TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 327, Number 2, October 1991

CONVEX OPTIMIZATION AND THE EPI-DISTANCE TOPOLOGY

GERALD BEER AND ROBERTO LUCCHETTI

ABSTRACT. Let $\Gamma(X)$ denote the proper, lower semicontinuous, convex functions on a Banach space X, equipped with the completely metrizable topology τ of uniform convergence of distance functions on bounded sets. A function f in $\Gamma(X)$ is called well-posed provided it has a unique minimizer, and each minimizing sequence converges to this minimizer. We show that well-posedness of $f \in \Gamma(X)$ is the minimal condition that guarantees strong convergence of approximate minima of τ-approximating functions to the minimum of f. Moreover, we show that most functions in $\langle \Gamma(X), \tau_{aw} \rangle$ are well-posed, and that this fails if $\Gamma(X)$ is topologized by the weaker topology of Mosco convergence, whenever X is infinite dimensional. Applications to metric projections are also given, including a fundamental characterization of approximative compactness.

1. INTRODUCTION

Let $\mathscr{C}(X)$ (resp. $\mathscr{C}_B(X)$) be the closed (resp. closed and bounded) nonempty convex sets in a normed linear space X. For over a half-century the basic topology on $\mathscr{C}_B(X)$ has been the well-known Hausdorff metric topology [15]. How should this topology be extended to $\mathscr{C}(X)$? The generally recognized [2, 37] successful solution in finite dimensions is the completely metrizable *Fell topology*, generated by all sets of the form $V^- \equiv \{A \in \mathscr{C}(X): A \cap V \neq \varnothing\}$ where V is open in X, and $(K^c)^+ \equiv \{A \in \mathscr{C}(X): A \subset K^c\}$ where K is a compact subset of X. Convergence of a sequence $\langle A_n \rangle$ to A in this topology in finite dimensions is equivalent to classical Kuratowski convergence of sets [27, §29]; alternatively, it is equivalent to the pointwise convergence of the associated sequence of distance functions $\langle d(\cdot, A_n)\rangle$ to $d(\cdot, A)$ (see, e.g., [11, 20]).

Certainly one of the most important features of the Fell topology on $\mathscr{C}(X)$ is its stability with respect to duality, as established by Wijsman [41], expressed by the continuity of the polar map $A \to A^\circ$, or in the case of proper lower

Received by the editors April 20, 1989 and, in revised form, September 5, 1989.
1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 49B50, 26B25; Secondary 54B20, 41A50.
Key words and phrases. Convex optimization, convex function, well-posed minimization problem, epi-distance topology, Mosco convergence, metric projection, approximative compactness, Chebyshev set.
Research of the second author was partially supported by C.N.R.

0002-9947/91 \$1.00 + \$.25 per page

semicontinuous convex functions as identified with their epigraphs, by the continuity of the conjugate map $f \to f^*$. Convergence of functions in this sense is called *epiconvergence* in the literature.

Intense efforts over the past twenty years have been focused on extending the basic results about the Fell topology for convex sets to infinite dimensions, requiring the introduction of topologies/convergence notions on $\mathscr{C}(X)$ which, in finite dimensions, reduce to convergence with respect to the Fell topology. Most prominent in this endeavor were the pioneering papers of Mosco [32, 33], where the definition of Kuratowski convergence for sequences of convex sets was modified in a most natural way: a sequence $\langle A_n \rangle$ in $\mathscr{C}(X)$ is declared *Mosco convergent* to $A \in \mathscr{C}(X)$ provided both of the following conditions are met:

(i) for each $a \in A$, there exists a sequence $\langle a_n \rangle$ strongly convergent to a such that for each n, we have $a_n \in A_n$;

(ii) whenever $n(1) < n(2) < \cdots$, and whenever $a_k \in A_{n(k)}$ for each $k \in Z^+$, then the weak convergence of $\langle a_k \rangle$ to $x \in X$ implies $x \in A$.

Mosco successfully extended Wijsman's sequential continuity results to the reflexive setting [33] (see also [8, 25]), but Mosco convergence is not stable with respect to duality in an arbitrary Banach space [10]. Moreover, Mosco convergence does not reduce to Hausdorff metric convergence on $\mathscr{C}_B(X)$ (in l_2, let A_n be the line segment joining the origin to e_n). The first tractable topology on $\mathscr{C}(X)$ stable with respect to duality [10, 16] in a general normed linear space that reduces to the Fell topology in finite dimensions has been seriously studied only recently [3, 4, 5, 6, 9, 10, 16]. Convergence of a sequence $\langle A_n \rangle$ to A in this topology means nothing more than uniform convergence of $\langle d(\cdot, A_n) \rangle$ to $d(\cdot, A)$ on bounded subsets of X. Uniform convergence of distance functions on bounded sets is compatible with a completely metrizable topology, provided X is complete (see §2 below). Although this convergence notion can be found in disguise in [32], we denote it by τ_{aw} in the sequel, in recognition of its development by Attouch and Wets for spaces of functions. Following these authors, we call τ_{aw} when restricted to the proper, lower semicontinuous, convex functions $\Gamma(X)$ on a normed linear space X the *epi-distance topology*.

In this paper we study convex minimization problems and the topology τ_{aw}. In many such problems, one is forced to approximate a given objective function $f \in \Gamma(X)$ by more tractable perturbed functions $\langle f_n \rangle$. Ideally, one would hope that minima (or approximate minima) of the perturbed sequence would converge to a minimum of f, and that $\inf_X f = \lim_{n\to\infty} \inf_X f_n$, as $\langle f_n \rangle$ converges to f. We show that when convergence means τ_{aw}-convergence, this behavior is guaranteed if and only if f is *well-posed* [40, 21, 29]: f has a unique minimizer x_0, and each minimizing sequence for f converges to x_0. Furthermore, we show that $\langle \Gamma(X), \tau_{aw} \rangle$ is completely metrizable whenever X is a Banach space, and that most convex functions (in the sense of Baire category) are well-posed. All of the above results fail for the topology of Mosco convergence. Finally, we obtain some basic results about metric projections, showing that in any reflexive space, most elements of $\langle\mathscr{C}(X), \tau_{aw}\rangle$ are Chebyshev, and that for most $(x, A) \in X \times \mathscr{C}(X)$, the associated metric projection program is well-posed.

2. Preliminaries

In the sequel we denote the unit ball and origin of our normed linear space X by U and θ, respectively. X^* will represent the continuous dual of X, with dual unit ball U^* and origin θ^*. We equip $X \times R$ with the box norm: $\|(x, \alpha)\| = \max\{\|x\|, |\alpha|\}$. If $f: X \to [-\infty, \infty]$ is any function, its *epigraph* is the following subset of $X \times R$: epi $f = \{(x, \alpha): x \in X, \alpha \in R$, and $\alpha \geq f(x)\}$. If epi $f \neq \varnothing$ and contains no vertical lines, we call f *proper*; f is *convex* (respectively *lower semicontinuous*) provided epi f is a convex (resp. closed) subset of $X \times R$. We write $v(f)$ for $\inf\{f(x): x \in X\}$, and $\arg\min f$ for the possibly empty set of points $\{x \in X: f(x) = v(f)\}$. For each $\alpha \in R$, we denote by $\operatorname{lev}(f; \alpha)$ the *sublevel set* of f at height α, that is, $\{x \in X: f(x) \leq \alpha\}$. Convexity (resp. lower semicontinuity) of f guarantees that each sublevel set is convex (resp. closed).

Again, $\Gamma(X)$ will denote the proper, lower semicontinuous, convex functions on X. If $A \in \mathscr{C}(X)$ we denote the *distance function* for A by $d(\cdot, A)$. It is standard to identify $A \in \mathscr{C}(X)$ with its *indicator function* $I(\cdot, A) \in \Gamma(X)$, defined by

$$I(x, A) = \begin{cases} 0, & \text{if } x \in A, \\ \infty, & \text{if } x \notin A. \end{cases}$$

For each $f \in \Gamma(X)$ its *conjugate* $f^* \in \Gamma(X^*)$ is defined by the familiar formula $f^*(y) = \sup\{\langle x, y\rangle - f(x): x \in X\}$. All of the above terminology is standard (see, e.g., [15, 23, 24, 36]).

We now turn to a discussion of τ_{aw} and the topology of Mosco convergence τ_M on $\mathscr{C}(X)$. These in turn give rise to topologies on $\Gamma(X)$, with functions identified with their epigraphs. On $\mathscr{C}(X)$, uniform convergence of distance functions on bounded subsets of X is formally convergence with respect to a topology on $\mathscr{C}(X)$, induced by a uniformity on $\mathscr{C}(X)$ with a (countable) base consisting of the following sets:

$$\Omega[n] = \left\{(A, B): \sup_{\|x\| \leq n} |d(x, A) - d(x, B)| < 1/n\right\} \qquad (n \in Z^+).$$

Denoting $\sup_{\|x\| \leq n} |d(x, A) - d(x, B)|$ by $d_n(A, B)$, it is clear that the induced topology is defined by the following metric on $\mathscr{C}(X)$:

$$m(A, B) = \sum_{n=1}^{\infty} 2^{-n}[d_n(A, B)/(1 + d_n(A, B))].$$

By Theorem 2.1 of [3], this metric is complete if X is a Banach space. We remark that in finite dimensions, uniform convergence of distance functions on bounded subsets is no stronger than their pointwise convergence, because distance functions are equicontinuous.

Recall [15] that the *excess* of a set A over a set B is defined by the formula $e(A, B) = \inf\{\varepsilon > 0: A \subset B + \varepsilon U\}$, and that the *Hausdorff distance* between A and B is given by $\text{haus}(A, B) = \max\{e(A, B), e(B, A)\}$. Since the Hausdorff distance between two sets A and B is nothing but $\sup_{x \in X} |d(x, A) - d(x, B)|$, it is not surprising that τ_{aw} admits a presentation akin to the standard presentation of Hausdorff distance. For each $\rho > 0$, we define the *ρ-Hausdorff distance* [3, 4, 10] between A and B by the formula

$$\text{haus}_\rho(A, B) = \max\{e(A \cap \rho U, B), e(B \cap \rho U, A)\}.$$

Apparently, the connection between τ_{aw} and the "distances" $\{\text{haus}\,\rho: \rho > 0\}$ was first observed in [6] and in [9]: $A = \tau_{aw}\text{-}\lim A_n$ if and only if for each $\rho > 0$, we have $\lim_{n\to\infty} \text{haus}_\rho(A_n, A) = 0$. Actually, to show that $A = \tau_{aw}\text{-}\lim A_n$, we need only show that $\lim_{n\to\infty} \text{haus}_\rho(A_n, A) = 0$ for all ρ beyond some fixed ρ_0 because haus_ρ increases with ρ. We will use this fact repeatedly. From the perspective of uniformities, the connection between τ_{aw} and $\{\text{haus}\,\rho: \rho > 0\}$ may be recast as follows [3, Theorem 1.2]: a (countable) base for another (weaker!) compatible uniformity for τ_{aw} consists all sets of the form:

$$\Sigma[n] = \{(A, B): \text{haus}_n(A, B) < 1/n\} \qquad (n \in Z^+).$$

Given that Mosco convergence was introduced twenty years ago and has been of great interest thereafter [2, 39], it is inexplicable why a simple topology compatible with Mosco convergence in any Banach space [6, Theorem 3.1] was not identified until very recently. This *Mosco topology* τ_M has as a subbase all sets of the form:

$$V^- \equiv \{A \in \mathscr{C}(X): A \cap V \neq \varnothing\} \qquad (V \text{ norm open}),$$
$$(K^c)^+ \equiv \{A \in \mathscr{C}(X): A \subset K^c\} \qquad (K \text{ weakly compact}).$$

Like Mosco convergence, this topology is well-behaved only when X is reflexive. In this setting, it is Hausdorff and completely regular, but it is metrizable if and only if X is separable [7]. Evidently, this topology reduces to the Fell topology in finite dimensions. As noted earlier, the Fell topology and the topology of pointwise convergence of distance functions agree in finite dimensions; so, $\tau_M = \tau_{aw}$ here. The Mosco topology τ_M is in general weaker than τ_{aw} [4, Proposition 4.5 and 10, Lemma 2.1], and in reflexive spaces, stronger than the topology of pointwise convergence of distance functions [7, Theorem 3.5; 39, p. II. 6]. That τ_M-convergence need not guarantee τ_{aw}-convergence is easy to see: in l_2, again let $A_n = \text{conv}\{\theta, e_n\}$. That pointwise convergence of distance functions in reflexive spaces need not guarantee τ_M-convergence is not easy to see, and requires that the dual norm for X^* fails to have the Kadec property [9, 14].

It is very well known [2, 32] that τ_M-convergence of a sequence $\langle f_n \rangle$ in $\Gamma(X)$ to $f \in \Gamma(X)$ is equivalent to the conjunction of the following two conditions:

(a) for each $x \in X$, there exists a sequence $\langle x_n \rangle$ strongly convergent to x for which $\lim_{n\to\infty} f_n(x_n) = f(x)$;

(b) for each $x \in X$, whenever $\langle x_n \rangle$ is weakly convergent to x, we have $f(x) \le \liminf_{n\to\infty} f_n(x_n)$.

For reflexive spaces, the conjugate map $f \to f^*$ is a homeomorphism of $\langle \Gamma(X), \tau_M \rangle$ onto $\langle \Gamma(X^*), \tau_M \rangle$ [8]; this is also true if the function spaces are equipped with the stronger epi-distance topology [4], even without reflexivity [10, 16]. More fundamentally, τ_M is the weakest topology on $\Gamma(X)$ such that the epigraphical multifunctions $f \to \operatorname{epi} f$ and $f \to \operatorname{epi} f^*$ are both lower semicontinuous [27, §18] as set-valued functions [12].

3. Some tool theorems for τ_{aw}-convergence

In this section we collect some basic facts about τ_{aw}-convergence of sets and functions that have not appeared in the literature. We will apply all of them in subsequent sections. In the process, we point out the favorable properties of this convergence that distinguish it from the less well-behaved Mosco convergence.

Evidently, a uniformly bounded sequence in $\mathscr{C}(X)$ is τ_{aw}-convergent if and only if it is convergent in Hausdorff distance. In fact, this equivalence holds assuming only that the limit set is bounded, a fact noted by Salinetti and Wets in finite dimensions [38]. As a result, the topology τ_{aw} reduces to the usually stronger Hausdorff metric topology when restricted to the bounded elements of $\mathscr{C}(X)$.

Lemma 3.1. *Let X be a normed linear space, and let $A \in \mathscr{C}(X)$ be bounded. Suppose $\langle A_n \rangle$ is a sequence in $\mathscr{C}(X)$ τ_{aw}-convergent to A. Then $\langle A_n \rangle$ is convergent to A in Hausdorff distance.*

Proof. Fix $a_0 \in A$ and choose $\rho > 0$ with $A \subset \rho U$. By the definition of τ_{aw}, there exists $N \in Z^+$ such that for each $n > N$ we have

$$\sup\{|d(x, A) - d(x, A_n)| : \|x\| \le \rho + 3\} < 1.$$

We claim that for all $n > N$, we have $A_n \subset (\rho + 2)U$. Suppose not; then there exists $a_n \in A_n$ with $\|a_n\| > \rho + 2$. Since $d(a_0, A_n) < 1$, there exists $b_n \in A_n$ with $\|b_n\| < \rho + 1$. As a result, some convex combination c_n of a_n and b_n has norm $\rho + 2$ so that $\|d(c_n, A_n) - d(c_n, A)\| = d(c_n, A) \ge 2$, a contradiction to the choice of N. Thus, $\langle A_n \rangle$ is uniformly bounded eventually, and converges to A in Hausdorff distance. □

We remark in passing that Lemma 3.1 remains valid for sequences of connected sets but fails in general. As an immediate corollary we have

Lemma 3.2. *Let X be a normed linear space; then $A \to \operatorname{diam} A$ is a continuous extended real valued functional on $\langle \mathscr{C}(X), \tau_{aw} \rangle$.*

Proof. Fix $A_0 \in \mathscr{C}(X)$ and let $\langle A_n \rangle$ be a sequence in $\mathscr{C}(X)$ τ_{aw}-convergent to A_0. We consider two cases: (i) $\operatorname{diam} A_0 = \infty$; (ii) $\operatorname{diam} A_0 < \infty$.

In case (i), upper semicontinuity of the diameter functional obviously holds at A_0. For lower semicontinuity, we show that for each $\alpha > 0$, there exists $N \in Z^+$ such that for each $n > N$ we have $\operatorname{diam} A_n > \alpha$. Choose points a and b in A_0 with $\|a-b\| > \alpha$. Let $\varepsilon = (\|a-b\|-\alpha)/2$ and let $\rho = \max\{\|a\|, \|b\|\}$. There exists $N \in Z^+$ with $\operatorname{haus}_\rho(A_n, A) < \varepsilon$ for each $n > N$. For each such n, there exist a_n and b_n in A_n with $\|a_n - a\| < \varepsilon$ and $\|b_n - b\| < \varepsilon$, and as a result, $\operatorname{diam} A_n > \alpha$. In case (ii), by Lemma 3.1, we have convergence of $\langle A_n \rangle$ to A_0 in Hausdorff distance, where continuity of the diameter functional is well known. □

Consideration of the example in l_2 presented in the introduction shows that Lemma 3.2 fails with τ_M-convergence replacing τ_{aw}-convergence. More precisely, on $\langle \mathscr{C}(X), \tau_M \rangle$, $A \to \operatorname{diam} A$ is only lower semicontinuous.

We use the next fact twice in the sequel.

Lemma 3.3. *Let $\langle A_n \rangle$ be an increasing sequence of closed (convex) subsets of a normed linear space X such that for each $\rho > 0$ there exists n with $\rho U \subset A_n$. Then for each $A \in \mathscr{C}(X)$, we have $A = \tau_{aw}\text{-}\lim A_n \cap A$.*

Proof. For each $\rho > 0$, there exists $N \in Z^+$ such that for all $n > N$, we have $A \cap \rho U = A_n \cap A \cap \rho U$. Thus, $\operatorname{haus}_\rho(A, A_n \cap A) = 0$ for each $n > N$. □

The next lemma may be found buried in the proof of Theorem 11 of [31]. Although it is very simple, it is, in our view, an important technical feature of τ_{aw}-convergence. It is an immediate consequence of the following version of the *Radström cancellation principle* [6], which itself follows easily from the separation theorem.

Radström cancellation principle. *Let A, B, and C be closed convex subsets of a normed linear space X with B bounded. Suppose $A + B \subset C + B$. Then $A \subset C$.*

Lemma 3.4. *Let X be a normed linear space. Suppose $B \in \mathscr{C}(X)$, $C \in \mathscr{C}(X)$, and* $\operatorname{haus}_\rho(C, B) < \delta$. *Then whenever $x+2\delta U \subset B \cap \rho U$, we have $x + \delta U \subset C$.*

Proof. We have the following inclusions:

$$(x + \delta U) + \delta U = x + 2\delta U \subset B \cap \rho U \subset C + \delta U.$$

By the Radström cancellation principle, this yields $x + \delta U \subset C$. □

An immediate consequence of Lemma 3.4 is that $\{A \in \mathscr{C}(X)\colon \operatorname{int} A \neq \varnothing\}$ is τ_{aw}-open. This fails for the Mosco topology.

Example. In l_2, the unit ball is the τ_M-limit of $\langle A_n \rangle$ where for each n

$$A_n = \left\{ x\colon \sum_{i=1}^{n} \langle x, e_i \rangle^2 \leq 1 \text{ and } \langle x, e_i \rangle = 0 \text{ for } i > n \right\}.$$

Evidently, each A_n has empty interior.

We now turn our attention to $\Gamma(X)$.

Lemma 3.5. *Let X be a Banach space. Then $\langle \Gamma(X), \tau_{aw}\rangle$ is completely metrizable.*

Proof. As mentioned in §2, the space $\langle \mathscr{C}(X\times R), \tau_{aw}\rangle$ is completely metrizable. Evidently, the lower semicontinuous convex functions on X (other than the function that is identically equal to ∞) may be described as the following set: $\mathscr{F} \equiv \{A \in \mathscr{C}(X \times R)\colon$ whenever $(x, \alpha) \in A$ and $\beta > \alpha$, then $(x, \beta) \in A\}$. It is routine to check that $\mathscr{F}$ is τ_{aw}-closed in $\mathscr{C}(X \times R)$, so that $\langle \mathscr{F}, \tau_{aw}\rangle$ is completely metrizable. Since a closed convex set in $X \times R$ that contains a vertical line must actually be a product of some closed subset of X with R, it is clear that $\Gamma(X) = \{f \in \mathscr{F}\colon$ there exists some $x \in X$ with $f(x)$ finite$\}$. We show $\Gamma(X)$ is an open subset of $\langle \mathscr{F}, \tau_{aw}\rangle$, yielding its complete metrizability by Alexandrov's Theorem [27, §33].

Fix $f_0 \in \Gamma(X)$ and $x_0 \in X$ with $f_0(x_0)$ finite. By lower semicontinuity of f_0 there exist $\delta \in (0, 1)$ such that for each $x \in x_0 + 3\delta U$ we have $f_0(x) > f_0(x_0) - 1$. Consider the following bounded subset B of $X \times R$:

$$B = (x_0 + \delta U) \times [f_0(x_0) - 4, f_0(x_0) - 2].$$

Notice that if $(x, \alpha) \in B$ then $d((x, \alpha), \operatorname{epi} f_0) > \delta$. Let $C = B \cup \{(x_0, f(x_0))\}$. Then

$$\mathscr{A} = \left\{ f \in \mathscr{F}\colon \sup_{(x,\alpha)\in C} |d((x, \alpha), \operatorname{epi} f) - d((x, \alpha), \operatorname{epi} f_0)| < \delta \right\}$$

is a τ_{aw}-neighborhood of f_0 in $\mathscr{F}$. We show that $\mathscr{A} \subset \Gamma(X)$. Fix $f \in \mathscr{A}$. Since $d((x_0, f(x_0)), \operatorname{epi} f_0) = 0$ we must have $d((x_0, f(x_0)), \operatorname{epi} f) < \delta$. This means that there exists $(x, \alpha) \in \operatorname{epi} f$ with $x \in x_0 + \delta U$ and $|\alpha - f(x_0)| < 1$. Now $(x, \alpha - 3) \in B$ so that

$$d((x, \alpha - 3), \operatorname{epi} f) > d((x, \alpha - 3), \operatorname{epi} f_0) - \delta > 0.$$

Thus $(x, \alpha-3) \notin \operatorname{epi} f$, and we have $\alpha-3 < f(x) \leq \alpha$. Thus f is somewhere finite and $f \in \Gamma(X)$. □

Similar arguments show that $\langle \Gamma(X), \tau_M\rangle$ is completely metrizable when X is reflexive and separable, for the function space will also be a G_δ subset of the completely metrizable space $\langle \mathscr{C}(X \times R), \tau_M\rangle$ [7, Theorem 4.3]. First, the set $\mathscr{F} = \{A \in \mathscr{C}(X \times R)\colon$ whenever $(x, \alpha) \in A$ and $\beta > \alpha$, then $(x, \beta) \in A\}$ is again easily shown to be τ_M-closed in $\mathscr{C}(X \times R)$. Second, $\Gamma(X)$ is again an open subset of $\mathscr{F}$. To see this, fix $f_0 \in \Gamma(X)$ and $x_0 \in X$ with $f_0(x_0)$ finite. There exists $\alpha \in R$ with $\inf_{\|x-x_0\|\leq 1} f(x) > \alpha$ because $x_0 + U$ is weakly compact and f is weakly lower semicontinuous. Let V be a norm open subset of $X \times R$ of diameter less than 1 containing $(x_0, f_0(x_0))$ such that V lies above $(x_0 + U) \times \{\alpha\}$. Evidently,

$$\mathscr{D} \equiv V^- \cap [((x_0 + U) \times \{\alpha\})^c]^+ \cap \mathscr{F}$$

is a τ_M-neighborhood of f_0 relative to $\mathscr{F}$, and if $f \in \mathscr{D}$, then there exists $x \in x_0 + U$ with $\alpha < f(x) < f_0(x_0) + 1$. Thus, $\mathscr{D} \subset \Gamma(X)$.

We anticipate that our next theorem will have numerous applications. An analagous result does hold for Mosco convergence, and the proof is much simpler (see, e.g., [8, 33]).

Theorem 3.6. *Let X be a normed linear space. Suppose $\langle f_n \rangle$ is a sequence in $\Gamma(X)$ with $f = \tau_{aw}$-$\lim f_n$. Then for each $\alpha > v(f)$, we have* $\operatorname{lev}(f;\alpha) = \tau_{aw}$-$\lim \operatorname{lev}(f_n;\alpha)$.

Proof. Choose β strictly between α and $v(f)$, and choose $\rho_0 > |\alpha|$ such that for some $x_0 \in \operatorname{int} \rho_0 U$ we have $f(x_0) < \beta$. Since $[\operatorname{int} \rho_0 U \times (-\infty, \beta)]^-$ is a τ_M-neighborhood of f, it is a τ_{aw}-neighborhood of f. Thus, there exists $N_1 \in Z^+$ such that $n > N_1$ implies $f_n \in [\operatorname{int} \rho_0 U \times (-\infty, \beta)]^-$. Fix $\rho > \rho_0$ and $\varepsilon > 0$; we produce $N \in Z^+$ such that for each $n > N$, both of the following conditions hold:

(1) $\operatorname{lev}(f;\alpha) \cap \rho U \subset \operatorname{lev}(f_n;\alpha) + \varepsilon U$;
(2) $\operatorname{lev}(f_n;\alpha) \cap \rho U \subset \operatorname{lev}(f;\alpha) + \varepsilon U$.

Choose $\delta > 0$ such that $\delta + 2\delta\rho/(\alpha - \beta + \delta) < \varepsilon$. Pick $N_2 \in Z^+$ so large that whenever $n > N_2$, we have $\operatorname{haus}_\rho(\operatorname{epi} f, \operatorname{epi} f_n) < \delta$. We claim that the choice $N = N_1 + N_2$ works. We verify that condition (1) holds; verification of (2) is exactly the same and is left to the reader.

Fix $n > N$ and $x \in \operatorname{lev}(f;\alpha) \cap \rho U$. Since $\rho > |\alpha|$, we obtain $(x, \alpha) \in \rho U \times [-\rho, \rho]$. Since $n > N_2$, there exists $(w_n, \alpha_n) \in \operatorname{epi} f_n$ with $\|w_n - x\| < \delta$ and $|\alpha_n - \alpha| < \delta$. Since $n > N_1$ and $\rho > \rho_0$, there exists $z_n \in \rho U$ with $f_n(z_n) < \beta$. Let $\lambda = (\alpha - \beta)/(\alpha - \beta + \delta)$; we will show

(i) $\lambda w_n + (1-\lambda) z_n \in \operatorname{lev}(f_n;\alpha)$;
(ii) $\|\lambda w_n + (1-\lambda) z_n - x\| \le \varepsilon$.

Condition (i) follows easily from $f_n(w_n) \le \alpha_n < \alpha + \delta$ and $f_n(z_n) < \beta$:

$$\begin{aligned} f_n(\lambda w_n + (1-\lambda) z_n) &\le \lambda f_n(w_n) + (1-\lambda) f_n(z_n) \\ &< \lambda(\alpha + \delta) + (1-\lambda)\beta \\ &= \frac{(\alpha - \beta)(\alpha + \delta) + \delta\beta}{(\alpha - \beta + \delta)} = \alpha. \end{aligned}$$

Condition (ii) is immediate from the choice of δ and the fact that both z_n and x lie in ρU:

$$\begin{aligned} \|\lambda w_n + (1-\lambda) z_n - x\| &\le \lambda \|w_n - x\| + (1-\lambda)\|z_n - x\| \\ &< \|w_n - x\| + \left(\frac{\delta}{\alpha - \beta + \delta}\right)\|z_n - x\| \\ &< \delta + 2\delta\rho/(\alpha - \beta + \delta) < \varepsilon. \end{aligned}$$

This completes the proof of τ_{aw}-convergence of sublevel sets at a fixed height above $v(f)$. □

It is easy to see that τ_{aw}-convergence of $\langle f_n \rangle$ to f does not guarantee that $\mathrm{lev}(f; v(f)) = \tau_{aw}\text{-}\lim \mathrm{lev}(f_n; v(f))$: take $f_n(x) \equiv 1/n$ and $f(x) \equiv 0$. As a first application of Theorem 3.6, we have

Theorem 3.7. *Let X be a normed linear space, and let $f, f_1, f_2, f_3, \ldots$ be functions in $\Gamma(X)$ with $f = \tau_{aw}\text{-}\lim f_n$. Then*

(a) $v(f) \geq \limsup_{n\to\infty} v(f_n)$;

(b) *if for some $\alpha > v(f)$, $\mathrm{lev}(f; \alpha)$ is bounded, then $v(f) = \lim_{n\to\infty} v(f_n)$.*

Proof. (a) Upper semicontinuity of the value function holds for the weaker Mosco topology; in fact, it holds for the topology generated by all sets of the form $\{f \in \Gamma(X): \mathrm{epi}\, f \cap V \neq \varnothing\} = \Gamma(X) \cap V^-$ where V is norm open in $X \times R$ (see, e.g., [2, p. 128]).

(b) Clearly, $v(f) \leq \liminf_{n\to\infty} v(f_n)$ holds if $v(f) = -\infty$. Suppose now that $v(f)$ is finite and for some $\varepsilon > 0$, we have $v(f_n) < v(f) - 3\varepsilon$ for each n in some infinite subset J of Z^+. For each $n \in J$, we may choose $c_n \in X$ with $f_n(c_n) < v(f) - 2\varepsilon$. By Lemma 3.1 and Theorem 3.6, there exists $N_1 \in Z^+$ and $\rho > |v(f)| + 3\varepsilon$ such that $\mathrm{lev}(f_n; v(f) - 2\varepsilon) \subset \rho U$ for each $n > N_1$. Pick $N_2 > N_1$ such that for each $n > N_2$ we have $\mathrm{haus}_\rho(\mathrm{epi}\, f, \mathrm{epi}\, f_n) < \varepsilon$. Fix $n \in J$ with $n > N_2$. Now

$$(c_n, v(f) - 2\varepsilon) \in \mathrm{epi}\, f_n \cap \rho U \times [-\rho, \rho];$$

so, there exists $(x, \alpha) \in \mathrm{epi}\, f$ with $\|(x, \alpha) - (c_n, v(f) - 2\varepsilon)\| < \varepsilon$. This means that $f(x) \leq \alpha < v(f) - \varepsilon$, a contradiction. We conclude that in this case, we also have $v(f) \leq \liminf_{n\to\infty} v(f_n)$. $\square$

Example. Part (b) of Theorem 3.7 fails for Mosco convergence, even in l_2. To see this, let f be the indicator function for the origin (a well-posed function), and for each n, let $f_n: l_2 \to R$ be defined by

$$f_n(x) = \begin{cases} \max\{-\alpha/n, -1\}, & \text{if } x = \alpha e_n \text{ and } \alpha \geq 0; \\ \infty, & \text{otherwise.} \end{cases}$$

Obviously, $v(f) = 0$ and $v(f_n) = -1$ for each n. Using the fact that the origin is the weak limit of $\langle e_n \rangle$, it is easy to see that $f = \tau_M\text{-}\lim f_n$.

We refer the reader to [5] for a sharper version of Theorem 3.7(b). The converse of Theorem 3.6 fails: for any X and any $A \in \mathscr{C}(X)$, let $f_n = I(\cdot, A) - n$ and let $f = I(\cdot, A)$. We do, however, have a partial converse for Theorem 3.6, which holds without any convexity assumptions whatsoever.

Theorem 3.8. *Let X be a normed linear space and let $f, f_1, f_2, \ldots$ be (convex) functions on X satisfying $\mathrm{lev}(f; \alpha) = \tau_{aw}\text{-}\lim \mathrm{lev}(f_n; \alpha)$ for each $\alpha > v(f)$. Then if $v(f) \leq \liminf_{n\to\infty} v(f_n)$, we have $f = \tau_{aw}\text{-}\lim f_n$.*

Proof. Let ρ_0 be a positive scalar exceeding $v(f)$. Fix $\rho > \rho_0$ and $\varepsilon > 0$. We produce $N \in Z^+$ such that for each $n > N$, $\mathrm{haus}_\rho(f, f_n) \leq \varepsilon$.

Choose $k \in Z^+$ with $4\rho/k < \varepsilon$. Also, choose $N_1 \in Z^+$ such that for each $n > N_1$ we have $v(f_n) \geq v(f) - 2\rho/k$. Write $\alpha_j = -\rho + j(2\rho/k)$ for each $j \in \{1, 2, 3, \dots, k+1\}$, and let j_0 be the smallest index in $\{1, 2, 3, \dots, k\}$ with $\alpha_{j_0} > v(f)$ (j_0 exists because $\rho > \rho_0 > v(f)$). By our assumption on convergence of sublevel sets, there exists $N_2 \in Z^+$ such that $\text{haus}_\rho(\text{lev}(f; \alpha_j), \text{lev}(f_n; \alpha_j)) < \varepsilon$ for each $n > N_2$ and for each j with $j_0 \leq j \leq k+1$. Fix $n > N \equiv N_1 + N_2$. We verify both of the following:

(i) $e(\text{epi} f \cap \rho U \times [-\rho, \rho], \text{epi} f_n) \leq \varepsilon$, and
(ii) $e(\text{epi} f_n \cap \rho U \times [-\rho, \rho], \text{epi} f) \leq \varepsilon$.

The cases are not entirely symmetric, (ii) being the more subtle. For (i), fix $(x, \alpha) \in \text{epi} f \cap \rho U \times [-\rho, \rho]$. Let $j \in \{1, 2, \dots, k+1\}$ be the minimal index with $\alpha < \alpha_j$. Then $x \in \text{lev}(f; \alpha_j)$, and since $n > N_2$, there exists $z_n \in \text{lev}(f_n; \alpha_j)$ with $\|z_n - x\| < \varepsilon$. Since $\alpha_j - \varepsilon < \alpha_j - 2\rho/k \leq \alpha < \alpha_j$, we have $\|(z_n, \alpha_j) - (x, \alpha)\| < \varepsilon$, and (i) follows.

To establish (ii), fix $(x_n, \beta_n) \in \text{epi} f_n \cap \rho U \times [-\rho, \rho]$. In this case, let $j \in \{1, 2, \dots, k\}$ be the minimal index with $\beta_n \leq \alpha_j$. Since $n > N_1$ and $j \leq k$, we have

$$v(f) \leq v(f_n) + 2\rho/k \leq \beta_n + 2\rho/k \leq \alpha_j + 2\rho/k = \alpha_{j+1} \leq \alpha_{k+1}.$$

In particular, $j_0 \leq j+1 \leq k+1$. Since $x_n \in \text{lev}(f_n; \alpha_{j+1})$ and $n > N_2$, there exists $z \in \text{lev}(f; \alpha_{j+1})$ with $\|z - x_n\| < \varepsilon$. Then $(z, \alpha_{j+1}) \in \text{epi} f$, and

$$\begin{aligned}\|(z, \alpha_{j+1}) - (x_n, \beta_n)\| &= \max\{\|z - x_n\|, |\alpha_{j+1} - \beta_n|\} \\ &< \max\{\varepsilon, 4\rho/k\} = \varepsilon.\end{aligned}$$

This establishes (ii), completing the proof. □

We note that the condition $\liminf_{n\to\infty} v(f_n) \geq v(f)$ in Theorem 3.8 is really no weaker than the condition $v(f) = \lim_{n\to\infty} v(f_n)$, by virtue of τ_{aw}-convergence of sublevel sets above height $v(f)$.

4. Well-posedness of convex functions

The main results of this section depend on the following characterization of well-posedness, due to Furi and Vignoli: a proper lower semicontinuous function on a complete metric space is *well-posed* if and only if $\inf\{\text{diam}\,\text{lev}(f; \alpha) : \alpha > v(f)\} = 0$ [20]. This has been used frequently in more classical settings (see, e.g., [17, 29, 34]).

Theorem 4.1. *Let X be a Banach space. Then $f \in \Gamma(X)$ is well-posed if and only if whenever $f = \tau_{aw}\text{-}\lim f_n$ and $x_n \in \text{lev}(f_n; v(f_n) + 1/n)$ for each n, then $\langle x_n \rangle$ is convergent (to the unique minimizer of f).*

Proof. For sufficiency, simply take $f_n = f$ for each n. For necessity, let x_0 be the unique minimizer of f, let $\langle f_n \rangle$ and $\langle x_n \rangle$ be as described above, and let $\varepsilon > 0$. By the Furi-Vignoli characterization of well-posedness, there exists $\delta > 0$ such that $\text{diam}\,\text{lev}(f; v(f) + \delta) < \varepsilon/2$. By Lemmas 3.1 and Theorem

3.6, $\langle \mathrm{lev}(f_n; v(f)+\delta)\rangle$ is actually convergent to $\mathrm{lev}(f; v(f)+\delta)$ in Hausdorff distance. By upper semicontinuity of the value function there exists $N \in Z^+$ such that $1/N < \delta/2$ and for each $n > N$ both $v(f_n) < v(f) + \delta/2$ and $\mathrm{haus}(\mathrm{lev}(f_n; v(f)+\delta), \mathrm{lev}(f; v(f)+\delta)) < \varepsilon/2$. As a result, for each $n > N$ we have $\|x_n - x_0\| < \varepsilon$. □

Immediate consequences of Theorem 3.7 and Theorem 4.1 are these: if f is well-posed, then τ_{aw}-convergence of $\langle f_n\rangle$ to f drives minimizers of $\langle f_n\rangle$ (if they exist) into the unique minimizer of f, and $\lim_{n\to\infty} v(f_n) = v(f)$. Theorem 4.1 fails for Mosco convergence in l_2. In fact one can construct a sequence $\langle f_n\rangle$ in $\Gamma(l_2)$ Mosco convergent to $f(x) = \|x\|^2$ such that for each n, $e_n \in \arg\min f_n$ (see [28, Example 4.9]).

There is a generalization of Theorem 4.1 which we will record without proof. Recall that a proper lower semicontinuous function f is *well-posed in the generalized sense* (g.w.p.) [28] provided each minimizing sequence for f contains a convergent subsequence. Easily, it is seen that f is g.w.p. if and only if $\arg\min f$ is nonempty and compact, and

$$\inf_{\alpha > v(f)} e(\mathrm{lev}(f; \alpha), \arg\min f) = 0.$$

Following the proof of Theorem 4.1, we obtain

Theorem 4.2. *Let X be a Banach space. Then $f \in \Gamma(X)$ is well-posed in the generalized sense if and only if whenever $f = \tau_{aw}\text{-}\lim f_n$ and for each n, $x_n \in \mathrm{lev}(f_n; v(f_n)+\varepsilon_n)$ where $\langle \varepsilon_n\rangle \to 0$, then $\langle x_n\rangle$ has a convergent subsequence (to a minimizer of f).*

Recall that a G_δ subset of a topological space is one that can be expressed as an intersection of a countable family of open sets. Now let X be an arbitrary Banach space. We intend to show that the well-posed functions form a dense and G_δ subset of the function space $\langle \Gamma(X), \tau_{aw}\rangle$. This result was first obtained in finite dimensions by Lucchetti and Patrone [30]; for refinements, the reader may consult [13, 34]. By the Baire category theorem, a dense and G_δ subset of a complete metric space is most of the space; so, we will be entitled to say *most functions in $\langle \Gamma(X), \tau_{aw}\rangle$ are well-posed*, by virture of Lemma 3.5.

We obtain τ_{aw}-density of the well-posed problems in $\Gamma(X)$ rather routinely through successive approximations.

Lemma 4.3. *Let X be a normed linear space. Suppose $f \in \Gamma(X)$ and* $\arg\min f$ *$\neq \varnothing$. Then there exists a sequence $\langle f_n\rangle$ of well-posed functions in $\Gamma(X)$ τ_{aw}-convergent to f.*

Proof. Fix $x_0 \in \arg\min f$, and for each $n \in Z^+$, let $f_n \in \Gamma(X)$ be defined by

$$f_n(x) = f(x) + \|x - x_0\|/n.$$

Evidently, each f_n is well-posed (with unique minimum at x_0).

Choose ρ_0 with $\|x_0\| < \rho_0$. Fix $\rho > \rho_0$. Since $\mathrm{epi}\, f_n \subset \mathrm{epi}\, f$, we have $e(\mathrm{epi}\, f_n \cap \rho U \times [-\rho, \rho], \mathrm{epi}\, f) = 0$. On the other hand, if $(x, \alpha) \in \mathrm{epi}\, f$ where

$\|x\| \le \rho$ and $|\alpha| \le \rho$, then $\|x - x_0\| \le 2\rho$ so that $(x, \alpha + 2\rho/n) \in \operatorname{epi} f_n$. Thus, $e(\operatorname{epi} f \cap \rho U \times [-\rho, \rho], \operatorname{epi} f_n) \le 2\rho/n$. As a result, $\operatorname{haus}_\rho(\operatorname{epi} f, \operatorname{epi} f_n) \le 2\rho/n$, and $f = \tau_{aw}\text{-}\lim f_n$. □

Lemma 4.4. *Let X be a normed linear space. Then the well-posed functions are dense in $\langle \Gamma(X), \tau_{aw} \rangle$.*

Proof. By Lemma 4.3, it suffices to show that the functions with nonempty arg min are dense in $\Gamma(X)$. For each $f \in \Gamma(X)$ with $\operatorname{arg\,min} f = \varnothing$, we produce a sequence $\langle f_n \rangle$ in $\Gamma(X)$ that is τ_{aw}-convergent to f with $\operatorname{arg\,min} f_n \ne \varnothing$ for each n.

Suppose first that f is unbounded below. For each n, let

$$f_n(x) = \max\{f(x), -n\}.$$

Evidently, $\operatorname{arg\,min} f_n = \{x : f(x) \le -n\}$, a nonempty set, and by Lemma 3.3, $f = \tau_{aw}\text{-}\lim f_n$. If f is bounded below but has no minimum value, for each $n \in Z^+$, pick $x_n \in X$ with $v(f) < f(x_n) < v(f) + 1/n$. For each n let f_n be the lower semicontinuous convex function whose epigraph is the closure of the convex hull of $(x_n, v(f))$ and $\operatorname{epi} f$. It is routine to check that $\operatorname{haus}(\operatorname{epi} f, \operatorname{epi} f_n) < 1/n$, and since the Hausdorff metric topology is stronger than the epi-distance topology, we have $f = \tau_{aw}\text{-}\lim f_n$. □

To show that the well-posed functions form a G_δ-subset of the function space, we apply Lemma 3.2 and Theorem 3.6.

Theorem 4.5. *Let X be a Banach space. If $\Gamma(X)$ is equipped with the epi-distance topology τ_{aw}, then in the sense of Baire category, most elements of $\Gamma(X)$ are well-posed.*

Proof. By Lemma 3.5, the function space is a Baire space; so the statement *most elements of* $\Gamma(X)$ *are well-posed* at least makes sense. By Lemma 4.4, the well-posed functions are dense in the function space. Since X is complete, by our remarks at the beginning of this section, $f \in \Gamma(X)$ is well-posed if and only if for each $i \in Z^+$ there exists $\alpha > v(f)$ with $\operatorname{diam} \operatorname{lev}(f; \alpha) < 1/i$. Thus, f is not well-posed if and only if $f \in \mathscr{F}_i$ for some $i \in Z^+$, where

$$\mathscr{F}_i = \{f \in \Gamma(X): \text{ for each } \alpha > v(f), \text{ we have } \operatorname{diam} \operatorname{lev}(f; \alpha) \ge 1/i\}.$$

It remains only to show that each $\mathscr{F}_i$ is τ_{aw}-closed. To this end, let $\langle f_n \rangle$ be a sequence in $\mathscr{F}_i$ τ_{aw}-convergent to f, and let $\alpha > v(f)$ be arbitrary. By Theorem 3.6, $\operatorname{lev}(f; \alpha) = \tau_{aw}\text{-}\lim \operatorname{lev}(f_n; \alpha)$; so, by Lemma 3.2 and Theorem 3.7(a), we have

$$\operatorname{diam} \operatorname{lev}(f; \alpha) = \lim_{n \to \infty} \operatorname{diam} \operatorname{lev}(f_n; \alpha) \ge 1/i.$$

We conclude that $f \in \mathscr{F}_i$, completing the proof. □

We note that Theorem 4.5 has a dual interpretation. It is well-known that $f \to f^*$ is a bijection from $\Gamma(X)$ to $\Gamma^*(X^*)$, where $\Gamma^*(X^*)$ denotes the proper, weak*-lower semicontinuous, convex functions on X^* [24, §14]. Since the conjugate map is a homeomorphism of $\langle \Gamma(X), \tau_{aw} \rangle$ onto $\langle \Gamma^*(X^*), \tau_{aw} \rangle$ for an arbitrary normed linear space [10], it follows from Lemma 3.5 that $\langle \Gamma^*(X^*), \tau_{aw} \rangle$ is completely metrizable when X is a Banach space. Now by the celebrated Asplund-Rockafellar Theorem [1], for any Banach space X, $f \in \Gamma(X)$ is well-posed if and only if f^* is Frechet differentiable at the origin. Putting this all together, we obtain

Theorem 4.6. *Let X be a Banach space, and let $\Gamma^*(X^*)$ be the proper, weak*-lower semicontinuous, convex functions on X^*. Then $\langle \Gamma^*(X^*), \tau_{aw} \rangle$ is completely metrizable, and most functions in $\langle \Gamma^*(X^*), \tau_{aw} \rangle$ are Frechet differentiable at the origin.*

If X is any normed linear space, then the nonconstant continuous affine functionals on X are open in the Hausdorff metric topology on $\Gamma(X)$. Thus, the well-posed functions cannot even be dense in $\Gamma(X)$ so topologized. Now let X be any infinite dimensional normed linear space. We intend to show that with respect to the Mosco topology on $\Gamma(X)$, the set of functions that are unbounded below forms a dense and G_δ subset of $\Gamma(X)$. We require Lemma 2.2 of [8], which we reproduce for the convenience of the reader (see also [2, §3.5.2]).

Lemma 4.7. *Let X be a normed linear space, and let K be a weakly compact subset of $X \times R$. Suppose $f \in \Gamma(X) \cap (K^c)^+$. Then there exists $\varepsilon > 0$ and a finite collection of pairs (y_1, μ_1), $(y_2, \mu_2), \ldots, (y_m, \mu_m)$ in $X^* \times R$ such that $u(\cdot) \equiv \sup_{k \le m} \langle \cdot, y_k \rangle - \mu_k \in (K^c)^+$, and for each $k \le m$, we have*

$$\inf_{x \in X} [f(x) - \langle x, y_k \rangle + \mu_k] > \varepsilon .$$

Our construction rests on the next elementary fact.

Lemma 4.8. *Let X be a normed linear space, and let $\{y_1, y_2, \ldots, y_m\}$ be vectors in X^*. The following are equivalent:*

(1) $\theta^* \notin \operatorname{conv}\{y_1, y_2, \ldots, y_m\}$;
(2) *there exists $x_0 \in X$ such that for each $k \le m$, we have $\langle x_0, y_k \rangle > 0$.*

Proof. (1) $\Rightarrow$ (2). Let $P = \operatorname{conv}\{y_1, y_2, \ldots, y_m\}$; since P is weak*-compact, by the separation theorem applied to θ^* and P, there exists $x_0 \in X$ such that $\langle x_0, \theta^* \rangle < \inf_{y \in P} \langle x_0, y \rangle$. In particular, $\langle x_0, y_k \rangle$ is positive for each index k.

(2) $\Rightarrow$ (1). Suppose (1) fails. If some y_k is the origin of X^*, then clearly (2) fails. Otherwise, there exist nonnegative scalars $\alpha_1, \ldots, \alpha_m$ summing to 1 at least two of which are positive such that $\alpha_1 y_1 + \alpha_2 y_2 + \cdots + \alpha_m y_m = \theta^*$. Without loss of generality, we may assume $\alpha_1 \ne 0$. Clearly, whenever $x \in X$ satisfies $\langle x, y_k \rangle > 0$ for $k = 2, 3, \ldots, m$, we must have $\langle x, y_1 \rangle < 0$, because $y_1 = (-1/\alpha_1)(\alpha_2 y_2 + \cdots + \alpha_m y_m)$, and for some $k \ge 2$, the coefficient α_k is positive. Thus, (2) again fails. □

Theorem 4.9. *Let X be an infinite dimensional normed linear space. Then $\mathscr{D} \equiv \{f \in \Gamma(X): f$ is unbounded below$\}$ is a dense and G_δ subset of $\langle \Gamma(X), \tau_M \rangle$.*

Proof. Evidently, $\mathscr{D} = \bigcap_{n \in Z^+} (\Gamma(X) \cap (X \times (-\infty, -n))^-)$; so, $\mathscr{D}$ is a G_δ subset of the function space. Only the density of $\mathscr{D}$ is in question. To establish density, let $\mathscr{U} \equiv V_1^- \cap V_2^- \cap \cdots \cap V_n^- \cap (K^c)^+$ be a basic τ_M-neighborhood of a proper, convex, lower semicontinuous function f on X. For each $i \leq n$, choose $(x_i, \alpha_i) \in \operatorname{epi} f \cap V_i$. Let $\{(y_k, \mu_k): 1 \leq k \leq m\}$ be the points in $X^* \times R$ whose existence is guaranteed by Lemma 4.7 with respect to the function f and the weakly compact set K. For each $k \leq m$, write $\beta_k = \mu_k - \varepsilon/2$, and let $h: X \to R$ be defined by

$$h(x) = \sup_{1 \leq k \leq m} \langle x, y_k \rangle - \beta_k .$$

Lemma 4.7 says that for each $(x, \alpha) \in K$, we have $\alpha < h(x) - \varepsilon/2$, and for each $(x, \alpha) \in \operatorname{epi} f$, we have $\alpha > h(x) + \varepsilon/2$ (in particular, $\alpha_i > h(x_i) + \varepsilon/2$ for $i = 1, 2, \ldots, n$).

Next, let K_0 be the projection of K onto X, and choose $\lambda > 0$ such that $K_0 \cup \{x_1, x_2, \ldots, x_n\} \subset \lambda U$. Since X is infinite dimensional, there exist linearly independent vectors $z_1, z_2, \ldots, z_m$ in X^* such that $\|z_k - y_k\| < \varepsilon/2\lambda$ for each index $k \leq m$. Consider the convex function $g: X \to R$ defined by

$$g(x) = \sup_{1 \leq k \leq m} \langle x, z_k \rangle - \beta_k .$$

By the choice of λ and the definition of h, it is clear that $\operatorname{epi} g \cap K = \emptyset$ and that $\alpha_i > g(x_i)$ for $i = 1, \ldots, n$, so that $(x_i, \alpha_i) \in \operatorname{epi} g$ for each i. This means that $g \in \mathscr{U}$. Invoking Lemma 4.8, there exists $x_0 \in X$ such that for each $k \leq m$, we have $\langle x_0, z_k \rangle > 0$. As a result, for each $k \leq m$, we have $\lim_{\alpha \to -\infty} \langle \alpha x_0, z_k \rangle - \beta_k = -\infty$, so that g fails to be bounded below when restricted to the ray $\{\alpha x_0: \alpha \leq 0\}$. This proves that the functions unbounded below are τ_M-dense in $\Gamma(X)$. $\square$

5. Applications to metric projections

Let $A \in \mathscr{C}(X)$ and $x \in X$ be arbitrary. Recall [23, 24] that the *metric projection* of x onto A is the possibly empty convex set

$$P(x, A) = \{a \in A: \|x - a\| = d(x, A)\}.$$

The convex set A is called *Chebyshev* provided $P(x, A)$ is a singleton for each $x \in X$. If X is reflexive then $P(x, A)$ is nonempty for each x and A; in fact, this property characterizes reflexivity. Are most elements of $\langle \mathscr{C}(X), \tau_{aw} \rangle$ Chebyshev in a reflexive space? The answer is affirmative!

Recall that a convex set is called a *convex body* provided its interior is nonempty, and *rotund* provided its boundary contains no line segments. By a well-known theorem of Klee [26], within the space of closed and bounded convex bodies topologized by Hausdorff distance, the sets that are rotund form a dense and G_δ subset. Each such set is obviously Chebyshev.

Theorem 5.1. *Let X be a reflexive Banach space. Then most elements of $\langle \mathscr{C}(X), \tau_{aw}\rangle$ are Chebyshev.*

Proof. Let $\mathscr{C}_0(X)$ denote the closed and bounded convex bodies in X. By Lemma 3.1, τ_{aw} reduces to the Hausdorff metric topology on $\mathscr{C}_0(X)$. Thus, by Klee's Theorem, it suffices to show that $\mathscr{C}_0(X)$ is dense and open in $\langle \mathscr{C}(X), \tau_{aw}\rangle$. Density is easy to see, for if $A \in \mathscr{C}(X)$, then $A = \tau_{aw}\text{-}\lim(A+(1/n)U)$, and by Lemma 3.3, for each $n \in Z^+$, $A+(1/n)U = \tau_{aw}\text{-}\lim_{k\to\infty} kU \cap (A+(1/n)U)$. By Lemma 3.4 the closed convex bodies are open in $\langle \mathscr{C}(X), \tau_{aw}\rangle$, and by Lemma 3.1 the bounded elements of $\mathscr{C}(X)$ are open in $\langle \mathscr{C}(X), \tau_{aw}\rangle$. Thus, $\mathscr{C}_0(X)$ is dense and open in $\langle \mathscr{C}(X), \tau_{aw}\rangle$, concluding the proof. □

This general line of reasoning was used in [11] to establish Theorem 5.1 in finite dimensions, before the author was aware of τ_{aw}. Theorem 5.1 fails with respect to the box norm in finite dimensions if τ_{aw} is replaced by the Hausdorff metric topology [17]. It is not known whether Theorem 5.1 holds even in a separable reflexive space if τ_{aw} is replaced by τ_M (but see Theorem 5.4 of [7]). On the negative side, $\mathscr{C}_0(X)$ is a set of first category in $\langle \mathscr{C}(X), \tau_M\rangle$ whenever X is infinite dimensional [7, p. 252].

For each $x \in X$ let $\varphi_x : X \to R$ be defined by $\varphi_x(z) = \|x - z\|$. For each $A \in \mathscr{C}(X)$, we clearly have $\arg\min \varphi_x + I(\cdot, A) = P(x, A)$. We intend to show that for each $x \in A$, $\varphi_x + I(\cdot, A)$ is well-posed for most $A \in \mathscr{C}(X)$, provided X is reflexive.

Lemma 5.2. *Let X be a normed linear space. Suppose $\langle x_n\rangle \to x$ and $A = \tau_{aw}\text{-}\lim A_n$. Then $\varphi_x + I(\cdot, A) = \tau_{aw}\text{-}\lim \varphi_{x_n} + I(\cdot, A_n)$.*

Proof. Let $\varepsilon > 0$, $\rho > 0$. Choose N so large that $n > N$ implies both $\|x_n - x\| < \varepsilon/2$ and $\text{haus}_\rho(A_n, A) < \varepsilon/2$. Fix $n > N$. We must show that both of the following conditions hold:

(i) $e(\text{epi}\,\varphi_{x_n} + I(\cdot, A_n) \cap \rho U \times [-\rho, \rho], \text{epi}\,\varphi_x + I(\cdot, A)) \le \varepsilon$;

(ii) $e(\text{epi}\,\varphi_x + I(\cdot, A) \cap \rho U \times [-\rho, \rho], \text{epi}\,\varphi_{x_n} + I(\cdot, A_n)) \le \varepsilon$.

We only establish (i), as the proofs are identical. Fix $a_n \in A_n \cap \rho U$ and $\alpha \ge \|a_n - x_n\|$ (we need not assume that $\alpha \in [-\rho, \rho]$). Since $\text{haus}_\rho(A, A_n) < \varepsilon/2$, there exists $c_n \in A$ with $\|c_n - a_n\| < \varepsilon/2$. Clearly,

$$\varphi_x(c_n) = \|c_n - x\| < \|a_n - x_n\| + \varepsilon = \varphi_{x_n}(a_n) + \varepsilon \le \alpha + \varepsilon.$$

Thus, $(c_n, \alpha+\varepsilon) \in \text{epi}\,\varphi_x + I(\cdot, A)$, and $\|(a_n, \alpha) - (c_n, \alpha+\varepsilon)\| \le \varepsilon$. □

In the literature a closed convex set A is called *approximatively compact* [19, 24] provided for each $x \in X$ the convex function $\varphi_x + I(\cdot, A)$ is well-posed in the generalized sense. Of course, this yields compactness of $P(x, A)$ for each x. By Theorem 4.2 and Lemma 5.2, approximative compactness for a convex set A in a Banach space may be characterized by this stronger property: at each $x \in X$, whenever $\langle x_n\rangle \to x$, $A = \tau_{aw}\text{-}\lim A_n$, and $\langle d(x, a_n)\rangle \to d(x, A)$ where for each n, $a_n \in A_n$, then $\langle a_n\rangle$ has a subsequence convergent to a point of $P(x, A)$.

It is worthwhile saying what approximative compactness yields for the metric projection, viewed as a bivariate set valued function, in the language of multifunctions. Recall that a set valued function F from a topological space T to a topological space Y is called *upper semicontinuous* [27, §18] at $t_0 \in T$ provided whenever V is a neighborhood of $F(t_0)$, there exists a neighborhood W of t_0 such that for each $t \in W$ we have $F(t) \subset V$ (in the notation of §2, for each neighborhood V of $F(t_0)$, $\{t: F(t) \in V^+\}$ contains a neighborhood of t_0). By the remarks above, $(x, A) \to P(x, A)$ is upper semicontinuous at each (x, A) where A is approximatively compact, provided we equip $\mathscr{C}(X)$ with τ_{aw}. It is known [7] that for reflexive X, the metric projection is weakly upper semicontinuous at each point of $X \times \mathscr{C}(X)$, provided $\mathscr{C}(X)$ is equipped with τ_M. This of course remains true if we equip $\mathscr{C}(X)$ with the stronger topology τ_{aw}.

Lemma 5.3. *Let X be a reflexive Banach space. Then for each $x \in X$, $\{A \in \mathscr{C}(X): \varphi_x + I(\cdot, A)$ is well-posed$\}$ is dense in $\mathscr{C}(X)$.*

Proof. Without loss of generality, we may assume $x = \theta$. If $\theta \in A$, then $\varphi_\theta + I(\cdot, A)$ is already well-posed. Otherwise, by Lemma 3.3, it suffices to assume A is weakly compact, because A can be τ_{aw}-approximated by $\langle A \cap nU \rangle$. Under this assumption, choose $a_0 \in A$ nearest θ. Obviously, $\langle \text{conv}(\{(1-1/n)a_0\} \cup A) \rangle$ τ_{aw}-approximates A; so, it suffices to show that $\varphi_\theta + I(\cdot, \text{conv}(\{\alpha a_0\} \cup A))$ is well-posed for each $\alpha \in (0, 1)$.

To see this, we draw on a construction presented in Lemma 5.2 of [7]. Fix $\alpha \in (0, 1)$, choose ρ such that $A \subset \rho U$, and fix $\lambda > 0$. In order to prove well-posedness, it suffices to show that there exists $\mu > 0$ depending only on λ, α, and ρ such that whenever $z \in \text{conv}(\{\alpha a_0\} \cup A)$ with $\|z - \alpha a_0\| > \lambda$, then $\|z\| > \alpha \|a_0\| + \mu$.

Choose $y \in X^*$ separating $\|a_0\| U$ from A such that $\langle a_0, y \rangle = \|a_0\|$ and $\|y\| = 1$ [24, p. 76]. Let $z \in \text{conv}(\{\alpha a_0\} \cup A)$ with $\|z - \alpha a_0\| > \lambda$ be otherwise arbitrary. For some $\beta \in [0, 1)$ and $a \in A$ we have $z = \beta(\alpha a_0) + (1 - \beta)a$. Notice that $\|z - \alpha a_0\| = (1 - \beta)\|a - \alpha a_0\|$, so that $(1 - \beta)2\rho > \lambda$. Let $a_1 = \beta a_0 + (1 - \beta)a$. Since $z = a_1 - \beta(1 - \alpha)a_0$ and since $a_1 \in A$, we have

$$\begin{aligned}
\|z\| = \|z\| \cdot \|y\| \geq \langle z, y \rangle &= \langle a_1, y \rangle - \beta(1 - \alpha)\langle a_0, y \rangle \\
&\geq \|a_0\| - \beta(1 - \alpha)\|a_0\| \\
&= \alpha\|a_0\| + (1 - \beta)(1 - \alpha)\|a_0\| \\
&> \alpha\|a_0\| + [\lambda(1 - \alpha)/2\rho] \cdot \|a_0\|.
\end{aligned}$$

This shows that $\varphi_\theta + I(\cdot, \text{conv}(\{\alpha a_0\} \cup A))$ is well-posed. □

It is known that if for each $A \in \mathscr{C}(X)$ the associated metric projection program $\varphi_\theta + I(\cdot, A)$ is well-posed, then X is an *E-space* in the sense of Holmes [24, §31]: X is reflexive, its unit ball U is rotund, and the norm of X is a Kadec norm. We have generically

Theorem 5.4. *Let X be a reflexive Banach space, with $\mathscr{C}(X)$ equipped with τ_{aw}. Then for each $x \in X$, $\{A \in \mathscr{C}(X): \varphi_x + I(\cdot, A)$ is well-posed$\}$ is a dense and G_δ subset of $\mathscr{C}(X)$.*

Proof. For each $n \in Z^+$, let $\mathscr{F}_n = \{A \in \mathscr{C}(X):$ for each $\alpha > d(x, A)$, we have $\operatorname{diam}\{a \in A: \|x - a\| \leq \alpha\} \geq 1/n\}$. Notice that $\{a \in A: \|x - a\| \leq \alpha\}$ is nothing but $\operatorname{lev}(\varphi_x + I(\cdot, A); \alpha)$. By Lemma 3.2, Theorem 3.6, and Lemma 5.2, each $\mathscr{F}_n$ is τ_{aw}-closed in $\mathscr{C}(X)$. The result now follows from Lemma 5.3. □

Corollary 5.5. *Let X be a separable reflexive Banach space. Then for most $A \in \langle \mathscr{C}(X), \tau_{aw} \rangle$, $\varphi_x + I(\cdot, A)$ is well-posed for most $x \in X$.*

Proof. This follows from Theorem 5.4 and the Kuratowski-Ulam Theorem [27, p. 247], which in this application only requires separability of X ($\langle \mathscr{C}(X), \tau_{aw} \rangle$ is not separable unless X is finite dimensional [3]). □

For related genericity results on metric projections, the reader may consult [22].

References

1. E. Asplund and R. T. Rockafellar, *Gradients of convex functions*, Trans. Amer. Math. Soc. **139** (1969), 443–467.
2. H. Attouch, *Variational convergence for functions and operators*, Pitman, Boston, Mass., 1984.
3. H. Attouch, R. Lucchetti, and R. Wets, *The topology of the ρ-Hausdorff distance*, Ann. Mat. Pura Appl. (to appear).
4. H. Attouch and R. Wets, *Quantitative stability of variational systems*: I. *The epigraphical distance*, Trans. Amer. Math. Soc. (to appear).
5. ——, *Quantitative stability of variational systems*; III. *Stability of ε-minimizers*, Working paper IIASA, Laxenburg, Austria, 1988.
6. D. Azé and J-P. Penot, *Operations on convergent families of sets and functions*, Optimization **21** (1990), 521–534.
7. G. Beer, *On Mosco convergence of convex sets*, Bull. Austral. Math. Soc. **38** (1988), 239–253.
8. ——, *On the Young-Fenchel transform for convex functions*, Proc. Amer. Math. Soc. **104** (1988), 1115–1123.
9. ——, *Convergence of continuous linear functionals and their level sets*, Arch. Math. **52** (1989), 482–491.
10. ——, *Conjugate convex functions and the epi-distance topology*, Proc. Amer. Math. Soc. **108** (1990), 117–126.
11. ——, *Metric projections and the Fell topology*, Boll. Un. Mat. Ital. (7) **3-B** (1989), 925–937.
12. ——, *Three characterizations of the topology of Mosco convergence for convex functions*, Arch. Math. **55** (1990), 285–292.
13. G. Beer and R. Lucchetti, *Minima of quasi-convex functions*, Optimization **20** (1989), 581–596.
14. J. Borwein and S. Fitzpatrick, *Mosco convergence and the Kadec property*, Proc. Amer. Math. Soc. **106** (1989), 843–850.
15. C. Castaing and M. Valadier, *Convex analysis and measurable multifunctions*, Lecture Notes in Math., vol. 580, Springer-Verlag, Berlin, 1975.

16. L. Contesse and J.-P. Penot, *Continuity of polarity and conjugacy for the epi-distance topology*, preprint.

17. F. DeBlasi and J. Myjak, *On the minimum distance to a closed convex set in a Banach space*, Bull. Acad. Sci. Polon. **29** (1981), 373–376.

18. ——, *Some generic properties in convex and nonconvex optimization theory*, Ann. Soc. Math. Polon. **24** (1984), 1–14.

19. N. Efimov and S. Stechkin, *Approximative compactness and Chebyshev sets*, Soviet Math. Dokl. **2** (1961), 1226–1228.

20. S. Francaviglia, A. Lechicki, and S. Levi, *Quasi-uniformization of hyperspaces and convergence of nets of semicontinuous multifunctions*, J. Math. Anal. Appl. **112** (1985), 347–370.

21. M. Furi and A. Vignoli, *About well-posed optimization problems for functionals in metric spaces*, J. Optim. Theory Appl. (1970), 225–229.

22. P. G. Georgiev, *The strong Ekeland variational principle, the strong drop theorem, and applications*, J. Math. Anal. Appl. **131** (1988), 1–21.

23. J. Giles, *Convex analysis with application in differentiation of convex functions*, Research Notes in Math., no. 58, Pitman, London, 1982.

24. R. Holmes, *A course in optimization and best approximation*, Lecture Notes in Math., vol. 257, Springer-Verlag, New York, 1972.

25. J. Joly, *Une famille de topologies sur l'ensemble des fonctions convexes pour lesquelles la polarité est bicontinue*, J. Math. Pures Appl. **52** (1973), 421–441.

26. V. Klee, *Some new results on smoothness and rotundity in normed linear spaces*, Math. Ann. **139** (1959), 51–63.

27. K. Kuratowski, *Topology*, Vol. 1, Academic Press, New York, 1966.

28. R. Lucchetti, *Some aspects of the connections between Hadamard and Tyhonov well-posedness of convex programs*, Boll. Un. Mat. Ital. Anal. **1** (1982), 337–344.

29. ——, *Hadamard and Tyhonov well-posedness in optimal control*, Methods Oper. Res. **45** (1983), 113–125.

30. R. Lucchetti and F. Patrone, *Sulla densità e genericità di alcuni problemi di minimo ben posti*, Boll. Un. Mat. Ital. **15B** (1978), 225–240.

31. R. Lucchetti, G. Salinetti, and R. Wets, *Uniform convergence of probability measures: topological criteria*; preprint.

32. U. Mosco, *Convergence of convex sets and of solutions of variational inequalities*, Adv. in Math. **3** (1969), 510–585.

33. ——, *On the continuity of the Young-Fenchel transform*, J. Math. Anal. Appl. **35** (1971), 518–535.

34. F. Patrone, *Most convex functions are nice*, Numer. Funct. Anal. Optim. **9** (1987), 359–369.

35. J. Revalski, *Generic well-posedness in some classes of optimization problems*, Acta Univ. Carolin. Math. Phys. **28** (1987), 117–125.

36. R. T. Rockafellar, *Convex analysis*, Princeton Univ. Press, Princeton, N.J., 1970.

37. R. T. Rockafellar and R. Wets, *Variational systems, an introduction*, Multifunctions and Integrands, G. Salinetti, Ed., Lecture Notes in Math., vol. 1091, Springer-Verlag, Berlin, 1984.

38. G. Salinetti and R. Wets, *On the convergence of sequences of convex sets in finite dimensions*, SIAM Rev. **21** (1979), 18–33.

39. Y. Sonntag, *Convergence au sens de Mosco; théorie et applications à l'approximation des solutions d'inéquations*, Thèse d'Etat, Université de Provence, Marseille, 1982.

40. A. Tyhonov, *On the stability of the functional optimization problem*, U.S.S.R. Comput. Math. and Math. Phys. **6** (1966), 28–33.

41. R. Wijsman, *Convergence of sequences of convex sets, cones, and functions* II, Trans. Amer. Math. Soc. **123** (1966), 32–45.

Department of Mathematics, California State University, Los Angeles, California 90032

Department of Mathematics, University of California, Davis, California 95616