Solution to Ex 1.3:

Actually we'll prove the lemma in a more general context.

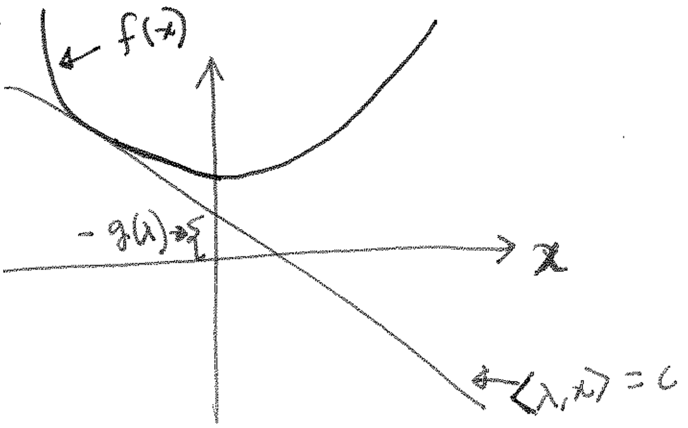
Lemma: Let $\mathcal{X}$ be a locally convex topological vector space. Let $f : \mathcal{X} \longrightarrow (-\infty, \infty]$ is a convex and lower semicontinuous function. For

$$g(\lambda) = \sup_{x \in \mathcal{X}} \{ \langle \lambda, x \rangle - f(x) \},$$

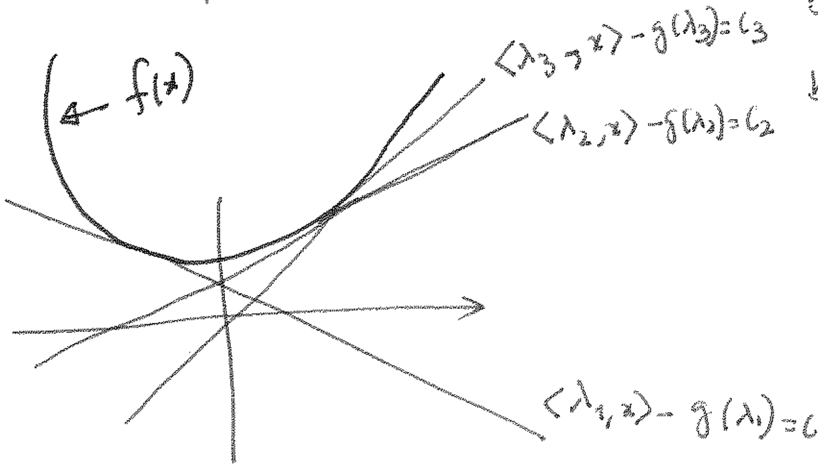
we have

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

Geometric interpretation:



For every hyperplane defined by λ, $g(\lambda)$ is the largest amount one needs to push up the tangent before it hits $f(\cdot)$ and becomes a hyperplane. The lemma says that to reconstruct f, one needs to find the tangent at x and push it down by $g(\lambda)$.



For the proof, see the following line of arguments (taken from [DZ98]):

(Lemma 4.5.8) to the convex rate function I. $\square$

Proof of Lemma 4.5.8: Consider the sets $\mathcal{X} \times \mathbb{R}$ and $\mathcal{X}^* \times \mathbb{R}$. Each of these can be made into a locally convex, Hausdorff topological vector space in the obvious way. If f is identically ∞, then g is identically $-\infty$ and the lemma trivially holds. Assume otherwise and define

$$\begin{aligned}\mathcal{E} &= \{(x, \alpha) \in \mathcal{X} \times \mathbb{R} : f(x) \leq \alpha\}\,, \\ \mathcal{E}^* &= \{(\lambda, \beta) \in \mathcal{X}^* \times \mathbb{R} : g(\lambda) \leq \beta\}\,.\end{aligned}$$

Note that for any $(\lambda, \beta) \in \mathcal{E}^*$ and any $x \in \mathcal{X}$,

$$f(x) \geq \langle \lambda, x \rangle - \beta\,.$$

Therefore, it also holds that

$$f(x) \geq \sup_{(\lambda,\beta)\in\mathcal{E}^*} \{\langle \lambda, x\rangle - \beta\} = \sup_{\lambda\in\mathcal{X}^*} \{\langle \lambda, x\rangle - g(\lambda)\} .$$

It thus suffices to show that for any $(x, \alpha) \notin \mathcal{E}$ (i.e., $f(x) > \alpha$), there exists a $(\lambda, \beta) \in \mathcal{E}^*$ such that

$$\langle \lambda, x\rangle - \beta > \alpha , \tag{4.5.17}$$

in order to complete the proof of the lemma.

Since f is a lower semicontinuous function, the set $\mathcal{E}$ is closed (alternatively, the set $\mathcal{E}^c$ is open). Indeed, whenever $f(x) > \gamma$, there exists a neighborhood V of x such that $\inf_{y\in V} f(y) > \gamma$, and thus $\mathcal{E}^c$ contains a neighborhood of (x, γ). Moreover, since $f(\cdot)$ is convex and not identically ∞, the set $\mathcal{E}$ is a non-empty convex subset of $\mathcal{X} \times \mathbb{R}$.

Fix $(x, \alpha) \notin \mathcal{E}$. The product space $\mathcal{X} \times \mathbb{R}$ is locally convex and therefore, by the Hahn–Banach theorem (Theorem B.6), there exists a hyperplane in $\mathcal{X} \times \mathbb{R}$ that strictly separates the non-empty, closed, and convex set $\mathcal{E}$ and the point (x, α) in its complement. Hence, as the topological dual of $\mathcal{X} \times \mathbb{R}$ is $\mathcal{X}^* \times \mathbb{R}$, for some $\mu \in \mathcal{X}^*$, $\rho \in \mathbb{R}$, and $\gamma \in \mathbb{R}$,

$$\sup_{(y,\xi)\in\mathcal{E}} \{\langle \mu, y\rangle - \rho\xi\} \leq \gamma < \langle \mu, x\rangle - \rho\alpha .$$

In particular, since f is not identically ∞, it follows that $\rho \geq 0$, for otherwise a contradiction results when $\xi \to \infty$. Moreover, by considering $(y, \xi) = (x, f(x))$, the preceding inequality implies that $\rho > 0$ whenever $f(x) < \infty$.

Suppose first that $\rho > 0$. Then, (4.5.17) holds for the point $(\mu/\rho,\ \gamma/\rho)$. This point must be in $\mathcal{E}^*$, for otherwise there exists a $y_0 \in \mathcal{X}$ such that $\langle \mu, y_0\rangle - \rho f(y_0) > \gamma$, contradicting the previous construction of the separating hyperplane (since $(y_0, f(y_0)) \in \mathcal{E}$). In particular, since $f(x) < \infty$ for some $x \in \mathcal{X}$ it follows that $\mathcal{E}^*$ is non-empty.

Now suppose that $\rho = 0$ so that

$$\sup_{\{y:f(y)<\infty\}} \{\langle \mu, y\rangle - \gamma\} \leq 0 ,$$

while $\langle \mu, x\rangle - \gamma > 0$. Consider the points

$$(\lambda_\delta, \beta_\delta) \stackrel{\triangle}{=} \left(\frac{\mu}{\delta} + \lambda_0, \frac{\gamma}{\delta} + \beta_0\right) , \quad \forall \delta > 0 ,$$

where (λ_0, β_0) is an arbitrary point in $\mathcal{E}^*$. Then, for all $y \in \mathcal{X}$,

$$\langle \lambda_\delta, y\rangle - \beta_\delta = \frac{1}{\delta}(\langle \mu, y\rangle - \gamma) + (\langle \lambda_0, y\rangle - \beta_0) \leq f(y) .$$

Therefore, $(\lambda_\delta, \beta_\delta) \in \mathcal{E}^*$ for any $\delta > 0$. Moreover,

$$\lim_{\delta \to 0} (\langle \lambda_\delta, x \rangle - \beta_\delta) = \lim_{\delta \to 0} \left\{ \frac{1}{\delta} (\langle \mu, x \rangle - \gamma) + (\langle \lambda_0, x \rangle - \beta_0) \right\} = \infty .$$

Thus, for any $\alpha < \infty$, there exists $\delta > 0$ small enough so that $\langle \lambda_\delta, x \rangle - \beta_\delta > \alpha$. This completes the proof of (4.5.17) and of Lemma 4.5.8. $\square$