Semicontinuous functions and convexity

Jordan Bell

April 3, 2014

1 Lattices

If $(A, \leq)$ is a partially ordered set and S is a subset of A, a *supremum of* S is an upper bound that is $\leq$ any upper bound of S, and an *infimum of* S is a lower bound that is $\geq$ any lower bound of S. Because a partial order is antisymmetric, if a supremum exists it is unique, and we denote it by $\bigvee S$, and if an infimum exists it is unique, and we denote it by $\bigwedge S$. If $(A, \leq)$ is a lattice, then one proves by induction that both the supremum and the infimum exist for every finite nonempty subset of A. Vacuously, every element of a partially ordered set is an upper bound for $\emptyset$ and a lower bound for $\emptyset$. Thus, if $\emptyset$ has a supremum x then $x \leq y$ for all $y \in A$, and if $\emptyset$ has an infimum x then $x \geq y$ for all $y \in A$. That is,

$$\bigvee \emptyset = \bigwedge A, \qquad \bigwedge \emptyset = \bigvee A.$$

If X is a set and $R \subseteq [-\infty, \infty]$, the set R^X is partially ordered where $f \leq g$ if $f(x) \leq g(x)$ for all $x \in X$. Moreover, R^X is a lattice:

$$(f \vee g)(x) = \max\{f(x), g(x)\}, \qquad (f \wedge g)(x) = \min\{f(x), g(x)\}.$$

2 Urysohn's lemma

A Hausdorff topological space (X, τ) is said to be *normal* if for every pair of disjoint closed sets E, F there are disjoint open sets U, V with $E \subset U$ and $F \subset V$. Every metrizable space is a normal topological space, but there are normal topological spaces that are not metrizable. A useful fact about normal topological spaces is *Urysohn's lemma*:[1] For each pair of disjoint nonempty closed sets E, F there is a continuous function $f : X \to [0, 1]$ such that $f(E) = 0$ and $f(F) = 1$.

A locally compact Hausdorff space need not be normal. For example, the real numbers with the *rational sequence topology* is Hausdorff and locally compact but is not normal. We say that a topological space is σ*-compact* if it is the union of countably many compact subsets. The following lemma states that if a locally compact Hausdorff space is σ-compact then it is normal.[2]

[1] Gert K. Pedersen, *Analysis Now*, revised printing, p. 24, Theorem 1.5.6.
[2] Gert K. Pedersen, *Analysis Now*, revised printing, p. 39, Proposition 1.7.8.

Lemma 1. *If a locally compact Hausdorff space is σ-compact, then it is normal.*

Urysohn's metrization theorem states that if a topological space is normal and is second-countable (its topology has a countable basis) then it is metrizable. However, a metric space need not be second-countable: a metric space is second-countable precisely when it is separable, and hence the converse of Urysohn's metrization theorem is false. The following lemma shows that a second-countable locally compact Hausdorff space is σ-compact, and hence metrizable by Urysohn's metrization theorem.

Lemma 2. *If a locally compact Hausdorff space is second-countable, then it is σ-compact.*

Proof. Let (X, τ) be a second-countable locally compact Hausdorff space. Because it is second-countable, there is a countable subset $\mathscr{B}$ of τ that is a basis for τ. If $x \in X$, then because X is locally compact there is an open set U containing x which is itself contained in a compact set K, and then there is some $V \in \mathscr{B}$ such that $x \in V \subseteq U$. The closure $\overline{V}$ of V is contained in K and hence $\overline{V}$ is compact. Defining $\mathscr{B}'$ to be those $V \in \mathscr{B}$ such that $\overline{V}$ is compact, it follows that $\mathscr{B}'$ is a basis for τ. The closures of the elements of $\mathscr{B}'$ are countably many compact sets whose union is equal to X, showing that X is σ-compact. $\square$

3 Lower semicontinuous functions

If (X, τ) is a topological space, then $f : X \to [-\infty, \infty]$ is said to be *lower semicontinuous* if $t \in \mathbb{R}$ implies that $f^{-1}(t, \infty] \in \tau$. We say that f is *finite* if $-\infty < f(x) < \infty$ for all $x \in X$.

If $A \subseteq X$ and $t \in \mathbb{R}$, then

$$\chi_A^{-1}(t, \infty] = \begin{cases} X & t < 0 \\ A & 0 \leq t < 1 \\ \emptyset & t \geq 1 \end{cases}$$

We see that the characteristic function of a set is lower semicontinuous if and only if the set is open.

The following theorem characterizes lower semicontinuous functions in terms of nets.[3]

Theorem 3. *If (X, τ) is a topological space and $f : X \to [-\infty, \infty]$ is a function, then f is lower semicontinuous if and only if $(x_\alpha)_{\alpha \in I}$ being a convergent net in X implies that*

$$f(\lim x_\alpha) \leq \liminf f(x_\alpha).$$

[3] Gert K. Pedersen, *Analysis Now*, revised printing, p. 26, Proposition 1.5.11.

Proof. Suppose that f is lower semicontinuous and $x_\alpha \to x$. Say $t < f(x)$. Because f is lower semicontinuous, $f^{-1}(t, \infty] \in \tau$. As $x \in f^{-1}(t, \infty]$ and $x_\alpha \to x$, there is some α_t such that $\alpha \geq \alpha_t$ implies $x_\alpha \in f^{-1}(t, \infty]$. That is, if $\alpha \geq \alpha_t$ then $f(x_\alpha) > t$. This implies that $\liminf f(x_\alpha) \geq t$. But this is true for all $t < f(x)$, hence $\liminf f(x_\alpha) \geq f(x)$.

Suppose that $x_\alpha \to x$ implies that $f(x) \leq \liminf f(x_\alpha)$. Let $t \in \mathbb{R}$ and let $F = f^{-1}(-\infty, t]$. If $x \in \overline{F}$, then there is a net $x_\alpha \in F$ with $x_\alpha \to x$. As $x_\alpha \to x$, by hypothesis $f(x) \leq \liminf f(x_\alpha)$. By definition of F we have $f(x_\alpha) \leq t$ for each α, and hence $f(x) \leq t$. This means that $x \in F$, showing that F is closed and so the complement of F is open. But the complement of F is $f^{-1}(t, \infty]$, showing that f is lower semicontinuous. □

Let $\mathrm{LSC}(X)$ be the set of all lower semicontinuous functions $X \to [-\infty, \infty]$. $\mathrm{LSC}(X)$ is a partially ordered set: $f \leq g$ means that $f(x) \leq g(x)$ for all $x \in X$. The following theorem shows that $\mathrm{LSC}(X)$ is a lattice that contains the supremum of each of its subsets.[4]

Theorem 4. *If (X, τ) is a topological space, then* $\mathrm{LSC}(X)$ *is a lattice, and if $\mathscr{F} \subseteq \mathrm{LSC}(X)$ then $g : X \to [-\infty, \infty]$ defined by*

$$g(x) = \sup_{f \in \mathscr{F}} f(x), \qquad x \in X$$

belongs to $\mathrm{LSC}(X)$*.*

Proof. If $\mathscr{F} = \emptyset$, then g is the constant function $x \mapsto -\infty$, which is lower semicontinuous as for any $t \in \mathbb{R}$, the inverse image of $(t, \infty]$ is $\emptyset$, which belongs to τ. Otherwise, let $t \in \mathbb{R}$. Saying $x \in g^{-1}(t, \infty]$ means that $g(x) > t$, and with $\epsilon = g(x) - t$ there is some some $f \in \mathscr{F}$ satisfying $f(x) > g(x) - \epsilon = t$. Therefore, if $x \in g^{-1}(t, \infty]$ then $x \in \bigcup_{f \in \mathscr{F}} f^{-1}(t, \infty]$. On the other hand, if $x \in \bigcup_{f \in \mathscr{F}} f^{-1}(t, \infty]$, then there is some $f \in \mathscr{F}$ with $f(x) > t$, and hence $g(x) > t$. Therefore

$$g^{-1}(t, \infty] = \bigcup_{f \in \mathscr{F}} f^{-1}(t, \infty].$$

But each $f \in \mathscr{F}$ is lower semicontinuous so the right-hand side is a union of elements of τ, and hence $g^{-1}(t, \infty] \in \tau$, showing that g is lower semicontinuous. Therefore every subset of $\mathrm{LSC}(X)$ has a supremum.

Suppose that $f_1, f_2 \in \mathrm{LSC}(X)$, and define $f : X \to [-\infty, \infty]$ by

$$f(x) = \min\{f_1(x), f_2(x)\}, \qquad x \in X.$$

For $t \in \mathbb{R}$, it is apparent that

$$f^{-1}(t, \infty] = f_1^{-1}(t, \infty] \cap f_2^{-1}(t, \infty].$$

As f_1 and f_2 are each lower semicontinuous, these two inverse images are each open sets, and so their intersection is an open set. Therefore f is lower semicontinuous, showing that $\mathrm{LSC}(X)$ is a lattice. □

[4] Gert K. Pedersen, *Analysis Now*, revised printing, p. 27, Proposition 1.5.12.

One is sometimes interested in lower semicontinuous functions that do not take the value $-\infty$. As the following theorem shows, the sum of two lower semicontinuous functions that do not take the value $-\infty$ is also a lower semicontinuous function.

Theorem 5. *If X is a topological space, if $f, g \in \mathrm{LSC}(X)$, and if $f, g > -\infty$, then $f + g \in \mathrm{LSC}(X)$, and if $r > 0$ and $f \in \mathrm{LSC}(X)$ then $rf \in \mathrm{LSC}(X)$.*

Proof. If $(x_\alpha)_{\alpha \in I}$ is a net that converges to $x \in X$, then, by Theorem 3,

$$(f+g)(x) \leq \liminf f(x_\alpha) + \liminf g(x_\alpha) \leq \liminf f(x_\alpha) + g(x_\alpha) = \liminf (f+g)(x_\alpha),$$

and by Theorem 3 this tells us that $f + g$ is lower semicontinuous. As well,

$$(rf)(x) = rf(x) \leq r \liminf f(x_\alpha) = \liminf rf(x_\alpha) = \liminf (rf)(x_\alpha),$$

showing that $rf \in \mathrm{LSC}(X)$. □

The following theorem shows that if $f_n \in \mathrm{LSC}(X)$ are each finite and converge uniformly on X to $f \in \mathbb{R}^X$, then $f \in \mathrm{LSC}(X)$.[5]

Theorem 6. *If (X, τ) is a topological space, if $f_n \in \mathrm{LSC}(X)$ are finite, and if f_n converge uniformly in X to $f \in \mathbb{R}^X$, then $f \in \mathrm{LSC}(X)$.*

Proof. If $\epsilon > 0$, then there is some N such that $n \geq N$ and $x \in X$ imply that $|f_n(x) - f(x)| < \epsilon$. Define

$$\delta_n = \sup\{|f_n(x) - f(x)| : x \in X\}.$$

Thus, for all $\epsilon > 0$ there is some N such that $n \geq N$ implies that $\delta_n \leq \epsilon$. If $(x_\alpha)_{\alpha \in I}$ is a convergent net in X, then, for all $n \geq N$,

$$\begin{aligned} f(\lim x_\alpha) &\leq \delta_n + f_n(\lim x_\alpha) \\ &\leq \delta_n + \liminf f_n(x_\alpha) \\ &\leq 2\delta_n + \liminf f(x_\alpha) \\ &\leq 2\epsilon + \liminf f(x_\alpha). \end{aligned}$$

This is true for all ϵ, so we get

$$f(\lim x_\alpha) \leq \liminf f(x_\alpha),$$

and therefore f is lower semicontinuous. □

The following theorem shows in particular that on a normal topological space X, any finite nonnegative lower semicontinuous function is the the supremum of the set of all continuous functions $X \to \mathbb{R}$ that are dominated by it.[6] To say that continuous functions $X \to [0, 1]$ *separate points and closed sets* means that if $x \in X$ and F is a disjoint closed set, then there is a continuous function $g : X \to [0, 1]$ such that $g(x) = 1$ and $g(F) = 0$. Urysohn's lemma states that in a normal topological space continuous functions separate closed sets, so in particular they separate points and closed sets.

[5] Gert K. Pedersen, *Analysis Now*, revised printing, p. 27, Proposition 1.5.12.
[6] Gert K. Pedersen, *Analysis Now*, revised printing, p. 27, Proposition 1.5.13.

Theorem 7. *If (X, τ) is a topological space such that continuous functions $X \to [0,1]$ separate points and closed sets, and $f \geq 0$ is a finite lower semicontinuous function on X, then f is the supremum of the set of continuous functions $g : X \to \mathbb{R}$ such that $g \leq f$.*

Proof. Let $\mathscr{M}(f)$ be the set of all continuous functions $g : X \to \mathbb{R}$ with $g \leq f$. As $f \geq 0$ we get $0 \in \mathscr{M}(f)$, and so $\mathscr{M}(f)$ is nonempty. If $x \in X$ and $\epsilon > 0$, let

$$F = f^{-1}(-\infty, f(x) - \epsilon].$$

$X \setminus F = f^{-1}(f(x) - \epsilon, \infty) \in \tau$, so F is closed. It is apparent that $x \notin F$. Therefore there is a continuous function $h : X \to [0,1]$ such that $h(x) = 1$ and $h(F) = 0$. If $f(x) - \epsilon \leq 0$ then, as $h \geq 0$ and $f \geq 0$, we have $(f(x) - \epsilon)h \leq f$. If $f(x) - \epsilon > 0$ and $y \in X$, then

$$(f(x) - \epsilon)h(y) = \begin{cases} 0 & y \in F \\ (f(x) - \epsilon)h(y) & y \notin F \end{cases} \leq \begin{cases} 0 & y \in F \\ f(x) - \epsilon & y \notin F \end{cases} \leq f(y).$$

Therefore $(f(x) - \epsilon)h \leq f$, and because the function $(f(x) - \epsilon)h$ is continuous we have $(f(x) - \epsilon)g \in \mathscr{M}(f)$ Because this is an element of $\mathscr{M}(f)$ we get

$$\left(\bigvee \mathscr{M}(f)\right)(x) \geq (f(x) - \epsilon)g(x) = f(x) - \epsilon.$$

As ϵ was arbitrary it follows that $\left(\bigvee \mathscr{M}(f)\right)(x) \geq f(x)$, and as x was arbitrary we have $\bigvee \mathscr{M}(f) \geq f$. But f is an upper bound for $\mathscr{M}(f)$, so $\bigvee \mathscr{M}(f) \leq f$. Therefore $f = \bigvee \mathscr{M}(f)$. $\square$

The following is a formulation of the extreme value theorem for lower semicontinuous functions on a compact topological space.

Theorem 8 (Extreme value theorem)**.** *If X is a compact topological space and if f is a lower semicontinuous function on X, then*

$$K = \left\{ x \in X : f(x) = \inf_{y \in X} f(y) \right\}$$

is a nonempty closed subset of X.

Proof. Let $C = f(X) \subseteq [-\infty, \infty]$, and for $c \in C$ let $F_c = \{x \in X : f(x) \leq c\}$. Because $F_c = X \setminus f^{-1}(c, \infty]$ and f is lower semicontinuous, F_c is a closed set. Suppose that $c_1, \ldots, c_n \in C$. Taking $c = \min\{c_k : 1 \leq k \leq n\} \in C$, we have

$$\bigcap_{k=1}^{n} F_{c_k} = F_c \neq \emptyset.$$

This shows that the collection $\{F_c : c \in C\}$ has the finite intersection property (the intersection of finitely many members of it is nonempty). But a topological space is compact if and only if for every collection of closed subsets with the finite intersection property the intersection of all the members of the collection is nonempty.[7] Applying this theorem, we get that

$$\bigcap_{c \in C} F_c \neq \emptyset,$$

and this intersection is closed because each member is closed.

Let $x \in \bigcap_{c \in C} F_c$. Then for all $c \in C$ we have $f(x) \leq c$, and because $C = f(X)$ this means that for all $y \in X$ we have $f(x) \leq f(y)$. Therefore $x \in K$. Let $x \in K$. Then for all $y \in X$ we have $f(x) \leq f(y)$, hence for all $c \in C$ we have $f(x) \leq c$, hence $x \in \bigcap_{c \in C} F_c$. Therefore $K = \bigcap_{c \in C} F_c$, which we have shown is nonempty and closed, proving the claim. $\square$

4 Upper semicontinuous functions

If (X, τ) is a topological space, then $f : X \to [-\infty, \infty]$ is said to be *upper semicontinuous* if $t \in \mathbb{R}$ implies that $f^{-1}[-\infty, t) \in \tau$. We denote by $\mathrm{USC}(X)$ the set of upper semicontinuous functions $X \to [-\infty, \infty]$. We say that f is *finite* if $-\infty < f(x) < \infty$ for all $x \in X$.

If $A \subseteq X$ and $t \in \mathbb{R}$, then

$$\chi_A^{-1}[t, \infty) = \begin{cases} X & t \leq 0 \\ A & 0 < t \leq 1 \\ \emptyset & t > 1 \end{cases}$$

But $\chi_A^{-1}[t, \infty)$ is closed if and only if $\chi_A^{-1}[-\infty, t)$ is open, hence χ_A is upper semicontinuous if and only if A is closed.

It is apparent that $f : X \to [-\infty, \infty]$ is upper semicontinuous if and only if $-f : X \to [-\infty, \infty]$ is lower semicontinuous. Because the set of all open intervals are a basis for the topology of $\mathbb{R}$, a function $X \to \mathbb{R}$ is continuous if and only if it is both lower semicontinuous and upper semicontinuous. That is,

$$C(X) = \mathbb{R}^X \cap \mathrm{LSC}(X) \cap \mathrm{USC}(X).$$

5 Approximating integrable functions

If X is a Hausdorff space, if $\mathfrak{M}$ is a σ-algebra on X that contains the Borel σ-algebra of X (equivalently, if every open set belongs to $\mathfrak{M}$), and if μ is a measure on $\mathfrak{M}$, we say that μ is *outer regular on* $E \in \mathfrak{M}$ if

$$\mu(E) = \inf\{\mu(V) : E \subseteq V \text{ and } V \text{ is open}\},$$

and we say that μ is *inner regular on* E if

$$\mu(E) = \sup\{\mu(K) : K \subseteq E \text{ and } K \text{ is compact}\}.$$

[7] James Munkres, *Topology*, second ed., p. 169, Theorem 26.9.

We state the following to motivate the conditions in Theorem 9. If X is a locally compact Hausdorff space and λ is a positive linear functional on $C_c(X)$ ($f \geq 0$ implies that $\lambda f \geq 0$), then the *Riesz-Markov theorem*[8] states that there is a σ-algebra $\mathfrak{M}$ on X that contains the Borel σ-algebra of X and there is a unique complete measure μ on $\mathfrak{M}$ that satisfies:

1. If $f \in C_c(X)$ then $\lambda f = \int_X f d\mu$.
2. If K is compact then $\mu(K) < \infty$.
3. μ is outer regular on all $E \in \mathfrak{M}$
4. μ is inner regular on all open sets and on all sets with finite measure.

The following theorem gives conditions under which we can bound an integrable function above and below by semicontinuous functions that can be chosen as close as we please in L^1 norm.[9]

Theorem 9 (Vitali-Carathéodory theorem)**.** *Let X be a locally compact Hausdorff space, let $\mathfrak{M}$ be a σ-algebra containing the Borel σ-algebra of X, and let μ be a complete measure on $\mathfrak{M}$ that satisfies $\mu(K) < \infty$ for compact K, that is outer regular on all measurable sets, and that is inner regular on open sets and on sets with finite measure. If $f \in L^1(\mu)$ is real valued and if $\epsilon > 0$, then there is some upper semicontinuous function u that is bounded above and some lower semicontinuous function v that is bounded below such that $u \leq f \leq v$ and such that*

$$\int_X (v-u) d\mu < \epsilon.$$

Proof. Let $g \in L^1(\mu)$ be ≥ 0 and let $\epsilon > 0$. There is a nondecreasing sequence of measurable simple functions s_n such that for all $x \in X$ we have $g(x) = \lim_{n\to\infty} s_n(x)$.[10] Writing $s_0 = 0$ and $t_n = s_n - s_{n-1}$, each t_n is a measurable simple function and is ≥ 0. Then, there are some $c_i \geq 0$ and measurable sets E_i such that for all $x \in X$ we have

$$g(x) = \lim_{n\to\infty} \sum_{i=1}^n t_i(x) = \sum_{i=1}^\infty t_i(x) = \sum_{i=1}^\infty c_i \chi_{E_i}(x).$$

Integrating this we get[11]

$$\int_X g d\mu = \sum_{i=1}^\infty \int_X c_i \chi_{E_i} d\mu = \sum_{i=1}^\infty c_i \mu(E_i).$$

[8] Walter Rudin, *Real and Complex Analysis*, third ed., p. 40, Theorem 2.14.

[9] Walter Rudin, *Real and Complex Analysis*, third ed., p. 56, Theorem 2.25.

[10] Walter Rudin, *Real and Complex Analysis*, third ed., p. 15, Theorem 1.17. A *simple function* is a finite linear combination of characteristic functions, either over $\mathbb{R}$ or $\mathbb{C}$.

[11] Walter Rudin, *Real and Complex Analysis*, third ed., p. 22, Theorem 1.27.

As $g \in L^1(\mu)$ the left-hand side is finite and so the right-hand side is too, hence there is then some N such that $\sum_{i=N+1}^{\infty} c_i \mu(E_i) < \frac{\epsilon}{2}$.

For each i, because μ is outer regular on E_i there is an open set V_i containing E_i such that $c_i \mu(V_i \setminus E_i) < 2^{-i-2}\epsilon$. Each E_i has finite measure so μ is inner regular on E_i, hence there is a compact set K_i contained in E_i such that $c_i \mu(E_i \setminus K_i) < 2^{-i-2}\epsilon$. Define

$$v = \sum_{i=1}^{\infty} c_i \chi_{V_i}, \qquad u = \sum_{i=1}^{N} c_i \chi_{K_i}.$$

Each V_i is open so the characteristic function χ_{V_i} is lower semicontinuous, and $c_i \geq 0$ so each of the functions $\sum_{i=1}^{n} c_i \chi_{V_i}$ is a sum of finitely many lower semicontinuous functions and hence is lower semicontinuous. But v is the supremum of the functions $\sum_{i=1}^{n} c_i \chi_{V_i}$, so v is lower semicontinuous. As each K_i is closed the characteristic function χ_{K_i} is upper semicontinuous, and $c_i \geq 0$ so u is a sum of finitely many upper semicontinuous functions and hence is upper semicontinuous. u is a finite sum so is bounded above, and v is a sum of nonnegative terms so is bounded below by 0.

Because $K_i \subseteq E_i \subseteq V_i$ we have $u \leq g \leq v$, and

$$\begin{aligned}
v - u &= \sum_{i=1}^{\infty} c_i \chi_{V_i} - \sum_{i=1}^{N} c_i \chi_{K_i} \\
&= \sum_{i=1}^{N} c_i (\chi_{V_i} - \chi_{K_i}) + \sum_{i=N+1}^{\infty} c_i \chi_{V_i} \\
&\leq \sum_{i=1}^{\infty} c_i (\chi_{V_i} - \chi_{K_i}) + \sum_{i=N+1}^{\infty} c_i \chi_{E_i};
\end{aligned}$$

the inequality is because $\chi_{V_i} - \chi_{K_i} + \chi_{E_i} \geq \chi_{V_i}$. Integrating,

$$\begin{aligned}
\int_X (v-u) d\mu &\leq \sum_{i=1}^{\infty} c_i \mu(V_i \setminus K_i) + \sum_{i=N+1}^{\infty} c_i \mu(E_i) \\
&= \sum_{i=1}^{\infty} (c_i \mu(V_i \setminus E_i) + c_i \mu(E_i \setminus V_i)) + \sum_{i=N+1}^{\infty} c_i \mu(E_i) \\
&\leq \sum_{i=1}^{\infty} (2^{-i-2}\epsilon + 2^{-i-2}\epsilon) + \frac{\epsilon}{2} \\
&= \epsilon.
\end{aligned}$$

Let $f = f^+ - f^-$, for $f^+, f^- \in L^1(\mu)$ with $f^+, f^- \geq 0$, and let $\epsilon > 0$. From what we have established above, there is an upper semicontinuous function u_1 that is bounded above and a lower semicontinuous function v_1 that is bounded below satisfying $u_1 \leq f^+ \leq v_1$ and

$$\int_X (v_1 - u_1) d\mu < \frac{\epsilon}{2},$$

and similarly u_2, v_2 with $u_2 \leq f^- \leq v_2$ and

$$\int_X (v_2 - u_2) d\mu < \frac{\epsilon}{2}.$$

We have

$$u_1 - v_2 \leq f^+ - f^- = f = f^+ - f^- \leq v_1 - u_2.$$

That v_2 is lower semicontinuous and bounded below means that $-v_2$ is upper semicontinuous and bounded above, hence $u_1 - v_2$ is upper semicontinuous and bounded above. That u_2 is upper semicontinuous and bounded above means that $-u_2$ is lower semicontinuous and bounded below, so $v_1 - u_2$ is lower semicontinuous and bounded below. Taking $u = u_1 - v_2$ and $v = v_1 - u_2$, we have $u \leq f \leq v$, and

$$\int_X (v-u) d\mu = \int_X (v_1 - u_2 - u_1 + v_2) d\mu = \int_X (v_1 - u_1) d\mu + \int_X (v_2 - u_2) d\mu < \epsilon.$$

□

6 Convex functions

If X is a set and $f : X \to [-\infty, \infty]$ is a function, its *epigraph* is the set

$$\operatorname{epi} f = \{(x, \alpha) \in X \times \mathbb{R} : \alpha \geq f(x)\}.$$

When X is a vector space, we say that f is *convex* if $\operatorname{epi} f$ is a convex subset of the vector space $X \times \mathbb{R}$. The *effective domain* of a convex function f is the set

$$\operatorname{dom} f = \{x \in X : f(x) < \infty\}.$$

To say that $x \in \operatorname{dom} f$ is to say that there is some $\alpha \in \mathbb{R}$ such that $(x, \alpha) \in \operatorname{epi} f$, from which it follows that if $f : X \to [-\infty, \infty]$ is a convex function then $\operatorname{dom} f$ is a convex subset of X. A convex function f is said to be *proper* if $\operatorname{dom} f \neq \emptyset$ and $f(x) > -\infty$ for all $x \in X$, i.e. if f does not only take the value ∞ and never takes the value $-\infty$.

If C is a nonempty convex subset of X and $f : C \to \mathbb{R}$ is a function, we extend f to X by defining $f(x) = \infty$ for $x \notin C$. One checks that this extension is a convex function if and only if

$$f((1-t)x + ty) \leq (1-t)f(x) + tf(y), \qquad x, y \in C, \quad 0 < t < 1,$$

and we call $f : C \to \mathbb{R}$ convex if $f : X \to (-\infty, \infty]$ is convex. If this extension is convex, then it has effective domain C and is proper.

The following lemma is straightforward to prove.[12]

[12] Charalambos D. Aliprantis and Kim C. Border, *Infinite Dimensional Analysis: A Hitchhiker's Guide*, third ed., p. 187, Lemma 5.41.

Lemma 10. *If X is a vector space, if C is a convex subset of X, if $f : C \to \mathbb{R}$ is convex, if $x, x+z, x-z \in C$, and if $0 \leq \delta \leq 1$, then*

$$|f(x+\delta z) - f(x)| \leq \delta \max\{f(x+z) - f(x), f(x-z) - f(x)\}.$$

The following lemma asserts that a convex function that is bounded above on some neighborhood of an interior point of a convex subset of a topological vector space is continuous at that point.[13]

Lemma 11. *If X is a topological vector space, if C is a convex subset of X, if $f : C \to \mathbb{R}$ is convex, if x is in the interior of C, and if f is bounded above on some neighborhood of x, then f is continuous at x.*

Proof. There is some neighborhood of x contained in C on which f is bounded above. Thus, there is some open neighborhood U of the origin such that $x+U \subseteq C$ and such that f is bounded above on $x+U$. Being bounded above means that there is some M such that $y \in x+U$ implies that $f(y) \leq M$. Any open neighborhood of 0 contains a balanced open neighborhood of 0,[14] so let V be a balanced open neighborhood of 0 contained in U.

Let $\epsilon > 0$ and take $\delta > 0$ small enough that $\delta(M - f(x)) < \epsilon$. For $y \in x+\delta V$, there is some $z \in V$ with $y = x + \delta z$, and because V is balanced we have $x+z, x-z \in x+V$. Then we can apply Lemma 10 to get

$$|f(y) - f(x)| \leq \delta \max\{f(x+z) - f(x), f(x-z) - f(x)\} \leq \delta(M - f(x)) < \epsilon.$$

But $x + \delta V$ is an open neighborhood of x (because scalar multiplication is continuous), so we have shown that if $\epsilon > 0$ then there is some open neighborhood of x such that y being in this neighborhood implies that $|f(y) - f(x)| < \epsilon$. This means that f is continuous at x. □

The following theorem shows that properties that by themselves are weaker than continuity on a set are equivalent to it for a convex function on an open convex set.[15] We prove three of the five implications because two are immediate.

Theorem 12. *If X is a topological vector space, if C is an open convex subset of X, and if $f : C \to \mathbb{R}$ is convex, then the following are equivalent:*

1. *f is continuous on C.*
2. *f is upper semicontinuous on C.*
3. *For each $x \in C$ there is some neighborhood of x on which f is bounded above.*

[13] Charalambos D. Aliprantis and Kim C. Border, *Infinite Dimensional Analysis: A Hitchhiker's Guide*, third ed., p. 188, Theorem 5.42.

[14] Walter Rudin, *Functional Analysis*, second ed., p. 12, Theorem 1.14. For a set V to be *balanced* means that $|\alpha| \leq 1$ implies that $\alpha V \subseteq V$.

[15] Charalambos D. Aliprantis and Kim C. Border, *Infinite Dimensional Analysis: A Hitchhiker's Guide*, third ed., p. 188, Theorem 5.43.

4. *There is some $x \in C$ and some neighborhood of x on which f is bounded above.*

5. *There is some $x \in C$ at which f is continuous.*

Proof. Suppose that f is upper semicontinuous on C, and say $x \in C$. Because f is upper semicontinuous and C is open, the set

$$U = \{y \in C : f(y) < f(x) + 1\} = C \cap f^{-1}(-\infty, f(x) + 1)$$

is open. $x \in U$ because $f(x) < f(x) + 1$, so U is a neighborhood of x, and f is bounded on U.

Suppose that $x \in C$ and U is a neighborhood of x on which f is bounded above. Lemma 11 states that f is continuous at x, showing that f is continuous at some point in C.

Suppose that f is continuous at some $x \in C$, and let y be another point in C. The function $t \mapsto x + t(y - x)$ is continuous $\mathbb{R} \to X$, and because it sends 1 to $y \in C$ and C is open, there is some $t > 1$ such that $x + t(y - x) \in C$. (That is, the line segment from x to y remains in C for some length past y.) Set $z = x + t(y - x)$, i.e. $z = (1 - t)x + ty$, i.e.

$$y = \left(1 - \frac{1}{t}\right) x + \frac{1}{t} z,$$

or

$$y = \lambda x + (1 - \lambda)z,$$

where $\lambda = 1 - \frac{1}{t}$ satisfies $0 < \lambda < 1$. f being continuous at x means that for every $\epsilon > 0$ there is some open neighborhood V of 0 such that $y \in x + V$ implies that $|f(y) - f(x)| < \epsilon$. Take $x + V \subseteq C$, which we can do because C is open. In particular, there is some M such that $f(w) \leq M$ for $w \in x + V$. If $v \in V$, then

$$y + \lambda v = \lambda x + (1 - \lambda)z + \lambda v = \lambda(x + v) + (1 - \lambda)z.$$

$x+v \in C$ and $z \in C$, and because C is convex this tells us $y+\lambda v \in C$. Therefore $y + \lambda V \subseteq C$. Because f is convex we have

$$f(y+\lambda v) = f(\lambda(x+v)+(1-\lambda)z) \leq \lambda f(x+v)+(1-\lambda)f(z) \leq \lambda M+(1-\lambda)f(z).$$

This holds for every $v \in V$, so f is bounded above by $\lambda M + (1 - \lambda)f(z)$ on $y + \lambda V$. $y + \lambda V$ is an open neighborhood of y on which f is bounded above, so we can apply Lemma 11, which tells us that f is continuous at y. Since f is continuous at each point in C, it is continuous on C. □

7 Convex hulls

If X is a topological vector space over $\mathbb{R}$, let X^* denote the set of continuous linear maps $X \to \mathbb{R}$. X^* is called the *dual space of* X and is a vector space. We call the bilinear map $\langle \cdot, \cdot \rangle : X \times X^* \to \mathbb{R}$ defined by

$$\langle x, \lambda \rangle = \lambda x, \qquad x \in X, \lambda \in X^*$$

the *dual pairing* of X and X^*. If X is locally convex, it follows from the Hahn-Banach separation theorem[16] that for distinct $x, y \in X$ there is some $\lambda \in X^*$ such that $\lambda x \neq \lambda y$. The *weak topology* on X is the initial topology for the set of functions X^*, and we denote the vector space X with the weak topology by X_w. X_w is a locally convex space whose dual space is X^*.[17]

If X is a vector space and E is a subset of X, the *convex hull* of E is the set of all convex combinations of finitely many points in E and is denoted by $\operatorname{co} E$. The convex hull $\operatorname{co} E$ is a convex set, and it is straightforward to prove that $\operatorname{co} E$ is equal to the intersection of all convex sets containing E. If X is a topological vector space, the *closed convex hull* of E is the closure of the convex hull $\operatorname{co} E$ and is denoted by $\overline{\operatorname{co}}\, E$. One proves that the closed convex hull $\overline{\operatorname{co}}\, E$ is equal to the intersection of all closed convex sets containing E.

A *closed half-space* in a locally convex space X over $\mathbb{R}$ is a set of the form $\{x \in X : \langle x, \lambda \rangle \leq \beta\}$, for $\beta \in \mathbb{R}$ and $\lambda \in X^*$ with $\lambda \neq 0$. (If X is merely a topological vector space then it may be the case that $X^* = \{0\}$, for example $L^p[0,1]$ for $0 < p < 1$.) If $\langle x, \lambda \rangle \leq \beta$, $\langle y, \lambda \rangle \leq \beta$, and $0 \leq t \leq 1$, then

$$\langle (1-t)x + ty, \lambda \rangle = (1-t)\langle x, \lambda \rangle + t \langle y, \lambda \rangle \leq (1-t)\beta + t\beta = \beta,$$

so a closed half-space is convex.

Lemma 13. *If X is a locally convex space over $\mathbb{R}$ and E is a subset of X, then $\overline{\operatorname{co}}\, E$ is the intersection of all closed half-spaces containing E.*

Proof. If $E = \emptyset$ then $\overline{\operatorname{co}}\, E = \emptyset$. But every closed half-space contains $\emptyset$ and the intersection of all of these is also $\emptyset$, so the claim is true in this case. If $\overline{\operatorname{co}}\, E = X$ then there are no closed half-spaces that contain E, and as an intersection over an empty index set is equal to the universe which is X here, so the claim is true in this case also. Otherwise, $\overline{\operatorname{co}}\, E \neq \emptyset, X$, and let $a \notin \overline{\operatorname{co}}\, E$. Because $\{a\}$ is a compact convex set and $\overline{\operatorname{co}}\, E$ is a disjoint nonempty closed convex set, we can apply the Hahn-Banach separation theorem,[18] which tells us that there is some $\lambda_a \in X^*$ and some $\gamma_a \in \mathbb{R}$ such that

$$\lambda_a x < \gamma_a < \lambda_a a, \qquad x \in \overline{\operatorname{co}}\, E.$$

$\overline{\operatorname{co}}\, E$ is contained in the closed half-space $\{x \in X : \langle x, \lambda_a \rangle \leq \gamma_a\}$ and a is not. Hence

$$\overline{\operatorname{co}}\, E = \bigcap_{a \notin \overline{\operatorname{co}}\, E} \{x \in X : \langle x, \lambda_a \rangle \leq \gamma_a\}.$$

This shows that $\overline{\operatorname{co}}\, E$ is equal to an intersection of closed half-spaces containing E. Since a closed half-space is closed and convex and $\overline{\operatorname{co}}\, E$ is the intersection of all closed convex sets containing E, it follows that $\overline{\operatorname{co}}\, E$ is equal to the intersection of all closed half-spaces containing E. □

[16] Walter Rudin, *Functional Analysis*, second ed., p. 59, Theorem 3.4.
[17] Walter Rudin, *Functional Analysis*, second ed., p. 64, Theorem 3.10.
[18] Walter Rudin, *Functional Analysis*, second ed., p. 59, Theorem 3.4.

If X is a vector convex space and $f : X \to [-\infty, \infty]$ is a function, the *convex hull of* f is the function $\operatorname{co} f : X \to [-\infty, \infty]$ defined by

$$\operatorname{co} f = \bigvee \{g \in [-\infty, \infty]^X : g \text{ is convex and } g \le f\}.$$

We have $\operatorname{co} f \le f$. The following lemma shows that the supremum of a set of convex functions is itself a convex function (because an intersection of convex sets is itself a convex set and for a function to be convex means that its epigraph is a convex set), and hence that the convex hull of a function is itself a convex function.

Lemma 14. *If X is a set and $\mathscr{F} \subseteq [-\infty, \infty]^X$, then $F = \bigvee \mathscr{F}$ satisfies*

$$\operatorname{epi} F = \bigcap_{f \in \mathscr{F}} \operatorname{epi} f.$$

Proof. If $\mathscr{F} = \emptyset$, then $\bigvee \mathscr{F}$ is the function $x \mapsto -\infty$, whose epigraph is $X \times \mathbb{R}$. The intersection over the empty set of subsets of $X \times \mathbb{R}$ is equal to the universe $X \times \mathbb{R}$, so the claim holds in this case. Otherwise, let $F = \bigvee \mathscr{F}$, which satisfies

$$F(x) = \sup\{f(x) : f \in \mathscr{F}\}, \qquad x \in X.$$

Suppose that $(x, \alpha) \in \operatorname{epi} F$. This means that $F(x) \le \alpha$, so $f(x) \le \alpha$ for all $f \in \mathscr{F}$, hence $(x, \alpha) \in \operatorname{epi} f$ for all $f \in \mathscr{F}$, and therefore

$$(x, \alpha) \in \bigcap_{f \in \mathscr{F}} \operatorname{epi} f.$$

Suppose that $x \in X$ and $\alpha \in \mathbb{R}$ and that $(x, \alpha) \in \operatorname{epi} f$ for all $f \in \mathscr{F}$. This means that $f(x) \le \alpha$ for all $f \in \mathscr{F}$, hence $F(x) \le \alpha$ and so $(x, \alpha) \in \operatorname{epi} F$. □

Lower semicontinuity of a function can also be expressed using the notion of epigraphs.

Lemma 15. *If X is a topological space and $f : X \to [-\infty, \infty]$ is a function, then f is lower semicontinuous if and only if* $\operatorname{epi} f$ *is a closed subset of* $X \times \mathbb{R}$.

Proof. Suppose that f is lower semicontinuous and let $(x_i, \alpha_i) \in \operatorname{epi} f$ be a net that converges to $(x, \alpha) \in X \times \mathbb{R}$. Then $x_i \to x$ and $\alpha_i \to \alpha$, and Theorem 3 gives us

$$f(x) \le \liminf f(x_i) \le \liminf \alpha_i = \lim \alpha_i = \alpha.$$

Hence $f(x) \le \alpha$, which means that $(x, \alpha) \in \operatorname{epi} f$. Therefore $\operatorname{epi} f$ is closed.

Suppose that $\operatorname{epi} f$ is closed and let $t \in \mathbb{R}$. The set

$$\operatorname{epi} f \cap (X \times \{t\}) = \{(x, t) : x \in X, t \ge f(x)\}$$

is a closed subset of $X \times \mathbb{R}$. This implies that $f^{-1}[-\infty, t]$ is a closed subset of X, which is equivalent to $f^{-1}(t, \infty]$ being an open subset of X. This is true for all $t \in \mathbb{R}$, so f is lower semicontinuous. □

The following lemma shows that if a convex lower semicontinuous function takes the value $-\infty$ then it is nowhere finite. This means that if there is some point at which a convex lower semicontinuous function takes a finite value then it does not take the value $-\infty$, namely, if a convex lower semicontinuous function takes a finite value at some point then it is proper.

Lemma 16. *If X is a topological vector space, if $f : X \to [-\infty, \infty]$ is convex and lower semicontinuous, and if there is some $x_0 \in X$ such that $f(x_0) = -\infty$, then $f(x) \in \{-\infty, \infty\}$ for all $x \in X$.*

Proof. Suppose by contradiction that there is some $x \in X$ such that $-\infty < f(x) < \infty$. Because f is convex, for every $\lambda \in (0, 1]$ we have

$$f((1-\lambda)x + \lambda x_0) \leq (1-\lambda)f(x) + \lambda f(x_0) = \text{finite} - \infty = -\infty,$$

hence $f((1-\lambda)x + \lambda x_0) = -\infty$ for all $\lambda \in (0, 1]$. Because f is lower semicontinuous,

$$f\Big(\lim_{\lambda\to 0}((1-\lambda)x + \lambda x_0)\Big) \leq \liminf_{\lambda \to 0} f((1-\lambda)x + \lambda x_0),$$

and hence

$$f(x) \leq -\infty,$$

a contradiction. Therefore there is no $x \in X$ such that $-\infty < f(x) < \infty$. $\square$

If X is a topological space and $f : X \to [-\infty, \infty]$ is a function, the *lower semicontinuous hull of f* is the function $\operatorname{lsc} f : X \to [-\infty, \infty]$ defined by

$$\operatorname{lsc} f = \bigvee \{g \in \mathrm{LSC}(X) : g \leq f\}.$$

By Theorem 4, $\operatorname{lsc} f \in \mathrm{LSC}(X)$. It is apparent that a function is lower semicontinuous if and only if it is equal to its lower semicontinuous hull. The following lemma shows that the epigraph of the lower semicontinuous hull of a function is equal to the closure of its epigraph.[19]

Lemma 17. *If X is a topological space and $f : X \to [-\infty, \infty]$ is a function, then*

$$\operatorname{epi} \operatorname{lsc} f = \overline{\operatorname{epi} f}$$

Proof. Check that $\overline{\operatorname{epi} f} = \operatorname{epi} g$ for $g : X \to [-\infty, \infty]$ defined by[20]

$$g(x) = \liminf_{y \to x} f(y), \qquad x \in X,$$

and that $\operatorname{lsc} f = g$. $\square$

[19] Jean-Paul Penot, *Calculus Without Derivatives*, p. 18, Proposition 1.21.

[20] Let $\mathscr{N}(x)$ be the neighborhood filter at x. $\liminf_{y \to x} f(y)$ is defined to be

$$\sup_{N \in \mathscr{N}(x)} \inf_{y \in N \setminus \{x\}} f(y).$$

The notion of being a convex function applies to functions on a vector space, and the notion of being lower semicontinuous applies to functions on a topological space. The following theorem shows that the lower semicontinuous hull of a convex function on a topological vector space is convex.[21]

Theorem 18. *If X is a topological vector space and $f : X \to [-\infty, \infty]$ is convex, then* $\operatorname{lsc} f$ *is convex.*

Proof. For $\operatorname{lsc} f$ to be a convex function means that $\operatorname{epi} \operatorname{lsc} f$ is a convex set. But Lemma 17 tells us that $\operatorname{epi} \operatorname{lsc} f = \overline{\operatorname{lsc} f}$. As f is convex, the epigraph $\operatorname{epi} f$ is convex, and the closure of a convex set is convex.[22] Hence $\operatorname{epi} \operatorname{lsc} f$ is convex, and so $\operatorname{lsc} f$ is a convex function. □

8 Extreme points

If X is a vector space and C is a subset of X, a nonempty subset of S of C is called an *extreme set of* C if $x, y \in C, 0 < t < 1$ and $(1-t)x + ty \in S$ together imply that $x, y \in S$. An *extreme point of* C is an element x of C such that the singleton $\{x\}$ is an extreme set of C. The set of extreme points of C is denoted by $\operatorname{ext} C$. If C is a convex set and S is an extreme set of C that is itself convex, then S is called a *face of* C.

Lemma 19. *If X is a vector space, if C is a convex subset of X, and if $a \in C$, then a is an extreme point of C if and only if $C \setminus \{a\}$ is a convex set.*

Proof. Suppose that a is an extreme point of C and let $x, y \in C \setminus \{a\}$ and $0 < t < 1$. Because C is convex, $(1-t)x + ty \in C$. If $(1-t)x + ty = a$, then because a is an extreme point of C we would have $x = a$ and $y = a$, contradicting $x, y \in C \setminus \{a\}$. Therefore $(1-t)x + ty \in C \setminus \{a\}$.

Suppose that $C \setminus \{a\}$ is a convex set. Suppose that $x, y \in C$, $0 < t < 1$, and $(1-t)x + ty = a$. Assume by contradiction that $x \neq a$. If $y = a$ then we get $(1-t)x + ta = a$, or $(1-t)x = (1-t)a$, hence $x = a$, a contradiction. If $y \neq a$, then using that $C \setminus \{a\}$ is convex, we have $(1-t)x + ty \in C \setminus \{a\}$, contradicting that $(1-t)x + ty = a$. Therefore $x = a$. We similarly show that $y = a$. Therefore a an extreme point of C. □

The following lemma is about the set of maximizers of a convex function, and does not involve a topology on the vector space.[23]

Lemma 20. *If X is a vector space, if C is a convex susbset of X, and if $f : C \to \mathbb{R}$ is convex, then*

$$F = \left\{ x \in C : f(x) = \sup_{y \in C} f(y) \right\}$$

is either an extreme set of C or is empty.

[21] R. Tyrrell Rockafellar, *Conjugate Duality and Optimization*, p. 15, Theorem 4.

[22] Walter Rudin, *Functional Analysis*, second ed., p. 11, Theorem 1.13.

[23] Charalambos D. Aliprantis and Kim C. Border, *Infinite Dimensional Analysis: A Hitchhiker's Guide*, third ed., p. 296, Lemma 7.64.

Proof. Suppose that there is some x_0 at which f is maximized, i.e. that F is nonempty, and let $M = f(x_0)$. Suppose that $x, y \in C$, $0 < t < 1$, and $(1-t)x+ty \in F$. If at least one of x, y do not belong to F, then, as $(1-t)x+ty \in F$ and as f is convex,

$$M = f((1-t)x+ty) \leq (1-t)f(x) + tf(y) < (1-t)M + tM = M,$$

a contradiction. Therefore $x, y \in F$, showing that F is an extreme set of C. $\square$

Elements of an extreme set need not be extreme points, but the following lemma shows that in a locally convex space if an extreme set is compact then it contains an extreme point.[24]

Lemma 21. *If X is a real locally convex space, if C is a subset of X, and if F is a compact extreme set of C, then*

$$F \cap \operatorname{ext} C \neq \emptyset.$$

Proof. Define $\mathscr{F} = \{G \subseteq F : G \text{ is a compact extreme set of } C\}$. $F \in \mathscr{F}$ so $\mathscr{F}$ is nonempty. $\mathscr{F}$ is a partially ordered set ordered by set inclusion. If T is a chain in $\mathscr{F}$, then the intersection of finitely many elements of T is equal to the minimum of these elements which is an extreme set and hence is nonempty. Therefore the chain T has the finite intersection property, and because the elements of T are closed subsets of F and F is compact, the intersection of all the elements of T is nonempty.[25] One checks that this intersection belongs to $\mathscr{F}$ (one must verify that it is an extreme set of C) and is a lower bound for T. We have shown that every chain in $\mathscr{F}$ has a lower bound in $\mathscr{F}$, and applying Zorn's lemma, there is a minimal element G in $\mathscr{F}$ (if an element of $\mathscr{F}$ is contained in G then it is equal to G).

Assume by contradiction that there are $a, b \in G$ with $a \neq b$. Then there is some $\lambda \in X^*$ such that $\lambda a > \lambda b$. G is compact and λ is continuous, so

$$G_0 = \left\{ c \in G : \lambda c = \sup_{y \in G} \lambda y \right\}$$

is nonempty and closed. We shall show that G_0 is an extreme set of G. Suppose that $x, y \in G$, $0 < t < 1$, and $(1-t)x + ty \in G_0$. Let $M = \sup_{y \in G} \lambda y$. If at least one of x, y do not belong to G_0, then

$$M = \lambda((1-t)x+ty) = (1-t)\lambda x + t\lambda y < (1-t)M + tM = M,$$

a contradiction. Hence $x, y \in G_0$, showing that G_0 is an extreme set of G. Check that G_0 being an extreme set of G implies that G_0 is an extreme set

[24] Charalambos D. Aliprantis and Kim C. Border, *Infinite Dimensional Analysis: A Hitchhiker's Guide*, third ed., p. 296, Lemma 7.65.
[25] James Munkres, *Topology*, second ed., p. 169, Theorem 26.9.

of C. Then $G_0 \in \mathscr{F}$, but as $\lambda a > \lambda b$ we have $b \notin G_0$, so that G_0 is strictly contained in G, contradicting that G is a minimal element of $\mathscr{G}$. Therefore G has a single element, as G being an extreme set means that it is nonempty. But G is an extreme set of C, and since G is a singleton this means that the single point it contains is an extreme point of C. $\square$

The following theorem gives conditions under which a function on a set has a maximizer that is an extreme point of the set.[26]

Theorem 22 (Bauer maximum principle)**.** *If X is a real locally convex space, if C is a compact convex subset of X, and if $f : C \to \mathbb{R}$ is upper semicontinuous, then there is a maximizer of f that belongs to* $\operatorname{ext} C$*.*

Proof. Because f is upper semicontinuous and C is compact, it follows from Theorem 8 that

$$F = \left\{ x \in C : f(x) = \sup_{y \in C} f(y) \right\}$$

is a nonempty closed subset of C. Since C is convex and f is a convex function, by Lemma 20, F is an extreme set of C. F is a closed subset of the compact set C, so F is compact. Hence F is a compact extreme set of C, and by Lemma 21 there is an extreme point in F, which was the claim. $\square$

9 Duality

Lemma 23. *A convex function on a real locally convex space is lower semicontinuous if and only if it is weakly lower semicontinuous.*

Proof. Let X be a locally convex space and let X_w denote this vector space with the weak topology, with which it is a locally convex space. As $\mathbb{R}$ is a locally convex space, the product $X \times \mathbb{R}$ is a locally convex space, and one checks that $X \times \mathbb{R}$ with the weak topology is $X_w \times \mathbb{R}$. Thus, to say that a subset of $X \times \mathbb{R}$ is weakly closed is equivalent to saying that it is closed in $X_w \times \mathbb{R}$. Furthermore, the closure of a convex set in a locally convex space is equal to its weak closure.[27] In particular, a convex subset of a locally convex space is closed if and only if it is weakly closed. Therefore, a convex subset of $X \times \mathbb{R}$ is closed if and only if it is closed in $X_w \times \mathbb{R}$.

By Lemma 15, a function $f : X \to [-\infty, \infty]$ is lower semicontinuous if and only if $\operatorname{epi} f$ is a closed subset of $X \times \mathbb{R}$, and is weakly lower semicontinuous if and only if $\operatorname{epi} f$ is a closed subset of $X_w \times \mathbb{R}$. $\square$

A topological vector space is said to have the *Heine-Borel property* if every closed and bounded subset of it is compact. The following theorem gives conditions under which a function is minimized on a set that is not necessarily compact but which is convex, closed, and bounded.

[26] Charalambos D. Aliprantis and Kim C. Border, *Infinite Dimensional Analysis: A Hitchhiker's Guide*, third ed., p. 298, Theorem 7.69.
[27] Walter Rudin, *Functional Analysis*, second ed., p. 66, Theorem 3.12.

Theorem 24. *If X is a real locally convex space such that X_w has the Heine-Borel property, if $f : X \to [-\infty, \infty]$ is a lower semicontinuous convex function, and if C is a convex closed bounded subset of X, then*

$$K = \left\{ x \in C : f(x) = \inf_{y \in C} f(y) \right\}$$

is a nonempty closed subset of X.

Proof. Because X is a locally convex space and C is convex, the fact that C is closed implies that it is weakly closed.[28] It is straightforward to prove that if a subset of a topological vector space is bounded then it is weakly bounded.[29] Thus, C is weakly closed and weakly bounded, and because X_w has the Heine-Borel property we get that C is weakly compact. In other words, C_w is compact, where C_w is the set C with the subspace topology inherited from X_w. By Lemma 23, f is weakly lower semicontinuous, i.e. $f : X_w \to [-\infty, \infty]$ is lower semicontinuous. Thus, the restriction of f to C_w is lower semicontinuous. We have established that C_w is compact and that the restriction of f to C_w is lower semicontinuous, so we can apply Theorem 8 (the extreme value theorem) to obtain that K is a nonempty closed subset of C_w. Finally, K being a closed subset of C_w implies that K is a closed subset of X. □

10 Convex conjugation

If X is a locally convex space and X^* is its dual space, the *strong dual topology* on X^* is the seminorm topology induced by the seminorms $\lambda \mapsto \sup_{x \in E} |\lambda x|$, where E are the bounded subsets of X. Because these seminorms are a separating family, X^* with the strong dual topology is a locally convex space. (If X is a normed space then the strong dual topology on X^* is equal to the operator norm topology on X^*.[30]) A locally convex space is said to be *reflexive* if the strong dual of its strong dual is isomorphic as a locally convex space to the original space.

If X is a real locally convex space, the *convex conjugate*[31] of a function $f : X \to [-\infty, \infty]$ is the function $f^* : X^* \to [-\infty, \infty]$ defined by

$$f^*(\lambda) = \sup\{\langle x, \lambda \rangle - f(x) : x \in X\} = \sup\{\langle x, \lambda \rangle - f(x) : x \in \operatorname{dom} f\}.$$

The *convex biconjugate* of f is the function $f^{**} : X \to [-\infty, \infty]$ defined by

$$f^{**}(x) = \sup\{\langle x, \lambda \rangle - f^*(\lambda) : \lambda \in X^*\} = \sup\{\langle x, \lambda \rangle - f^*(\lambda) : \lambda \in \operatorname{dom} f^*\}.$$

The convex biconjugate of a function on a real reflexive locally convex space is the convex conjugate of its convex conjugate. From the definition of f^* it is

[28] Walter Rudin, *Functional Analysis*, second ed., p. 66, Theorem 3.12.
[29] The converse is true in a locally convex space. Walter Rudin, *Functional Analysis*, second ed., p. 70, Theorem 3.18.
[30] Kôsaku Yosida, *Functional Analysis*, sixth ed., p. 111, Theorem 1.
[31] Also called the *Fenchel transform*.

apparent that for all $x \in X$ and $\lambda \in X^*$,

$$\langle x, \lambda \rangle \leq f(x) + f^*(\lambda), \tag{1}$$

called *Young's inequality.*

The following theorem establishes some properties of the convex conjugates and convex biconjugates of any function from a real locally convex space to $[-\infty, \infty]$.[32]

Theorem 25. *If X is a real locally convex space and $f : X \to [-\infty, \infty]$ is a function, then*

- *f^* is convex and weak-* lower semicontinuous,*
- *f^{**} is convex and weakly lower semicontinuous,*
- *$f^{**} \leq f$.*

If $f_1, f_2 : X \to [-\infty, \infty]$ are functions satisfying $f_1 \leq f_2$, then $f_1^* \geq f_2^*$.

Proof. For each $x \in X$, it is apparent that the function $\lambda \mapsto \langle x, \lambda \rangle$ is convex and weak-* continuous, and a fortiori is weak-* lower semicontinuous. Whether $f(x)$ is finite or infinite, the function $\lambda \mapsto \langle x, \lambda \rangle$ is weak-* lower semicontinuous. By Lemma 14, the supremum of a collection of convex functions is a convex function, and by Theorem 4 the supremum of a collection of lower semicontinuous functions is a lower semicontinuous function. f^* is the supremum of this set of functions, and therefore is convex and weak-* lower semicontinuous.

For each $\lambda \in X^*$, the function $x \mapsto \langle x, \lambda \rangle - f^*(\lambda)$ is convex and is weakly semicontinuous, and a fortiori is weakly lower semicontinuous. As f^{**} is the supremum of this set of functions, f^{**} is convex and weakly lower semicontinuous.

For every $x \in X$ and $\lambda \in X^*$ we have by Young's inequality (1),

$$\langle x, \lambda \rangle - f^*(\lambda) \leq f(x),$$

and hence for every $x \in X$,

$$f^{**}(x) = \sup_{\lambda \in X^*} (\langle x, \lambda \rangle - f^*(\lambda)) \leq f(x),$$

and this means that $f^{**} \leq f$.

For $\lambda \in X^*$, because $f_1 \leq f_2$,

$$f_2^*(\lambda) = \sup_{x \in X} (\langle x, \lambda \rangle - f_2(x)) \leq \sup_{x \in X} (\langle x, \lambda \rangle - f_1(x)) = f_1^*(\lambda),$$

which means that $f_1^* \geq f_2^*$. □

[32] Viorel Barbu and Teodor Precupanu, *Convexity and Optimization in Banach Spaces*, fourth ed., p. 77, Proposition 2.19.

We remind ourselves that to say that a convex function is *proper* means that at some point it takes a value other than ∞, and that it nowhere takes the value $-\infty$. The following lemma shows that any lower semicontinuous proper convex function on a real locally convex space is bounded below by a continuous affine functional.

Lemma 26. *If X is a real locally convex space and $f : X \to (-\infty, \infty]$ is a lower semicontinuous proper convex function, then there is some $\mu \in X^*$ and some $c \in \mathbb{R}$ such that $f \geq \mu + c$.*

Proof. The fact that f is a convex function tells us that $\operatorname{epi} f$ is a convex subset of $X \times \mathbb{R}$, and as f is proper, $\operatorname{dom} f \neq \emptyset$ and $x_0 \in \operatorname{dom} f$ satisfies $f(x_0) > -\infty$. The fact that f is lower semicontinuous tells us that $\operatorname{epi} f$ is a closed subset of $X \times \mathbb{R}$. Let $x_0 \in \operatorname{dom} f$. We have $(x_0, f(x_0) - 1) \notin \operatorname{epi} f$. The singleton $\{(x_0, f(x_0) - 1)\}$ is a compact convex set and $\operatorname{epi} f$ is a disjoint closed convex set, so we can apply the Hahn-Banach separation theorem to obtain that there is some $\Lambda \in (X \times \mathbb{R})^*$ and some $\gamma \in \mathbb{R}$ satisfying

$$\Lambda(x, \alpha) < \gamma < \Lambda(x_0, f(x_0) - 1), \qquad (x, \alpha) \in \operatorname{epi} f.$$

There is some $\lambda \in X^*$ and some $\beta \in \mathbb{R}^* = \mathbb{R}$ such that $\Lambda(x, \alpha) = \lambda x + \beta \alpha$ for all $(x, \alpha) \in X \times \mathbb{R}$. So we have

$$\lambda x + \beta \alpha < \gamma < \lambda x_0 + \beta(f(x_0) - 1), \qquad (x, \alpha) \in \operatorname{epi} f.$$

And $(x_0, f(x_0)) \in \operatorname{epi} f$, so

$$\lambda x_0 + \beta f(x_0) < \lambda x_0 + \beta(f(x_0) - 1),$$

hence $\beta < 0$. If $x \in \operatorname{dom} f$ then $(x, f(x)) \in \operatorname{epi} f$ and

$$\lambda x + \beta f(x) < \lambda x_0 + \beta(f(x_0) - 1).$$

Rearranging, and as $\beta < 0$,

$$f(x) > -\frac{1}{\beta}\lambda x + \frac{1}{\beta}\lambda x_0 + f(x_0) - 1, \qquad x \in \operatorname{dom} f.$$

If $x \notin \operatorname{dom} f$ then $f(x) = \infty$, for which the above inequality also holds. $\square$

We proved in Theorem 25 that the conjugate of any function is convex and weak-* lower semicontinuous, which a fortiori gives that it is lower semicontinuous. In the following lemma we show that a convex lower semicontinuous function is proper if and only if its convex conjugate is proper.[33]

Lemma 27. *If X is a locally convex space and $f : X \to [-\infty, \infty]$ is a lower semicontinuous convex function, then f is proper if and only if f^* is proper.*

[33] Viorel Barbu and Teodor Precupanu, *Convexity and Optimization in Banach Spaces*, fourth ed., p. 78, Corollary 2.21.

Proof. Suppose that f is proper. By Lemma 26 there is some $\mu \in X^*$ and some $c \in \mathbb{R}$ such that $f(x) \geq \mu x + c$ for all $x \in X$. For any $\lambda \in X^*$ we have

$$f^*(\lambda) = \sup_{x \in X}(\lambda x - f(x)) \leq \sup_{x \in X}(\lambda x - \mu x - c),$$

thus $f^*(\mu) = -c < \infty$, so $\operatorname{dom} f^* \neq \emptyset$. And there is some $x_0 \in X$ such that $f(x_0) \neq \infty$, giving $\sup_{x \in X}(\lambda x - f(x)) \geq \lambda x_0 - f(x_0) > -\infty$, showing that $f^*(\lambda) > -\infty$ for all $\lambda \in X^*$. Therefore f^* is proper.

Suppose that f^* is proper. If f took only the value ∞ then f^* would take only the value $-\infty$, and f^* being proper means that it in fact never takes the value $-\infty$. Let $x \in X$. As f^* is proper there is some $\lambda \in X^*$ for which $f^*(\lambda) < \infty$, and using Young's inequality (1) we get

$$f(x) \geq -f^*(\lambda) + \langle x, \lambda \rangle > -\infty.$$

Thus, for every $x \in X$ we have $f(x) > -\infty$, so we have verified that f is proper. $\square$

The following theorem is called the *Fenchel-Moreau theorem*, and gives necessary and sufficient conditions for a function to equal its convex biconjugate.[34]

Theorem 28 (Fenchel-Moreau theorem)**.** *If X is a real locally convex space and $f : X \to [-\infty, \infty]$ is a function, then $f = f^{**}$ if and only if one of the following three conditions holds:*

1. *f is a proper convex lower semicontinuous function*
2. *f is the constant function ∞*
3. *f is the constant function $-\infty$*

Proof. Suppose that f is a proper convex lower semicontinuous function. From Lemma 27, its convex conjugate f^* is a proper convex function. As f^* does not take only the value ∞ we get from the definition of f^{**} that $f^{**}(x) > -\infty$ for every $x \in X$. Theorem 25 tells us that $f^{**} \leq f$, and suppose by contradiction that there is some $x_0 \in X$ for which $-\infty < f^{**}(x_0) < f(x_0)$. For this x_0 we have $(x_0, f^{**}(x_0)) \notin \operatorname{epi} f$, so we can apply the Hahn-Banach separation theorem to the sets $\{(x_0, f^{**}(x_0))\}$ and $\operatorname{epi} f$ to get that there is some $\Lambda \in (X \times \mathbb{R})^*$ and some $\gamma \in \mathbb{R}$ for which

$$\Lambda(x, \alpha) < \gamma < \Lambda(x_0, f^{**}(x_0)), \qquad (x, \alpha) \in \operatorname{epi} f.$$

But $\Lambda(x, \alpha)$ can be written as $\lambda x + \beta \alpha$ for some $\lambda \in X^*$ and some $\beta \in \mathbb{R}^* = \mathbb{R}$, so

$$\lambda x + \beta \alpha < \gamma < \lambda x_0 + \beta f^{**}(x_0), \qquad (x, \alpha) \in \operatorname{epi} f. \tag{2}$$

If β were > 0 then the left-hand side of (2) would be ∞, because for $x \in \operatorname{dom} f$ there are arbitrarily large α such that $(x, \alpha) \in \operatorname{epi} f$. But the left-hand side is

[34] Viorel Barbu and Teodor Precupanu, *Convexity and Optimization in Banach Spaces*, fourth ed., p. 79, Theorem 2.22.

upper bounded by the constant right-hand side, so $\beta \leq 0$. Either $f(x_0) < \infty$ or $f(x_0) = \infty$. In the first case, $(x_0, f(x_0)) \in \operatorname{epi} f$, and applying (2) gives $\beta f(x_0) < \beta f^{**}(x_0)$. The assumption that $f^{**}(x_0) < f(x_0)$ then implies that $\beta < 0$. In the case that $f(x_0) = \infty$, assume by contradiction that $\beta = 0$. Then (2) becomes

$$\lambda x < \gamma < \lambda x_0, \qquad x \in \operatorname{dom} f. \tag{3}$$

Let $\mu \in \operatorname{dom} f^*$; f^* is proper so $\operatorname{dom} f^* \neq \emptyset$. For all $h > 0$ we have

$$\begin{aligned}
f^*(\mu + h\lambda) &= \sup\{\langle x, \mu + h\lambda\rangle - f(x) : x \in X\} \\
&= \sup\{\langle x, \mu + h\lambda\rangle - f(x) : x \in \operatorname{dom} f\} \\
&\leq \sup\{\langle x, \mu\rangle - f(x) : x \in \operatorname{dom} f\} + h \sup\{\langle x, \lambda\rangle : x \in \operatorname{dom} f\} \\
&= f^*(\mu) + h \sup\{\langle x, \lambda\rangle : x \in \operatorname{dom} f\}.
\end{aligned}$$

But using the definition of f^{**} we have $f^{**}(x_0) \geq \langle x_0, \mu + h\lambda\rangle - f^*(\mu + h\lambda)$, so

$$\begin{aligned}
f^{**}(x_0) &\geq \langle x_0, \mu\rangle + h\langle x_0, \lambda\rangle - f^*(\mu + h\lambda) \\
&\geq \langle x_0, \mu\rangle + h\langle x_0, \lambda\rangle - f^*(\mu) - h \sup\{\langle x, \lambda\rangle : x \in \operatorname{dom} f\} \\
&= \langle x_0, \mu\rangle - f^*(\mu) + h(\langle x_0, \lambda\rangle - \sup\{\langle x, \lambda\rangle : x \in \operatorname{dom} f\}).
\end{aligned}$$

But (3) tells us that $\langle x_0, \lambda\rangle - \sup\{\langle x, \lambda\rangle : x \in \operatorname{dom} f\} > 0$, and since the above inequality holds for arbitrarily large h we get $f^{**}(x_0) = \infty$, contradicting that $f^{**}(x_0) < f(x_0)$. Therefore, $\beta < 0$, and then we can divide (2) by β to obtain

$$\frac{1}{\beta}\lambda x + \alpha > \frac{\gamma}{\beta} > \frac{1}{\beta}\lambda x_0 + f^{**}(x_0), \qquad (x, \alpha) \in \operatorname{epi} f,$$

hence

$$\begin{aligned}
\lambda\left(-\frac{x_0}{\beta}\right) - f^{**}(x_0) &> \sup\left\{-\frac{1}{\beta}\lambda x - \alpha : (x, \alpha) \in \operatorname{epi} f\right\} \\
&= \sup\left\{-\frac{1}{\beta}\lambda x - f(x) : x \in \operatorname{dom} f\right\} \\
&= f^*\left(-\frac{1}{\beta}\lambda\right).
\end{aligned}$$

From the definition of f^{**} we have

$$f^{**}(x_0) \geq \left\langle x_0, -\frac{1}{\beta}\lambda\right\rangle - f^*\left(-\frac{1}{\beta}\lambda\right),$$

and this and the above give

$$\lambda\left(-\frac{x_0}{\beta}\right) - f^*\left(-\frac{1}{\beta}\lambda\right) > \left\langle x_0, -\frac{1}{\beta}\lambda\right\rangle - f^*\left(-\frac{1}{\beta}\lambda\right),$$

but the two sides are equal, a contradiction. Therefore, $f^{**}(x_0) = f(x_0)$.

Suppose that f is the constant function ∞. Then f^* is the constant function $-\infty$, and this means that f^{**} is the constant function ∞, giving $f = f^{**}$.

Suppose that f is the constant function $-\infty$. This implies that f^* is the constant function ∞, and so f^{**} is the constant function $-\infty$, giving $f = f^{**}$.

Suppose that $f = f^{**}$. By Theorem 25, f^{**} is convex and weakly lower semicontinuous, so a fortiori it is lower semicontinuous, hence f is convex and lower semicontinuous. Suppose that f is neither the constant function ∞ nor the constant function $-\infty$. If f took the value $-\infty$ then f^* would take only the value ∞, and then f^{**} would take only the value $-\infty$, contradicting that $f = f^{**}$ is not the constant function $-\infty$. Therefore, if $f = f^{**}$ then either f is the constant function ∞, or f is the constant function $-\infty$, or f is a proper convex lower semicontinuous function. $\square$

If X is a topological space and $f : X \to [-\infty, \infty]$ is a function, the *closure of* f is the function $\operatorname{cl} f : X \to [-\infty, \infty]$ that is defined to be $\operatorname{lsc} f$ if $(\operatorname{lsc} f)(x) > -\infty$ for all $x \in X$, and defined to be the constant function $-\infty$ if there is some $x \in X$ such that $(\operatorname{lsc} f)(x) = -\infty$. We say that a function is *closed* if it is equal to its closure, and thus to say that a function is closed is to say that it is lower semicontinuous, and either does not take the value $-\infty$ or only takes the value $-\infty$. One checks that $(\operatorname{cl} f)^* = f^*$, and combined with the Fenchel-Moreau theorem one can obtain the following.

Corollary 29. *If X is a real locally convex space and $f : X \to [-\infty, \infty]$ is a convex function, then $f^{**} = \operatorname{cl} f$.*