)] + [\pi(x^*)(y) - g(y)]\big| \\
&\le |g(x+y) - \pi(x^*)(x+y)| + |\pi(x^*)(x) - g(x)| + |\pi(x^*)(y) - g(y)| \\
&\le \varepsilon/3 + \varepsilon/3 + \varepsilon/3 = \varepsilon.
\end{aligned}
$$

Since $\varepsilon > 0$ was arbitrary, $g(x + y) = g(x) + g(y)$. Next, let $x \in X$, $\lambda \in \mathbb{R}$ and $\varepsilon > 0$. Then there exists $x^* \in B_{X^*}$ such that $|g(\lambda x) - \pi(x^*)(\lambda x)| < \varepsilon/2$ and $|g(x) - \pi(x^*)(x)| < \varepsilon/2(|\lambda| + 1)$. Then, since x^* is linear,

$$
\begin{aligned}
|g(\lambda x) - \lambda g(x)| &= \big|[g(\lambda x) - \pi(x^*)(\lambda x)] + [\pi(x^*)(\lambda x) - \lambda g(x)]\big| \\
&= \big|[g(\lambda x) - \pi(x^*)(\lambda x)] + [\lambda \pi(x^*)(x) - \lambda g(x)]\big| \\
&\le |g(\lambda x) - \pi(x^*)(\lambda x)| + |[\lambda \pi(x^*)(x) - \lambda g(x)]| \\
&\le \varepsilon/2 + |\lambda||\pi(x^*)(x) - g(x)| < \varepsilon/2 + \varepsilon/2 = \varepsilon.
\end{aligned}
$$

Since $\varepsilon > 0$ was arbitrary, $g(\lambda x) = \lambda g(x)$. Thus, if we define $y^* : X \to \mathbb{R}$ by, $y^*(x) := g(x)$ for all $x \in X$, then $y^* \in B_{X^*}$ and $g = \pi(y^*) \subseteq \pi(B_{X^*})$. □

Proposition 2.24. *Let $(X, \|\cdot\|)$ be a normed linear space. Then the relative weak topology and the relative weak* topology coincide on the subspace $\widehat{X}$ of X^{**}.*

Proof. It follows immediately from the definitions that each relatively weak* open subset of $\widehat{X}$ is open in the relative weak topology on $\widehat{X}$. So we need only consider the converse statement. Suppose that U is a relatively weak open subset of $\widehat{X}$. Let $\widehat{x}$ be any element of U. Then by, Proposition 2.5 there exists a finite subset $\{\mathscr{F}_1, \mathscr{F}_2, \ldots, \mathscr{F}_n\}$ of X^{***} and an $\varepsilon > 0$ such that $N(\widehat{x}, \mathscr{F}_1, \mathscr{F}_2, \ldots, \mathscr{F}_N, \varepsilon) \cap \widehat{X} \subseteq U$. For each $1 \le k \le n$ let $f_k : X \to \mathbb{R}$ be defined by, $f_k(x) := \mathscr{F}_k(\widehat{x})$. Then $f_k \in X^*$, in fact $\|f_k\| \le \|\mathscr{F}_k\|$ for each $1 \le k \le n$. We claim that

$$N(\widehat{x}, \widehat{f}_1, \widehat{f}_2, \ldots, \widehat{f}_n, \varepsilon) \cap \widehat{X} \subseteq N(\widehat{x}, \mathscr{F}_1, \mathscr{F}_2, \ldots, \mathscr{F}_N, \varepsilon) \cap \widehat{X} \subseteq U.$$

To see this, let $F \in N(\widehat{x}, \widehat{f}_1, \widehat{f}_2, \ldots, \widehat{f}_n, \varepsilon) \cap \widehat{X}$. Then $F = \widehat{y}$ for some $y \in X$ and $\widehat{y} \in N(\widehat{x}, \widehat{f}_1, \widehat{f}_2, \ldots, \widehat{f}_n, \varepsilon) \cap \widehat{X}$. Fix $1 \le k \le n$. Then

$$|\mathscr{F}_k(F) - \mathscr{F}_k(\widehat{x})| = |\mathscr{F}_k(\widehat{y}) - \mathscr{F}_k(\widehat{x})| = |f_k(y) - f_k(x)| = |\widehat{f}_k(\widehat{y}) - \widehat{f}_k(\widehat{x})| = |\widehat{f}_k(F) - \widehat{f}_k(\widehat{x})| < \varepsilon.$$

Therefore, $F \in N(\widehat{x}, \mathscr{F}_1, \mathscr{F}_2, \ldots, \mathscr{F}_N, \varepsilon) \cap \widehat{X}$; which completes the proof of the claim. The result now follows from Proposition 2.5. $\square$

In what follows we will often use (without saying) the fact that if (Z, τ) is a topological space and $A \subseteq Y \subseteq Z$, then A is compact in Z if, and only if, A is compact in Y, with respect to the relative topology on Y.

Remark 2.25. *Together, Theorem 2.23 and Proposition 2.24 are essential for our future endeavours, as they provide a method for showing that a closed and bounded convex subset C of a normed linear space $(X, \|\cdot\|)$ is weakly compact. Namely, to show that C is weakly compact it is sufficient (and necessary) to show that $\overline{\widehat{C}}^{w^*} \subseteq \widehat{X}$. The reason for this is as follows: $\overline{\widehat{C}}^{w^*}$ is a weak* compact subset of (X^{**}, weak^*), by Theorem 2.23, and hence compact with respect to the relative weak* topology on $\widehat{X}$. Therefore, by Proposition 2.24, $\overline{\widehat{C}}^{w^*}$ is compact with respect to the relative weak topology on $\widehat{X}$. Further, by Proposition 2.14, $\overline{\widehat{C}}^{w^*}$ is compact with respect to the $\sigma(\widehat{X}, (\widehat{X})^*)$-topology on $\widehat{X}$ (i.e., the weak topology on $\widehat{X}$). Let $j : X \to \widehat{X}$ be the linear isometry defined by, $j(x) := \widehat{x}$ for all $x \in X$. Then $j^{-1} : \widehat{X} \to X$ is also a linear isometry. Thus, by Proposition 2.10, $C \subseteq j^{-1}(\overline{\widehat{C}}^{w^*})$ is weakly compact. Since C is closed and convex it is closed in the weak topology on X (see, Proposition 2.18). Hence C is compact with respect to the weak topology on X.*

As an example of this approach, we will give our first characterisation of reflexivity in terms of the weak compactness of the unit ball.

Theorem 2.26 ([9, page 425]). *Let $(X, \|\cdot\|)$ be a normed linear space. Then X is reflexive (i.e., $X^{**} = \widehat{X}$) if, and only if, B_X is compact with respect to the weak topology on X.*

Proof. Suppose that B_X is compact with respect to the weak topology on X. Then, by Proposition 2.10, $B_{\widehat{X}}$ is compact with respect to the weak topology on X^{**} as: (i) the mapping, $x \mapsto \widehat{x}$, is a bounded linear operator from X into X^{**} and (ii) the general fact that the continuous image of a compact set is compact. Now, since the weak* topology on X^{**} is weaker (and certainly no stronger) than the weak topology on X^{**}, $B_{\widehat{X}}$ is compact with respect to the weak* topology on X^{**}. Furthermore, since the weak* topology is Hausdorff, $B_{\widehat{X}}$ is closed with respect to the weak* topology on X^{**}. Thus, by Goldstine's Theorem, (Theorem 2.22)

$$B_{\widehat{X}} = \overline{B_{\widehat{X}}}^{w^*} = B_{X^{**}}.$$

So, $X^{**} = \bigcup_{n \in \mathbb{N}} n B_{X^{**}} = \bigcup_{n \in \mathbb{N}} n B_{\widehat{X}} = \widehat{X}$.

Conversely, suppose that $X^{**} = \widehat{X}$. Then $B_{X^{**}} = B_{\widehat{X}}$ and so by Theorem 2.23, $(B_{\widehat{X}}, \text{weak}^*)$ is compact. Since $B_{\widehat{X}} \subseteq \widehat{X}$ we have by Proposition 2.24 that $(B_{\widehat{X}}, \text{weak})$ is compact. Finally, since $x \mapsto \widehat{x}$ is a linear isometry from X onto X^{**} (since we are assuming that $X^{**} = \widehat{X}$), its inverse is a continuous linear operator (in fact an isometry as well) and so by Proposition 2.10, (B_X, weak) is compact too, as the continuous image of a compact set is compact. $\square$

3 James' Theorem on weak compactness

In this section we will provide three proofs of James' Theorem on weak compactness, listing them in order of increasing generality. First we provide a proof that is valid in all separable Banach spaces, then we give a proof that is valid in any Banach space whose dual ball is weak* sequentially compact, and then finally, we will present a proof that holds in all Banach spaces. It is our hope that this incremental approach to the full James' Theorem will make the final proof more accessible and less intimidating to the reader.

3.1 James' Theorem on weak compactness: the separable case

Convexity is the key to all our proofs of James' Theorem.

Let A be a nonempty convex subset of a vector space V and let $\varphi : A \to \mathbb{R}$ be a function. We say that φ is *convex* if

$$\varphi(\lambda x + (1-\lambda)y) \leq \lambda\varphi(x) + (1-\lambda)\varphi(y)$$

for all $x, y \in A$ and all $0 \leq \lambda \leq 1$.

Lemma 3.1 ([37]). *Let $0 < \beta$, $0 < \beta'$ and suppose that $\varphi : [0, \beta + \beta'] \to \mathbb{R}$ is a convex function. Then*

$$\frac{\varphi(\beta) - \varphi(0)}{\beta} \leq \frac{\varphi(\beta + \beta') - \varphi(\beta)}{\beta'}.$$

Proof. The inequality given in the statement of the lemma follows by rearranging the inequality

$$\varphi(\beta) \leq \frac{\beta}{\beta + \beta'}\varphi(\beta + \beta') + \frac{\beta'}{\beta + \beta'}\varphi(0). \qquad \square$$

Our first application of convexity is given next. It plays an important role in all three proofs of James' theorem.

Lemma 3.2 ([37]). *Let V 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*

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

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

Proof. We proceed in two parts. Firstly we prove that if $\beta_n r + \varphi(u) < \inf_{a \in A_n} \varphi(u + \beta_n a)$ for some $n \in \mathbb{N}$ and some $u \in (\sum_{i=0}^{n-1} \beta_i)A_1$, where $\beta_0 := 0$, then there exists an $a_n \in A_n$, such that

$$\beta_{n+1} r + \varphi(u + \beta_n a_n) < \inf_{a \in A_n} \varphi(u + \beta_n a_n + \beta_{n+1} a).$$

To see this, suppose that $u \in (\sum_{i=0}^{n-1} \beta_i)A_1$ and that $\beta_n r + \varphi(u) < \inf_{a \in A_n} \varphi(u + \beta_n a)$. Then there exists an $\varepsilon > 0$ such that

$$r + 2\varepsilon < \frac{\inf_{a \in A_n} \varphi(u + \beta_n a) - \varphi(u)}{\beta_n}. \qquad (*)$$

So, choose $a_n \in A_n$ such that $\varphi(u + \beta_n a_n) < \inf_{a \in A_n} \varphi(u + \beta_n a) + \beta_{n+1}\varepsilon$. Let $a \in A_n$. Then $v := (\beta_n a_n + \beta_{n+1} a)/(\beta_n + \beta_{n+1}) \in A_n$ (since A_n is convex) and so,

$$\begin{aligned}
r + 2\varepsilon &< \frac{\varphi(u + \beta_n v) - \varphi(u + 0v)}{\beta_n} && \text{(by $(*)$ and the fact that $v \in A_n$)} \\
&\leq \frac{\varphi(u + (\beta_n + \beta_{n+1})v) - \varphi(u + \beta_n v)}{\beta_{n+1}}. && \text{(by Lemma 3.1.)}
\end{aligned}$$

13

Rearranging gives
$$\beta_{n+1}(r + \varepsilon) + [\varphi(u + \beta_n v) + \beta_{n+1}\varepsilon] < \varphi(u + \beta_n a_n + \beta_{n+1}a),$$
for all $a \in A_n$. Since $\varphi(u + \beta_n a_n) < [\varphi(u + \beta_n v) + \beta_{n+1}\varepsilon]$, the desired inequality follows.

From this, we may inductively construct a sequence $(a_n : n \in \mathbb{N})$ with the requisite properties (i) and (ii). For the first step, we set $u := 0$ and then, by hypothesis, we have that

$$\beta_1 r + \varphi(0) < \inf_{a \in A_1} \varphi(\beta_1 a) = \inf_{a \in A_1} \varphi(0 + \beta_1 a).$$

So, by the first result, there exists an $a_1 \in A_1$, such that $\beta_2 r + \varphi(\beta_1 a_1) < \inf_{a \in A_1} \varphi(\beta_1 a_1 + \beta_2 a)$.

For the nth step, set $u := \sum_{i=1}^{n-1} \beta_i a_i$. Since $A_n \subseteq A_{n-1}$, and by the way a_{n-1} was constructed, we have that

$$\beta_n r + \varphi(u) < \inf_{a \in A_{n-1}} \varphi(u + \beta_n a) \le \inf_{a \in A_n} \varphi(u + \beta_n a).$$

So, by the first result again, there exists $a_n \in A_n$, such that

$$\beta_{n+1}r + \varphi\left(\sum_{i=1}^n \beta_i a_i\right) < \inf_{a \in A_n} \varphi\left(\sum_{i=1}^n \beta_i a_i + \beta_{n+1}a\right)$$

which completes the induction. The sequence $(a_n : n \in \mathbb{N})$ has the properties claimed above. $\square$

For the proof of James' theorem we will only require the following special case of this lemma.

Lemma 3.3. *Let V be a vector space (over $\mathbb{R}$) and let $\varphi : V \to \mathbb{R}$ be a sublinear function. 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 $\sum_{i=1}^\infty \beta_i < \infty$ and*

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

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

In order to formulate our first version of James' theorem on weak compactness we need to introduce the following notions.

Let K be a weak* compact convex subset of the dual of a Banach space $(X, \|\cdot\|)$. A subset B of K is called a *boundary* of K if for every $\widehat{x} \in \widehat{X}$ there exists a $b^* \in B$ such that $\widehat{x}(b^*) = \sup\{\widehat{x}(y^*) : y^* \in K\}$. We shall say B, *(I)-generates* K, if for every countable cover $(C_n : n \in \mathbb{N})$ of B by weak* compact convex subsets of K, the convex hull of $\bigcup_{n \in \mathbb{N}} C_n$ is norm dense in K. The following proof is found in [37].

Theorem 3.4 ([13, 14]). *Let K be a weak* compact convex subset of the dual of a Banach space $(X, \|\cdot\|)$ and let B be a boundary of K. Then B, (I)-generates K.*

Proof. After possibly translating K, we may assume that $0 \in B$. Let $\{C_n : n \in \mathbb{N}\}$ be weak* compact, convex subsets of K such that $B \subseteq \bigcup_{n \in \mathbb{N}} C_n$ and suppose, for a contradiction, that $\text{co}[\bigcup_{n \in \mathbb{N}} C_n]$ is not norm dense in K. Then there must exist an $0 < \varepsilon$ and $y^* \in K$ such that

$$y^* \in K \backslash (\text{co}[\textstyle\bigcup_{n \in \mathbb{N}} C_n] + \varepsilon B_{X^*})$$

Since, for each $n \in \mathbb{N}$, $\text{co}[\bigcup_{j=1}^n C_j]$ is weak* compact and convex, there exist $(\widehat{x}_n : n \in \mathbb{N})$ in $S_{\widehat{X}}$ such that

$$\max\{\widehat{x}_n(x^*) : x^* \in \text{co}[\textstyle\bigcup_{j=1}^n C_j]\} + \varepsilon = \max\{\widehat{x}_n(x^*) : x^* \in \text{co}[\textstyle\bigcup_{j=1}^n C_j] + \varepsilon B_{X^*}\} < \widehat{x}_n(y^*). \qquad (*)$$

for all $n \in \mathbb{N}$. Now, $(\widehat{x}_n(y^*) : n \in \mathbb{N})$ is a bounded sequence of real numbers and thus, has a convergent subsequence $(\widehat{x}_{n_k}(y^*) : k \in \mathbb{N})$. Let $s := \lim_{k \to \infty} \widehat{x}_{n_k}(y^*)$. Then, $\varepsilon \le s$ as $0 \in C_N$ for some $N \in \mathbb{N}$ and

$\varepsilon = 0 + \varepsilon < \widehat{x}_n(y^*)$ for all $N \leq n$, by the inequality $(*)$. After relabelling the sequence $(\widehat{x}_n : n \in \mathbb{N})$ if necessary, we may assume that $\lim_{n \to \infty} \widehat{x}_n(y^*) = s$ and $|\widehat{x}_n(y^*) - s| < \varepsilon/3$ for all $n \in \mathbb{N}$. Importantly, we note that this relabelling does not disturb the inequality in $(*)$.

We define $A_n := \mathrm{co}\{\widehat{x}_k : n \leq k\}$ for all $n \in \mathbb{N}$ and note that: (i) $(A_n : n \in \mathbb{N})$ is a decreasing sequence of nonempty convex subsets of $\widehat{X}$ and (ii) if $N < n$ and $b^* \in C_N$ then

$$g(b^*) < [g(y^*) - \varepsilon] \qquad \text{for all } g \in A_n \qquad\qquad (**)$$

since, $\{\widehat{x}_k : n \leq k\} \subseteq \{\widehat{x} \in \widehat{X} : \widehat{x}(b^* - y^*) < -\varepsilon\}$; which is convex. Next, we define $p : \widehat{X} \to \mathbb{R}$ by,

$$p(\widehat{x}) := \sup_{x^* \in K} \widehat{x}(x^*) \quad \text{for all } \widehat{x} \in \widehat{X}.$$

Then p defines a sublinear functional on $\widehat{X}$. Moreover, for all $g \in A_1$, we have $(s - \varepsilon/3) < g(y^*)$ since $\{\widehat{x}_n : n \in \mathbb{N}\} \subseteq \{\widehat{x} \in \widehat{X} : (s - \varepsilon/3) < \widehat{x}(y^*)\}$; which is convex. Therefore, $(s - \varepsilon/3) < p(g)$ for all $g \in A_1$, since $y^* \in K$.

Let $(\beta_n : n \in \mathbb{N})$ be any sequence of positive numbers such that $\lim_{n \to \infty} \left(\sum_{i=n+1}^{\infty} \beta_i\right)/\beta_n = 0$. Now,

$$(s - \varepsilon/2) < (s - \varepsilon/3) \leq \inf_{g \in A_1} p(g).$$

Therefore, by Lemma 3.3, there exists a sequence $(g_n : n \in \mathbb{N})$ in $\widehat{X}$ such that $g_n \in A_n$ and

$$p(\textstyle\sum_{i=1}^{n} \beta_i g_i) + \beta_{n+1}(s - \varepsilon/2) < p(\textstyle\sum_{i=1}^{n+1} \beta_i g_i) \quad \text{for all } n \in \mathbb{N}. \qquad (***)$$

Note also, that since $\lim_{n \to \infty} \widehat{x}_n(y^*) = s$, $\lim_{n \to \infty} g_n(y^*) = s$. As $\|g_n\| \leq 1$ for all $n \in \mathbb{N}$, we have that $\sum_{i=1}^{\infty} \|\beta_i g_i\| \leq \sum_{i=1}^{\infty} \beta_i < \infty$. Since X is a Banach space, this implies that $g := \sum_{i=1}^{\infty} \beta_i g_i \in \widehat{X}$. Because p is continuous, this implies that $(p(\sum_{i=1}^{n} \beta_i g_i) : n \in \mathbb{N})$ is a convergent and hence a bounded sequence in $\mathbb{R}$. Moreover, by the inequality $(***)$, we have that $(p(\sum_{i=1}^{n} \beta_i g_i) : n \in \mathbb{N})$ is an increasing sequence. Therefore, by the Monotone Convergence Theorem, $(p(\sum_{i=1}^{n} \beta_i g_i) : n \in \mathbb{N})$ converges to its supremum. That is,

$$\sup_{n \in \mathbb{N}} p(\textstyle\sum_{i=1}^{n} \beta_i g_i) = \lim_{n \to \infty} p(\textstyle\sum_{i=1}^{n} \beta_i g_i) = p(\lim_{n \to \infty} \textstyle\sum_{i=1}^{n} \beta_i g_i) = p(g). \qquad (****)$$

Since $g \in \widehat{X}$ and B is a boundary for K, there must exist a $b^* \in B$ such that

$$g(b^*) = \sup\{g(x^*) : x^* \in K\} = p(g).$$

Let $n \in \mathbb{N}$, then

$$\begin{aligned}
\beta_n(s - \varepsilon/2) &< p(\textstyle\sum_{i=1}^{n} \beta_i g_i) - p\left(\textstyle\sum_{i=1}^{n-1} \beta_i g_i\right) \qquad \text{by } (***) \\
&\leq p(g) - p\left(\textstyle\sum_{i=1}^{n-1} \beta_i g_i\right) \qquad \text{by } (****) \\
&= g(b^*) - p\left(\textstyle\sum_{i=1}^{n-1} \beta_i g_i\right) \\
&\leq g(b^*) - (\textstyle\sum_{i=1}^{n-1} \beta_i g_i)(b^*) = \textstyle\sum_{i=n}^{\infty} \beta_i g_i(b^*).
\end{aligned}$$

Since $B \subseteq \bigcup_{n \in \mathbb{N}} C_n$, $b^* \in C_N$ for some $N \in \mathbb{N}$. Thus, if $N < n$, then

$$(s - \varepsilon/2) < \frac{1}{\beta_n}\left(\textstyle\sum_{i=n+1}^{\infty} \beta_i g_i(b^*)\right) + g_n(b^*) < \frac{1}{\beta_n}\left(\textstyle\sum_{i=n+1}^{\infty} \beta_i g_i(b^*)\right) + [g_n(y^*) - \varepsilon] \quad \text{by } (**),$$

since $g_n \in A_n$. By taking the limit, as n tends to infinity, we get that $(s - \varepsilon/2) \leq (s - \varepsilon)$; which is impossible. Therefore, B, (I)-generates K. $\square$

Remark 3.5. If $\beta_n := \dfrac{1}{n!}$ for all $n \in \mathbb{N}$ or, $\beta_n := \dfrac{1}{2^{n^2}}$ for all $n \in \mathbb{N}$, then $\lim_{n \to \infty} \dfrac{\sum_{i=n+1}^{\infty} \beta_i}{\beta_n} = 0$.

Theorem 3.6 (James' Theorem: version 1, [19]). *Let C be a closed and bounded convex subset of a Banach space $(X, \|\cdot\|)$. If C is separable and every continuous linear functional on X attains its supremum over C, then C is weakly compact.*

Proof. Let $K := \overline{\widehat{C}}^{w^*}$. To show that C is weakly compact it is sufficient to show $K \subseteq \widehat{X}$, (see Remark 2.25). In fact, since X is a Banach space and $x \mapsto \widehat{x}$ is a linear isometry, we have that $\widehat{X}$ is a Banach subspace of X^{**} and so a closed subspace of X^{**}. Therefore, it is sufficient to show that for every $0 < \varepsilon$, $K \subseteq \widehat{X} + 2\varepsilon B_{X^{**}}$. To this end, fix $0 < \varepsilon$ and let $\{x_n : n \in \mathbb{N}\}$ be a dense subset of C. For each $n \in \mathbb{N}$, let $K_n^\varepsilon := K \cap [\widehat{x_n} + \varepsilon B_{X^{**}}]$. Then $(K_n^\varepsilon : n \in \mathbb{N})$ is a cover of $\widehat{C}$ by weak* closed convex subsets of K. Since $\widehat{C}$ is a boundary of K, we have that $K \subseteq \overline{\mathrm{co}} \bigcup_{n \in \mathbb{N}} K_n^\varepsilon \subseteq \widehat{X} + 2\varepsilon B_{X^{**}}$; which completes the proof. $\square$

By working a bit harder, we could extend this approach to proving James' theorem, via (I)-generation, to spaces whose dual ball is weak* sequentially compact. Indeed, this is done in the paper [34]. However, in this paper we will take another tack. We will prove James' theorem, in the case when the dual ball is weak* sequentially compact, in a way that naturally extends to the general case, albeit requiring several extra technical steps regarding the extraction of subsequences with "small" sets of cluster points.

One of the strengths of Theorem 3.6 is that it essentially only relies upon a separation argument (Theorem 2.17) and Lemma 3.3. In this way we see that this proof is very elementary.

3.2 James' Theorem on weak compactness: the weak* sequentially compact case

We shall start this subsection with two simple preliminary results.

Proposition 3.7 ([11, Corollary 49]). *Let $(X, \|\cdot\|)$ be a normed linear space. Then every finite-dimensional subspace of X^* is weak*-closed.*

Proof. Suppose that $Y := \mathrm{span}\{x_1^*, \ldots, x_n^*\}$ is a finite-dimensional subspace of X^* and let $x_0^* \notin Y$. Then, by Lemma 2.7, we have that $\bigcap_{i=1}^n \ker(x_i^*) \not\subseteq \ker(x_0^*)$. So, let

$$x \in \bigcap_{i=1}^n \ker(x_i^*) \setminus \ker(x_0^*).$$

Then, taking $-x$ if need be, we may assume that $x_0^*(x) > 0$, while $x_i^*(x) = 0$ for all $1 \le i \le n$. Observe that $Y = \mathrm{span}\{x_1^*, \ldots, x_n^*\} \subseteq \ker(\widehat{x})$, since $\ker(\widehat{x})$ is a subspace and $x_i^* \in \ker(\widehat{x})$ for all $1 \le i \le n$. So $y^*(x) = 0$ for all $y^* \in Y$. Thus,

$$\{x^* \in X^* : x^*(x) > 0\} = \widehat{x}^{-1}(0, \infty)$$

is a weak*-open neighbourhood of x_0^*, which is disjoint from Y. Since x_0^* was arbitrary, we have that Y is weak*-closed. $\square$

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

Proof. Let Y be a finite-dimensional subspace of X^* such that $\mathrm{dist}(x^*, Y) > \varepsilon > 0$. Then we have that $x^* \notin Y + \varepsilon B_{X^*}$. Since Y is weak*-closed (Proposition 3.7) and convex, and B_{X^*} is weak*-compact (Theorem 2.23) and convex, we have that $Y + \varepsilon B_{X^*}$ is also weak*-closed and convex. Therefore, by Theorem 2.17, there exists an $x \in S_X$ such that

$$\begin{aligned}
x^*(x) &> \sup\{y^*(x) : y^* \in Y + \varepsilon B_{X^*}\} \\
&= \sup\{y^*(x) : y^* \in Y\} + \varepsilon \ge \varepsilon.
\end{aligned}$$

Finally observe that for this x, we have that $\widehat{x}(Y)$ is bounded above, and since Y is a subspace, the only way this is possible is if $\widehat{x}(y^*) = y^*(x) = 0$ for all $y^* \in Y$. $\square$

Theorem 3.9 (James' Theorem: version 2). *Let C be a closed and bounded convex subset of a Banach space $(X, \|\cdot\|)$. If (B_{X^*}, weak^*) is sequentially compact and every $x^* \in X^*$ attains its supremum over C, then C is weakly compact.*

Proof. To show that C is weakly compact, it is sufficient to show that $K := \overline{\widehat{C}}^{w*} \subseteq \widehat{X}$ (see, Remark 2.25). Suppose, for a contradiction, that this is not the case. Then, there exists an $F \in K \backslash \widehat{X}$. Since X is a Banach space, $\widehat{X}$ is a closed subspace of X^{**}, and so there must exist an $0 < \varepsilon < \text{dist}(F, \widehat{X})$. Let $(\beta_n : n \in \mathbb{N})$ be a sequence of strictly positive numbers such that $\lim_{n \to \infty} \frac{1}{\beta_n} \sum_{i=n+1}^{\infty} \beta_i = 0$.

Part I: Let $f_0 := 0$. We inductively create sequences $(f_n : n \in \mathbb{N})$ in S_{X^*} and $(\widehat{x}_n : n \in \mathbb{N})$ in $\widehat{C}$, such that the statements

(A_n) $|(F - \widehat{x}_n)(f_j)| < \varepsilon/2$ for all $0 \le j < n$.

(B_n) $F(f_n) > \varepsilon$ and $\widehat{x}_j(f_n) = 0$ for all $1 \le j \le n$.

are true for all $n \in \mathbb{N}$. For the first step, choose any $\widehat{x}_1 \in \widehat{C}$. Then it is clear that $|(F - \widehat{x}_1)(f_0)| = 0 < \varepsilon/2$. Now note that
$$\text{dist}(F, \text{span}\{\widehat{x}_1\}) \ge \text{dist}(F, \widehat{X}) > \varepsilon.$$
So, by Lemma 3.8, there exists $f_1 \in S_{X^*}$ such that $F(f_1) > \varepsilon$ and $\widehat{x}_1(f_1) = 0$. So the statements (A_1) and (B_1) hold.

Now fix $k \in \mathbb{N}$. Suppose that we have created $\{\widehat{x}_1, \ldots, \widehat{x}_k\}$ and $\{f_1, \ldots, f_k\}$ such that the statements (A_k) and (B_k) hold true. Then consider the set
$$W := \bigcap_{j=0}^{k} \{G \in X^{**} : |(F - G)(f_j)| < \varepsilon/2\}.$$

Since W is a weak*-open neighbourhood of F and $F \in \overline{\widehat{C}}^{w^*}$, we can choose $\widehat{x}_{k+1} \in \widehat{C}$ such that $\widehat{x}_{k+1} \in W$ i.e., such that the statement (A_{k+1}) holds. Next, observe that
$$\text{dist}(F, \text{span}\{\widehat{x}_1, \ldots, \widehat{x}_{k+1}\}) \ge \text{dist}(F, \widehat{X}) > \varepsilon.$$

So, by Lemma 3.8, there exists $f_{k+1} \in S_{X^*}$ such that $F(f_{k+1}) > \varepsilon$ and $\widehat{x}_j(f_{k+1}) = 0$ for all $1 \le j \le k+1$. Therefore the statement (B_{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 $f_k' := f_{n_k}$ and $x_k' := x_{n_k}$. Also define $f_0' := 0$. Then the sequences $(\widehat{x}_n' : n \in \mathbb{N})$ and $(f_n' : n \in \mathbb{N})$ still satisfy (A_n) and (B_n) for all $n \in \mathbb{N}$. Therefore, passing to a subsequence does not disturb the statements (A_n) and (B_n).

Now, as (B_{X^*}, weak^*) is sequentially compact, and $(f_n : n \in \mathbb{N})$ is a sequence in B_{X^*}, we have that $(f_n : n \in \mathbb{N})$ has a weak*-convergent subsequence. So, by passing to subsequences and relabelling if necessary, we may assume that $(f_n : n \in \mathbb{N})$ is weak*-convergent to some $f_\infty \in B_{X^*}$. By the above, we know that the statements (A_n) and (B_n) remain true for all $n \in \mathbb{N}$.

Part III: Let $k \in \mathbb{N}$. For any $n \ge k$, we have that $\widehat{x}_k(f_n) = 0$ by the statement (B_n). Therefore, it follows that $\widehat{x}_k(f_\infty) = 0$. Since k was arbitrary, this is true for all $k \in \mathbb{N}$.

On the other hand, let $k \in \mathbb{N}$ and let $n > k$. Then, by the statement (A_n), we have that $|(F - \widehat{x}_n)(f_k)| < \varepsilon/2$. Moreover, from (B_k), we know that $F(f_k) > \varepsilon$. Combining these, we get that
$$\widehat{x}_n(f_k) = F(f_k) + (\widehat{x}_n - F)(f_k) > \varepsilon/2$$
for all $n > k$. Therefore $\widehat{x}_n(f_k - f_\infty) > \varepsilon/2$, for all $n > k$.

Part IV: For each $n \in \mathbb{N}$, define $C_n := \text{co}\{f_k : k \ge n\} - f_\infty$ and note that $(C_n : n \in \mathbb{N})$ is a decreasing sequence of nonempty, convex subsets of X^*. Define $p : X^* \to \mathbb{R}$ to be $p(x^*) = \sup\{x^*(c) : c \in C\}$ for all $x^* \in X^*$. Then p is a sublinear function and $\inf_{f \in C_1} p(f) > \varepsilon/4$.

17

To confirm this second assertion consider the following. Let $f \in C_1$. Then $f = \sum_{i=1}^{k} \lambda_i f_{n_i} - f_\infty$ where $0 \leq \lambda_i$ for all $1 \leq i \leq k$ and $\sum_{i=1}^{k} \lambda_i = 1$. Let $m > \max\{n_1, \ldots, n_k\}$. Then

$$p(f) \geq f(x_m) = \widehat{x}_m\left(\sum_{i=1}^{k}\lambda_i f_{n_i} - f_\infty\right) = \sum_{i=1}^{k}\lambda_i\widehat{x}_m(f_{n_i} - f_\infty) > \varepsilon/2.$$

Therefore, since $f \in C_1$ was arbitrary, we have that $\inf_{f \in C_1} p(f) > \varepsilon/4$ as claimed. So, by Lemma 3.3, there exists a sequence $(g_n : n \in \mathbb{N})$ such that for all $n \in \mathbb{N}$:

(i) $g_n \in \mathrm{co}\{f_k : k \geq n\}$ and

(ii) $p(\sum_{i=1}^{n} \beta_i(g_i - f_\infty)) + \beta_{n+1}\varepsilon/4 < p(\sum_{i=1}^{n+1} \beta_i(g_i - f_\infty)).$ ($*$)

Part V: Now, since (X^*, weak^*) is a locally convex space $(g_n : n \in \mathbb{N})$ also converges to $g_\infty := f_\infty$. Indeed, if W is any convex weak* open neighbourhood of f_∞ then there exists an $N \in \mathbb{N}$ such that $f_n \in W$ for all $n \geq N$. Therefore, $\mathrm{co}\{f_i : i \geq N\} \subseteq W$. Since $g_k \in \mathrm{co}\{f_i : i \geq N\}$ for all $k \geq N$, we have that $g_k \in W$ for all $k \geq N$. This shows that $(g_k : k \in \mathbb{N})$ converges to $g_\infty = f_\infty$. Set $g := \sum_{i=1}^{\infty} \beta_i(g_i - f_\infty)$. Since $\|g_i - f_\infty\| \leq 2$ for all $i \in \mathbb{N}$, we have that

$$\sum_{i=1}^{\infty} \|\beta_i(g_i - f_\infty)\| = \sum_{i=1}^{\infty} \beta_i\|g_i - f_\infty\| \leq 2\sum_{i=1}^{\infty} \beta_i < \infty.$$

Therefore, $g \in X^*$ since X^* is a Banach space. As p is continuous, it is clear that $(p(\sum_{i=1}^{n} \beta_i(g_i - f_\infty)) : n \in \mathbb{N})$ is a convergent and hence a bounded sequence in $\mathbb{R}$. Moreover, the statement ($*$) gives that this is also an increasing sequence. Therefore, by the Monotone Convergence Theorem, $(p(\sum_{i=1}^{n} \beta_i(g_i - f_\infty)) : n \in \mathbb{N})$ converges to its supremum. That is,

$$\sup_{n \in \mathbb{N}} p(\sum_{i=1}^{n}\beta_i(g_i - f_\infty)) = \lim_{n \to \infty} p(\sum_{i=1}^{n}\beta_i(g_i - f_\infty)) = p(\lim_{n \to \infty}\sum_{i=1}^{n}\beta_i(g_i - f_\infty)) = p(g).$$

Part VI: Since $g \in X^*$, there exists a $c \in C$ such that $\widehat{c}(g) = g(c) = \sup\{g(x) : x \in C\} = p(g)$. Then, for any $n > 1$,

$$\begin{aligned}
\beta_n\varepsilon/4 &< p(\sum_{i=1}^{n}\beta_i(g_i - f_\infty)) - p(\sum_{i=1}^{n-1}\beta_i(g_i - f_\infty)) &&\text{by } (*) \\
&\leq p(g) - p(\sum_{i=1}^{n-1}\beta_i(g_i - f_\infty)) &&\text{since } p(g) = \sup\{p(\sum_{i=1}^{n}\beta_i(g_i - f_\infty)) : n \in \mathbb{N}\} \\
&= \widehat{c}(g) - p(\sum_{i=1}^{n-1}\beta_i(g_i - f_\infty)) \\
&\leq \widehat{c}(g) - \widehat{c}(\sum_{i=1}^{n-1}\beta_i(g_i - f_\infty)) = \widehat{c}(\sum_{i=n}^{\infty}\beta_i(g_i - f_\infty)) = \beta_n\widehat{c}(g_n - f_\infty) + \sum_{i=n+1}^{\infty}\beta_i\widehat{c}(g_i - f_\infty).
\end{aligned}$$

Rearranging gives that

$$\varepsilon/4 < \widehat{c}(g_n - f_\infty) + \frac{1}{\beta_n}\sum_{i=n+1}^{\infty}\beta_i\widehat{c}(g_i - f_\infty) \leq \widehat{c}(g_n - f_\infty) + \frac{2\|\widehat{c}\|}{\beta_n}\sum_{i=n+1}^{\infty}\beta_i.$$

Taking the limit, as n tends to infinity, we get that

$$\varepsilon/4 \leq \lim_{n \to \infty} g_n(c) - f_\infty(c) + 2\|\widehat{c}\|\left(\lim_{n \to \infty}\frac{1}{\beta_n}\sum_{i=n+1}^{\infty}\beta_i\right) = \lim_{n \to \infty} g_n(c) - f_\infty(c),$$

which contradicts the fact that $f_\infty(c) = \lim_{n \to \infty} g_n(c)$. Therefore, $K \subseteq \widehat{X}$ and so C is weakly compact. $\qquad\square$

The power of this result stems from the fact that the class of all Banach spaces whose dual ball is weak* sequentially compact is very large. Indeed, in addition to all the separable Banach spaces (whose dual ball is weak* metrisable, [9, page 426]), it contains all Asplund spaces, [31] (i.e., spaces in which every separable subspace has a separable dual space) and all spaces that admit an equivalent smooth norm, [18] (which includes all WCG spaces, [8]). In fact, it contains all Gateaux differentiability spaces, [31].

3.3 James' Theorem on weak compactness: the general case

The short-coming of the previous subsection is that the dual ball of a Banach space need not be weak* sequentially compact. For example, the dual ball of $(C(\beta\mathbb{N}), \|\cdot\|_\infty)$ is not weak* sequentially compact, as it contains a copy of $\beta\mathbb{N}$ - the Stone-Cech compactification of the natural numbers, endowed with the discrete topology, and this space is known to have no nontrivial (i.e., not eventually constant) convergent sequences, [10].

So the method of passing to a subsequence which is weak* convergent must be abandoned. However we can, by passing to a suitable subsequence, insist that $K := \bigcap_{n\in\mathbb{N}} \overline{\{f_k : k \geq n\}}^{w^*}$ is "small" in the sense that for countably many weak* lower semicontinuous real-valued functions $(p_n : n \in \mathbb{N})$, the sets $p_n(K)$ are singletons. In this way, the set K of all weak* cluster points of the sequence $(f_n : n \in \mathbb{N})$ "acts" like a singleton set in **Part V** and **Part VI** of the proof of Theorem 3.9.

So next we will show how to extract "nice" subsequences from a given sequence. The approach we adopt is very general and will provide much more than needed, but these technical results may possibly be of some independent interest.

We shall start with the precise definition 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.

Since we will be working extensively with subsequences we will introduce some concise notation for a subsequence of a given sequence. Let $\widetilde{x} : \mathbb{N} \to X$ denote the sequence $(x_n : n \in \mathbb{N})$ and let J be an infinite subset of $\mathbb{N}$. Then J can be *uniquely* enumerated as, $J = \{n_k : k \in \mathbb{N}\}$ with $n_k < n_{k+1}$ for all $k \in \mathbb{N}$. Then $\widetilde{x}|_J$ will denote the subsequence $(x_{n_k} : k \in \mathbb{N})$ of $\widetilde{x} = (x_n : n \in \mathbb{N})$. We will also be working with the set of all cluster points of a given sequence and so it is worth our while to introduce some notation for the set of all cluster points. Let (X, τ) be a linear topological space and let $\widetilde{x} : \mathbb{N} \to X$ be the sequence $(x_n : n \in \mathbb{N})$. We define

$$cl_\tau(\widetilde{x}) := \bigcap_{n=1}^{\infty} \overline{\{x_k : k \geq n\}}^\tau.$$

That is, $cl_\tau(\widetilde{x})$ is the set of all τ-cluster points of $\widetilde{x}$. Further, we define $K_\tau(\widetilde{x}) := \overline{co}^\tau(cl_\tau(\widetilde{x}))$. When there is no ambiguity concerning the topology, we will simply write $cl(\widetilde{x})$ and $K(\widetilde{x})$.

Lemma 3.10. *If $\varphi : A \to \mathbb{R}$ is a convex lower-semicontinuous function defined on a nonempty closed and convex subset A of a Hausdorff locally convex space $(X, +, \cdot, \tau)$, then for every sequence $\widetilde{x} := (x_n : n \in \mathbb{N})$ in A, either there exists a subsequence without any cluster points, or else, there exists a subsequence $\widetilde{x}|_J$ such that φ is bounded on $K(\widetilde{x}|_J)$.*

Proof. Let $\widetilde{x} := (x_n : n \in \mathbb{N})$ be a sequence in A and suppose that every subsequence of $\widetilde{x}$ has a cluster point. We shall construct a subsequence $\widetilde{x}|_J$ that that φ is bounded on $K(\widetilde{x}|_J)$, but first, we shall construct an infinite subset J' of $\mathbb{N}$ such that φ is bounded below on $K(\widetilde{x}|_{J'})$.

Let $x \in cl(\widetilde{x})$. Then $x \notin \varphi^{-1}(-\infty, \varphi(x) - 1]$, which is closed and convex. Therefore, there exists a closed and convex neighbourhood, N of x such that $N \cap \varphi^{-1}(-\infty, \varphi(x) - 1] = \varnothing$ i.e., $\varphi(N) \subseteq (\varphi(x) - 1, \infty)$. Since x is a cluster point of $\widetilde{x}$, we may choose an infinite set $J' \subseteq \mathbb{N}$ such that $x_j \in N$ for all $j \in J'$. Then, because N is closed and convex, $K(\widetilde{x}|_{J'}) \subseteq N$ and so $\varphi(K(\widetilde{x}|_{J'})) \subseteq \varphi(N) \subseteq (\varphi(x) - 1, \infty)$. Hence $\varphi(K(\widetilde{x}|_{J'}))$ is bounded below.

We now claim that J' possesses an infinite subset J such that $\varphi(K(\widetilde{x}|_J))$ is bounded above. Indeed, suppose in order to obtain a contradiction, that this is not the case. Then we inductively proceed as follows. First, there must be $x \in cl(\widetilde{x}|_{J'})$ with $\varphi(x) > 1$, otherwise $\varphi(K(\widetilde{x}|_{J'})) \subseteq (-\infty, 1]$ and we would be done. So, we may choose a closed, convex neighbourhood, N of x such that $N \cap \varphi^{-1}(-\infty, 1] = \varnothing$. Then, since x is a cluster point of $\widetilde{x}|_{J'}$, we can choose an infinite subset $J_1 \subseteq J'$ such that $x_j \in N$ for all $j \in J_1$. Because N is closed and convex, we have that $K(\widetilde{x}|_{J_1}) \subseteq N$ and so $K(\widetilde{x}|_{J_1}) \cap \varphi^{-1}(-\infty, 1] = \varnothing$.

In general, suppose that we have chosen infinite subsets $J_n \subseteq \cdots \subseteq J_1 \subseteq J'$ such that for all $1 \leq i \leq n$: $K(\widetilde{x}|_{J_i}) \cap \varphi^{-1}(-\infty, i] = \varnothing$.

For the $(n+1)^{th}$ step, we suppose that $cl(\widetilde{x}|_{J_n}) \not\subseteq \varphi^{-1}(-\infty, n+1]$, otherwise $\varphi(K(\widetilde{x}|_{J_n})) \subseteq (-\infty, n+1]$ is bounded above and we are done. Therefore, we can choose $x \in cl(\widetilde{x}|_{J_n})$ such that $\varphi(x) > n+1$, and a

19

closed, convex neighbourhood, N of x such that $N \cap \varphi^{-1}(-\infty, n+1] = \varnothing$. Then, since x is a cluster point of $\widetilde{x}|_{J_n}$, we can choose an infinite subset $J_{n+1} \subseteq J_n$ such that $x_j \in N$ for all $j \in J_{n+1}$. Because N is closed and convex, we have that $K(\widetilde{x}|_{J_{n+1}}) \subseteq N$ and so $K(\widetilde{x}|_{J_{n+1}}) \cap \varphi^{-1}(-\infty, n+1] = \varnothing$. This completes the induction.

Lastly, we apply the so-called diagonalisation argument. Define $J'' := \{n_k : k \in \mathbb{N}\} \subseteq \mathbb{N}$ such that $n_k < n_{k+1}$ and $n_k \in J_k$ for all $k \in \mathbb{N}$.

Consider the subsequence of $\widetilde{x}$ given by $\widetilde{x}|_{J''} = (x_{n_k} : k \in \mathbb{N})$. Then, since $J_{n+1} \subseteq J_n$ for all $n \in \mathbb{N}$, we have that $n_k \in J_m$ for all $k \geq m$. Let $m \in \mathbb{N}$. Then

$$K(\widetilde{x}|_{J''}) = K(\{x_{n_k} : k \geq m\}) \subseteq K(\widetilde{x}|_{J_m}) \subseteq X \setminus \varphi^{-1}(-\infty, m],$$

and so $K(\widetilde{x}|_{J''}) \cap \varphi^{-1}(-\infty, m] = \varnothing$. Since m was arbitrary, this holds for all $m \in \mathbb{N}$ and so we have that $\varphi(K(\widetilde{x}|_{J''})) = \varnothing$, which contradicts our original assumption. Thus, there exists a subsequence $\widetilde{x}|_J$ of $\widetilde{x}$ such that $\varphi(K(\widetilde{x}|_J))$ is bounded. $\square$

We can further refine Lemma 3.10 as follows.

Lemma 3.11. *If $\varphi : A \to \mathbb{R}$ is a convex lower-semicontinuous function defined on a nonempty closed and convex subset A of a Hausdorff locally convex space $(X, +, \cdot, \tau)$, then for every sequence $\widetilde{x} := (x_n : n \in \mathbb{N})$ in A, either there exists a subsequence without any cluster points, or else, there exists a subsequence $\widetilde{x}|_J$ such that φ is constant on $K(\widetilde{x}|_J)$.*

Proof. Let $\widetilde{x} := (x_n : n \in \mathbb{N})$ be a sequence in A and suppose that every subsequence of $\widetilde{x}$ has a cluster point. Then, by Lemma 3.10, and by passing to a subsequence if necessary, we may assume that φ is bounded on $K(\widetilde{x})$. Let $\alpha_1, \beta_1 \in \mathbb{R}$ denote $\inf \varphi(K(\widetilde{x}))$ and $\sup \varphi(K(\widetilde{x}))$ respectively and let $J_0 := \mathbb{N}$. Of course if $\alpha_1 = \beta_1$, then $\varphi(K(\widetilde{x}))$ is a singleton and we are done. If not, we inductively construct a decreasing sequence of infinite subsets $(J_n : n \in \mathbb{N})$ of $\mathbb{N}$ such that $\operatorname{diam}(\varphi(K(\widetilde{x}|_{J_n})) \leq (\beta_1 - \alpha_1)/2^n$ for all $n \in \mathbb{N}$.

We begin as follows. Set $\delta_1 := (\alpha_1 + \beta_1)/2$. Since φ is convex and lower-semicontinuous, we have that $\varphi^{-1}(-\infty, \delta_1]$ is a closed, convex set. Then, we can pick $x \in cl(\widetilde{x})$ such that $\delta_1 < \varphi(x) \leq \beta_1$. Indeed, if not, then $\varphi(cl(\widetilde{x})) \subseteq (-\infty, \delta_1]$ and so $\varphi(K(\widetilde{x})) \subseteq (-\infty, \delta_1]$ also. However, this contradicts the fact that $\beta_1 = \sup \varphi(K(\widetilde{x}))$.

Therefore $x \notin \varphi^{-1}(-\infty, \delta_1]$, and so there exists a closed, convex neighbourhood, N of x such that $N \cap \varphi^{-1}(-\infty, \delta_1] = \varnothing$. As $x \in cl(\widetilde{x})$, there is an infinite set $J_1 \subseteq \mathbb{N}$ such that $x_j \in N$ for all $j \in J_1$. In particular, $K(\widetilde{x}|_{J_1}) \subseteq N$, since N is closed and convex, and so $\inf \varphi(K(\widetilde{x}|_{J_1})) \geq \delta_1$. Also, because $\widetilde{x}|_{J_1}$ is a subsequence of $\widetilde{x}$, we have that $\sup \varphi(K(\widetilde{x}|_{J_1})) \leq \beta_1$. Therefore, $\operatorname{diam}(\varphi(K(\widetilde{x}|_{J_1})) \leq \beta_1 - \delta_1 = (\beta_1 - \alpha_1)/2$.

Suppose now that we have created the infinite subsets $J_n \subseteq J_{n-1} \subseteq \cdots \subseteq J_0$ such that

$$\operatorname{diam}(\varphi(K(\widetilde{x}|_{J_i})) \leq (\beta_1 - \alpha_1)/2^i \quad \text{for all } 1 \leq i \leq n.$$

Set $\alpha_n := \inf \varphi(K(\widetilde{x}|_{J_n}))$, $\beta_n := \sup \varphi(K(\widetilde{x}|_{J_n}))$ and $\delta_n := (\alpha_n + \beta_n)/2$. Then,

$$\operatorname{diam}(\varphi(K(\widetilde{x}|_{J_n})) = \beta_n - \alpha_n \leq (\beta_1 - \alpha_1)/2^n,$$

by construction. If $\alpha_n = \beta_n$ then let $J_{n+1} := J_n$ and we are done. Otherwise, we can choose (as above) $x \in cl(\widetilde{x}|_{J_n})$ such that $x \notin \varphi^{-1}(-\infty, \delta_n]$, because if not, $\varphi(K(\widetilde{x}|_{J_n})) \subseteq (-\infty, \delta_n]$, which contradicts the fact that $\beta_n = \sup \varphi(K(\widetilde{x}|_{J_n}))$. Therefore, there exists a closed, convex neighbourhood, N of x such that $N \cap \varphi^{-1}(-\infty, \delta_n] = \varnothing$. Since $x \in cl(\widetilde{x}|_{J_n})$, there is an infinite set $J_{n+1} \subseteq J_n$ such that $x_j \in N$ for all $j \in J_{n+1}$. In particular, since N is closed and convex, $K(\widetilde{x}|_{J_{n+1}}) \subseteq N$ and so $\inf \varphi(K(\widetilde{x}|_{J_{n+1}})) \geq \delta_n$. Therefore,

$$\operatorname{diam}(\varphi(K(\widetilde{x}|_{J_{n+1}})) \leq \beta_n - \delta_n = (\beta_n - \alpha_n)/2 \leq (\beta_1 - \alpha_1)/2^{n+1}.$$

Thus, by induction, we have created a decreasing sequence of infinite subsets $(J_n : n \in \mathbb{N})$ of $\mathbb{N}$ such that

$$\operatorname{diam}(\varphi(K(\widetilde{x}|_{J_n})) \leq (\beta_1 - \alpha_1)/2^n \quad \text{for all } n \in \mathbb{N}.$$

Lastly, define $J := \{n_k : k \in \mathbb{N}\}$ such that $n_k < n_{k+1}$ and $n_k \in J_k$ for all $k \in \mathbb{N}$. Consider the subsequence of $\widetilde{x}$ given by $\widetilde{x}|_J = (x_{n_k} : k \in \mathbb{N})$. Then, since $J_{n+1} \subseteq J_n$ for all $n \in \mathbb{N}$, we have that $n_k \in J_m$ for all $k \geq m$. Let $m \in \mathbb{N}$. Then,

$$K(\widetilde{x}|_J) = K(\{x_{n_k} : k \geq m\}) \subseteq K(\widetilde{x}|_{J_m}),$$

which gives that $\operatorname{diam}(\varphi(K(\widetilde{x}|_J))) \leq \operatorname{diam}(\varphi(K(\widetilde{x}|_{J_m}))) \leq (\beta_1 - \alpha_1)/2^m$. Since $m \in \mathbb{N}$ was arbitrary, we conclude that $\varphi(K(\widetilde{x}|_J))$ is a singleton as required. $\quad\square$

For our next result we need to recall the definition of the topology of pointwise convergence. If X is a nonempty set and A is a nonempty subset of X then we may put a topology on the vector space $\mathbb{R}^X$ of all real-valued functions defined on X endowed with pointwise addition and pointwise scalar multiplication. We will call the weak topology on $\mathbb{R}^X$ generated by $\{\delta_a : a \in A\}$ the *topology of pointwise convergence on A*, where for each $a \in A$, $\delta_a : \mathbb{R}^X \to \mathbb{R}$ is defined by, $\delta_a(f) := f(a)$. We shall denote the topology of pointwise convergence on A by $\tau_p(A)$.

Corollary 3.12. *For each $n \in \mathbb{N}$, let $\varphi_n : A \to \mathbb{R}$ be a convex lower-semicontinuous function defined on a nonempty closed and convex subset A of a Hausdorff locally convex space $(X, +, \cdot, \tau)$, then for every sequence $\widetilde{x} := (x_n : n \in \mathbb{N})$ in A, either there exists a subsequence without any cluster points, or else, there exists a subsequence $\widetilde{x}|_J$ such that φ is a constant on $K(\widetilde{x}|_J)$, for each $\varphi \in \overline{\{\varphi_n : n \in \mathbb{N}\}}^{\tau_p(A)}$.*

Proof. Let $\widetilde{x} := (x_n : n \in \mathbb{N})$ be a sequence in A and suppose that every subsequence of $\widetilde{x}$ has a cluster point. Let $J_0 := \mathbb{N}$. We inductively construct a decreasing sequence of infinite subsets $(J_n : n \in \mathbb{N})$ of $\mathbb{N}$ such that φ_n is a constant on $K(\widetilde{x}|_{J_n})$, for each $n \in \mathbb{N}$.

We begin as follows. Since φ_1 is convex and lower-semicontinuous, there exists, by Lemma 3.11, an infinite subset J_1 of $\mathbb{N}$ such that φ_1 is constant on $K(\widetilde{x}|_{J_1})$.

Now, suppose that we have created infinite subsets $J_n \subseteq J_{n-1} \subseteq \cdots \subseteq J_1 \subseteq \mathbb{N}$ such that φ_i is constant on $K(\widetilde{x}|_{J_i})$ for all $1 \leq i \leq n$.

Then, for the $(n+1)^{\text{th}}$ step choose, using Lemma 3.11, an infinite subset J_{n+1} of J_n such that φ_{n+1} is constant on $K(\widetilde{x}|_{J_{n+1}})$.

Now, define $J := \{n_k : k \in \mathbb{N}\} \subseteq \mathbb{N}$ such that $n_k < n_{k+1}$ and $n_k \in J_k$ for all $k \in \mathbb{N}$. Consider the subsequence of $\widetilde{x}$ given by $\widetilde{x}|_J = (x_{n_k} : k \in \mathbb{N})$. Then, since $J_{n+1} \subseteq J_n$ for all $n \in \mathbb{N}$, we have that $n_k \in J_m$ for all $m \leq k$. Let $m \in \mathbb{N}$. Then,

$$\varnothing \neq K(\widetilde{x}|_J) = K(\{x_{n_k} : k \geq m\}) \subseteq K(\widetilde{x}|_{J_m}),$$

and so $1 \leq |\varphi_m(K(\widetilde{x}|_J))| \leq |\varphi_m(K(\widetilde{x}|_{J_m}))| \leq 1$. Since m was arbitrary, this gives that φ_m is constant on $K(\widetilde{x}|_J)$, for all $m \in \mathbb{N}$.

Now, let $\varphi \in \overline{\{\varphi_n : n \in \mathbb{N}\}}^{\tau_p(A)}$ and let $x, y \in K(\widetilde{x}|_J)$. Suppose, for a contradiction, that $\varphi(x) > \varphi(y)$. Then

$$N := \left\{F \in \mathbb{R}^X : F(x) > (1/2)[\varphi(x) + \varphi(y)]\right\} \cap \left\{F \in \mathbb{R}^X : F(y) < (1/2)[\varphi(x) + \varphi(y)]\right\}$$

is a $\tau_p(A)$-neighbourhood of φ. Since $\varphi \in \overline{\{\varphi_n : n \in \mathbb{N}\}}^{\tau_p(A)}$ there must exist $k \in \mathbb{N}$ such that $\varphi_k \in N$. However, this is impossible as $\varphi_k(x) = \varphi_k(y)$ for all $k \in \mathbb{N}$, and so φ is constant on $K(\widetilde{x}|_J)$. $\quad\square$

By applying Corollary 3.12 we obtain the following technical result that is needed (i.e., provides the required subsequence) in the proof of the general version of James' weak compactness theorem.

Corollary 3.13. *Let $\varphi : X \to \mathbb{R}$ be a τ-continuous convex function defined on a locally convex space $(X, +, \cdot, \tau)$. If τ' is a Hausdorff locally convex topology on X such that (i) $\tau' \subseteq \tau$ and (ii) φ is τ'-lower semicontinuous then, for every sequence $\widetilde{x} := (x_n : n \in \mathbb{N})$ in X, either there exists a subsequence without any τ'-cluster points, or else, there exists a subsequence $\widetilde{x}|_J$ of $\widetilde{x}$ such that φ is constant on, $y - aK_{\tau'}(\widetilde{x}|_J)$ for all $y \in \operatorname{span}\{x_n : n \in \mathbb{N}\}$ and all $a \in \mathbb{R}$.*

Proof. Let $\widetilde{x} := (x_n : n \in \mathbb{N})$ be a sequence in X and suppose that every subsequence of $\widetilde{x}$ has a τ'-cluster point. Observe that $Y := \mathrm{span}\{x_n : n \in \mathbb{N}\}$ is τ-separable, so let $\{y_n : n \in \mathbb{N}\}$ be a countable, τ-dense subset of Y. Moreover, let $\{q_n : n \in \mathbb{N}\}$ be an enumeration of $\mathbb{Q}\backslash\{0\}$. Now for all $m, n \in \mathbb{N}$, define $\varphi_n^m : X \to \mathbb{R}$ by

$$\varphi_n^m(x) = \varphi(y_n - q_m x) \quad \text{for all } x \in X.$$

Since $x \mapsto (y_n - q_m x)$ is a τ'-continuous affine function and φ is convex and τ'-lower semicontinuous, we have that φ_n^m is τ'-lower-semicontinuous and convex for all $m, n \in \mathbb{N}$. Then, by Corollary 3.12, [applied in $(X, +, \cdot, \tau')$] there exists a subsequence, $\widetilde{x}|_J$, of $\widetilde{x}$ such that ψ is constant on $K_{\tau'}(\widetilde{x}|_J)$, for each ψ in the $\tau_p(X)$-closure of $\{\varphi_n^m : m, n \in \mathbb{N}\}$.

Now observe that, for all $a \in \mathbb{R}$ and all $y \in Y$, the function $\varphi_y^a : X \to \mathbb{R}$ given by $\varphi_y^a(x) := \varphi(y - ax)$ is in the $\tau_p(X)$-closure of $\{\varphi_n^m : m, n \in \mathbb{N}\}$. To see this, fix $y \in Y$ and $a \in \mathbb{R}$. Let $F := \{x_1, x_2, \ldots, x_k\}$ be a finite subset of X and let $\varepsilon > 0$. For each $1 \le i \le k$ choose a τ-neighbourhood U_i of y and an open interval V_i of a such that $|\varphi(y' - q'x_i) - \varphi(y - ax_i)| < \varepsilon$ for all $y' \in U_i$ and all $q' \in V_i$. Note that this is possible because φ is τ-continuous and the mapping $(y', q') \mapsto (y' - q'x_i)$ is τ-continuous. Let $U := \bigcap_{1 \le i \le k} U_i$ and $V := \bigcap_{1 \le i \le k} V_i$ then $(y, a) \in U \times V$. Since $\{y_n : n \in \mathbb{N}\}$ is τ-dense in Y there exists an $n \in \mathbb{N}$ such that $y_n \in U$ and since $\{q_m : m \in \mathbb{N}\}$ is dense in $\mathbb{R}$ there exists an $m \in \mathbb{N}$ such that $q_m \in V$. Therefore,

$$|\varphi_n^m(x_i) - \varphi_y^a(x_i)| = |\varphi(y_n - q_m x_i) - \varphi(y - ax_i)| < \varepsilon \quad \text{for all } 1 \le i \le k.$$

This shows that, φ_y^a is in the $\tau_p(X)$-closure of $\{\varphi_n^m : m, n \in \mathbb{N}\}$. Therefore, $\varphi_y^a(K_{\tau'}(\widetilde{x}|_J)) = \varphi(y - aK_{\tau'}(\widetilde{x}|_J))$ is a singleton for all $y \in \mathrm{span}\{x_n : n \in \mathbb{N}\}$ and all $a \in \mathbb{R}$, as required. $\square$

The last result we need before we can prove the full version of James' theorem concerns the convergence of the subsequences that we constructed in **Part IV** of the proof of Theorem 3.9.

Proposition 3.14. *Let (X, τ) be a locally convex space and let $\widetilde{x} := (x_n : n \in \mathbb{N})$ be a sequence in a τ-compact convex subset K of X. If $\widetilde{y} := (y_n : n \in \mathbb{N})$ is any sequence such that $y_k \in \mathrm{co}\{x_n : n \ge k\}$ for all $k \in \mathbb{N}$, then $cl(\widetilde{y}) \subseteq K(\widetilde{x})$.*

Proof. It is sufficient to show that for any open convex neighbourhood, W of 0, $cl(\widetilde{y}) \subseteq K(\widetilde{x}) + \overline{W}$. To this end, let W be an open convex neighbourhood of 0. Then note that for k sufficiently large,

$$\{x_n : n \ge k\} \subseteq cl(\widetilde{x}) + W.$$

Indeed if this is not the case, then we could construct a subsequence $(x_{n_k} : k \in \mathbb{N})$ of $(x_n : n \in \mathbb{N})$ such that $x_{n_k} \notin cl(\widetilde{x}) + W$ for all $k \in \mathbb{N}$, However, since $X \backslash [cl(\widetilde{x}) + W]$ is a closed set containing $\{x_{n_k} : k \in \mathbb{N}\}$ we have that $\overline{\{x_{n_k} : k \in \mathbb{N}\}} \cap [cl(\widetilde{x}) + W] = \varnothing$, but this is impossible since $\overline{\{x_{n_k} : k \in \mathbb{N}\}} \cap cl(\widetilde{x}) \ne \varnothing$. Thus, we have a contradiction. Therefore, if $y \in cl(\widetilde{y})$, then for k sufficiently large, we have that

$$y \in \overline{\{y_n : n \ge k\}} \subseteq \overline{\mathrm{co}\{x_n : n \ge k\}} \subseteq K(\widetilde{x}) + \overline{W} \quad \text{since, } K(\widetilde{x}) + \overline{W} \text{ is closed and convex.}$$

Hence, $cl(\widetilde{y}) \subseteq K(\widetilde{x}) + \overline{W}$ as required. $\square$

Theorem 3.15 (James' Theorem: version 3, [21]). *Let C be a closed and bounded convex subset of a Banach space $(X, \|\cdot\|)$. If every $x^* \in X^*$ attains its supremum over C, then C is weakly compact.*

Proof. To show that C is weakly compact, it suffices to show that $K := \overline{\widehat{C}}^{w^*} \subseteq \widehat{X}$ (see Remark 2.26). Suppose, for a contradiction, that this is not the case. Then there exists $F \in K\backslash\widehat{X}$. Since $\widehat{X}$ is a closed subspace of X^{**}, this means there must exist $0 < \varepsilon < \mathrm{dist}(F, \widehat{X})$. Let $(\beta_n : n \in \mathbb{N})$ be a sequence of strictly positive numbers such that $\lim_{n \to \infty} \frac{1}{\beta_n} \sum_{i=n+1}^{\infty} \beta_i = 0$.

Part I: We inductively create the two sequences $(\widehat{x}_n : n \in \mathbb{N})$ in $\widehat{C}$, and $(f_n : n \in \mathbb{N})$ in S_{X^*}, which satisfy the statements (A_n) and (B_n), exactly as in Part I of the proof of Theorem 3.9.

Part II: Define $p : X^* \to \mathbb{R}$ to be $p(x^*) = \sup\{x^*(c) : c \in C\}$ for all $x^* \in X^*$. Then p is norm-continuous, weak*-lower-semicontinuous and convex. Just as in Part II of the proof of Theorem 3.9, passing

22

to a subsequence does not disturb the statements (A_n) and (B_n).

So, by passing to a subsequence and relabelling if necessary, we may assume, by Corollary 3.13, that p is constant on $f - aK_{w^*}(f_n : n \in \mathbb{N})$ for all $f \in \mathrm{span}\{f_n : n \in \mathbb{N}\}$ and all $a \in \mathbb{R}$. Since $(f_n : n \in \mathbb{N})$ is a sequence in B_{X^*} (which is weak*-compact), it must a have a weak*-cluster point, call it f_∞.

Part III: This step is exactly the same as Part III of the proof of Theorem 3.9 - we deduce that $\widehat{x}_k(f_\infty) = 0$ for all $k \in \mathbb{N}$ and that $\widehat{x}_n(f_k) > \varepsilon/2$ for all $n > k$; from which we conclude that $\widehat{x}_n(f_k - f_\infty) > \varepsilon/2$ for all $n > k$.

Part IV: As in the proof of Theorem 3.9, we use Lemma 3.3 to construct a sequence $(g_n : n \in \mathbb{N})$ such that for all $n \in \mathbb{N}$:

(i) $g_n \in \mathrm{co}\{f_k : k \geq n\}$ and

(ii) $p(\sum_{i=1}^{n} \beta_i(g_i - f_\infty)) + \beta_{n+1}\varepsilon/4 < p(\sum_{i=1}^{n+1} \beta_i(g_i - f_\infty))$.

Part V: Since $(g_n : n \in \mathbb{N})$ is a sequence in B_{X^*} (which is weak*-compact), it must a have a weak*-cluster point, call it g_∞. Then, by Proposition 3.14, we have that $g_\infty \in K_{w^*}(f_n : n \in \mathbb{N})$. While it may no longer be the case that $f_\infty = g_\infty$ as in Theorem 3.9, we do have that, for all $n \in \mathbb{N}$,

$$p(\textstyle\sum_{i=1}^{n}\beta_i(g_i - g_\infty)) = p(\textstyle\sum_{i=1}^{n}\beta_i g_i - \textstyle\sum_{i=1}^{n}\beta_i \cdot g_\infty) = p(\textstyle\sum_{i=1}^{n}\beta_i g_i - \textstyle\sum_{i=1}^{n}\beta_i \cdot f_\infty) = p(\textstyle\sum_{i=1}^{n}\beta_i(g_i - f_\infty)) \qquad (**)$$

since $g_\infty \in K_{w^*}(f_n : n \in \mathbb{N})$ and for all $f \in \mathrm{span}\{f_n : n \in \mathbb{N}\}$ and all $a \in \mathbb{R}$, the set $p(f - aK_{w^*}(f_n : n \in \mathbb{N}))$ is a singleton. As in Part V of the proof of Theorem 3.9, we set $g := \sum_{i=1}^{\infty} \beta_i(g_i - g_\infty)$ and deduce that $g \in X^*$.

Part VI: This final step is almost the same as Part VI of the proof of Theorem 3.9, with two small changes that we note here. We may replace f_∞ with g_∞ throughout the inequalities, not because $f_\infty = g_\infty$ but because of statement $(**)$ above. Lastly, the final contradiction is not because $\lim_{n\to\infty} g_n(c) = g_\infty(c)$ necessarily, but because $\liminf_{n\to\infty} g_n(c) \leq g_\infty(c)$. This still gives a contradiction. $\square$

3.4 James' Theorem: applications

Theorem 3.16 ([20]). *Let $(X, \|\cdot\|)$ be a Banach space. Then X is reflexive if, and only if, every continuous linear functional x^* on X attains its norm (i.e., there exists an $x \in B_X$ such that $\|x^*\| = x^*(x)$).*

Proof. By Theorem 2.26, X is reflexive if, and only if, B_X is weakly compact. So the result now follows from Theorem 3.15 once one remembers that every continuous linear functional on X is continuous with respect to the weak topology on X. $\square$

Note: if X is reflexive then one can use the Hahn-Banach Theorem to directly show that every continuous linear functional on X attains its norm. Indeed, suppose that x^* is a nonzero continuous linear functional on X. Then by Corollary 2.13 there exists an $x^{**} \in S_{X^{**}}$ such that $x^{**}(x^*) = \|x^*\|$. However, since X is reflexive, $x^{**} = \widehat{x}$ for some $x \in S_X$. Hence, $\|x^*\| = x^{**}(x^*) = \widehat{x}(x^*) = x^*(x)$. This shows that x^* attains its norm.

We now recall a geometric concept in Banach space theory. We say that a Banach space, $(X, \|\cdot\|)$, is *uniformly convex* if, for any $\varepsilon > 0$, there exists $\delta_\varepsilon > 0$ with the following property: if $x, y \in B_X$ and $\|x + y\| > 2 - \delta_\varepsilon$, then $\|x - y\| < \varepsilon$.

Theorem 3.17 ([41]). *Let $(X, \|\cdot\|)$ be a Banach space. If $(X, \|\cdot\|)$ is uniformly convex, then $(X, \|\cdot\|)$ is reflexive.*

Proof. Let $x^* \in S_{X^*}$ and define $(x_n : n \in \mathbb{N})$ in B_X so that $x^*(x_n) > 1 - \frac{1}{n}$. Let $\varepsilon > 0$ and choose $\delta_\varepsilon > 0$ such that if $x, y \in B_X$ and $\|x + y\| > 2 - \delta_\varepsilon$, then $\|x - y\| < \varepsilon$ Then, for $n, m \in \mathbb{N}$ greater than $N_0 := 2/\delta_\varepsilon$, we have that $2 \geq \|x_n + x_m\| \geq x^*(x_n + x_m) > 2 - \delta_\varepsilon$. By the uniform convexity of X, this gives that for $n, m > N_0$, we have $\|x_n - x_m\| \leq \varepsilon$. So, $(x_n : n \in \mathbb{N})$ is a Cauchy sequence in X. Therefore, $(x_n : n \in \mathbb{N})$ is convergent to some $x \in B_X$. It is clear that for this x, $x^*(x) = 1 = \|x^*\|$. Since x^* was arbitrary in S_{X^*}, every x^* in X^* attains its norm, and so, by Theorem 3.16, X is reflexive. $\square$

Another interesting application of Theorem 3.15 is the Krein-Phillips theorem.

Corollary 3.18 (Krein-Phillips Theorem, [30, 43])**.** *Let C be a weakly compact subset of a Banach space $(X, \|\cdot\|)$. Then $\overline{co}(C)$ is also weakly compact.*

Proof. Let $K := \overline{co}(C)$. Since C is weakly compact, every $x^* \in X^*$ must attain its supremum over C i.e. for every $x^* \in X^*$, there exists $c \in C \subseteq K$ such that $x^*(c) = \sup_{x \in C} x^*(x)$. However, for every $x^* \in X^*$, it is a routine observation that

$$\sup_{x \in C} x^*(x) = \sup_{x \in co(C)} x^*(x) = \sup_{x \in K} x^*(x).$$

Thus, every $x^* \in X^*$ attains its supremum over K too. Therefore, by James' Theorem (Theorem 3.15), K is weakly compact. $\square$

Using Theorem 3.4 we can prove some well-known results of S. Simons, see [48]. For a detailed survey of Simons' results and applications thereof, see [5].

Theorem 3.19 (Simons, [48])**.** *Let K be a weak*-compact, convex subset of the dual of a Banach space $(X, \|\cdot\|)$, let B be a boundary for K, and let $f_n : K \to \mathbb{R}$ be weak*-lower-semicontinuous, convex functions, for all $n \in \mathbb{N}$. If $(f_n : n \in \mathbb{N})$ is equicontinuous with respect to the norm, and $\limsup_{n \to \infty} f_n(b^*) \leq 0$ for all $b^* \in B$, then $\limsup_{n \to \infty} f_n(x^*) \leq 0$ for all $x^* \in K$.*

Proof. Let $\varepsilon > 0$. For each $n \in \mathbb{N}$, define:

$$C_n := \bigcap_{k \geq n} \{y^* \in K : f_k(y^*) \leq \varepsilon/2\}.$$

Let $k \in \mathbb{N}$. Since $f_k : K \to \mathbb{R}$ is weak*-lower-semicontinuous and convex, the set $\{y^* \in K : f_k(y^*) \leq \varepsilon/2\}$ is weak*-closed and convex. It follows that for all $n \in \mathbb{N}$, C_n is the intersection of weak*-closed and convex sets, and so is weak*-closed and convex itself. Then, since $C_n \subseteq K$ for all $n \in \mathbb{N}$, we have that C_n is weak*-compact and convex for all $n \in \mathbb{N}$. Moreover, if $b^* \in B$, then $\limsup_{n \to \infty} f_n(b^*) \leq 0$ and so $b^* \in C_N$ for some $N \in \mathbb{N}$. Hence, $(C_n : n \in \mathbb{N})$ is a countable cover of B by weak*-compact, convex subsets of K.

Therefore, since B is a boundary for K, by Theorem 3.4 we have that $co[\bigcup_{n \in \mathbb{N}} C_n] = \bigcup_{n \in \mathbb{N}} C_n$ (since $C_n \subseteq C_{n+1}$ for all $n \in \mathbb{N}$) is norm-dense in K. Let $x^* \in K$. Since $(f_n : n \in \mathbb{N})$ is equicontinuous with respect to the norm, there exists a $\delta > 0$ such that $f_n(x^*) < f_n(y^*) + \varepsilon/2$ for all $n \in \mathbb{N}$ and all $y^* \in B(x^*, \delta)$.

However since $\bigcup_{n \in \mathbb{N}} C_n$ is norm-dense in K, there exists $N \in \mathbb{N}$ such that $B(x^*, \delta) \cap C_N \neq \varnothing$. Therefore, $f_n(x^*) < \varepsilon$ for all $n > N$ and so $\limsup_{n \to \infty} f_n(x^*) \leq \varepsilon$. Since $\varepsilon > 0$ and $x^* \in K$ were arbitrary, we have that $\limsup_{n \to \infty} f_n(x^*) \leq 0$ for all $x^* \in K$ as claimed. $\square$

Corollary 3.20. *Let K be a weak*-compact, convex subset of the dual of a Banach space $(X, \|\cdot\|)$ and let B be a boundary for K. Let $(x_n : n \in \mathbb{N})$ be a bounded sequence in X and let $x \in X$. If $\lim_{n \to \infty} b^*(x_n) = b^*(x)$ for all $b^* \in B$, then $\lim_{n \to \infty} x^*(x_n) = x^*(x)$ for all $x^* \in K$.*

Proof. For all $n \in \mathbb{N}$, define $f_n : K \to \mathbb{R}$ to be given by

$$f_n(x^*) := |x^*(x_n) - x^*(x)| = |\widehat{(x_n - x)}(x^*)| \quad \text{for all } x^* \in K.$$

Then $f_n : K \to \mathbb{R}$ is a weak*-lower-semicontinuous and convex function for all $n \in \mathbb{N}$, as $x^* \mapsto \widehat{(x_n - x)}(x^*)$ is weak* continuous and linear (into $\mathbb{R}$) and $r \mapsto |r|$ is continuous and convex. Furthermore, $(f_n : n \in \mathbb{N})$ is equicontinuous with respect to the norm. Finally, $\limsup_{n \to \infty} f_n(b^*) \leq 0$ for all $b^* \in B$ and so, by Theorem 3.19, $\limsup_{n \to \infty} f_n(x^*) \leq 0$ for all $x^* \in K$. From this it is clear that $\lim_{n \to \infty} x^*(x_n) = x^*(x)$ for all $x^* \in K$. $\square$

Sometimes called the Rainwater-Simons Theorem, Corollary 3.20 is due to S. Simons (although he proved it differently). It generalises a famous result of J. Rainwater, originally from [45].

Corollary 3.21 (Simons). *Let K be a weak*-compact convex subset of the dual of a Banach space $(X, \|\cdot\|)$, let B be a boundary for K, and let $(x_n : n \in \mathbb{N})$ be a bounded sequence in X. Then*

$$\sup_{b^* \in B} \left\{ \limsup_{n \to \infty} \widehat{x}_n(b^*) \right\} = \sup_{x^* \in K} \left\{ \limsup_{n \to \infty} \widehat{x}_n(x^*) \right\}.$$

Proof. Since $B \subseteq K$, clearly

$$\sup_{b^* \in B} \left\{ \limsup_{n \to \infty} \widehat{x}_n(b^*) \right\} \leq \sup_{x^* \in K} \left\{ \limsup_{n \to \infty} \widehat{x}_n(x^*) \right\}.$$

So it only remains to show that

$$\sup_{b^* \in B} \left\{ \limsup_{n \to \infty} \widehat{x}_n(b^*) \right\} \geq \sup_{x^* \in K} \left\{ \limsup_{n \to \infty} \widehat{x}_n(x^*) \right\}.$$

To this end, let

$$r := \sup_{b^* \in B} \left\{ \limsup_{n \to \infty} \widehat{x}_n(b^*) \right\},$$

and for each $n \in \mathbb{N}$, let $f_n : K \to \mathbb{R}$ be defined by $f_n(x^*) := \sup\{\widehat{x}_k(x^*) : k \geq n\} - r$ for all $x^* \in K$. Then, for all $n \in \mathbb{N}$, f_n is weak*-lower-semicontinuous and convex, as the pointwise supremum of a family of convex functions is again convex and the pointwise supremum of a family of lower semi-continuous functions is again lower semi-continuous. Furthermore, $(f_n : n \in \mathbb{N})$ is equicontinuous with respect to the norm and moreover, $\lim_{n \to \infty} f_n(b^*) \leq 0$ for all $b^* \in B$. Therefore, by Theorem 3.19, $\lim_{n \to \infty} f_n(x^*) \leq 0$ for all $x^* \in K$. From this, the result is immediate. $\square$

In the next part of this subsection we will show that in order to deduce that a closed and bounded convex subset C of Banach space $(X, \|\cdot\|)$ is weakly compact it is not necessary to show that **all** the elements of X^* attain their maximum value of C, but only a "large" subset of X^*. To achieve this goal we need some more definitions.

Let K be a subset of the dual of a normed linear space $(X, \|\cdot\|)$. A point $x^* \in K$ is called a *weak* exposed point of K* if there exists a $x \in X \setminus \{0\}$ such that $\widehat{x}(x^*) \geq \sup_{y^* \in K} \widehat{x}(y^*)$. There are some simple, but useful, facts that we can easily deduce about weak* exposed points.

Firstly, (i) if x^* is a weak* exposed point of K then λx^* is a weak* exposed point of λK for any $\lambda \in \mathbb{R} \setminus \{0\}$; (ii) if x^* is a weak* exposed point of K then $x^* + y^*$ is a weak* exposed point of $K + y^*$ for any $y^* \in X^*$; (iii) if $x^* \in A \subseteq K$ is a weak* exposed point of K then x^* is a weak* exposed point of A.

The next result shows that weak* exposed points are directly related to weak compactness.

Proposition 3.22. *Let K be a closed and convex subset of the dual of a Banach space $(X, \|\cdot\|)$. If $0 \in \mathrm{int}(K)$ and every point of $\mathrm{Bd}(K)$ is a weak* exposed point of K, then $K_\circ$ is a weakly compact subset of X.*

Proof. We shall appeal directly to Theorem 3.15. To this end, let $x^* \in X^* \setminus \{0\}$. We consider two cases.
Case (I) Suppose that for every $0 < \lambda$, $\lambda x^* \in K$. Let $k \in K_\circ$ and let $0 < \lambda$. Then

$$\lambda x^*(k) = (\lambda x^*)(k) \leq 1 \quad \text{since, } \lambda x^* \in K \subseteq (K_\circ)^\circ.$$

Therefore, $x^*(k) \leq \lambda^{-1}$. Since $0 < \lambda$ was arbitrary, $x^*(k) \leq 0$ and so x^* attains its maximum value over $K_\circ$ at $0 \in K_\circ$.

Case(II) Suppose that for some $0 < \lambda$, $(\lambda x^*) \notin K$. Let $\Lambda := \{r \in [0, \infty) : r x^* \in K\}$. Then Λ is a closed and bounded interval of $[0, \infty)$ since K is closed and convex and $\lambda \notin \Lambda$. Let $\lambda_0 := \max_{r \in \Lambda} r$. Then $\lambda_0 x^* \in \mathrm{Bd}(K)$. Hence there exists a $x \in X \setminus \{0\}$ such that

$$\lambda_0 \widehat{x}(x^*) = \widehat{x}(\lambda_0 x^*) = \sup_{y^* \in K} \widehat{x}(y^*) > 0 \quad \text{since, } 0 \in \mathrm{int}(K).$$

By replacing x by μx for some $\mu > 0$ and relabelling if necessary, we can assume that

$$1 = \widehat{x}(\lambda_0 x^*) = \sup_{y^* \in K} \widehat{x}(y^*) = \sup_{y^* \in K} y^*(x).$$

Therefore $x \in K_\circ$. On the other hand, since $\lambda_0 x^* \in K \subseteq (K_\circ)^\circ$ we have that

$$(\lambda_0 x^*)(k) \leq 1 = (\lambda_0 x^*)(x) \quad \text{for all } k \in K_\circ.$$

Therefore $\lambda_0 x^*$ attains its maximum value over $K_\circ$ at x, and hence so does x^*. Therefore, by Theorem 3.15, $K_\circ$ is weakly compact. $\quad\square$

Theorem 3.23 ([24]). *Let $(X, \|\cdot\|)$ be a Banach space. If there exists a weak* open subset U of X^* such that $\varnothing \neq S_{X^*} \cap U$ and every member of $S_{X^*} \cap U$ attains its norm on X, then X is reflexive.*

Proof. Suppose that $\{x_1, x_2, \ldots, x_n\} \subseteq X$, $\varepsilon > 0$ and $x_0^* \in X^*$ are chosen so that if

$$W' := \bigcap_{k=1}^n \{x^* \in X^* : |x^*(x_k) - x_0^*(x_k)| < \varepsilon\} \quad \text{and} \quad W := \bigcap_{k=1}^n \{x^* \in X^* : |x^*(x_k) - x_0^*(x_k)| \leq \varepsilon\}$$

then $\varnothing \neq S_{X^*} \cap W'$ and every member of $S_{X^*} \cap W$ attains its norm on X (i.e., every point of $S_{X^*} \cap W$ is a weak* exposed point of B_{X^*} - just consider the $x \in B_X$ such that $\widehat{x}(x^*) = \|x^*\| = 1$). Let $K' := W \cap B_{X^*}$. Then K' is closed and bounded and convex. Furthermore, $\text{int}(K') \neq \varnothing$. Let us now recall some basic facts from general topology. If A and B are closed subsets of a topological space (T, τ) then

$$\text{Bd}(A \cap B) \subseteq (\text{Bd}(A) \cap B) \cup (\text{Bd}(B) \cap A) \subseteq \text{Bd}(A) \cup \text{Bd}(B).$$

So $\text{Bd}(K') \subseteq [\text{Bd}(W) \cap B_{X^*}] \cup [S_{X^*} \cap W]$. We claim that every point of $\text{Bd}(K')$ is a weak* exposed point. To see this, suppose that $x^* \in \text{Bd}(W) \cap B_{X^*} \subseteq \text{Bd}(W)$. Then clearly, x^* is a weak* exposed point of the set W (exposed by $\widehat{x_k}$ for some $1 \leq k \leq n$). Then by property (iii) above, x^* is a weak* exposed point of $W \cap B_{X^*}$. If $x^* \in S_{X^*} \cap W$ then by the way W was chosen, x^* is a weak* exposed point of B_{X^*} and hence by property (iii) above, also a weak* exposed point of $B_{X^*} \cap W$. Choose $x^* \in \text{int}(K')$ and let $K := K' - x^*$. Then $0 \in \text{int}(K)$ and by property (ii) above, each point of $\text{Bd}(K)$ is a weak* exposed point. Thus, by Proposition 3.22, $K_\circ$ is weakly compact. Now since K is bounded, $0 \in \text{int}(K_\circ)$. Hence X is reflexive. $\quad\square$

The proof of the next theorem can be found in [36], (see also [39, 40]). Its conclusion is phrased in terms of the notion of "relative weak compactness". So in order to avoid any possible confusion we will give the precise definition of this notion here. A subset C of a Banach space $(X, \|\cdot\|)$ is said to be *relatively weakly compact* if its closure, with respect to the weak topology on X, is weakly compact.

Theorem 3.24 ([47]). *Let $(X, \|\cdot\|)$ be a Banach space and let $f : X \to \mathbb{R} \cup \{\infty\}$ be a proper function on X. If $f - x^*$ attains minimum for every $x^* \in X^*$ then for each $a \in \mathbb{R}$, $S(a) := \{(y, s) \in X \times \mathbb{R} : f(y) \leq s \leq a\}$ is relatively weakly compact.*

Proof. In this proof we will identify the dual of $X \times \mathbb{R}$ with $X^* \times \mathbb{R}$. We will also consider $X \times \mathbb{R}$ endowed with the norm $\|(x, r)\|_1 := \|x\| + |r|$ and note that with this norm, $(X \times \mathbb{R}, \|\cdot\|_1)$ is a Banach space. We shall apply James' theorem, (Theorem 3.15), in $X \times \mathbb{R}$. Let $H := \{(x, r) \in X \times \mathbb{R} : r = 0\}$ and define $T : (X \times \mathbb{R}) \setminus H \to (X \times \mathbb{R}) \setminus H$ by, $T(x, r) := r^{-1}(x, -1)$. Then T is a bijection. In fact, T is a homeomorphism when $(X \times \mathbb{R}) \setminus H$ is considered with the relative weak topology. Note that since f is bounded below we may assume, after possibly translating, that $1 = \inf_{x \in X} f(x)$. Our proof relies upon the *Fenchel conjugate* (see, [46, page 102] or [2, page 49]) $f^* : X^* \to \mathbb{R}$ *of f*, which is defined by,

$$f^*(x^*) := \sup_{x \in X}[x^*(x) - f(x)] = -\inf_{x \in X}[f(x) - x^*(x)] = -\min_{x \in X}[f(x) - x^*(x)] = \max_{x \in X}[x^*(x) - f(x)].$$

It is routine to check that f^* is convex on X^* (see, Proposition 5.2, part (i)). We claim that $\overline{\text{co}}[T(\text{epi}(f)) \cup \{(0,0)\}]$ is weakly compact. To show this, it is sufficient, because of James' theorem, to show that every non-zero continuous linear functional attains its maximum value over $T(\text{epi}(f)) \cup \{(0,0)\}$. To this end, let $(x^*, r) \in (X^* \times \mathbb{R}) \setminus \{(0,0)\}$. We consider two cases.

Case (I) Suppose that for every $0 < \lambda$, $f^*(\lambda x^*) \leq \lambda r$. Then $x^*(x) - \lambda^{-1} f(x) \leq r$ for all $x \in X$ and all $0 < \lambda$. Let $(y, s) \in \mathrm{epi}(f)$ and let $0 < \lambda$. Then,

$$(x^*, r)(T(y, s)) = s^{-1}(x^*(y) - r) \leq s^{-1}(x^*(y) - [x^*(y) - \lambda^{-1} f(y)]) = s^{-1} f(y) \lambda^{-1} \leq \lambda^{-1}$$

since $0 < 1 \leq f(y) \leq s$. As $0 < \lambda$ was arbitrary, $(x^*, r)(T(y, s)) \leq 0 = (x^*, r)(0, 0)$. Thus, (x^*, r) attains its maximum value over $T(\mathrm{epi}(f)) \cup \{(0, 0)\}$ at $(0, 0)$.

Case(II) Suppose that for some $0 < \lambda$, $\lambda r < f^*(\lambda x^*)$. Then, since the mapping, $\lambda' \mapsto f^*(\lambda' x^*)$, is real-valued and convex, it is continuous. Furthermore, it follows, from the intermediate value theorem applied to the function $g : [0, \lambda] \to \mathbb{R}$, defined by,

$$g(\lambda') := f^*(\lambda' x^*) - \lambda' r \quad \text{for all } \lambda' \in [0, \lambda],$$

that there exists a $0 < \mu < \lambda$ such that $g(\mu) = 0$, i.e., $f^*(\mu x^*) = \mu r$, since

$$g(0) = f^*(0x^*) = -\inf_{x \in X} f(x) = -1 < 0 < g(\lambda).$$

Thus, $\mu(x^*, r) = (\mu x^*, f^*(\mu x^*))$. Choose $z \in X$ such that $f^*(\mu x^*) = \mu x^*(z) - f(z)$. We claim that (x^*, r) attains its maximum value over $T(\mathrm{epi}(f)) \cup \{(0, 0)\}$ at $T(z, f(z)) = f(z)^{-1}(z, -1)$. Now,

$$
\begin{aligned}
(x^*, r)(T(z, f(z))) &= f(z)^{-1}(x^*(z) - r) = f(z)^{-1}(x^*(z) - [\mu^{-1} f^*(\mu x^*)]) \\
&= f(z)^{-1}(x^*(z) - [x^*(z) - \mu^{-1} f(z)]) = \mu^{-1} > 0.
\end{aligned}
$$

On the other hand, if $(y, s) \in \mathrm{epi}(f)$ then

$$
\begin{aligned}
(x^*, r)(T(y, s)) &= s^{-1}(x^*(y) - r) = s^{-1}(x^*(y) - [\mu^{-1} f^*(\mu x^*)]) \\
&\leq s^{-1}(x^*(y) - [x^*(y) - \mu^{-1} f(y)]) = s^{-1} f(y) \mu^{-1} \leq \mu^{-1} = (x^*, r)(T(z, f(z)))
\end{aligned}
$$

since $0 < 1 \leq f(y) \leq s$. Note also that $(x^*, r)(0, 0) = 0 < \mu^{-1} = (x^*, r)(T(z, f(z)))$. Therefore, by James' Theorem 3.15, $\overline{\mathrm{co}}[T(\mathrm{epi}(f)) \cup \{(0, 0)\}]$ is weakly compact.

Let $1 \leq a$, then $T(S(a)) \subseteq \overline{\mathrm{co}}[T(\mathrm{epi}(f)) \cup \{(0, 0)\}] \cap \{(x, r) \in X \times \mathbb{R} : r \leq -a^{-1}\}$; which is weakly compact and disjoint from H. Therefore,

$$S(a) \subseteq T^{-1}(\overline{\mathrm{co}}[T(\mathrm{epi}(f)) \cup \{(0, 0)\}] \cap \{(x, r) \in X \times \mathbb{R} : r \leq -a^{-1}\});$$

which completes the proof. $\square$

For each $a \in \mathbb{R}$, let $L(a) := \{x \in X : f(x) \leq a\}$. It follows from Theorem 3.24 that if $(X, \|\cdot\|)$ is a Banach space, $f : X \to \mathbb{R} \cup \{\infty\}$ is a proper function on X and $f - x^*$ attains minimum for every $x^* \in X^*$ then, for each $a \in \mathbb{R}$, $L(a)$ is relatively weakly compact, since $L(a) = \pi(S(a))$, where $\pi : X \times \mathbb{R} \to X$ is defined by, $\pi(x, r) := x$ for all $(x, r) \in X \times \mathbb{R}$ and is weak-to-weak continuous, (see Proposition 2.10).

An interesting corollary of this result is the following.

Corollary 3.25 ([47])**.** *Let $\varphi : U \to \mathbb{R}$ be a continuous convex function defined on a nonempty open convex subset U of a Banach space $(X, \|\cdot\|)$. If $\varphi - x^*$ attains minimum for every $x^* \in X^*$ then X is reflexive.*

Proof. For each $n \in \mathbb{N}$, let $F_n := \{x \in U : \varphi(x) \leq n\}$. Then each set F_n is closed and $U = \bigcup_{n \in \mathbb{N}} F_n$. Since X is a Banach space, U is of the second Baire category. Thus, there exists an $n_0 \in \mathbb{N}$ such that $\mathrm{int}(F_{n_0}) \neq \varnothing$. In particular, there exists an $x_0 \in F_{n_0}$ and a $\delta_0 > 0$ such that $B[x_0, \delta_0] \subseteq F_{n_0}$. Therefore, by Theorem 3.24, $B[x_0, \delta_0] = x_0 + \delta_0 B_X$ is compact with respect to the weak topology, and hence so is B_X. The result now follows from Theorem 2.26. $\square$

Lemma 3.26 ([35])**.** *Let $(Y, \|\cdot\|)$ be a Banach space and C be a nonempty bounded subset of $Y \times \mathbb{R}$, endowed with the norm $\|(y, r)\|_1 := \|y\| + |r|$. If for every $x^* \in Y^*$, $\max\{(x^*, -1)(y, s) : (y, s) \in C\}$ exists, then C is relatively weakly compact.*

Proof. Let $\pi : Y \times \mathbb{R} \to Y$ be defined by $\pi(y,r) := y$, $A := \pi(C)$ and $f : Y \to \mathbb{R} \cup \{\infty\}$ be defined by,

$$f(y) := \begin{cases} \inf\{s \in \mathbb{R} : (y,s) \in C\} & \text{if } y \in A \\ \infty & \text{if } y \notin A. \end{cases}$$

Then f is a proper function on Y and $x^* - f$ attains it maximum for every $x^* \in Y^*$. To confirm this second assertion, consider the following. Let $x^* \in Y^*$, then

$$\sup_{y \in Y}(x^* - f)(y) \;=\; \sup_{y \in A}(x^* - f)(y) = \sup_{y \in A}\left(x^*(y) - \inf_{(y,s) \in C} s\right) = \sup_{y \in A}\left(x^*(y) + \sup_{(y,s) \in C}(-s)\right)$$

$$= \sup_{y \in A}\left(\sup_{(y,s) \in C}(x^*(y) - s)\right) = \sup_{(y,s) \in C}(x^*(y) - s) = \sup_{(y,s) \in C}(x^*, -1)(y,s). \quad (\dagger)$$

Now, by assumption, there exists an element $(y_0, s_0) \in C$ such that $(x^*, -1)(y_0, s_0) = \sup_{(y,s) \in C}(x^*, -1)(y,s)$.

Thus,

$$\sup_{y \in Y}(x^* - f)(y) \geq (x^* - f)(y_0) \;=\; x^*(y_0) - f(y_0)$$

$$\geq \quad x^*(y_0) - s_0 \qquad \text{since } f(y_0) \leq s_0$$

$$= \quad \sup_{(y,s) \in C}(x^*, -1)(y,s)$$

$$= \quad \sup_{y \in Y}(x^* - f)(y) \qquad \text{by } (\dagger)$$

From this it follows that

$$(x^* - f)(y_0) = \sup_{y \in Y}(x^* - f)(y) \quad \text{and} \quad s_0 = f(y_0).$$

Therefore, by Theorem 3.24, for each $a \in \mathbb{R}$, $S(a) := \{(y,s) \in Y \times \mathbb{R} : f(y) \leq s \leq a\}$ is relatively weakly compact. Since C is bounded there exists an $a \in \mathbb{R}$ such that $C \subseteq S(a)$. $\quad\square$

Theorem 3.27 ([35])**.** *Let $(X, \|\cdot\|)$ be a Banach space and let A and B be nonempty closed and bounded convex sets with $0 < dist(A, B)$. If every $x^* \in X^*$ with $\sup\{(x^*(b) : b \in B\} < \inf\{x^*(a) : a \in A\}$ attains its infimum over A and its supremum over B, then both A and B are weakly compact.*

Proof. To show that both A and B are weakly compact it is sufficient (and necessary) to show that $B - A$ is weakly compact. This will be our approach. From the hypotheses it follows that if $C := \overline{B - A}$, then C is a bounded nonempty closed and convex subset of X with $0 \notin C$. Furthermore, it follows that each $x^* \in X^*$ with $\sup\{x^*(c) : c \in C\} < 0$ attains it supremum over C. Choose $y^* \in X^*$ such that $\sup\{y^*(c) : c \in C\} < 0$. Note that such a functional exists by the Hahn-Banach theorem. Let $Y := \ker(y^*)$ and choose $x_0 \in C$. Define $S : Y \times \mathbb{R} \to X$ by, $S(y,r) := y + rx_0$ and let us consider $Y \times \mathbb{R}$ endowed with the norm $\|(y,r)\|_1 := \|y\| + |r|$. Then S is an isomorphism and there exists an $0 < \varepsilon$ and an $\varepsilon < a$ such that

$$S^{-1}(C) \subseteq \{(y,r) \in Y \times \mathbb{R} : \varepsilon \leq r \leq a\}.$$

Moreover, each $(x^*, r) \in (Y \times \mathbb{R})^*$ with $\sup\{(x^*, r)(y,s) : (y,s) \in S^{-1}(C)\} < 0$ attains its supremum over $S^{-1}(C)$. Let $\pi : Y \times \mathbb{R} \to Y$ be defined by $\pi(y,r) := y$, $A := \pi(S^{-1}(C))$ and $f : Y \to \mathbb{R} \cup \{\infty\}$ be defined by,

$$f(y) := \begin{cases} \inf\{s \in \mathbb{R} : (y,s) \in S^{-1}(C)\} & \text{if } y \in A \\ \infty & \text{if } y \notin A. \end{cases}$$

Next, we define $T : Y \times (\mathbb{R} \setminus \{0\}) \to Y \times (\mathbb{R} \setminus \{0\})$ by $T(y,s) := s^{-1}(y, -1)$. Then T is a bijection. In fact, T is a homeomorphism when $Y \times (\mathbb{R} \setminus \{0\})$ is considered with the relative weak topology. Let $f^* : Y^* \to \mathbb{R}$ be defined by,

$$f^*(x^*) := \sup_{y \in Y}[x^*(y) - f(y)] = \sup\{(x^*, -1)(y,s) : (y,s) \in S^{-1}(C)\}, \quad \text{as in the proof of Lemma 3.26.}$$

28

It is routine to check that f^* is real-valued and convex on Y^* (see, Proposition 5.2, part (i)). To show that C is weakly compact it is sufficient to show $S^{-1}(C)$ is weakly compact, and to show that $S^{-1}(C)$ is weakly compact, it is sufficient to show that

$$T(S^{-1}(C)) \subseteq \overline{T(S^{-1}(C))}^{\text{weak}} \subseteq \{(y, r) \in Y \times \mathbb{R} : -1/\varepsilon \le r \le -1/a\}$$

is a relatively weakly compact subset of $Y \times \mathbb{R}$. To achieve this we appeal to Lemma 3.26. First note that $T(S^{-1}(C))$ is a nonempty bounded subset of $Y \times \mathbb{R}$. Then consider any $x^* \in Y^*$. We consider two cases.

Case (I) Suppose that for every $0 < \lambda$, $f^*(\lambda x^*) \le -\lambda$. Then $x^*(y) - \lambda^{-1} f(y) \le -1$ for all $y \in Y$ and all $0 < \lambda$. In particular, $-\lambda^{-1} f(0) \le -1$ for all $0 < \lambda$, i.e., $\lambda \le f(0)$ for all $0 < \lambda$. On the other hand, $S(0, 1) = x_0 \in C$, i.e., $(0, 1) \in S^{-1}(C)$ and so $f(0) \le 1$. Thus, Case (I) does not occur.

Case(II) Suppose that for some $0 < \lambda$, $-\lambda < f^*(\lambda x^*)$. Then, since the mapping, $\lambda' \mapsto f^*(\lambda' x^*)$, is real-valued and convex, it is continuous. Furthermore, it follows from the intermediate value theorem applied to the function $g : [0, \lambda] \to \mathbb{R}$, defined by,

$$g(\lambda') := f^*(\lambda' x^*) + \lambda' \quad \text{for all } \lambda' \in [0, \lambda],$$

that there exists a $0 < \mu < \lambda$ such that $g(\mu) = 0$, i.e., $f^*(\mu x^*) = -\mu$, since

$$g(0) = f^*(0 x^*) = - \inf_{y \in Y} f(y) \le -\varepsilon < -0 = 0 < g(\lambda).$$

Thus, $\mu(x^*, -1) = (\mu x^*, f^*(\mu x^*))$ and so $f^*(\mu x^*) = \sup\{(\mu x^*, -1)(y, s) : (y, s) \in S^{-1}(C)\} = -\mu < 0$.

Choose $(z, s_0) \in S^{-1}(C)$ such that $(\mu x^*, -1)(z, s_0) = \sup\{(\mu x^*, -1)(y, s) : (y, s) \in S^{-1}(C)\} = f^*(\mu x^*)$. Note that, as in the proof of Lemma 3.26, $z \in A$ and $s_0 = f(z)$. We claim that $(x^*, -1)$ attains its maximum value over $T(S^{-1}(C))$ at $T(z, f(z)) = f(z)^{-1}(z, -1)$. Now,

$$
\begin{aligned}
(x^*, -1)(T(z, f(z))) &= f(z)^{-1}(x^*(z) + 1) = f(z)^{-1}(x^*(z) - [\mu^{-1} f^*(\mu x^*)]) \\
&= f(z)^{-1}(x^*(z) - [x^*(z) - \mu^{-1} f(z)]) = \mu^{-1}.
\end{aligned}
$$

On the other hand, if $(y, s) \in S^{-1}(C)$ then

$$
\begin{aligned}
(x^*, -1)(T(y, s)) &= s^{-1}(x^*(y) + 1) = s^{-1}(x^*(y) - [\mu^{-1} f^*(\mu x^*)]) \\
&\le s^{-1}(x^*(y) - [x^*(y) - \mu^{-1} f(y)]) = s^{-1} f(y) \mu^{-1} \le \mu^{-1} = (x^*, -1)(T(z, f(z)))
\end{aligned}
$$

since $0 < \varepsilon \le f(y) \le s$. This completes the proof. $\quad \square$

Remark 3.28. *It might be interesting to note the following: If $(X, \|\cdot\|)$ is a Banach space, A and B are nonempty closed and bounded convex sets such that every $x^* \in X^*$ with $\inf\{x^*(a) : a \in A\} < \sup\{x^*(b) : b \in B\}$ attains its infimum over A and its supremum over B, then both A and B are weakly compact. To see this, note that $C := \text{co}[\{0\} \cup \overline{B - A}]$ is a closed and bounded convex subset of X with the property that every continuous linear function attains it supremum over C.*

A special case of the previous theorem was given in [4].

Example 3.29. *Let $(X, \|\cdot\|)$ be a non-trivial normed linear space. Then there exists an equivalent norm $\|\cdot\|$ on X and a nonempty open subset U of X^* such that every member of U attains its norm on $(X, \|\cdot\|)$.*

Proof. Choose $x_0 \in X$ with $\|x_0\| = 2$. Then, by the Hahn-Banach theorem, there exists a continuous linear functional $x^* \in S_{X^*}$ such that $x^*(x_0) = 2$. Let $U := \{y^* \in X^* : \|y^* - x^*\| < 1/3\}$ and let $B := \text{co}(B_X \cup \{x_0, -x_0\})$. Then B is convex, bounded, symmetric and $0 \in \text{int}(B)$. Therefore, B is the closed unit ball of some equivalent norm $\|\cdot\|$ on X. Furthermore, every member of U attains its maximum value over B at x_0. Indeed, if $y^* \in U$ then

$$y^*(x_0) = x^*(x_0) + [y^*(x_0) - x^*(x_0)] \ge 2 - \|y^* - x^*\| \|x_0\| > 4/3.$$

On the other hand, for any $x \in B_X$,

$$y^*(x) = x^*(x) + [y^*(x) - x^*(x)] \le 1 + \|y^* - x^*\| \|x\| < 1 + 1/3 = 4/3 < y^*(x_0)$$

and $y^*(-x_0) = -y^*(x_0) = -2 < y^*(x_0)$. Therefore, y^* attains its maximum value over B at x_0. $\quad \square$

Together, Example 3.29 and Theorem 3.23 give rise to the following conjecture.

Conjecture 3.30. *Let $(X, \|\cdot\|)$ be a Banach space. If there exists a weak open subset U of X^* such that $\varnothing \neq S_{X^*} \cap U$ and every member of $S_{X^*} \cap U$ attains its norm on X, then X is reflexive.*

A special case of this conjecture was proven in [24]. For some further results in this direction see [7].

4 Convex analysis and minimal uscos

In this section we prove a generalisation of James' weak compactness theorem. Unfortunately, to achieve this generalisation we will need to take an excursion into convex analysis and set-valued analysis. Hopefully, some of the results along the way are of some interest in their own right.

4.1 Convex functions and monotone operators

We shall need the following very important fact regarding the continuity of convex functions.

Proposition 4.1 ([42, Proposition 1.6])**.** *Let U be a nonempty open convex subset of a Banach space $(X, \|\cdot\|)$ and let $\varphi : U \to \mathbb{R}$ be a convex function. If φ is locally bounded above on U, that is, for every $x_0 \in U$ there exists an $M > 0$ and a $\delta > 0$ such that $B(x_0, \delta) \subseteq U$ and $\varphi(x) \leq M$ for all $x \in B(x_0, \delta)$, then it is locally Lipschitz on U; that is, for every $x_0 \in U$, there exists an $L > 0$ and $\delta > 0$ such that $B(x_0, \delta) \subseteq U$ and*

$$|\varphi(x) - \varphi(y)| \leq L\|x - y\|$$

for all $x, y \in B(x_0, \delta)$.

Proof. Let $x_0 \in U$. Choose $M^* > 0$ and $\delta > 0$ such that $B(x_0, 2\delta) \subseteq U$ and $\varphi(x) \leq M^*$ for all $x \in B(x_0, 2\delta)$. Then for all $x \in B(x_0, \delta)$ we have that $2x_0 - x = x_0 - (x - x_0) \in B(x_0, \delta)$ and $x_0 = (1/2)(2x_0 - x) + (1/2)x$. Hence,

$$\varphi(x_0) \leq \frac{\varphi(2x_0 - x) + \varphi(x)}{2} \leq \frac{M^* + \varphi(x)}{2},$$

so $-\varphi(x) \leq M^* + 2|\varphi(x_0)|$; that is, $|\varphi(x)| \leq (M^* + 2|\varphi(x_0)|) =: M'$ for all $x \in B(x_0, \delta)$. So $|\varphi|$ is bounded by M' on $B(x_0, \delta)$. Let $\delta' := \delta/2$. If x and y are distinct points in $B(x_0, \delta')$, let $\alpha := \|x - y\|$ and let $z := y + (\delta/\alpha)(y - x)$. Note that $z \in B(x_0, 2\delta')$. Since $y = [\alpha/(\alpha + \delta')]z + [\delta'/(\alpha + \delta')]x$ is a convex combination (lying in $B(x_0, 2\delta')$), we have that $\varphi(y) \leq [\alpha/(\alpha + \delta')]\varphi(z) + [\delta'/(\alpha + \delta')]\varphi(x)$ and so

$$\varphi(y) - \varphi(x) \leq [\alpha/(\alpha + \delta')](\varphi(z) - \varphi(x)) + [\delta'/(\alpha + \delta')](\varphi(x) - \varphi(x)) \leq (\alpha/\delta')2M' = (2M'/\delta')\|x - y\|.$$

Interchanging x and y gives the desired result, with $M := 2M'/\delta'$. $\square$

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

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

We can also define the subdifferential in terms of the right-hand derivative of f. Suppose that $f : U \to \mathbb{R}$ is a convex function defined on a nonempty open convex subset U of a normed linear space $(X, \|\cdot\|)$. Let $x_0 \in U$ and let $v \in X$. Then the *right-hand directional derivative of f, at the point $x_0 \in U$, in the direction v*, is defined to be

$$f'_+(x_0; v) := \lim_{\lambda \to 0^+} \frac{f(x_0 + \lambda v) - f(x_0)}{\lambda}.$$

Now there is a subtlety that we have overlooked. Namely, how do we know if the limit exists? Well, if we revisit Lemma 3.1, then we can see why. So suppose f, x_0 and $v \neq 0$ are as in the definition of $f'_+(x_0; v)$

and suppose that $0 < \beta$ and $0 < \beta'$ Then,

$$
\begin{aligned}
\frac{f(x_0 + (\beta + \beta')v) - f(x_0)}{\beta + \beta'} &= \frac{[f(x_0 + (\beta + \beta')v) - f(x_0 + \beta v)] + [f(x_0 + \beta v) - f(x_0)]}{\beta + \beta'} \\
&\geq \frac{1}{\beta + \beta'} \left(\frac{\beta'}{\beta}[f(x_0 + \beta v) - f(x_0)] + [f(x_0 + \beta v) - f(x_0)] \right) \quad \text{by Lemma 3.1.} \\
&= \frac{1}{\beta + \beta'} \left(\frac{\beta + \beta'}{\beta}[f(x_0 + \beta v) - f(x_0)] \right) \\
&= \frac{f(x_0 + \beta v) - f(x_0)}{\beta}.
\end{aligned}
$$

Therefore, $t \mapsto \frac{f(x_0 + tv) - f(x_0)}{t}$ is an increasing function over $(0, \delta)$ for some $\delta > 0$ small enough so that $x_0 + tv \in U$ whenever $0 < t < \delta$. Since one can also use Lemma 3.1 to show that

$$
\frac{f(x_0 + sv) - f(x_0)}{s} \leq \frac{f(x_0 + tv) - f(x_0)}{t} \quad \text{for any } s < 0 \text{ and } 0 < t, \text{ (but small enough to stay in } U),
$$

we see that the limit in the definition of the right-hand directional derivative always exists.

We can now give the basic properties of the subdifferential mapping $x \mapsto \partial\varphi(x)$.

Lemma 4.2 ([42, Proposition 1.11])**.** *Let U be a nonempty open and convex subset of a normed linear space $(X, \|\cdot\|)$ and let $\varphi : U \to \mathbb{R}$ be a continuous convex function. If $x_0 \in U$ then $\partial\varphi(x_0) \neq \varnothing$.*

Proof. Let $x_0 \in U$ and define $p : X \to \mathbb{R}$ by, $p(x) := f'_+(x_0; x)$ for all $x \in X$. Note that p is well-defined. Let $0 < \mu < \infty$ and let $x \in X$ then

$$
\begin{aligned}
p(\mu x) &= \lim_{\lambda \to 0^+} \frac{\varphi(x_0 + \lambda(\mu x)) - \varphi(x_0)}{\lambda} \\
&= \mu \lim_{\lambda \to 0^+} \frac{\varphi(x_0 + (\lambda\mu)x) - \varphi(x_0)}{\lambda\mu} \\
&= \mu \lim_{\lambda' \to 0^+} \frac{\varphi(x_0 + \lambda'x) - \varphi(x_0)}{\lambda'} \quad \text{(where, } \lambda' := \lambda\mu.) \\
&= \mu p(x).
\end{aligned}
$$

So p is positively homogeneous on X. Next, choose $\delta > 0$ such that $B[x_0, \delta] \subseteq U$. We claim that p is convex on $B[0, \delta]$. Fix $n \in \mathbb{N}$ and define $p_n : B[0, \delta] \to \mathbb{R}$ by,

$$
p_n(x) := \frac{\varphi(x_0 + (1/n)x) - \varphi(x_0)}{(1/n)} \quad \text{for all } x \in B[0, \delta].
$$

Since, $x \mapsto x_0 + (1/n)x$, is an affine map, $x \mapsto \varphi(x_0 + (1/n)x)$, is convex, and so p_n is also convex. Now, $p(x) = \lim_{n \to \infty} p_n(x)$ for each $x \in B[0, \delta]$. Therefore, $p|_{B[0,\delta]}$ is convex, as the pointwise limit of convex functions is again convex. Since p is also positively homogeneous on X it is an easy exercise to show that p is sublinear on X.

Let y_0 be any element of S_X and define $f : \text{span}\{y_0\} \to \mathbb{R}$ by, $f(\lambda y_0) := \lambda p(y_0)$ for all $\lambda \in \mathbb{R}$. Then $f(\lambda y_0) = \lambda p(y_0) = p(\lambda y_0) \leq p(\lambda y_0)$ for all $0 < \lambda < \infty$. Now, fix $0 < \lambda < \infty$, then

$$
0 = p(0) = p((-\lambda)y_0 + \lambda y_0) \leq p((-\lambda)y_0) + p(\lambda y_0).
$$

Therefore, $(-\lambda)p(y_0) = -p(\lambda y_0) \leq p((-\lambda)y_0)$. Thus,

$$
f((-\lambda)y_0) = (-\lambda)p(y_0) \leq p((-\lambda)y_0).
$$

Hence, $f(\lambda y_0) \leq p(\lambda y_0)$ for all $\lambda \in \mathbb{R}$. Thus, by the Hahn-Banach Theorem (Theorem 2.11) there exists a linear functional $F : X \to \mathbb{R}$ such that $F(x) \leq p(x)$ for all $x \in X$. Note also, that by Proposition 4.1 and

31

the definition of p, there exists an $L > 0$ such that $F(x) \leq p(x) \leq L\|x\|$ for all $x \in X$. Thus, $F \in X^*$. We claim that $F \in \partial\varphi(x_0)$. To see this, let $x \in U$ then

$$
\begin{aligned}
F(x) - F(x_0) &= F(x - x_0) \\
&\leq p(x - x_0) \\
&= \lim_{\lambda \to 0^+} \frac{\varphi(x_0 + \lambda(x - x_0)) - \varphi(x_0)}{\lambda} \\
&\leq \frac{\varphi(x_0 + 1(x - x_0)) - \varphi(x_0)}{1} \quad \text{since, } \lambda \mapsto \frac{\varphi(x_0 + \lambda(x - x_0)) - \varphi(x_0)}{\lambda}, \text{ is increasing over } (0, 1]. \\
&= \varphi(x) - \varphi(x_0).
\end{aligned}
$$

This completes the proof. $\square$

Proposition 4.3 ([42, Proposition 1.11]). *Let U be a nonempty open and convex subset of a normed linear space $(X, \|\cdot\|)$ and let $\varphi : U \to \mathbb{R}$ be a continuous convex function. If $x_0 \in U$, then $\partial\varphi(x_0)$ is a weak*-compact convex subset of X^*. Moreover, the map $x \mapsto \partial\varphi(x)$ is locally bounded at x_0. That is, there exists an $L > 0$ and a $\delta > 0$ such that $B(x_0, \delta) \subseteq U$ and $\|x^*\| \leq L$ whenever $x \in B(x_0, \delta)$ and $x^* \in \partial\varphi(x)$.*

Proof. For each $x \in U$, let $F_x := \{x^* \in X^* : x^*(x - x_0) \leq \varphi(x) - \varphi(x_0)\} = (\widehat{x - x_0})^{-1}(-\infty, \varphi(x) - \varphi(x_0)]$. Thus, each set F_x is weak* closed and convex. Now, $\partial\varphi(x_0) = \bigcap_{x \in U} F_x$. Therefore, $\partial\varphi(x_0)$ is weak* closed and convex. Let us now show that, $x \mapsto \partial\varphi(x)$, is locally bounded at x_0 (Note: this will then automatically show that $\partial\varphi(x)$ is weak* compact, by Theorem 2.23). By Proposition 4.1, there exists an $L > 0$ and a $\delta > 0$ such that $B(x_0, \delta) \subseteq U$ and $|\varphi(x) - \varphi(y)| \leq L\|x - y\|$ for all $x, y \in B(x_0, \delta)$. We claim that $\|x^*\| \leq L$ whenever $x \in B(x_0, \delta)$ and $x^* \in \partial\varphi(x)$. To this end, let $x \in B(x_0, \delta)$ and $x^* \in \partial\varphi(x)$. Let $v \in S_X$ and choose $0 < \mu$ such that $x + \mu v \in B(x_0, \delta)$. Then,

$$
x^*(v) = \frac{x^*((x + \mu v) - x)}{\mu} \leq \frac{\varphi(x + \mu v) - \varphi(x)}{\mu} \leq \frac{L\|\mu v\|}{\mu} = L.
$$

Thus, $\|x^*\| \leq L$. Note: we used here the simple fact that if $x^*(v) \leq L$ for all $v \in S_X$ then $\|x^*\| \leq L$. $\square$

One of the most important features of the subdifferential mapping of a convex function is that it belongs to a much studied class of set-valued mappings called "monotone operators".

Let $T : X \to 2^{X^*}$ be a set-valued mapping from $(X, \|\cdot\|)$ be a normed linear space into subsets of its dual X^*. T is said to be a *monotone operator* provided $(x^* - y^*)(x - y) \geq 0$ whenever $x, y \in X$ and $x^* \in T(x)$, $y^* \in T(y)$.

Proposition 4.4 ([42, Example 2.2]). *If $\varphi : U \to \mathbb{R}$ be a continuous convex function defined on a nonempty open convex subset U of a normed linear space $(X, \|\cdot\|)$ then $T : X \to 2^{X^*}$ defined by,*

$$
T(x) := \begin{cases} \partial\varphi(x), & x \in U \\ \varnothing, & x \notin U \end{cases}
$$

is a monotone operator on X.

Proof. Let $x^*, y^* \in X^*$ and suppose that $x^* \in T(x)$ and $y^* \in T(y)$ for some $x, y \in X$. Then $x, y \in U$ since $T(x) \neq \varnothing$ and $T(y) \neq \varnothing$. In fact, $T(x) = \partial\varphi(x)$ and $T(y) = \partial\varphi(y)$. Therefore,

$$
x^*(y - x) \leq \varphi(y) - \varphi(x) \quad \text{and} \quad (-y^*)(y - x) = y^*(x - y) \leq \varphi(x) - \varphi(y).
$$

If we add these two inequalities together we get $(x^* - y^*)(y - x) \leq 0$ and so $(x^* - y^*)(x - y) \geq 0$. Hence, T is indeed a monotone operator. $\square$

4.2 Minimal Uscos

In order to prove our final "convex analysts" proof of James' theorem, we will need to briefly consider some notions from set-valued analysis.

A set-valued mapping φ from a topological space A 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 $\varphi(x_0)$, there exists an open neighbourhood U of x_0 such that $\varphi(U) \subseteq W$. If φ is τ-upper semicontinuous at each point of A then we say that φ is τ-*upper semicontinuous on A*. In the case when φ also has nonempty compact images then we call φ a τ-*usco* mapping. Finally, if (X, τ) is a linear topological space then we call a τ-usco mapping into convex subsets of X a τ-*cusco* mapping.

Our interest in cusco mappings is revealed in the next proposition.

Proposition 4.5 ([42, Proposition 2.5]). *If $\varphi : U \to \mathbb{R}$ is a continuous convex function defined on a nonempty open convex subset U of a normed linear space $(X, \| \cdot \|)$, then the subdifferential mapping, $x \mapsto \partial\varphi(x)$, is a weak*-cusco on U.*

Proof. It follows from Lemma 4.2 and Proposition 4.3 that we need only show that, $x \mapsto \partial\varphi(x)$, is weak*-upper semicontinuous on U. So suppose, in order to obtain a contradiction, that $\partial\varphi$ is not weak* upper semicontinuous at some point $x_0 \in U$. Then there exists a weak* open subset W of X^*, containing $\partial\varphi(x_0)$, such that for every $0 < \delta$, $\partial\varphi(B(x_0, \delta)) \not\subseteq W$. Therefore, in particular, there exist sequences $(x_n : n \in \mathbb{N})$ in U and $(x_n^* : n \in \mathbb{N})$ in X^* such that $\lim_{n \to \infty} x_n = x_0$ and $x_n^* \in \partial\varphi(x_n) \backslash W$. Furthermore, by Proposition 4.3, we can assume that the sequence $(x_n^* : n \in \mathbb{N})$ is norm bounded in X^*. That is, there exists an $0 < L$ such that $\|x_n^*\| \leq L$ for all $n \in \mathbb{N}$. Hence, by the Banach-Alaoglu Theorem, [1] the sequence $(x_n^* : n \in \mathbb{N})$ has a weak* cluster-point x_∞^*, which must lie in $X^* \backslash W$. We will obtain our desired contradiction by showing that $x_\infty^* \in \partial\varphi(x_0) \subseteq W$. To this end, fix $x \in U$ and $0 < \varepsilon$. Since φ is continuous at x_0 there exists an $N \in \mathbb{N}$ such that $\|x_n - x_0\| < \varepsilon/L$ and $|\varphi(x_n) - \varphi(x_0)| < \varepsilon$ for all $N < n$. Let $N < n$ then,

$$
\begin{aligned}
(\widehat{x - x_0})(x_n^*) = x_n^*(x - x_0) &= x_n^*(x - x_n) + x_n^*(x_n - x_0) \\
&\leq \varphi(x) - \varphi(x_n) + \|x_n^*\|\|x_n - x_0\| \\
&\leq [\varphi(x) - \varphi(x_0)] + [\varphi(x_0) - \varphi(x_n)] + L\|x_n - x_0\| \\
&\leq [\varphi(x) - \varphi(x_0)] + \varepsilon + \varepsilon.
\end{aligned}
$$

Therefore, $x_\infty^*(x - x_0) = (\widehat{x - x_0})(x_\infty^*) \leq [\varphi(x) - \varphi(x_0)] + 2\varepsilon$. Since $0 < \varepsilon$ was arbitrary, we have that $x_\infty^*(x - x_0) \leq \varphi(x) - \varphi(x_0)$. Since $x \in U$ was arbitrary, we have that $x_\infty^* \in \partial\varphi(x_0)$, as desired. $\square$

Among the class of usco (cusco) mappings, special attention is given to the so-called minimal usco (minimal cusco) mappings.

An usco (cusco) from a topological space A into subsets of a topological space X (linear topological space X) 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.

It is not immediately obvious from this definition that there are any interesting minimal usco mappings at all, apart from single-valued continuous functions (e.g. $f : A \to X$), which are trivially minimal uscos once one replaces $f(x)$ with $\{f(x)\}$ - to make them set-valued mappings. So our first task is to show that there are always many minimal uscos.

Proposition 4.6 ([6]). *Suppose that (X, τ) and (Y, τ') are topological spaces and $\varphi : X \to 2^Y$ is an usco on X. If (Y, τ') is Hausdorff then there exists a minimal usco mapping $\Psi : X \to 2^Y$ such that $\mathrm{Gr}(\Psi) \subseteq \mathrm{Gr}(\varphi)$ (i.e., every usco contains a minimal usco).*

Proof. Let $\mathcal{U}$ denote the family of all usco mappings defined on X whose graphs are contained in the graph of φ. Obviously $\mathcal{U} \neq \varnothing$ as the mapping φ is contained in $\mathcal{U}$. We may now partially order $\mathcal{U}$ as follows. If Ψ_1 and Ψ_2 are members of $\mathcal{U}$, then we write $\Psi_1 \leq \psi_2$ if $\Psi_1(x) \subseteq \Psi_2(x)$ for each $x \in X$. Next, we apply Zorn's lemma to show that $(\mathcal{U}, \leq)$ possesses a minimal element. To this end, let $\{\Psi_\gamma : \gamma \in \Gamma\}$ be a totally ordered subset of $\mathcal{U}$ and let $\varphi_M : X \to 2^Y$ be defined by, $\varphi_M(x) := \bigcap\{\Psi_\gamma(x) : \gamma \in \Gamma\}$. Since each $\Psi_\gamma(x)$ is

nonempty and compact, $\varphi_M(x)$ too is nonempty and compact. Let W be an open subset of Y and consider $U := \{x \in X : \varphi_M(x) \subseteq W\}$. We need to show that U is open in X. We may, without loss of generality, assume that $U \neq \varnothing$ and consider $x_0 \in U$. By the finite intersection property, there exists some $\gamma_0 \in \Gamma$ such that $\Psi_{\gamma_0}(x_0) \subseteq W$. Hence there exists an open neighbourhood U_0 of x_0 such that $\Psi_{\gamma_0}(U_0) \subseteq W$, which means that $\varphi_M(U_0) \subseteq W$. Therefore $x_0 \in U_0 \subseteq U$ and so U is open in X. From this, it follows that $\varphi_M \in \mathcal{U}$ and $\varphi_M \leq \Psi_\gamma$ for each $\gamma \in \Gamma$. Thus, by Zorn's lemma, $(\mathcal{U}, \leq)$ possesses a minimal element. It is now easy to see that this element is in fact a minimal usco. $\square$

A similar argument shows that every cusco contains a minimal cusco. However, there is a much more concrete supply of minimal cuscos.

Proposition 4.7. *Let $\varphi : A \to 2^{X^*}$ be a weak*-cusco defined on a nonempty open subset A of a normed linear space $(X, \|\cdot\|)$. If the mapping $T : X \to 2^{X^*}$ defined by, $T(x) := \varphi(x)$ if $x \in A$ and by $T(x) := \varnothing$ if $x \in X \setminus A$, is a monotone operator, then φ is a minimal weak*-cusco.*

Proof. Suppose, in order to obtain a contradiction, that φ is not a minimal weak*-cusco. Then there exists a weak*-cusco $\Psi : A \to 2^{X^*}$ such that $\Psi(x) \subseteq \varphi(x)$ for all $x \in A$, but $\Psi(x_0) \neq \varphi(x_0)$ for some $x_0 \in A$. Choose $x_0^* \in \varphi(x_0) \setminus \Psi(x_0) = T(x_0) \setminus \Psi(x_0)$. By the Separation Theorem (Theorem 2.17), applied in (X^*, weak^*), there exists a $y \in X$ such that $\sup_{y^* \in \Psi(x_0)} \widehat{y}(y^*) < \widehat{y}(x_0^*)$. Let $W := \{x^* \in X^* : \widehat{y}(x^*) < \widehat{y}(x_0^*)\}$. Then W is a weak*-open subset of X^*, containing $\Psi(x_0)$. Therefore, there exists an open neighbourhood $U \subseteq A$ of x_0 such that $\Psi(U) \subseteq W$. Choose $0 < t < \infty$ such that $x_0 + ty \in U$. Let $y^* \in \Psi(x_0 + ty) \subseteq \varphi(x_0 + ty) = T(x_0 + ty)$. Since T is a monotone operator, $x_0^* \in T(x_0)$ and $y^* \in T(x_0 + ty)$, we have that:

$$t(y^* - x_0^*)(y) = (x_0^* - y^*)(-ty) = (x_0^* - y^*)(x_0 - (x_0 + ty)) \geq 0;$$

which implies that $y^*(y) \geq x_0^*(y)$. However, this contradicts the fact that $y^* \in W$, i.e., $\widehat{y}(y^*) < \widehat{y}(x_0^*)$. Thus, φ must be a minimal weak*-cusco on A. $\square$

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

Proof. By Proposition 4.5 we have that, $x \mapsto \partial\varphi(x)$, is a weak*-cusco on U. So the result follows from Proposition 4.4 and Proposition 4.7. $\square$

We will end our detour into set-valued analysis by giving two more results concerning uscos. The first one shows that minimal usco behave a lot like quasi-continuous mappings, see [12, page 52], while the last result shows how to convert an usco into a cusco.

Proposition 4.9 ([6, 12]). *Let $\varphi : A \to 2^X$ be a minimal τ-usco acting from a topological space A into nonempty subsets of a topological space (X, τ). Then, for every pair of open subsets U of A and W of X such that $\varphi(U) \cap W \neq \varnothing$, there exists a nonempty open subset V of U such that $\varphi(V) \subseteq W$.*

Proof. Let U be an open subset of A and let W be an open subset of X such that $\varphi(U) \cap W \neq \varnothing$. We consider two cases.

Case(I): If there exists a $x \in U$ such that $\varphi(x) \subseteq W$, then the result follows directly from the τ-upper semicontinuity of φ.

Case(II): Suppose that for each $x \in U$, $\varphi(x) \not\subseteq W$. Let $\Psi : A \to 2^X$ be defined by, $\Psi(x) := \varphi(x) \cap (X \setminus W)$ if $x \in U$ and by $\Psi(x) := \varphi(x)$ if $x \notin U$. Then, by assumption, Ψ has nonempty compact images. In fact, we claim that Ψ is a τ-usco on A. To show this, we need only show that Ψ is τ-upper semicontinuous. Let $x_0 \in A$ and let W' be a τ-open set in X containing $\Psi(x_0)$. If $x_0 \notin U$ then clearly there exists an open neighbourhood V of x_0 such that $\Psi(V) \subseteq W'$ since, in this case, $\varphi(x_0) = \Psi(x_0) \subseteq W'$ and $\Psi(x) \subseteq \varphi(x)$ for all $x \in A$. So we are left to consider the case when $x_0 \in U$. Suppose $x_0 \in U$. Then $\varphi(x_0) \subseteq W' \cup W$, since $\varphi(x_0) \cap (X \setminus W) = \Psi(x_0) \subseteq W'$. Since φ is τ-upper semicontinuous there exists an open neighbourhood V of x_0, contained in U, such that $\varphi(V) \subseteq W' \cup W$. Therefore,

$$\Psi(V) = \varphi(V) \cap (X \setminus W) \subseteq (W' \cup W) \cap (X \setminus W) = W' \cap (X \setminus W) \subseteq W'.$$

This shows that Ψ is an τ-usco. Since, φ is a minimal τ-usco, we must have that $\varphi = \Psi$, but then $\varphi(U) = \Psi(U) \subseteq (X \setminus W)$, which contradicts our original assumption that $\varphi(U) \cap W \neq \varnothing$. Therefore, Case(II) does not occur, and so the result follows from Case(I). $\square$

Proposition 4.10 ([25, 42]). *Suppose that $\varphi : A \to 2^X$ is a τ-usco acting from a topological space A into nonempty subsets of a locally convex space $(X, +, \cdot, \tau)$. If for each $t \in A$, $\overline{co}^\tau \varphi(t)$ is a compact subset of X, then the mapping $\Psi : A \to 2^X$ defined by, $\Psi(t) := \overline{co}^\tau \varphi(t)$ for all $t \in A$, is a τ-cusco on A.*

Proof. Clearly, Ψ has nonempty, compact convex images. So it is sufficient to show that Ψ is τ-upper semicontinuous on A. Let $x_0 \in A$ and let W be a τ-open subset of X, containing $\Psi(x_0)$. Since vector addition is continuous, for each $x \in \Psi(x_0)$ there exist τ-open convex neighbourhoods U_x of x and V_x of 0 such that $x = x + 0 \subseteq U_x + V_x \subseteq W$. Since linear topological spaces are also regular we can assume, by possibly making V_x smaller, that $U_x + \overline{V_x}^\tau \subseteq W$. Now, $\{U_x : x \in \Psi(x_0)\}$ is an open cover of $\Psi(x_0)$. Therefore, there exists a finite subcover $\{U_{x_k} : 1 \leq k \leq n\}$ of $\{U_x : x \in \Psi(x_0)\}$. Let $V := \bigcap_{1 \leq k \leq n} V_{x_k}$. Then V is a convex open neighbourhood of 0 and futhermore,

$$\Psi(x_0) + \overline{V}^\tau \subseteq (\bigcup_{1 \leq k \leq n} U_{x_k}) + \overline{V}^\tau = \bigcup_{1 \leq k \leq n}(U_{x_k} + \overline{V}^\tau) = \subseteq \bigcup_{1 \leq k \leq n}(U_{x_k} + \overline{V_{x_k}}^\tau) \subseteq W.$$

Since $\varphi(x_0) \subseteq \Psi(x_0) + V$, which is τ-open, there exists an open neighbourhood U of x_0 such that $\varphi(U) \subseteq \Psi(x_0) + V$. Let $x \in U$. Then

$$\Psi(x) = \overline{co}^\tau \varphi(x) \subseteq \overline{co}^\tau(\Psi(x_0) + V) \subseteq \Psi(x_0) + \overline{V}^\tau \subseteq W \quad \text{since, } \Psi(x_0) + \overline{V}^\tau \text{ is closed and convex.}$$

Here we used the fact that the sum of a closed set with a compact set is closed. $\square$

4.3 A generalisation of James' Theorem

By making an obvious modification to Corollary 3.13, we obtain the following lemma.

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, τ) 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 $\widetilde{x} := (x_n : n \in \mathbb{N})$ in T, either there exists a subsequence without any τ'-cluster points, or else, there exists a subsequence, $\widetilde{x}|_J$, of $\widetilde{x}$ such that φ is constant on, $y - aK_{\tau'}(\widetilde{x}|_J)$ for all $y \in S$ and all $a \in [0, 1]$.*

The following lemma shows us that in Theorem 4.13 we get the weak* lower semicontinuity of φ for free.

Lemma 4.12. *Let $(X, \|\cdot\|)$ be a Banach space and let A be a nonempty open convex subset of X^*. If $\varphi : A \to \mathbb{R}$ is a continuous convex function and $\partial\varphi(x^*) \cap \widehat{X} \neq \varnothing$ for all $x^* \in A$, then φ is weak*-lower-semicontinuous on A.*

Proof. Let $x_0^* \in A$ and let $\varepsilon > 0$. Then, there exists an $\widehat{x} \in \partial\varphi(x_0^*) \cap \widehat{X}$. Define $h : A \to \mathbb{R}$ to be

$$h(x^*) := \widehat{x}(x^*) - \widehat{x}(x_0^*) + \varphi(x_0^*) \quad \text{for all } x^* \in A.$$

Then observe that, since $\widehat{x} \in \partial\varphi(x_0^*)$, we have $h(x^*) \leq \varphi(x^*)$ for all $x^* \in A$. Now the set

$$U := \{x^* \in A : |\widehat{x}(x^* - x_0^*)| < \varepsilon\},$$

is a weak*-open neighbourhood of x_0^*, and for all $x^* \in U$, we have that

$$\varphi(x_0^*) - \varepsilon < \widehat{x}(x^*) - \widehat{x}(x_0^*) + \varphi(x_0^*) = h(x^*) \leq \varphi(x^*).$$

Therefore, φ is weak*-lower-semicontinuous at x_0^*. Since x_0^* was arbitrary, we conclude that φ is weak*-lower-semicontinuous on A. $\square$

At last, we can present our "convex analysts" version of James' weak compactness theorem.

Theorem 4.13. *Let $(X, \|\cdot\|)$ be a Banach space and let A be a nonempty open convex subset of X^*. If $\varphi : A \to \mathbb{R}$ is a continuous convex function and $\partial\varphi(x^*) \cap \widehat{X} \neq \varnothing$ for all $x^* \in A$, then $\partial\varphi(x^*) \subseteq \widehat{X}$ for all $x^* \in A$.*

Proof. Let $x_0^* \in A$. Without loss of generality, we may assume that $x_0^* = 0$. Indeed, if not, we consider the function $\psi : (A - x_0^*) \to \mathbb{R}$ given by $\psi(x^*) := \varphi(x^* + x_0^*)$. Note that ψ is continuous and convex and that $\partial\varphi(x^* + x_0^*) = \partial\psi(x^*)$ for all $x^* \in A - x_0^*$. In particular, $\partial\psi(x^*) \cap \widehat{X} \neq \varnothing$ for all $x^* \in (A - x_0^*)$ and $\partial\varphi(x_0^*) = \partial\psi(0)$. So, if $x_0^* \neq 0$, we can simply translate φ and use the argument at 0.

Since A is open and since $x^* \mapsto \partial\varphi(x^*)$ is locally bounded (Proposition 4.3), there exist $m, L > 0$ such that $mB_{X^*} \subseteq A$ and $\|x^{**}\| \leq L$ for all $x^{**} \in \partial\varphi(mB_{X^*})$. Let $(\beta_n : n \in \mathbb{N})$ be a sequence of strictly positive numbers such that $\sum_{n=1}^\infty \beta_n < m/2$ and $\lim_{n\to\infty} \frac{1}{\beta_n} \sum_{i=n+1}^\infty \beta_i = 0$.

Since $x^* \mapsto \partial\varphi(x^*)$ is a minimal weak*-cusco, (see, Corollary 4.8) we know that there exists a minimal weak*-usco, $M : A \to 2^{X^{**}}$, such that $M(x^*) \subseteq \partial\varphi(x^*)$ for all $x^* \in A$, by Proposition 4.6. In fact, by Proposition 4.10, we know that $\partial\varphi(x^*) = \overline{\mathrm{co}}^{w^*}[M(x^*)]$ for all $x^* \in A$.

Therefore, to show that $\partial\varphi(0) \subseteq \widehat{X}$, it suffices to show that $M(0) \subseteq \widehat{X}$. This is because if $M(0) \subseteq \widehat{X}$, then $M(0)$ is weakly compact (see Remark 2.25) and then, by the Krein-Phillips Theorem (Corollary 3.18), $\overline{\mathrm{co}}[M(0)]$ is also weakly compact. Since the weak* topology is weaker than the weak topology, $\overline{\mathrm{co}}[M(0)]$ is clearly weak*-compact and hence weak*-closed. Therefore,

$$\partial\varphi(0) = \overline{\mathrm{co}}^{w^*}[M(0)] = \overline{\mathrm{co}}[M(0)] \subseteq \widehat{X}.$$

So suppose, for a contradiction, that $M(0) \nsubseteq \widehat{X}$. Then there exists an $F \in M(0) \setminus \widehat{X}$. Since $\widehat{X}$ is a closed subspace of X^{**}, this means there must exist an $0 < \varepsilon < \mathrm{dist}(F, \widehat{X})$.

Part I: Let $f_0 := 0$. We inductively create sequences $(f_n : n \in \mathbb{N})$ in S_{X^*}, $(v_n : n \in \mathbb{N})$ in $mB_{X^*} \subseteq A$, and $(\widehat{x}_n : n \in \mathbb{N})$ in $\widehat{X}$, such that the statements

(A_n) $\|v_n\| < m/n$ and $\widehat{x}_n \in \partial\varphi(v_n)$.

(B_n) $|(F - \widehat{x}_n)(f_j)| \leq \varepsilon/2$ for all $0 \leq j < n$.

(C_n) $F(f_n) > \varepsilon$ and $\widehat{x}_j(f_n) = 0$ 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 X^*, centred at 0, of radius $r > 0$.] For the first step, choose any $v_1 \in B(0, m) \subseteq A$. Then $\partial\varphi(v_1) \cap \widehat{X} \neq \varnothing$ and so we may choose $\widehat{x}_1 \in \partial\varphi(v_1) \cap \widehat{X}$ which clearly satisfies $|(F - \widehat{x}_1)(f_0)| = 0 \leq \varepsilon/2$. Now note that

$$\mathrm{dist}(F, \mathrm{span}\{\widehat{x}_1\}) \geq \mathrm{dist}(F, \widehat{X}) > \varepsilon,$$

and so, by Lemma 3.8, there exists $f_1 \in S_{X^*}$ such that $F(f_1) > \varepsilon$ and $\widehat{x}_1(f_1) = 0$. So the statements (A_1), (B_1) and (C_1) hold.

Now fix $k \in \mathbb{N}$. Suppose that we have created $\{v_1, \ldots, v_k\}$, $\{\widehat{x}_1, \ldots, \widehat{x}_k\}$ and $\{f_1, \ldots, f_k\}$ such that the statements (A_k), (B_k) and (C_k) hold true. Then consider the set

$$W := \bigcap_{j=0}^k \{G \in X^{**} : |(F - G)(f_j)| < \varepsilon/2\}.$$

Note that $F \in M(0) \cap W$ and so $M(B(0, \frac{m}{k+1})) \cap W \neq \varnothing$. Therefore, by the minimality of M and Proposition 4.9, 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 $\partial\varphi(v_{k+1}) \cap \widehat{X} \neq \varnothing$, and so we may choose $\widehat{x}_{k+1} \in \widehat{X}$ such that

$$\widehat{x}_{k+1} \in \partial\varphi(v_{k+1}) = \overline{\mathrm{co}}^{w^*}[M(v_{k+1})] \subseteq \overline{\mathrm{co}}^{w^*}(W) \subseteq \bigcap_{j=0}^k \{G \in X^{**} : |(F - G)(f_j)| \leq \varepsilon/2\}.$$

36

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

$$\text{dist}(F, \text{span}\{\widehat{x}_1, \ldots, \widehat{x}_{k+1}\}) \geq \text{dist}(F, \widehat{X}) > \varepsilon,$$

and so, by Lemma 3.8, there exists $f_{k+1} \in S_{X^*}$ such that $F(f_{k+1}) > \varepsilon$ and $\widehat{x}_j(f_{k+1}) = 0$ 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 $x'_k := x_{n_k}$ and $f'_k := f_{n_k}$. Also define $f'_0 := 0$. Then the sequences $(v'_n : n \in \mathbb{N})$, $(\widehat{x}'_n : n \in \mathbb{N})$ and $(f'_n : n \in \mathbb{N})$ 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 $\partial\varphi(x^*) \cap \widehat{X} \neq \varnothing$ for all $x^* \in A$, we have that φ is weak*-lower-semicontinuous on A, by Lemma 4.12. Let $S := \frac{m}{2}B_{X^*} \cap \text{span}\{f_n : n \in \mathbb{N}\}$ and $T := \frac{m}{2}B_{X^*}$ and note that $S - T \subseteq mB_{X^*} \subseteq A$. Then, by passing to a subsequence and relabelling if necessary, we may assume, by Lemma 4.11, that φ is constant on $f - a[(m/2)K_{w^*}(f_n : n \in \mathbb{N})]$ for all $0 \leq a \leq 1$ and all $f \in S$. Since $(f_n : n \in \mathbb{N})$ is a sequence in B_{X^*} (which is weak*-compact), it must have a weak*-cluster point, call it f_∞.

Part III: As in Part III of the proof of Theorem 3.9, we derive from the statements (B_n) and (C_n) that $\widehat{x}_k(f_\infty) = 0$ for all $k \in \mathbb{N}$ and that $\widehat{x}_n(f_k) > \varepsilon/2$ for all $n > k$; from which we conclude that $\widehat{x}_n(f_k - f_\infty) > \varepsilon/2$ for all $n > k$. We also note that, from the statement (A_n), we have $v_n \to 0$ in norm. Therefore, since φ is norm-continuous, there exists $N_0 \in \mathbb{N}$ such that

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

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

$$|\widehat{x}_n(v_n)| \leq \|\widehat{x}_n\| \|v_n\| \leq \frac{Lm}{n} < \frac{\beta_1 \varepsilon}{8}.$$

Part IV: For each $n \in \mathbb{N}$, let $K_n := \text{co}\{f_k : k \geq n\} - f_\infty$ and note that $(K_n : n \in \mathbb{N})$ is a decreasing sequence of nonempty convex subsets of X^*. Set $r := \varepsilon/8 > 0$. Then we have that

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

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

$$\begin{aligned}
\varphi(\beta_1 f) - \varphi(0) &= [\varphi(\beta_1 f) - \varphi(v_N)] + [\varphi(v_N) - \varphi(0)] \\
&\geq \widehat{x}_N(\beta_1 f) - \widehat{x}_N(v_N) + [\varphi(v_N) - \varphi(0)] \quad \text{since } \widehat{x}_N \in \partial\varphi(v_N) \\
&> \beta_1 \widehat{x}_N(f) - \widehat{x}_N(v_N) - \beta_1 \varepsilon/8 \\
&> \beta_1 \widehat{x}_N(f) - \beta_1 \varepsilon/8 - \beta_1 \varepsilon/8 \\
&= \beta_1(\widehat{x}_N(\sum_{i=1}^k \lambda_i f_{n_i} - f_\infty)) - \beta_1 \varepsilon/4 \\
&= \beta_1 \sum_{i=1}^k \lambda_i \widehat{x}_N(f_{n_i} - f_\infty) - \beta_1 \varepsilon/4 > \beta_1 \varepsilon/4.
\end{aligned}$$

Therefore, since $f \in K_1$ was arbitrary, we have that $\beta_1 r + \varphi(0) < \inf_{f \in K_1} \varphi(\beta_1 f)$ as claimed. So, by Lemma 3.2, there exists a sequence $(g_n : n \in \mathbb{N})$ such that for all $n \in \mathbb{N}$:

(i) $g_n \in \text{co}\{f_k : k \geq n\}$ and

(ii) $\varphi(\sum_{i=1}^n \beta_i(g_i - f_\infty)) + \beta_{n+1} r < \varphi(\sum_{i=1}^{n+1} \beta_i(g_i - f_\infty))$. $\quad (*)$

Part V: Since $(g_n : n \in \mathbb{N})$ is a sequence in B_{X^*}, it must have a weak*-cluster point. So, let g_∞ be a weak*-cluster point of $(g_n : n \in \mathbb{N})$. Then, by Proposition 3.14, we have that $g_\infty \in K_{w^*}(f_n : n \in \mathbb{N})$. Since for all $n \in \mathbb{N}$, we have that $\sum_{i=1}^n \beta_i g_i \in S = \frac{m}{2}B_{X^*} \cap \text{span}\{f_n : n \in \mathbb{N}\}$, and $0 \leq \sum_{i=1}^n \beta_i \leq m/2$, then

$$\varphi(\sum_{i=1}^n \beta_i(g_i - g_\infty)) = \varphi(\sum_{i=1}^n \beta_i g_i - \sum_{i=1}^n \beta_i \cdot g_\infty) = \varphi(\sum_{i=1}^n \beta_i g_i - \sum_{i=1}^n \beta_i \cdot f_\infty) = \varphi(\sum_{i=1}^n \beta_i(g_i - f_\infty)), \quad (**)$$

37

by the observation made in Part II. As in Part V of the proof of Theorem 3.9, we set $g := \sum_{i=1}^{\infty} \beta_i(g_i - g_\infty)$ and deduce that $g \in X^*$.

Part VI: Lastly note that $\|g\| \leq 2\sum_{i=1}^{\infty} \beta_i \leq m$ and so $g \in mB_{X^*} \subseteq A$. Therefore, there exists $\widehat{x} \in \widehat{X}$ such that $\widehat{x} \in \partial\varphi(g)$. Then, if $n > 1$,

$$\begin{aligned}
\beta_n r &< \varphi(\textstyle\sum_{i=1}^{n}\beta_i(g_i - f_\infty)) - \varphi(\textstyle\sum_{i=1}^{n-1}\beta_i(g_i - f_\infty)) \qquad \text{by } (*) \\
&= \varphi(\textstyle\sum_{i=1}^{n}\beta_i(g_i - g_\infty)) - \varphi(\textstyle\sum_{i=1}^{n-1}\beta_i(g_i - g_\infty)) \qquad \text{by } (**) \\
&\leq \varphi(g) - \varphi(\textstyle\sum_{i=1}^{n-1}\beta_i(g_i - g_\infty)) \qquad \text{since } \varphi(g) = \sup_{n\in\mathbb{N}}\varphi(\textstyle\sum_{i=1}^{n}\beta_i(g_i - g_\infty)) \\
&\leq \widehat{x}(g) - \widehat{x}(\textstyle\sum_{i=1}^{n-1}\beta_i(g_i - g_\infty)) \qquad \text{since } \widehat{x} \in \partial\varphi(g) \\
&= \widehat{x}(\textstyle\sum_{i=n}^{\infty}\beta_i(g_i - g_\infty)) = \beta_n\widehat{x}(g_n - g_\infty) + \textstyle\sum_{i=n+1}^{\infty}\beta_i\widehat{x}(g_i - g_\infty).
\end{aligned}$$

Rearranging gives us that

$$r < \widehat{x}(g_n - g_\infty) + \frac{1}{\beta_n}\textstyle\sum_{i=n+1}^{\infty}\beta_i\widehat{x}(g_i - g_\infty) \leq \widehat{x}(g_n - g_\infty) + \frac{2\|\widehat{x}\|}{\beta_n}\textstyle\sum_{i=n+1}^{\infty}\beta_i.$$

Taking the limit, as $n \to \infty$, we get that

$$r \leq \liminf_{n\to\infty} g_n(x) - g_\infty(x) + 2\|\widehat{x}\|\left(\lim_{n\to\infty}\frac{1}{\beta_n}\textstyle\sum_{i=n+1}^{\infty}\beta_i\right) = \liminf_{n\to\infty} g_n(x) - g_\infty(x),$$

which contradicts the inequality $\liminf\limits_{n\to\infty} g_n(x) \leq g_\infty(x)$. Thus, our assumption that $M(0) \not\subseteq \widehat{X}$ was false. This completes the proof. $\square$

Remark 4.14. *To see that Theorem 4.13 is indeed a generalisation of Theorem 3.15 consider the following. Suppose that C is a nonempty closed and bounded convex subset of a Banach space $(X, \|\cdot\|)$ with $0 \in C$. Define $p : X^* \to \mathbb{R}$ by, $p(x^*) := \sup_{c\in C} x^*(c)$ for all $x^* \in X^*$. Then $\widehat{C} \subseteq \partial p(0)$. If every $x^* \in X^*$ attains its supremum over C then $\partial p(x^*) \cap \widehat{X} \neq \varnothing$ for every $x^* \in X^*$. This last fact follows because, if $x^* \in X^* \setminus \{0\}$, $c \in C$ and $p(x^*) = x^*(c)$ then $\widehat{c} \in \partial p(x^*)$. Thus, by Theorem 4.13,*

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

Hence, C is weakly compact by Remark 2.25. Let us also note that an earlier version of Theorem 4.13 appeared in [33, Theorem 2.2].

5 Variational Principle

The corner stone of this section is the Brøndsted-Rockafellar Theorem which gives the existence of subgradients for lower semicontinuous convex functions defined on Banach spaces. The key notion behind this theorem is the notion of an "ε-subgradient". 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 $\varepsilon > 0$, we define the ε-*subdifferential* $\partial_\varepsilon f(x)$ by,

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

Theorem 5.1 ([3]). *Suppose that $f : X \to \mathbb{R} \cup \{\infty\}$ is a convex proper lower semicontinuous function on a Banach space $(X, \|\cdot\|)$. Then, given any point $x_0 \in \text{Dom}(f)$, $\varepsilon > 0$ and any $x_0^* \in \partial_\varepsilon f(x_0)$, there exists $x \in \text{Dom}(f)$ and $x^* \in X^*$ such that $x^* \in \partial f(x)$, $\|x - x_0\| \leq \sqrt{\varepsilon}$ and $\|x^* - x_0^*\| \leq \sqrt{\varepsilon}$.*

We will use the Brøndsted-Rockafellar Theorem (Theorem 5.1) to show that certain functions attain their maximum value in a rather strong way, that we now make precise. We shall say that a function $f : X \to [-\infty, \infty)$ defined on a normed linear space $(X, \|\cdot\|)$ attains a (or has a) *strong maximum at $x_0 \in X$* if, $f(x_0) = \sup_{x\in X} f(x)$ and $\lim_{n\to\infty} x_n = x_0$ 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 addition to the Brøndsted-Rockafellar Theorem and the definition of a strong maximum, we shall require one more definition. Let $\varphi : X \to 2^Y$ be a set-valued mapping acting between a topological space (X, τ) and a normed linear space $(Y, \|\cdot\|)$. Then we say that φ is *single-valued and norm upper semicontinuous at* $x_0 \in X$ if, $\varphi(x_0) =: \{y_0\}$ is a singleton subset of Y and for each $\varepsilon > 0$ there exists an open neighbourhood U of x_0 such that $\varphi(U) \subseteq B[y_0, \varepsilon]$.

We shall now combine the Brøndsted-Rockafellar Theorem with these definitions in order to obtain the following preliminary result.

Proposition 5.2. *Suppose that* $f : X \to \mathbb{R} \cup \{\infty\}$ *is a proper function on a Banach space* $(X, \|\cdot\|)$ *and suppose that* $f^* : X^* \to \mathbb{R} \cup \{\infty\}$ *(the Fenchel conjugate of* f*), see [2, 46], is defined by,*

$$f^*(x^*) := \sup_{x \in X} [x^*(x) - f(x)] \ = \ \sup_{x \in \mathrm{Dom}(f)} [x^*(x) - f(x)].$$

Then,

(i) f^* *is a convex and weak* lower semicontinuous function on* $\mathrm{Dom}(f^*)$*;*

(ii) f^* *is continuous on* $\mathrm{int}(\mathrm{Dom}(f^*))$*;*

(iii) *if* $x^* \in \mathrm{Dom}(f^*)$ *and* $x \in \mathrm{argmax}(x^* - f)$ *then* $\widehat{x} \in \partial f^*(x^*)$*;*

(iv) *if* $\varepsilon > 0$*,* $x^* \in \mathrm{Dom}(f^*)$*,* $x \in X$ *and* $f^*(x^*) - \varepsilon < x^*(x) - f(x)$ *then* $\widehat{x} \in \partial_\varepsilon f^*(x^*)$*;*

(v) *if* $x_0^* \in \mathrm{int}(\mathrm{Dom}(f^*))$*,* $x \in \mathrm{argmax}(x_0^* - f)$ *and* $x^* \mapsto \partial f^*(x^*)$ *is single-valued and norm upper semicontinuous at* x_0^**, then* $x_0^* - f$ *has a strong maximum at* x*.*

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 $g_x : X^* \to \mathbb{R}$ by, $g_x(x^*) := \widehat{x}(x^*) - f(x)$. Then each function g_x is weak* continuous and affine. Now for each $x^* \in X^*$, $f^*(x^*) = \sup_{x \in \mathrm{Dom}(f)} g_x(x^*)$. Thus, as the pointwise supremum of a family of weak* continuous affine mappings, the Fenchel conjugate of f, is itself convex and weak* 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 $\mathrm{int}(\mathrm{Dom}(f^*)) = \varnothing$, we will assume that $\mathrm{int}(\mathrm{Dom}(f^*))$ is nonempty. Let us first recall that by Proposition 4.1, the fact that f^* is convex, and the fact that $\mathrm{int}(\mathrm{Dom}(f^*))$ is also convex, it is sufficient to show that f^* is locally bounded above on $\mathrm{int}(\mathrm{Dom}(f^*))$. In fact, as we shall now show, it is sufficient to show that f^* is locally bounded above at a single point $x_0^* \in \mathrm{int}(\mathrm{Dom}(f^*))$. To this end, suppose that f^* is locally bounded above at $x_0^* \in \mathrm{int}(\mathrm{Dom}(f^*))$. Then there exist an $0 < M$ and a $0 < \delta$ such that $f^*(y^*) \leq M$ for all $y^* \in B[x_0^*, \delta]$. Let x^* be any point in $\mathrm{int}(\mathrm{Dom}(f^*))$. Since $\mathrm{int}(\mathrm{Dom}(f^*))$ is an open convex set, there exists a point $y^* \in \mathrm{int}(\mathrm{Dom}(f^*))$ and a $0 < \lambda < 1$ such that $x^* = \lambda y^* + (1 - \lambda)x_0^*$. Let $M^* := \max\{M, f^*(y^*)\}$ and note that

$$x^* \in B[x^*, (1 - \lambda)\delta] = \lambda y^* + (1 - \lambda)B[x_0^*, \delta] \subseteq \mathrm{int}(\mathrm{Dom}(f^*)), \quad \text{since } \mathrm{int}(\mathrm{Dom}(f^*)) \text{ is convex.}$$

We claim that f^* 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 $w^* \in B[x_0^*, \delta]$ since,

$$B[x^*, (1 - \lambda)\delta] = x^* + (1 - \lambda)B[0, \delta] = \lambda y^* + (1 - \lambda)x_0^* + (1 - \lambda)B[0, \delta] = \lambda y^* + (1 - \lambda)B[x_0^*, \delta].$$

Therefore,

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

Next, we will use that fact that since $\mathrm{int}(\mathrm{Dom}(f^*))$ 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

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

39

Since f^* is weak* 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 $\text{int}(\text{Dom}(f^*))$. Since $\text{int}(\text{Dom}(f^*)) = \bigcup_{n \in \mathbb{N}} F_n$, there exists an $n_0 \in \mathbb{N}$ such that $\text{int}(F_{n_0}) \neq \varnothing$. Hence, f^* is locally bounded above at each point of $\text{int}(F_{n_0})$. This completes the proof of part (ii).

(iii) Let y^* be any element of $\text{Dom}(f^*)$. Then,

$$\widehat{x}(y^*) - \widehat{x}(x^*) = y^*(x) - x^*(x) = [y^*(x) - f(x)] - [x^*(x) - f(x)] = [y^*(x) - f(x)] - f^*(x^*) \leq f^*(y^*) - f^*(x^*).$$

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

(iv) Let y^* be any element of $\text{Dom}(f^*)$. Then,

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

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

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

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

We will show that $(x_n : n \in \mathbb{N})$ converges to x. Let $\varepsilon > 0$. By (iii) and the assumption that $\partial f^*(x_0^*)$ is a singleton we have that $\partial f^*(x_0^*) = \{\widehat{x}\}$. Since, $x^* \mapsto \partial f^*(x^*)$, is norm upper semicontinuous at x_0^* there exists a $0 < \delta < \varepsilon$ such that if $\|x^* - x_0^*\| \leq \delta$ then $\|F - \widehat{x}\| < \varepsilon$ for all $F \in \partial f^*(x^*)$. Choose $N \in \mathbb{N}$ such that $(x_0^* - f)(x_n) > f^*(x_0) - \delta^2$ for all $n > N$. Then, by (iv), $\widehat{x_n} \in \partial_{\delta^2} f^*(x_0^*)$ for all $n > N$. Let $n > N$. Then, by the Brøndsted-Rockafellar Theorem, there exist $x_n^* \in \text{Dom}(f^*)$ and $F_n \in X^{**}$ such that $F_n \in \partial f^*(x_n^*)$, $\|x_n^* - x_0^*\| \leq \delta$ and $\|F_n - \widehat{x_n}\| \leq \delta < \varepsilon$. Therefore,

$$\|x_n - x\| = \|\widehat{x_n} - \widehat{x}\| \leq \|\widehat{x_n} - F_n\| + \|F_n - \widehat{x}\| \leq \varepsilon + \varepsilon = 2\varepsilon.$$

This completes the proof. $\square$

Our first variational principle applies to dual differentiation spaces, [15]. Recall that a Banach space $(X, \|\cdot\|)$ is called a *dual differentiability space (or DD-space for short)* if every continuous convex function $\varphi : A \to \mathbb{R}$ defined on a nonempty open convex subset A of X^* such that $\{x^* \in A : \partial \varphi(x^*) \cap \widehat{X} \neq \varnothing\}$ contains a dense and G_δ subset of A, has the property that its subdifferential mapping $\partial \varphi : A \to 2^{X^{**}}$ is single-valued and norm upper semicontinuous at each point of a dense and G_δ subset of A (or equivalently, φ is Fréchet differentiable at the points of a dense and G_δ subset of A, [42, Proposition 2.8]).

Theorem 5.3. *Let $f : X \to \mathbb{R} \cup \{\infty\}$ be a proper function on a dual differentiation space $(X, \|\cdot\|)$. If there exists a nonempty open subset A of $\text{Dom}(f^*)$ and a dense and G_δ subset R of A such that $\text{argmax}(x^* - f) \neq \varnothing$ for each $x^* \in R$, 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 for each $x^* \in R'$. In addition, if $0 \in A$ and $\varepsilon > 0$ 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.*

Proof. Consider $\partial f^* : A \to 2^{X^{**}}$. Then by Proposition 5.2 part (iii)

$$R_1 := \{x^* \in A : \partial f^*(x^*) \cap \widehat{X} \neq \varnothing\}$$

contains a dense and G_δ subset of A. Since X is a dual differentiation space,

$$R_2 := \{x^* \in A : \partial f^* \text{ is single-valued and norm upper semicontinuous at } x^*\}$$

contains a dense and G_δ subset of A. Let $R' := R_1 \cap R_2$. Then R' contains a dense and G_δ subset of A and by Proposition 5.2 part (v), $(x^* - f)$ has a strong maximum for each $x^* \in R'$. $\square$

Remark 5.4. *There are two main weaknesses with this theorem: (i) although it is known that many Banach spaces (e.g. all spaces with the Radon-Nikodým property, [15], all weakly Lindelöf spaces, [29], all spaces that admit an equivalent locally uniformly rotund norm [16] and all spaces whose dual space X^* is weak Asplund, [15]) are dual differentiation spaces, it is still an open question as to whether every Banach space is a dual differentiation space; (ii) it is not clear how one would go about showing that there exists a "large" subset R of* $\mathrm{int}(\mathrm{Dom}(f^*))$ *with the property that* $\mathrm{argmax}(x^* - f) \neq \varnothing$ *for each* $x^* \in R$.

For our next result, and main variational principle, we address concern (i) of Remark 5.4 by giving a variational principle that holds in all Banach spaces. Unfortunately, there is a cost for this level of generality. Namely, we need to impose a strong assumption upon the mapping $x^* \mapsto \mathrm{argmax}(x^* - f)$. We also need to employ the following non-trivial result concerning minimal weak cuscos, which was first proved by J. Christensen in [6], using topological games (in the domain space), and later rephrased in [15].

Theorem 5.5. *A minimal weak* cusco $\varphi : A \to 2^{X^{**}}$ from a complete metric space A into subsets of the second dual X^{**} of a Banach space $(X, \|\cdot\|)$, where the set $\{x \in A : \varphi(x) \subseteq \widehat{X}\}$ contains a dense and G_δ subset of A, is single-valued and norm upper semicontinuous at the points of a dense and G_δ subset of A.*

In order to extend the applicability of Theorem 5.5, we will show that some sets that are not necessarily complete metric spaces under their given metrics can be "re-metrized" to become a complete metric space under a new metric, while retaining the same topology. Indeed, suppose that A is a nonempty open subset of a complete metric space (M, d). Then $(M \times \mathbb{R}, \rho)$ is also a complete metric space under the metric,

$$\rho((x_1, r_1), (x_2, r_2)) := d(x_1, x_2) + |r_1 - r_2|.$$

Let $f : A \to \mathbb{R}$ be defined by, $f(x) := \inf\{d(x, y) \in \mathbb{R} : y \in M \setminus A\} = \mathrm{dist}(x, M \setminus A)$. Note that f is continuous on A. Let $G := \{(x, r) \in M \times \mathbb{R} : x \in A$ and $r = 1/f(x)\}$. Then G is a closed subset of $M \times \mathbb{R}$, and hence is a complete metric space with respect to the restriction of the metric ρ to G. Finally, let us note that G is homeomorphic to A. Indeed, the mapping $\pi : G \to A$ defined by, $\pi(x, r) := x$, is such a homeomorphism. Thus, a nonempty open subset of a complete metric space is "completely metrisable".

Theorem 5.6. *Let $f : X \to \mathbb{R} \cup \{\infty\}$ be a proper function on a Banach space $(X, \|\cdot\|)$. If there exists a nonempty open subset A of $\mathrm{Dom}(f^*)$ such that $\mathrm{argmax}(x^* - f) \neq \varnothing$ 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 for each $x^* \in R'$. In addition, if $0 \in A$ and $\varepsilon > 0$ 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.*

Proof. Consider $\partial f^* : A \to 2^{X^{**}}$. Then, by Proposition 5.2 part (iii), $\partial f^*(x^*) \cap \widehat{X} \neq \varnothing$ for all $x^* \in A$. Thus, by Theorem 4.13, $\partial f^*(x^*) \subseteq \widehat{X}$ for all $x^* \in A$. Furthermore, by Proposition 4.5, $x^* \mapsto \partial f^*(x^*)$, is a minimal weak* cusco on A. Therefore, by Theorem 5.5, there exists a dense and G_δ subset R' of A such that ∂f^* is single-valued and norm upper semicontinuous at each point of R'. So, by Proposition 5.2 part (v), $(x^* - f)$ has a strong maximum for each $x^* \in R'$. $\quad\square$

Note that the conclusion of this theorem is identical to that of Stegall's variational principle, see [49].

Question 5.7. *Is every Banach space $(X, \|\cdot\|)$ a dual differentiation space?*

If the answer to this question is "yes" then Theorem 5.3 will supersede Theorem 5.6.

Index of notation and assumed knowledge

- The *natural numbers*, $\mathbb{N} := \{1, 2, 3, \ldots\}$.

- The *integers*, $\mathbb{Z} := \{\ldots, -2, -1, 0, 1, 2 \ldots\}$.

- The *rational numbers*, $\mathbb{Q} := \{a/b : a, b \in \mathbb{Z}, b \neq 0\}$.

- The *real numbers*, $\mathbb{R}$.

- For any set X, $\mathcal{P}(X)$ is the set of all subsets of X.

- For any subset A of a topological space (X, τ), we define

  - $\mathrm{int}(A)$, called the *interior* of A, is the union of all open sets contained in A;
  - $\overline{A}$, called the *closure* of A, is the intersection of all closed sets containing A;
  - $\mathrm{Bd}(A)$, called the *boundary* of A, is $\overline{A} \setminus \mathrm{int}(A)$,

- For any points x and y in a vector space X, we define the following intervals:

  - $[x, y] := \{x + \lambda(y - x) : 0 \le \lambda \le 1\}$;
  - $(x, y) := \{x + \lambda(y - x) : 0 < \lambda < 1\}$;
  - $[x, y) := \{x + \lambda(y - x) : 0 \le \lambda < 1\}$;
  - $(x, y] := \{x + \lambda(y - x) : 0 < \lambda \le 1\}$.

- For any normed linear space $(X, \|\cdot\|\cdot)$, we define

  - $B[x, r] := \{y \in X : \|x - y\| \le r\}$, for any $x \in X$ and $r > 0$;
  - $B(x; r) := \{y \in X : \|x - y\| < r\}$, for any $x \in X$ and $r > 0$;
  - $B_X := B[0, 1]$;
  - $S_X := \{x \in X : \|x\| = 1\}$.

- Given a compact Hausdorff space K, we write $C(K)$ for the set of all real-valued continuous functions on K. This is a vector space under the operations of pointwise addition and pointwise scalar multiplication. $C(K)$ becomes a Banach space when equipped with the uniform norm $\|\cdot\|_\infty$, defined by

$$\|f\|_\infty := \sup_{x \in K} |f(x)|, \quad \text{for all } f \in C(K).$$

- Let A and B be sets. Given a function $f : A \to B$, we define $f(A) := \bigcup_{a \in A}\{f(x)\}$. Similarly, given a set valued mapping $\varphi : A \to \mathcal{P}(B)$, we define $\varphi(A) := \bigcup_{a \in A} \varphi(x)$.

- For a normed linear space $(X, \|\cdot\|\cdot)$, X^*, the set of bounded linear maps from X to $\mathbb{R}$, is called the *dual space* of X. X^* is a Banach space when equipped with the operator norm, given by

$$\|f\| := \sup_{x \in B_X} |f(x)|, \quad \text{for all } f \in X^*.$$

- Let X be a set and Y a totally ordered set. For any function $f : X \to Y$ we define

$$\mathrm{argmax}(f) := \{x \in X : f(y) \le f(x) \text{ for all } y \in X\},$$
$$\mathrm{argmin}(f) := \{x \in X : f(x) \le f(y) \text{ for all } y \in X\}.$$

- Let A be a subset of a vector space X. Then the *convex hull* of A, denoted by $\mathrm{co}(A)$, is defined to be the intersection of all convex subsets of X that contain A.

- Let X be a set and let $f : X \to \mathbb{R} \cup \{\infty\}$ a function. Then

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

We say that the function f is a *proper function* if $\mathrm{Dom}(f) \ne \varnothing$.

- Let $(X, \|\cdot\|)$ be a normed linear space and $f : X \to [-\infty, \infty]$. Then the *Fenchel conjugate* of f is the function $f^* : X^* \to [-\infty, \infty]$ defined by,

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

The function f^* is convex and if f is a proper function then f^* never takes the value $-\infty$.

- If f is a convex function defined on a nonempty convex subset K of a normed linear space $(X, \|\cdot\|\cdot)$ and $x \in K$, then we define the *subdifferential of f at x* to be the set $\boldsymbol{\partial f(x)}$ of all $x^* \in X^*$ satisfying

$$x^*(y - x) \leq f(y) - f(x) \quad \text{for all } y \in K.$$

- It is assumed that the reader has a basic working knowledge of metric spaces, normed linear spaces and even basic general topology. In particular, it is assumed that the reader is familiar with Tychonoff's theorem.

  Theorem (Tychonoff's Theorem [10]). *The Cartesian product $\prod_{s \in S} X_s$, where $X_s \neq \varnothing$ for all $s \in S$, is compact if, and only if, all spaces X_s are compact.*

References

[1] L. Alaoglu, *Weak topologies of normed linear spaces*, Ann. of Math. (2) **41** (1940), 252–267.

[2] J. M. Borwein and A. S. Lewis, *Convex analysis and nonlinear optimization*, second ed., CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC, vol. 3, Springer, New York, 2006, Theory and examples. MR 2184742

[3] A. Brøndsted and R. T. Rockafellar, *On the subdifferentiability of convex functions*, Proc. Amer. Math. Soc. **16** (1965), 605–611.

[4] B. Cascales, J. Orihuela, and A. Pérez, *One-sided James' compactness theorem*, J. Math. Anal. Appl. **445** (2017), no. 2, 1267–1283.

[5] B. Cascales, J. Orihuela, and M. Ruiz Galán, *Compactness, optimality, and risk*, Computational and analytical mathematics, Springer Proc. Math. Stat., vol. 50, Springer, New York, 2013, pp. 161–218. MR 3108428

[6] 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.

[7] G. Debs, G. Godefroy, and J. Saint-Raymond, *Topological properties of the set of norm-attaining linear functionals*, Canad. J. Math. **47** (1995), no. 2, 318–329. MR 1335081

[8] R. Deville, G. Godefroy, and V. Zizler, *Smoothness and renormings in Banach spaces*, Pitman Monographs and Surveys in Pure and Applied Mathematics, vol. 64, Longman Scientific & Technical, Harlow; copublished in the United States with John Wiley & Sons, Inc., New York, 1993. MR 1211634

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

[10] R. Engelking, *General topology*, Sigma Series in Pure Mathematics, vol. 6, Heldermann Verlag, Berlin, 1989.

[11] M. Fabian, P. Habala, P. Hájek, V. M. Santalucia, J. Pelant, and V. Zizler, *Functional analysis and infinite-dimensional geometry*, CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC, vol. 8, Springer-Verlag, New York, 2001. MR 1831176

[12] 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

[13] V. P. Fonf and J. Lindenstrauss, *Boundaries and generation of convex sets*, Israel J. Math. **136** (2003), 157–172.

[14] V. P. Fonf, J. Lindenstrauss, and R. R. Phelps, *Infinite dimensional convexity*, Handbook of the geometry of Banach spaces, Vol. I, North-Holland, Amsterdam, 2001, pp. 599–670.

[15] 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.

[16] J. R. Giles and W. B. Moors, *A continuity property related to Kuratowski's index of noncompactness, its relevance to the drop property, and its implications for differentiability theory*, J. Math. Anal. Appl. **178** (1993), no. 1, 247–268. MR 1231740

[17] G. Godefroy, *Boundaries of a convex set and interpolation sets*, Math. Ann. **277** (1987), no. 2, 173–184. MR 886417

[18] J. Hagler and F. Sullivan, *Smoothness and weak* sequential compactness*, Proc. Amer. Math. Soc. **78** (1980), no. 4, 497–503. MR 556620

[19] R. C. James, *Reflexivity and the supremum of linear functionals*, Ann. of Math. (2) **66** (1957), 159–169.

[20] _____, *Characterizations of reflexivity*, Studia Math. **23** (1963/1964), 205–216.

[21] _____, *Weakly compact sets*, Trans. Amer. Math. Soc. **113** (1964), 129–140.

[22] _____, *A counterexample for a* sup *theorem in normed spaces*, Israel J. Math. **9** (1971), 511–512. MR 0279565

[23] _____, *Reflexivity and the sup of linear functionals*, Israel J. Math. **13** (1972), 289–300 (1973). MR 0338742

[24] M. Jiménez Sevilla and J. P. Moreno, *A note on norm attaining functionals*, Proc. Amer. Math. Soc. **126** (1998), no. 7, 1989–1997.

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

[26] O. F. K. Kalenda, *(I)-envelopes of closed convex sets in Banach spaces*, Israel J. Math. **162** (2007), 157–181. MR 2365858

[27] _____, *(I)-envelopes of unit balls and James' characterization of reflexivity*, Studia Math. **182** (2007), no. 1, 29–40. MR 2326490

[28] J. L. Kelley, *General topology*, D. Van Nostrand Company, Inc., Toronto-New York-London, 1955. MR 0070144

[29] P. S. Kenderov and W. B. Moors, *Separate continuity, joint continuity and the Lindelöf property*, Proc. Amer. Math. Soc. **134** (2006), no. 5, 1503–1512.

[30] 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.

[31] D. G. Larman and R. R. Phelps, *Gâteaux differentiability of convex functions on Banach spaces*, J. London Math. Soc. (2) **20** (1979), no. 1, 115–127. MR 545208

[32] S. Mazur, *Über schwache Konvergentz in den Räumen (L_p) (in German)*, Studia Math. **4** (1933), 128–133.

[33] W. B. Moors, *A selection theorem for weak upper semi-continuous set-valued mappings*, Bull. Austral. Math. Soc. **53** (1996), no. 2, 213–227.

[34] _____, *An elementary proof of James' characterization of weak compactness*, Bull. Aust. Math. Soc. **84** (2011), no. 1, 98–102.

[35] _____, *On a one-sided James' theorem*, J. Math. Anal. Appl. **449** (2017), no. 1, 528–530. MR 3595216

[36] _____, *Weak compactness of sublevel sets*, Proc. Amer. Math. Soc. **145** (2017), 3377–3379.

[37] 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.

[38] M. Morillon, *A new proof of James' sup theorem*, Extracta Math. **20** (2005), no. 3, 261–271. MR 2243342

[39] J. Orihuela and M. Ruiz Galán, *A coercive James's weak compactness theorem and nonlinear variational problems*, Nonlinear Anal. **75** (2012), no. 2, 598–611.

[40] ———, *Lebesgue property for convex risk measures on Orlicz spaces*, Math. Financ. Econ. **6** (2012), no. 1, 15–35.

[41] B. J. Pettis, *A proof that every uniformly convex space is reflexive*, Duke Math. J. **5** (1939), no. 2, 249–253.

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

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

[44] J. D. Pryce, *Weak compactness in locally convex spaces*, Proc. Amer. Math. Soc. **17** (1966), 148–155.

[45] J. Rainwater, *Weak convergence of bounded sequences*, Proc. Amer. Math. Soc. **14** (1963), 999. MR 0155171

[46] R. T. Rockafellar, *Convex analysis*, Princeton Landmarks in Mathematics, Princeton University Press, Princeton, NJ, 1997, Reprint of the 1970 original, Princeton Paperbacks. MR 1451876

[47] J. Saint Raymond, *Weak compactness and variational characterization of the convexity*, Mediterr. J. Math. **10** (2013), no. 2, 927–940.

[48] S. Simons, *A convergence theorem with boundary*, Pacific J. Math. **40** (1972), 703–708. MR 0312193

[49] C. Stegall, *Optimization of functions on certain subsets of Banach spaces*, Math. Ann. **236** (1978), no. 2, 171–176.

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.