9.1. Metrizability and weak* topology

Let $(X, \|\cdot\|_X)$ be a separable normed $\mathbb{K}$-vector space (with $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$). Prove that the weak* topology on the unit ball $B^* := \{\varphi \in X^* \colon \|\varphi\|_{X^*} \leq 1\}$ of X^* is metrizable.

Solution: Let $(x_n)_{n\in\mathbb{N}} \subseteq X$ be a dense subset of the unit ball $B := \{x \in X \colon \|x\|_X \leq 1\}$ in X. The fact that $\sup_{n\in\mathbb{N}}\|x_n\|_X \leq 1$ ensures that the mapping $d\colon B^* \times B^* \to [0, \infty)$, given by

$$d(\varphi, \psi) = \sum_{n=1}^{\infty} 2^{-n}|\varphi(x_n) - \psi(x_n)| \quad \text{for all } \varphi, \psi \in B^*,$$

is well-defined. Indeed:

$$\begin{aligned} 0 &\leq \sum_{n=1}^{\infty} 2^{-n}|\varphi(x_n) - \psi(x_n)| \leq \sum_{n=1}^{\infty} 2^{-n}\|\varphi - \psi\|_{X^*}\|x_n\|_X \\ &\leq \sum_{n=1}^{\infty} 2^{-n}\|\varphi - \psi\|_{X^*} \leq \|\varphi - \psi\|_{X^*} \quad \text{for all } \varphi, \psi \in B^*. \end{aligned}$$

We claim that d is a metric on B^*. For this, note that symmetry is clear. Moreover, for all $\varphi, \psi, \xi \in B^*$, we obtain

$$\begin{aligned} d(\varphi, \xi) &= \sum_{n=1}^{\infty} 2^{-n}|\varphi(x_n) - \xi(x_n)| \\ &\leq \sum_{n=1}^{\infty} 2^{-n}|\varphi(x_n) - \psi(x_n)| + \sum_{n=1}^{\infty} 2^{-n}|\psi(x_n) - \xi(x_n)| \\ &= d(\varphi, \psi) + d(\psi, \xi), \end{aligned}$$

that is, the triangle inequality holds. Finally, for $\varphi, \psi \in B^*$ we can infer from $d(\varphi, \psi) = 0$ that $\varphi(x_n) = \psi(x_n)$ for all $n \in \mathbb{N}$. Hence, any $\varphi, \psi \in B^*$ with $d(\varphi, \psi) = 0$ have to coincide on $\operatorname{span}\{x_n \mid n \in \mathbb{N}\}$ because of linearity and even on $\overline{\operatorname{span}\{x_n \mid n \in \mathbb{N}\}}$ because of continuity. As $\overline{\operatorname{span}\{x_n \mid n \in \mathbb{N}\}} = X$ due to $(x_n)_{n\in\mathbb{N}}$ lying dense in the unit ball B of X, we obtain that any $\varphi, \psi \in B^*$ with $d(\varphi, \psi) = 0$ have to be identical.

All of the above is useless if we cannot show that the weak* topology τ_{w^*} on B^* is equal to the topology τ_{d} on B^* which is induced by the metric d. Next, we are going to show that $\tau_{\mathrm{d}} \subseteq \tau_{\mathrm{w}^*}$ and $\tau_{\mathrm{w}^*} \subseteq \tau_{\mathrm{d}}$.

"$\tau_{\mathrm{d}} \subseteq \tau_{\mathrm{w}^*}$": Let $O \in \tau_{\mathrm{d}}$ and $\varphi \in O$ be arbitrary. Then there exists $\varepsilon \in (0, \infty)$ such that $\{\psi \in B^* \mid d(\varphi, \psi) < \varepsilon\} \subseteq O$. With $N \in \mathbb{N}$ so that $2^{-N} < \frac{\varepsilon}{4}$, we get that

$$\begin{aligned} \sum_{n=N+1}^{\infty} 2^{-n}|\varphi(x_n) - \psi(x_n)| &\leq \sum_{n=N+1}^{\infty} 2^{-n}(\|\varphi\|_{X^*} + \|\psi\|_{X^*}) \\ &\leq \sum_{n=N+1}^{\infty} 2^{-n+1} = 2^{-N+1} < \frac{\varepsilon}{2} \quad \text{for all } \psi \in B^*. \end{aligned}$$

This implies in particular that

$$\left\{\psi \in B^* \mid \forall\, n \in \{1,2,\dots,N\} \colon |\varphi(x_n) - \psi(x_n)| < \frac{\varepsilon}{2}\right\} \subseteq O.$$

As $\varphi \in O$ was arbitrary, this ensures that $O \in \tau_{\mathrm{w}^*}$. As $O \in \tau_{\mathrm{d}}$ was arbitrary, we've arrived at showing $\tau_{\mathrm{d}} \subseteq \tau_{\mathrm{w}^*}$.

"$\tau_{\mathrm{w}^*} \subseteq \tau_{\mathrm{d}}$": Let $O \in \tau_{\mathrm{w}^*}$ and $\varphi \in O$ be arbitrary. Then there exist $N \in \mathbb{N}$, $\varepsilon \in (0,\infty)$ and $y_1, y_2, \dots, y_N \in X$ satisfying that

$$\{\psi \in B^* \mid \forall\ n \in \{1,2,\dots,N\} \colon |\psi(y_n) - \varphi(y_n)| < \varepsilon\} \subseteq O.$$

W.l.o.g. we may assume that $\sup_{n\in\mathbb{N}} \|y_n\|_X \leq 1$ (otherwise, replace y_n by $\frac{y_n}{\|y_n\|_X}$ if $\|y_n\|_X > 1$). Since $(x_n)_{n\in\mathbb{N}} \subseteq B$ is dense in B, there exist $k_1, k_2, \dots, k_N \in \mathbb{N}$ such that

$$\|y_n - x_{k_n}\|_X < \frac{\varepsilon}{4} \quad \text{for all } n \in \{1,2,\dots,N\}.$$

Thus, with $\mathcal{N} := \max_{1\leq i\leq N} k_i \in \mathbb{N}$, we have

$$\begin{aligned} &\left\{\psi \in B^* \mid \forall n \in \{1,2,\dots,\mathcal{N}\} \colon |\psi(x_n) - \varphi(x_n)| < \frac{\varepsilon}{2}\right\} \\ &\subseteq \{\psi \in B^* \mid \forall n \in \{1,2,\dots,N\} \colon |\psi(y_n) - \varphi(y_n)| < \varepsilon\} \end{aligned} \tag{1}$$

since, if $\psi \in B^*$ satisfies $|\psi(x_n) - \varphi(x_n)| < \frac{\varepsilon}{2}$ for all $n \in \{1,2,\dots,\mathcal{N}\}$, then it holds in particular for all $n \in \{1,2,\dots,N\}$ that

$$\begin{aligned} |\psi(y_n) - \varphi(y_n)| &\leq |\psi(y_n) - \psi(x_{k_n})| + |\psi(x_{k_n}) - \varphi(x_{k_n})| + |\varphi(x_{k_n}) - \varphi(y_n)| \\ &\leq \|\psi\|_{X^*}\|y_n - x_{k_n}\|_X + |\psi(x_{k_n}) - \varphi(x_{k_n})| + \|\varphi\|_{X^*}\|x_{k_n} - y_n\|_X \\ &\leq 2\|y_n - x_{k_n}\|_X + \frac{\varepsilon}{2} < \varepsilon. \end{aligned}$$

But now we are done since for all $\psi \in B^*$ with $d(\psi,\varphi) < 2^{-\mathcal{N}}\frac{\varepsilon}{2}$ it holds that

$$|\psi(x_n) - \varphi(x_n)| \leq 2^n d(\psi,\varphi) < \frac{\varepsilon}{2} \quad \text{for all } n \in \{1,2,\dots,\mathcal{N}\},$$

which implies (having (1) in mind) that

$$\begin{aligned} &\left\{\psi \in B^* \mid d(\psi,\varphi) < 2^{-\mathcal{N}}\frac{\varepsilon}{2}\right\} \\ &\subseteq \left\{\psi \in B^* \mid \forall n \in \{1,2,\dots,\mathcal{N}\} \colon |\psi(x_n) - \varphi(x_n)| < \frac{\varepsilon}{2}\right\} \\ &\subseteq \{\psi \in B^* \mid \forall n \in \{1,2,\dots,N\} \colon |\psi(y_n) - \varphi(y_n)| < \varepsilon\}. \end{aligned}$$

As $\phi \in O$ was arbitrary, we demonstrated that $O \in \tau_{\mathrm{d}}$. As $O \in \tau_{\mathrm{w}^*}$ was arbitrary, we showed $\tau_{\mathrm{w}^*} \subseteq \tau_{\mathrm{d}}$.

9.2. Weak convergence in Hilbert spaces

Let $(H, (\cdot,\cdot)_H)$ be an infinite-dimensional $\mathbb{K}$-Hilbert space (with $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$).

(a) Let $(x_n)_{n\in\mathbb{N}} \subseteq H$ and $x_\infty \in H$ satisfy that $x_n \xrightarrow{w} x_\infty$ in H and $\|x_n\|_H \to \|x_\infty\|_H$ in $\mathbb{R}$ as $n \to \infty$. Prove that $x_n \to x_\infty$ in H as $n \to \infty$, i. e. $\limsup_{n\to\infty} \|x_n - x_\infty\|_H = 0$.

Solution: Since $(H \ni y \mapsto (y, x_\infty)_H \in \mathbb{K}) \in H^*$, the weak convergence of $(x_n)_{n\in\mathbb{N}}$ to x_∞ implies

$$\lim_{n\to\infty} (x_n, x_\infty)_H = (x_\infty, x_\infty)_H = \|x_\infty\|_H^2 \quad \text{and} \quad \lim_{n\to\infty} \operatorname{Re}(x_n, x_\infty) = \|x_\infty\|^2.$$

Combining this with the assumption that $\|x_n\|_H \to \|x_\infty\|_H$ as $n \to \infty$, we obtain

$$\begin{aligned}\limsup_{n\to\infty} \|x_n - x_\infty\|_H^2 &= \limsup_{n\to\infty} (x_n - x_\infty, x_n - x_\infty)_H \\ &= \limsup_{n\to\infty} \Big[\|x_n\|_H^2 - 2\operatorname{Re}(x_\infty, x_n)_H + \|x_\infty\|_H^2\Big] = 0.\end{aligned}$$

(b) Suppose $(x_n)_{n\in\mathbb{N}}, (y_n)_{n\in\mathbb{N}} \subseteq H$ and $x_\infty, y_\infty \in H$ satisfy that $x_n \xrightarrow{w} x_\infty$ and $\|y_n - y_\infty\|_H \to 0$ as $n \to \infty$. Prove that $(x_n, y_n)_H \to (x_\infty, y_\infty)_H$ as $n \to \infty$.

Solution: Weak convergence $x_n \xrightarrow{w} x_\infty$ implies in particular that $(x_n, y_\infty)_H \to (x_\infty, y_\infty)_H$ as $n \to \infty$ and that there exists a finite constant C such that $\|x_n\|_H \le C$ for all $n \in \mathbb{N}$. Thus,

$$\begin{aligned}&\limsup_{n\to\infty} |(x_n, y_n)_H - (x_\infty, y_\infty)_H| \\ &= \limsup_{n\to\infty} |(x_n, y_n - y_\infty)_H + (x_n, y_\infty)_H - (x_\infty, y_\infty)_H| \\ &\le \limsup_{n\to\infty} \Big[C\|y_n - y_\infty\|_H + |(x_n, y_\infty)_H - (x_\infty, y_\infty)_H|\Big] = 0.\end{aligned}$$

(c) Let $(e_n)_{n\in\mathbb{N}}$ be an orthonormal system of $(H, (\cdot,\cdot)_H)$. Prove $e_n \xrightarrow{w} 0$ as $n \to \infty$.

Solution: Note that Bessel's inequality, i.e.,

$$\sum_{n=0}^{\infty} |(x, e_n)_H|^2 \le \|x\|_H^2 \quad \text{for all } x \in H,$$

implies that $(x, e_n)_H \to 0$ as $n \to \infty$ for any $x \in H$. Since by the Riesz representation theorem any $f \in H^*$ satisfies $f(e_n) = (e_n, x)_H$ for a unique $x \in H$, we obtain $e_n \xrightarrow{w} 0$ as $n \to \infty$.

(d) Given any $x_\infty \in H$ with $\|x_\infty\|_H \le 1$, prove that there exists a sequence $(x_n)_{n\in\mathbb{N}}$ in H satisfying $\|x_n\|_H = 1$ for all $n \in \mathbb{N}$ and $x_n \overset{w}{\rightharpoonup} x_\infty$ as $n \to \infty$.

Solution: If $x_\infty = 0$, then any orthonormal system converges weakly to x_∞ by (c). If $x_\infty \neq 0$, then an orthonormal system $(e_n)_{n\in\mathbb{N}}$ of H with $e_1 = \|x_\infty\|_H^{-1} x_\infty$ can be constructed via the Gram–Schmidt algorithm. For $n \in \mathbb{N}$, let

$$x_n := x_\infty + \Big(\sqrt{1 - \|x_\infty\|_H^2}\Big) e_{n+1}.$$

Then, since $x_\infty \perp e_{n+1}$ for every $n \in \mathbb{N}$, we have $\|x_n\|^2 = \|x_\infty\|_H^2 + (1 - \|x_\infty\|_H^2) = 1$ for every $n \in \mathbb{N}$. Moreover, $x_n \overset{w}{\rightharpoonup} x_\infty$ follows from $e_{n+1} \overset{w}{\rightharpoonup} 0$ as $n \to \infty$ by (c).

(e) Let the functions $f_n \colon [0, 2\pi] \to \mathbb{R}$ be given by $f_n(t) = \sin(nt)$ for $n \in \mathbb{N}$. Prove the Riemann–Lebesgue Lemma: $f_n \overset{w}{\rightharpoonup} 0$ in $L^2([0, 2\pi], \mathbb{R})$ as $n \to \infty$.

Solution: Let $f_n \colon [0, 2\pi] \to \mathbb{R}$ be given by $f_n(t) = \sin(nt)$ for $n \in \mathbb{N}$. Then, $(\sqrt{\frac{1}{\pi}} f_n)_{n\in\mathbb{N}}$ is an orthonormal system of $L^2([0, 2\pi], \mathbb{R})$, because

$$\begin{aligned}\int_0^{2\pi} \sin(mt)\sin(nt)\,dt &= \frac{1}{2}\int_0^{2\pi} \Big[\cos\big((m-n)t\big) - \cos\big((m+n)t\big)\Big]\,dt \\ &= \begin{cases} 0, & \text{if } m \neq n, \\ \pi, & \text{if } m = n \end{cases}\end{aligned}$$

holds for any $m, n \in \mathbb{N}$. By (c) weak convergence $f_n \overset{w}{\rightharpoonup} 0$ as $n \to \infty$ follows.

Remark. The assumption that H is infinite-dimensional was only used in (c) and (d). As weak and strong convergence are equivalent in finite-dimensional spaces, adaptions of (c) and (d) to the finite-dimensional situation are necessarily wrong. (a) and (b), however, hold in any Hilbert space. (b) can even be formulated so that weak convergence of $x_n \to x_\infty$ in a Banach space X and strong convergence of $\varphi_n \to \varphi_\infty$ in the dual space X^* imply the convergence $\varphi_n(x_n) \to \varphi_\infty(x_\infty)$.

9.3. Annihilating annihilators

Let X be a normed $\mathbb{K}$-vector space (with $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$).

- For every set $U \subseteq X$ let $U^\perp \subseteq X^*$ be defined by $U^\perp = \{\varphi \in X^* \colon \varphi(u) = 0 \text{ for all } u \in U\}$.
- For every set $\Phi \subseteq X^*$ let ${}^\perp\Phi \subseteq X$ be defined by ${}^\perp\Phi = \{x \in X \colon \varphi(x) = 0 \text{ for all } \varphi \in \Phi\}$.

Prove for all $\emptyset \neq U \subseteq X$ and $\emptyset \neq \Phi \subseteq X^*$ that ${}^\perp(U^\perp) = \overline{\operatorname{span}(U)}$ and $\overline{\operatorname{span}(\Phi)} \subseteq ({}^\perp\Phi)^\perp$.

Solution: Let $\emptyset \neq U \subseteq X$. Then it holds for all $\varphi \in U^\perp$ that $\varphi(u) = 0$ for all $u \in U$. By linearity, this extends to $\varphi(u) = 0$ for all $u \in \operatorname{span}(U)$, $\varphi \in U^\perp$ and by continuity, we even get $\varphi(u) = 0$ for all $u \in \overline{\operatorname{span}(U)}$, $\varphi \in U^\perp$. Hence, we obtain $\overline{\operatorname{span}(U)} \subseteq {}^\perp(U^\perp)$. For the opposite inclusion, let us consider an arbitrary $u \in {}^\perp(U^\perp) \setminus \overline{\operatorname{span}(U)}$ (if existent). Note that $A = \{u\}$ is a non-empty, convex and compact set while $B = \overline{\operatorname{span}(U)}$ is a non-empty, convex and closed set. Since, in addition, $A \cap B = \emptyset$, there exist $\varphi \in X^*$, $\lambda \in \mathbb{R}$ such that $\varphi(u) < \lambda \leq \inf_{b\in B} \varphi(b)$. As B is a linear space, $\inf_{b\in B} \varphi(b)$ can only be 0 (in which case $\varphi|_B \equiv 0$) or $-\infty$, the latter being impossible as $\varphi|_B$ is bounded below by $\varphi(u)$. Long story short, there exists $\varphi \in X^*$ such that $\varphi(u) \neq 0$ but $\varphi|_B \equiv 0$ (and, in particular, $\varphi|_U \equiv 0$). In other words, there exists $\varphi \in U^\perp$ with $\varphi(u) \neq 0$, which proves that $u \notin {}^\perp(U^\perp)$. Thus, we have shown that ${}^\perp(U^\perp) \subseteq \overline{\operatorname{span}(U)}$, which concludes the proof of ${}^\perp(U^\perp) = \overline{\operatorname{span}(U)}$.

For the second claim, let $\emptyset \neq \Phi \subseteq X^*$. Then it holds for all $u \in {}^\perp\Phi$ that $\varphi(u) = 0$ for all $\varphi \in \Phi$. By linearity, this extends to $\varphi(u) = 0$ for all $\varphi \in \operatorname{span}(\Phi)$ and by continuity, we get $\varphi(u) = 0$ for all $u \in {}^\perp\Phi$, $\varphi \in \overline{\operatorname{span}(\Phi)}$. Thus, $\overline{\operatorname{span}(\Phi)} \subseteq ({}^\perp\Phi)^\perp$.

9.4. Duals and quotient spaces

Let $(X, \|\cdot\|_X)$ be a normed $\mathbb{K}$-vector space (with $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$) and $U \subseteq X$ a closed subspace.

(a) Prove that $(X/U)^*$ is isometrically isomorphic to $U^\perp$.

Solution: Let $\pi := (X \ni x \mapsto x + U \in X/U)$ denote the canonical projection. As π is a linear and continuous mapping from X to X/U (i.e., $\pi \in L(X, X/U)$), it holds for every $\Phi \in (X/U)^*$ that $\Phi \circ \pi \in X^*$. Hence, the mapping $T\colon (X/U)^* \to X^*$, defined by

$$T\Phi = \Phi \circ \pi \quad \text{for all } \Phi \in (X/U)^*,$$

is well-defined. T is clearly a linear mapping. Moreover, for all $\Phi \in (X/U)^*$, $x \in X$ it holds that

$$|(T\Phi)(x)| = |\Phi(\pi(x))| \leq \|\Phi\|_{(X/U)^*}\|\pi(x)\|_{X/U} \leq \|\Phi\|_{(X/U)^*}\|x\|_X,$$

that is, $\|T\Phi\|_{X^*} \leq \|\Phi\|_{(X/U)^*}$ for all $\Phi \in (X/U)^*$. On the other hand, for every $\Phi \in (X/U)^*$ we can find $(x_n)_{n\in\mathbb{N}} \subseteq X$ such that

- $\|\pi(x_n)\|_{(X/U)^*} = 1$ for all $n \in \mathbb{N}$,
- $\lim_{n\to\infty} \Phi(\pi(x_n)) = \|\Phi\|_{(X/U)*}$,

- $\|x_n\|_X < \|\pi(x_n)\|_{(X/U)^*} + \frac{1}{n}$ for every $n \in \mathbb{N}$,

which implies that

$$\|T\Phi\|_{X^*} \geq \sup_{n\in\mathbb{N}} \frac{(T\Phi)(x_n)}{\|x_n\|_X} = \sup_{n\in\mathbb{N}} \frac{\Phi(\pi(x_n))}{1+\frac{1}{n}} \geq \lim_{n\to\infty} \frac{\Phi(\pi(x_n))}{1+\frac{1}{n}} = \|\Phi\|_{(X/U)^*}.$$

Thus, we obtain for all $\Phi \in (X/U)^*$ that $\|T\Phi\|_{X^*} = \|\Phi\|_{(X/U)^*}$. In other words, T is an isometry (and, in particular, injective). In the following we are going to show that $\operatorname{im}(T) \subseteq U^\perp$ and $U^\perp \subseteq \operatorname{im}(T)$, which will complete the proof as it shows that the range of the isometry T is $U^\perp$.

"$\operatorname{im}(T) \subseteq U^\perp$": From $\pi(u) = 0$ for all $u \in U$ we get that $(T\Phi)(u) = 0$ for all $\Phi \in (X/U)^*$, $u \in U$. Hence, $T\Phi \in U^\perp$ for all $\Phi \in (X/U)^*$, which shows $\operatorname{im}(T) \subseteq U^\perp$.

"$U^\perp \subseteq \operatorname{im}(T)$": Define the mapping $S\colon U^\perp \to (X/U)^*$ via

$$(S\varphi)(x+U) = \varphi(x) \quad \text{for all } x \in X.$$

Since for all $\varphi \in U^\perp$ and $x, y \in X$ with $x + U = y + U$ it holds (as $x - y \in U$) that $\varphi(x-y) = 0$, we obtain that, for every $\varphi \in U^\perp$, $S\varphi\colon (X/U) \to \mathbb{R}$ is a well-defined mapping. Moreover, by linearity of π, every $\varphi \in U^\perp$ gives rise to a linear function $S\varphi$. Next, since for all $x \in X$ it holds that

$$\begin{aligned} |(S\varphi)(\pi(x))| &= \inf_{y\in\pi^{-1}(\pi(x))} |(S\varphi)(\pi(y))| = \inf_{y\in\pi^{-1}(\pi(x))} |\varphi(y)| \\ &\leq \inf_{y\in\pi^{-1}(\pi(x))} \|\varphi\|_{X^*}\|y\|_X = \|\varphi\|_{X^*}\|\pi(x)\|_{(X/U)^*}, \end{aligned}$$

we finally get that $S\colon U^\perp \to (X/U)^*$ is indeed well-defined. In addition, for all $\varphi \in U^\perp$, $x \in X$ it holds that

$$(TS\varphi)(x) = (S\varphi)(\pi(x)) = \varphi(x),$$

which proves that $U^\perp \subseteq \operatorname{im}(T)$.

(b) Prove that U^* is isometrically isomorphic to $X^*/U^\perp$.

Solution: Let $\Pi := X^* \ni x^* \mapsto x^* + U^\perp \in X^*/U^\perp$ be the canonical projection. Define the mapping $T\colon X^*/U^\perp \to U^*$ by

$$T(x^* + U^\perp) = x^*|_U \quad \text{for all } x^* \in X^*.$$

T is well-defined as for all $x^*, y^* \in X^*$ with $x^* + U^\perp = y^* + U^\perp$ it holds that $x^* - y^* \in U^\perp$ and therefore $(x^* - y^*)|_U \equiv 0$. Also, T is clearly a linear mapping.

Moreover, $x^*|_U$ belongs clearly to U^* if $x^* \in X^*$. Next, note that for all $x^* \in X^*$ it holds that

$$\|T(x^* + U^\perp)\|_{U^*} = \|x^*|_U\|_{U^*} \le \|x^*\|_{X^*}.$$

Hence,

$$\|T(x^* + U^\perp)\|_{U^*} \le \inf_{y^* \in x^* + U^\perp} \|y^*\|_{X^*} = \|x^* + U^\perp\|_{X^*/U^\perp} \quad \text{for all } x^* \in X^*.$$

Note that, according to the Hahn–Banach theorem, for every $u^* \in U^*$, there exists $x^* \in X^*$ with $x^*|_U = u^*$ and $\|x^*\|_{X^*} = \|u^*\|_{U^*}$. This implies that T is surjective and that

$$\|T(x^* + U^\perp)\|_{U^*} = \|x^*|_U\|_{U^*} \ge \|x^* + U^*\|_{X^* \setminus U^\perp} \quad \text{for every } x^* \in X^*.$$

Putting everything together, we have that T is a surjective isometry, which completes our proof.

(c) Prove that reflexivity of X implies reflexivity of U (in other words, closed subspaces of reflexive spaces are reflexive).

Solution: Let $u^{**} \in U^{**}$ be arbitrary but fixed. The map $X^* \ni x^* \mapsto u^{**}(x^*|_U) \in \mathbb{K}$ is clearly linear and bounded and therefore belongs to X^{**}. By the reflexivity of X, there exists $x \in X$ such that

$$x^*(x) = u^{**}(x^*|_U) \quad \text{for all } x^* \in X^*.$$

In particular, it holds for all $x^* \in U^\perp$ that $x^*(x) = 0$. Therefore, $x \in {}^\perp(U^\perp) = \overline{U} = U$ by Problem 9.3 (*Annihilating annihilators*). Since for every $u^* \in U^*$ there exists $x^* \in X^*$ with $x^*|_U = u^*$ by the Hahn–Banach theorem, we have that

$$u^*(x) = u^{**}(u^*) \quad \text{for all } u^* \in U^*.$$

Thus, the canonical embedding $\iota_U := (U \ni u \mapsto (U^* \ni u^* \mapsto u^*(u) \in \mathbb{K}) \in U^{**})$ is surjective. This proves that U is reflexive.

9.5. Invariant measures à la Krylov–Bogolioubov

Let (K, d) be a non-empty compact metric space and let $T\colon K \to K$ be continuous. Prove that there exists a Borel probability measure $\mu \in \mathcal{P}(K)$ on K satisfying for all Borel sets $A \subseteq K$ that $\mu(T^{-1}(A)) = \mu(A)$.

Hint: Use Problem 7.3 (*Banach limits*) to show that there exists $\varphi \in (C(K,\mathbb{R}))^*$ satisfying $\varphi \ge 0$, $\|\varphi\|_{(C(K,\mathbb{R}))^*} = 1$ and $\varphi(f) = \varphi(f \circ T)$ for all $f \in C(K,\mathbb{R})$. Conclude recalling **Riesz's representation theorem**:

With (K,d) being a compact metric space and with $\mathcal{M}(K)$ denoting the set of Borel regular finite signed measures on K, $\mathcal{M}(K)$ is isometrically isomorphic to $(C(K,\mathbb{R}))^*$ via the mapping $\Phi\colon \mathcal{M}(K) \to (C(K,\mathbb{R}))^*$, defined by

$$[\Phi(\mu)](f) = \int_K f\,d\mu \quad \text{for all } \mu \in \mathcal{M}(K), f \in C(K,\mathbb{R}).$$

In particular, the positive regular Borel measures correspond to the positive continuous linear functionals.

Solution: Let $\mathcal{T}\colon \ell^\infty \to \ell^\infty$ denote the left shift, i.e., $\mathcal{T}x = (x_{n+1})_{n\in\mathbb{N}}$ for all $x = (x_n)_{n\in\mathbb{N}} \in \ell^\infty$. From Problem 7.3 (*Banach limits*) we know that there exists $L \in (\ell^\infty)^*$ such that

- $\|L\|_{(\ell^\infty)^*} = 1$,
- $Lx \geq 0$ for all $x = (x_n)_{n\in\mathbb{N}} \in \ell^\infty$ satisfying $x_n \geq 0$ for all $n \in \mathbb{N}$,
- $Lx = L(\mathcal{T}x)$ for all $(x_n)_{n\in\mathbb{N}} \in \ell^\infty$.

Now, fix an arbitrary $x \in K$ and define the mapping $S\colon C(K,\mathbb{R}) \to \ell^\infty$ by

$$S(f) = ((f \circ T^n)(x))_{n\in\mathbb{N}} \quad \text{for all } f \in C(K,\mathbb{R}).$$

Note that S is well-defined because $\|S(f)\|_{\ell^\infty} \leq \sup_{x\in K}|f(x)| < \infty$ for every $f \in C(K,\mathbb{R})$ by compactness of K. Moreover, S is clearly linear and – by $\|S(f)\|_{\ell^\infty} \leq \sup_{x\in K}|f(x)|$ for all $f \in C(K,\mathbb{R})$ – bounded. Thus, $\varphi := L \circ S \in (C(K,\mathbb{R}))^*$. In addition, for all $f \in C(K,\mathbb{R})$ with $f \geq 0$ it holds that $S(f) \geq 0$ in ℓ^∞ and therefore also $\varphi(f) \geq 0$. Riesz‘s representation theorem therefore ensures that there exists a finite positive Borel regular measure μ on K such that for all $f \in C(K,\mathbb{R})$ it holds that $\varphi(f) = \int_K f\,d\mu$. Since

$$\mu(K) = \int_K 1\,d\mu = \varphi(K \ni x \mapsto 1 \in \mathbb{R}) = L((1)_{n\in\mathbb{N}}) = 1,$$

we obtain that μ is a probability measure. Furthermore, it holds for all $f \in C(K,\mathbb{R})$ that

$$\varphi(f \circ T) = L(S(f \circ T)) = L(\mathcal{T}S(f)) = L(S(f)) = \varphi(f).$$

This implies that

$$\int_K f\,d\mu = \int_K f \circ T\,d\mu \quad \text{for all } f \in C(K,\mathbb{R}).$$

It follows by standard measure-theoretic arguments that $\int_K f\,d\mu = \int_K f \circ T\,d\mu$ for all bounded Borel measurable $f\colon K \to \mathbb{R}$. In particular, for all Borel sets $A \subseteq K$, we get $\mu(A) = \mu(T^{-1}(A))$.

9.6. Optimal transport à la Kantorovich

Let (X, d_X) and (Y, d_Y) be non-empty compact metric spaces, let $c\colon X\times Y \to \mathbb{R}\cup\{\infty\}$ be lower semi-continuous, and let $\mu \in \mathcal{P}(X)$ and $\nu \in \mathcal{P}(Y)$ be probability measures on X and Y, respectively. We denote by $\Gamma(\mu,\nu)$ the set of probability measures on $X \times Y$ with first marginal μ and second marginal ν, i.e.,

$$\Gamma(\mu,\nu) = \left\{\gamma \in \mathcal{P}(X\times Y)\colon \begin{array}{c}\gamma(A\times Y) = \mu(A), \gamma(X\times B) = \nu(B) \\ \text{for all Borel sets } A \subseteq X, B \subseteq Y\end{array}\right\}.$$

Prove that there exists $\gamma \in \Gamma(\mu,\nu)$ satisfying that

$$\int_{X\times Y} c(x,y)\, d\gamma(x,y) = \inf_{\eta\in\Gamma(\mu,\nu)} \int_{X\times Y} c(x,y)\, d\eta(x,y).$$

Hint: Assume first that c is continuous. For general lower semi-continuous c, use that c can be written as pointwise limit of an increasing sequence $(f_k)_{k\in\mathbb{N}} \subseteq C(X\times Y,\mathbb{R})$.

Solution: Since $\mu\otimes\nu \in \Gamma(\mu,\nu)$, we know that $\Gamma(\mu,\nu)\neq\emptyset$. Since $X\times Y$ is compact and c is lower semi-continuous, $\inf_{(x,y)\in X\times Y} c(x,y) > -\infty$. Consequentially, we obtain for all $\eta\in\Gamma(\mu,\nu)$ that

$$\int_{X\times Y} c(x,y)\, d\eta(x,y) \geq \int_{X\times Y} \inf_{X\times Y} c\, d\eta(x,y) = \inf_{X\times Y} c > -\infty.$$

Let $(\gamma_n)_{n\in\mathbb{N}} \subseteq \Gamma(\mu,\nu)$ be a sequence satisfying

$$\lim_{n\to\infty} \int_{X\times Y} c(x,y)\, d\gamma_n(x,y) = \inf_{\eta\in\Gamma(\mu,\nu)} \int_{X\times Y} c(x,y)\, d\eta(x,y).$$

Since $\Gamma(\mu,\nu) \subseteq \mathcal{P}(X\times Y) \hookrightarrow (C(X\times Y,\mathbb{R}))^*$ and $C(X\times Y,\mathbb{R})$ is separable, we may by the Banach–Alaoglu theorem assume w.l.o.g. that $\gamma_n \xrightarrow{w^*} \gamma_\infty \in (C(K,\mathbb{R}))^*$. By the Riesz representation theorem we may (and will) interpret γ_∞ as an element of $\mathcal{M}(X\times Y)$. Due to

$$\int_{X\times Y} f(x,y)\, d\gamma_\infty(x,y) = \lim_{n\to\infty} \int_{X\times Y} f(x,y)\, d\gamma_n(x,y) \quad \text{for all } f \in C(X\times Y,\mathbb{R})$$

we get (by applying the above with $f \geq 0$ and with $f = (X\times Y \ni (x,y) \mapsto 1 \in \mathbb{R})$ respectively) that $\gamma_\infty \in \mathcal{P}(X\times Y)$. Moreover, for all $f\in C(X,\mathbb{R})$ and all $g \in C(Y,\mathbb{R})$ we have that

$$\int_{X\times Y} f(x)\, d\gamma_\infty(x,y) = \lim_{n\to\infty}\int_{X\times Y} f(x)\, d\gamma_n(x,y) = \int_X f(x)\, d\mu(x)$$

and

$$\int_{X\times Y} g(y)\, d\gamma_\infty(x,y) = \lim_{n\to\infty}\int_{X\times Y} g(y)\, d\gamma_n(x,y) = \int_Y g(y)\, d\nu(y),$$

i.e., $\gamma_\infty \in \Gamma(\mu,\nu)$. In the case that $c\colon X\times Y\to\mathbb{R}$ was continuous (and not only lower semi–continuous), we obtain

$$\int_{X\times Y} c(x,y)\,d\gamma_\infty(x,y) = \lim_{n\to\infty}\int_{X\times Y} c(x,y)\,d\gamma_n(x,y) = \inf_{\eta\in\Gamma(\mu,\nu)}\int_{X\times Y} c(x,y)\,d\eta(x,y).$$

In the general case, there exist (Lipschitz) continuous functions $(f_n)_{n\in\mathbb{N}}\subseteq C(X\times Y,\mathbb{R})$ with $f_n \geq \inf_{(x,y)\in X\times Y} c(x,y)$ for all $n\in\mathbb{N}$ and with $c(x,y)=\sup_{n\in\mathbb{N}} f_n(x,y)$ for all $(x,y)\in X\times Y$. With this, we obtain for every $m\in\mathbb{N}$ that

$$\begin{aligned}\int_{X\times Y} f_m(x,y)\,d\gamma_\infty(x,y) &= \lim_{n\to\infty}\int_{X\times Y} f_m(x,y)\,d\gamma_n(x,y)\\ &\leq \limsup_{n\to\infty}\int_{X\times Y} c(x,y)\,d\gamma_n(x,y)\\ &= \inf_{\eta\in\Gamma(\mu,\nu)}\int_{X\times Y} c(x,y)\,d\eta(x,y)\end{aligned}$$

Lebesgue's monotone convergence theorem implies that the left hand side converges to $\int_{X\times Y} c(x,y)\,d\gamma_\infty(x,y)$ as $m\to\infty$.

9.7. Minimal Energy

Let $m\in\mathbb{N}$ and let $\Omega\subseteq\mathbb{R}^m$ be a bounded measurable set with $|\Omega|>0$. For $g\in L^2(\mathbb{R}^m,\mathbb{R})$, we define the map

$$\begin{aligned} V\colon L^2(\Omega,\mathbb{R}) &\to \mathbb{R}\\ f &\mapsto \int_\Omega\int_\Omega g(x-y)f(x)f(y)\,dy\,dx\end{aligned}$$

and given $h\in L^2(\Omega,\mathbb{R})$, we define the map

$$\begin{aligned} E\colon L^2(\Omega,\mathbb{R}) &\to \mathbb{R}\\ f &\mapsto \|f-h\|^2_{L^2(\Omega,\mathbb{R})} + V(f).\end{aligned}$$

(a) Prove that V is weakly sequentially continuous.

Solution: Given a bounded measurable $\Omega\subseteq\mathbb{R}^m$ and $g\in L^2(\mathbb{R}^m,\mathbb{R})$, the goal is weak sequential continuity of the map

$$\begin{aligned} V\colon L^2(\Omega,\mathbb{R}) &\to \mathbb{R}\\ f &\mapsto \int_\Omega\int_\Omega g(x-y)f(x)f(y)\,dy\,dx.\end{aligned}$$

Claim 1. The linear operator $T\colon L^2(\Omega,\mathbb{R})\to L^2(\Omega,\mathbb{R})$ mapping $f\mapsto Tf$ given by

$$(Tf)(x)=\int_\Omega g(x-y)f(y)\,dy$$

is well-defined.

Proof. Let $f\in L^2(\Omega,\mathbb{R})$. Note that $(Tf)(x)$ is well-defined for every $x\in\Omega$ by the Cauchy–Schwarz inequality. Since $\Omega\subseteq\mathbb{R}^m$, being a bounded set, has finite volume $|\Omega|<\infty$, we obtain in addition that $Tf\in L^2(\Omega,\mathbb{R})$:

$$\begin{aligned}\|Tf\|^2_{L^2(\Omega,\mathbb{R})}&=\int_\Omega|(Tf)(x)|^2\,dx=\int_\Omega\biggl|\int_\Omega g(x-y)f(y)\,dy\biggr|^2\,dx\\&\le\int_\Omega\biggl(\int_\Omega|g(x-y)f(y)|\,dy\biggr)^2\,dx\le\int_\Omega\biggl(\int_\Omega|g(x-y)|^2\,dy\biggr)\|f\|^2_{L^2(\Omega,\mathbb{R})}\,dx\\&\le\int_\Omega\|g\|^2_{L^2(\mathbb{R}^m,\mathbb{R})}\|f\|^2_{L^2(\Omega,\mathbb{R})}\,dx\le|\Omega|\|g\|^2_{L^2(\mathbb{R}^m,\mathbb{R})}\|f\|^2_{L^2(\Omega,\mathbb{R})}<\infty.\end{aligned}$$ □

Claim 2. Let $(f_k)_{k\in\mathbb{N}}$ be a sequence in $L^2(\Omega,\mathbb{R})$ such that $f_k\xrightharpoonup{w}f$ in $L^2(\Omega,\mathbb{R})$ as $k\to\infty$. Then, $\|Tf_k-Tf\|_{L^2(\Omega,\mathbb{R})}\to0$ as $k\to\infty$, where T is as in Claim 1.

Proof of 2. Since the sequence $(f_k)_{k\in\mathbb{N}}$ is weakly convergent, it is bounded (by Banach–Steinhaus): $\exists C\in[0,\infty)\ \forall k\in\mathbb{N}:\ \|f_k\|_{L^2(\Omega,\mathbb{R})}\le C$. For every fixed $x_0\in\Omega$ and $k\in\mathbb{N}$, there holds

$$\begin{aligned}|(Tf_k)(x_0)|&\le\int_\Omega|g(x_0-y)f_k(y)|\,dy\le\biggl(\int_\Omega|g(x_0-y)|^2\,dy\biggr)^{\frac12}\biggl(\int_\Omega|f_k(y)|^2\,dy\biggr)^{\frac12}\\&\le\|g\|_{L^2(\mathbb{R}^m,\mathbb{R})}\|f_k\|_{L^2(\Omega,\mathbb{R})}.\end{aligned}$$

In particular, the map $f_k\mapsto(Tf_k)(x_0)$ is a linear continuous functional $L^2(\Omega,\mathbb{R})\to\mathbb{R}$. Therefore, weak convergence $f_k\xrightharpoonup{w}f$ implies $(Tf_k)(x_0)\to(Tf)(x_0)$ as $k\to\infty$. In other words, Tf_k converges pointwise to Tf. Moreover,

$$\sup_{k\in\mathbb{N}}|(Tf_k)(x_0)|\le\sup_{k\in\mathbb{N}}\Bigl(\|g\|_{L^2(\mathbb{R}^m,\mathbb{R})}\|f_k\|_{L^2(\Omega,\mathbb{R})}\Bigr)\le C\|g\|_{L^2(\mathbb{R}^m,\mathbb{R})}.$$

Since Ω is bounded, the constant $C\|g\|_{L^2(\mathbb{R}^m,\mathbb{R})}$ on the right right hand side belongs $L^2(\Omega,\mathbb{R})$. Hence, the claim follows by Lebesgue's dominated convergence theorem. □

Claim 3. Let $(f_k)_{k\in\mathbb{N}}$ be a sequence in $L^2(\Omega,\mathbb{R})$ such that $f_k\xrightharpoonup{w}f$ in $L^2(\Omega,\mathbb{R})$ as $k\to\infty$. Then, $V(f_k)\to V(f)$ as $k\to\infty$, i. e. V is weakly sequentially continuous.

Proof. Let T be as in Claim 1. Since $f_k \xrightarrow{w} f$ and $\|Tf_k - Tf\|_{L^2(\Omega,\mathbb{R})} \to 0$ as $k \to \infty$ by claim 2, we conclude

$$V(f_k) = \int_\Omega f_k(x) \int_\Omega g(x-y) f_k(y)\, dy\, dx = \langle f_k, Tf_k \rangle_{L^2(\Omega)} \xrightarrow{k\to\infty} \langle f, Tf \rangle = V(f),$$

using the continuity property of scalar products proven in Problem 9.2 (b). ☐

(b) Under the assumption $g \geq 0$ almost everywhere, prove that E restricted to

$$L^2_+(\Omega,\mathbb{R}) := \{f \in L^2(\Omega,\mathbb{R}) \mid f(x) \geq 0 \text{ for almost every } x \in \Omega\}$$

attains a global minimum.

Solution: In the case that $0 \leq g \in L^2(\mathbb{R}^m,\mathbb{R})$ and $h \in L^2(\Omega,\mathbb{R})$ the claim is that the map

$$\begin{aligned} E\colon L^2(\Omega,\mathbb{R}) &\to \mathbb{R} \\ f &\mapsto \|f-h\|^2_{L^2(\Omega,\mathbb{R})} + V(f) \end{aligned}$$

restricted to $L^2_+(\Omega,\mathbb{R})$ attains a global minimum. Since $L^2(\Omega,\mathbb{R})$ is reflexive (being a Hilbert space), we may invoke the direct method in the calculus of variations if we prove the following claims.

Claim 4. $L^2_+(\Omega,\mathbb{R})$ is non-empty and weakly sequentially closed.

Proof. Clearly, $L^2_+(\Omega,\mathbb{R}) \ni 0$ is non-empty. Let $(f_k)_{k\in\mathbb{N}}$ be a sequence in $L^2_+(\Omega,\mathbb{R})$ such that $f_k \xrightarrow{w} f$ for some $f \in L^2(\Omega,\mathbb{R})$ as $k \to \infty$. Suppose $f \notin L^2_+(\Omega,\mathbb{R})$. Then there exists $U \subseteq \Omega$ with positive measure such that $f|_U < 0$. In particular, we can test the weak convergence with the characteristic function χ_U to obtain the contradiction

$$0 > \langle f, \chi_U \rangle_{L^2(\Omega,\mathbb{R})} = \lim_{k\to\infty} \langle f_k, \chi_U \rangle \geq 0.$$ ☐

Claim 5. $E\colon L^2_+(\Omega,\mathbb{R}) \to \mathbb{R}$ is coercive and weakly sequentially lower semi-continuous.

Proof of Claim 5. Since $V(f) \geq 0$ if both $g \geq 0$ and $f \geq 0$ almost everywhere, we have

$$\begin{aligned} E(f) \geq \|f-h\|^2_{L^2(\Omega,\mathbb{R})} &\geq \|f\|^2_{L^2(\Omega,\mathbb{R})} - 2\|f\|_{L^2(\Omega,\mathbb{R})}\|h\|_{L^2(\Omega,\mathbb{R})} + \|h\|^2_{L^2(\Omega,\mathbb{R})} \\ &\geq \frac{1}{2}\|f\|^2_{L^2(\Omega,\mathbb{R})} - \|h\|^2_{L^2(\Omega,\mathbb{R})} \end{aligned}$$

for every $f \in L^2_+(\Omega,\mathbb{R})$ as we have by Young's inequality that $2ab \leq \frac{1}{2}a^2 + 2b^2$ for all $a, b \in \mathbb{R}$. Since $h \in L^2(\Omega,\mathbb{R})$ is fixed, E is coercive.

By part (a), $L^2(\Omega,\mathbb{R}) \ni f \mapsto V(f) \in \mathbb{R}$ is weakly sequentially lower semi-continuous. Moreover, every term on the right hand side of

$$\|f-h\|^2_{L^2(\Omega,\mathbb{R})} = \|f\|^2_{L^2(\Omega,\mathbb{R})} - 2\langle f,h\rangle_{L^2(\Omega,\mathbb{R})} + \|h\|^2_{L^2(\Omega,\mathbb{R})}$$

is weakly sequentially lower semi-continuous in f since h is fixed. This proves the claim. ☐

9.8. A result by Lions-Stampacchia

Let $(H,(\cdot,\cdot)_H)$ be a real Hilbert space and let $a\colon H\times H\to\mathbb{R}$ be a bilinear map so that:

(i) $a(x,y)=a(y,x)$ for every $x,y\in H$,

(ii) there exists $\Lambda\in(0,\infty)$ so that $|a(x,y)|\le\Lambda\|x\|_H\|y\|_H$ for every $x,y\in H$,

(iii) there exists $\lambda\in(0,\infty)$ so that $a(x,x)\ge\lambda\|x\|^2_H$ for every $x\in H$.

Let moreover $f\colon H\to\mathbb{R}$ be a continuous linear functional. Consider the map $J\colon H\to\mathbb{R}$ given by

$$J(x)=a(x,x)-2f(x).$$

Finally, let $K\subseteq H$ be a non-empty closed convex subset.

(a) Prove that there exists a *unique* $y_0\in K$ such that $J(y_0)\le J(z)$ for every $z\in K$.

Solution: The special structure of the terms involved allows to give here a solution based on Problem 5.6 (*Projections on closed convex sets*). A standard argument in the spirit of the direct method of the calculus of variations would of course be possible as well.

Claim 1. Given $f\in H^*$, there exists a unique $x_0\in H$ such that for all $x\in H$

$$J(x):=a(x,x)-2f(x)=a(x-x_0,x-x_0)-a(x_0,x_0).$$

Proof. Since a is bilinear and satisfies (ii) and (iii) the Lax–Milgram theorem applies ((ii) implies continuity of a). In particular, since $f\in H^*$, there exists a unique $x_0\in H$ satisfying $a(x_0,x)=f(x)$ for all $x\in H$. (The same follows from claim 2 below and the Riesz representation theorem applied in $(H,a(\cdot,\cdot))$). Moreover,

$$\begin{aligned} J(x)=a(x,x)-2f(x) &= a(x,x)-2a(x_0,x)\\ &= a(x-x_0,x)-a(x_0,x)\\ &= a(x-x_0,x-x_0)+a(x-x_0,x_0)-a(x,x_0)\\ &= a(x-x_0,x-x_0)-a(x_0,x_0) \end{aligned}$$ ☐

for all $x\in H$, as claimed.

Claim 2. $(H, a(\cdot,\cdot))$ is a Hilbert space.

Proof. By assumption (i) the bilinear map a is symmetric. By (ii) and (iii), we have

$$\lambda\|x\|_H^2 \leq a(x,x) \leq \Lambda\|x\|_H^2 \tag{$*$}$$

which shows $a(x,x) \geq 0$ and $a(x,x) = 0 \Leftrightarrow x = 0$. Therefore, $a(\cdot,\cdot)$ is a scalar product on H. In fact, $(*)$ implies that the induced norm $\|\cdot\|_a = \sqrt{a(\cdot,\cdot)}$ is *equivalent* to $\|\cdot\|_H$. It is easy to check that equivalent norms have the same Cauchy-sequences and induce the same notion of convergence. Therefore, $(H, \|\cdot\|_a)$ is complete since $(H, \|\cdot\|_H)$ is complete and the claim follows. □

By assumption, the set $\emptyset \neq K \subseteq H$ is convex and closed in $(H, \|\cdot\|_H)$. Since the two norms are equivalent, K is also closed in $(H, \|\cdot\|_a)$ and we can apply the result of part (a) of Problem 5.6 (*Projections on closed convex sets*) in the $\mathbb{R}$-Hilbert space $(H, a(\cdot,\cdot))$ with the point x_0 from claim 1. That is: there exists a unique $y_0 \in K$ satisfying

$$\|x_0 - y_0\|_a = \inf_{y\in K}\|x_0 - y\|_a. \tag{$\dagger$}$$

By Claim 1 we have for arbitrary $y \in K$

$$J(y_0) = \|y_0 - x_0\|_a^2 - \|x_0\|_a^2 \leq \|y - x_0\|_a^2 - \|x_0\|_a^2 = J(y).$$

Moreover, since y_0 is the unique element of K satisfying $(\dagger)$, it is also the unique element of K satisfying $J(y_0) \leq J(y)$ for all $y \in K$.

(b) Prove that the unique minimizer y_0 from (a) is also the unique element of K satisfying $a(y_0, z - y_0) \geq f(z - y_0)$ for every $z \in K$.

Solution: We saw in part (a) that y_0 is the unique element of K with $\|x_0 - y_0\|_a = \inf_{y\in K}\|x_0 - y\|_a$. By part (b) of Problem 5.6 (*Projections on closed convex sets*) y_0 is therefore the unique element of K which satisfies

$$a(x_0 - y_0, z - y_0) \leq 0 \quad \text{for all } z \in K.$$

This and the fact that $a(x_0, x) = f(x)$ for all $x \in H$ implies that y_0 is the unique element of K such that

$$f(z - y_0) = a(x_0, z - y_0) \leq a(y_0, z - y_0) \quad \text{for all } z \in K.$$