PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 108, Number 1, January 1990

CONJUGATE CONVEX FUNCTIONS AND THE EPI-DISTANCE TOPOLOGY

GERALD BEER

(Communicated by R. Daniel Mauldin)

ABSTRACT. Let $\Gamma(X)$ denote the proper, lower semicontinuous, convex functions on a normed linear space, and let $\Gamma^*(X^*)$ denote the proper, weak*-lower semicontinuous, convex functions on the dual X^* of X. It is well-known that the Young-Fenchel transform (conjugate operator) is bicontinuous when X is reflexive and both $\Gamma(X)$ and $\Gamma^*(X^*)$ are equipped with the topology of Mosco convergence. We show that without reflexivity, the transform is bicontinuous, provided we equip both $\Gamma(X)$ and $\Gamma^*(X^*)$ with the (metrizable) epi-distance topology of Attouch and Wets. Convergence of a sequence of convex functions $\langle f_n \rangle$ to f in this topology means uniform convergence on bounded subsets of the associated sequence of distance functionals $\langle d(\cdot, \operatorname{epi} f_n)\rangle$ to $d(\cdot, \operatorname{epi} f)$.

1. INTRODUCTION

Let X be a normed linear space with continuous dual X^*, and let $\Gamma(X)$ (resp. $\Gamma^*(X^*)$) be the proper, lower semicontinuous (resp. weak*-lower semicontinuous) convex functions on X (resp. X^*). A fundamental construction in convex analysis [14], [20] is the *Young-Fenchel transform* $f \to f^*$ from $\Gamma(X)$ to $\Gamma^*(X^*)$, where f^* is defined by the familiar formula

$$f^*(y) = \sup_{x \in X} \langle x, y\rangle - f(x) \qquad (y \in X^*).$$

The function f^* is called the *conjugate* of f. As is well known [14, §14], $f \to f^*$ is an order reversing bijection of $\Gamma(X)$ onto $\Gamma^*(X^*)$. Specifically, the inverse $h \to h^\#$ of $f \to f^*$ from $\Gamma^*(X^*)$ to $\Gamma(X)$—called the *second conjugate map*—is defined at each $h \in \Gamma^*(X^*)$ by

$$h^\#(x) = \sup_{y \in X^*} \langle x, y\rangle - h(y) \qquad (x \in X).$$

Received by the editors October 17, 1988 and, in revised form, February 7, 1989.
1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 26B25; Secondary 54B20, 46B10.
Key words and phrases. Convex function, Young-Fenchel transform, conjugate convex function, polar, epi-distance topology, uniform convergence on bounded sets, Mosco convergence.

A great deal of interest over the last twenty years has been focused on Mosco convergence of sequences in $\Gamma(X)$ [1], [8], [16], [17], [19], [21]. Given a sequence $f, f_1, f_2, \ldots$ in $\Gamma(X)$, $\langle f_n \rangle$ is declared *Mosco convergent* of f provided at each $x \in X$, both of the following conditions hold:

(i) there exists a sequence $\langle x_n \rangle$ in X convergent strongly to x for which $f(x) = \lim f_n(x_n)$;
(ii) whenever $\langle x_n \rangle$ converges weakly to x, then $f(x) \le \liminf f_n(x_n)$.

The key fact about Mosco convergence is the "sequential bicontinuity" of the Young-Fenchel transform in reflexive spaces: if $f, f_1, f_2, \ldots$ is a sequence in $\Gamma(X)$, then $\langle f_n \rangle$ is convergent to f in the sense of Mosco if and only if $\langle f_n^* \rangle$ is also convergent to f^* in the sense of Mosco. This result was established in infinite dimensions by Mosco in [19], but first in finite dimensions by Wijsman [24].

There are two reservations one might have about Mosco's result. First, it is not a bona fide continuity theorem, in that Mosco convergence of sequences is not compatible with a first countable topology on $\Gamma(X)$ unless X is separable (see [1], [7], [22]). Second, its validity as stated seems limited to reflexive spaces. In [8], this author addressed the first shortcoming, producing a simple Vietoris-type topology on $\Gamma(X)$ (identifying functions with their epigraphs) compatible with Mosco convergence of sequences with respect to which $f \to f^*$ is actually a homeomorphism (see more generally [5], [11], 15]). On the other hand, continuity of $f \to f^*$ with respect to Mosco convergence seems to require reflexivity.

It is the purpose of this article to produce a stronger and even more tractible topology on $\Gamma(X)$—which coincides with the topology of [8] in finite dimensions —with respect to which $f \to f^*$ is a homeomorphism of $\Gamma(X)$ onto $\Gamma^*(X^*)$ for *any* normed linear space X.

2. Preliminaries

In the sequel, the origin and unit ball of X (resp. X^*) will be denoted by θ and U (resp. θ^* and U^*). We agree to equip $X \times \mathbf{R}$ with the box norm: $\|(x, \alpha)\| = \max\{\|x\|, |\alpha|\}$ for each $x \in X$ and $\alpha \in \mathbf{R}$. A similar convention applies to $X^* \times \mathbf{R}$.

The collection of closed nonempty convex subsets of X will be denoted by $\mathscr{C}(X)$. For each $A \in \mathscr{C}(X)$, the polar A° of A is the following weak*-closed convex subset of X^*: $A^\circ = \{y \in X^* : \text{for each } a \in A,\ \langle a, y \rangle \le 1\}$.

If $f: X \to (-\infty, \infty]$, its *effective domain*, denoted by $\operatorname{dom} f$, is the set of points x in X where $f(x)$ is finite. Its *epigraph* is the following subset of $X \times \mathbf{R}$:

$$\operatorname{epi} f \equiv \{(x, \alpha) : x \in X, \alpha \in \mathbf{R}, \text{ and } \alpha \ge f(x)\}.$$

As is well known [13], [20], f is convex (resp. lower semicontinuous) provided $\operatorname{epi} f$ is a convex (resp. closed) subset of $X \times \mathbf{R}$. If $A \in \mathscr{C}(X)$, the *distance functional* $d(\cdot, A): X \to \mathbf{R}$ for A, defined by $d(x, A) = \inf_{a \in A} \|x - a\|$, is a continuous convex function. Another basic proper lower semicontinuous convex function on X is the *indicator function* $I(\cdot, A)$ associated with an element A of $\mathscr{C}(X)$, defined by

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

The conjugate of $I(\cdot, A)$ is the *support functional* $s(\cdot, A)$ of A, defined at each $y \in X^*$ by $s(y, A) = \sup_{a\in A}\langle a, y\rangle$.

Mosco convergence of sequences in $\Gamma(X)$ is a particular instance of Mosco convergence of closed convex sets, identifying functions with their epigraphs [18, Lemma 1.10]. Specifically, given $A, A_1, A_2, \ldots$ in $\mathscr{C}(X)$, $\langle A_n\rangle$ is declared *Mosco convergent* to A provided

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

(ii) whenever $n(1) < n(2) < n(3) < \cdots$, and $a_{n(k)} \in A_{n(k)}$ for each index k, then the weak convergence of $\langle a_{n(k)}\rangle$ to $x \in X$ implies $x \in A$.

In any Banach space, Mosco convergence of convex sets is compatible with the topology τ_M on $\mathscr{C}(X)$ generated by all sets of the form

$$(K^c)^+ \equiv \{A \in \mathscr{C}(X) : A \cap K = \varnothing\},$$
$$V^- \equiv \{A \in \mathscr{C}(X) : A \cap V \neq \varnothing\},$$

where K ranges over the weakly compact subsets of X and V ranges over the norm open subsets of X [7, Theorem 3.1]. When X is reflexive, this topology is the weakest topology on $\mathscr{C}(X)$ such that for each weakly compact subset K of X, $A \to \inf\{\|a - k\| : a \in A, k \in K\}$ is continuous on $\mathscr{C}(X)$ [7, Theorem 3.3]. Most importantly, when X is reflexive, the conjugate operator from $\langle\Gamma(X), \tau_M\rangle$ to $\langle\Gamma(X^*), \tau_M\rangle$ is a homeomorphism, under the usual identification of elements of $\Gamma(X)$ and $\Gamma(X^*)$ with their epigraphs [8].

The topology of interest here was also formally introduced (but not studied in depth) by Mosco in [18]. It is implicit in a slightly earlier paper of Walkup and Wets [23]. As indicated by recent work of Attouch and Wets [3, 4], it has considerable potential in numerical analysis: it seems a promising tool to obtain convergence rates for sequences of functions and quantitative rather than topological results with respect to the stability of solutions to optimization problems.

To describe this topology, we follow the approach of this author in [9]. For a normed linear space X, let $C_B(X, \mathbf{R})$ denote the vector space of continuous real functions on X that are bounded on bounded subsets of X. We make $C_B(X, \mathbf{R})$ a metrizable locally convex space by imposing the following sequence $\langle p_n\rangle$ of seminorms on $C_B(X, \mathbf{R})$:

$$p_n(f) = \sup\{|f(x)| : \|x\| \leq n\} \qquad (f \in C_B(X, \mathbf{R})).$$

Clearly, convergence of a net in $C_B(X, \mathbf{R})$ with respect to these seminorms means uniform convergence on bounded subsets of X. By the topology τ of *uniform convergence of distance functions on bounded subsets* for $\mathscr{C}(X)$, we mean $\mathscr{C}(X)$ as identified with $\{d(\cdot, A): A \in \mathscr{C}(X)\}$ as a subset of this locally convex space, equipped with the relative topology. Evidently, $\langle \mathscr{C}(X), \tau\rangle$ is metrizable; in fact, it is completely metrizable [2]. We find it worthwhile to record the following facts, essentially observed in sequential/functional form in [3], as a lemma.

Lemma 2.1. *Let $\mathscr{C}(X)$ be the nonempty closed convex subsets of a normed linear space X. Then on $\mathscr{C}(X)$, $\tau_M \subset \tau$, and if X is finite dimensional, the two topologies coincide.*

Proof. To show $\tau_M \subset \tau$, we show that $(K^c)^+ \in \tau$ for each weakly compact subset K of X, and $V^- \in \tau$ for each norm open subset V of X. To this end, suppose first that $A_0 \in (K^c)^+$ where K is weakly compact. Since A_0 is weakly closed, there exists $\delta > 0$ such that for each $a \in A_0$ and for each $k \in K$, we have $\|a-k\| > \delta$. Choose $n \in \mathbf{Z}^+$ with $K \subset nU$. Then

$$\begin{aligned} A_0 &\in \{A \in \mathscr{C}(X): p_n(d(\cdot, A_0) - d(\cdot, A)) < \delta\} \\ &\subset \{A \in \mathscr{C}(X): \text{for each } k \in K, |d(k, A_0) - d(k, A)| < \delta\} \subset (K^c)^+. \end{aligned}$$

This proves that $(K^c)^+$ is τ-open. On the other hand, if $A_0 \in V^-$ where V is norm open, choose $a_0 \in A_0$ and $\varepsilon > 0$ such that $a_0 + \varepsilon U \subset V^-$. With $n > \|a_0\|$, we have

$$\begin{aligned} A_0 &\in \{A \in \mathscr{C}(X): p_n(d(\cdot, A_0) - d(\cdot, A)) < \varepsilon\} \\ &\subset \{A \in \mathscr{C}(X): |d(a_0, A_0) - d(a_0, A)| < \varepsilon\} \\ &\subset \{A \in \mathscr{C}(X): A \cap (a_0 + \varepsilon U) \neq \varnothing\} \subset V^-. \end{aligned}$$

This proves that V^- is τ-open.

When X is finite dimensional, it is known ([6, Theorem 1], [12, p. 356]) that $\langle \mathscr{C}(X), \tau_M\rangle$ is homeomorphic to $\{d(\cdot, A): A \in \mathscr{C}(X)\}$, equipped with the topology of uniform convergence on compacta. But in a space in which closed and bounded sets are compact, this topology coincides with the topology of uniform convergence on bounded subsets. □

Surely the simplest example showing that τ-convergence of sequences may be properly stronger than τ_M-convergence is the following: in l_2 with the standard orthonormal base $\{\mathbf{e}_n: n \in \mathbf{Z}^+\}$, take $A_n = \operatorname{conv}\{\theta, \mathbf{e}_n\}$ and take $A = \{\theta\}$. Although $A = \tau_M - \lim A_n$, $\langle d(\cdot, A_n)\rangle$ fails to converge uniformly to $d(\cdot, A)$ on the unit ball, since for each n, $|d(\mathbf{e}_n, A_n) - d(\mathbf{e}_n, A)| = 1$.

It is well known that convergence in $\mathscr{C}(X)$ with respect to the familiar Hausdorff metric H [10] amounts to global uniform convergence of distance functionals. In fact, for each A and B in $\mathscr{C}(X)$, we have

$$\begin{aligned} H(A, B) &= \inf\{\varepsilon: A \subset B + \varepsilon U \text{ and } B \subset A + \varepsilon U\} \\ &= \sup_{x \in X} |d(x, A) - d(x, B)|. \end{aligned}$$

Thus, the topology induced by H on $\mathscr{C}(X)$ is finer than τ. It is noteworthy that τ-convergence of a sequence is very close to Hausdorff metric convergence on bounded subsets of X, as described by [9, Lemma 3.1]: $A = \tau - \lim A_n$ if and only if for each $\rho > 0$ and each $\varepsilon > 0$, there exists an integer N such that for each $n > N$, we have both

$$A \cap \rho U \subset A_n + \varepsilon U \quad \text{and} \quad A_n \cap \rho U \subset A + \varepsilon U .$$

Following Attouch and Wets [3], we write $\mathrm{haus}_\rho(A, B)$ for $\inf\{\varepsilon\colon A \cap \rho U \subset B + \varepsilon U$ and $B \cap \rho U \subset A + \varepsilon U\}$. In view of [9, Lemma 3.1], we have $A = \tau - \lim A_n$ if and only if for each $\rho > 0$, $\lim \mathrm{haus}_\rho(A, A_n) = 0$. In fact, $A = \tau - \lim A_n$ if and only if $\lim \mathrm{haus}_\rho(A, A_n) = 0$ for all $\rho > \rho_0$ where ρ_0 is a fixed positive constant. We use this fact in the proof of our main result.

We may, of course, view $\mathbf{\Gamma}(X)$ (resp. $\mathbf{\Gamma}^*(X^*)$) as metric subspaces of $\langle \mathscr{C}(X \times \mathbf{R}), \tau\rangle$ (resp. $\langle \mathscr{C}(X^* \times \mathbf{R}), \tau\rangle$). Again following [3], we call τ so restricted to either $\mathbf{\Gamma}(X)$ or $\mathbf{\Gamma}^*(X^*)$ the *epi-distance topology*. In view of the seminorm presentation of τ, a local base for the epi-distance topology at $f \in \mathbf{\Gamma}(X)$ consists of all sets of the form

$$\mathbf{\Omega}(f; \rho, \varepsilon) \equiv \{g \in \mathbf{\Gamma}(X)\colon \sup_{\|(x,\alpha)\| \le \rho} |d((x,\alpha), \operatorname{epi} g) - d((x,\alpha), \operatorname{epi} f)| < \varepsilon\}.$$

3. Results

The proof of bicontinuity of the Young-Fenchel transform with respect to the epi-distance topology requires some involved numerical estimates in conjunction with a simple, yet unorthodox, application of the separation theorem. It will be executed by combining four elementary lemmas that now follow.

Lemma 3.1. *Let X be a normed linear space. Suppose B is a bounded subset of $X \times \mathbf{R}$ and ρ is a positive constant. Then there exists $\delta > 0$ such that for each $(x, \lambda) \in B$ and each $(y, \alpha) \in X^* \times \mathbf{R}$ with $\|y\| \le \rho$ and $|\alpha| < \rho + 1$, we have $\lambda - (\langle x, y\rangle - \alpha) < \delta$.*

Proof. Choose $\mu > 0$ with $B \subset \mu U \times [-\mu, \mu]$. We have

$$\lambda - [\langle x, y\rangle - \alpha] < \mu - (-\rho\mu - (\rho + 1)) = (\mu + 1)(\rho + 1)$$

provided $(x, \lambda) \in B$, $\|y\| \le \rho$ and $|\alpha| < \rho + 1$. □

Lemma 3.2. *Let X be a normed linear space and let μ, ρ and ε be positive constants with $\rho > 1$. Suppose $f \in \mathbf{\Gamma}(X)$ and $(y, \alpha) \in \operatorname{epi} f^*$ with $\|(y, \alpha)\| \le \rho$. Then for each $g \in \mathbf{\Omega}(f; \rho\mu + \rho + \varepsilon; \varepsilon(\rho + 1)^{-1})$ and for each $x \in \mu U$, we have $\langle x, y\rangle - \alpha - \varepsilon < g(x)$.*

Proof. If not, then there exists $(x, \lambda) \in \operatorname{epi} g$ such that $x \in \mu U$ and $\langle x, y\rangle - \alpha - \varepsilon = \lambda$. Since $\rho > 1$, we have $\|x\| < \rho\mu + \rho + \varepsilon$, and

$$|\lambda| = |\langle x, y\rangle - \alpha - \varepsilon| \le \rho\mu + \rho + \varepsilon .$$

As a result, $d((x,\lambda), \operatorname{epi} f) = |d((x,\lambda), \operatorname{epi} f) - d((x,\lambda), \operatorname{epi} g)| < \varepsilon(\rho+1)^{-1}$, and there exists $(z,\beta) \in \operatorname{epi} f$ with $\|z-x\| < \varepsilon(\rho+1)^{-1}$ and $\beta < \lambda+\varepsilon(\rho+1)^{-1}$. Also,

$$\langle z, y\rangle \geq \langle x, y\rangle - \|z-x\| \cdot \|y\| > \langle x, y\rangle - \varepsilon(\rho+1)^{-1}\rho$$

so that

$$\begin{aligned} f(z) - (\langle z, y\rangle - \alpha) &\leq \beta - (\langle z, y\rangle - \alpha) \\ &< \lambda + \varepsilon(\rho+1)^{-1} - (\langle z, y\rangle - \alpha) \\ &< \lambda + \varepsilon(\rho+1)^{-1} - \langle x, y\rangle + \varepsilon(\rho+1)^{-1}\rho + \alpha = 0. \end{aligned}$$

This contradicts $(y,\alpha) \in \operatorname{epi} f^*$. $\square$

The next lemma is implicit in the proof of [8, Theorem 3.1]. Despite its simplicity, it is the key tool in our results here, as well as the results of [8].

Lemma 3.3. *Let X be a normed linear space and let $\mu > 0$. Suppose $g \in \Gamma(X)$ and* $\operatorname{dom} g \cap \operatorname{int} \mu U \neq \varnothing$. *Further, suppose there exists $(y,\alpha) \in X^* \times \mathbf{R}$ such that for each $x \in \mu U$ we have $\langle x, y\rangle - \alpha \leq g(x)$. Then there exists (y_1, α_1) $\in$* epi g^* *such that for each $x \in \mu U$, we have $\langle x, y\rangle - \alpha \leq \langle x, y_1\rangle - \alpha_1$.*

Proof. Let C be the following closed convex subset of $X \times \mathbf{R}$:

$$C = \{(x,\beta): x \in \mu U \text{ and } \beta \leq \langle x, y\rangle - \alpha\}.$$

Since $\operatorname{int} C = \{(x,\beta): \|x\| < \mu \text{ and } \beta < \langle x, y\rangle - \alpha\} \neq \varnothing$ and $\operatorname{epi} g \cap \operatorname{int} C = \varnothing$, by the usual separation theorem [13, p. 66] there exists a norm closed hyperplane H in $X \times \mathbf{R}$ that separates $\operatorname{epi} g$ from C. Pick $x_0 \in \operatorname{dom} g$ with $\|x_0\| < \mu$. Now if H were vertical it would necessarily contain both $(x_0, g(x_0))$ and $(x_0, \langle x_0, y\rangle - \alpha)$ and would therefore pass through the interior of C, incompatible with separation. Thus, H is the graph of a norm continuous affine function on X, say $x \to \langle x, y_1\rangle - \alpha_1$. By the definition of conjugate convex function, $(y_1, \alpha_1) \in \operatorname{epi} g^*$, and separation ensures that $\langle x, y\rangle - \alpha \leq \langle x, y_1\rangle - \alpha_1$ for each $x \in \mu U$. $\square$

Lemma 3.4. *Let X be a normed linear space and let $\mu > 0$. Suppose $h \in \Gamma^*(X^*)$ and* $\operatorname{dom} h \cap \mu U^* \neq \varnothing$. *Further, suppose there exists $(x,\alpha) \in X \times \mathbf{R}$ such that for each $y \in \mu U^*$ we have $\langle x, y\rangle - \alpha < h(y)$. Then there exists $(x_1, \alpha_1) \in$* epi $h^{\#}$ *such that for each $y \in \mu U^*$, we have $\langle x, y\rangle - \alpha < \langle x_1, y\rangle - \alpha_1$.*

Proof. The functional $y \to \langle x, y\rangle - \alpha$ is weak*-continuous on the weak*-compact set μU^*, whence its graph K is a weak*-compact subset of $X^* \times \mathbf{R}$. Since $\operatorname{epi} h \cap K = \varnothing$, the sets $\operatorname{epi} h$ and K can be strongly separated by a weak*-closed hyperplane H in $X^* \times \mathbf{R}$ [13, p. 70]. Choose $y_0 \in \operatorname{dom} h \cap \mu U^*$; since $(y_0, h(y_0)) \in \operatorname{epi} h$ and $(y_0, \langle x, y_0\rangle - \alpha) \in K$, H cannot be vertical; so, H is the graph of a weak*-continuous affine functional on X^*, say $y \to \langle x_1, y\rangle - \alpha_1$ with $x_1 \in X$. The result now follows as in the proof of Lemma 3.3. $\square$

Theorem 3.5. *Let X be a normed linear space and let τ be the epi-distance topology. Then the Young-Fenchel transform is a homeomorphism from $\langle \mathbf{\Gamma}(X), \tau \rangle$ to $\langle \mathbf{\Gamma}^*(X^*), \tau \rangle$.*

Proof. Fix $f_0 \in \mathbf{\Gamma}(X)$, and let $\rho > 1$ and $\varepsilon < 1$ be otherwise arbitrary positive constants. To prove continuity of $f \to f^*$ at f_0, we produce $\sigma > 0$ and $\beta > 0$ such that whenever $g \in \mathbf{\Omega}(f_0; \sigma; \beta)$, then $\operatorname{haus}_\rho(\operatorname{epi} g^*, \operatorname{epi} f_0^*) \le \varepsilon$.

Pick $(x_0, \lambda_0) \in \operatorname{epi} f_0$, and let δ be the positive constant whose existence is guaranteed by Lemma 3.1 with respect to the constant ρ and the bounded subset B of $X \times \mathbf{R}$ given by

$$B \equiv (\|x_0\| + 1)U \times [-|\lambda_0| - 1, |\lambda_0| + 1].$$

Also, specify a scalar μ by the formula

$$\mu = \max\{\|x_0\| + 1 + \delta/\varepsilon, |\lambda_0| + 1\}.$$

We are now ready to fix our parameters σ and β:

$$\sigma = \rho\mu + \rho + 1, \qquad \beta = \varepsilon(\rho + 1)^{-1}.$$

By the choice of μ, it is evident that

$$(x_0, \lambda_0) \in B \subset \mu U \times [-\mu, \mu].$$

In particular, $(x_0, \lambda_0) \in \sigma U \times [-\sigma, \sigma]$; so, for each $g \in \mathbf{\Omega}(f_0; \sigma; \beta)$, we have

$$\begin{aligned} d[(x_0, \lambda_0), \operatorname{epi} g] &< d[(x_0, \lambda_0), \operatorname{epi} f_0] + \varepsilon(\rho + 1)^{-1} \\ &= 0 + \varepsilon(\rho + 1)^{-1} < 1. \end{aligned}$$

This means that $\operatorname{epi} g \cap \operatorname{int} B \neq \varnothing$, provided $g \in \mathbf{\Omega}(f_0; \sigma; \beta)$.

To show that $\operatorname{haus} \rho(\operatorname{epi} g^*, \operatorname{epi} f_0^*) \le \varepsilon$ whenever $g \in \mathbf{\Omega}(f_0; \sigma; \beta)$, we must establish both the following for $g \in \mathbf{\Omega}(f_0; \sigma; \beta)$:

(i) $\operatorname{epi} f_0^* \cap (\rho U^* \times [-\rho, \rho]) \subset \operatorname{epi} g^* + (\varepsilon U^* \times [-\varepsilon, \varepsilon])$,
(ii) $\operatorname{epi} g^* \cap (\rho U^* \times [-\rho, \rho]) \subset \operatorname{epi} f_0^* + (\varepsilon U^* \times [-\varepsilon, \varepsilon])$.

Since exactly the same argument applies in each case, we just verify (i). Fix $(y, \alpha) \in \operatorname{epi} f_0^*$ with $\|y\| \le \rho$ and $|\alpha| \le \rho$. By Lemma 3.2, whenever $\|x\| < \mu$, we have $\langle x, y \rangle - \alpha - \varepsilon < g(x)$. Since $\operatorname{epi} g \cap \operatorname{int} B \neq \varnothing$ and $B \subset \mu U \times [-\mu, \mu]$, we have $\operatorname{dom} g \cap \operatorname{int} \mu U \neq \varnothing$. By Lemma 3.3, there exists $(y_1, \alpha_1) \in \operatorname{epi} g^*$ such that for each $x \in \mu U$,

$$\langle x, y \rangle - \alpha - \varepsilon \le \langle x, y_1 \rangle - \alpha_1. \tag{1}$$

The last inequality is valid when $x = \theta$ so that $\alpha_1 \le \alpha + \varepsilon$. Thus, $(y_1, \alpha + \varepsilon) \in \operatorname{epi} g^*$. We will have shown that $(y, \alpha) \in \operatorname{epi} g^* + (\varepsilon U^* \times [-\varepsilon, \varepsilon])$, provided we can show that $\|y - y_1\| \le \varepsilon$.

Suppose this is not so. Then for some norm one element w of X, we have $\langle w, y_1 - y \rangle < -\varepsilon$. Pick $(x_1, \lambda_1) \in \operatorname{epi} g \cap B$. Since $|\alpha + \varepsilon| < \rho + 1$, the choice of δ guarantees that $\lambda_1 - (\langle x_1, y \rangle - \alpha - \varepsilon) < \delta$. Furthermore, since $(x_1, \lambda_1) \in \operatorname{epi} g$

and $(y_1, \alpha_1) \in \operatorname{epi} g^*$, we have $\langle x_1, y_1\rangle - \alpha_1 \le \lambda_1$. Combining these last two inequalities yields

$$\langle x_1, y_1\rangle - \alpha_1 < \langle x_1, y\rangle - \alpha - \varepsilon + \delta. \tag{2}$$

From the choice of w, we now obtain

$$\begin{aligned}
&\langle x_1 + (\delta/\varepsilon)w, y_1\rangle - \alpha_1 - (\langle x_1 + (\delta/\varepsilon)w, y\rangle - \alpha - \varepsilon) \\
&\quad = \langle x_1, y_1\rangle - \alpha_1 - (\langle x_1, y\rangle - \alpha - \varepsilon) + (\delta/\varepsilon)\langle w, y_1 - y\rangle \\
&\quad < \langle x_1, y_1\rangle - \alpha_1 - \langle x_1, y\rangle + \alpha + \varepsilon - \delta.
\end{aligned}$$

By inequality (2) above, this last quantity is negative, whence

$$\langle x_1 + (\delta/\varepsilon)w, y_1\rangle - \alpha_1 - (\langle x_1 + (\delta/\varepsilon)w, y\rangle - \alpha - \varepsilon)$$

is also negative. However, this contradicts inequality (1) with $x = x_1 + (\delta/\varepsilon)w$, for by the choice of μ we have $x_1 + (\delta/\varepsilon)w \in \mu U$.

Proof of the continuity of the second conjugate map from $\langle \mathbf{\Gamma}^*(X^*), \tau\rangle$ to $\langle \mathbf{\Gamma}(X), \tau\rangle$ is identical to the proof of the continuity of the first conjugate map, modulo applying Lemma 3.4 at the place where Lemma 3.3 was applied in the above argument. □

To derive τ_M-sequential continuity of the polar operator $A \to A^\circ$ in the reflexive setting [19, Theorem 3.2], Mosco used τ_M-sequential continuity of $f \to f^*$ in conjunction with the following convergence result for sublevel sets [19, Lemma 3.1]: if $f = \tau_M - \lim f_n$, then for each $\alpha > \inf f$, we have

$$\{x : f(x) \le \alpha\} = \tau_M - \lim\{x : f_n(x) \le \alpha\}.$$

Although we could establish the analogous fact with respect to τ-convergence of sublevel sets to yield τ-continuity of $A \to A^\circ$, the proof involves some ugly estimates. Instead, we choose to produce a short, direct proof of τ-continuity of the polar operator (without reflexivity).

Theorem 3.6. *Let X be a normed linear space and let $A, A_1, A_2, \ldots$ be a sequence in $\mathscr{C}(X)$ with $A = \tau - \lim A_n$. Then $A^\circ = \tau - \lim A_n^\circ$.*

Proof. For each pair of nonempty closed convex subsets A and B of X and each $\rho > 0$, we have $\operatorname{haus}_\rho(A, B) = \operatorname{haus}_\rho(\operatorname{epi} I(\cdot, A), \operatorname{epi} I(\cdot, B))$. Thus, $A = \tau - \lim A_n$ immediately implies $I(\cdot, A) = \tau - \lim I(\cdot, A_n)$. By Theorem 3.5, we conclude that $s(\cdot, A) = \tau - \lim s(\cdot, A_n)$. Fix $\rho > 1$ and $\varepsilon > 0$; it suffices to show that $\operatorname{haus}_\rho(A^\circ, A_n^\circ) \le \varepsilon$, provided $\operatorname{haus}_\rho(\operatorname{epi} s(\cdot, A), \operatorname{epi} s(\cdot, A_n)) < \lambda$, where $\lambda > 0$ is chosen so that $\lambda + \lambda(1+\lambda)^{-1}\rho < \varepsilon$.

Fix n with $\operatorname{haus}_\rho(\operatorname{epi} s(\cdot, A), \operatorname{epi} s(\cdot, A_n)) < \lambda$. We must show that

$$A^\circ \cap \rho U^* \subset A_n^\circ + \varepsilon U^* \quad \text{and} \quad A_n^\circ \cap \rho U^* \subset A^\circ + \varepsilon U^*.$$

Fix $y \in A^\circ \cap \rho U^*$. By the definition of the polar operator and the support functional, this means that $(y, 1) \in \operatorname{epi} s(\cdot, A)$. By assumption, $\operatorname{haus}_\rho(\operatorname{epi} s(\cdot, A), \operatorname{epi} s(\cdot, A_n)) < \lambda$, and since $\rho > 1$, there exists $(y_n, \alpha) \in \operatorname{epi} s(\cdot, A_n)$ with

$\|y_n - y\| < \lambda$ and $|\alpha - 1| < \lambda$. If $\alpha \le 1$, then already we have $y_n \in A_n^\circ$, and $y \in A_n^\circ + \lambda U^* \subset A_n^\circ + \varepsilon U^*$. Otherwise, write $\alpha = 1 + \delta$. Since $\operatorname{epi} s(\cdot, A_n)$ is a convex cone, $((1+\delta)^{-1} y_n, 1) \in \operatorname{epi} s(\cdot, A_n)$. Furthermore,

$$\begin{aligned}\|y - (1+\delta)^{-1} y_n\| &\le (1+\delta)^{-1}\|y - y_n\| + \delta(1+\delta)^{-1}\|y\| \\ &< \lambda + \lambda(1+\lambda)^{-1}\rho < \varepsilon.\end{aligned}$$

Since $(1+\delta)^{-1} y_n \in A_n^\circ$, we again have $y \in A_n^\circ + \varepsilon U^*$. This establishes condition (i). Verification of (ii) is achieved in exactly the same way. □

Curiously, there is no example in the literature of a Mosco convergent sequence $\langle A_n \rangle$ of closed convex sets in a nonreflexive space for which the associated polar sequence $\langle A_n^\circ \rangle$ fails to converge in the same sense. We provide such an example here. Our construction rests on the following fact: if X is a Banach space, and $y, y_1, y_2, y_3, \ldots$ is a sequence of nonzero vectors in X^*, then $\langle y_n \rangle$ converges weak* to y if and only if for each scalar α, $\langle \{x \in X : \langle x, y_n \rangle = \alpha\} \rangle$ converges to $\{x \in X : \langle x, y \rangle = \alpha\}$ in the classical sense of Kuratowski [9, Theorem 4.1]. We will now show that τ_M-continuity of $A \to A^\circ$ fails for $X = l_1$. The key consideration here is that strong and weak convergence agree for sequences in l_1, by Schur's theorem. Thus, Kuratowski and Mosco convergence agree for sequences of closed convex sets. Consider this sequence in l_∞ (the dual of l_1): $\langle \mathbf{e}_1 + \mathbf{e}_n \rangle$. Evidently, this converges weak* to $\mathbf{e}_1$. For each n let $A_n = \{x \in l_1 : \langle x, \mathbf{e}_1 + \mathbf{e}_n \rangle = 1\}$ and let $A = \{x \in l_1 : \langle x, \mathbf{e}_1 \rangle = 1\}$. By the above theorem, and the equivalence of Mosco and Kuratowski convergence, we get $A = \tau_M - \lim A_n$. For each n, we have $A_n^\circ = \{\lambda \mathbf{e}_1 + \lambda \mathbf{e}_n : \lambda \le 1\}$. Clearly, $\langle A_n^\circ \rangle$ fails to Mosco converge to $A^\circ = \{\lambda \mathbf{e}_1 : \lambda \le 1\}$, because $\mathbf{e}_1$ is not the strong limit of a sequence of points taken from $\langle A_n^\circ \rangle$.

REFERENCES

1. H. Attouch, *Variational convergence for functions and operators*, Pitman, Boston, 1984.
2. H. Attouch, R. Lucchetti, and R. Wets, *The topology of the ρ-Hausdorff distance*, (to appear Ann. Mat. Pura. Appl.).
3. H. Attouch and R. Wets, *Quantitative stability of variational systems*: I. *The epigraphical distance*, working paper IIASA, Laxenburg, Austria, 1988.
4. ——, *Quantitative statibility of variational systems*: II. *A framework for nonlinear conditioning*, working paper, IIASA, Laxenburg, Austria, 1988.
5. K. Back, *Continuity of the Fenchel transform of convex functions*, Proc. Amer. Math. Soc. **97** (1986), 661–667.
6. G. Beer, *An embedding theorem for the Fell topology*, Michigan Math. J. **35** (1988), 3–9.
7. ——, *On Mosco convergence of convex sets*, Bull. Australian 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*, Archiv der Math., **52** (1989), 482–491.
10. C. Castaing and M. Valadier, *Convex analysis and measurable multifunctions*, Lecture Notes in Math. #580, Springer-Verlag, Berlin, 1975.

11. S. Dolecki, *Continuity of bilinear and non-bilinear polarities*, in Optimization and related fields (proceedings, Erice, 1984), R. Conti, E. DiGiorgi and F. Giannessi, eds., Lecture Notes in Mathematics #1190, Springer-Verlag, Berlin, 1986.

12. 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.

13. J. Giles, *Convex analysis with application in differentiation of convex functions*, Res. Notes in Math. #58, Pitman, London, 1982.

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

15. 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.

16. R. Lucchetti and F. Patrone, *Hadamard and Tyhonov well-posedness of a certain class of convex functions*, J. Math. Anal. Appl. **88** (1982), 204–215.

17. L. Mclinden and R. Bergstrom, *Preservation of convergence of sets and functions in finite dimensions*, Trans. Amer. Math. Soc. **268** (1981), 127–142.

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

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

20. R. T. Rockafellar, *Convex analysis*, Princeton Univ. Press, Princeton, New Jersey, 1970.

21. G. Salinetti and R. Wets, *On the relation between two types of convergence for convex functions*, J. Math. Anal. Appl. **60** (1977), 211–226.

22. M. Tsukada, *Convergence of best approximations in a smooth Banach space*, J. Approximation Theory **40** (1984), 301–309.

23. D. Walkup and R. Wets, *Continuity of some convex-cone valued mappings*, Proc. Amer. Math. Soc. **18** (1967), 229-235.

24. 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, Los Angeles, California 90064