2

Continuity and $\Gamma(X)$

Par délicatesse j'ai perdu ma vie
(A. Rimbaud, "Chanson de la plus haute tour")

Continuity, Lipschitz behavior, existence of directional derivatives, and differentiability are, of course, topics of the utmost importance in analysis. Thus the next two chapters will be dedicated to a description of the special features of convex functions from this point of view. Specifically, in this chapter we analyze the continuity of the convex functions and their Lipschitz behavior.

The first results show that a convex function which is bounded above around a point is continuous at that point, and that if it is at the same time lower and upper bounded on a ball centered at some point x, then it is Lipschitz in every smaller ball centered at x. The above continuity result entails also that a convex function is continuous at the interior points of its effective domain. It follows, in particular, that a convex, real valued function defined on a Euclidean space is everywhere continuous. This is no longer true in infinite dimensions.

We then introduce the notion of lower semicontinuity, and we see that if we require this additional property, then a real valued convex function is everywhere continuous in general Banach spaces. Lower semicontinuity, on the other hand, has a nice geometrical meaning, since it is equivalent to requiring that the epigraph of f, and all its level sets, are closed sets: one more time we relate an analytical property to a geometrical one. It is then very natural to introduce, for a Banach space X, the fundamental class $\Gamma(X)$ of convex, lower semicontinuous functions whose epigraph is nonempty (closed, convex) and does not contain vertical lines.

The chapter ends with a very fundamental characterization of a function in $\Gamma(X)$: it is the pointwise supremum of all affine functions minorizing it. Its proof relies, quite naturally, on the Hahn-Banach separation theorems recalled in Appendix A.

2.1 Continuity and Lipschitz behavior

Henceforth, as we shall deal with topological issues, every linear space will be endowed with a norm.

Convex functions have remarkable continuity properties. A key result is the following lemma, asserting that continuity at a point is implied by upper boundedness in a neighborhood of the point.

Lemma 2.1.1 *Let $f\colon X \to [-\infty, \infty]$ be convex, let $x_0 \in X$. Suppose there are a neighborhood V of x_0 and a real number a such that $f(x) \le a \ \forall x \in V$. Then f is continuous at x_0.*

Proof. We show the case when $f(x_0) \in \mathbb{R}$. By a translation of coordinates, which obviously does not affect continuity, we can suppose $x_0 = 0 = f(0)$. We can also suppose that V is a symmetric neighborhood of the origin. Suppose $x \in \varepsilon V$. Then $\frac{x}{\varepsilon} \in V$ and we get

$$f(x) \le (1-\varepsilon)f(0) + \varepsilon f(\frac{x}{\varepsilon}) \le \varepsilon a.$$

Now, write $0 = \frac{\varepsilon}{1+\varepsilon}(-\frac{x}{\varepsilon}) + \frac{1}{1+\varepsilon}x$ to get

$$0 \le \frac{\varepsilon}{1+\varepsilon} f(-\frac{x}{\varepsilon}) + \frac{1}{1+\varepsilon} f(x),$$

whence

$$f(x) \ge -\varepsilon f(-\frac{x}{\varepsilon}) \ge -\varepsilon a.$$

□

From the previous result, it is easy to get the fundamental

Theorem 2.1.2 *Let $f \in \mathcal{F}(X)$. The following are equivalent:*

(i) *There are a nonempty open set O and a real number a such that $f(x) \le a$ $\forall x \in O$;*

(ii) $\operatorname{int}\operatorname{dom} f \ne \emptyset$, *and f is continuous at all points of* $\operatorname{int}\operatorname{dom} f$.

Proof. The only nontrivial thing to show is that, whenever (i) holds, f is continuous at each point $x \in \operatorname{int}\operatorname{dom} f$. We shall exploit boundedness of f in O to find a nonempty open set I containing x where f is upper bounded. Suppose $f(z) \le a \ \forall z \in O$ and, without loss of generality, that $x = 0$. Fix a point $v \in O$. There exists $t > 0$ such that $-tv \in \operatorname{int}\operatorname{dom} f$. Now, let $h(y) := \frac{t+1}{t}y + v$. Then $h(0) = v$ and $I = h^{-1}(O)$ is a neighborhood of $x = 0$. Let $y \in I$. Then $y = \frac{t}{t+1}h(y) + \frac{1}{t+1}(-tv)$ and

$$f(y) \le \frac{t}{t+1}a + \frac{1}{t+1}f(-tv) \le a + f(-tv).$$

We found an upper bound for f in I, and this concludes the proof.

□

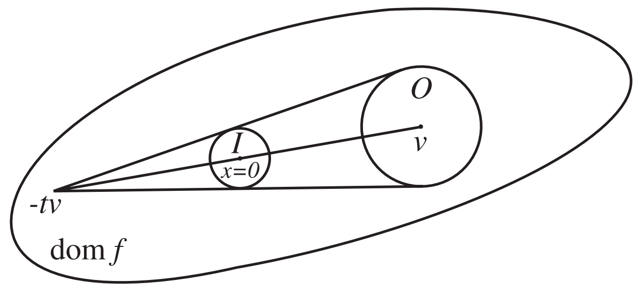

Figure 2.1.

Corollary 2.1.3 *Let $f \in \mathcal{F}(\mathbb{R}^n)$. Then f is continuous at each point of* $\operatorname{int}\operatorname{dom} f$. *In particular, if f is real valued, then it is everywhere continuous.*

Proof. If $x \in \operatorname{int}\operatorname{dom} f$, to show that f is upper bounded in a neighborhood of x, it is enough to observe that x can be put in the interior of a simplex, where f is bounded above by the maximum value assumed by f on the vertices of the simplex (see Exercise 1.2.26). □

Remark 2.1.4 The continuity of f at the boundary points of $\operatorname{dom} f$ is a more delicate issue. For instance, the function

$$f(x) = \begin{cases} 0 & \text{if } |x| < 1, \\ 1 & \text{if } |x| = 1, \\ \infty & \text{if } |x| > 1, \end{cases}$$

is convex and at the boundary points does not fulfill any continuity condition.

The next exercise characterizes the continuity of a sublinear function.

Exercise 2.1.5 Show the following:

Proposition 2.1.6 *Let $h\colon X \to (-\infty, \infty]$ be a sublinear function. Then the following are equivalent:*

(i) *h is finite at a point $x_0 \neq 0$ and continuous at $-x_0$;*
(ii) *h is upper bounded on a neighborhood of zero;*
(iii) *h is continuous at zero;*
(iv) *h is everywhere continuous.*

Hint. To show that (i) implies (ii), observe that $h(x_0) < \infty$ and $h(x) \leq h(x - x_0) + h(x_0)$. Moreover, observe that (iii) implies that h is everywhere real valued.

Exercise 2.1.7 Referring to Exercise 1.2.15, show that the Minkowski functional is continuous if and only if C is an absorbing set.

We saw that upper boundedness around a point guarantees continuity; the next lemma shows that a convex function is Lipschitz around a point if it is upper and lower bounded near that point.

Lemma 2.1.8 *Let $f \in \mathcal{F}(X)$, and let $x_0 \in X, R > 0$, $m, M \in \mathbb{R}$. Suppose $m \leq f(x) \leq M, \forall x \in B(x_0; R)$. Then f is Lipschitz on $B(x_0; r)$, for all $r < R$, with Lipschitz constant $\frac{M-m}{R-r}$.*

Proof. Let $x, y \in B(x_0; r)$ and let $z = y + \frac{R-r}{\|y-x\|}(y-x)$. Then $z \in B(x_0; R)$, hence $f(z) \leq M$. Moreover y is a convex combination of x and z:

$$y = \frac{\|y-x\|}{R-r+\|y-x\|}z + \frac{R-r}{R-r+\|y-x\|}x.$$

Hence

$$f(y) - f(x) \leq \frac{\|y-x\|}{R-r+\|y-x\|}M - \frac{\|y-x\|}{R-r+\|y-x\|}m \leq \frac{M-m}{R-r}\|y-x\|.$$

By interchanging the roles of x and y we get the result. □

2.2 Lower semicontinuity and $\Gamma(X)$

Let X be a topological space. Let $f\colon X \to (-\infty, \infty]$, $x \in X$, and denote by $\mathcal{N}$ the family of all neighborhoods of x. Remember that

$$\liminf_{y \to x} f(y) = \sup_{W \in \mathcal{N}} \inf_{y \in W \setminus \{x\}} f(y).$$

Definition 2.2.1 Let $f\colon X \to (-\infty, \infty]$. f is said to be *lower semicontinuous* if $\operatorname{epi} f$ is a closed subset of $X \times \mathbb{R}$. Given $x \in X$, f is said to be *lower semicontinuous at x* if

$$\liminf_{y \to x} f(y) \geq f(x).$$

Exercise 2.2.2 A subset E of $X \times \mathbb{R}$ is an epigraph if and only if $(x, a) \in E$ implies $(x, b) \in E$ for all $b \geq a$. If E is an epigraph, then $\operatorname{cl} E = \operatorname{epi} f$ with $f(x) = \inf\{a : (x, a) \in E\}$, and f is lower semicontinuous.

Definition 2.2.3 Let $f\colon X \to (-\infty, \infty]$. The *lower semicontinuous regularization* of f is the function $\bar{f}$ such that

$$\operatorname{epi} \bar{f} := \operatorname{cl} \operatorname{epi} f.$$

The definition above is consistent because $\operatorname{cl} \operatorname{epi}(f)$ is an epigraph, as is easy to prove (see Exercise 2.2.2). Moreover, it is obvious that $\bar{f}$ is the greatest lower semicontinuous function minorizing f: if $g \leq f$ and g is lower semicontinuous, then $g \leq \bar{f}$. Namely, $\operatorname{epi} g$ is a closed set containing $\operatorname{epi} f$, and thus it contains its closure too.

Exercise 2.2.4 Show that f is lower semicontinuous if and only if it is lower semicontinuous at $x, \forall x \in X$. Show that f is lower semicontinuous at x if and only if $f(x) = \bar{f}(x)$.

Hint. Let $l = \liminf_{y \to x} f(y)$. Show that $(x, l) \in \operatorname{cl} \operatorname{epi} f$. If f is everywhere lower semicontinuous, show that if $(x, r) \in \operatorname{cl} \operatorname{epi} f$, $\forall \varepsilon > 0$, $\forall W$ neighborhood of x, there is $y \in W$ such that $f(y) < r + \varepsilon$. Next, suppose f lower semicontinuous at x, observe that $(x, \bar{f}(x)) \in \operatorname{cl} \operatorname{epi} f$ and see that this implies $f(x) \leq \bar{f}(x)$. Finally, to see that $f(x) = \bar{f}(x)$ implies f lower semicontinuous at x, observe that $f(y) \geq \bar{f}(y)$ $\forall y \in X$ and use the definition.

Proposition 2.2.5 *Let $f\colon X \to (-\infty, \infty]$. Then f is lower semicontinuous if and only if f^a is a closed set $\forall a \in \mathbb{R}$.*

Proof. Let $x_0 \notin f^a$. Then $(x_0, a) \notin \operatorname{epi} f$. Thus there is an open set W containing x_0 such that $f(x) > a$ $\forall x \in W$. This shows that $(f^a)^c$ is open. Suppose, by way of contradiction, f^a closed for all a, and let $(x, b) \notin \operatorname{epi} f$. Then there is $\varepsilon > 0$ such that $f(x) > b + \varepsilon$, so that $x \notin f^{b+\varepsilon}$. Then there exists an open set W containing x such that $\forall y \in W$ $f(y) \geq b + \varepsilon$. Thus $W \times (-\infty, b + \varepsilon) s \cap \operatorname{epi} f = \emptyset$, which means that $(\operatorname{epi} f)^c$ is open and this ends the proof. □

When X is first countable, for instance a metric space, then lower semicontinuity of f at x can be given in terms of sequences: f is lower semicontinuous at x if and only if $\forall x_n \to x$,

$$\liminf_{n \to \infty} f(x_n) \geq f(x).$$

Example 2.2.6 I_C is lower semicontinuous if and only if C is a closed set.

Remark 2.2.7 Let $f\colon \mathbb{R} \to (-\infty, \infty]$ be convex. Then $\operatorname{dom} f$ is an interval, possibly containing its endpoints. If f is lower semicontinuous, then f restricted to $\operatorname{cl} \operatorname{dom} f$ is continuous.

We saw in Corollary 2.1.3 that a real valued convex function defined on a finite-dimensional space is everywhere continuous. The result fails in infinite dimensions. To see this, it is enough to consider a linear functional which is not continuous. However continuity can be recovered by assuming that f is lower semicontinuous. The following result holds:

Theorem 2.2.8 *Let X be a Banach space and let $f\colon X \to (-\infty, \infty]$ be a convex and lower semicontinuous function. Then f is continuous at the points of* $\operatorname{int} \operatorname{dom} f$.

Proof. Suppose $0 \in \operatorname{int} \operatorname{dom} f$, let $a > f(0)$ and let V be the closure of an open neighborhood of the origin which is contained in $\operatorname{dom} f$. Let us see that the closed convex set $f^a \cap V$ is absorbing (in V). Let $x \in V$. Then $g(t) := f(tx)$ defines a convex function on the real line. We have that $[-b, b] \in \operatorname{dom} g$ for some $b > 0$. Then g is continuous at $t = 0$, and thus it follows that there is

$\bar{t} > 0$ such that $\bar{t}x \in f^a$. By convexity and since $0 \in f^a$, we then have that $x \in nf^a$, for some large n. Thus

$$V = \bigcup_{n=1}^{\infty} n(f^a \cap V).$$

As a consequence of Baire's theorem (see Proposition B.1.1), $f^a \cap V$ is a neighborhood of the origin, (in V, and so in X), where f is upper bounded. Then f is continuous at the points of $\operatorname{int}\operatorname{dom} f$, see Theorem 2.1.2. $\square$

The family of convex, lower semicontinuous functions plays a key role in optimization, so that now we shall focus our attention on this class. For a Banach space X, we denote by $\Gamma(X)$ the set

$$\Gamma(X) := \{f \in \mathcal{F}(X) : f \text{ is lower semicontinuous}\}.$$

In other words, $\Gamma(X)$ is the subset of $\mathcal{F}(X)$ of the functions with a nonempty closed convex epigraph not containing vertical lines.

Example 2.2.9 $I_C \in \Gamma(X)$ if and only if C is a nonempty closed convex set.

Exercise 2.2.10 Verify that

$$f(x,y) := \begin{cases} \frac{y^2}{x} & \text{if } x > 0, y > 0, \\ 0 & \text{if } x \geq 0, y = 0, \\ \infty & \text{otherwise} \end{cases}$$

belongs to $\Gamma(\mathbb{R}^2)$. Verify also that f does not assume a maximum on the (compact, convex) set $C = \{(x,y) \in \mathbb{R}^2 : 0 \leq x, y \leq 1,\ y \leq \sqrt{x} - x^2\}$.

Hint. Consider the sequence $\{(1/n, (1/\sqrt{n} - 1/n^2))\}$.

The example above shows that $f(\operatorname{dom} f \cap C))$ need not be closed, even if C is compact. The next exercise highlights the structure of the image of a convex set by a function in $\Gamma(X)$.

Exercise 2.2.11 Prove that for $f \in \Gamma(X)$, $f(\operatorname{dom} f \cap C))$ is an interval for every (closed) convex set C.

Hint. Let $a, b \in f(\operatorname{dom} f \cap C)$. Then there exist $x \in C$, $y \in C$ such that $f(x) = a$, $f(y) = b$. Now consider $g(t) = f(tx + (1-t)y)$, $t \in [0,1]$.

We see now that $\Gamma(X)$ is an (essentially) stable family with respect to some operations.

Proposition 2.2.12 *Let $f_i \in \Gamma(X), \forall i = 1, \dots, n$ and let $t_1, \dots, t_n > 0$. If for some $x_0 \in X$ $f_i(x_0) < \infty\ \forall i$, then $(\sum_{i=1}^n t_i f_i) \in \Gamma(X)$.*

Proof. From Proposition 1.2.18 and because for $a, b > 0$, $f, g \in \Gamma(X)$, $x \in X$, W a neighborhood of x,

$$\inf_{y \in W \setminus \{x\}} f(y) + g(y) \geq \inf_{y \in W \setminus \{x\}} f(y) + \inf_{y \in W \setminus \{x\}} g(y).$$

Thus

$$\begin{aligned}\sup_W \inf_{y \in W \setminus \{x\}} af(y) + bg(y) &\geq \sup_W \bigl(a \inf_{y \in W \setminus \{x\}} f(y) + b \inf_{y \in W \setminus \{x\}} g(y)\bigr) \\ &= a \sup_W \inf_{y \in W \setminus \{x\}} f(y) + b \sup_W \inf_{y \in W \setminus \{x\}} g(y).\end{aligned}$$

□

Proposition 2.2.13 *Let $f_i \in \Gamma(X), \forall i \in J$, where J is an arbitrary index set. If for some $x_0 \in X$ $\sup_{i \in J} f_i(x_0) < \infty$, then $(\sup_{i \in J} f_i) \in \Gamma(X)$.*

Proof. $\operatorname{epi}(\sup_{i \in J} f_i) = \bigcap_{i \in J} \operatorname{epi} f_i$. □

The following Example shows that $\Gamma(X)$ is not closed with respect to the inf-convolution operation.

Example 2.2.14 Let C_1, C_2 be closed convex sets. Then $I_{C_1} \nabla I_{C_2} = I_{C_1 + C_2}$ (see Exercise 1.2.24). On the other hand, the function I_C is lower semicontinuous if and only if C is a closed convex set. Taking

$$C_1 := \{(x, y) \in \mathbb{R}^2 : x \leq 0 \text{ and } y \geq 0\}$$

and

$$C_2 := \{(x, y) \in \mathbb{R}^2 : x \geq 0 \text{ and } y \geq \frac{1}{x}\},$$

since $C_1 + C_2$ is not a closed set, then $I_{C_1} \nabla I_{C_2} \notin \Gamma(X)$.

Remark 2.2.15 An example as above cannot be constructed for functions defined on the real line. Actually, in this case the inf-convolution of two convex lower semicontinuous functions is lower semicontinuous. It is enough to observe that the effective domain of $(f \nabla g)$ is an interval. Let us consider, for instance, its right endpoint b, assuming that $(f \nabla g)(b) \in \mathbb{R}$ (the other case is left for the reader). Then if b_1 is the right endpoint of $\operatorname{dom} f$ and b_2 is the right endpoint of $\operatorname{dom} g$, it follows that

$$(f \nabla g)(b) = f(b_1) + g(b_2),$$

and if $x_k \to b^-$, taking x_k^1, x_k^2 with $x_k^1 + x_k^2 = x_k$ and $f(x_k^1) + g(x_k^2) \leq (f \nabla g)(x_k) + \frac{1}{k}$, then $x_k^1 \to b_1^-$, $x_k^2 \to b_2^-$ and

$$\begin{aligned}(f \nabla g)(b) = f(b_1) + g(b_2) &\leq \liminf (f(x_k^1) + g(x_k^2)) \\ &\leq \liminf ((f \nabla g)(x_k) + \frac{1}{k}) = \liminf (f \nabla g)(x_k).\end{aligned}$$

We intend now to prove a fundamental result for functions in $\Gamma(X)$. We start with some preliminary facts. Let X be a Banach space and denote by X^* its topological dual space, the space of all real valued linear continuous functionals defined on X. Then X^* is a Banach space, when endowed with the canonical norm $\|x^*\|_* = \sup\{\langle x^*, x\rangle : \|x\| = 1\}$.

Lemma 2.2.16 *Let $f \in \Gamma(X)$, $x_0 \in \operatorname{dom} f$ and $k < f(x_0)$. Then there are $y^* \in X^*$ and $q \in \mathbb{R}$ such that the affine function $l(x) = \langle y^*, x\rangle + q$ fulfills*

$$f(x) \geq l(x), \forall x \in X, l(x_0) > k.$$

Proof. In $X \times \mathbb{R}$, let us consider the closed convex set $\operatorname{epi} f$ and the point (x_0, k). They can be separated by a closed hyperplane (Theorem A.1.6): there are $x^* \in X^*$, $r, c \in \mathbb{R}$ such that

$$\langle x^*, x\rangle + rb > c > \langle x^*, x_0\rangle + rk, \forall x \in \operatorname{dom} f, \forall b \geq f(x).$$

With the choice of $x = x_0, b = f(x_0)$ in the left part of the above formula, we get $r(f(x_0) - k) > 0$, and so $r > 0$. Let us consider the affine function $l(x) = \langle y^*, x\rangle + q$, with $y^* = \frac{-x^*}{r}$, $q = \frac{c}{r}$. It is then easy to see that $l(x) \leq f(x)\, \forall x \in X$ and that $l(x_0) > k$. □

Corollary 2.2.17 *Let $f \in \Gamma(X)$. Then there exists an affine function minorizing f.*

Corollary 2.2.18 *Let $f \in \Gamma(X)$. Then f is lower bounded on bounded sets.*

Corollary 2.2.19 *Let $f \in \Gamma(X)$ be upper bounded on a neighborhood of $x \in X$. Then f is locally Lipschitz around x.*

Proof. From the previous Corollary and Lemma 2.1.8. □

Remark 2.2.20 The conclusion of Corollary 2.2.19 can be strengthened if X is finite-dimensional and f is real valued. In this case f is Lipschitz on all bounded sets. This is no longer true in infinite dimensions, because then it can happen that f is not upper bounded on all bounded sets, as the following example shows. Consider a separable Hilbert space X, and let $\{e_n\}$ be an orthonormal basis. Consider the function

$$f(x) = \sum_{n=1}^{\infty} n(x, e_n)^{2n}.$$

Then f is not upper bounded on the unit ball.

Theorem 2.2.21 *Let $f\colon X \to (-\infty, \infty]$ be not identically ∞. Then $f \in \Gamma(X)$ if and only if, $\forall x \in X$*

$$f(x) = \sup\{\langle x^*, x\rangle + a : x^* \in X^*, a \in \mathbb{R}, f(x) \geq \langle x^*, x\rangle + a\}.$$

Proof. Denote by $h(x)$ the function $h(x) = \sup\{\langle x^*, x\rangle + a : x^* \in X^*, a \in \mathbb{R}, f(x) \geq \langle x^*, x\rangle + a\}$. Then $h(x) \leq f(x)$ and, being the pointwise supremum of affine functions, $h \in \Gamma(X)$ (see Proposition 2.2.13); this provides one of the implications. As far as the other one is concerned, let us consider $x_0 \in X$, $k < f(x_0)$ and prove that $h(x_0) > k$. Lemma 2.2.16 shows that $h(x_0) > k$ if $x_0 \in \operatorname{dom} f$. We then consider the case $f(x_0) = \infty$.

Recalling the proof of Lemma 2.2.16, we can claim existence of $x^* \in X^*$, $r, c \in \mathbb{R}$ such that

$$\langle x^*, x\rangle + rb > c > \langle x^*, x_0\rangle + rk, \forall x \in \operatorname{dom} f, \forall b \geq f(x).$$

If $r \neq 0$, we conclude as in Lemma 2.2.16. If $r = 0$, which geometrically means that the hyperplane separating $\operatorname{epi} f$ and (x_0, k) is vertical, then

$$\langle x^*, x\rangle > c > \langle x^*, x_0\rangle, \forall x \in \operatorname{dom} f.$$

Calling $l(x) = \langle -x^*, x\rangle + c$, we have $l(x_0) > 0$ and $l(x) < 0, \forall x \in \operatorname{dom} f$. From Corollary 2.2.17, there exists an affine function $m(x) := \langle y^*, x\rangle + q$ with the property that $f(x) \geq m(x), \forall x \in X$. Hence, $\forall h > 0$, $m(x) + hl(x) \leq f(x), \forall x \in \operatorname{dom} f$, whence $m(x) + hl(x) \leq f(x), \forall x \in X$. On the other hand, as $l(x_0) > 0$, for a sufficiently large h, $(m + hl)(x_0) > k$, and this concludes the proof. $\square$

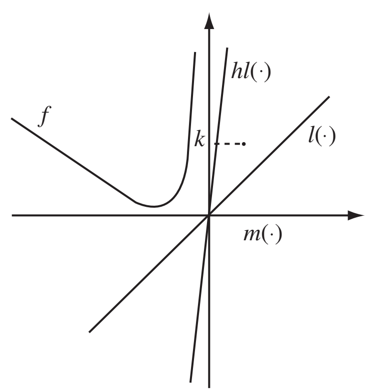

Figure 2.2.

The previous theorem can be refined if f is also a positively homogeneous function.

Corollary 2.2.22 *Let $h \in \Gamma(X)$ be sublinear. Then*

$$h(x) = \sup\{\langle x^*, x\rangle : x^* \in X^*, h(x) \geq \langle x^*, x\rangle\}.$$

Proof. It is enough to show that if the *affine* function $\langle x^*, \cdot\rangle + c$ minorizes h, then the *linear* function $\langle x^*, \cdot\rangle$ minorizes h. Now, since h is positively homogeneous, $\forall x \in X$, $\forall t > 0$,

$$\langle x^*, \frac{x}{t}\rangle + \frac{c}{t} \leq h\left(\frac{x}{t}\right),$$

i.e.,

$$\langle x^*, y\rangle + \frac{c}{t} \leq h(y),$$

$\forall y \in X$. We conclude now by letting t go to ∞. □

Exercise 2.2.23 Let C be a nonempty closed convex set. Let $d(\cdot\,, C)$ be the distance function from C: $d(x, C) = \inf_{c\in C} \|x - c\|$. Then d is 1-Lipschitz.