Chapter 4: The cornerstones of Functional Analysis

4.1 Hahn-Banach Theorem

Recall:

Definition Let M be a set, and $D \subseteq M \times M$. Let $\prec$ ("sub") be the associated binary relation: $x \prec y :\Leftrightarrow (x,y) \in D$.

(a) $\prec$ is a partial ordering (or, M is a partially ordered set)
$:\Leftrightarrow \forall x,y,z \in M$:
(i) $x \prec x$ (reflexive)
(ii) $(x \prec y \ \& \ y \prec x) \Rightarrow (x = y)$ (antisymmetry)
(iii) $(x \prec y \ \& \ y \prec z) \Rightarrow (x \prec z)$ (transitive)

(b) $x, y \in M$ are comparable $:\Leftrightarrow x \prec y$ or $y \prec x$
x, y are incomparable $:\Leftrightarrow x, y$ are not comparable.

(c) $\prec$ is a total ordering $:\Leftrightarrow$
(i) $\prec$ is a partial ordering
(ii) x, y are comparable for all $x, y \in M$

(d) Let $W \subseteq M$. Then $u \in M$ is an upper bound of/for W
$:\Leftrightarrow w \prec u \ \forall w \in W$.

(e) $u \in W$ is a maximal element of W : $\Leftrightarrow$ The following implication holds:
$$u \prec w \text{ for } w \in W \Rightarrow u = w.$$

(Note: A maximal element need not be an upper bound and vice versa - see examples below).

Examples (a) "$\leq$" is a total ordering on $\mathbb{R}$.
$W = [0,1)$ has no maximal element, but any $u \geq 1$ is an upper bound for W

(b) "$\subseteq$" is a partial ordering on $\mathcal{P}(X)$, but not a total ordering.

c) $(x_1, x_2) \prec (y_1, y_2) :\Leftrightarrow x_j \leq y_j$, $j = 1,2$,
is a partial ordering (but not total) on $M = \mathbb{R}^2$.
$W := \{(0,0), (1,0), (0,1)\}$ has 2 (!) maximal elements,
but none of them is an upper bound for W.

The following axiom is equivalent to the axiom of choice:

Axiom 4.1 (Zorn's Lemma) Let $M \neq \emptyset$ be a partially ordered set. Assume every totally ordered subset of M has an upper bound. Then M has a maximal element.

We now finish the

Proof of Thm. 2.53 (Remains: Every Hilbert space $X \neq \{0\}$ has an orthonormal basis)

Let $M := \{E \subseteq X \mid E \text{ orthonormal}\}$. Then:

(i) $M \neq \emptyset$

(ii) "$\subseteq$" is partial ordering on M

(iii) Let W be a totally ordered subset of M.

Then $U := \bigcup_{E \in W} E \in M$:

If $x_1, x_2 \in U$ then $\exists E_j \in W : x_j \in E_j$, $j = 1,2$.

As W is totally ordered, $E_1 \subseteq E_2$ (w.l.o.g.)

Hence, $x_1, x_2 \in E_2 \Rightarrow x_1 \perp x_2 \Rightarrow U$ orthonormal

That is, U is an upper bound for W.

By Zorn: $\exists$ maximal element $N \in M$.

Claim: $\mathcal{M}$ is an orthonormal basis for $\overline{X}$

(i) Orthonormal clear, since $\mathcal{M} \in \mathcal{M}$

(ii) Assume $\mathcal{M}$ is not complete, i.e., not a basis.

Then (by Def. 2.47(d)), there is some $0 \neq x \in \overline{X}$ such that $x \perp u$ $\forall u \in \mathcal{M}$. Hence,

$$\mathcal{M}' := \left\{ \frac{x}{\|x\|} \right\} \cup \mathcal{M} \in \mathcal{M} \quad \text{(since } \mathcal{M}' \text{ is orthonormal)}$$

so $\mathcal{M} \subsetneq \mathcal{M}'$ which contradicts $\mathcal{M}$ being maximal ∎

Remark 4.2 Similar arguments prove Thm. 2.4 (existence of Hamel basis; see exercise)

Theorem 4.3 (Hahn-Banach (H-B)) Let X be a vectorspace, let $p: X \to \mathbb{R}$ be convex, i.e. $\forall x, x' \in X$, $\forall \alpha \in [0,1]$:

$$p(\alpha x + (1-\alpha)x') \leq \alpha p(x) + (1-\alpha) p(x') \quad (*)$$

Let Y be a subspace, and let $\lambda: Y \to \mathbb{K}$ be linear, with

$$\operatorname{Re} \lambda(y) \leq p(y) \quad \forall y \in Y \quad (\Pi)$$

Then there exists $\Lambda: X \to \mathbb{K}$ linear such that:

(i) $\Lambda|_Y = \lambda$ (i.e. Λ is extension of λ)

(ii) $\operatorname{Re} \Lambda(x) \leq p(x)$ $\forall x \in X$ (i.e. (Π) is preserved)

If, in addition, p also satisfies

$(**)$ $\quad p(\alpha x) \leq p(x)$ $\forall x \in X$, $\forall \alpha \in \mathbb{K}$ with $|\alpha| = 1$

then we even have $|\Lambda(x)| \leq p(x)$ $\forall x \in X$.

Remark 4.4 (a) Condition $(**)$ is equivalent to:

$$p(\alpha x) = p(x) \quad \forall x \in \mathbb{K} \text{ with } |\alpha| = 1 \quad \text{(check!)}$$

(b) If p is a (semi-) norm on X, then p satisfies $(*)$ & $(**)$.

Pf (of 4.3): 4 Steps: Steps 1 & 2 prove the main part for $\mathbb{R}$-vectorspaces, step 3 for $\mathbb{C}$-vector-spaces, and step 4 proves the addendum under add. condition (**).
Wlog: $Y \subsetneq X$ (otherwise trivial).

Step 1: Case $\mathbb{K}=\mathbb{R}$. Extend by one dimension – a preparation for Step 2.
By assumption: $\exists\, z \in X \setminus Y$ (so $z \neq 0$). Let $\tilde{Y} := \operatorname{span}(Y, \{z\})$.
For every $\tilde{y} \in \tilde{Y}$, $\exists!$ decomposition $\tilde{y} = y + \alpha z$, $y \in Y$, $\alpha \in \mathbb{R}$
Candidate for extension $\tilde{\lambda}$ of λ to $\tilde{Y}$:

$$\tilde{\lambda}(\tilde{y}) := \lambda(y) + \alpha\xi \qquad \text{for some } \xi \in \mathbb{R} \text{ (to be chosen)}$$

Interpretation: $\xi = \tilde{\lambda}(z)$. Clearly, $\tilde{\lambda}$ is ($\mathbb{R}$-) linear on $\tilde{Y}$, and $\tilde{\lambda}|_Y = \lambda$. Will choose ξ s.t. $\tilde{\lambda} \leq p$ on $\tilde{Y}$ (recall (iii)):
Let $\beta_1, \beta_2 > 0$, $y_1, y_2 \in Y$, then

$$\beta_1\lambda(y_1) + \beta_2\lambda(y_2) = (\beta_1+\beta_2)\,\lambda\Big(\underbrace{\tfrac{\beta_1}{\beta_1+\beta_2}y_1 + \tfrac{\beta_2}{\beta_1+\beta_2}y_2}\Big) \quad (\nabla)$$

$$\overset{\text{by (II)}}{\leq} p\Big(\underbrace{\tfrac{\beta_1}{\beta_1+\beta_2}y_1 + \tfrac{\beta_2}{\beta_1+\beta_2}y_2}\Big)$$

$$= \tfrac{\beta_1}{\beta_1+\beta_2}(y - \beta_2 z) + \tfrac{\beta_2}{\beta_1+\beta_2}(y_2 + \beta_1 z)$$

Hence,

$$(\nabla) \overset{p \text{ convex}}{\leq} \beta_1 p(y_1 - \beta_2 z) + \beta_2 p(y_2 + \beta_1 z).$$

Re-arranging, we get

$$\frac{1}{\beta_2}\big(\lambda(y_1) - p(y_1 - \beta_2 z)\big) \leq \frac{1}{\beta_1}\big(p(y_2 + \beta_1 z) - \lambda(y_2)\big)$$

$$\forall\, \beta_1, \beta_2 > 0,\ \forall\, y_1, y_2 \in Y$$

Hence, there exists $a \in \mathbb{R}$:

$$(\nabla\nabla)\quad \sup_{\substack{\beta_2>0\\ \gamma_1\in Y}} \left[\frac{1}{\beta_2}(\lambda(\gamma_1)-p(\gamma_1-\beta_2 z)\right] \le a \le \inf_{\substack{\beta_1>0\\ \gamma_2\in Y}} \left[\frac{1}{\beta_1}(p(\gamma_2+\beta_1 z)-\lambda(\gamma_2)\right]$$

Set $\xi = \tilde\lambda(z) := a$. Then, $\forall\, \tilde\gamma = \gamma + \alpha z \in \tilde Y$ with $\alpha > 0$, the right inequality in $(\nabla\nabla)$ implies

$$\tilde\lambda(z) \le \frac{1}{\alpha}(p(\tilde\gamma) - \lambda(\gamma)), \text{ so } \tilde\lambda(\tilde\gamma) \le p(\tilde\gamma).$$

If $\tilde\gamma = \gamma - \alpha z$ with $\alpha > 0$, use the left ineq. in $(\nabla\nabla)$ instead. Therefore, $\tilde\lambda \le p$ on $\tilde Y$.

Step 2: Case $\mathbb{K} = \mathbb{R}$. Idea: Use Zorn to construct the extension.

Let $M := \{(\mathbb{R}\text{-})$ linear extensions e of λ with $e \le p$ on $\mathrm{dom}(e)\}$

We have

(i) $M \neq \emptyset$ since $\lambda \in M$

(ii) Define partial ordering $\prec$ on M:

$$e_1 \prec e_2 :\Leftrightarrow \mathrm{dom}(e_1) \subseteq \mathrm{dom}(e_2) \wedge e_2|_{\mathrm{dom}(e_1)} = e_1.$$

(iii) Let $W \subseteq M$ be totally ordered.

Define $u : \bigcup_{e\in W} \mathrm{dom}(e) \to \mathbb{R}$, $x \mapsto \tilde e(x)$ where $\tilde e$ is any element of W such that $x \in \mathrm{dom}(\tilde e)$.

u is well-defined: (i.e. indep. of the chosen $\tilde e$ among allowed elements):
Let $x \in \mathrm{dom}(\tilde e_1) \cap \mathrm{dom}(\tilde e_2)$. Since W is totally ordered, we have (wlog.) $\tilde e_2$ is an extension of $\tilde e_1$. Hence, $\tilde e_2(x) = \tilde e_1(x)$.

u is $(\mathbb{R}\text{-})$ linear: Let $x_1, x_2 \in \bigcup_{e\in W} \mathrm{dom}(e)$, $\gamma \in \mathbb{R}$. Hence, $\exists\, \tilde e_j \in W : x_j \in \mathrm{dom}(\tilde e_j)$, $j = 1, 2$. Since W is totally ordered, (wlog.) $\tilde e_2$ is extension of $\tilde e_1$. Then $x_1, x_2, x_1 + \gamma x_2 \in \mathrm{dom}(\tilde e_2)$ & linearity of u follows from that of $\tilde e_2$.

Lastly: $u(x) = \tilde{e}(x) \leq p(x) \quad \forall x \in \bigcup_{e \in W} \operatorname{dom}(e)$.

Hence: $u \in M$, and (check!) u is an upper bound for W.

By Zorn's Lemma: M has a maximal element Λ.

Claim: $\operatorname{dom}(\Lambda) = X$: Suppose $\operatorname{dom}(\Lambda) \subsetneq X$. Then there is some $0 \neq z \in X \setminus \operatorname{dom}(\Lambda)$. Let $\tilde{Y} := \operatorname{span}(\operatorname{dom}(\Lambda), \{z\})$.

By Step 1, Λ has some ($\mathbb{R}$-) linear extension $\tilde{\Lambda} \in M$ to $\tilde{Y}$, which contradicts M being maximal.

Hence, the main part of the thm. follows for $\mathbb{K} = \mathbb{R}$.

Step 3: Case $\mathbb{K} = \mathbb{C}$. Reduce to the real case.

Define $\ell(y) := \operatorname{Re}\lambda(y) \quad \forall y \in Y$. Then $\ell: X \to \mathbb{R}$ is $\mathbb{R}$-linear, and $\ell \leq p$ on Y. So, Step 2 implies: There exists $\mathbb{R}$-linear functional $L: X \to \mathbb{R}$ with $L|_Y = \ell$ and $L \leq p$ on X.

Note: $\lambda(y) = \ell(y) + i\ell(-iy) \quad \forall y \in Y$ since:

$$\ell(-iy) = \operatorname{Re}\lambda(-iy) \overset{\lambda\ \mathbb{C}\text{-lin}}{=} \operatorname{Re}[-i\lambda(y)] = \operatorname{Im}\lambda(y)$$

Define $\Lambda(x) := L(x) + iL(-ix)$, $x \in X$.

Then, by Step 2,

(i) Λ is $\mathbb{R}$-linear on X

(ii) $\Lambda|_Y = \lambda$

(iii) $\operatorname{Re}\Lambda \overset{L(-ix) \in \mathbb{R}}{=} L \leq p$ on X

Also, Λ is $\mathbb{C}$-linear, since if is $\mathbb{R}$-linear and

$$\Lambda(ix) = L(ix) + iL(x) \overset{L\ \mathbb{R}\text{-lin}}{=} i(L(x) + iL(-ix)) = i\Lambda(x),$$

proving the main part for $\mathbb{K} = \mathbb{C}$.

Step 4: Addendum: We fix $x \in X$ and use the polar repr.

$$\Lambda(x) = |\Lambda(x)| e^{i\theta(x)} \quad [\text{If } \mathbb{K} = \mathbb{R} \text{ then } e^{i\theta(x)} \in \{-1, 1\}]$$

Then $|\Lambda(x)| = e^{-i\theta(x)}\Lambda(x) \overset{\mathbb{K}\text{-lin}}{=} \Lambda(e^{-i\theta(x)}x) \overset{\text{LHS} \in \mathbb{R}}{=} \operatorname{Re}\Lambda(e^{-i\theta(x)}x)$

$\overset{\text{main part}}{\leq} p(e^{-i\theta(x)}x) \overset{(**)}{\leq} p(x)$ ∎

Corollary 4.5 Let X be a normed space, $Y \subseteq X$ a subspace, and $\varphi \in Y^*$. Then there exists $f \in X^*$ with $f|_Y = \varphi$ and $\|f\|_{X^*} = \|\varphi\|_{Y^*}$.

Pf: Apply Thm. 4.3 with $p(x) := \|\varphi\|_{Y^*} \cdot \|x\|$ for $x \in X$ (fulfills assumptions - check!) to $\lambda = \varphi$. Then $\exists f: X \to \mathbb{K}$ linear with $|f(x)| \leq \|\varphi\|_{Y^*}\|x\|$, hence $\|f\|_{X^*} \leq \|\varphi\|_{Y^*}$.

But $$\|f\|_{X^*} = \sup_{0 \neq x \in X} \frac{|f(x)|}{\|x\|} \geq \sup_{0 \neq x \in Y} \frac{|\varphi(x)|}{\|x\|} = \|\varphi\|_{Y^*}$$ ∎

Corollary 4.6 Let X be a normed space and let $0 \neq x_0 \in X$. Then there exists $f \in X^*$ with $f(x_0) = \|x_0\|$ and $\|f\|_{X^*} = 1$.

Pf: Let $Y = \operatorname{span}\{x_0\}$. If $y \in Y$ then $y = \alpha x_0$ for some unique $\alpha \in \mathbb{K}$. Define φ on Y by $\varphi(y) := \alpha \cdot \|x_0\|$.

This implies $\varphi(x_0) = \|x_0\|$ and $\varphi \in Y^*$ with $\|\varphi\|_{Y^*} = 1$, since $|\varphi(y)| = \|y\|$. By Cor. 4.5, $\exists f: X \to \mathbb{K}$ linear, with $f|_Y = \varphi$ (in particular, $f(x_0) = \|x_0\|$) and $\|f\|_{X^*} = \|\varphi\|_{Y^*} = 1$ ∎

Corollary 4.7 Let X be a normed space, $Z \subseteq X$ a closed subspace, and $x_0 \in X \setminus Z$ with $0 < \operatorname{dist}(x_0, Z) =: d$. Then there exists $f \in X^*$ with $f|_Z = 0$, $f(x_0) = d$ and $\|f\|_{X^*} = 1$.

Proof: See exercise.