Functional Analysis

Notes de Cours / Lecture Notes

Graduate Level — 2025–2026

Yaë Ulrich Gaba

From Banach spaces to operator algebras: a journey through modern functional analysis.

March 25, 2026

Contents

Chapter 1

Banach Spaces — Foundations

Contents

Functional analysis extends the ideas of linear algebra and real analysis to infinite-dimensional spaces. The natural framework is that of a *normed vector space*, and the single most important property such a space can enjoy is *completeness*. A complete normed space is called a *Banach space*, after Stefan Banach, who laid the foundations of the subject in his 1932 monograph.

Throughout this chapter $\mathbb{K}$ denotes either $\mathbb{R}$ or $\mathbb{C}$, and all vector spaces are over $\mathbb{K}$ unless stated otherwise.

1.1 Normed vector spaces

Definition 1.1 (Normed vector space)**.** Let E be a vector space over $\mathbb{K}$. A **norm** on E is a function $\|\cdot\| : E \to [0, +\infty)$ satisfying, for all $x, y \in E$ and $\lambda \in \mathbb{K}$:

(N1) $\|x\| = 0 \iff x = 0$ (positive definiteness);

(N2) $\|\lambda x\| = |\lambda| \, \|x\|$ (absolute homogeneity);

(N3) $\|x + y\| \leq \|x\| + \|y\|$ (triangle inequality).

The pair $(E, \|\cdot\|)$ is called a **normed vector space** (or simply a **normed space**).

Every norm induces a metric $d(x, y) = \|x - y\|$ and hence a topology on E. We write $B(x, r) = \{y \in E : \|y - x\| < r\}$ for the open ball and $\overline{B}(x, r) = \{y \in E : \|y - x\| \leq r\}$ for the closed ball. The **unit ball** is $B_E = \overline{B}(0, 1)$.

Definition 1.2 (Seminorm)**.** A function $p : E \to [0, +\infty)$ satisfying (N2) and (N3) (but not necessarily (N1)) is called a **seminorm**.

Remark 1.3. If p is a seminorm, then $N = \{x \in E : p(x) = 0\}$ is a vector subspace, and p induces a norm on the quotient E/N.

1.1.1 Fundamental examples

Example 1.4 (Euclidean space $\mathbb{R}^n$)**.** On $\mathbb{R}^n$ (or $\mathbb{C}^n$) the most common norms are, for $x = (x_1, \ldots, x_n)$:

$$\|x\|_p = \left(\sum_{k=1}^n |x_k|^p \right)^{1/p}, \qquad 1 \leq p < \infty, \qquad (1.1)$$

$$\|x\|_\infty = \max_{1 \leq k \leq n} |x_k|. \qquad\qquad\qquad (1.2)$$

The case $p = 2$ gives the standard Euclidean norm. That $\|\cdot\|_p$ is indeed a norm for $p \geq 1$ follows from **Minkowski's inequality** (see definition 1.9 below).

Example 1.5 (Sequence spaces ℓ^p)**.** For $1 \leq p < \infty$ define

$$\ell^p = \Big\{ x = (x_n)_{n \geq 1} \subset \mathbb{K} : \sum_{n=1}^{\infty} |x_n|^p < \infty \Big\}, \qquad \|x\|_{\ell^p} = \Big(\sum_{n=1}^{\infty} |x_n|^p \Big)^{1/p}.$$

For $p = \infty$:

$$\ell^{\infty} = \Big\{ x = (x_n)_{n \geq 1} \subset \mathbb{K} : \sup_n |x_n| < \infty \Big\}, \qquad \|x\|_{\ell^{\infty}} = \sup_{n \geq 1} |x_n|.$$

The subspace $c_0 \subset \ell^{\infty}$ consists of sequences converging to zero; $c \subset \ell^{\infty}$ consists of all convergent sequences. Both are closed subspaces of ℓ^{∞}.

Example 1.6 (Space of continuous functions $C(K)$)**.** Let K be a compact topological space. The space

$$C(K) = \{ f : K \to \mathbb{K} \mid f \text{ is continuous} \}$$

equipped with the **supremum norm** $\|f\|_{\infty} = \sup_{t \in K} |f(t)|$ is a normed space. For $K = [a, b] \subset \mathbb{R}$ we often write $C([a, b])$.

Example 1.7 (Preview: L^p spaces)**.** Let $(\Omega, \mathcal{A}, \mu)$ be a measure space. For $1 \leq p < \infty$ one defines

$$L^p(\Omega, \mu) = \Big\{ f : \Omega \to \mathbb{K} \text{ measurable} : \int_{\Omega} |f|^p \, \mathrm{d}\mu < \infty \Big\} \Big/ \sim,$$

where $f \sim g$ iff $f = g$ μ-a.e. The norm is $\|f\|_{L^p} = \left(\int |f|^p \, \mathrm{d}\mu \right)^{1/p}$. These spaces will be discussed in detail later; for now they serve as motivation.

1.1.2 Hölder's and Minkowski's inequalities

The fundamental tool for proving that $\|\cdot\|_p$ is a norm is the following pair of classical inequalities.

Proposition 1.8 (Hölder's inequality)**.** *Let* $1 \leq p \leq \infty$ *and* $\frac{1}{p} + \frac{1}{q} = 1$

(with $q = \infty$ when $p = 1$). For any sequences $(a_n), (b_n) \in \mathbb{K}^{\mathbb{N}}$:

$$\sum_{n=1}^{\infty} |a_n b_n| \leq \left(\sum_{n=1}^{\infty} |a_n|^p \right)^{1/p} \left(\sum_{n=1}^{\infty} |b_n|^q \right)^{1/q}.$$

In particular, if $a \in \ell^p$ and $b \in \ell^q$ then $ab \in \ell^1$.

Proof. The cases $p = 1$ or $p = \infty$ are immediate. Assume $1 < p < \infty$. Recall Young's inequality: for $\alpha, \beta \geq 0$,

$$\alpha\beta \leq \frac{\alpha^p}{p} + \frac{\beta^q}{q}.$$

If $A = \|a\|_p = 0$ or $B = \|b\|_q = 0$ the result is trivial. Otherwise set $\alpha_n = |a_n| / A$ and $\beta_n = |b_n| / B$. Then

$$\frac{|a_n b_n|}{AB} = \alpha_n \beta_n \leq \frac{\alpha_n^p}{p} + \frac{\beta_n^q}{q}.$$

Summing over n:

$$\frac{1}{AB} \sum_n |a_n b_n| \leq \frac{1}{p} \sum_n \frac{|a_n|^p}{A^p} + \frac{1}{q} \sum_n \frac{|b_n|^q}{B^q} = \frac{1}{p} + \frac{1}{q} = 1. \qquad \square$$

Proposition 1.9 (Minkowski's inequality). *For $1 \leq p \leq \infty$ and $a, b \in \ell^p$:*

$$\|a + b\|_p \leq \|a\|_p + \|b\|_p.$$

Proof. For $p = 1$ and $p = \infty$ this is clear. Let $1 < p < \infty$. We have

$$\sum_n |a_n + b_n|^p \leq \sum_n |a_n + b_n|^{p-1} \left(|a_n| + |b_n| \right)$$

$$= \sum_n |a_n + b_n|^{p-1} |a_n| + \sum_n |a_n + b_n|^{p-1} |b_n|.$$

Apply Hölder's inequality to each sum with exponents q and p (noting $(p - 1)q = p$):

$$\sum_n |a_n + b_n|^p \leq \left(\sum_n |a_n + b_n|^p \right)^{1/q} \left(\|a\|_p + \|b\|_p \right).$$

Dividing both sides by $\left(\sum |a_n + b_n|^p \right)^{1/q}$ gives the result (the case when the sum is zero being trivial). $\qquad \square$

1.2 Convergence and series in normed spaces

Definition 1.10 (Convergence). A sequence (x_n) in a normed space $(E, \|\cdot\|)$ **converges** to $x \in E$ if $\|x_n - x\| \to 0$. We write $x_n \to x$ or $\lim_{n\to\infty} x_n = x$.

Definition 1.11 (Cauchy sequence). A sequence (x_n) in E is a **Cauchy sequence** if for every $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that $\|x_m - x_n\| < \varepsilon$ for all $m, n \geq N$.

Every convergent sequence is Cauchy. The converse is the content of completeness.

Definition 1.12 (Series). Let $(x_n)_{n\geq 1}$ be a sequence in E. The **series** $\sum_{n=1}^{\infty} x_n$ is said to **converge** if the sequence of partial sums $S_N = \sum_{n=1}^{N} x_n$ converges in E. The series **converges absolutely** if $\sum_{n=1}^{\infty} \|x_n\| < \infty$.

Remark 1.13. In general, absolute convergence does *not* imply convergence. In fact, the converse characterises Banach spaces (see definition 1.29 below).

1.3 Banach spaces: definition and completeness

Definition 1.14 (Banach space). A normed vector space $(E, \|\cdot\|)$ is called a **Banach space** if it is *complete*, i.e., every Cauchy sequence in E converges in E.

Theorem 1.15 (Completeness of ℓ^p). *For every $1 \leq p \leq \infty$, the space ℓ^p is a Banach space.*

Proof. We give the proof for $1 \leq p < \infty$; the case $p = \infty$ is analogous (and simpler).

Let $(x^{(k)})_{k \geq 1}$ be a Cauchy sequence in ℓ^p, where $x^{(k)} = \left(x_n^{(k)}\right)_{n \geq 1}$. For each fixed n,

$$\left|x_n^{(k)} - x_n^{(j)}\right| \leq \left\|x^{(k)} - x^{(j)}\right\|_{\ell^p} \to 0 \qquad \text{as } k, j \to \infty,$$

so $(x_n^{(k)})_{k \geq 1}$ is Cauchy in $\mathbb{K}$, hence converges to some $x_n \in \mathbb{K}$. Set $x = (x_n)_{n \geq 1}$.

Step 1: $x \in \ell^p$. Let $\varepsilon > 0$ and choose K such that $\left\|x^{(k)} - x^{(j)}\right\|_p < \varepsilon$ for all $k, j \geq K$. For any $N \in \mathbb{N}$:

$$\sum_{n=1}^{N} \left|x_n^{(k)} - x_n^{(j)}\right|^p < \varepsilon^p \quad \text{for all } k, j \geq K.$$

Letting $j \to \infty$ (with N fixed): $\sum_{n=1}^{N} \left|x_n^{(k)} - x_n\right|^p \leq \varepsilon^p$. Since this holds for all N,

$$\left\|x^{(k)} - x\right\|_p \leq \varepsilon \qquad \text{for all } k \geq K.$$

In particular $x = x^{(K)} + (x - x^{(K)}) \in \ell^p$.

Step 2: $x^{(k)} \to x$ in ℓ^p. The inequality above shows $\left\|x^{(k)} - x\right\|_p \leq \varepsilon$ for $k \geq K$, which is exactly $x^{(k)} \to x$. $\qquad\square$

> **Theorem 1.16** (Completeness of $C(K)$). *If K is a compact topological space, then $(C(K), \|\cdot\|_\infty)$ is a Banach space.*

Proof. Let (f_n) be a Cauchy sequence in $C(K)$. For each $t \in K$, $|f_m(t) - f_n(t)| \leq \|f_m - f_n\|_\infty \to 0$, so $(f_n(t))$ converges in $\mathbb{K}$. Define $f(t) = \lim_{n \to \infty} f_n(t)$.

Given $\varepsilon > 0$, choose N so that $\|f_m - f_n\|_\infty < \varepsilon$ for $m, n \geq N$. Letting $m \to \infty$ we get $|f(t) - f_n(t)| \leq \varepsilon$ for all $t \in K$ and $n \geq N$, i.e., the convergence is *uniform*. As a uniform limit of continuous functions, f is continuous, so $f \in C(K)$. Moreover $\|f - f_n\|_\infty \leq \varepsilon$ for $n \geq N$. $\qquad\square$

> *Remark* 1.17. The space $C([0,1])$ equipped with the L^1-norm $\|f\|_1 = \int_0^1 |f(t)| \, \mathrm{d}t$ is *not* complete. One can construct a Cauchy sequence of continuous functions whose L^1-limit is a discontinuous function (e.g., a step function).

1.4   Finite-dimensional normed spaces

Finite-dimensional spaces are the "tame" case: they enjoy several remarkable properties that fail spectacularly in infinite dimensions.

> **Theorem 1.18** (Equivalence of norms in finite dimension). *On a finite-dimensional vector space E, all norms are equivalent: if $\|\cdot\|_a$ and $\|\cdot\|_b$ are two norms on E with $\dim E = n < \infty$, there exist constants $0 < \alpha \leq \beta$ such that*
>
> $$\alpha \|x\|_a \leq \|x\|_b \leq \beta \|x\|_a \qquad \forall\, x \in E.$$

Proof. It suffices to show that every norm $\|\cdot\|$ on E is equivalent to the ℓ^1-norm relative to a basis. Fix a basis $e_1, \ldots, e_n$ and write $x = \sum_{i=1}^n \xi_i e_i$; set $\|x\|_1 = \sum_{i=1}^n |\xi_i|$.

Upper bound. By the triangle inequality and homogeneity,

$$\|x\| \leq \sum_{i=1}^n |\xi_i|\, \|e_i\| \leq M \|x\|_1, \qquad M = \max_i \|e_i\|.$$

Lower bound. The map $\varphi : (\mathbb{K}^n, \|\cdot\|_1) \to \mathbb{R}$ defined by $\varphi(\xi) = \|\sum_i \xi_i e_i\|$ is continuous (by the upper bound just proved) and positive on the unit sphere $S = \{x \in \mathbb{K}^n : \|x\|_1 = 1\}$. Since S is compact in $\mathbb{K}^n$ (here we use finite dimension), φ attains a minimum $m = \min_S \varphi > 0$. Hence for $x \neq 0$:

$$\|x\| = \|x\|_1 \cdot \varphi\big(x/\|x\|_1\big) \geq m \|x\|_1.$$

Thus $m \|x\|_1 \leq \|x\| \leq M \|x\|_1$. $\qquad \square$

> **Corollary 1.19.** *Every finite-dimensional normed space is a Banach space.*

Proof. By definition 1.18, any norm on $E \cong \mathbb{K}^n$ is equivalent to the Euclidean norm. Since $(\mathbb{K}^n, \|\cdot\|_2)$ is complete, so is $(E, \|\cdot\|)$. $\qquad \square$

> **Theorem 1.20** (Compactness of the closed unit ball in finite dimension). *A normed space E has the property that $B_E = \overline{B}(0,1)$ is compact if and only if $\dim E < \infty$.*

Proof. ($\Rightarrow$) If $\dim E < \infty$, then by equivalence of norms B_E is homeomorphic to the closed unit ball in $(\mathbb{K}^n, \|\cdot\|_2)$, which is compact by the Heine–Borel theorem.

($\Leftarrow$) Suppose $\dim E = \infty$. We construct a sequence in B_E with no convergent subsequence, using Riesz's lemma (definition 1.21 below). Start with $x_1 \in E$, $\|x_1\| = 1$. Set $E_1 = \operatorname{span}\{x_1\}$. By Riesz's lemma with $\theta = 1/2$,

there exists x_2 with $\|x_2\| = 1$ and $\|x_2 - y\| \geq 1/2$ for all $y \in E_1$. Set $E_2 = \mathrm{span}\{x_1, x_2\}$ and repeat. By induction we obtain $(x_n) \subset B_E$ with $\|x_m - x_n\| \geq 1/2$ for $m \neq n$, so no subsequence can be Cauchy. $\qquad\square$

> **Lemma 1.21** (Riesz's lemma). *Let E be a normed space, $F \subsetneq E$ a closed proper subspace, and $0 < \theta < 1$. Then there exists $x_\theta \in E$ with $\|x_\theta\| = 1$ and*
> $$\mathrm{dist}(x_\theta, F) = \inf_{y \in F} \|x_\theta - y\| \geq \theta.$$

Proof. Pick $z \in E \setminus F$. Since F is closed, $d = \mathrm{dist}(z, F) > 0$. Choose $y_0 \in F$ with $\|z - y_0\| \leq d/\theta$ (possible since $d/\theta > d$). Set

$$x_\theta = \frac{z - y_0}{\|z - y_0\|}.$$

Then $\|x_\theta\| = 1$, and for any $y \in F$:

$$\|x_\theta - y\| = \frac{\|z - y_0 - \|z - y_0\| \, y\|}{\|z - y_0\|} = \frac{\left\| z - \underbrace{(y_0 + \|z - y_0\| \, y)}_{\in F} \right\|}{\|z - y_0\|} \geq \frac{d}{\|z - y_0\|} \geq \frac{d}{d/\theta} = \theta.$$

$\qquad\square$

> *Remark* 1.22. One cannot take $\theta = 1$ in general. However, when F has finite codimension, the infimum $d = \mathrm{dist}(z, F)$ is attained and one can find x with $\mathrm{dist}(x, F) = 1$.

1.5   Continuous linear maps and the operator norm

> **Definition 1.23** (Bounded linear map). Let $(E, \|\cdot\|_E)$ and $(F, \|\cdot\|_F)$ be normed spaces. A linear map $T : E \to F$ is **bounded** if there exists $C \geq 0$ such that
> $$\|Tx\|_F \leq C \, \|x\|_E \qquad \forall\, x \in E.$$

> **Proposition 1.24** (Continuity = boundedness). *For a linear map $T : E \to F$ between normed spaces, the following are equivalent:*

(i) T *is continuous;*

(ii) T *is continuous at* 0*;*

(iii) T *is bounded;*

(iv) T *is Lipschitz.*

Proof. (i)$\Rightarrow$(ii) is trivial.

(ii)$\Rightarrow$(iii): Continuity at 0 gives $\delta > 0$ with $\|Tx\|_F < 1$ whenever $\|x\|_E < \delta$. For $x \neq 0$, set $y = \frac{\delta}{2\|x\|_E} x$; then $\|y\|_E < \delta$, so $\|Ty\|_F < 1$, i.e., $\|Tx\|_F < \frac{2}{\delta} \|x\|_E$.

(iii)$\Rightarrow$(iv): $\|Tx - Ty\|_F = \|T(x - y)\|_F \leq C \|x - y\|_E$.

(iv)$\Rightarrow$(i) is clear. $\qquad\square$

Definition 1.25 (Operator norm). For $T \in \mathcal{L}(E, F)$ (the space of bounded linear maps $E \to F$), the **operator norm** is

$$\|T\|_{\mathcal{L}(E,F)} = \sup_{\substack{x \in E \\ x \neq 0}} \frac{\|Tx\|_F}{\|x\|_E} = \sup_{\|x\|_E \leq 1} \|Tx\|_F = \sup_{\|x\|_E = 1} \|Tx\|_F .$$

Notation 1.26. We write $\mathcal{L}(E, F)$ for the space of bounded linear maps $E \to F$, and $E^* = \mathcal{L}(E, \mathbb{K})$ for the **(continuous) dual space**. When $E = F$ we write $\mathcal{L}(E) = \mathcal{L}(E, E)$.

Theorem 1.27 ($\mathcal{L}(E, F)$ is Banach when F is). *If F is a Banach space then $\mathcal{L}(E, F)$ is a Banach space. In particular the dual E^* is always a Banach space.*

Proof. Let (T_n) be a Cauchy sequence in $\mathcal{L}(E, F)$. For each $x \in E$,

$$\|T_m x - T_n x\|_F \leq \|T_m - T_n\| \, \|x\|_E \to 0,$$

so $(T_n x)$ is Cauchy in F, hence converges to some element we call Tx.

Linearity. For $\alpha, \beta \in \mathbb{K}$ and $x, y \in E$: $T(\alpha x + \beta y) = \lim_n T_n(\alpha x + \beta y) = \alpha \lim_n T_n x + \beta \lim_n T_n y = \alpha Tx + \beta Ty$.

Boundedness. Since (T_n) is Cauchy, it is bounded: say $\|T_n\| \leq M$ for all n. Then $\|Tx\|_F = \lim_n \|T_n x\|_F \leq M \|x\|_E$, so $T \in \mathcal{L}(E, F)$.

Convergence. Given $\varepsilon > 0$, pick N such that $\|T_m - T_n\| < \varepsilon$ for $m, n \geq N$. For any x with $\|x\| \leq 1$ and $n \geq N$:

$$\|T_m x - T_n x\|_F \leq \varepsilon.$$

Letting $m \to \infty$: $\|Tx - T_n x\|_F \leq \varepsilon$. Taking the sup over $\|x\| \leq 1$: $\|T - T_n\| \leq \varepsilon$. $\square$

Corollary 1.28. *For any normed space E, the dual space E^* is a Banach space.*

1.6 Absolutely convergent series and completeness

The following elegant characterisation of Banach spaces reduces completeness to a statement about series.

Theorem 1.29 (Characterisation of completeness)**.** *A normed space E is a Banach space if and only if every absolutely convergent series in E converges.*

Proof. ($\Rightarrow$) Suppose E is Banach and $\sum \|x_n\| < \infty$. The partial sums $S_N = \sum_{n=1}^{N} x_n$ form a Cauchy sequence since for $M > N$:

$$\|S_M - S_N\| = \left\| \sum_{n=N+1}^{M} x_n \right\| \leq \sum_{n=N+1}^{M} \|x_n\| \to 0 \quad \text{as } N, M \to \infty.$$

Hence (S_N) converges.

($\Leftarrow$) Let (y_n) be a Cauchy sequence. Choose a subsequence (y_{n_k}) with $\left\| y_{n_{k+1}} - y_{n_k} \right\| < 2^{-k}$. Set $x_k = y_{n_{k+1}} - y_{n_k}$. Then $\sum \|x_k\| < \sum 2^{-k} = 1 < \infty$, so by hypothesis $\sum x_k$ converges, i.e., the telescoping partial sums $y_{n_{k+1}} - y_{n_1}$ converge. Thus (y_{n_k}) converges, and since (y_n) is Cauchy with a convergent subsequence, (y_n) itself converges. $\square$

1.7 Quotient spaces

Definition 1.30 (Quotient norm)**.** Let E be a normed space and $F \subset E$ a *closed* subspace. On the quotient vector space E/F we define

$$\|\dot{x}\|_{E/F} = \inf_{y \in F} \|x - y\|_E = \operatorname{dist}(x, F), \qquad \dot{x} = x + F.$$

> **Proposition 1.31.** *The quotient norm is indeed a norm on E/F (closedness of F is essential for positive definiteness). Moreover, if E is a Banach space, so is E/F.*

Proof. **Norm axioms.**

- *Positive definiteness:* $\|\dot{x}\| = 0$ means $\text{dist}(x, F) = 0$, so $x \in \overline{F} = F$ (since F is closed), i.e., $\dot{x} = \dot{0}$.

- *Homogeneity:* $\|\lambda\dot{x}\| = \inf_{y \in F} \|\lambda x - y\| = |\lambda| \inf_{y \in F} \|x - y/\lambda\| = |\lambda| \, \|\dot{x}\|$ for $\lambda \neq 0$.

- *Triangle inequality:* For any $y_1, y_2 \in F$: $\|(x_1 + x_2) - (y_1 + y_2)\| \leq \|x_1 - y_1\| + \|x_2 - y_2\|$. Taking infima over $y_1, y_2 \in F$ gives $\|\dot{x}_1 + \dot{x}_2\| \leq \|\dot{x}_1\| + \|\dot{x}_2\|$.

Completeness. We use definition 1.29. Let $\sum_n \|\dot{x}_n\|_{E/F} < \infty$. For each n, pick a representative $z_n \in \dot{x}_n$ with $\|z_n\|_E < \|\dot{x}_n\|_{E/F} + 2^{-n}$. Then $\sum \|z_n\|_E < \infty$, so since E is Banach, $S = \sum_n z_n$ converges in E. The partial sums of $\sum \dot{x}_n$ equal $\dot{S}_N$ (the class of $\sum_{k=1}^N z_k$), and $\left\|\dot{S}_N - \dot{S}\right\| \leq \|S_N - S\|_E \to 0$. $\qquad\square$

The canonical projection $\pi : E \to E/F$, $\pi(x) = \dot{x}$ is a bounded linear map with $\|\pi\| = 1$ (when $F \neq E$).

1.8 Completion of a normed space

> **Theorem 1.32** (Completion). *Let $(E, \|\cdot\|)$ be a normed space. There exists a Banach space $\widehat{E}$ and an isometric linear injection $\iota : E \hookrightarrow \widehat{E}$ such that $\iota(E)$ is dense in $\widehat{E}$. The pair $(\widehat{E}, \iota)$ is unique up to isometric isomorphism.*

Proof (sketch). Consider the set $\mathcal{C}$ of all Cauchy sequences in E. Define an equivalence relation $(x_n) \sim (y_n)$ iff $\|x_n - y_n\| \to 0$. Set $\widehat{E} = \mathcal{C}/\sim$. Vector space operations are defined componentwise, and the norm $\|[(x_n)]\|_{\widehat{E}} = \lim_n \|x_n\|_E$ (which exists since $(\|x_n\|)$ is Cauchy in $\mathbb{R}$) is well-defined. The map $\iota : E \to \widehat{E}$ sends x to the class of the constant sequence $(x, x, x, \dots)$. One verifies that $\widehat{E}$ is complete and $\iota(E)$ is dense; we leave the details as an exercise.

For uniqueness, if $(\widetilde{E}, j)$ is another completion, the map $j \circ \iota^{-1} : \iota(E) \to \widetilde{E}$ is an isometry between dense subsets of Banach spaces; by the BLT theorem (Bounded Linear Transformation extension), it extends to an isometric isomorphism $\widehat{E} \to \widetilde{E}$. $\qquad\square$

The following commutative diagram summarises the universal property of the completion:

$$
\begin{array}{ccc}
E & \overset{\iota}{\longhookrightarrow} & \widehat{E} \\
 & {}_{T}\searrow & \downarrow {}^{\widehat{T}} \\
 & & F
\end{array}
$$

Given any bounded linear map $T : E \to F$ into a Banach space F, there exists a unique $\widehat{T} : \widehat{E} \to F$ with $\widehat{T} \circ \iota = T$ and $\left\|\widehat{T}\right\| = \|T\|$.

1.9   Exercises for Chapter 1

Exercise 1.1 ($\star$). Show that the norm $\|\cdot\| : E \to \mathbb{R}$ is a (Lipschitz) continuous function: $\big|\|x\| - \|y\|\big| \le \|x - y\|$.

Exercise 1.2 ($\star$). Let E be a Banach space and $F \subset E$ a subspace. Prove that F is a Banach space (with the induced norm) if and only if F is closed in E.

Exercise 1.3 ($\star$). Prove that c_0 is a closed subspace of ℓ^∞, hence a Banach space.

Exercise 1.4 ($\star\star$). Show that ℓ^p is separable for $1 \le p < \infty$, but ℓ^∞ is not separable.

Exercise 1.5 ($\star\star$). Let E, F, G be normed spaces, $T \in \mathcal{L}(E, F)$ and $S \in \mathcal{L}(F, G)$. Prove that $S \circ T \in \mathcal{L}(E, G)$ and $\|S \circ T\| \le \|S\| \, \|T\|$.

Exercise 1.6 ($\star\star$). Let E be a Banach space and $T \in \mathcal{L}(E)$ with

$\|T\| < 1$. Prove that $\mathrm{Id} - T$ is invertible in $\mathcal{L}(E)$ and

$$(\mathrm{Id} - T)^{-1} = \sum_{n=0}^{\infty} T^n,$$

with $\|(\mathrm{Id} - T)^{-1}\| \leq \frac{1}{1 - \|T\|}$. *Hint:* use the characterisation of completeness via absolutely convergent series.

Exercise 1.7 ($\star\star$)**.** Deduce from the previous exercise that the set $GL(E)$ of invertible elements of $\mathcal{L}(E)$ is open, and that the map $T \mapsto T^{-1}$ is continuous on $GL(E)$.

Exercise 1.8 ($\star\star$)**.** Let $\|\cdot\|_a$ and $\|\cdot\|_b$ be two norms on a vector space E. Show they are equivalent if and only if the identity map $\mathrm{Id} : (E, \|\cdot\|_a) \to (E, \|\cdot\|_b)$ is a homeomorphism.

Exercise 1.9 ($\star\star\star$)**.** Prove that the dual of ℓ^1 is isometrically isomorphic to ℓ^∞. *Hint:* given $\varphi \in (\ell^1)^*$, define $a_n = \varphi(e_n)$ and show $a \in \ell^\infty$ with $\|a\|_\infty = \|\varphi\|$.

Exercise 1.10 ($\star\star\star$)**.** Let E be a normed space and F a closed subspace. Prove that the completion of E/F is isometrically isomorphic to $\widehat{E}/\overline{F}$, where $\overline{F}$ denotes the closure of $\iota(F)$ in $\widehat{E}$.

Exercise 1.11 ($\star$)**.** Show that a norm $\|\cdot\|$ on a vector space E satisfies the **parallelogram law**

$$\|x + y\|^2 + \|x - y\|^2 = 2\big(\|x\|^2 + \|y\|^2\big)$$

for all $x, y \in E$ if and only if it arises from an inner product via $\|x\| = \sqrt{\langle x, x \rangle}$. This is the **Jordan–von Neumann theorem**.

Exercise 1.12 ($\star\star$)**.** Show that ℓ^p for $p \neq 2$ does not satisfy the parallelogram law, hence is not a Hilbert space.

Exercise 1.13 ($\star\star$). Give an example of a Banach space E and a closed subspace $F \subsetneq E$ such that there is no x with $\|x\| = 1$ and $\mathrm{dist}(x, F) = 1$. *Hint:* consider c_0 inside ℓ^∞.

Exercise 1.14 ($\star\star\star$). Let $C_c(\mathbb{R})$ be the space of continuous functions with compact support, equipped with the L^p-norm. Show that $C_c(\mathbb{R})$ is not complete for any $1 \le p < \infty$, and that its completion is $L^p(\mathbb{R})$.

Chapter 2

Hilbert Spaces

Contents

Hilbert spaces are Banach spaces whose norm arises from an inner product. This extra structure is enormously powerful: it gives rise to orthogonal projections, the Riesz representation theorem, and the theory of orthonormal bases, bringing the subject much closer to the intuition of finite-dimensional linear algebra. The archetypal Hilbert space is L^2, and the theory developed here underpins quantum mechanics, signal processing, and the modern theory of PDEs.

2.1   Inner products and the Cauchy–Schwarz inequality

Definition 2.1 (Inner product). Let H be a vector space over $\mathbb{K}$ ($= \mathbb{R}$ or $\mathbb{C}$). An **inner product** (or **scalar product**) is a map $\langle \cdot, \cdot \rangle : H \times H \to \mathbb{K}$ satisfying:

(IP1) $\langle \alpha x_1 + \beta x_2, y \rangle = \alpha \langle x_1, y \rangle + \beta \langle x_2, y \rangle$ for all $\alpha, \beta \in \mathbb{K}$, $x_1, x_2, y \in H$ (linearity in the first variable);

(IP2) $\langle x, y \rangle = \overline{\langle y, x \rangle}$ (conjugate symmetry);

(IP3) $\langle x, x \rangle \geq 0$, with equality iff $x = 0$ (positive definiteness).

The pair $(H, \langle \cdot, \cdot \rangle)$ is called a **pre-Hilbert space** (or **inner product space**).

Remark 2.2. We adopt the convention that the inner product is linear in the *first* argument (the "physicist's convention" is the opposite). Note that (IP1) and (IP2) together give conjugate-linearity in the second argument: $\langle x, \alpha y \rangle = \bar{\alpha} \langle x, y \rangle$.

Notation 2.3. Every inner product induces a norm:

$$\|x\| = \sqrt{\langle x, x \rangle}.$$

We write $x \perp y$ ("x is orthogonal to y") when $\langle x, y \rangle = 0$.

Theorem 2.4 (Cauchy–Schwarz inequality). *For all x, y in a pre-Hilbert space H:*
$$|\langle x, y \rangle| \leq \|x\| \, \|y\|,$$
with equality if and only if x and y are linearly dependent.

Proof. If $y = 0$ both sides vanish. Assume $y \neq 0$ and let $\lambda \in \mathbb{K}$. Then

$$0 \leq \|x - \lambda y\|^2 = \|x\|^2 - \lambda \langle y, x \rangle - \bar{\lambda} \langle x, y \rangle + |\lambda|^2 \|y\|^2.$$

Choose $\lambda = \langle x, y \rangle / \|y\|^2$. Then

$$0 \leq \|x\|^2 - \frac{|\langle x, y \rangle|^2}{\|y\|^2} - \frac{|\langle x, y \rangle|^2}{\|y\|^2} + \frac{|\langle x, y \rangle|^2}{\|y\|^4} \|y\|^2 = \|x\|^2 - \frac{|\langle x, y \rangle|^2}{\|y\|^2}.$$

Rearranging gives $|\langle x, y \rangle|^2 \leq \|x\|^2 \|y\|^2$. Equality holds iff $\|x - \lambda y\| = 0$, i.e., $x = \lambda y$. $\qquad \square$

> **Corollary 2.5.** *The function* $\|x\| = \sqrt{\langle x, x \rangle}$ *is indeed a norm.*

Proof. Only the triangle inequality needs verification:

$$\|x + y\|^2 = \|x\|^2 + 2 \operatorname{Re} \langle x, y \rangle + \|y\|^2$$
$$\leq \|x\|^2 + 2 \|x\| \|y\| + \|y\|^2 = (\|x\| + \|y\|)^2. \qquad \square$$

> **Proposition 2.6** (Polarization identity). *The inner product can be recovered from the norm. Over* $\mathbb{R}$:
>
> $$\langle x, y \rangle = \frac{1}{4} \left(\|x + y\|^2 - \|x - y\|^2 \right).$$
>
> *Over* $\mathbb{C}$:
>
> $$\langle x, y \rangle = \frac{1}{4} \sum_{k=0}^{3} i^k \|x + i^k y\|^2.$$

Proof. Direct expansion using $\|x + y\|^2 = \|x\|^2 + 2 \operatorname{Re} \langle x, y \rangle + \|y\|^2$. For the complex case, the four terms $\|x + y\|^2$, $i \|x + iy\|^2$, $- \|x - y\|^2$, $-i \|x - iy\|^2$ yield $4 \langle x, y \rangle$ when summed. $\qquad \square$

> **Proposition 2.7** (Parallelogram law). *In any inner product space:*
>
> $$\|x + y\|^2 + \|x - y\|^2 = 2 \left(\|x\|^2 + \|y\|^2 \right).$$
>
> *Conversely, a norm satisfying this identity comes from an inner product (Jordan–von Neumann theorem).*

Proof. Expanding both sides: $\|x + y\|^2 + \|x - y\|^2 = (\|x\|^2 + 2 \operatorname{Re} \langle x, y \rangle + \|y\|^2) + (\|x\|^2 - 2 \operatorname{Re} \langle x, y \rangle + \|y\|^2) = 2 \|x\|^2 + 2 \|y\|^2$. $\qquad \square$

2.2   Hilbert spaces: definition and examples

Definition 2.8 (Hilbert space)**.** A **Hilbert space** is a pre-Hilbert space that is complete with respect to the norm induced by its inner product.

Example 2.9 ($\mathbb{K}^n$)**.** The space $\mathbb{K}^n$ with $\langle x, y \rangle = \sum_{k=1}^{n} x_k \bar{y}_k$ is a Hilbert space of dimension n.

Example 2.10 (ℓ^2)**.** The space ℓ^2 with inner product $\langle x, y \rangle = \sum_{n=1}^{\infty} x_n \bar{y}_n$ is a Hilbert space. The series converges absolutely by the Cauchy–Schwarz inequality for sequences. Completeness was proved in definition 1.15.

Example 2.11 ($L^2(\Omega, \mu)$)**.** For a measure space $(\Omega, \mathcal{A}, \mu)$, the space $L^2(\Omega, \mu)$ with inner product

$$\langle f, g \rangle = \int_\Omega f \, \bar{g} \, \mathrm{d}\mu$$

is a Hilbert space (the Fischer–Riesz theorem, definition 2.30).

Example 2.12 (Sobolev space $H^1(\Omega)$ — preview)**.** Let $\Omega \subset \mathbb{R}^n$ be open. The Sobolev space

$$H^1(\Omega) = \left\{ f \in L^2(\Omega) : \partial_i f \in L^2(\Omega) \text{ for } i = 1, \ldots, n \right\}$$

(derivatives in the distributional sense) is a Hilbert space with inner product

$$\langle f, g \rangle_{H^1} = \int_\Omega f \bar{g} \, \mathrm{d}x + \sum_{i=1}^{n} \int_\Omega (\partial_i f) \overline{(\partial_i g)} \, \mathrm{d}x.$$

We will return to Sobolev spaces in later chapters.

2.3   Projection onto closed convex sets

The following theorem is the cornerstone of Hilbert space geometry.

Theorem 2.13 (Projection onto a closed convex set). *Let H be a Hilbert space and $C \subset H$ a non-empty, closed, convex set. For every $x \in H$ there exists a unique $p \in C$ such that*

$$\|x - p\| = \operatorname{dist}(x, C) := \inf_{y \in C} \|x - y\|.$$

Moreover, p is characterised by:

$$p \in C \quad \text{and} \quad \operatorname{Re} \langle x - p, y - p \rangle \leq 0 \quad \forall\, y \in C. \tag{2.1}$$

Proof. Set $d = \operatorname{dist}(x, C)$ and choose a minimising sequence $(y_n) \subset C$ with $\|x - y_n\| \to d$.

Existence (via the parallelogram law). For $m, n \in \mathbb{N}$, since C is convex, $\frac{y_m + y_n}{2} \in C$, so $\left\|x - \frac{y_m + y_n}{2}\right\| \geq d$. The parallelogram law gives:

$$
\begin{aligned}
\|y_m - y_n\|^2 &= \|(y_m - x) - (y_n - x)\|^2 \\
&= 2\|y_m - x\|^2 + 2\|y_n - x\|^2 - \|(y_m - x) + (y_n - x)\|^2 \\
&= 2\|y_m - x\|^2 + 2\|y_n - x\|^2 - 4\left\|\frac{y_m + y_n}{2} - x\right\|^2 \\
&\leq 2\|y_m - x\|^2 + 2\|y_n - x\|^2 - 4d^2.
\end{aligned}
$$

As $m, n \to \infty$, the right side tends to $2d^2 + 2d^2 - 4d^2 = 0$. Hence (y_n) is Cauchy, and since H is complete and C is closed, $y_n \to p \in C$ with $\|x - p\| = d$.

Uniqueness. If p, p' both achieve the minimum, apply the parallelogram law to $y_m = p$, $y_n = p'$ to get $\|p - p'\|^2 \leq 0$.

Characterisation. ($\Rightarrow$) Let $y \in C$ and $t \in (0, 1]$. Since C is convex, $p + t(y - p) \in C$, so

$$d^2 \leq \|x - p - t(y - p)\|^2 = d^2 - 2t \operatorname{Re} \langle x - p, y - p \rangle + t^2 \|y - p\|^2.$$

Hence $2 \operatorname{Re} \langle x - p, y - p \rangle \leq t \|y - p\|^2$. Letting $t \to 0^+$ gives $\operatorname{Re} \langle x - p, y - p \rangle \leq 0$.

($\Leftarrow$) If p satisfies (2.1), then for any $y \in C$:

$$\|x - y\|^2 = \|(x - p) - (y - p)\|^2 = \|x - p\|^2 - 2 \operatorname{Re} \langle x - p, y - p \rangle + \|y - p\|^2 \geq \|x - p\|^2.$$

$\square$

We denote the projection of x onto C by $P_C(x)$.

Proposition 2.14 (Projection is a contraction). *The map $P_C : H \to C$ is a contraction: $\|P_C(x) - P_C(y)\| \leq \|x - y\|$ for all $x, y \in H$.*

Proof. Set $p = P_C(x), q = P_C(y)$. By the characterisation: $\text{Re} \langle x - p, q - p \rangle \leq 0$ and $\text{Re} \langle y - q, p - q \rangle \leq 0$. Adding:

$$\text{Re} \langle (x - p) - (y - q), q - p \rangle \leq 0,$$

i.e., $\text{Re} \langle (x - y) - (p - q), q - p \rangle \leq 0$, hence $\text{Re} \langle x - y, q - p \rangle \leq \text{Re} \langle p - q, q - p \rangle = - \|p - q\|^2$. Therefore $\|p - q\|^2 \leq \text{Re} \langle x - y, p - q \rangle \leq \|x - y\| \ \|p - q\|$. $\square$

2.3.1   Projection onto a closed subspace

When $C = F$ is a *closed subspace*, the characterisation simplifies considerably.

Theorem 2.15 (Projection onto a closed subspace). *Let F be a closed subspace of a Hilbert space H. For every $x \in H$:*

(i) *There exists a unique $p \in F$ nearest to x, characterised by*

$$p \in F \quad and \quad x - p \in F^{\perp}, \tag{2.2}$$

*where $F^{\perp} = \{z \in H : \langle z, y \rangle = 0 \ \forall y \in F\}$ is the **orthogonal complement** of F.*

(ii) *$H = F \oplus F^{\perp}$ (algebraic and topological direct sum).*

(iii) *The projection $P_F : H \to F$ is a bounded linear operator with $\|P_F\| = 1$ (when $F \neq \{0\}$), $P_F^2 = P_F$, and $\langle P_F x, y \rangle = \langle x, P_F y \rangle$ for all $x, y \in H$ (i.e., P_F is self-adjoint).*

Proof. (i) Since F is a closed convex set, definition 2.13 gives a unique $p \in F$ minimising $\|x - p\|$. The characterisation (2.1) gives $\text{Re} \langle x - p, y - p \rangle \leq 0$ for all $y \in F$. Since F is a subspace, for any $z \in F$ and $t \in \mathbb{R}$ we may take $y = p + tz$ to get $t \, \text{Re} \langle x - p, z \rangle \leq 0$ for all $t \in \mathbb{R}$, forcing $\text{Re} \langle x - p, z \rangle = 0$. Replacing z by iz (in the complex case) gives $\text{Im} \langle x - p, z \rangle = 0$ too, so $\langle x - p, z \rangle = 0$ for all $z \in F$, i.e., $x - p \in F^{\perp}$.

(ii) Every $x \in H$ decomposes as $x = P_F x + (x - P_F x)$ with $P_F x \in F$ and $x - P_F x \in F^{\perp}$. If $z \in F \cap F^{\perp}$ then $\langle z, z \rangle = 0$, so $z = 0$. Hence the sum is direct.

(iii) Linearity: let $x_1, x_2 \in H$ and $\alpha, \beta \in \mathbb{K}$. Set $p_i = P_F x_i$. Then $\alpha p_1 + \beta p_2 \in F$ and $(\alpha x_1 + \beta x_2) - (\alpha p_1 + \beta p_2) = \alpha(x_1 - p_1) + \beta(x_2 - p_2) \in F^\perp$, so $P_F(\alpha x_1 + \beta x_2) = \alpha p_1 + \beta p_2$.

Boundedness: $\|P_F x\|^2 = \langle P_F x, P_F x \rangle = \langle P_F x, x \rangle \leq \|P_F x\| \, \|x\|$, giving $\|P_F x\| \leq \|x\|$. For $x \in F$, $P_F x = x$, so $\|P_F\| = 1$.

Idempotence: $P_F^2 = P_F$ since $P_F x \in F$ implies $P_F(P_F x) = P_F x$.

Self-adjointness: $\langle P_F x, y \rangle = \langle P_F x, P_F y + (y - P_F y) \rangle = \langle P_F x, P_F y \rangle$ (since $P_F x \in F$ and $y - P_F y \in F^\perp$). By symmetry, $\langle x, P_F y \rangle = \langle P_F x, P_F y \rangle$. $\qquad\square$

Corollary 2.16. *For any closed subspace F of H: $(F^\perp)^\perp = F$.*

Proof. $F \subset (F^\perp)^\perp$ is clear. If $x \in (F^\perp)^\perp$, decompose $x = p + q$ with $p \in F$, $q \in F^\perp$. Then $q = x - p \in (F^\perp)^\perp$ (since both x and p lie there), so $q \in F^\perp \cap (F^\perp)^\perp$, giving $q = 0$ and $x = p \in F$. $\qquad\square$

Corollary 2.17. *A subspace F is dense in H if and only if $F^\perp = \{0\}$.*

Proof. $\overline{F}$ is a closed subspace and $\overline{F}^\perp = F^\perp$. If $F^\perp = \{0\}$, then $\overline{F} = (\overline{F}^\perp)^\perp = \{0\}^\perp = H$. Conversely, if $\overline{F} = H$ and $z \in F^\perp$, then $\langle z, y \rangle = 0$ for all $y \in F$; by continuity, $\langle z, y \rangle = 0$ for all $y \in H$, forcing $z = 0$. $\qquad\square$

2.4   The Riesz representation theorem

Theorem 2.18 (Riesz–Fréchet representation theorem). *Let H be a Hilbert space. For every continuous linear functional $\varphi \in H^*$ there exists a unique $u \in H$ such that*

$$\varphi(x) = \langle x, u \rangle \qquad \forall\, x \in H.$$

Moreover, $\|\varphi\|_{H^*} = \|u\|_H$. The map $J : H \to H^*$ defined by $Ju = \langle \cdot, u \rangle$ is a conjugate-linear isometric isomorphism (an isometric isomorphism in the real case).

Proof. **Existence.** If $\varphi = 0$, take $u = 0$. Assume $\varphi \neq 0$. Set $F = \operatorname{Ker} \varphi$; this is a closed hyperplane (i.e., $\operatorname{codim} F = 1$). Since $F \neq H$, we have $F^\perp \neq \{0\}$; pick $v \in F^\perp \setminus \{0\}$.

For any $x \in H$, the element $x - \frac{\varphi(x)}{\varphi(v)} v$ lies in F (check: φ applied to it gives 0). Hence

$$0 = \left\langle x - \frac{\varphi(x)}{\varphi(v)} v, v \right\rangle = \langle x, v \rangle - \frac{\varphi(x)}{\varphi(v)} \|v\|^2.$$

Solving: $\varphi(x) = \left\langle x, \frac{\overline{\varphi(v)}}{\|v\|^2} v \right\rangle$. Set $u = \frac{\overline{\varphi(v)}}{\|v\|^2} v$.

Uniqueness. If $\langle x, u \rangle = \langle x, u' \rangle$ for all x, then $\langle x, u - u' \rangle = 0$ for all x; taking $x = u - u'$ gives $u = u'$.

Isometry. On one hand, $|\varphi(x)| = |\langle x, u \rangle| \leq \|x\| \, \|u\|$, so $\|\varphi\| \leq \|u\|$. On the other hand, $\varphi(u) = \|u\|^2$, so $\|\varphi\| \geq |\varphi(u/\|u\|)| = \|u\|$.

Conjugate-linearity. $J(\alpha u_1 + \beta u_2)(x) = \langle x, \alpha u_1 + \beta u_2 \rangle = \bar{\alpha} \langle x, u_1 \rangle + \bar{\beta} \langle x, u_2 \rangle = \bar{\alpha} \, J u_1(x) + \bar{\beta} \, J u_2(x)$.

Surjectivity was established in the existence part. $\square$

Remark 2.19. The Riesz theorem says that a Hilbert space is (conjugate-linearly) isometrically isomorphic to its own dual: $H \cong H^*$. This **self-duality** is a key distinction from general Banach spaces.

Corollary 2.20 (Lax–Milgram theorem — a preview). *Let $a : H \times H \to \mathbb{K}$ be a continuous, coercive sesquilinear form on a Hilbert space H. Then for every $\varphi \in H^*$ there exists a unique $u \in H$ with $a(u, v) = \varphi(v)$ for all $v \in H$.*

This will be proved in full generality later; it is the foundation of the variational approach to elliptic PDEs.

2.5   Orthonormal systems and Hilbert bases

Definition 2.21 (Orthonormal system). A family $(e_i)_{i \in I}$ in a Hilbert space H is called an **orthonormal system** (ONS) if

$$\langle e_i, e_j \rangle = \delta_{ij} \qquad \forall \, i, j \in I,$$

where δ_{ij} is the Kronecker delta.

Definition 2.22 (Hilbert basis). An orthonormal system $(e_i)_{i \in I}$ is a **Hilbert basis** (or **complete orthonormal system**) if it is *maximal,*

i.e., no proper orthonormal system contains it. Equivalently, $\mathrm{span}\{e_i : i \in I\}$ is dense in H.

Definition 2.23 (Fourier coefficients)**.** Given an orthonormal system $(e_i)_{i \in I}$ and $x \in H$, the scalars $\hat{x}_i = \langle x, e_i \rangle$ are called the **Fourier coefficients** of x.

Lemma 2.24 (Best approximation)**.** *Let $e_1, \ldots, e_n$ be orthonormal and $x \in H$. Then*

$$\min_{\alpha_1,\ldots,\alpha_n \in \mathbb{K}} \left\| x - \sum_{k=1}^n \alpha_k e_k \right\|^2 = \|x\|^2 - \sum_{k=1}^n |\langle x, e_k \rangle|^2 ,$$

and the minimum is attained at $\alpha_k = \langle x, e_k \rangle$.

Proof. Set $s = \sum_{k=1}^n \alpha_k e_k$ and $\hat{x}_k = \langle x, e_k \rangle$. Then

$$\|x - s\|^2 = \|x\|^2 - \sum_k \bar{\alpha}_k \langle x, e_k \rangle - \sum_k \alpha_k \overline{\langle x, e_k \rangle} + \sum_k |\alpha_k|^2$$
$$= \|x\|^2 - \sum_k |\hat{x}_k|^2 + \sum_k |\alpha_k - \hat{x}_k|^2 .$$

This is minimised when $\alpha_k = \hat{x}_k$, and the minimum value is $\|x\|^2 - \sum_k |\hat{x}_k|^2$. $\quad\square$

Theorem 2.25 (Bessel's inequality)**.** *Let $(e_i)_{i \in I}$ be an orthonormal system in H. For every $x \in H$:*

$$\sum_{i \in I} |\langle x, e_i \rangle|^2 \le \|x\|^2 .$$

In particular, at most countably many Fourier coefficients are nonzero.

Proof. For any finite subset $J \subset I$, definition 2.24 gives $\sum_{i \in J} |\langle x, e_i \rangle|^2 \le \|x\|^2$. Taking the supremum over all finite J gives the result. The set $\{i \in I : \langle x, e_i \rangle \ne 0\} = \bigcup_{n \ge 1} \{i : |\langle x, e_i \rangle| > 1/n\}$; each set in the union is finite (by Bessel), hence the union is countable. $\quad\square$

> **Theorem 2.26** (Parseval's theorem). *Let $(e_i)_{i \in I}$ be an orthonormal system in a Hilbert space H. The following are equivalent:*
>
>   (i) *$(e_i)_{i \in I}$ is a Hilbert basis;*
>
>   (ii) *$\overline{\operatorname{span}}\{e_i : i \in I\} = H$;*
>
>   (iii) *For every $x \in H$: $x = \sum_{i \in I} \langle x, e_i \rangle \, e_i$ (convergence in norm);*
>
>   (iv) *(Parseval's identity) $\|x\|^2 = \sum_{i \in I} |\langle x, e_i \rangle|^2$ for every $x \in H$;*
>
>   (v) *(Generalised Parseval) $\langle x, y \rangle = \sum_{i \in I} \langle x, e_i \rangle \overline{\langle y, e_i \rangle}$ for all $x, y \in H$;*
>
>   (vi) *$\langle x, e_i \rangle = 0$ for all $i \in I$ implies $x = 0$.*

Proof. We prove the cycle (i)$\Rightarrow$(vi)$\Rightarrow$(ii)$\Rightarrow$(iii)$\Rightarrow$(iv)$\Rightarrow$(v)$\Rightarrow$(ii), plus (ii)$\Leftrightarrow$(i).

   (i)$\Rightarrow$(vi): If $\langle x, e_i \rangle = 0$ for all i and $x \neq 0$, then $e_i \cup \{x/\|x\|\}$ would be a strictly larger ONS, contradicting maximality.

   (vi)$\Rightarrow$(ii): If $F = \overline{\operatorname{span}}\{e_i\} \neq H$, pick $x \in F^\perp \setminus \{0\}$. Then $\langle x, e_i \rangle = 0$ for all i, contradicting (vi).

   (ii)$\Rightarrow$(iii): Let $x \in H$ and $\varepsilon > 0$. By (ii), there exist finitely many indices $i_1, \ldots, i_n$ and scalars α_k with $\|x - \sum_k \alpha_k e_{i_k}\| < \varepsilon$. By the best approximation lemma, replacing α_k by $\langle x, e_{i_k} \rangle$ can only decrease the error. Since Bessel's inequality guarantees $\sum_{i \in I} |\langle x, e_i \rangle|^2 < \infty$, we can arrange the nonzero Fourier coefficients in a sequence and the partial sums converge to x.

   (iii)$\Rightarrow$(iv): $\|x\|^2 = \langle x, x \rangle = \left\langle \sum_i \langle x, e_i \rangle e_i, \sum_j \langle x, e_j \rangle e_j \right\rangle = \sum_i |\langle x, e_i \rangle|^2$ (the last step uses continuity of the inner product).

   (iv)$\Rightarrow$(v): Apply the polarization identity to $\|x + y\|^2$, $\|x - y\|^2$, etc.

   (v)$\Rightarrow$(ii): If $x \perp \overline{\operatorname{span}}\{e_i\}$, then $\langle x, e_i \rangle = 0$ for all i, so (v) gives $\langle x, x \rangle = 0$, i.e., $x = 0$. By definition 2.17, $\overline{\operatorname{span}}\{e_i\} = H$.

   (ii)$\Leftrightarrow$(i): $\overline{\operatorname{span}}\{e_i\} = H$ iff $\{e_i\}^\perp = \{0\}$ iff (e_i) is maximal (any vector orthogonal to all e_i must be zero, so no ONS properly contains (e_i)). $\square$

2.6   Separable spaces and countable bases

> **Theorem 2.27** (Existence of Hilbert bases). *Every Hilbert space admits a Hilbert basis.*

Proof. Apply Zorn's lemma to the partially ordered set (by inclusion) of orthonormal systems. Every chain has an upper bound (its union), so there exists a maximal element. □

Theorem 2.28. *A Hilbert space H is separable if and only if it admits a countable (or finite) Hilbert basis.*

Proof. ($\Leftarrow$) If $(e_n)_{n\geq 1}$ is a countable Hilbert basis, the set $D = \{\sum_{k=1}^{N} q_k e_k : N \in \mathbb{N}, q_k \in \mathbb{Q} + i\mathbb{Q}\}$ is countable and dense.

($\Rightarrow$) If H is separable, every ONS is countable (distinct elements are distance $\sqrt{2}$ apart, so each lives in a distinct member of a countable open cover). By definition 2.27, a Hilbert basis exists, and it must be countable. □

Theorem 2.29 (Isomorphism theorem)**.** *Every separable infinite-dimensional Hilbert space is isometrically isomorphic to ℓ^2.*

Proof. Let $(e_n)_{n\geq 1}$ be a Hilbert basis. Define $U : H \to \ell^2$ by $Ux = (\langle x, e_n \rangle)_{n\geq 1}$. By Parseval, $\|Ux\|_{\ell^2} = \|x\|_H$, so U is an isometry (hence injective). For surjectivity, given $(\alpha_n) \in \ell^2$, the series $\sum_n \alpha_n e_n$ converges in H (partial sums are Cauchy by Bessel/Parseval) to some x with $Ux = (\alpha_n)$. □

2.7 The Fischer–Riesz theorem

Theorem 2.30 (Fischer–Riesz)**.** *For any measure space $(\Omega, \mathcal{A}, \mu)$, the space $L^2(\Omega, \mu)$ is a Hilbert space.*

Proof. We need to show completeness. Let (f_n) be a Cauchy sequence in L^2. Choose a subsequence (f_{n_k}) with $\left\|f_{n_{k+1}} - f_{n_k}\right\|_2 < 2^{-k}$. Define

$$g_N = \sum_{k=1}^{N} \left|f_{n_{k+1}} - f_{n_k}\right|, \qquad g = \sum_{k=1}^{\infty} \left|f_{n_{k+1}} - f_{n_k}\right|.$$

By Minkowski's inequality in L^2:

$$\|g_N\|_2 \leq \sum_{k=1}^{N} 2^{-k} < 1.$$

By monotone convergence, $\|g\|_2 \leq 1$, so $g < \infty$ a.e. This means the telescoping series

$$f_{n_1} + \sum_{k=1}^{\infty} (f_{n_{k+1}} - f_{n_k})$$

converges absolutely a.e. to a function f. We have $f_{n_k} \to f$ a.e. Moreover $|f_{n_k} - f_{n_1}|^2 \leq g^2 \in L^1$, so by dominated convergence, $\|f - f_{n_k}\|_2 \to 0$. Since (f_n) is Cauchy and a subsequence converges to f, the entire sequence converges to f in L^2. □

Remark 2.31. The same argument (with p replacing 2) proves that L^p is complete for all $1 \leq p < \infty$. The case $p = \infty$ requires a different (simpler) argument.

2.8   Hilbert direct sum

Definition 2.32 (Hilbert direct sum)**.** Let $(H_i)_{i \in I}$ be a family of Hilbert spaces. The **Hilbert direct sum** is

$$\bigoplus_{i \in I}^{2} H_i = \left\{ (x_i)_{i \in I} : x_i \in H_i, \ \sum_{i \in I} \|x_i\|_{H_i}^2 < \infty \right\}$$

with inner product $\langle (x_i), (y_i) \rangle = \sum_{i \in I} \langle x_i, y_i \rangle_{H_i}$.

Proposition 2.33. *The Hilbert direct sum is a Hilbert space.*

Proof. The inner product is well-defined by the Cauchy–Schwarz inequality applied to $\ell^2(I)$. Completeness follows from the completeness of each H_i and of $\ell^2(I)$: if $(x^{(n)})$ is Cauchy in $\bigoplus^2 H_i$, then for each i, $(x_i^{(n)})$ is Cauchy in H_i, hence converges to some x_i. A standard $\varepsilon/3$ argument (as in the proof of completeness of ℓ^p) shows $x = (x_i) \in \bigoplus^2 H_i$ and $x^{(n)} \to x$. □

Example 2.34. $\ell^2(I) = \bigoplus_{i \in I}^2 \mathbb{K}$ is the Hilbert direct sum of copies of the scalar field.

2.9 Operators on Hilbert spaces: the adjoint

Definition 2.35 (Adjoint operator)**.** Let H, K be Hilbert spaces and $T \in \mathcal{L}(H, K)$. The **adjoint** of T is the unique operator $T^* \in \mathcal{L}(K, H)$ satisfying

$$\langle Tx, y \rangle_K = \langle x, T^*y \rangle_H \qquad \forall\, x \in H,\ y \in K.$$

Proof of existence and uniqueness. For fixed $y \in K$, the map $x \mapsto \langle Tx, y \rangle_K$ is a continuous linear functional on H (with norm $\leq \|T\|\,\|y\|$). By the Riesz representation theorem (definition 2.18), there exists a unique $z \in H$ with $\langle Tx, y \rangle = \langle x, z \rangle$ for all x. Define $T^*y = z$. Linearity and boundedness of T^* follow from uniqueness and the estimate $\|T^*y\| \leq \|T\|\,\|y\|$. $\square$

Proposition 2.36 (Properties of the adjoint)**.** *Let $S, T \in \mathcal{L}(H, K)$ and $\alpha \in \mathbb{K}$.*

(i) $(S + T)^* = S^* + T^*$ *and* $(\alpha T)^* = \bar{\alpha}\, T^*$.

(ii) $(T^*)^* = T$.

(iii) $\|T^*\| = \|T\|$.

(iv) $\|T^*T\| = \|T\|^2$.

(v) *If $R \in \mathcal{L}(K, L)$, then $(RT)^* = T^*R^*$.*

(vi) $\operatorname{Ker} T^* = (\operatorname{ran} T)^\perp$ *and* $\overline{\operatorname{ran} T} = (\operatorname{Ker} T^*)^\perp$.

Proof. We prove the most important parts.

(ii) For all $x \in H$, $y \in K$: $\langle x, T^{**}y \rangle = \langle T^*x, y \rangle = \overline{\langle y, T^*x \rangle} = \overline{\langle Ty, x \rangle} = \langle x, Ty \rangle$, so $T^{**} = T$ (here we identify H with H^{**} via the inner product).

Wait — let us be more careful. We have $T^* \in \mathcal{L}(K, H)$, so $T^{**} = (T^*)^* \in \mathcal{L}(H, K)$. For $x \in H$, $y \in K$: $\langle T^{**}x, y \rangle_K = \langle x, T^*y \rangle_H = \langle Tx, y \rangle_K$. Since this holds for all y, $T^{**}x = Tx$.

(iii) $\|T^*\| \leq \|T\|$ since $|\langle x, T^*y \rangle| = |\langle Tx, y \rangle| \leq \|T\|\,\|x\|\,\|y\|$. By (ii), $\|T\| = \|T^{**}\| \leq \|T^*\|$.

(iv) $\|T^*T\| \leq \|T^*\|\,\|T\| = \|T\|^2$. Conversely, $\|Tx\|^2 = \langle Tx, Tx \rangle = \langle T^*Tx, x \rangle \leq \|T^*Tx\|\,\|x\| \leq \|T^*T\|\,\|x\|^2$, so $\|T\|^2 \leq \|T^*T\|$.

(vi) $y \in \operatorname{Ker} T^*$ iff $T^*y = 0$ iff $\langle x, T^*y \rangle = 0$ for all x iff $\langle Tx, y \rangle = 0$ for all x iff $y \in (\operatorname{ran} T)^\perp$. By definition 2.16, $\overline{\operatorname{ran} T} = ((\operatorname{ran} T)^\perp)^\perp = (\operatorname{Ker} T^*)^\perp$. $\square$

Definition 2.37 (Self-adjoint, normal, unitary)**.** Let $T \in \mathcal{L}(H)$.

- T is **self-adjoint** (or **Hermitian**) if $T^* = T$.

- T is **normal** if $T^*T = TT^*$.

- T is **unitary** if $T^*T = TT^* = \mathrm{Id}$ (equivalently, T is a surjective isometry).

Proposition 2.38. *If $T = T^*$, then $\langle Tx, x \rangle \in \mathbb{R}$ for all $x \in H$. Conversely, if $\mathbb{K} = \mathbb{C}$ and $\langle Tx, x \rangle \in \mathbb{R}$ for all x, then $T = T^*$.*

Proof. If $T = T^*$: $\overline{\langle Tx, x \rangle} = \langle x, Tx \rangle = \langle T^*x, x \rangle = \langle Tx, x \rangle$, so $\langle Tx, x \rangle \in \mathbb{R}$.

Conversely, if $\langle Tx, x \rangle \in \mathbb{R}$ for all x, then $\langle Tx, x \rangle = \overline{\langle Tx, x \rangle} = \langle x, Tx \rangle = \langle T^*x, x \rangle$. Hence $\langle (T - T^*)x, x \rangle = 0$ for all x. Over $\mathbb{C}$, the polarization identity implies $T - T^* = 0$. □

Remark 2.39. Over $\mathbb{R}$, the condition $\langle Tx, x \rangle = 0$ for all x does *not* imply $T = 0$ (consider a 90° rotation in $\mathbb{R}^2$). The implication relies on the full complex polarization identity.

2.10　Exercises for Chapter 2

Exercise 2.1 ($\star$)**.** Show that the inner product $\langle \cdot, \cdot \rangle : H \times H \to \mathbb{K}$ is (jointly) continuous. Specifically, if $x_n \to x$ and $y_n \to y$, then $\langle x_n, y_n \rangle \to \langle x, y \rangle$.

Exercise 2.2 ($\star$)**.** Prove the **Pythagorean theorem**: if $x \perp y$ then $\|x + y\|^2 = \|x\|^2 + \|y\|^2$. Generalise to finite sums.

Exercise 2.3 ($\star$)**.** Show that if $S \subset H$ is any subset, then $S^\perp$ is a closed subspace of H.

Exercise 2.4 ($\star\star$)**.** Let $F = \mathrm{span}\{e_1, \ldots, e_n\}$ where $e_1, \ldots, e_n$ are orthonormal. Show that the orthogonal projection onto F is given by

$P_F x = \sum_{k=1}^{n} \langle x, e_k \rangle \, e_k.$

Exercise 2.5 ($\star\star$). Let $(v_n)_{n \geq 1}$ be a linearly independent sequence in a Hilbert space H. Describe the **Gram–Schmidt orthonormalization** process and prove it produces an orthonormal system (e_n) with $\mathrm{span}\{e_1, \ldots, e_n\} = \mathrm{span}\{v_1, \ldots, v_n\}$ for each n.

Exercise 2.6 ($\star\star$). Show that the system $\left(\frac{1}{\sqrt{2\pi}} e^{int}\right)_{n \in \mathbb{Z}}$ is an orthonormal system in $L^2([-\pi, \pi])$. Proving completeness requires the Stone–Weierstrass theorem or Fejér's theorem; state this as a fact and deduce Parseval's identity for Fourier series:

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} |f(t)|^2 \, \mathrm{d}t = \sum_{n \in \mathbb{Z}} \left| \hat{f}(n) \right|^2.$$

Exercise 2.7 ($\star\star$). Let $H = L^2([0, 1])$ and define $\varphi(f) = \int_0^1 t \, f(t) \, \mathrm{d}t$. Show that $\varphi \in H^*$ and find the element $u \in H$ given by the Riesz representation theorem such that $\varphi(f) = \langle f, u \rangle$. Compute $\|\varphi\|$.

Exercise 2.8 ($\star\star$). Let $T : \mathbb{K}^n \to \mathbb{K}^m$ be the linear map represented by the matrix A. Show that T^* is represented by $A^* = \bar{A}^T$ (the conjugate transpose).

Exercise 2.9 ($\star\star$). Show that $T \in \mathcal{L}(H)$ is normal if and only if $\|Tx\| = \|T^*x\|$ for all $x \in H$.

Exercise 2.10 ($\star\star$). Let $U \in \mathcal{L}(H)$. Prove the following are equivalent: (a) U is unitary; (b) U is surjective and $\langle Ux, Uy \rangle = \langle x, y \rangle$ for all x, y; (c) U is surjective and $\|Ux\| = \|x\|$ for all x.

Exercise 2.11 ($\star\star\star$). A sequence (x_n) in a Hilbert space H **converges weakly** to x, written $x_n \rightharpoonup x$, if $\langle x_n, y \rangle \to \langle x, y \rangle$ for all $y \in H$.

(a) Show that strong convergence implies weak convergence but not

conversely ($e_n \rightharpoonup 0$ in ℓ^2 but $\|e_n\| = 1$).

(b) Prove that a weakly convergent sequence is bounded. *Hint:* use the Riesz theorem to identify H with H^*, then apply the uniform boundedness principle (accept it for now).

(c) Show that $x_n \rightharpoonup x$ and $\|x_n\| \to \|x\|$ together imply $x_n \to x$ (strong convergence).

Exercise 2.12 ($\star \star \star$). Let $(H_i)_{i \in I}$ be a family of closed, mutually orthogonal subspaces of a Hilbert space H with $H = \overline{\bigoplus_{i \in I} H_i}$. Show that H is isometrically isomorphic to $\bigoplus_{i \in I}^2 H_i$.

Exercise 2.13 ($\star\star\star$). Let H be a separable Hilbert space with Hilbert basis (e_n). Define $T : H \to H$ by $Te_n = \frac{1}{n}e_n$.

(a) Show that T is bounded, self-adjoint, and $\|T\| = 1$.

(b) Show that T maps bounded sets to sets with compact closure. (Such an operator is called **compact**.)

(c) Show that T is *not* surjective.

(d) Find $\operatorname{Ker} T$ and $\overline{\operatorname{ran} T}$.

Exercise 2.14 ($\star$). Let F, G be closed subspaces of a Hilbert space. Show that $(F + G)^{\perp} = F^{\perp} \cap G^{\perp}$. Is it always true that $(F \cap G)^{\perp} = F^{\perp} + G^{\perp}$?

Exercise 2.15 ($\star\star$). A normed space is **strictly convex** if $\|x\| = \|y\| = 1$ and $x \neq y$ imply $\|x + y\| < 2$. Show that Hilbert spaces are strictly convex, and deduce uniqueness of nearest points in closed convex sets without using the parallelogram law directly.

Exercise 2.16 ($\star \star \star$). Prove the Lax–Milgram theorem (definition 2.20): if $a : H \times H \to \mathbb{K}$ is a sesquilinear form satisfying $|a(u, v)| \leq M \|u\| \|v\|$ (continuity) and $\operatorname{Re} a(u, u) \geq \alpha \|u\|^2$ (coerciv-

ity) for some $\alpha > 0$, then for every $\varphi \in H^*$ there exists a unique $u \in H$ with $a(u,v) = \varphi(v)$ for all $v \in H$, and $\|u\| \leq \frac{1}{\alpha}\|\varphi\|$.

Hint: By Riesz, write $a(u,v) = \langle Au, v \rangle$ for some $A \in \mathcal{L}(H)$. Show A is injective with closed range, then surjective.

Chapter 3

Bounded Linear Operators

Contents

The study of bounded linear operators between Banach spaces lies at the heart of functional analysis. In this chapter we develop the algebraic and topological theory of the space $\mathcal{L}(E, F)$, introduce compact operators, and build the spectral theory of bounded operators. The chapter culminates with the rich additional structure available when the underlying spaces are Hilbert spaces: the adjoint, and the classes of normal, self-adjoint, and unitary operators.

3.1 The space $\mathcal{L}(E, F)$ revisited

Throughout this section, $(E, \|\cdot\|_E)$ and $(F, \|\cdot\|_F)$ denote normed spaces over the field $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$.

Definition 3.1 (Bounded linear operator). A linear map $T\colon E \to F$ is *bounded* if there exists $C \geq 0$ such that

$$\|Tx\|_F \leq C \, \|x\|_E \qquad \text{for all } x \in E.$$

The infimum of all such constants is the *operator norm*

$$\|T\| \;=\; \sup_{\|x\|_E \leq 1} \|Tx\|_F \;=\; \sup_{\|x\|_E = 1} \|Tx\|_F \;=\; \sup_{x \neq 0} \frac{\|Tx\|_F}{\|x\|_E}.$$

We denote by $\mathcal{L}(E, F)$ the vector space of all bounded linear operators from E to F, equipped with the operator norm. When $E = F$ we write $\mathcal{L}(E) = \mathcal{L}(E, E)$.

Proposition 3.2 (Completeness of $\mathcal{L}(E, F)$). *If F is a Banach space, then $\mathcal{L}(E, F)$ is a Banach space.*

Proof. Let (T_n) be a Cauchy sequence in $\mathcal{L}(E, F)$. For each $x \in E$,

$$\|T_n x - T_m x\|_F \leq \|T_n - T_m\| \, \|x\|_E \to 0,$$

so $(T_n x)$ is Cauchy in F, hence convergent. Define $Tx = \lim_{n \to \infty} T_n x$. Linearity of T is clear. Since (T_n) is Cauchy, it is bounded: $\sup_n \|T_n\| = M < \infty$. Then $\|Tx\|_F = \lim_n \|T_n x\|_F \leq M \|x\|_E$, so $T \in \mathcal{L}(E, F)$. Finally, given $\varepsilon > 0$, choose N so that $\|T_n - T_m\| < \varepsilon$ for $n, m \geq N$. Letting $m \to \infty$, $\|T_n - T\| \leq \varepsilon$ for $n \geq N$. $\qquad\square$

Example 3.3 (Multiplication operators). Let $\varphi \in L^\infty(\mu)$ for a measure space (Ω, Σ, μ). The operator $M_\varphi\colon L^p(\mu) \to L^p(\mu)$ defined by $M_\varphi f = \varphi f$ satisfies $\|M_\varphi\| = \|\varphi\|_\infty$.

Example 3.4 (Integral operators). Let $K \in L^2([0, 1]^2)$. The operator $T_K\colon L^2([0, 1]) \to L^2([0, 1])$ defined by

$$(T_K f)(x) = \int_0^1 K(x, y) \, f(y) \, \mathrm{d}y$$

satisfies $\|T_K\| \leq \|K\|_{L^2([0,1]^2)}$ by the Cauchy–Schwarz inequality.

Proposition 3.5 (Composition). *Let E, F, G be normed spaces and $T \in \mathcal{L}(E, F)$, $S \in \mathcal{L}(F, G)$. Then $ST \in \mathcal{L}(E, G)$ and $\|ST\| \leq \|S\| \, \|T\|$.*

Proof. For any $x \in E$, $\|STx\|_G \leq \|S\| \, \|Tx\|_F \leq \|S\| \, \|T\| \, \|x\|_E$. $\square$

Corollary 3.6. *If E is a Banach space, then $\mathcal{L}(E)$ is a unital Banach algebra with product given by composition and unit Id_E.*

3.2 Finite-rank and compact operators

Definition 3.7 (Finite-rank operator). An operator $T \in \mathcal{L}(E, F)$ is of *finite rank* if $\dim(\operatorname{ran} T) < \infty$. We write $\mathcal{F}(E, F)$ for the set of finite-rank operators, which is a linear subspace of $\mathcal{L}(E, F)$. The *rank* of T is $\dim(\operatorname{ran} T)$.

Example 3.8. Let $f \in E'$ and $y \in F$. The operator $T = y \otimes f$ defined by $Tx = f(x)\, y$ has rank at most one and $\|T\| = \|f\| \, \|y\|$. Every finite-rank operator is a finite sum of such rank-one operators.

Definition 3.9 (Compact operator). An operator $T \in \mathcal{L}(E, F)$ is *compact* if $T(B_E)$ is relatively compact in F, where $B_E = \{x \in E : \|x\| \leq 1\}$. Equivalently, for every bounded sequence (x_n) in E, the sequence (Tx_n) has a convergent subsequence in F. We write $\mathcal{K}(E, F)$ for the set of compact operators.

Proposition 3.10 (Basic properties of compact operators). *Let E, F, G be Banach spaces.*

(i) *Every finite-rank operator is compact.*

(ii) *$\mathcal{K}(E, F)$ is a closed linear subspace of $\mathcal{L}(E, F)$.*

(iii) *If $T \in \mathcal{K}(E, F)$ and $S \in \mathcal{L}(F, G)$, then $ST \in \mathcal{K}(E, G)$.*

(iv) *If $T \in \mathcal{L}(E, F)$ and $S \in \mathcal{K}(F, G)$, then $ST \in \mathcal{K}(E, G)$.*

Thus $\mathcal{K}(E, F)$ is a closed two-sided ideal in $\mathcal{L}(E)$ when $E = F = G$.

Proof. *(i)* If T has finite rank, then $T(B_E)$ is a bounded subset of a finite-dimensional subspace, hence relatively compact.

(ii) LINEARITY. If $T_1, T_2 \in \mathcal{K}(E, F)$ and $\alpha \in \mathbb{K}$, take any bounded sequence (x_n). Extract a subsequence (x_{n_k}) so that $(T_1 x_{n_k})$ converges, then a further subsequence so that $(T_2 x_{n_{k_l}})$ converges. Then $((T_1 + \alpha T_2) x_{n_{k_l}})$ converges.

CLOSEDNESS. Let $T_n \to T$ in $\mathcal{L}(E, F)$ with each T_n compact. Let (x_k) be a bounded sequence in E with $\|x_k\| \leq M$. By a diagonal argument, we extract a subsequence (x_{k_j}) such that $(T_n x_{k_j})_j$ converges for every n. Given $\varepsilon > 0$, choose n_0 with $\|T - T_{n_0}\| < \varepsilon/(3M)$. Choose J so that $\left\| T_{n_0} x_{k_i} - T_{n_0} x_{k_j} \right\| < \varepsilon/3$ for $i, j \geq J$. Then

$$\left\| T x_{k_i} - T x_{k_j} \right\| \leq 2 \left\| T - T_{n_0} \right\| M + \left\| T_{n_0} x_{k_i} - T_{n_0} x_{k_j} \right\| < \varepsilon.$$

So $(T x_{k_j})$ is Cauchy, hence convergent.

(iii) If (x_n) is bounded, extract a subsequence with $(T x_{n_k})$ convergent. Then $(S T x_{n_k})$ converges by continuity of S.

(iv) If (x_n) is bounded, then $(T x_n)$ is bounded in F, so compactness of S yields a convergent subsequence of $(S T x_n)$. $\square$

> **Example 3.11** (Hilbert–Schmidt operators are compact)**.** The integral operator T_K from Example 3.4 is compact: it is the limit (in operator norm) of finite-rank operators obtained by truncating the singular value decomposition (or an orthonormal expansion of K).

> *Remark* 3.12. Not every compact operator has finite rank. The diagonal operator $T \colon \ell^2 \to \ell^2$ defined by $T e_n = \frac{1}{n} e_n$ is compact (being the norm-limit of the finite-rank operators T_N that agree with T on $e_1, \dots, e_N$ and vanish on e_n for $n > N$) but has infinite-dimensional range.

3.3   Spectrum and resolvent

In this section, E is a Banach space over $\mathbb{C}$ and $T \in \mathcal{L}(E)$.

> **Definition 3.13** (Resolvent set and spectrum)**.** The *resolvent set* of T is
> $$\rho(T) = \big\{ \lambda \in \mathbb{C} : \lambda \operatorname{Id} - T \text{ is bijective in } \mathcal{L}(E) \big\}.$$
> The *spectrum* of T is $\sigma(T) = \mathbb{C} \setminus \rho(T)$. For $\lambda \in \rho(T)$, the *resolvent*

operator is $R(\lambda, T) = (\lambda \operatorname{Id} - T)^{-1} \in \mathcal{L}(E)$.

Definition 3.14 (Classification of the spectrum). The spectrum is partitioned as $\sigma(T) = \sigma_p(T) \cup \sigma_c(T) \cup \sigma_r(T)$ where:

  (i) $\sigma_p(T)$: the *point spectrum* (eigenvalues): $\lambda \operatorname{Id} - T$ is not injective.

  (ii) $\sigma_c(T)$: the *continuous spectrum*: $\lambda \operatorname{Id} - T$ is injective, $\operatorname{ran}(\lambda \operatorname{Id} - T)$ is dense but not all of E.

  (iii) $\sigma_r(T)$: the *residual spectrum*: $\lambda \operatorname{Id} - T$ is injective, $\operatorname{ran}(\lambda \operatorname{Id} - T)$ is not dense.

Remark 3.15. By the open mapping theorem (Theorem 6.3 or the relevant theorem from Chapter 2), if $\lambda \operatorname{Id} - T$ is bijective, then $(\lambda \operatorname{Id} - T)^{-1}$ is automatically bounded. Hence the definition of $\rho(T)$ simply requires bijectivity.

3.4   The Neumann series

Theorem 3.16 (Neumann series). *Let E be a Banach space and $T \in \mathcal{L}(E)$ with $\|T\| < 1$. Then $\operatorname{Id} - T$ is invertible in $\mathcal{L}(E)$ and*

$$(\operatorname{Id} - T)^{-1} = \sum_{n=0}^{\infty} T^n,$$

with $\|(\operatorname{Id} - T)^{-1}\| \leq \dfrac{1}{1 - \|T\|}$.

Proof. Since $\|T^n\| \leq \|T\|^n$ and $\|T\| < 1$, the series $S_N = \sum_{n=0}^{N} T^n$ is absolutely convergent in the Banach space $\mathcal{L}(E)$:

$$\sum_{n=0}^{\infty} \|T^n\| \leq \sum_{n=0}^{\infty} \|T\|^n = \frac{1}{1 - \|T\|} < \infty.$$

Let $S = \sum_{n=0}^{\infty} T^n$. Then

$$(\operatorname{Id} - T)S_N = \sum_{n=0}^{N} T^n - \sum_{n=1}^{N+1} T^n = \operatorname{Id} - T^{N+1},$$

and $\left\|T^{N+1}\right\| \leq \|T\|^{N+1} \to 0$. So $(\mathrm{Id} - T)S = \mathrm{Id}$. Similarly $S(\mathrm{Id} - T) = \mathrm{Id}$, whence $(\mathrm{Id} - T)^{-1} = S$.   □

Example 3.17 (Application: Fredholm integral equations of the second kind). Consider the equation $f - \lambda K f = g$ in $L^2([0, 1])$, where K is the integral operator with kernel $k \in L^2([0, 1]^2)$ and $\lambda \in \mathbb{C}$. If $|\lambda| < \|K\|^{-1}$, the Neumann series gives the unique solution

$$f = \sum_{n=0}^{\infty} \lambda^n K^n g = g + \lambda K g + \lambda^2 K^2 g + \cdots.$$

The partial sums $f_N = \sum_{n=0}^{N} \lambda^n K^n g$ converge to f in $L^2([0, 1])$ with rate $\|f - f_N\| \leq \frac{(|\lambda|\|K\|)^{N+1}}{1 - |\lambda|\|K\|} \|g\|$.

Corollary 3.18 (Invertible operators form an open set). *The set* $\mathrm{GL}(E)$ *of invertible elements of* $\mathcal{L}(E)$ *is open in the operator-norm topology, and the map* $T \mapsto T^{-1}$ *is continuous (in fact, analytic).*

Proof. Let $S \in \mathrm{GL}(E)$ and let $T \in \mathcal{L}(E)$ with $\|T - S\| < \|S^{-1}\|^{-1}$. Write

$$T = S\big(\mathrm{Id} - S^{-1}(S - T)\big).$$

Since $\|S^{-1}(S - T)\| \leq \|S^{-1}\| \, \|S - T\| < 1$, the Neumann series shows that $\mathrm{Id} - S^{-1}(S - T)$ is invertible, hence T is invertible. Moreover,

$$\left\|T^{-1} - S^{-1}\right\| = \left\|T^{-1}(S - T)S^{-1}\right\| \leq \left\|T^{-1}\right\| \, \|S - T\| \, \left\|S^{-1}\right\|,$$

which tends to zero as $T \to S$ (since $\|T^{-1}\|$ remains bounded).   □

Corollary 3.19 (Spectrum is compact). *For every* $T \in \mathcal{L}(E)$, *the spectrum* $\sigma(T)$ *is a non-empty compact subset of* $\{\lambda \in \mathbb{C} : |\lambda| \leq \|T\|\}$.

Proof. BOUNDEDNESS. If $|\lambda| > \|T\|$, write $\lambda \, \mathrm{Id} - T = \lambda(\mathrm{Id} - \lambda^{-1}T)$. Since $\|\lambda^{-1}T\| < 1$, the Neumann series shows $\lambda \in \rho(T)$.

CLOSEDNESS. The resolvent set is the preimage of $\mathrm{GL}(E)$ under the continuous map $\lambda \mapsto \lambda \, \mathrm{Id} - T$. Since $\mathrm{GL}(E)$ is open (Corollary 3.18), $\rho(T)$ is open, so $\sigma(T)$ is closed. Being closed and bounded, it is compact.

NON-EMPTINESS. Suppose $\sigma(T) = \emptyset$. Then $R(\lambda, T)$ is an entire $\mathcal{L}(E)$-valued function. For any $f \in E'$ and $x \in E$, the function $\lambda \mapsto f(R(\lambda, T)x)$ is entire. For $|\lambda| > 2\|T\|$,

$$\|R(\lambda, T)\| \leq \frac{1}{|\lambda| - \|T\|} \leq \frac{1}{|\lambda|/2} \to 0 \quad \text{as } |\lambda| \to \infty.$$

By Liouville's theorem, $f(R(\lambda, T)x) = 0$ for all λ. Since f and x were arbitrary, $R(\lambda, T) = 0$, a contradiction since $R(\lambda, T)(\lambda \operatorname{Id} - T) = \operatorname{Id}$. $\qquad\square$

3.5 Spectral radius

Definition 3.20 (Spectral radius). The *spectral radius* of $T \in \mathcal{L}(E)$ is
$$r(T) = \sup\{|\lambda| : \lambda \in \sigma(T)\}.$$

Theorem 3.21 (Spectral radius formula / Gelfand's formula). *For every $T \in \mathcal{L}(E)$,*
$$r(T) = \lim_{n \to \infty} \|T^n\|^{1/n} = \inf_{n \geq 1} \|T^n\|^{1/n}.$$

Proof. Write $\alpha = \inf_{n \geq 1} \|T^n\|^{1/n}$ and $\beta = \limsup_{n \to \infty} \|T^n\|^{1/n}$.

STEP 1: $r(T) \leq \alpha$. We first show $\alpha = \inf$ equals $\liminf$, using the submultiplicativity $\|T^{m+n}\| \leq \|T^m\| \|T^n\|$. The sequence $a_n = \log \|T^n\|$ is subadditive, so by Fekete's lemma, $\lim_n a_n/n = \inf_n a_n/n$. Exponentiating, $\lim_n \|T^n\|^{1/n} = \alpha$. Hence $\beta = \alpha$ and the limit exists.

STEP 2: $r(T) \leq \alpha$. If $\lambda \in \sigma(T)$, then $\lambda^n \in \sigma(T^n)$ (since $\lambda^n \operatorname{Id} - T^n = (\lambda \operatorname{Id} - T)\left(\sum_{k=0}^{n-1} \lambda^{n-1-k} T^k\right)$; if $\lambda^n \operatorname{Id} - T^n$ were invertible, so would $\lambda \operatorname{Id} - T$ be). Thus $|\lambda|^n \leq \|T^n\|$, so $|\lambda| \leq \|T^n\|^{1/n}$ for all n. Taking the infimum over n gives $|\lambda| \leq \alpha$, hence $r(T) \leq \alpha$.

STEP 3: $\alpha \leq r(T)$. For $|\lambda| > r(T)$, $\lambda \in \rho(T)$ and the resolvent has the Laurent expansion
$$R(\lambda, T) = \sum_{n=0}^{\infty} \frac{T^n}{\lambda^{n+1}},$$

convergent in $\mathcal{L}(E)$ for $|\lambda| > \|T\|$. By analytic continuation, this series converges for all $|\lambda| > r(T)$.

For any $f \in E'$ and $x \in E$, the scalar power series $\sum_{n=0}^{\infty} f(T^n x)/\lambda^{n+1}$ converges for $|\lambda| > r(T)$. By the root test, $\limsup_n |f(T^n x)|^{1/n} \leq r(T)$ for all f, x.

By the uniform boundedness principle, for each x the sequence $(\|T^n x\|^{1/n})$ satisfies $\limsup_n \|T^n x\|^{1/n} \leq r(T)$. Applying the uniform boundedness principle again (to the operators T^n/r^n for any $r > r(T)$) shows $\sup_n \|T^n\|/r^n < \infty$, hence $\limsup_n \|T^n\|^{1/n} \leq r$. Since $r > r(T)$ was arbitrary, $\alpha = \beta \leq r(T)$. $\square$

> **Example 3.22.** Let $T \colon \mathbb{C}^n \to \mathbb{C}^n$ be nilpotent ($T^n = 0$). Then $\|T^n\|^{1/n} = 0$ for all large n, so $r(T) = 0$ and $\sigma(T) = \{0\}$.

> **Example 3.23** (Shift operator). The right shift $R \colon \ell^2 \to \ell^2$, $(x_1, x_2, \dots) \mapsto (0, x_1, x_2, \dots)$ satisfies $\|R^n\| = 1$ for all n, so $r(R) = 1$. One verifies $\sigma(R) = \overline{\mathbb{D}}$ (the closed unit disc).

3.6   Perturbation of invertible operators

> **Theorem 3.24** (Perturbation bound). *Let $S \in \mathrm{GL}(E)$ and $T \in \mathcal{L}(E)$ with $\|T - S\| < \|S^{-1}\|^{-1}$. Then $T \in \mathrm{GL}(E)$ and*
>
> $$\left\|T^{-1}\right\| \leq \frac{\|S^{-1}\|}{1 - \|S^{-1}\|\,\|T - S\|}, \qquad \left\|T^{-1} - S^{-1}\right\| \leq \frac{\|S^{-1}\|^2\,\|T - S\|}{1 - \|S^{-1}\|\,\|T - S\|}.$$

Proof. Write $T = S(\mathrm{Id} - S^{-1}(S - T))$ and set $U = S^{-1}(S - T)$. Then $\|U\| \leq \|S^{-1}\|\,\|S - T\| < 1$, so $\mathrm{Id} - U$ is invertible by the Neumann series with $\|(\mathrm{Id} - U)^{-1}\| \leq (1 - \|U\|)^{-1}$. Hence $T^{-1} = (\mathrm{Id} - U)^{-1} S^{-1}$ and

$$\left\|T^{-1}\right\| \leq \frac{\|S^{-1}\|}{1 - \|S^{-1}\|\,\|S - T\|}.$$

For the second inequality,

$$T^{-1} - S^{-1} = T^{-1}(S - T)S^{-1},$$

so $\|T^{-1} - S^{-1}\| \leq \|T^{-1}\|\,\|S - T\|\,\|S^{-1}\|$, and we substitute the bound on $\|T^{-1}\|$. $\square$

Example 3.25 (Perturbation of the identity). If $T \in \mathcal{L}(E)$ with $\|T\| < 1$, then $\mathrm{Id} - T$ is invertible and

$$\left\|(\mathrm{Id} - T)^{-1} - \mathrm{Id}\right\| = \left\|(\mathrm{Id} - T)^{-1} - (\mathrm{Id} - T)^{-1}(\mathrm{Id} - T)\right\| = \left\|(\mathrm{Id} - T)^{-1}T\right\| \leq \frac{\|T\|}{1 - \|T\|}.$$

This shows that invertibility is "stable": operators close to the identity remain invertible, with a quantitative bound on how far the inverse is from the identity.

Remark 3.26 (Analytic dependence of the resolvent). The resolvent map $\lambda \mapsto R(\lambda, T)$ is analytic on $\rho(T)$ as a function valued in $\mathcal{L}(E)$. Indeed, for $\lambda_0 \in \rho(T)$, writing $\lambda \, \mathrm{Id} - T = (\lambda_0 \, \mathrm{Id} - T)\big(\mathrm{Id} + (\lambda - \lambda_0)R(\lambda_0, T)\big)$, the Neumann series gives a power series expansion in $(\lambda - \lambda_0)$ convergent for $|\lambda - \lambda_0| < \|R(\lambda_0, T)\|^{-1}$.

Corollary 3.27 (Resolvent identity). *For $\lambda, \mu \in \rho(T)$,*

$$R(\lambda, T) - R(\mu, T) = (\mu - \lambda) \, R(\lambda, T) \, R(\mu, T).$$

In particular, $R(\lambda, T)$ and $R(\mu, T)$ commute.

Proof. Multiply both sides on the left by $(\lambda \, \mathrm{Id} - T)$ and on the right by $(\mu \, \mathrm{Id} - T)$:

$$(\lambda \, \mathrm{Id} - T)\big[R(\lambda, T) - R(\mu, T)\big](\mu \, \mathrm{Id} - T) = (\mu \, \mathrm{Id} - T) - (\lambda \, \mathrm{Id} - T)$$
$$= (\mu - \lambda) \, \mathrm{Id} \,. \qquad \square$$

3.7 Operators on Hilbert spaces: the adjoint

Throughout this section, H and K denote Hilbert spaces over $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$.

Theorem 3.28 (Existence of the adjoint). *For every $T \in \mathcal{L}(H, K)$ there exists a unique operator $T^* \in \mathcal{L}(K, H)$ such that*

$$\langle Tx, y \rangle_K = \langle x, T^*y \rangle_H \qquad \text{for all } x \in H, \, y \in K.$$

Moreover, $\|T^*\| = \|T\|$.

Proof. Fix $y \in K$. The map $\varphi_y \colon H \to \mathbb{K}$ defined by $\varphi_y(x) = \langle Tx, y \rangle_K$ is a bounded linear functional with $\|\varphi_y\| \leq \|T\| \, \|y\|$. By the Riesz representation theorem, there is a unique $z_y \in H$ with $\varphi_y(x) = \langle x, z_y \rangle_H$ for all x, and $\|z_y\| = \|\varphi_y\| \leq \|T\| \, \|y\|$. Define $T^* y = z_y$.

LINEARITY. For $y_1, y_2 \in K$ and $\alpha \in \mathbb{K}$,

$$\langle x, T^*(y_1 + \alpha y_2) \rangle_H = \langle Tx, y_1 + \alpha y_2 \rangle_K = \langle Tx, y_1 \rangle_K + \overline{\alpha} \, \langle Tx, y_2 \rangle_K = \langle x, T^* y_1 + \overline{\alpha} T^* y_2 \rangle_H.$$

Wait—in a complex Hilbert space, $\langle Tx, \alpha y \rangle = \overline{\alpha} \, \langle Tx, y \rangle$, so

$$\langle x, T^*(\alpha y) \rangle_H = \overline{\alpha} \, \langle x, T^* y \rangle_H = \langle x, \alpha T^* y \rangle_H.$$

Hence $T^*(\alpha y) = \alpha T^* y$ and T^* is linear.

BOUNDEDNESS AND NORM. We have $\|T^*\| \leq \|T\|$ from the bound $\|T^* y\| \leq \|T\| \, \|y\|$. For the reverse inequality,

$$\|Tx\|_K = \sup_{\|y\| \leq 1} |\langle Tx, y \rangle_K| = \sup_{\|y\| \leq 1} |\langle x, T^* y \rangle_H| \leq \|x\| \, \|T^*\|,$$

whence $\|T\| \leq \|T^*\|$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

> **Proposition 3.29** (Properties of the adjoint). *Let $S, T \in \mathcal{L}(H, K)$, $U \in \mathcal{L}(K, L)$, and $\alpha \in \mathbb{K}$.*
>
> *(i)* $(T + \alpha S)^* = T^* + \overline{\alpha} S^*$.
>
> *(ii)* $(UT)^* = T^* U^*$.
>
> *(iii)* $T^{**} = T$.
>
> *(iv)* $\|T^* T\| = \|T T^*\| = \|T\|^2$.

Proof. *(i)* Direct from the inner product definition.

(ii) For all $x \in H$ and $z \in L$,

$$\langle UTx, z \rangle_L = \langle Tx, U^* z \rangle_K = \langle x, T^* U^* z \rangle_H.$$

By uniqueness, $(UT)^* = T^* U^*$.

(iii) For all $x \in H$ and $y \in K$, $\langle y, T^{**} x \rangle_K = \langle T^* y, x \rangle_H = \overline{\langle x, T^* y \rangle_H} = \overline{\langle Tx, y \rangle_K} = \langle y, Tx \rangle_K$, so $T^{**} = T$.

(iv) We have $\|T^* T\| \leq \|T^*\| \, \|T\| = \|T\|^2$. Conversely,

$$\|Tx\|^2 = \langle Tx, Tx \rangle_K = \langle T^* Tx, x \rangle_H \leq \|T^* Tx\| \, \|x\| \leq \|T^* T\| \, \|x\|^2.$$

Taking the supremum over $\|x\| \leq 1$ gives $\|T\|^2 \leq \|T^* T\|$. The equality $\|T T^*\| = \|T\|^2$ follows by replacing T with T^*. $\qquad\qquad\qquad\square$

> *Remark* 3.30 (The C^*-identity). Property (iv) is the C^*-*identity*. It shows that $\mathcal{L}(H)$ with the adjoint operation is a C^*-algebra.

3.8   Normal, self-adjoint, and unitary operators

> **Definition 3.31.** Let $T \in \mathcal{L}(H)$.
>
>   (i) T is *normal* if $T^*T = TT^*$.
>
>   (ii) T is *self-adjoint* (or *Hermitian*) if $T^* = T$.
>
>   (iii) T is *unitary* if $T^*T = TT^* = \mathrm{Id}$.
>
>   (iv) T is a *(orthogonal) projection* if $T^2 = T$ and $T^* = T$.

> **Proposition 3.32** (Characterizations of normality). *The following are equivalent for $T \in \mathcal{L}(H)$:*
>
>   *(i) T is normal.*
>
>   *(ii) $\|Tx\| = \|T^*x\|$ for all $x \in H$.*
>
>   *(iii) $T = A + iB$ where A, B are self-adjoint and $AB = BA$.*

Proof. (i)$\Rightarrow$(ii): $\|Tx\|^2 = \langle T^*Tx, x \rangle = \langle TT^*x, x \rangle = \|T^*x\|^2$.

(ii)$\Rightarrow$(i): Polarization gives $\langle T^*Tx, y \rangle = \langle TT^*x, y \rangle$ for all x, y, so $T^*T = TT^*$.

(i)$\Leftrightarrow$(iii): Set $A = (T + T^*)/2$, $B = (T - T^*)/(2i)$, both self-adjoint. Then $T = A + iB$ and

$$TT^* - T^*T = 2i(AB - BA),$$

so normality is equivalent to $AB = BA$. $\square$

> **Proposition 3.33** (Properties of self-adjoint operators). *Let $T \in \mathcal{L}(H)$ be self-adjoint. Then:*
>
>   *(i) $\langle Tx, x \rangle \in \mathbb{R}$ for all $x \in H$.*
>
>   *(ii) $\|T\| = \sup_{\|x\| \leq 1} |\langle Tx, x \rangle|$.*

(iii) If $\mathbb{K} = \mathbb{C}$ and $\langle Tx, x \rangle = 0$ for all x, then $T = 0$.

Proof. *(i)* $\overline{\langle Tx, x \rangle} = \langle x, Tx \rangle = \langle T^* x, x \rangle = \langle Tx, x \rangle$.

(ii) Let $M = \sup_{\|x\| \leq 1} |\langle Tx, x \rangle|$. Clearly $M \leq \|T\|$. For the converse, the parallelogram-type identity

$$4 \operatorname{Re} \langle Tx, y \rangle = \langle T(x + y), x + y \rangle - \langle T(x - y), x - y \rangle$$

(using $T^* = T$) gives $|\operatorname{Re} \langle Tx, y \rangle| \leq M \cdot \frac{1}{4}(\|x + y\|^2 + \|x - y\|^2) = \frac{M}{2}(\|x\|^2 + \|y\|^2)$. Taking x, y on the unit sphere and choosing the phase of y to make $\langle Tx, y \rangle$ real and positive, $|\langle Tx, y \rangle| \leq M$. Hence $\|Tx\| = \sup_{\|y\|=1} |\langle Tx, y \rangle| \leq M$ and $\|T\| \leq M$.

(iii) By polarization, $\langle Tx, y \rangle = 0$ for all x, y, so $T = 0$. $\qquad\square$

Theorem 3.34 (Spectrum of a self-adjoint operator is real)**.** *Let H be a complex Hilbert space and $T \in \mathcal{L}(H)$ self-adjoint. Then $\sigma(T) \subset \mathbb{R}$. More precisely, $\sigma(T) \subset [m, M]$ where $m = \inf_{\|x\|=1} \langle Tx, x \rangle$ and $M = \sup_{\|x\|=1} \langle Tx, x \rangle$, and both $m, M \in \sigma(T)$.*

Proof. STEP 1: $\sigma(T) \subset \mathbb{R}$. Let $\lambda = \alpha + i\beta$ with $\beta \neq 0$. For $x \in H$ with $\|x\| = 1$,

$$\|(\lambda \operatorname{Id} - T)x\|^2 = \|(\alpha \operatorname{Id} - T)x\|^2 + \beta^2 \|x\|^2 \geq \beta^2 > 0.$$

(Here we used $\operatorname{Re} \langle (\alpha \operatorname{Id} - T)x, i\beta x \rangle = \beta \operatorname{Im} \langle (\alpha \operatorname{Id} - T)x, x \rangle = 0$ since $\langle (\alpha \operatorname{Id} - T)x, x \rangle \in \mathbb{R}$.) Hence $\lambda \operatorname{Id} - T$ is bounded below. Similarly, $\overline{\lambda} \operatorname{Id} - T^* = \overline{\lambda} \operatorname{Id} - T$ is bounded below, so $\operatorname{ran}(\lambda \operatorname{Id} - T)$ is dense (if $y \perp \operatorname{ran}(\lambda \operatorname{Id} - T)$, then $(\overline{\lambda} \operatorname{Id} - T)y = 0$, contradicting $\overline{\lambda} \operatorname{Id} - T$ bounded below). By the open mapping theorem, $\lambda \operatorname{Id} - T$ is invertible.

STEP 2: $\sigma(T) \subset [m, M]$. If $\lambda > M$, then $\langle (\lambda \operatorname{Id} - T)x, x \rangle \geq (\lambda - M) \|x\|^2$, so $\lambda \operatorname{Id} - T$ is bounded below and, since self-adjoint, invertible. Similarly for $\lambda < m$.

STEP 3: $m, M \in \sigma(T)$. We show $M \in \sigma(T)$; the argument for m is analogous. By definition of M, there exists a sequence (x_n) with $\|x_n\| = 1$ and $\langle Tx_n, x_n \rangle \to M$. Then

$$\|(M \operatorname{Id} - T)x_n\|^2 = M^2 - 2M \langle Tx_n, x_n \rangle + \|Tx_n\|^2.$$

Since $\langle Tx_n, x_n \rangle \to M$ and $\|Tx_n\|^2 = \langle T^2 x_n, x_n \rangle \leq \|T^2\| \leq \|T\|^2$, while also $\|Tx_n\|^2 \geq \langle Tx_n, x_n \rangle^2 \to M^2$, we deduce $\|Tx_n\|^2 \to M^2$. Hence $\|(M \operatorname{Id} - T)x_n\|^2 \to M^2 - 2M \cdot M + M^2 = 0$. So $M \operatorname{Id} - T$ is not bounded below, hence $M \in \sigma(T)$. $\qquad\square$

Proposition 3.35 (Spectral properties of unitary operators). *If $U \in \mathcal{L}(H)$ is unitary, then $\sigma(U) \subset \{z \in \mathbb{C} : |z| = 1\}$.*

Proof. Since $\|U\| = 1$, $\sigma(U) \subset \overline{\mathbb{D}}$. Since $U^{-1} = U^*$ with $\|U^*\| = 1$, $\sigma(U^{-1}) \subset \overline{\mathbb{D}}$. But $\lambda \in \sigma(U)$ iff $\lambda^{-1} \in \sigma(U^{-1})$, so $|\lambda| \le 1$ and $|\lambda|^{-1} \le 1$, giving $|\lambda| = 1$. $\qquad\square$

Proposition 3.36 (Orthogonal projections). *Let $P \in \mathcal{L}(H)$. Then P is an orthogonal projection if and only if there exists a closed subspace $M \subset H$ such that P is the orthogonal projection onto M. In that case, $\sigma(P) \subset \{0, 1\}$.*

Proof. If $P^2 = P = P^*$, set $M = \operatorname{ran} P = \operatorname{Ker}(\operatorname{Id} - P)$. Since P is continuous, M is closed. For $x \in H$, write $x = Px + (x - Px)$. We have $Px \in M$ and $P(x - Px) = Px - P^2 x = 0$, so $x - Px \in \operatorname{Ker} P$. Moreover $\langle Px, x - Px \rangle = \langle Px, x - Px \rangle = \langle P^2 x, x - Px \rangle = \langle Px, P(x - Px) \rangle = 0$. So P is the orthogonal projection onto M.

Conversely, if P is the orthogonal projection onto a closed subspace M, then $P^2 = P$ and $\langle Px, y \rangle = \langle Px, Py + (y - Py) \rangle = \langle Px, Py \rangle = \langle x, Py \rangle$ (since $Px \perp (y - Py)$ and $x - Px \perp Py$), so $P^* = P$.

For the spectrum: if $\lambda \notin \{0, 1\}$, then $(\lambda \operatorname{Id} - P)^{-1} = \frac{1}{\lambda} \operatorname{Id} + \frac{1}{\lambda(\lambda - 1)} P$. $\qquad\square$

Proposition 3.37 (Kernel and range of the adjoint). *Let $T \in \mathcal{L}(H, K)$. Then:*

  (i) $\operatorname{Ker} T^* = (\operatorname{ran} T)^\perp$.

  (ii) $\operatorname{Ker} T = (\operatorname{ran} T^*)^\perp$.

  (iii) $\overline{\operatorname{ran} T} = (\operatorname{Ker} T^*)^\perp$.

  (iv) $\overline{\operatorname{ran} T^*} = (\operatorname{Ker} T)^\perp$.

Proof. (i) $y \in \operatorname{Ker} T^*$ iff $T^* y = 0$ iff $\langle x, T^* y \rangle = 0$ for all $x \in H$ iff $\langle Tx, y \rangle = 0$ for all $x \in H$ iff $y \in (\operatorname{ran} T)^\perp$.

(ii) Apply (i) to T^*, noting $T^{**} = T$.

(iii) and (iv) follow from (i) and (ii) using $M^{\perp\perp} = \overline{M}$. $\qquad\square$

Example 3.38 (Matrix operators). If $H = \mathbb{C}^n$ and T is represented by a matrix $A = (a_{ij})$, then T^* corresponds to $A^* = (\overline{a_{ji}})$, the conjugate

transpose. Self-adjoint means $A = A^*$ (Hermitian matrix), unitary means $A^*A = I$ (unitary matrix), and normal means $A^*A = AA^*$.

Example 3.39 (Adjoint of a multiplication operator). On $L^2(\mu)$, the multiplication operator M_φ with $\varphi \in L^\infty(\mu)$ satisfies $M_\varphi^* = M_{\overline{\varphi}}$. Indeed,

$$\langle M_\varphi f, g \rangle = \int \varphi f \, \overline{g} \, \mathrm{d}\mu = \int f \, \overline{\overline{\varphi} g} \, \mathrm{d}\mu = \langle f, M_{\overline{\varphi}} g \rangle \,.$$

Thus M_φ is self-adjoint iff φ is real-valued a.e., normal always (since multiplication operators commute), and unitary iff $|\varphi| = 1$ a.e.

Example 3.40 (Adjoint of the right shift). The right shift $R\colon \ell^2 \to \ell^2$, $R(x_1, x_2, \dots) = (0, x_1, x_2, \dots)$, has adjoint $R^* = L$, the left shift: $L(x_1, x_2, x_3, \dots) = (x_2, x_3, \dots)$. Indeed, $\langle Rx, y \rangle = \sum_{n=1}^\infty x_n \overline{y_{n+1}} = \langle x, Ly \rangle$. Note that $R^*R = \mathrm{Id}$ but $RR^* \neq \mathrm{Id}$ (it is the projection onto $\{0\} \oplus \ell^2$), so R is an isometry but not unitary.

3.9 Exercises

Exercise 3.1 ($\star$). Let $A\colon \mathbb{R}^n \to \mathbb{R}^n$ be a linear map represented by a matrix (a_{ij}). Show that $\|A\| \leq \left(\sum_{i,j} a_{ij}^2\right)^{1/2}$ (the Frobenius norm).

Exercise 3.2 ($\star$). Let $(d_n) \subset \mathbb{C}$ be a bounded sequence and define $T\colon \ell^2 \to \ell^2$ by $Te_n = d_n e_n$. Show that T is compact if and only if $d_n \to 0$.

Exercise 3.3 ($\star\star$). Let R be the unilateral right shift on $\ell^2(\mathbb{N})$. Compute $\sigma(R)$, $\sigma_p(R)$, $\sigma_c(R)$, and $\sigma_r(R)$. Compute the same for the left shift $L = R^*$.

Exercise 3.4 ($\star$). Let $T \in \mathcal{L}(E)$ be nilpotent of order N ($T^N = 0$ but $T^{N-1} \neq 0$). Compute $(\mathrm{Id} - T)^{-1}$ explicitly.

Exercise 3.5 ($\star\star$). Let $T \in \mathcal{K}(E)$ be a compact operator on an infinite-dimensional Banach space. Show that $0 \in \sigma(T)$.

Exercise 3.6 ($\star\star$). Let $T \in \mathcal{L}(H)$. Prove that:

(a) $\operatorname{Ker} T^* = (\operatorname{ran} T)^\perp$.

(b) $\overline{\operatorname{ran} T} = (\operatorname{Ker} T^*)^\perp$.

(c) T is injective if and only if $\operatorname{ran} T^*$ is dense.

Exercise 3.7 ($\star\star$). Let $T \in \mathcal{L}(H)$ be normal. Show that eigenspaces corresponding to distinct eigenvalues are orthogonal. Show that $\operatorname{Ker}(T - \lambda \operatorname{Id}) = \operatorname{Ker}(T^* - \overline{\lambda} \operatorname{Id})$.

Exercise 3.8 ($\star\star\star$). Consider the Volterra operator $V \colon L^2([0,1]) \to L^2([0,1])$, $(Vf)(x) = \int_0^x f(t)\,\mathrm{d}t$.

(a) Show that V is compact.

(b) Show that V has no eigenvalues.

(c) Conclude $\sigma(V) = \{0\}$ and compute $r(V)$.

(d) Compute $\|V\|$. (*Hint:* use V^* and the C^*-identity.)

Exercise 3.9 ($\star$). An operator $T \in \mathcal{L}(H)$ is *positive* (written $T \geq 0$) if $T = T^*$ and $\langle Tx, x \rangle \geq 0$ for all x. Show that $T^*T \geq 0$ for any $T \in \mathcal{L}(H)$.

Exercise 3.10 ($\star\star\star$). Let $T \in \mathcal{L}(H)$ be self-adjoint. Show that every $\lambda \in \sigma(T)$ is an *approximate eigenvalue*: there exists (x_n) with $\|x_n\| = 1$ and $\|Tx_n - \lambda x_n\| \to 0$.

Exercise 3.11 ($\star\star$). Let $U \in \mathcal{L}(H)$. Show the following are equivalent: (i) U is unitary; (ii) U is surjective and $\|Ux\| = \|x\|$ for all x; (iii) U is surjective and $\langle Ux, Uy \rangle = \langle x, y \rangle$ for all x, y.

Exercise 3.12 ($\star$). Show that for any $T \in \mathcal{L}(E)$ and $\lambda \in \rho(T)$, $\|R(\lambda, T)\| \geq \frac{1}{\text{dist}(\lambda, \sigma(T))}$.

Exercise 3.13 ($\star\star$). Let $T \in \mathcal{K}(E)$ where E is an infinite-dimensional Banach space.

(a) Show that every nonzero $\lambda \in \sigma(T)$ is an eigenvalue.

(b) Show that $\sigma(T) \setminus \{0\}$ is at most countable, with 0 as the only possible accumulation point.

(*Hint:* use the Riesz lemma for part (a) and the fact that eigenspaces for distinct eigenvalues of a compact operator are finite-dimensional.)

Exercise 3.14 ($\star\star\star$). The *numerical range* of $T \in \mathcal{L}(H)$ is $W(T) = \{\langle Tx, x \rangle : \|x\| = 1\}$.

(a) Show that $W(T)$ is convex (Toeplitz–Hausdorff theorem; this is hard).

(b) Show that $\sigma(T) \subset \overline{W(T)}$.

(c) Show that if T is normal, $\overline{W(T)} = \text{conv}(\sigma(T))$.

Exercise 3.15 ($\star\star$). Let $T \in \mathcal{K}(H)$ be self-adjoint on a separable Hilbert space H. Show that either $\|T\|$ or $-\|T\|$ is an eigenvalue of T. (*Hint:* use the fact that $\|T\| = \sup_{\|x\|=1} |\langle Tx, x \rangle|$ and extract a convergent subsequence.)

Chapter 4

The Hahn–Banach Theorem and Consequences

Contents

The Hahn–Banach theorem is one of the three pillars of functional analysis (together with the uniform boundedness principle and the open mapping/closed graph theorems). It guarantees that normed spaces have a rich supply of continuous linear functionals, and its geometric formulations provide powerful convex separation results. This chapter presents the analytic and geometric forms, and develops fundamental consequences: the characterisation of the dual space, reflexivity, and the Banach–Alaoglu theorem.

4.1 Analytic form: real case

Definition 4.1 (Sublinear functional). A function $p\colon E \to \mathbb{R}$ on a real vector space E is *sublinear* if

(i) $p(\lambda x) = \lambda\, p(x)$ for all $\lambda \geq 0$ and $x \in E$ (positive homogeneity), and

(ii) $p(x + y) \leq p(x) + p(y)$ for all $x, y \in E$ (subadditivity).

Theorem 4.2 (Hahn–Banach, real analytic form). *Let E be a real vector space, $p\colon E \to \mathbb{R}$ a sublinear functional, $G \subset E$ a linear subspace, and $g\colon G \to \mathbb{R}$ a linear functional satisfying $g(x) \leq p(x)$ for all $x \in G$. Then there exists a linear functional $f\colon E \to \mathbb{R}$ extending g (i.e., $f|_G = g$) such that $f(x) \leq p(x)$ for all $x \in E$.*

Proof. STEP 1: EXTENSION BY ONE DIMENSION. Let $x_0 \in E \setminus G$ and set $G_1 = G \oplus \mathbb{R}x_0$. We seek $c \in \mathbb{R}$ such that $f_1(x + tx_0) = g(x) + tc$ satisfies $f_1 \leq p$ on G_1, i.e.,

$$g(x) + tc \leq p(x + tx_0) \qquad \text{for all } x \in G,\ t \in \mathbb{R}.$$

For $t > 0$: $c \leq \frac{1}{t}\big(p(x+tx_0)-g(x)\big) = p(x/t+x_0)-g(x/t)$, so $c \leq \inf_{u \in G}\big(p(u+x_0) - g(u)\big)$.

For $t < 0$ (set $s = -t > 0$): $-c \leq p(x - sx_0)/s - g(x/s) = p(v-x_0)-g(v)$ for $v = x/s$, so $c \geq \sup_{v \in G}\big(g(v) - p(v - x_0)\big)$.

We must verify the sup $\leq$ the inf. For any $u, v \in G$:

$$g(v) - p(v - x_0) \leq g(u) + p(u + x_0) - p(v - x_0) - g(u)$$

holds if $g(v) + g(u) \leq p(v - x_0) + p(u + x_0)$. But $g(v) + g(u) = g(v + u) \leq p(v+u) \leq p(v-x_0)+p(u+x_0)$ by subadditivity. So the choice of c is possible and the one-step extension works.

STEP 2: APPLICATION OF ZORN'S LEMMA. Consider the set

$$\mathcal{P} = \big\{(H, h) : G \subset H \subset E \text{ subspace}, h\colon H \to \mathbb{R} \text{ linear}, h|_G = g, h \leq p \text{ on } H\big\},$$

partially ordered by $(H_1, h_1) \leq (H_2, h_2)$ iff $H_1 \subset H_2$ and $h_2|_{H_1} = h_1$.

Every chain $\{(H_i, h_i)\}_{i \in I}$ has an upper bound: set $H = \bigcup_i H_i$ and $h(x) = h_i(x)$ if $x \in H_i$. This is well-defined, linear, extends g, and satisfies $h \leq p$.

By Zorn's lemma, $\mathcal{P}$ has a maximal element (H_0, f). If $H_0 \neq E$, then Step 1 would yield a strict extension, contradicting maximality. Hence $H_0 = E$. $\qquad\square$

4.2 Analytic form: complex case

> **Theorem 4.3** (Hahn–Banach, complex analytic form)**.** *Let E be a complex vector space, $p\colon E \to [0, \infty)$ a seminorm, $G \subset E$ a linear subspace, and $g\colon G \to \mathbb{C}$ a linear functional satisfying $|g(x)| \le p(x)$ for all $x \in G$. Then there exists a linear functional $f\colon E \to \mathbb{C}$ extending g such that $|f(x)| \le p(x)$ for all $x \in E$.*

Proof. Write $g = g_1 + i\, g_2$ where $g_1 = \operatorname{Re} g$ and $g_2 = \operatorname{Im} g$ are real-linear functionals on G (viewed as a real vector space). Note that $g(ix) = ig(x)$ gives $g_1(ix) + ig_2(ix) = -g_2(x) + ig_1(x)$, so

$$g_2(x) = -g_1(ix). \tag{4.1}$$

Moreover, $g_1(x) = \operatorname{Re} g(x) \le |g(x)| \le p(x)$, and p is a sublinear functional on $E_\mathbb{R}$ (the real vector space underlying E).

By the real Hahn–Banach theorem (Theorem 4.2), extend g_1 to a real-linear functional $f_1\colon E \to \mathbb{R}$ with $f_1 \le p$ on E. Define

$$f(x) = f_1(x) - i\, f_1(ix).$$

Then f is complex-linear: $f(ix) = f_1(ix) - if_1(-x) = f_1(ix) + if_1(x) = if(x)$. Also $f|_G = g$ by (4.1).

For the bound, write $f(x) = |f(x)|\, e^{i\theta}$. Then $|f(x)| = f(e^{-i\theta}x) = f_1(e^{-i\theta}x) \le p(e^{-i\theta}x) = p(x)$. $\qquad\square$

4.3 Consequences: extension of functionals, separation of points

> **Corollary 4.4** (Extension of bounded linear functionals)**.** *Let E be a normed space, $G \subset E$ a subspace, and $g \in G'$. Then there exists $f \in E'$ with $f|_G = g$ and $\|f\|_{E'} = \|g\|_{G'}$.*

Proof. Apply Theorem 4.3 (or Theorem 4.2 in the real case) with $p(x) = \|g\|\,\|x\|$. The extension satisfies $|f(x)| \le \|g\|\,\|x\|$, so $\|f\| \le \|g\|$. Since $f|_G = g$, $\|f\| \ge \|g\|$. $\qquad\square$

Corollary 4.5 (Separation of points by functionals). *Let E be a normed space and $x_0 \in E$ with $x_0 \neq 0$. Then there exists $f \in E'$ with $\|f\| = 1$ and $f(x_0) = \|x_0\|$.*

Proof. On $G = \mathbb{K}x_0$, define $g(\alpha x_0) = \alpha \|x_0\|$. Then $\|g\| = 1$. Extend by Corollary 4.4. □

Corollary 4.6 (The dual separates points). *If $x, y \in E$ and $f(x) = f(y)$ for all $f \in E'$, then $x = y$. Equivalently, $\|x\| = \sup\{|f(x)| : f \in E', \|f\| \leq 1\}$.*

Proof. Apply Corollary 4.5 to $x - y$. □

Corollary 4.7 (Functionals that vanish on a subspace). *Let G be a closed subspace of a normed space E and $x_0 \notin G$. Then there exists $f \in E'$ with $f|_G = 0$, $\|f\| = 1$, and $f(x_0) = \operatorname{dist}(x_0, G) > 0$.*

Proof. Set $d = \operatorname{dist}(x_0, G) > 0$ (since G is closed). On $H = G \oplus \mathbb{K}x_0$, define $g(y + \alpha x_0) = \alpha d$. Then $|g(y + \alpha x_0)| = |\alpha|\, d \leq |\alpha|\, \|x_0 - (-y/\alpha)\| = \|\alpha x_0 + y\|$ when $\alpha \neq 0$, and $g(y) = 0$ when $\alpha = 0$. So $\|g\| \leq 1$, and in fact $\|g\| = 1$ since for $y_n \in G$ with $\|x_0 - y_n\| \to d$, $|g(x_0 - y_n)| / \|x_0 - y_n\| = d / \|x_0 - y_n\| \to 1$. Extend by Corollary 4.4. □

4.4 Geometric forms: separation of convex sets

Definition 4.8 (Hyperplane). A *(closed) hyperplane* in E is a set of the form $H = \{x \in E : f(x) = \alpha\}$ where $f \in E' \setminus \{0\}$ and $\alpha \in \mathbb{R}$. A *supporting hyperplane* at a point x_0 of a convex set C is a hyperplane H containing x_0 such that C lies in one of the closed half-spaces determined by H.

Definition 4.9 (Minkowski functional). Let C be a convex subset of a real vector space E with $0 \in \operatorname{int}(C)$ (the algebraic interior, or core). The *Minkowski functional* (or *gauge*) of C is

$$p_C(x) = \inf\{t > 0 : x \in tC\} = \inf\{t > 0 : x/t \in C\}.$$

Proposition 4.10 (Properties of the Minkowski functional). *Let C be a convex set in a real normed space E with $0 \in \text{int}(C)$. Then:*

(i) p_C is sublinear (positively homogeneous and subadditive).

(ii) If C is open, then $C = \{x : p_C(x) < 1\}$.

(iii) If C is symmetric ($-C = C$), then p_C is a seminorm.

(iv) If C is open, bounded, and symmetric, then p_C is a norm equivalent to $\|\cdot\|$.

Proof. (i) Positive homogeneity is clear. For subadditivity: if $x \in sC$ and $y \in tC$, then $x + y = (s + t)\left(\frac{s}{s+t} \cdot \frac{x}{s} + \frac{t}{s+t} \cdot \frac{y}{t}\right) \in (s+t)C$ by convexity. So $p_C(x + y) \leq s + t$; taking infima gives $p_C(x + y) \leq p_C(x) + p_C(y)$.

(ii) If $p_C(x) < 1$, there exists $t < 1$ with $x \in tC \subset C$ (since $0 \in C$ and C is convex: $x = t(x/t) + (1 - t) \cdot 0 \in C$). Conversely, if $x \in C$ and C is open, there exists $\varepsilon > 0$ with $(1 + \varepsilon)x \in C$, so $p_C(x) \leq (1 + \varepsilon)^{-1} < 1$.

(iii) $p_C(-x) = \inf\{t > 0 : -x \in tC\} = \inf\{t > 0 : x \in t(-C)\} = \inf\{t > 0 : x \in tC\} = p_C(x)$.

(iv) Since C is bounded and open with $0 \in C$, there exist $r, R > 0$ with $B(0, r) \subset C \subset B(0, R)$. Then $\|x\| / R \leq p_C(x) \leq \|x\| / r$. $\qquad\square$

Theorem 4.11 (Separation of a point from a convex set). *Let E be a real normed space, $C \subset E$ a non-empty open convex set, and $x_0 \in E \setminus C$. Then there exists $f \in E' \setminus \{0\}$ and $\alpha \in \mathbb{R}$ such that*

$$f(x) < \alpha \leq f(x_0) \qquad \text{for all } x \in C.$$

Proof. Without loss of generality assume $0 \in C$ (translate if necessary). Let $p = p_C$ be the Minkowski functional. Since $x_0 \notin C$ and C is open, $p(x_0) \geq 1$ (by Proposition 4.10(ii)).

On $G = \mathbb{R}x_0$, define $g(tx_0) = t\, p(x_0)$. For $t \geq 0$, $g(tx_0) = tp(x_0) = p(tx_0)$. For $t < 0$, $g(tx_0) = tp(x_0) < 0 \leq p(tx_0)$. So $g \leq p$ on G.

By the real Hahn–Banach theorem, extend g to $f \colon E \to \mathbb{R}$ with $f \leq p$. For $x \in C$, $f(x) \leq p(x) < 1$ (since C is open). Also $f(x_0) = g(x_0) = p(x_0) \geq 1$. Set $\alpha = 1$.

It remains to verify f is continuous. Since $f \leq p$ and $p(x) \leq \|x\| / r$ for some $r > 0$ (as $B(0, r) \subset C$), $f(x) \leq \|x\| / r$. Also $f(-x) \leq p(-x)$; if C is not symmetric we still have $-C$ containing a ball, giving $f(-x) \leq \|x\| / r'$. In general, since $0 \in \text{int}(C)$, there exists $r > 0$ with $B(0, r) \subset C$, so $p(x) \leq$

$\|x\|/r$ and $f(x) \leq \|x\|/r$. Similarly $-f(x) = f(-x) \leq p(-x) \leq \|-x\|/r$, giving $|f(x)| \leq \|x\|/r$. (For the bound on $p(-x)$: since $0 \in C$ and C is open, $B(0, r) \subset C$ for some $r > 0$, and also $-r' \cdot x/\|x\| \in C$ for some $r' > 0$, but more directly, for x with $\|x\| \leq r$, $-x \in B(0, r) \subset C$ so $p(-x) < 1 \leq \|-x\|/r \cdot r/\|x\|$ gives what we need.) Hence $f \in E'$. □

> **Theorem 4.12** (Strict separation of convex sets)**.** *Let E be a real normed space, $A \subset E$ compact and convex, $B \subset E$ closed and convex, with $A \cap B = \emptyset$. Then there exist $f \in E' \setminus \{0\}$ and $\alpha, \beta \in \mathbb{R}$ with $\alpha < \beta$ such that*
>
> $$f(a) \leq \alpha < \beta \leq f(b) \qquad \text{for all } a \in A,\ b \in B.$$

Proof. STEP 1. The set $A - B = \{a - b : a \in A, b \in B\}$ is closed and convex (closed because A is compact and B is closed: if $a_n - b_n \to z$, extract $a_{n_k} \to a \in A$, then $b_{n_k} \to a - z \in B$, so $z = a - (a - z) \in A - B$). Also $0 \notin A - B$ since $A \cap B = \emptyset$.

STEP 2. Since $A - B$ is closed and $0 \notin A - B$, there exists $r > 0$ with $B(0, r) \cap (A - B) = \emptyset$. Then $C = (A - B) + B(0, r)$ is an open convex set not containing 0 (if $0 \in C$, then there exist $a \in A$, $b \in B$, u with $\|u\| < r$ and $a - b + u = 0$, so $a - b = -u$ and $\|a - b\| < r$, contradicting $\text{dist}(A, B) \geq r$). Actually, let us re-examine: the distance $d = \inf\{\|a - b\| : a \in A, b \in B\} > 0$ by compactness of A and closedness of B. Set $C = A - B + B(0, d/2)$; this is open, convex, and $0 \notin C$.

STEP 3. By Theorem 4.11 applied to C and $x_0 = 0$: there exists $f \in E' \setminus \{0\}$ and γ with $f(x) < \gamma \leq f(0) = 0$ for all $x \in C$. In particular, for all $a \in A$ and $b \in B$, $f(a - b) < 0$, i.e., $f(a) < f(b)$. Then $\sup_{a \in A} f(a) \leq \inf_{b \in B} f(b)$. Since A is compact, the sup is a max; set $\alpha = \max_A f$ and $\beta = \inf_B f$. We need $\alpha < \beta$. Indeed, $f(a - b + u) < 0$ for all $\|u\| < d/2$, so taking $u \to 0$, $f(a) \leq f(b)$ with equality only if $f(a - b) = 0$. But then $f(a - b + u) = f(u) < 0$ for suitable u, contradicting that f can take positive values on $B(0, d/2)$ (since $f \neq 0$). More precisely: since $f \neq 0$, $\sup_{\|u\| < d/2} f(u) > 0$, so for any $a \in A$, $b \in B$: $f(a) + \sup f(u) \leq f(b)$, giving $f(a) < f(b)$. Hence $\alpha < \beta$. □

> *Remark* 4.13. In a complex normed space, the geometric Hahn–Banach theorem applies to the real parts: if $f \in E'$ (complex), then $\text{Re}\, f$ separates, and f is determined by $\text{Re}\, f$.

> **Corollary 4.14** (Supporting hyperplane theorem)**.** *Let C be a closed convex subset of a normed space E and $x_0 \in \partial C$. Then there exists a supporting hyperplane to C at x_0: a functional $f \in E' \setminus \{0\}$ with $\operatorname{Re} f(x) \leq \operatorname{Re} f(x_0)$ for all $x \in C$.*

Proof. Since $x_0 \in \partial C$, there exist $y_n \notin C$ with $y_n \to x_0$. For each n, by Theorem 4.11 (applied to $\operatorname{int}(C)$ if it is nonempty, or a slight adaptation otherwise), there exists $f_n \in E'$ with $\|f_n\| = 1$ and $\operatorname{Re} f_n(x) \leq \operatorname{Re} f_n(y_n)$ for all $x \in C$. By weak*-compactness of the unit ball (Banach–Alaoglu, Theorem 8.20 below), extract a weak*-convergent subnet $f_{n_\alpha} \overset{*}{\rightharpoonup} f$ with $\|f\| \leq 1$. Then $\operatorname{Re} f(x) \leq \operatorname{Re} f(x_0)$ for all $x \in C$, and $f \neq 0$ since $\operatorname{Re} f(x_0) = \lim \operatorname{Re} f_{n_\alpha}(y_{n_\alpha}) \geq \limsup \operatorname{Re} f_{n_\alpha}(x)$ for any interior point x. $\square$

4.5 Applications to dual spaces

> **Theorem 4.15** (Dual of ℓ^p)**.** *For $1 \leq p < \infty$, the map $\Phi\colon \ell^q \to (\ell^p)'$ defined by*
>
> $$\Phi(y)(x) = \sum_{n=1}^{\infty} x_n y_n, \qquad y = (y_n) \in \ell^q,\ x = (x_n) \in \ell^p,$$
>
> *where $\frac{1}{p} + \frac{1}{q} = 1$ (with $q = \infty$ when $p = 1$), is an isometric isomorphism.*

Proof. Φ IS ISOMETRIC. By Hölder's inequality, $|\Phi(y)(x)| \leq \|x\|_p \|y\|_q$, so $\|\Phi(y)\| \leq \|y\|_q$. For the reverse, when $1 < p < \infty$, take $x_n = |y_n|^{q/p} \operatorname{sgn}(\overline{y_n})$. Then $\|x\|_p^p = \sum |y_n|^q = \|y\|_q^q$ and $\Phi(y)(x) = \sum |y_n|^{q/p+1} = \|y\|_q^q$, so $\|\Phi(y)\| \geq \|y\|_q^q / \|y\|_q^{q/p} = \|y\|_q$. The case $p = 1$ is similar.

 Φ IS SURJECTIVE. Let $f \in (\ell^p)'$. Set $y_n = f(e_n)$ where e_n is the n-th standard basis vector. For any finite set $F \subset \mathbb{N}$, the choice $x = \sum_{n \in F} |y_n|^{q/p} \operatorname{sgn}(\overline{y_n})\, e_n$ gives

$$\sum_{n \in F} |y_n|^q = f(x) \leq \|f\|\, \|x\|_p = \|f\| \left(\sum_{n \in F} |y_n|^q \right)^{1/p},$$

so $\left(\sum_{n \in F} |y_n|^q \right)^{1/q} \leq \|f\|$. Thus $y \in \ell^q$ with $\|y\|_q \leq \|f\|$. By density of finitely supported sequences, $f = \Phi(y)$. $\square$

Theorem 4.16 (Riesz–Markov representation, statement)**.** *The dual of $C([0,1])$ is isometrically isomorphic to $\mathcal{M}([0,1])$, the space of finite signed (or complex) Radon measures on $[0,1]$, via the pairing*

$$\mu \mapsto \left(f \mapsto \int_0^1 f \, d\mu\right).$$

Remark 4.17. The proof of the Riesz–Markov theorem requires measure-theoretic tools and will be presented in Chapter 9 when we study the theory of distributions and measures on locally compact spaces.

4.6 Bidual and canonical embedding

Definition 4.18 (Bidual)**.** The *bidual* of a normed space E is $E'' = (E')'$, the dual of the dual space.

Theorem 4.19 (Canonical embedding is isometric)**.** *The map $J\colon E \to E''$ defined by*

$$J(x)(f) = f(x), \qquad x \in E, \ f \in E',$$

is a linear isometric embedding (i.e., $\|J(x)\|_{E''} = \|x\|_E$ for all $x \in E$).

Proof. Linearity is clear. For the norm: $|J(x)(f)| = |f(x)| \leq \|f\| \, \|x\|$, so $\|J(x)\| \leq \|x\|$. Conversely, by Corollary 4.5, there exists $f_0 \in E'$ with $\|f_0\| = 1$ and $f_0(x) = \|x\|$. Then $|J(x)(f_0)| = \|x\|$, so $\|J(x)\| \geq \|x\|$. □

Definition 4.20 (Reflexive space)**.** A Banach space E is *reflexive* if $J\colon E \to E''$ is surjective (and hence an isometric isomorphism).

Example 4.21. (i) Every Hilbert space is reflexive (by the Riesz representation theorem).

(ii) ℓ^p and $L^p(\mu)$ are reflexive for $1 < p < \infty$.

(iii) ℓ^1, ℓ^∞, c_0, $L^1(\mu)$, $L^\infty(\mu)$, and $C([0,1])$ are not reflexive.

Proposition 4.22. *A closed subspace of a reflexive Banach space is reflexive.*

Proof. Let F be a closed subspace of a reflexive space E. Let $\xi \in F''$. Define $\Phi \in E''$ by $\Phi(f) = \xi(f|_F)$ for $f \in E'$. Since E is reflexive, $\Phi = J_E(x)$ for some $x \in E$, i.e., $f(x) = \xi(f|_F)$ for all $f \in E'$.

We claim $x \in F$. If $x \notin F$, by Corollary 4.7 there exists $f \in E'$ with $f|_F = 0$ and $f(x) \neq 0$. But then $f(x) = \xi(f|_F) = \xi(0) = 0$, a contradiction.

So $x \in F$ and for any $g \in F'$, extend g to $f \in E'$ (by Corollary 4.4). Then $\xi(g) = \xi(f|_F) = f(x) = g(x) = J_F(x)(g)$. Hence $\xi = J_F(x)$ and J_F is surjective. $\qquad\square$

Proposition 4.23. *A quotient of a reflexive Banach space by a closed subspace is reflexive.*

Proof. Let E be reflexive and $G \subset E$ closed. Let $\pi \colon E \to E/G$ be the quotient map. Given $\xi \in (E/G)''$, define $\Psi \in E''$ by $\Psi(f) = \xi(f \circ \pi^{-1})$ — more precisely, note that $\pi' \colon (E/G)' \to E'$ defined by $\pi'(g) = g \circ \pi$ is an isometric embedding with range $G^{\perp}$ (Exercise 4.4). Define $\Psi \in E''$ by $\Psi(f) = \xi((\pi')^{-1}(f))$ for $f \in G^{\perp}$ and extend to E'. Since E is reflexive, $\Psi = J_E(x)$ for some $x \in E$. Then for any $g \in (E/G)'$, $\xi(g) = \Psi(\pi'g) = (\pi'g)(x) = g(\pi(x)) = J_{E/G}(\pi(x))(g)$. Hence $\xi = J_{E/G}(\pi(x))$ and the quotient is reflexive. $\qquad\square$

Remark 4.24 (Characterizations of reflexivity). To summarize, for a Banach space E the following are equivalent:

(i) E is reflexive (i.e., J is surjective).

(ii) B_E is weakly compact (Corollary 4.33).

(iii) Every bounded sequence in E has a weakly convergent subsequence (Eberlein–Šmulian theorem, Chapter 8).

(iv) Every $f \in E'$ attains its norm on B_E (James' theorem).

4.7 The subdifferential

Definition 4.25 (Subdifferential). Let E be a real normed space and $\varphi \colon E \to \mathbb{R} \cup \{+\infty\}$ a convex function. The *subdifferential* of φ at a point $x_0 \in E$ where $\varphi(x_0) < +\infty$ is

$$\partial\varphi(x_0) = \big\{ f \in E' : \varphi(x) - \varphi(x_0) \geq f(x - x_0) \text{ for all } x \in E \big\}.$$

Elements of $\partial\varphi(x_0)$ are called *subgradients*.

Proposition 4.26 (Non-emptiness of the subdifferential). *If $\varphi \colon E \to \mathbb{R} \cup \{+\infty\}$ is convex, lower semicontinuous, and $\varphi(x_0) \in \mathbb{R}$, then $\partial\varphi(x_0) \neq \emptyset$.*

Proof. The epigraph $\mathrm{epi}(\varphi) = \{(x, t) : t \geq \varphi(x)\}$ is a closed convex subset of $E \times \mathbb{R}$, and $(x_0, \varphi(x_0)) \in \partial\, \mathrm{epi}(\varphi)$. By the supporting hyperplane theorem (Corollary 4.14), there exist $f \in E'$ and $\beta \in \mathbb{R}$, not both zero, such that

$$f(x) + \beta t \leq f(x_0) + \beta\varphi(x_0) \qquad \text{for all } (x, t) \in \mathrm{epi}(\varphi).$$

Since $(x_0, t) \in \mathrm{epi}(\varphi)$ for all $t \geq \varphi(x_0)$, we need $\beta t \leq \beta\varphi(x_0)$ for large t, so $\beta \leq 0$. If $\beta = 0$, then $f(x) \leq f(x_0)$ for all x in $\mathrm{dom}\,\varphi$, which (if $\mathrm{dom}\,\varphi$ contains a neighborhood of x_0) forces $f = 0$. Since φ is continuous at x_0 (being convex and finite, it is continuous on the interior of its effective domain), $\beta < 0$. Dividing by $-\beta > 0$:

$$\varphi(x) \geq \varphi(x_0) + (-f/\beta)(x - x_0),$$

so $-f/\beta \in \partial\varphi(x_0)$. $\qquad\square$

Example 4.27 (Subdifferential of the norm). For $\varphi(x) = \|x\|$ on a normed space E:

(i) If $x \neq 0$: $\partial\varphi(x) = \{ f \in E' : \|f\| = 1,\ f(x) = \|x\| \}$.

(ii) $\partial\varphi(0) = \{ f \in E' : \|f\| \leq 1 \} = B_{E'}$.

4.8 The Banach–Alaoglu theorem

Recall that the *weak* topology* on E' is the coarsest topology making all evaluation maps $f \mapsto f(x)$ $(x \in E)$ continuous.

Notation 4.28. We write $f_\alpha \xrightarrow{*} f$ to denote convergence in the weak* topology: $f_\alpha(x) \to f(x)$ for all $x \in E$.

Theorem 4.29 (Banach–Alaoglu). *Let E be a normed space. Then the closed unit ball $B_{E'} = \{f \in E' : \|f\| \leq 1\}$ is compact in the weak* topology.*

Remark 4.30. The proof uses Tychonoff's theorem and will be given in Chapter 8 (Weak Topologies). Here we state it and develop some immediate consequences.

Corollary 4.31. *Every bounded net (f_α) in E' has a weak*-convergent subnet.*

Corollary 4.32 (Goldstine's theorem). *Let E be a normed space. Then $J(B_E)$ is weak*-dense in $B_{E''}$.*

Proof. $J(B_E)$ is a convex subset of E''. If it were not weak*-dense in $B_{E''}$, there would exist $\xi_0 \in B_{E''}$ not in the weak*-closure of $J(B_E)$. By the geometric Hahn–Banach theorem in the weak* topology, there would exist $f \in E'$ and α such that $\operatorname{Re} J(x)(f) \leq \alpha < \operatorname{Re} \xi_0(f)$ for all $x \in B_E$. But $\operatorname{Re} J(x)(f) = \operatorname{Re} f(x)$, so $\sup_{\|x\| \leq 1} \operatorname{Re} f(x) = \|f\| \leq \alpha < \operatorname{Re} \xi_0(f) \leq \|\xi_0\| \, \|f\| \leq \|f\|$, a contradiction. $\qquad \square$

Corollary 4.33. *A Banach space E is reflexive if and only if B_E is weakly compact.*

Proof. If E is reflexive, then $J(B_E) = B_{E''}$ is weak*-compact (Banach–Alaoglu applied to E'), and since J is a homeomorphism from the weak topology on E to the weak* topology on E'', B_E is weakly compact.

Conversely, if B_E is weakly compact, then $J(B_E)$ is weak*-compact in E'', hence weak*-closed. By Goldstine's theorem, $J(B_E)$ is weak*-dense in $B_{E''}$, so $J(B_E) = B_{E''}$ and J is surjective. $\qquad \square$

4.9 Exercises

Exercise 4.1 ($\star$). Let E be a real vector space and $p \colon E \to \mathbb{R}$ a sublinear functional. Show that there exists a linear functional $f \colon E \to \mathbb{R}$ with $f \leq p$. (*Hint:* apply Theorem 4.2 with $G = \{0\}$.)

Exercise 4.2 ($\star$). Give a proof of the Hahn–Banach theorem in finite dimensions without Zorn's lemma (by induction on codimension).

Exercise 4.3 ($\star\star$). Show that $(c_0)' \cong \ell^1$ isometrically, where c_0 is the space of sequences converging to zero with the sup norm.

Exercise 4.4 ($\star$). Let E be a normed space and $G \subset E$ a subspace. The *annihilator* is $G^\perp = \{f \in E' : f|_G = 0\}$.

(a) Show $G^\perp$ is a closed subspace of E'.

(b) Show $(E/\overline{G})' \cong G^\perp$ isometrically.

(c) Show $(\overline{G})' \cong E'/G^\perp$ isometrically.

Exercise 4.5 ($\star\star$). Let C be a closed convex subset of a normed space E and $x_0 \notin C$. Show that there exists $f \in E'$ such that $\operatorname{Re} f(x_0) > \sup_{x \in C} \operatorname{Re} f(x)$.

Exercise 4.6 ($\star\star$). (a) Show that c_0 is not reflexive by showing that $J(c_0) \neq (\ell^1)' \cong \ell^\infty$.

(b) Show that ℓ^1 is not reflexive.

Exercise 4.7 ($\star\star\star$). Let E be a Banach space. Show that if every $f \in E'$ attains its norm on B_E (i.e., there exists x with $\|x\| = 1$ and $f(x) = \|f\|$), then E is reflexive. (This is James' theorem; proving it is quite hard—outline the argument using the characterisation of reflexivity via weak compactness.)

Exercise 4.8 (⋆⋆). Let $C = \{x \in \mathbb{R}^n : \langle a_i, x \rangle \leq b_i, \ i = 1, \ldots, m\}$ be a convex polytope containing the origin in its interior. Compute the Minkowski functional p_C explicitly.

Exercise 4.9 (⋆⋆). Compute $\partial\varphi(x)$ for $\varphi(x) = \|x\|_1$ on ℓ^1 (the ℓ^1-norm). Characterize the subgradients at the origin and at a point $x = (x_n)$ with all $x_n \neq 0$.

Exercise 4.10 (⋆). Use the Hahn–Banach theorem to show that for any closed subspace G of a Banach space E and any $x \notin G$,

$$\text{dist}(x, G) = \max\{f(x) : f \in G^\perp, \ \|f\| \leq 1\}.$$

Exercise 4.11 (⋆⋆). Let E be a separable normed space. Show that the weak* topology on $B_{E'}$ is metrizable. (*Hint:* if (x_n) is dense in B_E, use $d(f, g) = \sum_{n=1}^\infty 2^{-n} |f(x_n) - g(x_n)|$.)

Exercise 4.12 (⋆⋆⋆). Let E be a normed space and G a subspace. Show that every $g \in G'$ has a *unique* Hahn–Banach extension to E if and only if the dual norm on $E'/G^\perp$ is strictly convex. (A norm is *strictly convex* if $\|x\| = \|y\| = \|(x+y)/2\| = 1$ implies $x = y$.)

Exercise 4.13 (⋆⋆). Identify $(\ell^1)''$ explicitly. Show that the canonical embedding $J \colon \ell^1 \to (\ell^1)''$ corresponds to the natural inclusion $\ell^1 \hookrightarrow \ell^\infty \cong (\ell^1)'$ followed by the canonical embedding. Conclude again that ℓ^1 is not reflexive.

Exercise 4.14 (⋆). Show that if E is reflexive, then E' is also reflexive.

Exercise 4.15 (⋆⋆). Let E be a normed space and $C \subset E$ a convex set. Using the Hahn–Banach theorem, show that the closure of C in the norm topology equals the closure of C in the weak topology. (This is Mazur's theorem.)

Exercise 4.16 ($\star\star$). Let A and B be non-empty disjoint convex subsets of a normed space E, with A open. Show that there exist $f \in E' \setminus \{0\}$ and $\alpha \in \mathbb{R}$ such that $\operatorname{Re} f(a) < \alpha \leq \operatorname{Re} f(b)$ for all $a \in A$ and $b \in B$. Give an example showing that strict separation may fail without compactness of one of the sets.

Exercise 4.17 ($\star\star$). A *Banach limit* is a linear functional $L\colon \ell^\infty \to \mathbb{R}$ satisfying: (i) $L \geq 0$ (i.e., $x_n \geq 0$ for all n implies $L(x) \geq 0$); (ii) $L((x_{n+1})) = L((x_n))$ (shift-invariance); (iii) $L((1,1,1,\dots)) = 1$.

(a) Use the Hahn–Banach theorem to show that Banach limits exist. (*Hint:* on the subspace of convergent sequences, take $L = \lim$, and use the sublinear functional $p(x) = \limsup_{n\to\infty} \frac{1}{n}\sum_{k=1}^n x_k$.)

(b) Show that if (x_n) converges, then $L((x_n)) = \lim_{n\to\infty} x_n$ for every Banach limit L.

Exercise 4.18 ($\star\star\star$). A subset $S \subset E'$ is *total* (or *separating*) if $f(x) = 0$ for all $f \in S$ implies $x = 0$. Show that S is total if and only if $\operatorname{span}(S)$ is weak*-dense in E'. (*Hint:* use the Hahn–Banach theorem in the weak*-topology, identifying $(E', w^*)' \cong E$ via the canonical embedding.)

Exercise 4.19 ($\star$). Let $E = c_0$ and $f = (f_n) \in \ell^1 = (c_0)'$. Show that f attains its norm on B_{c_0} if and only if the supremum $\|f\|_1 = \sum |f_n|$ is attained in a suitable sense. Give an explicit $f \in (c_0)'$ that attains its norm and one that does not (impossible in this case—show all functionals in ℓ^1 attain their norm on B_{c_0}).

Remark 4.34 (Historical note). The Hahn–Banach theorem was proved independently by Hans Hahn (1927) for separable spaces and by Stefan Banach (1929) in full generality. The geometric forms were developed by various authors, including Mazur, Eidelheit, and Dieudonné. The use of Zorn's lemma (or equivalently, the axiom of choice) is essential: there exist models of set theory (without full choice) in which the Hahn–Banach theorem fails. However, weaker forms of choice such as the Boolean prime ideal theorem suffice.

Chapter 5

The Uniform Boundedness Principle

Contents

The *uniform boundedness principle*, also known as the *Banach–Steinhaus theorem*, is one of the cornerstones of functional analysis. It asserts that a family of bounded linear operators that is pointwise bounded must in fact be uniformly bounded in the operator norm. The proof rests on the *Baire category theorem*, a fundamental topological result concerning complete metric spaces. In this chapter we develop the Baire category theorem from scratch,

prove the uniform boundedness principle in its full generality, and present a wealth of applications ranging from Fourier analysis to the theory of divergent series.

5.1   The Baire category theorem

5.1.1   Meagre and residual sets

Definition 5.1 (Nowhere dense set)**.** Let (X, d) be a metric space. A subset $A \subset X$ is called **nowhere dense** if its closure $\overline{A}$ has empty interior:
$$\mathrm{int}(\overline{A}) = \varnothing.$$

Equivalently, A is nowhere dense if and only if for every nonempty open set $U \subset X$, there exists a nonempty open set $V \subset U$ such that $V \cap A = \varnothing$.

Example 5.2.   (i) The set $\mathbb{Z}$ is nowhere dense in $\mathbb{R}$ with the usual metric.

   (ii) Any finite subset of $\mathbb{R}$ is nowhere dense.

   (iii) The Cantor set $\mathcal{C} \subset [0, 1]$ is closed with empty interior, hence nowhere dense.

   (iv) In any normed space X, a proper closed subspace $Y \subsetneq X$ is nowhere dense. Indeed, if $\mathrm{int}(\overline{Y}) = \mathrm{int}(Y) \neq \varnothing$, then Y contains an open ball, and by the linearity of Y this forces $Y = X$.

Definition 5.3 (Meagre and residual sets)**.** Let (X, d) be a metric space.

   (i) A set $M \subset X$ is **meagre** (or **of the first category**) if it can be written as a countable union of nowhere dense sets:
$$M = \bigcup_{n=1}^{\infty} A_n, \qquad \text{each } A_n \text{ nowhere dense.}$$

   (ii) A set $R \subset X$ is **residual** (or **comeagre**) if its complement $X \setminus R$ is meagre.

(iii) A set $S \subset X$ is **of the second category** if it is not meagre.

Remark 5.4. The terminology "first" and "second" category is due to Baire. Meagre sets are "topologically small," analogous to measure-zero sets in measure theory. However, these notions do not coincide: there exist meagre sets of full Lebesgue measure and residual sets of measure zero.

5.1.2   Statement and proof of the Baire category theorem

Theorem 5.5 (Baire category theorem)**.** *Let (X, d) be a complete metric space. Then:*

(a) X is of the second category in itself, i.e., X is not meagre.

(b) Every countable intersection of dense open sets is dense in X.

(c) Equivalently, a countable union of closed sets with empty interior has empty interior.

Proof. We prove statement (b); the equivalence with (a) and (c) follows by taking complements.

Let $\{U_n\}_{n=1}^{\infty}$ be a sequence of dense open subsets of X. We must show that $G = \bigcap_{n=1}^{\infty} U_n$ is dense in X. Let $W \subset X$ be an arbitrary nonempty open set; we need to show $G \cap W \neq \varnothing$.

Step 1. Since U_1 is dense and open, $U_1 \cap W$ is a nonempty open set. Choose $x_1 \in U_1 \cap W$ and $r_1 > 0$ such that

$$\overline{B(x_1, r_1)} \subset U_1 \cap W, \qquad 0 < r_1 < 1.$$

This is possible because $U_1 \cap W$ is open.

Step 2. Since U_2 is dense, $U_2 \cap B(x_1, r_1)$ is nonempty and open. Choose $x_2 \in U_2 \cap B(x_1, r_1)$ and $r_2 > 0$ such that

$$\overline{B(x_2, r_2)} \subset U_2 \cap B(x_1, r_1), \qquad 0 < r_2 < \tfrac{1}{2}.$$

Induction. Proceeding inductively, suppose we have chosen x_n and r_n with $0 < r_n < 1/n$ and

$$\overline{B(x_n, r_n)} \subset U_n \cap B(x_{n-1}, r_{n-1}).$$

Since U_{n+1} is dense, $U_{n+1} \cap B(x_n, r_n)$ is nonempty and open. Choose $x_{n+1} \in U_{n+1} \cap B(x_n, r_n)$ and $0 < r_{n+1} < 1/(n+1)$ with

$$\overline{B(x_{n+1}, r_{n+1})} \subset U_{n+1} \cap B(x_n, r_n).$$

Step 3 (Cauchy sequence). The sequence $(x_n)_{n \geq 1}$ is Cauchy: for $m > n$, we have $x_m \in B(x_n, r_n)$, so $d(x_m, x_n) < r_n < 1/n \to 0$. Since X is complete, $x_n \to x^*$ for some $x^* \in X$.

Step 4 (Limit lies in the intersection). For each fixed n, we have $x_m \in \overline{B(x_n, r_n)}$ for all $m \geq n$, and since $\overline{B(x_n, r_n)}$ is closed, $x^* \in \overline{B(x_n, r_n)} \subset U_n$. Hence $x^* \in U_n$ for every n, so $x^* \in G$. Also, $x^* \in \overline{B(x_1, r_1)} \subset W$, so $x^* \in G \cap W$.

Since W was an arbitrary nonempty open set, G is dense in X.

For (a) $\Leftrightarrow$ (b): suppose $X = \bigcup_n F_n$ with each F_n nowhere dense. Then $U_n = X \setminus \overline{F_n}$ is dense and open, but $\bigcap_n U_n = X \setminus \bigcup_n \overline{F_n} \subset X \setminus \bigcup_n F_n = \varnothing$, contradicting (b). For (b) $\Leftrightarrow$ (c): if $X = \bigcup_n F_n$ with each F_n closed and $\mathrm{int}(F_n) = \varnothing$, then $U_n = X \setminus F_n$ is dense and open and $\bigcap_n U_n = \varnothing$, contradicting (b). □

> *Remark* 5.6. The Baire category theorem also holds for locally compact Hausdorff spaces. The proof is essentially the same, replacing completeness by local compactness: one chooses the closed balls $\overline{B(x_n, r_n)}$ to be compact, and the nested intersection of nonempty compact sets is nonempty by the finite intersection property.

Proposition 5.7. *Let X be a complete metric space.*

(i) *If $X = \bigcup_{n=1}^{\infty} F_n$ with each F_n closed, then at least one F_n has nonempty interior.*

(ii) *Every residual subset of X is dense.*

(iii) *A nonempty complete metric space with no isolated points is uncountable.*

Proof.   (i) This is the contrapositive of Theorem A.1(c).

(ii) If R is residual, then $X \setminus R$ is meagre, so R contains a dense G_δ set by (b) of the theorem.

(iii) If $X = \{x_1, x_2, \dots\}$ were countable, then each singleton $\{x_n\}$ is closed with empty interior (since x_n is not isolated), so $X = \bigcup_n \{x_n\}$ is meagre, contradicting Theorem A.1(a). □

Exercise 5.1 ($\mathbb{Q}$ is not a G_δ; $\star$)**.** Show that $\mathbb{Q}$ is *not* a G_δ set in $\mathbb{R}$; that is, $\mathbb{Q}$ cannot be written as a countable intersection of open sets. *Hint:* If $\mathbb{Q} = \bigcap_n U_n$ with each U_n open and dense, then $V_n = U_n \setminus \{q_n\}$ for an enumeration $\{q_n\}$ of $\mathbb{Q}$ would also be open and dense, but $\bigcap_n V_n \cap \bigcap_n (X \setminus \{q_n\}) = \varnothing$, which contradicts Baire.

Exercise 5.2 (Osgood's theorem; $\star\star$)**.** Let $f_n\colon [a, b] \to \mathbb{R}$ be a sequence of continuous functions converging pointwise to $f\colon [a, b] \to \mathbb{R}$. Use the Baire category theorem to show that f is continuous on a dense G_δ subset of $[a, b]$. *Hint:* For each $k \geq 1$, let $F_{k,m} = \{x \in [a, b] : |f_n(x) - f_p(x)| \leq 1/k \text{ for all } n, p \geq m\}$. Show that $\bigcup_m F_{k,m} = [a, b]$ and apply Baire.

5.2 The uniform boundedness principle

We now turn to the main theorem of this chapter. Throughout, X denotes a Banach space and Y a normed space.

Theorem 5.8 (Banach–Steinhaus / Uniform Boundedness Principle)**.**
*Let X be a Banach space, Y a normed space, and $\{T_\alpha\}_{\alpha \in \mathcal{A}} \subset \mathcal{L}(X, Y)$ a family of bounded linear operators. Suppose the family is **pointwise bounded**: for every $x \in X$,*

$$\sup_{\alpha \in \mathcal{A}} \|T_\alpha x\|_Y < \infty.$$

*Then the family is **uniformly bounded**:*

$$\sup_{\alpha \in \mathcal{A}} \|T_\alpha\|_{\mathcal{L}(X,Y)} < \infty.$$

Proof. For each $n \in \mathbb{N}$, define the set

$$F_n = \Big\{x \in X : \sup_{\alpha \in \mathcal{A}} \|T_\alpha x\|_Y \leq n\Big\} = \bigcap_{\alpha \in \mathcal{A}} \{x \in X : \|T_\alpha x\|_Y \leq n\}.$$

Each set $\{x : \|T_\alpha x\|_Y \leq n\}$ is closed (it is the preimage of $[0, n]$ under the continuous map $x \mapsto \|T_\alpha x\|$), so F_n is an intersection of closed sets, hence closed.

The pointwise boundedness hypothesis gives $X = \bigcup_{n=1}^{\infty} F_n$. Since X is a Banach space (hence a complete metric space), the Baire category theorem (Proposition 5.7(i)) guarantees that some F_{n_0} has nonempty interior. Hence there exist $x_0 \in X$ and $r > 0$ such that $B(x_0, r) \subset F_{n_0}$, meaning

$$\|x - x_0\| < r \implies \sup_{\alpha} \|T_\alpha x\| \le n_0.$$

Now let $x \in X$ with $\|x\| \le 1$. Then $x_0 + rx/2 \in B(x_0, r) \subset F_{n_0}$, and $x_0 \in F_{n_0}$. Therefore,

$$\left\|T_\alpha\left(\tfrac{r}{2}x\right)\right\| = \left\|T_\alpha(x_0 + \tfrac{r}{2}x) - T_\alpha x_0\right\| \le \left\|T_\alpha(x_0 + \tfrac{r}{2}x)\right\| + \left\|T_\alpha x_0\right\| \le n_0 + n_0 = 2n_0.$$

Hence $\|T_\alpha x\| \le 4n_0/r$ for all α and all $\|x\| \le 1$. This gives

$$\sup_{\alpha \in \mathcal{A}} \|T_\alpha\| \le \frac{4n_0}{r} < \infty. \qquad \square$$

> *Remark* 5.9. The completeness of X is essential. Consider $X = c_{00}$, the space of finitely supported sequences with the ℓ^∞ norm. Define $T_n \colon c_{00} \to \mathbb{K}$ by $T_n(x) = n\, x_n$. Then $\sup_n |T_n(x)| < \infty$ for each $x \in c_{00}$ (since x has finite support), but $\|T_n\| = n \to \infty$.

> **Corollary 5.10.** *Let X be a Banach space, Y a normed space, and $(T_n)_{n \ge 1} \subset \mathcal{L}(X, Y)$ a sequence such that the limit*
>
> $$Tx = \lim_{n \to \infty} T_n x$$
>
> *exists for every $x \in X$. Then:*
>
> *(i)* $\sup_n \|T_n\| < \infty$.
>
> *(ii) The limit operator $T \colon X \to Y$ is bounded, and $\|T\| \le \liminf_{n \to \infty} \|T_n\|$.*

Proof. (i) A convergent sequence in Y is bounded, so $\sup_n \|T_n x\| < \infty$ for each x. The Banach–Steinhaus theorem gives $\sup_n \|T_n\| = M < \infty$.

(ii) T is linear (being the pointwise limit of linear maps). For any $x \in X$,

$$\|Tx\| = \lim_{n \to \infty} \|T_n x\| \le \liminf_{n \to \infty} \|T_n\|\, \|x\|,$$

so T is bounded with $\|T\| \le \liminf_n \|T_n\|$. $\qquad \square$

5.3   Applications to sequences of functionals

Proposition 5.11 (Pointwise bounded functionals). *Let X be a Banach space and $(f_n)_{n \geq 1} \subset X^*$ a sequence of bounded linear functionals. If $\sup_n |f_n(x)| < \infty$ for every $x \in X$, then $\sup_n \|f_n\|_{X^*} < \infty$.*

Proof. This is the special case $Y = \mathbb{K}$ of Theorem 5.8. $\qquad\square$

Example 5.12 (Weak convergence implies boundedness). Let X be a Banach space and suppose $x_n \rightharpoonup x$ in X (i.e., $f(x_n) \to f(x)$ for all $f \in X^*$). Then $\sup_n \|x_n\| < \infty$.

To see this, consider the canonical images $\hat{x}_n \in X^{**}$ defined by $\hat{x}_n(f) = f(x_n)$. The hypothesis says that for each $f \in X^*$, the sequence $(\hat{x}_n(f))_n$ is convergent, hence bounded. Since X^* is a Banach space, the uniform boundedness principle applied to $\{\hat{x}_n\} \subset \mathcal{L}(X^*, \mathbb{K})$ gives

$$\sup_n \|\hat{x}_n\|_{X^{**}} = \sup_n \|x_n\|_X < \infty,$$

where the equality uses the isometric embedding $X \hookrightarrow X^{**}$.

Proposition 5.13 (Weak* convergence implies boundedness). *Let X be a Banach space and suppose $f_n \xrightarrow{*} f$ in X^* (i.e., $f_n(x) \to f(x)$ for all $x \in X$). Then $\sup_n \|f_n\|_{X^*} < \infty$ and $\|f\| \leq \liminf_n \|f_n\|$.*

Proof. Apply Corollary 5.10 with $T_n = f_n$ and $Y = \mathbb{K}$. $\qquad\square$

5.4 Strong convergence versus uniform convergence

Definition 5.14. Let X, Y be normed spaces and $(T_n)_{n \geq 1} \subset \mathcal{L}(X, Y)$.

(i) (T_n) converges **strongly** (or **pointwise**) to $T \in \mathcal{L}(X, Y)$ if $T_n x \to Tx$ for every $x \in X$. We write $T_n \xrightarrow{s} T$.

(ii) (T_n) converges **uniformly** (or **in norm**) to T if $\|T_n - T\| \to 0$. We write $T_n \to T$.

Remark 5.15. Uniform convergence implies strong convergence, since $\|T_n x - Tx\| \leq \|T_n - T\| \|x\|$. The converse is false in general.

Example 5.16. On $X = \ell^2(\mathbb{N})$, let P_n be the orthogonal projection onto $\operatorname{span}\{e_1, \ldots, e_n\}$. Then $P_n \xrightarrow{s} \operatorname{Id}$ (since every $x \in \ell^2$ is the sum of its Fourier series), but $\|P_n - \operatorname{Id}\| = 1$ for all n, so the convergence is not uniform.

Theorem 5.17 (Banach–Steinhaus for strong convergence)**.** *Let X be a Banach space, Y a normed space, and $(T_n) \subset \mathcal{L}(X, Y)$ a sequence converging strongly to T. Then:*

(i) $\sup_n \|T_n\| < \infty$ and $T \in \mathcal{L}(X, Y)$.

(ii) If $x_n \to x$ in X, then $T_n x_n \to Tx$ in Y.

Proof. (i) follows from Corollary 5.10. Let $M = \sup_n \|T_n\|$.
    (ii) We write

$$T_n x_n - Tx = T_n(x_n - x) + (T_n x - Tx).$$

For the first term: $\|T_n(x_n - x)\| \leq M \|x_n - x\| \to 0$. For the second term: $T_n x \to Tx$ by hypothesis. Hence $T_n x_n \to Tx$. $\qquad\square$

Exercise 5.3 (Strong convergence and equicontinuity; $\star$)**.** Let X be a Banach space, Y a normed space, and $(T_n) \subset \mathcal{L}(X, Y)$ with $\sup_n \|T_n\| \leq M$. Show that $T_n \xrightarrow{s} T$ if and only if $T_n x \to Tx$ for all x in a dense subset of X.

5.5   Application: divergent Fourier series

One of the most striking applications of the uniform boundedness principle is the proof that there exist continuous functions whose Fourier series diverges at a given point. This result, due to du Bois-Reymond (1876), was one of the motivations for the development of the Baire category approach.

5.5.1 The Dirichlet kernel and partial sums

Let $f \colon [-\pi, \pi] \to \mathbb{C}$ be a 2π-periodic integrable function. The N-th partial sum of its Fourier series at the point $t = 0$ is

$$S_N f(0) = \sum_{k=-N}^{N} \hat{f}(k), \qquad \hat{f}(k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \, e^{-ikt} \, dt.$$

One computes

$$S_N f(0) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \, D_N(t) \, dt,$$

where D_N is the **Dirichlet kernel**:

$$D_N(t) = \sum_{k=-N}^{N} e^{ikt} = \frac{\sin\left(\left(N + \frac{1}{2}\right)t\right)}{\sin(t/2)}.$$

Lemma 5.18. *The L^1 norm of the Dirichlet kernel satisfies*

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} |D_N(t)| \, dt = \frac{4}{\pi^2} \ln N + O(1) \quad \text{as } N \to \infty.$$

In particular, $\|D_N\|_{L^1} \to \infty$.

Proof. We have

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} |D_N(t)| \, dt = \frac{1}{\pi} \int_{0}^{\pi} \frac{\left|\sin\left(\left(N + \frac{1}{2}\right)t\right)\right|}{\sin(t/2)} \, dt$$

$$\geq \frac{1}{\pi} \int_{0}^{\pi} \frac{\left|\sin\left(\left(N + \frac{1}{2}\right)t\right)\right|}{t/2} \, dt \quad (\text{since } \sin(t/2) \leq t/2)$$

$$= \frac{2}{\pi} \int_{0}^{(N+1/2)\pi} \frac{|\sin u|}{u} \, du \quad (u = (N + \tfrac{1}{2})t)$$

$$\geq \frac{2}{\pi} \sum_{k=1}^{N} \int_{(k-1)\pi}^{k\pi} \frac{|\sin u|}{u} \, du \geq \frac{2}{\pi} \sum_{k=1}^{N} \frac{1}{k\pi} \int_{0}^{\pi} \sin u \, du = \frac{4}{\pi^2} \sum_{k=1}^{N} \frac{1}{k}.$$

Since $\sum_{k=1}^{N} 1/k = \ln N + O(1)$, we get the lower bound. A similar calculation gives a matching upper bound. $\qquad \square$

5.5.2   The du Bois-Reymond theorem

Consider the Banach space $X = C([-\pi, \pi])$ of continuous 2π-periodic functions with the supremum norm. For each N, the N-th partial sum operator at $t = 0$ defines a bounded linear functional:

$$\Lambda_N \colon C([-\pi, \pi]) \to \mathbb{C}, \qquad \Lambda_N(f) = S_N f(0) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \, D_N(t) \, dt.$$

Its operator norm is

$$\|\Lambda_N\| = \frac{1}{2\pi} \int_{-\pi}^{\pi} |D_N(t)| \, dt = \frac{4}{\pi^2} \ln N + O(1) \to \infty.$$

> **Theorem 5.19** (du Bois-Reymond)**.** *There exists a continuous 2π-periodic function f whose Fourier series diverges at $t = 0$, i.e., $\sup_N |S_N f(0)| = \infty$.*
> *More precisely, the set of functions $f \in C([-\pi, \pi])$ for which $\sup_N |S_N f(0)| = \infty$ is a dense G_δ subset of $C([-\pi, \pi])$; in particular, it is residual.*

Proof. We apply the *contrapositive* of the Banach–Steinhaus theorem (Theorem 5.8). Consider the functionals $\Lambda_N \in C([-\pi, \pi])^*$ defined above. We have shown that $\|\Lambda_N\| \to \infty$.

If $\sup_N |\Lambda_N(f)| < \infty$ were to hold for *every* $f \in C([-\pi, \pi])$, then the uniform boundedness principle would give $\sup_N \|\Lambda_N\| < \infty$, a contradiction. Hence there exists $f \in C([-\pi, \pi])$ with $\sup_N |S_N f(0)| = \infty$.

For the second statement, define $G_k = \{f \in C([-\pi, \pi]) : \sup_N |S_N f(0)| > k\}$. Each G_k is open (as a union of the open sets $\{|\Lambda_N(f)| > k\}$ over N), and we need to show each G_k is dense. If G_k were not dense for some k, there would exist a ball $\overline{B(g, \varepsilon)}$ with $\sup_N |\Lambda_N(f)| \le k$ for all $f \in \overline{B(g, \varepsilon)}$. Then for $\|h\| \le 1$:

$$|\Lambda_N(\varepsilon h)| \le |\Lambda_N(g + \varepsilon h)| + |\Lambda_N(g)| \le 2k,$$

giving $\|\Lambda_N\| \le 2k/\varepsilon$ for all N, contradicting $\|\Lambda_N\| \to \infty$. Hence each G_k is open and dense, and $\bigcap_k G_k = \{f : \sup_N |S_N f(0)| = \infty\}$ is a dense G_δ by the Baire category theorem. $\qquad\square$

> *Remark* 5.20. The theorem shows that "most" continuous functions (in the Baire category sense) have divergent Fourier series at any given point. This stands in sharp contrast to Carleson's celebrated theorem (1966) that the Fourier series of any L^2 function converges almost

everywhere.

5.6   The Gibbs phenomenon

The Gibbs phenomenon describes the overshoot of partial Fourier sums near a jump discontinuity. While not directly a consequence of the uniform boundedness principle, it illustrates the behavior of the Dirichlet kernel and complements the discussion of Fourier series.

Example 5.21 (Gibbs phenomenon for the square wave). Consider the 2π-periodic function

$$f(t) = \begin{cases} 1 & \text{if } 0 < t < \pi, \\ -1 & \text{if } -\pi < t < 0, \\ 0 & \text{if } t \in \{0, \pm\pi\}. \end{cases}$$

Its Fourier series is $f(t) \sim \frac{4}{\pi} \sum_{k=0}^{\infty} \frac{\sin(2k+1)t}{2k+1}$. The N-th partial sum at $t = \pi/(2N+1)$ satisfies

$$S_N f\left(\frac{\pi}{2N+1}\right) \to \frac{2}{\pi} \int_0^\pi \frac{\sin u}{u}\, du = \frac{2}{\pi} \operatorname{Si}(\pi) \approx 1.17898,$$

where Si is the sine integral. This represents an overshoot of approximately 8.95% above the function value 1, and this overshoot does not diminish as $N \to \infty$ — it merely moves closer to the discontinuity.

Proposition 5.22. *Let f be a piecewise smooth 2π-periodic function with a jump discontinuity at t_0. Then the partial sums $S_N f$ exhibit an overshoot near t_0 whose magnitude approaches*

$$\frac{f(t_0^+) - f(t_0^-)}{2} \cdot \left(\frac{2}{\pi} \int_0^\pi \frac{\sin u}{u}\, du - 1\right) \approx 0.0895 \cdot \left(f(t_0^+) - f(t_0^-)\right).$$

5.7 The Banach–Steinhaus condensation theorem

The following refinement of the uniform boundedness principle describes the structure of the "exceptional set" where pointwise boundedness fails.

Theorem 5.23 (Condensation of singularities / Banach–Steinhaus). *Let X be a Banach space, Y a normed space, and let $(T_{n,m})_{n,m\geq 1} \subset \mathcal{L}(X,Y)$ be a double sequence of bounded linear operators. Suppose that for each $n \geq 1$,*

$$\sup_{m\geq 1} \|T_{n,m}\| = \infty.$$

Then the set

$$\mathcal{R} = \big\{x \in X : \sup_{m\geq 1} \|T_{n,m}x\| = \infty \text{ for all } n \geq 1\big\}$$

is a dense G_δ subset of X (hence residual).

Proof. For each $n, k \geq 1$, define

$$G_{n,k} = \big\{x \in X : \exists\, m \text{ with } \|T_{n,m}x\| > k\big\} = \bigcup_{m=1}^{\infty} \big\{x \in X : \|T_{n,m}x\| > k\big\}.$$

Each set $\{x : \|T_{n,m}x\| > k\}$ is open (preimage of (k,∞) under a continuous map), so $G_{n,k}$ is open.

Claim: $G_{n,k}$ is dense for each n, k.

Suppose not: there exist $n_0, k_0, x_0 \in X$, and $r > 0$ with $B(x_0,r) \cap G_{n_0,k_0} = \varnothing$. This means $\|T_{n_0,m}x\| \leq k_0$ for all m and all $x \in B(x_0,r)$. For $\|h\| < r$:

$$\|T_{n_0,m}h\| \leq \|T_{n_0,m}(x_0 + h)\| + \|T_{n_0,m}x_0\| \leq 2k_0,$$

so $\|T_{n_0,m}\| \leq 2k_0/r$ for all m, contradicting $\sup_m \|T_{n_0,m}\| = \infty$.

Now observe that

$$\mathcal{R} = \bigcap_{n=1}^{\infty} \bigcap_{k=1}^{\infty} G_{n,k}.$$

This is a countable intersection of dense open sets. By the Baire category theorem (Theorem A.1), $\mathcal{R}$ is a dense G_δ. $\qquad\square$

> **Corollary 5.24.** *There exists a continuous 2π-periodic function whose Fourier series diverges at every rational multiple of π. More generally, given any countable set $S = \{t_n\}_{n\geq 1} \subset [-\pi, \pi]$, the set of functions $f \in C([-\pi, \pi])$ whose Fourier series diverges simultaneously at every $t_n \in S$ is residual in $C([-\pi, \pi])$.*

Proof. For each n, let $T_{n,m}(f) = S_m f(t_n)$ be the m-th partial Fourier sum at t_n. Then $\|T_{n,m}\| = \frac{1}{2\pi} \int_{-\pi}^{\pi} |D_m(t)|\, dt \to \infty$ as $m \to \infty$. The condensation theorem (Theorem 5.23) gives the result. $\qquad\square$

5.8 Uniform boundedness for families of operators

The uniform boundedness principle extends naturally to arbitrary (not necessarily countable) families of operators.

> **Theorem 5.25** (Uniform boundedness for operator families). *Let X be a Banach space, Y a Banach space, and let $\{T_\alpha\}_{\alpha \in \mathcal{A}} \subset \mathcal{L}(X, Y)$ and $\{S_\beta\}_{\beta \in \mathcal{B}} \subset \mathcal{L}(Y, \mathbb{K})$ be families of operators such that*
>
> $$\sup_{\alpha \in \mathcal{A},\ \beta \in \mathcal{B}} |S_\beta(T_\alpha x)| < \infty \qquad \text{for every } x \in X.$$
>
> *Then $\sup_{\alpha,\beta} \|S_\beta \circ T_\alpha\| < \infty$.*

Proof. For each fixed $x \in X$, the family $\{S_\beta \circ T_\alpha\}$ is pointwise bounded on X. First, for each α, the family $\{S_\beta(T_\alpha x)\}_\beta$ is bounded. Since Y is a Banach space, the Banach–Steinhaus theorem applied to $\{S_\beta\}_\beta \subset Y^*$ gives $C_\alpha := \sup_\beta \|S_\beta(T_\alpha x)\| < \infty$ for each x, and moreover $\sup_\beta \|S_\beta\| < \infty$ (call this bound K).

Now $\sup_\alpha \|T_\alpha x\| \leq \sup_\alpha \sup_\beta |S_\beta(T_\alpha x)| / \inf_{\|S_\beta\| \neq 0} \cdots$ — but more directly, since for each x, $\sup_\alpha \sup_\beta |S_\beta(T_\alpha x)| < \infty$, we have in particular $\sup_\alpha \|T_\alpha x\|_Y < \infty$ (by Hahn–Banach applied in Y). The Banach–Steinhaus theorem then gives $\sup_\alpha \|T_\alpha\| < \infty$, say $\leq M$. Hence $\|S_\beta \circ T_\alpha\| \leq \|S_\beta\| \|T_\alpha\| \leq KM$. $\qquad\square$

> **Proposition 5.26** (Principle of condensation for bilinear forms). *Let X, Y be Banach spaces and $B\colon X \times Y \to \mathbb{K}$ a separately continuous*

bilinear form. Then B is jointly bounded:

$$|B(x,y)| \leq M \, \|x\| \, \|y\| \qquad \text{for some } M > 0 \text{ and all } x \in X, \ y \in Y.$$

Proof. For each $x \in X$, the map $y \mapsto B(x,y)$ is a bounded linear functional on Y, so there exists $T_x \in Y^*$ with $T_x(y) = B(x,y)$ and $\|T_x\| = \sup_{\|y\| \leq 1} |B(x,y)|$. The map $x \mapsto T_x$ is a linear operator $T \colon X \to Y^*$.

For each y, the map $x \mapsto B(x,y)$ is bounded, so $|B(x,y)| \leq C_y \|x\|$. In particular, $\sup_{\|x\| \leq 1} |T_x(y)| \leq C_y < \infty$. Since Y is a Banach space, the uniform boundedness principle (applied to $\{T_x\}_{\|x\| \leq 1} \subset Y^*$) gives that $x \mapsto T_x$ maps the unit ball of X to a bounded subset of Y^*. But this means T is bounded, say $\|T\| \leq M$, and

$$|B(x,y)| = |T_x(y)| \leq \|T_x\| \, \|y\| \leq M \, \|x\| \, \|y\| . \qquad \square$$

5.9   Exercises for Chapter 5

Exercise 5.4 (Nowhere dense subsets; $\star$). Show that a finite union of nowhere dense sets is nowhere dense. Give an example showing that a countable union of nowhere dense sets need not be nowhere dense.

Exercise 5.5 (Baire category in ℓ^p; $\star$). Let $1 \leq p < \infty$. Show that $\ell^q \subset \ell^p$ is meagre in ℓ^p for $q < p$. *Hint:* Write $\ell^q = \bigcup_n \{x \in \ell^p : \|x\|_q \leq n\}$.

Exercise 5.6 (Pointwise but not uniform convergence; $\star$). Let $T_n \colon \ell^2 \to \ell^2$ be the right shift applied n times: $T_n(x_1, x_2, \ldots) = (0, \ldots, 0, x_1, x_2, \ldots)$ with n leading zeros. Show that $T_n \xrightarrow{s} 0$ but $\|T_n\| = 1$ for all n. Reconcile this with the Banach–Steinhaus theorem.

Exercise 5.7 (Divergence on a dense set; $\star\star$). Let X be a Banach space and $(T_n) \subset X^*$ with $\sup_n \|T_n\| = \infty$. Show that the set $\{x \in X : \sup_n |T_n(x)| = \infty\}$ is dense in X. Can you show it is residual?

Exercise 5.8 (Unbounded Toeplitz matrix; ⋆⋆). A matrix $A = (a_{nk})_{n,k \geq 0}$ defines a *summability method*: given a sequence (s_k), we set $\sigma_n = \sum_k a_{nk} s_k$ (when convergent). Show that if the Toeplitz conditions hold ($\lim_n a_{nk} = 0$ for each k, $\sup_n \sum_k |a_{nk}| < \infty$, $\lim_n \sum_k a_{nk} = 1$), then $\sigma_n \to s$ whenever $s_k \to s$. *Hint:* Apply the Banach–Steinhaus theorem to functionals on c (the space of convergent sequences).

Exercise 5.9 (Principle of condensation application; ⋆⋆). Let $(a_n)_{n \geq 0}$ be a sequence of complex numbers such that $\sum_n a_n z^n$ has radius of convergence 1. Use the condensation theorem to show that the set of points $e^{i\theta}$ on the unit circle where $\sum_n a_n e^{in\theta}$ diverges is either empty or residual in the unit circle.

Exercise 5.10 (Continuous nowhere differentiable functions; ⋆ ⋆ ⋆). Use the Baire category theorem to show that the set of continuous functions $f \in C([0,1])$ that are nowhere differentiable is residual in $C([0,1])$.
Hint: For each $n \geq 1$, let A_n be the set of $f \in C([0,1])$ such that there exists $x \in [0,1]$ with $|f(x+h) - f(x)| \leq n\,|h|$ for all sufficiently small h. Show that each A_n is closed and has empty interior.

Exercise 5.11 (Banach–Steinhaus for Hilbert spaces; ⋆). Let H be a Hilbert space and $(T_n) \subset \mathcal{L}(H)$ a sequence of self-adjoint operators with $\sup_n |\langle T_n x, x \rangle| < \infty$ for every $x \in H$. Show that $\sup_n \|T_n\| < \infty$. *Hint:* Use the polarization identity and the fact that $\|T\| = \sup_{\|x\|=1} |\langle Tx, x \rangle|$ for self-adjoint T.

Exercise 5.12 (Weak* sequential completeness; ⋆⋆). Let X be a Banach space and $(f_n) \subset X^*$ a weak* Cauchy sequence (i.e., $(f_n(x))$ is Cauchy in $\mathbb{K}$ for each $x \in X$). Show that there exists $f \in X^*$ such that $f_n \overset{*}{\rightharpoonup} f$. Is the analogous statement true for weak Cauchy sequences in X?

Exercise 5.13 (Superdense orbits; ⋆⋆⋆). Let X be a separable Banach space and $T \in \mathcal{L}(X)$ a *hypercyclic* operator (there exists $x \in X$ such that $\{T^n x : n \geq 0\}$ is dense in X). Use the Baire category theorem

to show that the set of hypercyclic vectors for T is either empty or a dense G_δ.

Exercise 5.14 (Uniform boundedness and integration; $\star\star$)**.** Let (f_n) be a sequence in $L^1([0,1])$ such that for every $g \in L^\infty([0,1])$, the sequence $\left(\int_0^1 f_n g \, dx \right)$ is bounded. Show that $\sup_n \|f_n\|_{L^1} < \infty$.

Chapter 6

The Open Mapping Theorem and Closed Graph Theorem

Contents

This chapter is devoted to two further pillars of functional analysis: the *open mapping theorem* (also called the *Banach–Schauder theorem*) and the

closed graph theorem. Like the uniform boundedness principle, both rest upon the Baire category theorem and require completeness in an essential way. Together with the Hahn–Banach theorem and the uniform boundedness principle, these form the "big four" theorems that underpin much of linear functional analysis.

6.1 The open mapping theorem

6.1.1 Preliminary: the open mapping lemma

> **Definition 6.1.** A map $T\colon X \to Y$ between topological spaces is called **open** if it maps open sets to open sets: $U \subset X$ open implies $T(U) \subset Y$ open.

The key technical step is the following lemma, which uses the Baire category theorem.

> **Lemma 6.2** (Open mapping lemma)**.** *Let X and Y be Banach spaces and $T \in \mathcal{L}(X, Y)$ a surjective bounded linear operator. Then there exists $\delta > 0$ such that*
>
> $$B_Y(0, \delta) \subset \overline{T(B_X(0, 1))}.$$
>
> *That is, the closure of the image of the open unit ball of X contains an open ball in Y.*

Proof. Since T is surjective, $Y = T(X) = \bigcup_{n=1}^{\infty} T(B_X(0, n)) = \bigcup_{n=1}^{\infty} n\, T(B_X(0, 1))$. Taking closures, $Y = \bigcup_{n=1}^{\infty} n\, \overline{T(B_X(0, 1))}$.

Since Y is a Banach space (hence a complete metric space), the Baire category theorem (Proposition 5.7(i)) implies that some $n_0\, \overline{T(B_X(0, 1))}$ has nonempty interior. Since multiplication by n_0 is a homeomorphism, $\overline{T(B_X(0, 1))}$ itself has nonempty interior. Hence there exist $y_0 \in Y$ and $r > 0$ with $B_Y(y_0, r) \subset \overline{T(B_X(0, 1))}$.

Since $\overline{T(B_X(0, 1))}$ is symmetric (if $y \in \overline{T(B_X(0, 1))}$ then $-y \in \overline{T(B_X(0, 1))}$, because $B_X(0, 1)$ is symmetric) and convex (since $T(B_X(0, 1))$ is convex), we can average:

$$B_Y(0, r) = \tfrac{1}{2}\big(B_Y(y_0, r) + B_Y(-y_0, r)\big) \subset \tfrac{1}{2}\big(\overline{T(B_X(0, 1))} + \overline{T(B_X(0, 1))}\big) \subset \overline{T(B_X(0, 1))}.$$

The last inclusion uses the convexity: if $u, v \in \overline{T(B_X(0, 1))}$, then $\frac{u+v}{2} \in \overline{T(B_X(0, 1))}$ since $B_X(0, 1)$ is convex and T is linear. Set $\delta = r$. $\qquad\square$

6.1.2 The main theorem

> **Theorem 6.3** (Open Mapping Theorem / Banach–Schauder). *Let X and Y be Banach spaces and $T \in \mathcal{L}(X, Y)$ surjective. Then T is an open mapping. More precisely, there exists $\delta > 0$ such that*
>
> $$B_Y(0, \delta) \subset T(B_X(0, 1)).$$

Proof. **Step 1 (Closure contains a ball).** By Lemma 6.2, there exists $\delta > 0$ such that $B_Y(0, \delta) \subset \overline{T(B_X(0, 1))}$. By scaling, for every $\varepsilon > 0$,

$$B_Y(0, \varepsilon\delta) \subset \overline{T(B_X(0, \varepsilon))}. \tag{6.1}$$

Step 2 (Image itself contains a ball — the approximation argument). We show that $B_Y(0, \delta) \subset T(B_X(0, 2))$. Let $y \in B_Y(0, \delta)$. We construct a sequence (x_n) in X such that

$$\left\| y - \sum_{k=0}^{n} T x_k \right\| < \frac{\delta}{2^{n+1}}, \qquad \|x_n\| < \frac{1}{2^n}.$$

Base case: By (6.1) with $\varepsilon = 1$, the element $y \in B_Y(0, \delta) \subset \overline{T(B_X(0, 1))}$, so there exists $x_0 \in X$ with $\|x_0\| < 1$ and $\|y - T x_0\| < \delta/2$.

Inductive step: Having chosen $x_0, \ldots, x_n$, the residual $y - \sum_{k=0}^{n} T x_k \in B_Y(0, \delta/2^{n+1})$. By (6.1) with $\varepsilon = 1/2^{n+1}$, there exists x_{n+1} with $\|x_{n+1}\| < 1/2^{n+1}$ and $\left\| y - \sum_{k=0}^{n+1} T x_k \right\| < \delta/2^{n+2}$.

Step 3 (Convergence). Set $s_n = \sum_{k=0}^{n} x_k$. Since $\sum_{k=0}^{\infty} \|x_k\| < \sum_{k=0}^{\infty} 2^{-k} = 2$, the series converges absolutely, so (s_n) is Cauchy in the Banach space X. Let $x = \sum_{k=0}^{\infty} x_k$; then $\|x\| < 2$. By continuity of T,

$$Tx = T\left(\sum_{k=0}^{\infty} x_k \right) = \lim_{n \to \infty} \sum_{k=0}^{n} T x_k = y.$$

Hence $y = Tx$ with $x \in B_X(0, 2)$, proving $B_Y(0, \delta) \subset T(B_X(0, 2))$.

Replacing by $B_X(0, 1)$ via scaling: $B_Y(0, \delta/2) \subset T(B_X(0, 1))$.

Step 4 (Openness). Let $U \subset X$ be open and $y_0 = T x_0 \in T(U)$. Choose $r > 0$ with $B_X(x_0, r) \subset U$. Then

$$T(U) \supset T(B_X(x_0, r)) = T x_0 + T(B_X(0, r)) \supset y_0 + B_Y(0, r\delta/2),$$

so $T(U)$ contains an open ball around each of its points, hence $T(U)$ is open. $\qquad \square$

Remark 6.4. The completeness of *both* X and Y is used: completeness of Y in the Baire category argument (Step 1), and completeness of X for the convergence of the series in Step 3. Without completeness, the theorem can fail: consider the identity map $\mathrm{Id}\colon (C([0,1]), \|\cdot\|_{C^1}) \to (C([0,1]), \|\cdot\|_\infty)$, which is continuous and bijective but not open.

6.2 The Banach isomorphism theorem

Corollary 6.5 (Banach isomorphism theorem). *Let X and Y be Banach spaces and $T \in \mathcal{L}(X,Y)$ a continuous linear bijection. Then $T^{-1} \in \mathcal{L}(Y,X)$; that is, T^{-1} is automatically continuous. Hence T is a topological isomorphism.*

Proof. Since T is a bijection, $T^{-1}\colon Y \to X$ is a well-defined linear map. By the open mapping theorem (Theorem 6.3), T is open, so for any open $U \subset X$, $T(U)$ is open in Y. But $(T^{-1})^{-1}(U) = T(U)$, so T^{-1} is continuous. $\qquad\square$

Example 6.6. Let X be a vector space equipped with two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ such that $(X, \|\cdot\|_1)$ and $(X, \|\cdot\|_2)$ are both Banach spaces. If there exists $C > 0$ with $\|x\|_2 \leq C \|x\|_1$ for all x, then the norms are equivalent, i.e., there exists $c > 0$ with $c\|x\|_1 \leq \|x\|_2$.
Indeed, the identity $\mathrm{Id}\colon (X, \|\cdot\|_1) \to (X, \|\cdot\|_2)$ is continuous and bijective. By the Banach isomorphism theorem, Id^{-1} is continuous, giving the reverse inequality.

Example 6.7. Consider $\ell^1(\mathbb{N})$ with its standard norm and the map $T\colon \ell^1 \to c_0$ defined by $T(x)_n = \sum_{k=n}^\infty x_k$ (partial tail sums). One can verify that T is a bounded linear bijection. The Banach isomorphism theorem guarantees that T^{-1} is bounded, even though an explicit formula for T^{-1} may be complicated: $(T^{-1}y)_n = y_n - y_{n+1}$.

Proposition 6.8 (Bounded below implies open range). *Let X be a Banach space, Y a normed space, and $T \in \mathcal{L}(X,Y)$. The following are equivalent:*

(i) *T is bounded below: there exists $c > 0$ with $\|Tx\| \geq c\|x\|$ for all $x \in X$.*

(ii) T is injective and $T(X)$ is closed in Y, and $T^{-1}\colon T(X) \to X$ is bounded.

If in addition Y is a Banach space and T is surjective, then (i) and (ii) are equivalent to T being a topological isomorphism.

Proof. (i) $\Rightarrow$ (ii): Boundedness below gives injectivity ($Tx = 0 \Rightarrow x = 0$) and $\|T^{-1}\| \leq 1/c$. To see $T(X)$ is closed, let $Tx_n \to y$. Then (Tx_n) is Cauchy, so $\|x_n - x_m\| \leq c^{-1}\|Tx_n - Tx_m\| \to 0$, hence (x_n) is Cauchy. Since X is complete, $x_n \to x$ and $y = Tx \in T(X)$.

(ii) $\Rightarrow$ (i): $T(X)$ with the norm inherited from Y is a Banach space (closed subspace of a — possibly incomplete — normed space, but $T(X) \cong X/\operatorname{Ker} T = X$ since T is injective). The boundedness of T^{-1} gives $\|x\| = \|T^{-1}(Tx)\| \leq \|T^{-1}\|\,\|Tx\|$, so $\|Tx\| \geq c\,\|x\|$ with $c = 1/\|T^{-1}\|$. $\qquad\square$

6.3   The closed graph theorem

6.3.1   The graph of a linear operator

Definition 6.9. Let X and Y be normed spaces and $T\colon D(T) \to Y$ a (possibly unbounded) linear operator defined on a subspace $D(T) \subset X$ (the **domain** of T). The **graph** of T is

$$\Gamma(T) = \{(x, Tx) : x \in D(T)\} \subset X \times Y.$$

We equip $X \times Y$ with the product norm $\|(x, y)\| = \|x\|_X + \|y\|_Y$ (or equivalently $\max\{\|x\|, \|y\|\}$, as these are equivalent).

Definition 6.10 (Closed operator)**.** An operator $T\colon D(T) \to Y$ is **closed** if its graph $\Gamma(T)$ is a closed subset of $X \times Y$. Equivalently, T is closed if and only if: whenever $(x_n) \subset D(T)$ with $x_n \to x$ in X and $Tx_n \to y$ in Y, then $x \in D(T)$ and $Tx = y$.

Remark 6.11. A continuous operator with closed domain is always closed (if $x_n \to x$ and $D(T)$ is closed, then $x \in D(T)$ and $Tx_n \to Tx$ by continuity). However, a closed operator need not be continuous — this is the key distinction when dealing with unbounded operators.

Example 6.12. On $X = C([0,1])$ with the supremum norm, consider the operator $T = d/dx$ with domain $D(T) = C^1([0,1])$. Then T is closed: if $f_n \to f$ uniformly and $f_n' \to g$ uniformly, then f is differentiable and $f' = g$. However, T is not bounded on $(D(T), \|\cdot\|_\infty)$.

6.3.2   Statement and proof of the closed graph theorem

Theorem 6.13 (Closed Graph Theorem). *Let X and Y be Banach spaces and $T\colon X \to Y$ a linear operator (defined on all of X). If T is closed (i.e., $\Gamma(T)$ is closed in $X \times Y$), then T is bounded.*

Proof. Since X and Y are Banach spaces, the product $X \times Y$ (with the norm $\|(x,y)\| = \|x\| + \|y\|$) is a Banach space. The graph $\Gamma(T)$ is a closed linear subspace of $X \times Y$, hence itself a Banach space.

Consider the projections:

$$\pi_1 \colon \Gamma(T) \to X, \quad \pi_1(x, Tx) = x,$$
$$\pi_2 \colon \Gamma(T) \to Y, \quad \pi_2(x, Tx) = Tx.$$

Both are bounded: $\|\pi_1(x, Tx)\| = \|x\| \le \|x\| + \|Tx\|$ and similarly for π_2.

The projection π_1 is a bounded linear bijection from the Banach space $\Gamma(T)$ to the Banach space X. By the Banach isomorphism theorem (Corollary 6.5), $\pi_1^{-1} \colon X \to \Gamma(T)$ is bounded. Therefore $T = \pi_2 \circ \pi_1^{-1}$ is bounded, being the composition of two bounded operators. $\square$

Remark 6.14. The proof of the closed graph theorem via the Banach isomorphism theorem is remarkably elegant: the entire argument reduces to the observation that two natural projections and one inverse are bounded.

Theorem 6.15 (Closed graph theorem — equivalent formulation). *Let X and Y be Banach spaces and $T\colon X \to Y$ a linear operator. The following are equivalent:*

(i) *T is bounded (i.e., continuous).*

(ii) *T is closed.*

(iii) *For every sequence (x_n) in X: if $x_n \to 0$ and $Tx_n \to y$, then $y = 0$.*

Proof. (i) $\Rightarrow$ (ii): If T is continuous and $x_n \to x$, $Tx_n \to y$, then $Tx_n \to Tx$ by continuity, so $y = Tx$ and $\Gamma(T)$ is closed.

(ii) $\Rightarrow$ (iii): Immediate from the definition, with $x = 0$.

(iii) $\Rightarrow$ (ii): Suppose $x_n \to x$ and $Tx_n \to y$. Set $z_n = x_n - x$; then $z_n \to 0$ and $Tz_n = Tx_n - Tx \to y - Tx$. By (iii), $y - Tx = 0$, so $y = Tx$.

(ii) $\Rightarrow$ (i): This is Theorem 6.13. $\qquad\square$

> **Example 6.16** (Automatic continuity). Let H be a Hilbert space and $T\colon H \to H$ a linear operator satisfying $\langle Tx, y \rangle = \langle x, Ty \rangle$ for all $x, y \in H$ (i.e., T is formally symmetric and defined on all of H). Then T is bounded.
>
> *Proof:* We verify condition (iii). Suppose $x_n \to 0$ and $Tx_n \to y$. For any $z \in H$:
>
> $$\langle y, z \rangle = \lim_n \langle Tx_n, z \rangle = \lim_n \langle x_n, Tz \rangle = 0.$$
>
> Since this holds for all z, $y = 0$. By the closed graph theorem, T is bounded. This is the *Hellinger–Toeplitz theorem*.

6.4 Applications of the closed graph theorem

6.4.1 Unbounded operators and their domains

> **Definition 6.17.** An **unbounded operator** from X to Y is a linear map $T\colon D(T) \to Y$ where $D(T)$ is a (not necessarily closed) linear subspace of X. The operator is **densely defined** if $D(T)$ is dense in X.

> **Proposition 6.18.** *Let X and Y be Banach spaces. If $T\colon D(T) \to Y$ is a closed operator and $D(T)$ is a closed subspace of X, then T is bounded.*

Proof. Since $D(T)$ is a closed subspace of a Banach space, it is itself a Banach space. The operator $T\colon D(T) \to Y$ is closed and defined on the entire Banach space $D(T)$. By the closed graph theorem, T is bounded. $\qquad\square$

> **Corollary 6.19.** *If $T\colon D(T) \to Y$ is closed and unbounded, then $D(T)$ cannot be a closed subspace of X. In particular, if T is densely defined,*

closed, and unbounded, then $D(T) \neq X$.

Example 6.20. The Laplacian $\Delta = \sum_{k=1}^{n} \partial^2/\partial x_k^2$ on $L^2(\mathbb{R}^n)$ is a densely defined closed operator with domain $D(\Delta) = H^2(\mathbb{R}^n)$ (the Sobolev space of functions with two weak derivatives in L^2). By the corollary, $H^2(\mathbb{R}^n)$ is not closed in $L^2(\mathbb{R}^n)$ (which is obvious since $H^2 \neq L^2$ but H^2 is dense).

Proposition 6.21 (Graph norm). *Let $T\colon D(T) \to Y$ be a closed operator. The **graph norm** on $D(T)$ is defined by*

$$\|x\|_T = \|x\|_X + \|Tx\|_Y .$$

Then $(D(T), \|\cdot\|_T)$ is a Banach space, and $T\colon (D(T), \|\cdot\|_T) \to Y$ is bounded with $\|T\|_{\|\cdot\|_T \to Y} \leq 1$.

Proof. The map $\Phi\colon (D(T), \|\cdot\|_T) \to \Gamma(T)$ defined by $\Phi(x) = (x, Tx)$ is an isometric isomorphism. Since T is closed, $\Gamma(T)$ is a closed subspace of the Banach space $X \times Y$, hence complete. Thus $(D(T), \|\cdot\|_T)$ is complete. The boundedness of T follows from $\|Tx\| \leq \|x\| + \|Tx\| = \|x\|_T$. $\square$

6.4.2 Application: automatic continuity of homomorphisms

Theorem 6.22. *Let A and B be Banach algebras and let $\varphi\colon A \to B$ be an algebra homomorphism that is also a surjection. If B is semisimple (i.e., its Jacobson radical is zero), then φ is continuous.*

Remark 6.23. This deep result, due to B. E. Johnson (1967), illustrates how the closed graph theorem can be used to derive automatic continuity in algebraic settings. The proof uses the closed graph theorem together with properties of the spectrum. We omit the full proof and refer to [1] for details.

6.5 The closed range theorem

Definition 6.24. Let X be a Banach space. For $S \subset X$, the **annihilator** of S is

$$S^\perp = \{f \in X^* : f(x) = 0 \text{ for all } x \in S\}.$$

For $W \subset X^*$, the **preannihilator** of W is

$$^\perp W = \{x \in X : f(x) = 0 \text{ for all } f \in W\}.$$

Theorem 6.25 (Closed range theorem). *Let X and Y be Banach spaces and $T \in \mathcal{L}(X,Y)$. The following conditions are equivalent:*

(i) $\mathrm{ran}(T) = T(X)$ *is closed in Y.*

(ii) $\mathrm{ran}(T^*) = T^*(Y^*)$ *is closed in X^*.*

(iii) $\mathrm{ran}(T) = {}^\perp(\mathrm{Ker}\, T^*)$.

(iv) $\mathrm{ran}(T^*) = (\mathrm{Ker}\, T)^\perp$.

Proof. We prove the cycle (i) $\Rightarrow$ (iv) $\Rightarrow$ (ii) $\Rightarrow$ (iii) $\Rightarrow$ (i).

(i) $\Rightarrow$ (iv): The inclusion $\mathrm{ran}(T^*) \subset (\mathrm{Ker}\, T)^\perp$ is clear: if $f = T^*g$ and $x \in \mathrm{Ker}\, T$, then $f(x) = g(Tx) = 0$.

For the reverse, let $f \in (\mathrm{Ker}\, T)^\perp$. Since $\mathrm{ran}(T)$ is closed, the quotient map induces an isomorphism $\tilde{T}\colon X/\mathrm{Ker}\, T \to \mathrm{ran}(T)$ defined by $\tilde{T}([x]) = Tx$, and $\tilde{T}$ is bounded below (by the open mapping theorem applied to $\tilde{T}$, since both $X/\mathrm{Ker}\, T$ and $\mathrm{ran}(T)$ are Banach spaces). Define $\tilde{f}\colon X/\mathrm{Ker}\, T \to \mathbb{K}$ by $\tilde{f}([x]) = f(x)$ (well-defined since $f \in (\mathrm{Ker}\, T)^\perp$). Then $\tilde{f}$ is bounded. Define $h\colon \mathrm{ran}(T) \to \mathbb{K}$ by $h(Tx) = f(x)$. Then h is well-defined and bounded (since $|h(Tx)| = |f(x)| \le \|f\| \|x\|$ and $\|x\| \le C\|Tx\|$ by the bounded below property on $X/\mathrm{Ker}\, T$, we get $|h(y)| \le \|f\| C \|y\|$ for $y \in \mathrm{ran}(T)$). Extend h to $g \in Y^*$ by Hahn–Banach. Then $T^*g = f$, so $f \in \mathrm{ran}(T^*)$.

(iv) $\Rightarrow$ (ii): Immediate, since $(\mathrm{Ker}\, T)^\perp$ is a closed subspace of X^*.

(ii) $\Rightarrow$ (iii): The inclusion $\mathrm{ran}(T) \subset {}^\perp(\mathrm{Ker}\, T^*)$ always holds. For the reverse, suppose $y \in {}^\perp(\mathrm{Ker}\, T^*)$, i.e., $g(y) = 0$ for all $g \in \mathrm{Ker}\, T^*$. We must show $y \in \overline{\mathrm{ran}(T)} = \mathrm{ran}(T)$ (we need to show the range is closed). Define $\Phi\colon \mathrm{ran}(T^*) \to \mathbb{K}$ by $\Phi(T^*g) = g(y)$. This is well-defined: if $T^*g_1 = T^*g_2$, then $g_1 - g_2 \in \mathrm{Ker}\, T^*$, so $(g_1 - g_2)(y) = 0$, i.e., $g_1(y) = g_2(y)$.

Since $\operatorname{ran}(T^*)$ is closed (hypothesis (ii)), it is a Banach space. The functional Φ is bounded: by the open mapping theorem, the induced map $\tilde{T}^* \colon Y^*/\operatorname{Ker} T^* \to \operatorname{ran}(T^*)$ is an isomorphism, so there exists $C > 0$ with $\inf_{h \in \operatorname{Ker} T^*} \|g - h\| \le C \|T^*g\|$. Then $|\Phi(T^*g)| = |g(y)| = \inf_{h \in \operatorname{Ker} T^*} |(g - h)(y)| \le \|y\| \inf_h \|g - h\| \le C \|y\| \|T^*g\|$.

By Hahn–Banach, extend Φ to $F \in (X^*)^* = X^{**}$. Under the canonical embedding $X \hookrightarrow X^{**}$, we have $F(T^*g) = g(y)$ for all g. However, we need $y \in \operatorname{ran}(T)$, not just $y \in {}^\perp(\operatorname{Ker} T^*)$.

Consider the restriction $\hat{T} \colon X \to \overline{\operatorname{ran}(T)}$. Then (ii) for T gives $\operatorname{ran}(T^*) = \operatorname{ran}(\hat{T}^*)$ is closed. Repeating the argument of (i) $\Rightarrow$ (iv) with $\overline{\operatorname{ran}(T)}$ in place of Y: since $\operatorname{ran}(\hat{T}^*)$ is closed, the induced map is bounded below, and the construction shows $y \in \operatorname{ran}(T)$. Hence $\operatorname{ran}(T) = {}^\perp(\operatorname{Ker} T^*)$ and in particular $\operatorname{ran}(T)$ is closed.

(iii) $\Rightarrow$ (i): ${}^\perp(\operatorname{Ker} T^*)$ is an intersection of kernels of continuous functionals, hence closed. So $\operatorname{ran}(T) = {}^\perp(\operatorname{Ker} T^*)$ is closed. $\qquad\square$

Corollary 6.26. *Let $T \in \mathcal{L}(X, Y)$ with X, Y Banach spaces. Then T is surjective if and only if T^* is bounded below, i.e., there exists $c > 0$ with $\|T^*g\| \ge c \|g\|$ for all $g \in Y^*$.*

Proof. If T is surjective, then $\operatorname{ran}(T) = Y$ is closed, and the closed range theorem gives $\operatorname{ran}(T^*) = (\operatorname{Ker} T)^\perp$. Moreover, $\operatorname{Ker} T^* = (\operatorname{ran} T)^\perp = Y^\perp = \{0\}$, so T^* is injective with closed range, hence bounded below by Proposition 6.8.

Conversely, if T^* is bounded below, then T^* is injective and $\operatorname{ran}(T^*)$ is closed. By the closed range theorem, $\operatorname{ran}(T) = {}^\perp(\operatorname{Ker} T^*) = {}^\perp\{0\} = Y$. $\qquad\square$

6.6 Characterization of operators with closed range

Theorem 6.27 (Characterization via minimum modulus)**.** *Let X and Y be Banach spaces and $T \in \mathcal{L}(X, Y)$. Define the **minimum modulus** of T as*

$$\gamma(T) = \inf\left\{ \frac{\|Tx\|}{\operatorname{dist}(x, \operatorname{Ker} T)} : x \notin \operatorname{Ker} T \right\}.$$

Then $\operatorname{ran}(T)$ is closed if and only if $\gamma(T) > 0$.

Proof. The quotient space $X/\operatorname{Ker} T$ is a Banach space (since $\operatorname{Ker} T$ is closed). The induced map $\hat{T}\colon X/\operatorname{Ker} T \to Y$ defined by $\hat{T}([x]) = Tx$ is a well-defined injective bounded operator with $\operatorname{ran}(\hat{T}) = \operatorname{ran}(T)$. Moreover,

$$\gamma(T) = \inf_{\|[x]\|=1} \left\| \hat{T}([x]) \right\|.$$

If $\gamma(T) > 0$, then $\hat{T}$ is bounded below by $\gamma(T)$, so $\operatorname{ran}(\hat{T})$ is closed (Proposition 6.8).

Conversely, if $\operatorname{ran}(T)$ is closed, then $\hat{T}\colon X/\operatorname{Ker} T \to \operatorname{ran}(T)$ is a bounded bijection between Banach spaces. By the Banach isomorphism theorem, $\hat{T}^{-1}$ is bounded, say $\left\| \hat{T}^{-1} \right\| = C$. Then $\|[x]\| \leq C \|Tx\|$ for all x, giving $\gamma(T) \geq 1/C > 0$. $\qquad\square$

> **Proposition 6.28.** *Let $T \in \mathcal{L}(X,Y)$ have closed range with $\gamma(T) > 0$. If $S \in \mathcal{L}(X,Y)$ with $\|S\| < \gamma(T)$, then $T + S$ also has closed range.*

Proof. For $x \notin \operatorname{Ker}(T + S)$:

$$\|(T + S)x\| \geq \|Tx\| - \|Sx\| \geq \gamma(T)\operatorname{dist}(x, \operatorname{Ker} T) - \|S\|\,\|x\|$$
$$\geq \gamma(T)\operatorname{dist}(x, \operatorname{Ker}(T + S)) - \|S\|\,\|x\|.$$

A more careful argument using the quotient space shows that $\gamma(T + S) \geq \gamma(T) - \|S\| > 0$, hence $\operatorname{ran}(T + S)$ is closed. $\qquad\square$

6.7 Factorization principles

> **Theorem 6.29** (Factorization through a closed subspace)**.** *Let X, Y, Z be Banach spaces and $T \in \mathcal{L}(X,Y)$, $S \in \mathcal{L}(X,Z)$. Suppose $\operatorname{Ker} T \subset \operatorname{Ker} S$. Then there exists a unique linear operator $R\colon \operatorname{ran}(T) \to Z$ such that $S = R \circ T$ on X. If $\operatorname{ran}(T)$ is closed, then R is bounded.*

Proof. Define $R(Tx) = Sx$. This is well-defined: if $Tx_1 = Tx_2$, then $x_1 - x_2 \in \operatorname{Ker} T \subset \operatorname{Ker} S$, so $Sx_1 = Sx_2$. Linearity is clear.

If $\operatorname{ran}(T)$ is closed, then $\operatorname{ran}(T)$ is a Banach space. We verify R is closed: suppose $Tx_n \to y \in \operatorname{ran}(T)$ and $R(Tx_n) = Sx_n \to z$. Since $\operatorname{ran}(T)$ is closed, $y = Tx_0$ for some x_0. By the open mapping theorem applied to $\hat{T}\colon X/\operatorname{Ker} T \to \operatorname{ran}(T)$, there exist $\tilde{x}_n$ with $T\tilde{x}_n = Tx_n$ and $\|\tilde{x}_n - \tilde{x}_0\| \to 0$ (choosing coset representatives). Since $\operatorname{Ker} T \subset \operatorname{Ker} S$, $S\tilde{x}_n = Sx_n \to z$ and $S\tilde{x}_n \to S\tilde{x}_0$, so $z = Sx_0 = R(Tx_0) = R(y)$. By the closed graph theorem, R is bounded. $\qquad\square$

> **Corollary 6.30** (Open mapping theorem — alternative formulation).
>
> *Let $T \in \mathcal{L}(X, Y)$ be surjective. Then T factors as $X \xrightarrow{\pi} X/\operatorname{Ker} T \xrightarrow{\hat{T}} Y$, where π is the quotient map and $\hat{T}$ is a topological isomorphism.*

Proof. $\hat{T}$ is a bijective bounded operator from $X/\operatorname{Ker} T$ (a Banach space) to Y (a Banach space). By the Banach isomorphism theorem, $\hat{T}$ is an isomorphism. $\qquad\square$

6.8 Fréchet spaces and the generalized closed graph theorem

The open mapping and closed graph theorems generalize beyond Banach spaces to the setting of *Fréchet spaces*.

> **Definition 6.31.** A **Fréchet space** is a topological vector space X whose topology is induced by a *complete, translation-invariant metric* d and by a countable family of seminorms $\{p_n\}_{n \geq 1}$ such that:
>
> $$d(x, y) = \sum_{n=1}^{\infty} \frac{1}{2^n} \frac{p_n(x - y)}{1 + p_n(x - y)}.$$
>
> Equivalently, X is a complete metrizable locally convex space.

> **Example 6.32.**   (i) Every Banach space is a Fréchet space (with a single seminorm equal to the norm).
>
> (ii) $C^{\infty}([0, 1])$, the space of smooth functions on $[0, 1]$, with seminorms $p_n(f) = \max_{0 \leq k \leq n} \left\| f^{(k)} \right\|_{\infty}$.
>
> (iii) The Schwartz space $\mathcal{S}(\mathbb{R}^n)$ of rapidly decreasing functions, with seminorms $p_{\alpha, \beta}(f) = \sup_x \left| x^{\alpha} D^{\beta} f(x) \right|$.
>
> (iv) $\mathcal{H}(\Omega)$, the space of holomorphic functions on an open set $\Omega \subset \mathbb{C}$, with seminorms $p_n(f) = \sup_{K_n} |f|$ for a compact exhaustion $\{K_n\}$ of Ω.
>
> (v) $\omega = \mathbb{K}^{\mathbb{N}}$, the space of all sequences, with seminorms $p_n(x) = |x_n|$.

Theorem 6.33 (Open mapping theorem for Fréchet spaces)**.** *Let X and Y be Fréchet spaces and $T\colon X \to Y$ a continuous linear surjection. Then T is open.*

Proof sketch. The proof follows the same pattern as for Banach spaces. The Baire category theorem applies since Y is a complete metric space. The approximation argument in Step 2 of Theorem 6.3 is adapted to the metric d (or equivalently to the countable family of seminorms), using the completeness of X to sum the convergent series. □

Theorem 6.34 (Closed graph theorem for Fréchet spaces)**.** *Let X and Y be Fréchet spaces and $T\colon X \to Y$ a linear operator with closed graph. Then T is continuous.*

Proof. The product $X \times Y$ is a Fréchet space. The graph $\Gamma(T)$ is a closed subspace, hence a Fréchet space. The projection $\pi_1\colon \Gamma(T) \to X$ is a continuous bijection between Fréchet spaces. By Theorem 6.33, π_1 is open, hence π_1^{-1} is continuous. Then $T = \pi_2 \circ \pi_1^{-1}$ is continuous. □

Remark 6.35. There are further generalizations. De Wilde (1971) proved the closed graph theorem for operators from an *ultrabornological* space to a *webbed space*. The class of webbed spaces includes all Fréchet spaces, countable inductive limits of Fréchet spaces (LF-spaces), and their closed subspaces and quotients. This provides the most general framework in which a "closed graph implies continuous" theorem holds without additional hypotheses on the operator.

Example 6.36 (Failure of the closed graph theorem)**.** The closed graph theorem fails if X is not complete. Let $X = (C^1([0,1]), \|\cdot\|_\infty)$ and $Y = (C([0,1]), \|\cdot\|_\infty)$. The differentiation operator $T = d/dx\colon X \to Y$ is closed (as shown in Example 6.12) but unbounded. Here X is not complete in the $\|\cdot\|_\infty$ norm.

6.9   Exercises for Chapter 6

Exercise 6.1 (Open mapping theorem application; $\star$). Let X be a Banach space and M a closed subspace. Show that the quotient map $\pi \colon X \to X/M$ is an open mapping. Conclude that the quotient topology on X/M coincides with the topology induced by the quotient norm.

Exercise 6.2 (Two Banach space norms; $\star$). Let $(X, \|\cdot\|_1)$ and $(X, \|\cdot\|_2)$ be Banach spaces with the same underlying vector space. Suppose that $x_n \to 0$ in $\|\cdot\|_1$ and $x_n \to x$ in $\|\cdot\|_2$ implies $x = 0$. Show that the two norms are equivalent.
Hint: Consider the identity map and use the closed graph theorem.

Exercise 6.3 (Projection operators; $\star$). Let X be a Banach space and $P \in \mathcal{L}(X)$ a projection ($P^2 = P$). Show that $X = \operatorname{ran}(P) \oplus \operatorname{Ker}(P)$ and both $\operatorname{ran}(P)$ and $\operatorname{Ker}(P)$ are closed. Conversely, if $X = M \oplus N$ with M, N closed subspaces, show that the projection onto M along N is bounded.
Hint for the converse: Use the closed graph theorem or the Banach isomorphism theorem on the map $M \times N \to X$, $(m, n) \mapsto m + n$.

Exercise 6.4 (Closed graph and sequential characterization; $\star$). Let X and Y be Banach spaces and $T \colon X \to Y$ a linear operator. Show that T is bounded if and only if $x_n \rightharpoonup x$ in X implies $Tx_n \rightharpoonup Tx$ in Y (i.e., T is weakly sequentially continuous).

Exercise 6.5 (Hellinger–Toeplitz; $\star\star$). Let H be a Hilbert space. Suppose $T \colon H \to H$ is a linear map satisfying $\langle Tx, y \rangle = \langle x, Ty \rangle$ for all $x, y \in H$. Prove that T is bounded *without* using the closed graph theorem, by directly applying the uniform boundedness principle.
Hint: Fix $y \in H$ and consider the functionals $f_y(x) = \langle Tx, y \rangle$.

Exercise 6.6 (Inverse of a perturbation; $\star\star$). Let X be a Banach space

and $T \in \mathcal{L}(X)$ with $\|T\| < 1$. Show that $\mathrm{Id} - T$ is invertible with

$$(\mathrm{Id} - T)^{-1} = \sum_{n=0}^{\infty} T^n$$

(the **Neumann series**), and $\|(\mathrm{Id} - T)^{-1}\| \leq 1/(1 - \|T\|)$. Use this to show: if $S \in \mathcal{L}(X)$ is invertible and $\|U - S\| < 1/\|S^{-1}\|$, then U is invertible.

Exercise 6.7 (Closed range and duality; ⋆⋆)**.** Let $T \in \mathcal{L}(X, Y)$ with X, Y Banach spaces.

(a) Show that if T has closed range, then $\mathrm{ran}(T)$ is a Banach space isomorphic to $X / \mathrm{Ker}\, T$.

(b) Show that T has closed range if and only if there exists $C > 0$ such that for every $y \in \mathrm{ran}(T)$, there exists $x \in X$ with $Tx = y$ and $\|x\| \leq C \|y\|$.

(c) Deduce that T is surjective if and only if T^* is injective and has closed range.

Exercise 6.8 (Fréchet space of smooth functions; ⋆⋆)**.** Let $X = C^\infty(\mathbb{R})$ with the family of seminorms $p_{n,K}(f) = \sup_{x \in [-n,n]} |f^{(k)}(x)|$ for $k \leq n$.

(a) Show that X is a Fréchet space.

(b) Show that the differentiation operator $D \colon X \to X$, $Df = f'$, is continuous.

(c) Show that D is surjective (every smooth function has a smooth antiderivative) and apply the open mapping theorem.

Exercise 6.9 (Schauder's theorem; ⋆⋆⋆)**.** Let X, Y be Banach spaces and $T \in \mathcal{L}(X, Y)$. Prove that T is compact if and only if T^* is compact. *Hint:* For one direction, use the Arzelà–Ascoli theorem. For the other, use the closed range theorem and properties of compact operators.

Exercise 6.10 (Operators on ℓ^p; ⋆⋆). Let $1 \le p \le \infty$ and let $A = (a_{ij})_{i,j \ge 1}$ be an infinite matrix defining a linear map $T_A \colon \ell^p \to \ell^p$ by $(T_A x)_i = \sum_j a_{ij} x_j$.

(a) Use the closed graph theorem to show: if T_A maps ℓ^p into ℓ^p (i.e., the series converges and the result is in ℓ^p for every $x \in \ell^p$), then T_A is automatically bounded.

(b) For $p = 1$, show directly that $\|T_A\| = \sup_j \sum_i |a_{ij}|$.

(c) For $p = \infty$, show that $\|T_A\| = \sup_i \sum_j |a_{ij}|$.

Exercise 6.11 (Banach isomorphism and complementation; ⋆⋆⋆). Let X be a Banach space and M, N closed subspaces with $M \cap N = \{0\}$ and $M + N = X$. Prove:

(a) The map $\Phi \colon M \times N \to X$, $\Phi(m, n) = m + n$ is a topological isomorphism.

(b) There exists $C > 0$ such that $\|m\| + \|n\| \le C \|m + n\|$ for all $m \in M$, $n \in N$.

(c) Give an example of a Banach space X and a closed subspace M that is *not* complemented. *Hint:* c_0 in ℓ^∞.

Exercise 6.12 (The three-space property; ⋆⋆⋆). Let X be a Banach space and M a closed subspace such that both M and X/M are isomorphic (as Banach spaces) to Hilbert spaces. Must X be isomorphic to a Hilbert space? *Hint:* Consider the Enflo–Lindenstrauss–Pisier example, or prove the positive result under additional hypotheses.

Exercise 6.13 (Uniform boundedness via open mapping; ⋆⋆). Derive the uniform boundedness principle as a consequence of the open mapping theorem.
Hint: Given pointwise bounded $(T_n) \subset \mathcal{L}(X, Y)$, consider the operator $\Phi \colon X \to \ell^\infty(Y)$ (or $Y^{\mathbb{N}}$) defined by $\Phi(x) = (T_n x)_n$, show it has closed graph, and apply the closed graph theorem.

Exercise 6.14 (Quotient maps are open; ⋆). Let X be a Banach space and $M \subset X$ a closed subspace. Prove directly (without using the open mapping theorem) that the canonical projection $\pi \colon X \to X/M$ is an open map. Then give a second proof using the open mapping theorem.

Exercise 6.15 (Isomorphic preduals; ⋆⋆⋆). Let X and Y be Banach spaces such that $X^* \cong Y^*$ (isometrically isomorphic). Must $X \cong Y$? Investigate this question for the following cases:

(a) $X = \ell^1$ and $Y = \ell^1$. (Here $X^* = Y^* = \ell^\infty$.)

(b) $X = L^1([0,1])$. What is X^*? Find a non-isomorphic space with the same dual.

Exercise 6.16 (Closed graph for multilinear maps; ⋆⋆). Let $X_1, \ldots, X_n, Y$ be Banach spaces and $T \colon X_1 \times \cdots \times X_n \to Y$ an n-linear map. Show that if T is *separately* continuous in each variable, then T is jointly continuous.
Hint: Proceed by induction on n. For $n = 1$, this is trivial. For the inductive step, fix all but one variable and apply the uniform boundedness principle.

Exercise 6.17 (Surjectivity and the open mapping theorem; ⋆⋆). Let X, Y be Banach spaces and $T \in \mathcal{L}(X, Y)$ with $\mathrm{ran}(T)$ of second category in Y. Show that T is surjective.
Hint: Revisit the proof of the open mapping lemma. If $\mathrm{ran}(T)$ is of second category, the same Baire argument shows that the closure of the image of the unit ball has nonempty interior *in Y*.

Exercise 6.18 (Application to differential equations; ⋆⋆). Consider the Banach space $X = C([0,1])$ and the operator $T \colon D(T) \to X$ defined by $Tf = f'' + f$ with $D(T) = \{f \in C^2([0,1]) : f(0) = f(1) = 0\}$.

(a) Show that T is a closed operator (with $D(T)$ equipped with the supremum norm from X).

(b) Show that $D(T)$ is not closed in X.

(c) Equip $D(T)$ with the graph norm. Show that T becomes a

bounded operator from $(D(T), \|\cdot\|_T)$ to X.

(d) Determine $\mathrm{Ker}(T)$ explicitly.

Exercise 6.19 (The Mittag-Leffler theorem; $\star\star\star$). *(A topological version.)* Let $(X_n, d_n)_{n \geq 0}$ be a sequence of complete metric spaces and $\varphi_n \colon X_{n+1} \to X_n$ continuous maps with dense range. Define the projective limit $X = \varprojlim X_n = \{(x_n) \in \prod X_n : \varphi_n(x_{n+1}) = x_n\}$. Prove that the projection $\pi_n \colon X \to X_n$ has dense range for each n.
Hint: Use a Baire-category-style nested ball argument, constructing compatible approximations in each X_n.

Exercise 6.20 (Open mapping for non-complete spaces; $\star\star$). Let $X = (\ell^1, \|\cdot\|_1)$ and define $Y = (\ell^1, \|\cdot\|_2)$ where $\|x\|_2 = \left(\sum_n |x_n|^2\right)^{1/2}$. Note that Y is *not* complete.

(a) Show that the identity $\mathrm{Id} \colon X \to Y$ is a well-defined, continuous, linear bijection.

(b) Show that Id^{-1} is *not* continuous, hence Id is not open.

(c) Explain which hypothesis of the open mapping theorem fails.

Exercise 6.21 (Fredholm operators; $\star\star\star$). Let X, Y be Banach spaces. An operator $T \in \mathcal{L}(X, Y)$ is called **Fredholm** if $\mathrm{Ker}\, T$ is finite-dimensional, $\mathrm{ran}(T)$ is closed, and $\mathrm{codim}(\mathrm{ran}\, T) < \infty$. The **Fredholm index** is $\mathrm{ind}(T) = \dim \mathrm{Ker}\, T - \mathrm{codim}\, \mathrm{ran}\, T$.

(a) Show that T is Fredholm if and only if T^* is Fredholm, and $\mathrm{ind}(T^*) = -\mathrm{ind}(T)$.

(b) Show that if T is Fredholm and $K \in \mathcal{L}(X, Y)$ is compact, then $T + K$ is Fredholm with $\mathrm{ind}(T + K) = \mathrm{ind}(T)$.

(c) Show that the set of Fredholm operators is open in $\mathcal{L}(X, Y)$ and the index is locally constant.

Chapter 7

Dual Spaces and Reflexivity

Contents

7.1   The Topological Dual

Throughout this chapter, E denotes a normed space over the field $\mathbb{K}$ ($= \mathbb{R}$ or $\mathbb{C}$) unless otherwise stated.

> **Definition 7.1** (Topological dual)**.** The **topological dual** (or simply

dual) of a normed space E is the Banach space

$$E' = \mathcal{L}(E, \mathbb{K}) = \big\{ f : E \to \mathbb{K} \mid f \text{ is continuous and linear} \big\},$$

equipped with the operator norm

$$\|f\|_{E'} = \sup_{\|x\| \leq 1} |f(x)|.$$

Elements of E' are called **continuous linear functionals** (or simply **functionals**).

Remark 7.2. Since $\mathbb{K}$ is complete, $E' = \mathcal{L}(E, \mathbb{K})$ is always a Banach space, regardless of whether E itself is complete.

We recall the fundamental extension theorem whose consequences pervade the entire theory.

Theorem 7.3 (Hahn–Banach, analytic form)**.** *Let E be a real or complex vector space, $p : E \to \mathbb{R}$ a sublinear functional (or seminorm in the complex case), $F \subset E$ a subspace, and $f : F \to \mathbb{K}$ a linear functional satisfying $|f(x)| \leq p(x)$ for all $x \in F$. Then there exists a linear functional $\tilde{f} : E \to \mathbb{K}$ extending f such that $\left| \tilde{f}(x) \right| \leq p(x)$ for all $x \in E$.*

Corollary 7.4. *Let F be a subspace of a normed space E and let $f \in F'$. Then there exists $\tilde{f} \in E'$ such that $\tilde{f}\big|_F = f$ and $\left\| \tilde{f} \right\|_{E'} = \|f\|_{F'}$.*

Corollary 7.5. *For every $x \in E$ with $x \neq 0$, there exists $f \in E'$ with $\|f\| = 1$ and $f(x) = \|x\|$. In particular, E' separates points of E:*

$$\big(\forall f \in E', \; f(x) = 0 \big) \implies x = 0.$$

Proof. Apply Corollary 7.4 to the functional $f_0 : \mathbb{K}x \to \mathbb{K}$ defined by $f_0(\lambda x) = \lambda \|x\|$, which satisfies $\|f_0\| = 1$. $\qquad\square$

7.2 Duals of Classical Sequence Spaces

7.2.1 The dual of ℓ^p for $1 \leq p < \infty$

We denote by p' the conjugate exponent of p, defined by $\frac{1}{p} + \frac{1}{p'} = 1$ with the convention $1' = \infty$.

Theorem 7.6 (Dual of ℓ^p, $1 \leq p < \infty$). *For $1 \leq p < \infty$, the map*

$$\Phi : \ell^{p'} \longrightarrow (\ell^p)', \qquad \Phi(y)(x) = \sum_{n=1}^{\infty} x_n y_n,$$

is an isometric isomorphism.

Proof. We divide the proof into three steps.

Step 1: Φ is well-defined and $\|\Phi(y)\| \leq \|y\|_{p'}$. By Hölder's inequality, for every $y \in \ell^{p'}$ and $x \in \ell^p$,

$$\left| \sum_{n=1}^{\infty} x_n y_n \right| \leq \sum_{n=1}^{\infty} |x_n| \, |y_n| \leq \|x\|_p \, \|y\|_{p'}.$$

Hence $\Phi(y) \in (\ell^p)'$ and $\|\Phi(y)\| \leq \|y\|_{p'}$.

Step 2: $\|\Phi(y)\| = \|y\|_{p'}$ (isometry). We exhibit an element $x \in \ell^p$ with $\|x\|_p = 1$ that achieves equality.

Case $1 < p < \infty$ (so $1 < p' < \infty$). Define

$$x_n = \begin{cases} |y_n|^{p'-2} \, \overline{y_n} \, \|y\|_{p'}^{1-p'} & \text{if } y_n \neq 0, \\ 0 & \text{if } y_n = 0. \end{cases}$$

Then $|x_n|^p = |y_n|^{(p'-1)p} \, \|y\|_{p'}^{(1-p')p} = |y_n|^{p'} \, \|y\|_{p'}^{-p'}$, so $\|x\|_p^p = \|y\|_{p'}^{p'} \, \|y\|_{p'}^{-p'} = 1$. Moreover,

$$\Phi(y)(x) = \sum_n x_n y_n = \|y\|_{p'}^{1-p'} \sum_n |y_n|^{p'} = \|y\|_{p'}^{1-p'} \, \|y\|_{p'}^{p'} = \|y\|_{p'}.$$

Case $p = 1$ (so $p' = \infty$). For any $\varepsilon > 0$, pick n_0 with $|y_{n_0}| > \|y\|_{\infty} - \varepsilon$. Set $x = \overline{\operatorname{sgn}(y_{n_0})} \, e_{n_0}$ where e_{n_0} is the n_0-th standard basis vector. Then $\|x\|_1 = 1$ and $\Phi(y)(x) = |y_{n_0}| > \|y\|_{\infty} - \varepsilon$. Letting $\varepsilon \to 0$ gives $\|\Phi(y)\| \geq \|y\|_{\infty}$.

Step 3: Φ is surjective. Let $f \in (\ell^p)'$. Define $y_n = f(e_n)$ where $(e_n)_{n \geq 1}$ is the standard basis.

Case $1 < p < \infty$. For each N, define $x^{(N)} \in \ell^p$ by

$$x_n^{(N)} = \begin{cases} |y_n|^{p'-2}\,\overline{y_n} & \text{if } 1 \le n \le N \text{ and } y_n \ne 0, \\ 0 & \text{otherwise.} \end{cases}$$

Then $\left\| x^{(N)} \right\|_p^p = \sum_{n=1}^N |y_n|^{(p'-1)p} = \sum_{n=1}^N |y_n|^{p'}$ and

$$f(x^{(N)}) = \sum_{n=1}^N |y_n|^{p'}.$$

By the bound $|f(x^{(N)})| \le \|f\| \left\| x^{(N)} \right\|_p$,

$$\sum_{n=1}^N |y_n|^{p'} \;\le\; \|f\| \left(\sum_{n=1}^N |y_n|^{p'} \right)^{1/p},$$

hence $\left(\sum_{n=1}^N |y_n|^{p'} \right)^{1-1/p} \le \|f\|$, i.e. $\left(\sum_{n=1}^N |y_n|^{p'} \right)^{1/p'} \le \|f\|$. Letting $N \to \infty$, we get $y \in \ell^{p'}$ and $\|y\|_{p'} \le \|f\|$.

Now for any $x \in \ell^p$, the partial sums $s_N = \sum_{n=1}^N x_n e_n$ converge to x in ℓ^p, so

$$f(x) = \lim_{N \to \infty} f(s_N) = \lim_{N \to \infty} \sum_{n=1}^N x_n y_n = \sum_{n=1}^\infty x_n y_n = \Phi(y)(x).$$

Hence $f = \Phi(y)$.

Case $p = 1$. We have $|y_n| = |f(e_n)| \le \|f\| \, \|e_n\|_1 = \|f\|$ for all n, so $y \in \ell^\infty$ with $\|y\|_\infty \le \|f\|$. The same density argument as above shows $f = \Phi(y)$. $\qquad\square$

7.2.2   The dual of c_0

Theorem 7.7 (Dual of c_0). *The map*

$$\Phi : \ell^1 \longrightarrow c_0', \qquad \Phi(y)(x) = \sum_{n=1}^\infty x_n y_n,$$

is an isometric isomorphism.

Proof. **Well-defined and contractive.** For $y \in \ell^1$ and $x \in c_0$ (hence $x \in \ell^\infty$),

$$\left| \Phi(y)(x) \right| \le \sum_{n=1}^\infty |x_n|\,|y_n| \le \|x\|_\infty \, \|y\|_1 \,,$$

so $\Phi(y) \in c_0'$ with $\|\Phi(y)\| \leq \|y\|_1$.

Isometry. For $\varepsilon > 0$, define $x^{(N)} \in c_0$ by

$$x_n^{(N)} = \begin{cases} \overline{\mathrm{sgn}(y_n)} & \text{if } 1 \leq n \leq N, \\ 0 & \text{if } n > N. \end{cases}$$

Then $\left\|x^{(N)}\right\|_\infty \leq 1$ and $\Phi(y)(x^{(N)}) = \sum_{n=1}^{N} |y_n| \to \|y\|_1$ as $N \to \infty$. Hence $\|\Phi(y)\| \geq \|y\|_1$.

Surjectivity. Let $f \in c_0'$ and set $y_n = f(e_n)$. For each N and each choice of signs $\theta_n \in \{z \in \mathbb{K} : |z| = 1\}$ with $\theta_n = \overline{\mathrm{sgn}(y_n)}$, the vector $x^{(N)} = \sum_{n=1}^{N} \theta_n e_n \in c_0$ satisfies $\left\|x^{(N)}\right\|_\infty \leq 1$ and

$$\sum_{n=1}^{N} |y_n| = f(x^{(N)}) \leq \|f\|.$$

Letting $N \to \infty$ gives $y \in \ell^1$ with $\|y\|_1 \leq \|f\|$. Since finite sequences are dense in c_0, the same approximation argument as in Theorem 7.6 shows $f = \Phi(y)$. $\qquad\square$

7.2.3   The dual of L^p

Theorem 7.8 (Dual of L^p, Riesz representation). *Let $(\Omega, \mathcal{A}, \mu)$ be a σ-finite measure space and $1 \leq p < \infty$. Then the map*

$$\Phi : L^{p'}(\mu) \longrightarrow (L^p(\mu))', \qquad \Phi(g)(f) = \int_\Omega f \, g \, \mathrm{d}\mu,$$

is an isometric isomorphism.

Remark 7.9. The complete proof uses the Radon–Nikodým theorem and will be given in Chapter 11. For $p = 2$, the result follows immediately from the Riesz representation theorem for Hilbert spaces.

7.3   The Canonical Injection and Bidual

Definition 7.10 (Bidual and canonical injection)**.** The **bidual** of E is $E'' = (E')'$. The **canonical injection** (or **evaluation map**) is

$$J : E \longrightarrow E'', \qquad J(x)(f) = f(x), \quad x \in E, \ f \in E'.$$

Proposition 7.11. *The canonical injection J is a linear isometry from E into E''.*

Proof. Linearity is clear. For each $x \in E$,

$$\|J(x)\|_{E''} = \sup_{\|f\|_{E'} \leq 1} |J(x)(f)| = \sup_{\|f\| \leq 1} |f(x)|.$$

On the one hand, $|f(x)| \leq \|f\| \, \|x\| \leq \|x\|$ for $\|f\| \leq 1$, so $\|J(x)\| \leq \|x\|$. On the other hand, Corollary 7.5 provides $f_0 \in E'$ with $\|f_0\| = 1$ and $f_0(x) = \|x\|$, so $\|J(x)\| \geq |f_0(x)| = \|x\|$. $\qquad\qquad\square$

Remark 7.12. The map J is always injective (since it is an isometry), but it need not be surjective. When J is surjective we say E is *reflexive*.

7.4   Reflexive Spaces

Definition 7.13 (Reflexive space)**.** A Banach space E is called **reflexive** if the canonical injection $J : E \to E''$ is surjective (hence an isometric isomorphism).

Remark 7.14. It is essential that the isomorphism onto E'' be the *canonical* one. James constructed a Banach space J that is isometrically isomorphic to J'', yet is not reflexive because the canonical injection is not surjective.

Example 7.15.

(i) Every finite-dimensional Banach space is reflexive (since $\dim E = \dim E' = \dim E''$).

(ii) Every Hilbert space is reflexive (by the Riesz representation theorem).

(iii) $L^p(\mu)$ is reflexive for $1 < p < \infty$ (as we shall prove below).

(iv) ℓ^p is reflexive for $1 < p < \infty$.

(v) $\ell^1, \ell^\infty, c_0, L^1(\mu)$, and $L^\infty(\mu)$ are **not** reflexive (for non-trivial μ).

Theorem 7.16 (ℓ^p is reflexive for $1 < p < \infty$). *For $1 < p < \infty$, the space ℓ^p is reflexive.*

Proof. Let $p' = p/(p-1)$ be the conjugate exponent, so that $1 < p' < \infty$ and $(p')' = p$. By Theorem 7.6, the isometric isomorphisms

$$(\ell^p)' \cong \ell^{p'}, \qquad (\ell^p)'' = ((\ell^p)')' \cong (\ell^{p'})' \cong \ell^{(p')'} = \ell^p$$

hold. We must verify that the composite isomorphism $(\ell^p)'' \cong \ell^p$ is precisely the canonical injection J.

Let $x = (x_n) \in \ell^p$. Under the identification $\Phi : \ell^{p'} \xrightarrow{\sim} (\ell^p)'$, a functional $f \in (\ell^p)'$ corresponds to $y = (y_n) \in \ell^{p'}$ via $f(x) = \sum x_n y_n$. Then

$$J(x)(f) = f(x) = \sum_{n=1}^{\infty} x_n y_n = \Psi(x)(y),$$

where $\Psi : \ell^p \to (\ell^{p'})'$ is the canonical identification from Theorem 7.6 applied to $\ell^{p'}$. Under the identification $(\ell^p)'' \cong (\ell^{p'})' \cong \ell^p$, the element $J(x)$ corresponds to x itself. Since J is an isometry and the identification maps to all of ℓ^p, J is surjective. $\square$

Theorem 7.17 (ℓ^1 is not reflexive). *The space ℓ^1 is not reflexive.*

Proof. We have $(\ell^1)' \cong \ell^\infty$ by Theorem 7.6. If ℓ^1 were reflexive, then $(\ell^1)'' \cong \ell^1$ canonically, whence $(\ell^\infty)' \cong \ell^1$. But ℓ^∞ is non-separable (it contains the uncountable set $\{(\varepsilon_n) : \varepsilon_n \in \{0,1\}\}$ whose pairwise distances are all 1), while the dual of a separable space need not be separable, but the predual of a separable space must be separable (since a separable dual implies a separable predual is false in general).

More directly: we exhibit a functional in $(\ell^\infty)'$ not in $J(\ell^1)$. Let $\mathcal{U}$ be a free ultrafilter on $\mathbb{N}$. Define $\Lambda : \ell^\infty \to \mathbb{K}$ by $\Lambda(x) = \lim_{\mathcal{U}} x_n$. Then Λ is a bounded linear functional with $\|\Lambda\| = 1$, and $\Lambda(e_n) = 0$ for every standard

basis vector e_n. If $\Lambda = J(a)$ for some $a \in \ell^1$, then $a_n = \Lambda(e_n) = 0$ for all n, hence $a = 0$, contradicting $\|\Lambda\| = 1$. $\qquad\square$

Theorem 7.18 (c_0 is not reflexive). *The space c_0 is not reflexive.*

Proof. We have $c_0' \cong \ell^1$ (Theorem 7.7) and $c_0'' \cong (\ell^1)' \cong \ell^\infty$. The canonical injection $J : c_0 \to \ell^\infty$ is the inclusion map. Since $c_0 \subsetneq \ell^\infty$ (for instance, the constant sequence $(1, 1, 1, \ldots) \in \ell^\infty \setminus c_0$), J is not surjective. $\qquad\square$

7.5   James' Characterization of Reflexivity

The following deep result provides a purely geometric characterization of reflexivity.

Theorem 7.19 (James, 1964). *A Banach space E is reflexive if and only if every continuous linear functional $f \in E'$ attains its norm, i.e., there exists $x_0 \in E$ with $\|x_0\| = 1$ and $|f(x_0)| = \|f\|$.*

Remark 7.20. The "only if" direction is straightforward: if E is reflexive, then B_E is weakly compact (Kakutani's theorem, Theorem 8.26), and any $f \in E'$ is weakly continuous, hence attains its supremum on B_E. The "if" direction is considerably harder; we refer to [2] for the proof.

Corollary 7.21. *The space c_0 is not reflexive because the functional $f(x) = \sum_{n=1}^{\infty} x_n/2^n$ satisfies $\|f\| = 1$ but does not attain its norm on the closed unit ball of c_0.*

Proof. We have $\|f\| = \sum 2^{-n} = 1$. If $x \in B_{c_0}$ with $\|x\|_\infty \le 1$, then $|f(x)| \le \sum |x_n|/2^n$. Equality $|f(x)| = 1$ requires $|x_n| = 1$ for all n, but then $x_n \not\to 0$, contradicting $x \in c_0$. $\qquad\square$

7.6 Subspaces and Quotients of Reflexive Spaces

Theorem 7.22. *Every closed subspace of a reflexive Banach space is reflexive.*

Proof. Let E be reflexive and $F \subset E$ a closed subspace. We must show that the canonical injection $J_F : F \to F''$ is surjective.

Let $\xi \in F''$. Define $\eta \in E''$ by

$$\eta(f) = \xi(f\big|_F), \qquad f \in E'.$$

This is well-defined: $f\big|_F \in F'$ for every $f \in E'$, and the restriction map $R : E' \to F'$, $R(f) = f\big|_F$, is bounded with $\|R\| \leq 1$ (by Hahn–Banach it is surjective). Moreover, $|\eta(f)| = |\xi(f\big|_F)| \leq \|\xi\| \, \|f\big|_F\|_{F'} \leq \|\xi\| \, \|f\|_{E'}$, so $\eta \in E''$ with $\|\eta\| \leq \|\xi\|$.

Since E is reflexive, $\eta = J_E(x)$ for some $x \in E$, meaning $f(x) = \eta(f) = \xi(f\big|_F)$ for all $f \in E'$.

We claim $x \in F$. If $x \notin F$, by Hahn–Banach there exists $f \in E'$ with $f\big|_F = 0$ and $f(x) = 1$. But then $1 = f(x) = \xi(f\big|_F) = \xi(0) = 0$, a contradiction.

So $x \in F$, and for every $g \in F'$, extend g to $\tilde{g} \in E'$ (Hahn–Banach):

$$J_F(x)(g) = g(x) = \tilde{g}(x) = \xi(\tilde{g}\big|_F) = \xi(g).$$

Hence $J_F(x) = \xi$, proving surjectivity. $\qquad\square$

Theorem 7.23. *If E is a reflexive Banach space and $F \subset E$ is a closed subspace, then the quotient space E/F is reflexive.*

Proof. Let $\pi : E \to E/F$ denote the quotient map. The adjoint $\pi' : (E/F)' \to E'$ is an isometric embedding whose image is $F^\perp = \{f \in E' : f\big|_F = 0\}$.

Let $\xi \in (E/F)''$. Define $\eta \in E''$ by $\eta(f) = \xi(\psi)$ where $\psi \in (E/F)'$ is the unique functional with $\pi'(\psi) = f$, whenever $f \in F^\perp$. For general $f \in E'$, we note that the composition $f \mapsto \xi \circ (\pi')^{-1}$ is defined on $F^\perp$ and extends by Hahn–Banach; however, a cleaner approach is as follows.

Define $T : (E/F)' \to \mathbb{K}$ by $T = \xi$. Consider the adjoint $\pi'' : (E/F)'' \to E''$ of π', defined by $\pi''(\xi)(f) = \xi((\pi')^{-1}(f))$ for $f \in \operatorname{ran}(\pi') = F^\perp$—but this requires care. Instead, we use the transpose construction.

Let $\eta = \xi \circ \pi^t$ where $\pi^t : E' \to (E/F)'$ is defined by $\pi^t(f)(\dot{x}) = f(x)$, valid when $f \in F^\perp$. Since π^t is not defined on all of E', we proceed differently.

Consider the map $\alpha : E' \to (E/F)'$ given by $\alpha(f) = 0$ if we compose differently. Let us use the direct approach.

Define $\zeta \in (F^\perp)'$ by $\zeta(f) = \xi(\psi_f)$ where $\psi_f \in (E/F)'$ corresponds to $f \in F^\perp$ via the isometry π'. Then ζ is bounded, and by Hahn–Banach we extend ζ to $\eta \in E''$. Since E is reflexive, $\eta = J_E(x)$ for some $x \in E$, meaning $f(x) = \zeta(f)$ for all $f \in F^\perp$.

Set $\dot{x} = \pi(x) = x + F \in E/F$. For any $\psi \in (E/F)'$, let $f = \pi'(\psi) = \psi \circ \pi \in F^\perp$. Then

$$J_{E/F}(\dot{x})(\psi) = \psi(\dot{x}) = \psi(\pi(x)) = f(x) = \zeta(f) = \xi(\psi).$$

Hence $J_{E/F}(\dot{x}) = \xi$ and E/F is reflexive. $\qquad\qquad\square$

> **Corollary 7.24.** *A Banach space E is reflexive if and only if E' is reflexive.*

Proof. ($\Rightarrow$) Suppose E is reflexive, so $J_E : E \xrightarrow{\sim} E''$. Let $\Phi \in E'''$. Define $f \in E'$ by $f(x) = \Phi(J_E(x))$ (well-defined since J_E is surjective). Then for any $\xi \in E''$, writing $\xi = J_E(x)$, we get $J_{E'}(f)(\xi) = \xi(f) = J_E(x)(f) = f(x) = \Phi(J_E(x)) = \Phi(\xi)$. So $J_{E'}(f) = \Phi$.

($\Leftarrow$) Suppose E' is reflexive. Since $J_E(E)$ is a closed subspace of $E'' = (E')'$ (being the isometric image of a Banach space), and E' is reflexive, Theorem 7.22 shows every closed subspace of E'' is reflexive. But if $J_E(E) \neq E''$, then by Hahn–Banach there exists $\Phi \in E''' \setminus \{0\}$ vanishing on $J_E(E)$. Since E' is reflexive, $\Phi = J_{E'}(f)$ for some $f \in E' \setminus \{0\}$. Then for all $x \in E$: $0 = \Phi(J_E(x)) = J_{E'}(f)(J_E(x)) = J_E(x)(f) = f(x)$, so $f = 0$, contradicting $f \neq 0$. Hence $J_E(E) = E''$. $\qquad\square$

7.7   The Eberlein–Šmulian Theorem

> **Theorem 7.25** (Eberlein–Šmulian). *Let E be a Banach space and $A \subset E$. The following are equivalent:*
>
> (i) *A is relatively weakly compact (its weak closure is weakly compact).*
>
> (ii) *A is relatively weakly sequentially compact (every sequence in A has a weakly convergent subsequence with limit in E).*
>
> (iii) *A is weakly limit-point compact (every infinite subset of A has a weak limit point in E).*

Remark 7.26. This is remarkable because the weak topology on an infinite-dimensional space is never metrizable on the whole space, so one cannot expect sequential and topological compactness to coincide in general. The Eberlein–Šmulian theorem shows they do coincide for the weak topology of a Banach space. The proof is technical; see [3] or [4] for details.

7.8 Uniformly Convex Spaces and the Milman–Pettis Theorem

Definition 7.27 (Uniformly convex space)**.** A normed space E is **uniformly convex** if for every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\|x\| \leq 1, \ \|y\| \leq 1, \ \|x - y\| \geq \varepsilon \quad \Longrightarrow \quad \left\|\frac{x+y}{2}\right\| \leq 1 - \delta.$$

The function $\delta_E(\varepsilon) = \inf\left\{1 - \left\|\frac{x+y}{2}\right\| : \|x\| \leq 1, \|y\| \leq 1, \|x - y\| \geq \varepsilon\right\}$ is called the **modulus of convexity** of E.

Example 7.28.

(i) Every Hilbert space is uniformly convex. Indeed, by the parallelogram law, $\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2 \leq 4$, so $\left\|\frac{x+y}{2}\right\|^2 \leq 1 - \frac{\varepsilon^2}{4}$, giving $\delta(\varepsilon) \geq 1 - \sqrt{1 - \varepsilon^2/4}$.

(ii) $L^p(\mu)$ and ℓ^p are uniformly convex for $1 < p < \infty$ (Clarkson's inequalities).

(iii) L^1, L^∞, ℓ^1, ℓ^∞ are *not* uniformly convex.

Proposition 7.29. *Every uniformly convex space is strictly convex: if* $\|x\| = \|y\| = 1$ *and* $x \neq y$, *then* $\left\|\frac{x+y}{2}\right\| < 1$.

Proof. Set $\varepsilon = \|x - y\| > 0$. By uniform convexity, $\left\|\frac{x+y}{2}\right\| \leq 1 - \delta(\varepsilon) < 1$. $\square$

Theorem 7.30 (Milman–Pettis, 1939)**.** *Every uniformly convex Banach space is reflexive.*

Proof. Let E be a uniformly convex Banach space. We must show that $J : E \to E''$ is surjective. Let $\xi \in E''$ with $\|\xi\| = 1$. We will find $x \in E$ with $J(x) = \xi$.

Step 1. By definition of the norm in E'' and the fact that $J(B_E)$ is dense in $B_{E''}$ for the weak-$*$ topology of E'' (this is Goldstine's theorem, Theorem 8.24, proved in Chapter 8), for each $n \geq 1$ there exists $x_n \in B_E$ such that

$$|\xi(f) - f(x_n)| < \frac{1}{n} \qquad \text{for } f \text{ in a finite set } F_n \subset E'. \tag{7.1}$$

However, we need a cleaner approach. We use the following key fact.

Step 1 (revised). Since $\|\xi\|_{E''} = 1$, for each n there exists $f_n \in E'$ with $\|f_n\| = 1$ and $\xi(f_n) > 1 - \frac{1}{n}$. Choose $x_n \in E$ with $\|x_n\| \leq 1$ and $f_n(x_n) > 1 - \frac{1}{n}$.

Step 2: (x_n) is Cauchy. Let $\varepsilon > 0$ and let $\delta = \delta_E(\varepsilon) > 0$ be the modulus of uniform convexity. We claim that for m, n large enough, $\|x_m - x_n\| < \varepsilon$.

Suppose for contradiction that $\|x_m - x_n\| \geq \varepsilon$ for infinitely many pairs. Then $\left\|\frac{x_m + x_n}{2}\right\| \leq 1 - \delta$. But for any $f \in E'$ with $\|f\| = 1$,

$$f\left(\frac{x_m + x_n}{2}\right) = \frac{f(x_m) + f(x_n)}{2}.$$

In particular, choosing appropriately: we take $f = f_n$ for large n and observe that we need f to test *both* x_m and x_n. Let us refine the argument.

We use a single functional. Since $\|\xi\| = 1$, there exists a net (x_α) in B_E with $J(x_\alpha) \to \xi$ weak-$*$ in E''. Instead, we use the following direct approach.

For m, n large, pick $g \in E'$ with $\|g\| = 1$ and $\xi(g) > 1 - \delta/2$. Then for n large enough (depending on g), since $J(B_E)$ is weak-$*$ dense in $B_{E''}$, we can ensure $g(x_n) > 1 - \delta/2$ and similarly $g(x_m) > 1 - \delta/2$. Hence

$$g\left(\frac{x_m + x_n}{2}\right) > 1 - \frac{\delta}{2}.$$

But $\left\|\frac{x_m + x_n}{2}\right\| \leq 1 - \delta$ (by uniform convexity if $\|x_m - x_n\| \geq \varepsilon$), which gives $g\left(\frac{x_m + x_n}{2}\right) \leq 1 - \delta$, a contradiction.

More precisely: we construct the sequence as follows. Pick $f \in E'$ with $\|f\| = 1$ and $\xi(f) > 1 - \delta/4$. By Goldstine's theorem, for each n there exists $x_n \in B_E$ with $|f(x_n) - \xi(f)| < \frac{1}{n}$ (and also with $|g(x_n) - \xi(g)| < \frac{1}{n}$ for finitely many g's we pre-select). In particular, $f(x_n) \to \xi(f) > 1 - \delta/4$, so for large n, $\operatorname{Re} f(x_n) > 1 - \delta/2$.

If $\|x_m - x_n\| \geq \varepsilon$, then $\left\|\frac{x_m + x_n}{2}\right\| \leq 1 - \delta$, so

$$\operatorname{Re} f\left(\frac{x_m + x_n}{2}\right) \leq \left\|\frac{x_m + x_n}{2}\right\| \leq 1 - \delta.$$

But $\operatorname{Re} f(\frac{x_m + x_n}{2}) = \frac{1}{2}(\operatorname{Re} f(x_m) + \operatorname{Re} f(x_n)) > 1 - \delta/2$ for m, n large, a contradiction.

Step 3: Conclusion. Since E is complete, $x_n \to x$ for some $x \in E$ with $\|x\| \le 1$. For any $f \in E'$, $J(x)(f) = f(x) = \lim f(x_n)$. By construction (choosing the sequence so that $f(x_n) \to \xi(f)$ for all f in a countable dense subset of E', and using the fact that $\|J(x_n) - \xi\| \to 0$ in E'' follows from the Cauchy property), we obtain $J(x) = \xi$.

Let us make Step 3 precise. We have shown that any sequence $(x_n) \subset B_E$ with $f_n(x_n) \to 1$ for some sequence (f_n) in the unit sphere of E' with $\xi(f_n) \to 1$ must be Cauchy. Now use a diagonal argument: let $(g_k)_{k \ge 1}$ be a sequence dense in $S_{E'}$ (or any countable family). By Goldstine's theorem, for each n we can find $x_n \in B_E$ with $|g_k(x_n) - \xi(g_k)| < \frac{1}{n}$ for $k = 1, \ldots, n$. The Cauchy argument shows $x_n \to x$. Then $g_k(x) = \lim g_k(x_n) = \xi(g_k)$ for all k. By density and continuity, $f(x) = \xi(f)$ for all $f \in E'$, i.e. $J(x) = \xi$.

(Note: if E' is not separable, one modifies the argument to work with nets, or uses a different approach. The argument works for arbitrary E because uniform convexity forces Cauchy behavior without needing separability of E'; one applies the argument to any single $\xi \in E''$.) $\qquad\square$

> **Corollary 7.31.** *For $1 < p < \infty$, the spaces $L^p(\mu)$ and ℓ^p are reflexive.*

Proof. These spaces are uniformly convex by Clarkson's inequalities, hence reflexive by the Milman–Pettis theorem. $\qquad\square$

7.9 Exercises

> **Exercise 7.1** ($\star$). Let E be a finite-dimensional normed space. Show that $E' \cong E$ (as vector spaces) and that every linear functional on E is continuous.

> **Exercise 7.2** ($\star$). Let E be a normed space and $A \subset E$. The *annihilator* of A is $A^\perp = \{f \in E' : f(a) = 0 \ \forall a \in A\}$.
>
> (a) Show that $A^\perp$ is a closed subspace of E'.
>
> (b) If F is a closed subspace of E, show that $F^\perp \cong (E/F)'$ isometrically.
>
> (c) Show that $E/F^\perp \cong F'$ if E is reflexive.

Exercise 7.3 ($\star\star$). Prove that if E' is separable, then E is separable. *Hint:* For each f_n in a dense subset of $S_{E'}$, choose x_n with $|f_n(x_n)| > \frac{1}{2}\|f_n\|$. Show that $\mathrm{span}\{x_n\}$ is dense.

Exercise 7.4 ($\star\star$). Show that a reflexive Banach space is separable if and only if its dual is separable.

Exercise 7.5 ($\star$). Let E be a Banach space and F a closed subspace. Show that E is reflexive if and only if both F and E/F are reflexive.

Exercise 7.6 ($\star\star$). (a) Show that ℓ^∞ is not separable.

  (b) Deduce that ℓ^1 is not reflexive (without using ultrafilters). *Hint:* Use Exercise 7.3.

Exercise 7.7 ($\star\star$). Let E be a Banach space and $f \in E'$ with $\|f\| = 1$. We say f *attains its norm* if there exists $x \in B_E$ with $|f(x)| = 1$.

  (a) Show that in a reflexive space, every $f \in E'$ attains its norm.

  (b) Find an explicit $f \in c_0'$ that does not attain its norm.

Exercise 7.8 ($\star\star$). Let E be a uniformly convex Banach space and $C \subset E$ a nonempty closed convex subset. Show that for every $x \in E$ there exists a unique $y \in C$ with $\|x - y\| = \inf_{z \in C} \|x - z\|$.

Exercise 7.9 ($\star\star\star$). (Day) Show that ℓ^1 admits no equivalent uniformly convex norm. *Hint:* Use the fact that ℓ^1 is not reflexive and the Milman–Pettis theorem.

Exercise 7.10 ($\star$). Let E_1, E_2 be Banach spaces and $E = E_1 \oplus_p E_2$ with the ℓ^p-norm ($1 \le p < \infty$). Show that $E' \cong E_1' \oplus_{p'} E_2'$ isometrically.

Exercise 7.11 ($\star\star$). Let E be a normed space. Show that $J(E)$ is dense in E'' if and only if E' is separating (which it always is, by

Hahn–Banach). Conclude that $J(E)$ is always dense in E'' for the weak-$*$ topology $\sigma(E'', E')$.

Exercise 7.12 $(\star\star\star)$. (James' space) Consider the space $\mathcal{J}$ of all sequences $x = (x_n)$ with $x_n \to 0$ and finite *James norm*:

$$\|x\|_{\mathcal{J}} = \sup\left\{ \left(\sum_{i=1}^{k-1}(x_{p_{i+1}} - x_{p_i})^2\right)^{1/2} : k \geq 2,\ p_1 < p_2 < \cdots < p_k\right\}.$$

(a) Verify that $\mathcal{J}$ is a Banach space.

(b) Show that $\mathcal{J}$ is isometrically isomorphic to $\mathcal{J}''$ but is not reflexive (codim $J(\mathcal{J}) = 1$ in $\mathcal{J}''$).

Chapter 8

Weak Topologies

Contents

8.1 Initial Topologies: Review

We begin with a review of the general construction that underlies both the weak topology and the weak-$*$ topology.

Definition 8.1 (Initial topology). Let X be a set and $\{(Y_i, \tau_i)\}_{i \in I}$ a family of topological spaces. For each $i \in I$, let $\varphi_i : X \to Y_i$ be a map. The **initial topology** on X with respect to the family $(\varphi_i)_{i \in I}$ is the coarsest topology on X making each φ_i continuous. A subbasis for this

topology is given by

$$\left\{ \varphi_i^{-1}(U_i) : i \in I,\ U_i \in \tau_i \right\}.$$

Proposition 8.2. *In the initial topology, a net $(x_\alpha)_{\alpha \in A}$ converges to x in X if and only if $\varphi_i(x_\alpha) \to \varphi_i(x)$ in Y_i for every $i \in I$.*

Proposition 8.3. *The initial topology is Hausdorff if and only if the family $(\varphi_i)_{i \in I}$ separates points of X: for every $x \neq y$ in X, there exists $i \in I$ with $\varphi_i(x) \neq \varphi_i(y)$.*

8.2   The Weak Topology $\sigma(E, E')$

Definition 8.4 (Weak topology)**.** Let E be a normed space. The **weak topology** on E, denoted $\sigma(E, E')$, is the initial topology on E with respect to the family of all continuous linear functionals:

$$\sigma(E, E') = \text{initial topology w.r.t. } \{f : E \to \mathbb{K}\}_{f \in E'}.$$

Proposition 8.5 (Subbasis and neighborhood base)**.** *A subbasis of open sets for $\sigma(E, E')$ is given by*

$$\left\{ f^{-1}(U) : f \in E',\ U \subset \mathbb{K}\ open \right\}.$$

A base of neighborhoods of a point $x_0 \in E$ is given by the sets

$$V(x_0; f_1, \ldots, f_n; \varepsilon) = \left\{ x \in E : |f_k(x) - f_k(x_0)| < \varepsilon,\ k = 1, \ldots, n \right\}$$

where $n \geq 1$, $f_1, \ldots, f_n \in E'$, and $\varepsilon > 0$.

Theorem 8.6. *The weak topology $\sigma(E, E')$ is Hausdorff.*

Proof. By Proposition 8.3, it suffices to show that E' separates points of E. This is exactly the content of Corollary 7.5 (a consequence of the Hahn–Banach theorem). $\qquad\square$

Proposition 8.7 (Weak topology is coarser)**.** *The weak topology $\sigma(E, E')$ is coarser than (or equal to) the norm topology. If $\dim E = \infty$, the inclusion is strict: $\sigma(E, E') \subsetneq \tau_{\|\cdot\|}$.*

Proof. Every $f \in E'$ is norm-continuous, so every weakly open set is norm-open. For the strict inclusion, note that the open unit ball $B(0, 1)$ is norm-open. If it were weakly open, it would contain a basic weak neighborhood $V(0; f_1, \ldots, f_n; \varepsilon)$. But $\bigcap_{k=1}^{n} \operatorname{Ker} f_k$ is a subspace of codimension at most n, hence of infinite dimension when $\dim E = \infty$. Any nonzero v in this intersection satisfies $tv \in V(0; f_1, \ldots, f_n; \varepsilon)$ for all t, so V contains the entire line through v, which is not contained in $B(0, 1)$. $\qquad\square$

Proposition 8.8. *If E is an infinite-dimensional normed space, the weak topology on E is not metrizable (on the whole space E).*

Proof. In a metrizable space, a point in the closure of a set A is the limit of a sequence from A. Consider the weak topology on an infinite-dimensional space. Every weak neighborhood of 0 contains a subspace of finite codimension (as shown above), hence is unbounded. In particular, B_E is not weakly open.

More formally: if the weak topology were metrizable, then since E is a topological vector space under the weak topology, it would be first countable. But a locally convex space is first countable (at the origin) if and only if its topology is generated by countably many seminorms, which for $\sigma(E, E')$ would require E' to be the span of countably many functionals. When $\dim E = \infty$, E' is infinite-dimensional (by Hahn–Banach), and in fact E' cannot be the countable union of finite-dimensional subspaces (by Baire's theorem if E' is a Banach space—which it is—then it cannot be countable-dimensional). $\qquad\square$

8.3 Weak Convergence

Definition 8.9 (Weak convergence of sequences)**.** A sequence (x_n) in a normed space E **converges weakly** to $x \in E$ if

$$f(x_n) \to f(x) \qquad \text{for every } f \in E'.$$

We write $x_n \rightharpoonup x$.

Proposition 8.10.

(i) *Norm convergence implies weak convergence: if $\|x_n - x\| \to 0$, then $x_n \rightharpoonup x$.*

(ii) *The converse is false in general (in infinite-dimensional spaces).*

(iii) *In finite-dimensional spaces, weak and norm convergence coincide.*

Proof. (i) If $\|x_n - x\| \to 0$ and $f \in E'$, then $|f(x_n) - f(x)| \leq \|f\| \, \|x_n - x\| \to 0$.

(ii) In ℓ^2, the standard basis vectors $e_n \rightharpoonup 0$ (since for any $f \in (\ell^2)' \cong \ell^2$ given by $y = (y_k)$, $f(e_n) = y_n \to 0$), but $\|e_n\| = 1 \not\to 0$.

(iii) In $\mathbb{R}^d$, the functionals $f_k(x) = x_k$ $(k = 1, \ldots, d)$ span E', so weak convergence is equivalent to coordinate-wise convergence, which is equivalent to norm convergence.   □

Example 8.11.

(i) In ℓ^p $(1 < p < \infty)$: $e_n \rightharpoonup 0$ but $\|e_n\|_p = 1$.

(ii) In $L^2([0,1])$: $\sqrt{2} \sin(2\pi n \cdot) \rightharpoonup 0$ (Riemann–Lebesgue lemma).

(iii) In ℓ^1: $e_n \rightharpoonup 0$ as well (for any $y \in \ell^\infty = (\ell^1)'$, $y_n = y(e_n)$; but since we only need $y_n \to 0$ for $y \in c_0$... actually $(\ell^1)' = \ell^\infty$ so we need $y_n \to 0$ for all $y \in \ell^\infty$, which fails for $y = (1,1,1,\ldots)$). Hence $e_n \not\rightharpoonup 0$ in ℓ^1.

(iv) In ℓ^1: the sequence $x_n = e_1 + e_2 + \cdots + e_n$ does not converge weakly (take $f = (1,1,1,\ldots) \in \ell^\infty$).

The next result is fundamental and illustrates the power of the Banach–Steinhaus theorem in the study of weak convergence.

Theorem 8.12 (Weakly convergent sequences are bounded). *Let E be a Banach space and (x_n) a sequence in E with $x_n \rightharpoonup x$. Then:*

(i) $\sup_n \|x_n\| < \infty$.

(ii) $\|x\| \leq \liminf_{n \to \infty} \|x_n\|$.

Proof. (i) Consider the family $(J(x_n))_{n\geq 1}$ in $E'' = \mathcal{L}(E', \mathbb{K})$. For each $f \in E'$, the sequence $(J(x_n)(f))_{n\geq 1} = (f(x_n))_{n\geq 1}$ converges (to $f(x)$), hence is bounded: $\sup_n |J(x_n)(f)| < \infty$. By the Banach–Steinhaus theorem (uniform boundedness principle), applied to the family of bounded linear operators $J(x_n) : E' \to \mathbb{K}$,

$$\sup_n \|J(x_n)\|_{E''} < \infty.$$

Since J is an isometry, $\|J(x_n)\| = \|x_n\|$, giving $\sup_n \|x_n\| < \infty$.
  (ii) For every $f \in E'$ with $\|f\| \leq 1$,

$$|f(x)| = \lim_n |f(x_n)| \leq \liminf_n \|f\| \, \|x_n\| \leq \liminf_n \|x_n\| .$$

Taking the supremum over $\|f\| \leq 1$ gives $\|x\| \leq \liminf_n \|x_n\|$. □

Remark 8.13. Part (ii) says that the norm is *weakly lower semicontinuous*. This is extremely useful in the calculus of variations.

Proposition 8.14. *Let E be a Banach space. If $x_n \rightharpoonup x$ and $\|x_n\| \to \|x\|$, and if E is uniformly convex, then $x_n \to x$ in norm.*

Proof. If $\|x\| = 0$, then $\|x_n\| \to 0$ and we are done. Assume $\|x\| > 0$. Let $u_n = x_n/\|x_n\|$ and $u = x/\|x\|$ (for n large enough). Then $\|u_n\| = \|u\| = 1$. Since E is uniformly convex, if $u_n \nrightarrow u$ in norm, there exist $\varepsilon > 0$ and a subsequence with $\|u_{n_k} - u\| \geq \varepsilon$, giving $\left\|\frac{u_{n_k}+u}{2}\right\| \leq 1 - \delta$.
  By Hahn–Banach, pick $f \in E'$ with $\|f\| = 1$ and $f(u) = 1$. Then $f(u_{n_k}) \to f(u) = 1$ (weak convergence and $\|x_n\| \to \|x\|$), so $f(\frac{u_{n_k}+u}{2}) \to 1$, contradicting $\left\|\frac{u_{n_k}+u}{2}\right\| \leq 1 - \delta$. □

8.4 The Weak-∗ Topology $\sigma(E', E)$

Definition 8.15 (Weak-∗ topology)**.** Let E be a normed space. The **weak-∗ topology** on E', denoted $\sigma(E', E)$, is the initial topology on E' with respect to the family of evaluation maps $\{\hat{x} : E' \to \mathbb{K}\}_{x \in E}$, where $\hat{x}(f) = f(x)$.

Remark 8.16. The weak-∗ topology on E' is coarser than or equal to the weak topology $\sigma(E', (E')') = \sigma(E', E'')$ on E'. In $\sigma(E', E)$, one only

requires convergence on evaluation functionals $\hat{x}$ for $x \in E$, whereas $\sigma(E', E'')$ requires convergence against all of E''. They coincide if and only if E is reflexive.

Definition 8.17 (Weak-$*$ convergence). A net (f_α) in E' converges to $f \in E'$ in the weak-$*$ topology if and only if $f_\alpha(x) \to f(x)$ for every $x \in E$. For sequences, we write $f_n \overset{*}{\rightharpoonup} f$.

Proposition 8.18. *The weak-$*$ topology is Hausdorff.*

Proof. The evaluation maps $\hat{x} : E' \to \mathbb{K}$ separate points of E': if $f \neq g$, there exists $x \in E$ with $f(x) \neq g(x)$, i.e. $\hat{x}(f) \neq \hat{x}(g)$. $\qquad\square$

Proposition 8.19. *A base of neighborhoods of $f_0 \in E'$ for the weak-$*$ topology is*

$$W(f_0; x_1, \ldots, x_n; \varepsilon) = \left\{ f \in E' : |f(x_k) - f_0(x_k)| < \varepsilon, \ k = 1, \ldots, n \right\},$$

where $n \geq 1$, $x_1, \ldots, x_n \in E$, $\varepsilon > 0$.

8.5   The Banach–Alaoglu Theorem

This is one of the most important compactness results in functional analysis.

Theorem 8.20 (Banach–Alaoglu, 1940)**.** *Let E be a normed space. The closed unit ball*

$$B_{E'} = \{ f \in E' : \|f\| \leq 1 \}$$

is compact in the weak-$*$ topology $\sigma(E', E)$.

Proof. For each $x \in E$, define

$$D_x = \{ \lambda \in \mathbb{K} : |\lambda| \leq \|x\| \},$$

which is a compact subset of $\mathbb{K}$. Consider the product space

$$P = \prod_{x \in E} D_x,$$

which is compact by Tychonoff's theorem. An element of P is a function $\omega : E \to \mathbb{K}$ with $|\omega(x)| \leq \|x\|$ for all x.

Define the map

$$\iota : B_{E'} \longrightarrow P, \qquad \iota(f) = (f(x))_{x \in E}.$$

This is well-defined since $|f(x)| \leq \|f\| \, \|x\| \leq \|x\|$ for $f \in B_{E'}$.

Step 1: ι is injective. If $\iota(f) = \iota(g)$, then $f(x) = g(x)$ for all $x \in E$, so $f = g$.

Step 2: ι is a homeomorphism onto its image. The product topology on P is the topology of pointwise convergence. The weak-$*$ topology on $B_{E'}$ is also the topology of pointwise convergence (convergence at each $x \in E$). Hence ι and ι^{-1} (on $\iota(B_{E'})$) are both continuous.

Step 3: $\iota(B_{E'})$ is closed in P. Let (ω_α) be a net in $\iota(B_{E'})$ converging to $\omega \in P$. Each $\omega_\alpha = \iota(f_\alpha)$ for some $f_\alpha \in B_{E'}$. Convergence in P means $f_\alpha(x) \to \omega(x)$ for every $x \in E$.

We must show ω is linear. For $x, y \in E$ and $\lambda \in \mathbb{K}$:

$$\omega(\lambda x + y) = \lim_\alpha f_\alpha(\lambda x + y) = \lim_\alpha \big[\lambda f_\alpha(x) + f_\alpha(y)\big]$$
$$= \lambda \lim_\alpha f_\alpha(x) + \lim_\alpha f_\alpha(y) = \lambda \, \omega(x) + \omega(y).$$

So ω is linear and $|\omega(x)| \leq \|x\|$ for all x (since $\omega \in P$), hence $\omega \in B_{E'}$ and $\iota(\omega) = \omega$.

Conclusion. $\iota(B_{E'})$ is a closed subset of the compact space P, hence compact. Since ι is a homeomorphism onto its image, $B_{E'}$ is weak-$*$ compact. $\quad \square$

Corollary 8.21. *For any $r > 0$, the ball $\{f \in E' : \|f\| \leq r\}$ is weak-$*$ compact.*

Proof. Apply the theorem to the rescaled space, or note that the map $f \mapsto rf$ is a weak-$*$ homeomorphism. $\quad \square$

Corollary 8.22. *If E is a separable normed space, then every bounded sequence in E' has a weak-$*$ convergent subsequence.*

Proof. Let $(f_n) \subset E'$ with $\sup_n \|f_n\| \leq M$. Then $(f_n) \subset M B_{E'}$, which is weak-$*$ compact by Corollary 8.21.

It remains to show that the weak-$*$ topology on $MB_{E'}$ is metrizable when E is separable. Let $(x_k)_{k \geq 1}$ be a dense sequence in E. Define

$$d(f, g) = \sum_{k=1}^{\infty} \frac{1}{2^k} \frac{|f(x_k) - g(x_k)|}{1 + |f(x_k) - g(x_k)|}.$$

This is a metric on E', and on bounded subsets of E', it induces the weak-$*$ topology. Indeed, if $f_n \overset{*}{\rightharpoonup} f$ in E' with $\|f_n\| \leq M$, then $f_n(x_k) \to f(x_k)$ for each k, hence $d(f_n, f) \to 0$ by dominated convergence. Conversely, if $d(f_n, f) \to 0$, then $f_n(x_k) \to f(x_k)$ for each k; by density and uniform boundedness, $f_n(x) \to f(x)$ for all $x \in E$.

Since $MB_{E'}$ is weak-$*$ compact and metrizable, it is sequentially compact. $\qquad\square$

Remark 8.23. The Banach–Alaoglu theorem can be viewed as an infinite-dimensional analogue of the Heine–Borel theorem: closed bounded sets in $\mathbb{R}^n$ are compact, and the unit ball in E' is compact— but only in the weak-$*$ topology, not the norm topology (unless E is finite-dimensional).

8.6   Goldstine's Theorem

Theorem 8.24 (Goldstine, 1938)**.** *Let E be a normed space and $J : E \to E''$ the canonical injection. Then $J(B_E)$ is weak-$*$ dense in $B_{E''}$:*

$$\overline{J(B_E)}^{\sigma(E'', E')} = B_{E''}.$$

Proof. Let $\xi \in B_{E''}$ and let $W = W(\xi; f_1, \ldots, f_n; \varepsilon)$ be a basic weak-$*$ neighborhood of ξ in E'':

$$W = \left\{ \eta \in E'' : |\eta(f_k) - \xi(f_k)| < \varepsilon, \ k = 1, \ldots, n \right\}.$$

We need to find $x \in B_E$ with $J(x) \in W$, i.e., $|f_k(x) - \xi(f_k)| < \varepsilon$ for $k = 1, \ldots, n$.

Consider the linear map $T : E \to \mathbb{K}^n$ defined by $T(x) = (f_1(x), \ldots, f_n(x))$. Set $a = (\xi(f_1), \ldots, \xi(f_n)) \in \mathbb{K}^n$. We need to show $a \in T(B_E) + B(0, \varepsilon)$ in $\mathbb{K}^n$, or equivalently that a is in the closure of $T(B_E)$ in $\mathbb{K}^n$.

Claim: $a \in \overline{T(B_E)}$ (closure in $\mathbb{K}^n$).

Since T is linear and continuous, $T(B_E)$ is a convex subset of $\mathbb{K}^n$. In finite-dimensional spaces, the closure of a convex set equals its weak closure, so it suffices to show $a \in \overline{T(B_E)}$.

Suppose not. Then by the Hahn–Banach separation theorem in $\mathbb{K}^n$ (a finite-dimensional space), there exists a linear functional $\ell : \mathbb{K}^n \to \mathbb{R}$ and $\gamma \in \mathbb{R}$ with $\ell(a) > \gamma \geq \ell(T(x))$ for all $x \in B_E$.

Writing $\ell(\lambda_1, \ldots, \lambda_n) = \mathrm{Re} \sum_{k=1}^n \alpha_k \lambda_k$ for some $\alpha_k \in \mathbb{K}$, and setting $g = \sum_{k=1}^n \alpha_k f_k \in E'$, we get

$$\mathrm{Re}\ g(x) = \ell(T(x)) \leq \gamma < \ell(a) = \mathrm{Re} \sum_k \alpha_k \xi(f_k) = \mathrm{Re}\ \xi(g)$$

for all $x \in B_E$. Taking the supremum over $x \in B_E$:

$$\|g\| = \sup_{\|x\| \leq 1} \mathrm{Re}\ g(x) \leq \gamma < \mathrm{Re}\ \xi(g) \leq |\xi(g)| \leq \|\xi\|\ \|g\| \leq \|g\|,$$

a contradiction. Hence $a \in \overline{T(B_E)}$, and we can find $x \in B_E$ with $T(x)$ arbitrarily close to a, which means $J(x) \in W$. $\square$

> **Corollary 8.25.** $J(E)$ *is weak-$*$ dense in E''.*

Proof. For any $\xi \in E''$ with $\|\xi\| = r > 0$, applying Goldstine's theorem to $\xi/r \in B_{E''}$ and scaling back gives $\xi \in \overline{J(rB_E)}^{w^*} \subset \overline{J(E)}^{w^*}$. $\square$

8.7 Kakutani's Theorem

> **Theorem 8.26** (Kakutani, 1938)**.** *A Banach space E is reflexive if and only if B_E is weakly compact.*

Proof. ($\Rightarrow$) Assume E is reflexive, so $J : E \to E''$ is a surjective isometry. Under J, B_E is mapped onto $B_{E''}$. By the Banach–Alaoglu theorem (Theorem 8.20), $B_{E''}$ is weak-$*$ compact in $\sigma(E'', E')$. Since J is surjective, the weak topology $\sigma(E, E')$ on E corresponds (via J) to the weak-$*$ topology $\sigma(E'', E')$ on E''. More precisely, for a net (x_α) in E:

$$
\begin{aligned}
x_\alpha \xrightarrow{\sigma(E,E')} x \iff & f(x_\alpha) \to f(x)\ \forall f \in E' \\
\iff & J(x_\alpha)(f) \to J(x)(f)\ \forall f \in E' \\
\iff & J(x_\alpha) \xrightarrow{\sigma(E'',E')} J(x).
\end{aligned}
$$

So $J : (E, \sigma(E, E')) \to (E'', \sigma(E'', E'))$ is a homeomorphism, and $J(B_E) = B_{E''}$ is $\sigma(E'', E')$-compact. Hence B_E is $\sigma(E, E')$-compact.

($\Leftarrow$) Assume B_E is weakly compact. Then $J(B_E)$ is compact in $\sigma(E'', E')$ (since J is a weak-to-weak-$*$ homeomorphism onto its image, as shown above). In particular, $J(B_E)$ is $\sigma(E'', E')$-closed.

By Goldstine's theorem, $J(B_E)$ is $\sigma(E'', E')$-dense in $B_{E''}$. Since $J(B_E)$ is also $\sigma(E'', E')$-closed, we conclude $J(B_E) = B_{E''}$. This implies $J(E) = E''$ (since every $\xi \in E''$ with $\|\xi\| = r$ satisfies $\xi/r \in B_{E''} = J(B_E)$), so E is reflexive. $\qquad\square$

> *Remark* 8.27. Kakutani's theorem converts the algebraic condition of reflexivity (J surjective) into a topological condition (weak compactness of B_E). This is extremely useful because compactness arguments (extracting convergent subsequences, attaining suprema of continuous functions, etc.) become available.

8.8   Mazur's Theorem

> **Theorem 8.28** (Mazur, 1933). *Let E be a normed space and $C \subset E$ a convex set. Then C is norm-closed if and only if C is weakly closed. Equivalently,*
> $$\overline{C}^{\,\|\cdot\|} = \overline{C}^{\,\sigma(E,E')}.$$

Proof. Since the weak topology is coarser than the norm topology, every weakly closed set is norm-closed. So we only need to show: if C is convex and norm-closed, then C is weakly closed.

Let $x_0 \notin C$. We show there is a weakly open set containing x_0 that is disjoint from C. Since C is norm-closed and convex, by the geometric Hahn–Banach theorem (strict separation of a point from a closed convex set), there exist $f \in E'$ and $\gamma \in \mathbb{R}$ such that

$$\operatorname{Re} f(x_0) > \gamma > \sup_{c \in C} \operatorname{Re} f(c).$$

The set $U = \{x \in E : \operatorname{Re} f(x) > \gamma\}$ is weakly open (since f is weakly continuous), contains x_0, and satisfies $U \cap C = \emptyset$. Hence C^c is weakly open, so C is weakly closed.

For the equivalent formulation: the norm closure $\overline{C}$ is convex (closure of a convex set is convex) and norm-closed, hence weakly closed. So $\overline{C}^w \subset \overline{C}$. The reverse inclusion $\overline{C} \subset \overline{C}^w$ holds because the weak topology is coarser. $\qquad\square$

Corollary 8.29 (Weakly convergent sequences and convex combinations)**.** *Let E be a Banach space and $x_n \rightharpoonup x$. Then x is in the norm-closed convex hull of $\{x_n : n \geq 1\}$. In other words, there exist convex combinations $y_k = \sum_{n \in F_k} \lambda_n^{(k)} x_n$ (with F_k finite, $\lambda_n^{(k)} \geq 0$, $\sum \lambda_n^{(k)} = 1$) such that $y_k \to x$ in norm.*

Proof. Let $C = \text{conv}\{x_n : n \geq 1\}$. Since $x_n \rightharpoonup x$, the point x belongs to the weak closure of $\{x_n\}$, hence to the weak closure of C. By Mazur's theorem, $\overline{C}^w = \overline{C}^{\|\cdot\|}$, so $x \in \overline{C}^{\|\cdot\|} = \overline{\text{conv}\{x_n\}}^{\|\cdot\|}$. $\qquad\square$

Remark 8.30. Mazur's theorem is the key bridge between weak and norm topologies for convex sets. It says that for convex sets, the distinction between weak and norm topologies disappears. This is not true for non-convex sets: for instance, the unit sphere $S_E = \{x : \|x\| = 1\}$ is norm-closed but weakly dense in B_E when $\dim E = \infty$ (a consequence of the fact that no weakly open set is norm-bounded).

8.9 Application: Existence of Minimizers

Theorem 8.31 (Direct method of the calculus of variations)**.** *Let E be a reflexive Banach space, $C \subset E$ a nonempty, bounded, closed, and convex subset, and $\Phi : C \to \mathbb{R}$ a function that is sequentially weakly lower semicontinuous (i.e., $x_n \rightharpoonup x$ implies $\Phi(x) \leq \liminf_n \Phi(x_n)$). Then Φ attains its minimum on C: there exists $x_0 \in C$ with $\Phi(x_0) = \inf_{x \in C} \Phi(x)$.*

Proof. Let $m = \inf_{x \in C} \Phi(x) \in [-\infty, +\infty)$. Choose a minimizing sequence $(x_n) \subset C$ with $\Phi(x_n) \to m$.

Since E is reflexive and C is bounded, closed, and convex, C is weakly compact by Kakutani's theorem (Theorem 8.26) and the fact that a closed convex subset of a weakly compact set is weakly compact (closed in the weak topology by Mazur's theorem, hence weakly compact as a weakly closed subset of the weakly compact set $\overline{B(0,R)}^w$ for some R large enough).

By the Eberlein–Šmulian theorem (Theorem 7.25), (x_n) has a subsequence (x_{n_k}) with $x_{n_k} \rightharpoonup x_0$ for some $x_0 \in C$ (since C is weakly closed, hence weakly sequentially closed).

By weak lower semicontinuity,

$$\Phi(x_0) \leq \liminf_{k \to \infty} \Phi(x_{n_k}) = m.$$

Since $x_0 \in C$, $\Phi(x_0) \geq m$, hence $\Phi(x_0) = m$.   □

Example 8.32. Let E be a reflexive Banach space and $C \subset E$ a
nonempty, closed, convex subset. For any $y \in E$, the function $\Phi(x) = \|x - y\|$ is weakly lower semicontinuous (Theorem 8.12(ii)). If C is
bounded, the direct method gives a minimizer. Even if C is unbounded,
one can restrict to $C \cap \overline{B}(y, r)$ for $r = \inf_{c \in C} \|c - y\| + 1$, and the direct
method still applies. Hence every nonempty closed convex subset of a
reflexive Banach space has a *nearest point* to any given y.

Example 8.33 (Minimization of an energy functional). Let $\Omega \subset \mathbb{R}^d$
be a bounded open set and consider the Sobolev space $H_0^1(\Omega)$, which
is a reflexive Banach space (in fact a Hilbert space). For $f \in L^2(\Omega)$,
define the energy functional

$$\mathcal{E}(u) = \frac{1}{2} \int_\Omega |\nabla u|^2 \, dx - \int_\Omega f u \, dx.$$

Then $\mathcal{E}$ is weakly lower semicontinuous and coercive on $H_0^1(\Omega)$ (by the
Poincaré inequality, $\mathcal{E}(u) \to +\infty$ as $\|u\|_{H_0^1} \to \infty$), so $\mathcal{E}$ attains its
minimum. The minimizer is the weak solution of $-\Delta u = f$.

8.10   Additional Results

Proposition 8.34 (Characterization of weak-$*$ continuous functionals). *A linear functional* $\Phi : E' \to \mathbb{K}$ *is* $\sigma(E', E)$-*continuous if and only
if there exists* $x \in E$ *such that* $\Phi(f) = f(x)$ *for all* $f \in E'$.

Proof. The "if" direction is clear: $\Phi = \hat{x}$ is weak-$*$ continuous by definition.

For the "only if" direction: if Φ is $\sigma(E', E)$-continuous, there exist $x_1, \ldots, x_n \in E$ and $\varepsilon > 0$ such that $|\Phi(f)| < 1$ whenever $|f(x_k)| < \varepsilon$ for $k = 1, \ldots, n$. This
implies $\operatorname{Ker}\Phi \supset \bigcap_{k=1}^n \operatorname{Ker}\hat{x}_k$. By a standard lemma on linear functionals,
$\Phi \in \operatorname{span}\{\hat{x}_1, \ldots, \hat{x}_n\}$, i.e., $\Phi = \sum_{k=1}^n \alpha_k \hat{x}_k = \widehat{\sum \alpha_k x_k}$. Setting $x = \sum \alpha_k x_k$
gives $\Phi(f) = f(x)$.   □

Proposition 8.35 (Weak-∗ compactness and reflexivity). *Let E be a Banach space. Then B_E is weak-∗ compact in E'' (i.e., $J(B_E)$ is $\sigma(E'', E')$-compact) if and only if E is reflexive.*

Proof. If E is reflexive, $J(B_E) = B_{E''}$ is weak-∗ compact by Banach–Alaoglu. Conversely, if $J(B_E)$ is $\sigma(E'', E')$-compact, then it is $\sigma(E'', E')$-closed. By Goldstine's theorem, $J(B_E)$ is dense in $B_{E''}$, so $J(B_E) = B_{E''}$, giving reflexivity. $\square$

Theorem 8.36 (Weak-∗ sequential compactness in separable duals). *If E is a separable Banach space, then every bounded sequence in E' has a weak-∗ convergent subsequence.*

Proof. This was established in Corollary 8.22: when E is separable, the weak-∗ topology on bounded subsets of E' is metrizable, and bounded subsets are relatively weak-∗ compact by Banach–Alaoglu, hence sequentially compact. $\square$

Proposition 8.37 (Weak-∗ closed subspaces). *A subspace $F \subset E'$ is $\sigma(E', E)$-closed if and only if $F = G^\perp$ for some subset $G \subset E$, where $G^\perp = \{f \in E' : f(x) = 0 \ \forall x \in G\}$.*

Proof. ($\Leftarrow$) $G^\perp = \bigcap_{x \in G} \operatorname{Ker} \hat{x}$, which is an intersection of weak-∗ closed sets, hence weak-∗ closed.

($\Rightarrow$) Let F be a weak-∗ closed subspace. Set $G = {}^\perp F = \{x \in E : f(x) = 0 \ \forall f \in F\}$. Then $F \subset G^\perp$ and we claim $F = G^\perp$. If $g \in G^\perp \setminus F$, then by the Hahn–Banach theorem for the locally convex space $(E', \sigma(E', E))$, there exists a $\sigma(E', E)$-continuous functional Φ with $\Phi\big|_F = 0$ and $\Phi(g) \neq 0$. By Proposition 8.34, $\Phi = \hat{x}$ for some $x \in E$. Then $f(x) = 0$ for all $f \in F$, so $x \in {}^\perp F = G$, giving $g(x) = 0$ (since $g \in G^\perp$), i.e. $\Phi(g) = 0$, a contradiction. $\square$

8.11 Exercises

Exercise 8.1 ($\star$). Show that $e_n \rightharpoonup 0$ in ℓ^p for $1 < p < \infty$ but $e_n \not\rightharpoonup 0$ in ℓ^1.

Exercise 8.2 ($\star$)**.** Let E be a normed space.

(a) Show that the weak topology is the coarsest topology making every $f \in E'$ continuous.

(b) Show that a linear map $T : E \to F$ between normed spaces is weakly continuous (continuous from $(E, \sigma(E, E'))$ to $(F, \sigma(F, F'))$) if and only if it is norm-continuous.

Exercise 8.3 ($\star\star$)**.** (This exercise shows that in non-separable spaces, weak sequential closure may differ from weak closure.) Let $E = \ell^\infty$ and consider the set $A = \{e_n : n \geq 1\}$.

(a) Show that 0 is in the weak closure of A.

(b) Show that no subsequence of (e_n) converges weakly to 0 in ℓ^∞.

Hint for (b): Use the fact that $(\ell^\infty)'$ is much larger than ℓ^1.

Exercise 8.4 ($\star$)**.**   (a) In $\ell^\infty = (\ell^1)'$, show that $e_n \overset{*}{\rightharpoonup} 0$.

(b) In $L^\infty([0, 1]) = (L^1([0, 1]))'$, show that $\sin(2\pi n \cdot) \overset{*}{\rightharpoonup} 0$.

Exercise 8.5 ($\star\star$)**.** Show that the closed unit ball of an infinite-dimensional Banach space is **not** compact in the norm topology. *Hint:* Construct an infinite sequence with no convergent subsequence (e.g., using Riesz's lemma).

Exercise 8.6 ($\star\star$)**.** Let (x_n) be a sequence in a Banach space E with $x_n \rightharpoonup x$.

(a) Use Mazur's theorem to show that there exist convex combinations $y_N = \sum_{n=N}^{M_N} \lambda_n x_n$ (with $\lambda_n \geq 0$, $\sum \lambda_n = 1$) such that $\|y_N - x\| \to 0$.

(b) Give an explicit construction when $E = L^2([0, 1])$ and $x_n(t) = \sqrt{2} \sin(2\pi n t)$.

Exercise 8.7 ($\star\star$). Let E be a Banach space. Show that the following are equivalent:

  (i) E is reflexive.

  (ii) The weak and weak-$*$ topologies on E' coincide: $\sigma(E', E) = \sigma(E', E'')$.

  (iii) $B_{E'}$ is weakly compact (in the weak topology $\sigma(E', (E')')$).

Exercise 8.8 ($\star\star\star$). (Schur's theorem) Show that in ℓ^1, weak convergence implies norm convergence: if $x_n \rightharpoonup x$ in ℓ^1, then $\|x_n - x\|_1 \to 0$. *Hint:* Argue by contradiction. If $\|x_n - x\|_1 \not\to 0$, pass to a subsequence and use a gliding hump argument to construct $f \in \ell^\infty = (\ell^1)'$ with $f(x_n) \not\to f(x)$.

Exercise 8.9 ($\star\star$). Use Kakutani's theorem to give another proof that c_0 is not reflexive. *Hint:* Show that the sequence $(e_1 + e_2 + \cdots + e_n)/n$ is in B_{c_0} but has no weakly convergent subsequence, using Schur's theorem and the identification $c_0'' \cong \ell^\infty$.

Exercise 8.10 ($\star\star$). Let E be a Banach space and $f \in E'$. Use Goldstine's theorem to show that $\|f\|_{E'} = \sup\{|f(x)| : x \in B_E\}$ can also be written as $\|f\|_{E'} = \sup\{|\xi(f)| : \xi \in B_{E''}\}$.

Exercise 8.11 ($\star\star$). Show that every weakly compact subset of a normed space is bounded. *Hint:* Use the uniform boundedness principle.

Exercise 8.12 ($\star\star\star$). (Dunford–Pettis criterion) Let (Ω, μ) be a finite measure space and $K \subset L^1(\mu)$. State (without proof) the Dunford–Pettis criterion for K to be relatively weakly compact. Verify it for $K = \{f_n\}$ where $f_n = n\,\mathbf{1}_{[0,1/n^2]}$ in $L^1([0,1])$.

Exercise 8.13 ($\star$). Give an example of a subset of a Banach space that is norm-closed but not weakly closed. *Hint:* Consider the unit

sphere in an infinite-dimensional space.

Exercise 8.14 ($\star\star$)**.** Let E be a Banach space and $\Phi : E \to \mathbb{R} \cup \{+\infty\}$ a convex, lower semicontinuous (for the norm topology) function. Show that Φ is weakly lower semicontinuous. *Hint:* The sublevel sets $\{x : \Phi(x) \leq \alpha\}$ are convex and norm-closed; apply Mazur's theorem.

Exercise 8.15 ($\star\star\star$)**.** Let $1 < p < \infty$, $\Omega \subset \mathbb{R}^d$ bounded and open, $g \in L^{p'}(\Omega)$. Consider the problem of minimizing

$$\Phi(u) = \frac{1}{p} \int_\Omega |u|^p \, \mathrm{d}x - \int_\Omega gu \, \mathrm{d}x$$

over $u \in L^p(\Omega)$.

(a) Show that Φ is weakly lower semicontinuous and coercive.

(b) Prove that a minimizer exists and is unique.

(c) Find the minimizer explicitly.

Exercise 8.16 ($\star\star$)**.** Let E, F be Banach spaces and $T \in \mathcal{L}(E, F)$. Show that T is compact if and only if $x_n \rightharpoonup 0$ in E implies $Tx_n \to 0$ in norm in F. *Hint:* Use the Eberlein–Šmulian theorem and the fact that T maps weakly convergent sequences to weakly convergent sequences.

Exercise 8.17 ($\star$)**.** Let E be a Banach space.

(a) Show that E is isometrically embedded in E'', which is isometrically embedded in $E^{(4)} = (E'')''$, etc.

(b) If E is reflexive, show that all iterated duals $E^{(n)}$ are reflexive and canonically isomorphic to E (for n even) or E' (for n odd).

(c) If E is not reflexive, show that $\dim E^{(2n)} < \dim E^{(2n+2)}$ for all $n \geq 0$ (in the sense that the canonical injection is not surjective).

Chapter 9

Spectral Theory of Compact Operators

Contents

9.1 Compact Operators: Review and Examples

We begin with a systematic study of the spectral properties of compact operators, building on the foundational Banach and Hilbert space theory developed in earlier chapters. The class of compact operators is arguably the most tractable infinite-dimensional analogue of matrices, and their spectral theory mirrors the finite-dimensional situation remarkably closely.

> **Definition 9.1** (Compact operator)**.** Let E and F be Banach spaces. A bounded linear operator $T\colon E \to F$ is called *compact* (or *completely*

continuous) if for every bounded sequence $(x_n)_{n \geq 1}$ in E, the sequence $(Tx_n)_{n \geq 1}$ has a convergent subsequence in F.

Equivalently, T is compact if and only if the image $T(B_E)$ of the closed unit ball $B_E = \{x \in E : \|x\| \leq 1\}$ is relatively compact in F (i.e., $\overline{T(B_E)}$ is compact).

We denote by $\mathcal{K}(E, F)$ the set of all compact operators from E to F, and write $\mathcal{K}(E) = \mathcal{K}(E, E)$.

Example 9.2 (Finite-rank operators). An operator $T \colon E \to F$ is of *finite rank* if $\dim(\operatorname{im} T) < \infty$. Every finite-rank bounded operator is compact: if (x_n) is bounded in E, then (Tx_n) lies in the finite-dimensional subspace $\operatorname{im} T$, and by the Bolzano–Weierstrass theorem it admits a convergent subsequence.

Example 9.3 (Integral operators). Let $\Omega \subset \mathbb{R}^d$ be a bounded measurable set and let $K \in L^2(\Omega \times \Omega)$. The operator $T_K \colon L^2(\Omega) \to L^2(\Omega)$ defined by

$$(T_K f)(x) = \int_\Omega K(x, y) \, f(y) \, \mathrm{d}y$$

is compact. Indeed, T_K is a Hilbert–Schmidt operator with $\|T_K\|_{\mathrm{HS}} = \|K\|_{L^2(\Omega \times \Omega)}$, and every Hilbert–Schmidt operator is compact (see §9.6).

Example 9.4 (Diagonal operators). Let $H = \ell^2(\mathbb{N})$ and let $(\lambda_n)_{n \geq 1}$ be a bounded sequence of scalars. Define $T \colon H \to H$ by $T(e_n) = \lambda_n e_n$, where (e_n) is the standard orthonormal basis. Then T is compact if and only if $\lambda_n \to 0$ as $n \to \infty$.

Example 9.5 (The inclusion $\ell^1 \hookrightarrow \ell^2$ is not compact). The natural inclusion $\iota \colon \ell^1(\mathbb{N}) \hookrightarrow \ell^2(\mathbb{N})$ is bounded (with $\|\iota\| = 1$) but not compact. The standard basis vectors e_n satisfy $\|e_n\|_{\ell^1} = 1$ and $\|e_m - e_n\|_{\ell^2} = \sqrt{2}$ for $m \neq n$, so (e_n) has no convergent subsequence in ℓ^2.

9.2 The Space of Compact Operators

Proposition 9.6. *Let E, F, and G be Banach spaces.*

(i) $\mathcal{K}(E,F)$ is a closed linear subspace of $\mathcal{L}(E,F)$.

(ii) If $T \in \mathcal{K}(E,F)$ and $S \in \mathcal{L}(F,G)$, then $ST \in \mathcal{K}(E,G)$.

(iii) If $T \in \mathcal{L}(E,F)$ and $S \in \mathcal{K}(F,G)$, then $ST \in \mathcal{K}(E,G)$.

In particular, $\mathcal{K}(E)$ is a closed two-sided ideal in the Banach algebra $\mathcal{L}(E)$.

Proof. Linearity of $\mathcal{K}(E,F)$ and properties (ii) and (iii) are straightforward exercises. We prove that $\mathcal{K}(E,F)$ is closed.

Let $(T_n)_{n \geq 1} \subset \mathcal{K}(E,F)$ with $T_n \to T$ in $\mathcal{L}(E,F)$. We show T is compact. Let $(x_k)_{k \geq 1}$ be a bounded sequence in E with $\|x_k\| \leq M$ for all k.

We use a diagonal argument. Since T_1 is compact, extract a subsequence $(x_k^{(1)})$ of (x_k) such that $(T_1 x_k^{(1)})$ converges. From $(x_k^{(1)})$, extract a further subsequence $(x_k^{(2)})$ such that $(T_2 x_k^{(2)})$ converges, and so on. The diagonal sequence $y_k := x_k^{(k)}$ satisfies: $(T_n y_k)_{k \geq 1}$ converges for every $n \geq 1$.

We claim $(T y_k)$ is Cauchy. Given $\varepsilon > 0$, choose n so that $\|T - T_n\| < \varepsilon/(3M)$. Since $(T_n y_k)$ converges, there exists K such that for all $j, k \geq K$, $\|T_n y_j - T_n y_k\| < \varepsilon/3$. Then for $j, k \geq K$:

$$
\begin{aligned}
\|T y_j - T y_k\| &\leq \|T y_j - T_n y_j\| + \|T_n y_j - T_n y_k\| + \|T_n y_k - T y_k\| \\
&\leq \|T - T_n\| \, \|y_j\| + \frac{\varepsilon}{3} + \|T_n - T\| \, \|y_k\| \\
&< \frac{\varepsilon}{3M} \cdot M + \frac{\varepsilon}{3} + \frac{\varepsilon}{3M} \cdot M = \varepsilon.
\end{aligned}
$$

Since F is complete, $(T y_k)$ converges, so T is compact. $\square$

Remark 9.7. A Banach space E has the *approximation property* if every compact operator $T \in \mathcal{K}(E)$ is the operator-norm limit of finite-rank operators. All Hilbert spaces and all L^p spaces ($1 \leq p \leq \infty$) have the approximation property. Enflo (1973) constructed a separable Banach space failing this property.

9.3   The Fredholm Alternative

The Fredholm alternative is the cornerstone result connecting the solvability of equations involving compact operators to a finite-dimensional eigenvalue problem.

> **Theorem 9.8** (Fredholm alternative). *Let E be a Banach space and $T \in \mathcal{K}(E)$. Exactly one of the following holds:*
>
> *(a) The operator $\mathrm{Id} - T$ is bijective, $(\mathrm{Id} - T)^{-1} \in \mathcal{L}(E)$, and for every $y \in E$ the equation $x - Tx = y$ has a unique solution.*
>
> *(b) $\mathrm{Ker}(\mathrm{Id} - T) \neq \{0\}$, i.e., $\lambda = 1$ is an eigenvalue of T. In this case, $\dim \mathrm{Ker}(\mathrm{Id} - T) < \infty$, $\mathrm{im}(\mathrm{Id} - T)$ is closed with $\mathrm{codim}\, \mathrm{im}(\mathrm{Id} - T) = \dim \mathrm{Ker}(\mathrm{Id} - T^*) = \dim \mathrm{Ker}(\mathrm{Id} - T) \geq 1$.*

We prepare the proof with several lemmas.

> **Lemma 9.9.** *If $T \in \mathcal{K}(E)$, then $\mathrm{Ker}(\mathrm{Id} - T)$ is finite-dimensional.*

Proof. Let $N = \mathrm{Ker}(\mathrm{Id} - T)$. On N we have $T|_N = \mathrm{Id}_N$, so T maps the unit ball $B_N = B_E \cap N$ onto itself. Since $T(B_N)$ is relatively compact and $T(B_N) = B_N$, the closed unit ball of N is compact. By Riesz's theorem, $\dim N < \infty$. $\qquad\square$

> **Lemma 9.10.** *If $T \in \mathcal{K}(E)$, then $\mathrm{im}(\mathrm{Id} - T)$ is closed in E.*

Proof. Let $y_n = (\mathrm{Id} - T)x_n \to y$. We must find x with $(\mathrm{Id} - T)x = y$. Let $N = \mathrm{Ker}(\mathrm{Id} - T)$ and let $d_n = d(x_n, N)$. Since N is finite-dimensional (hence closed), choose $z_n \in N$ with $\|x_n - z_n\| \leq d_n + 1/n$ and set $\tilde{x}_n = x_n - z_n$. Then $(\mathrm{Id} - T)\tilde{x}_n = y_n$.

   Claim: $(\tilde{x}_n)$ is bounded. Suppose not; then (passing to a subsequence) $\|\tilde{x}_n\| \to \infty$. Set $u_n = \tilde{x}_n / \|\tilde{x}_n\|$. Then $(\mathrm{Id} - T)u_n = y_n / \|\tilde{x}_n\| \to 0$. Since T is compact and (u_n) is bounded, a subsequence satisfies $Tu_{n_k} \to w$. Hence $u_{n_k} = (\mathrm{Id} - T)u_{n_k} + Tu_{n_k} \to 0 + w = w$, so $\|w\| = 1$ and $(\mathrm{Id} - T)w = 0$, i.e., $w \in N$. But

$$d(u_n, N) = \frac{d(\tilde{x}_n, N)}{\|\tilde{x}_n\|} = \frac{d_n}{\|\tilde{x}_n\|} \geq \frac{d_n}{d_n + 1/n} \to 1$$

(since $d_n \geq \|\tilde{x}_n\| - \|0\| \to \infty$ when $\|\tilde{x}_n\| \to \infty$ and $\tilde{x}_n$ has distance d_n to N with $\|\tilde{x}_n\| \leq d_n + 1/n$). Yet $u_{n_k} \to w \in N$ gives $d(u_{n_k}, N) \to 0$, a contradiction. So $(\tilde{x}_n)$ is bounded.

Since $(\tilde{x}_n)$ is bounded and T is compact, passing to a subsequence, $T\tilde{x}_{n_k} \to v$. Then $\tilde{x}_{n_k} = y_{n_k} + T\tilde{x}_{n_k} \to y + v =: x$, and $(\mathrm{Id} - T)x = y$. $\square$

Lemma 9.11 (Riesz's lemma on ascending chains). *Let $T \in \mathcal{K}(E)$ and define $N_k = \mathrm{Ker}(\mathrm{Id} - T)^k$ for $k \geq 0$. Then $N_0 \subset N_1 \subset N_2 \subset \cdots$ and there exists $p \geq 0$ such that $N_p = N_{p+1} = N_{p+2} = \cdots$*

Proof. The inclusions are clear. Suppose for contradiction that $N_k \subsetneq N_{k+1}$ for all k. By Riesz's lemma, for each $k \geq 1$ there exists $x_k \in N_k$ with $\|x_k\| = 1$ and $d(x_k, N_{k-1}) \geq 1/2$. For $j < k$, one checks that

$$Tx_k - Tx_j = x_k - \left[(\mathrm{Id} - T)x_k + x_j - (\mathrm{Id} - T)x_j\right],$$

and the bracketed term lies in N_{k-1}. Hence $\|Tx_k - Tx_j\| \geq d(x_k, N_{k-1}) \geq 1/2$, so (Tx_k) has no convergent subsequence, contradicting compactness. $\square$

Proof of Theorem 9.8. Finite-dimensionality of $\mathrm{Ker}(\mathrm{Id} - T)$ is Lemma 9.9, and closedness of $\mathrm{im}(\mathrm{Id} - T)$ is Lemma 9.10.

Claim: If $\mathrm{Id} - T$ is injective, then it is surjective.

Set $R_n = \mathrm{im}(\mathrm{Id} - T)^n$. Then $E = R_0 \supset R_1 \supset R_2 \supset \cdots$ and each R_n is closed. If $\mathrm{Id} - T$ is not surjective, then $R_0 \supsetneq R_1$, and injectivity forces $R_n \supsetneq R_{n+1}$ for all n. By Riesz's lemma, pick $x_n \in R_n$ with $\|x_n\| = 1$ and $d(x_n, R_{n+1}) \geq 1/2$. Then for $m > n$, $Tx_n - Tx_m = x_n - [(\mathrm{Id} - T)x_n + x_m - (\mathrm{Id} - T)x_m]$ where the bracket lies in R_{n+1}, so $\|Tx_n - Tx_m\| \geq 1/2$. This contradicts compactness.

Index zero: When E is a Hilbert space H, note that T^* is also compact. Since $\mathrm{im}(\mathrm{Id} - T)$ is closed, $\mathrm{im}(\mathrm{Id} - T)^\perp = \mathrm{Ker}(\mathrm{Id} - T^*)$. Hence $\mathrm{codim}\,\mathrm{im}(\mathrm{Id} - T) = \dim \mathrm{Ker}(\mathrm{Id} - T^*)$. The Fredholm index $\mathrm{ind}(\mathrm{Id} - T) = \dim \mathrm{Ker}(\mathrm{Id} - T) - \dim \mathrm{Ker}(\mathrm{Id} - T^*)$ equals zero; this follows by a homotopy argument: the map $t \mapsto \mathrm{ind}(\mathrm{Id} - tT)$ is continuous and integer-valued (hence constant) for $t \in [0, 1]$, and at $t = 0$ the index is 0.

For general Banach spaces, one uses the dual $T^* \in \mathcal{K}(E^*)$ and the Hahn–Banach theorem to establish $\mathrm{im}(\mathrm{Id} - T) = {}^\perp\mathrm{Ker}(\mathrm{Id} - T^*)$ and the analogous index computation. $\square$

Corollary 9.12. *Let $T \in \mathcal{K}(E)$ and $\lambda \neq 0$. Then either $\lambda\,\mathrm{Id} - T$ is bijective (and $(\lambda\,\mathrm{Id} - T)^{-1} \in \mathcal{L}(E)$), or λ is an eigenvalue of T with*

> *finite-dimensional eigenspace. The equation $\lambda x - Tx = y$ is solvable for all y if and only if $\lambda x = Tx$ implies $x = 0$.*

Proof. Apply Theorem 9.8 to the compact operator T/λ. $\qquad\square$

9.4   Riesz–Schauder Theory

> **Theorem 9.13** (Riesz–Schauder). *Let E be an infinite-dimensional Banach space and $T \in \mathcal{K}(E)$. Then:*
>
> *(i) $0 \in \sigma(T)$.*
>
> *(ii) $\sigma(T) \setminus \{0\} = \sigma_p(T) \setminus \{0\}$: every nonzero spectral value is an eigenvalue.*
>
> *(iii) Each nonzero eigenvalue has finite multiplicity.*
>
> *(iv) $\sigma(T)$ is at most countable, with 0 as the only possible accumulation point.*
>
> *(v) For every $\varepsilon > 0$, only finitely many eigenvalues satisfy $|\lambda| \geq \varepsilon$.*

Proof. **(i)** If T were invertible, $\mathrm{Id} = T^{-1}T$ would be compact (as a composition of a bounded with a compact operator), contradicting infinite-dimensionality of E.

(ii) If $\lambda \neq 0$ and $\lambda \in \sigma(T)$, then $\lambda\,\mathrm{Id} - T$ is not bijective. By Corollary 9.12, it is not injective, so $\lambda \in \sigma_p(T)$.

(iii) Follows from Lemma 9.9 applied to T/λ.

(iv)–(v) Suppose there are infinitely many distinct eigenvalues $(\lambda_n)_{n \geq 1}$ with $|\lambda_n| \geq \varepsilon > 0$. Let x_n be a corresponding unit eigenvector and $E_n = \mathrm{span}\{x_1, \ldots, x_n\}$. Since eigenvalues are distinct, the eigenvectors are linearly independent and $E_{n-1} \subsetneq E_n$. By Riesz's lemma, there exist $y_n \in E_n$ with $\|y_n\| = 1$ and $d(y_n, E_{n-1}) \geq 1/2$.

Now $Ty_n = \lambda_n y_n + w_n$ where $w_n \in E_{n-1}$ (since T maps E_n into E_n and the "leading term" is $\lambda_n y_n$). For $m < n$:

$$Ty_n - Ty_m = \lambda_n y_n + (\text{element of } E_{n-1}),$$

hence $\|Ty_n - Ty_m\| \geq |\lambda_n|\, d(y_n, E_{n-1}) \geq \varepsilon/2 > 0$. Thus (Ty_n) has no convergent subsequence, contradicting compactness. $\qquad\square$

9.5 Spectral Theorem for Compact Self-Adjoint Operators

We now specialize to Hilbert spaces and self-adjoint compact operators.

Theorem 9.14 (Spectral theorem for compact self-adjoint operators).
Let H be a separable Hilbert space and $T \in \mathcal{K}(H)$ self-adjoint. Then there exists an orthonormal basis $(e_n)_{n \geq 1}$ of H consisting of eigenvectors of T:

$$Te_n = \lambda_n e_n, \qquad \lambda_n \in \mathbb{R}, \qquad \lambda_n \to 0.$$

For every $x \in H$:

$$Tx = \sum_{n=1}^{\infty} \lambda_n \langle x, e_n \rangle \, e_n.$$

The eigenvalues are real, and eigenspaces for distinct eigenvalues are orthogonal.

Lemma 9.15. *If $T \in \mathcal{L}(H)$ is self-adjoint, then $\|T\| = \sup_{\|x\|=1} |\langle Tx, x \rangle|$. If additionally T is compact and $T \neq 0$, then either $\|T\|$ or $-\|T\|$ is an eigenvalue of T.*

Proof. Set $m = \sup_{\|x\|=1} |\langle Tx, x \rangle|$. Clearly $m \leq \|T\|$. For the reverse, the polarization identity gives

$$4 \operatorname{Re} \langle Tx, y \rangle = \langle T(x+y), x+y \rangle - \langle T(x-y), x-y \rangle.$$

Taking $\|x\| = \|y\| = 1$ and using $|\langle Tu, u \rangle| \leq m \|u\|^2$ yields $|\operatorname{Re} \langle Tx, y \rangle| \leq m$ after optimizing over the phase of y, hence $\|T\| \leq m$.

Now let T be compact with $T \neq 0$. Choose (x_n) with $\|x_n\| = 1$ and $\langle Tx_n, x_n \rangle \to \mu$ where $|\mu| = \|T\|$. By compactness, pass to a subsequence with $Tx_n \to y$. Compute:

$$\|Tx_n - \mu x_n\|^2 = \|Tx_n\|^2 - 2\mu \langle Tx_n, x_n \rangle + \mu^2 \to \|y\|^2 - 2\mu^2 + \mu^2.$$

Since $\|y\| = \lim \|Tx_n\| \leq \|T\| = |\mu|$ and $\|Tx_n\|^2 = \langle Tx_n, Tx_n \rangle = \langle T^2 x_n, x_n \rangle$ with $|\langle T^2 x_n, x_n \rangle| \leq \|T\|^2$ (by applying $m \leq \|T\|$ to T^2, but more carefully: $\|Tx_n\|^2 \leq \|T\| \|Tx_n\| \|x_n\|$ so $\|Tx_n\| \leq \|T\|$), we get $\|Tx_n\|^2 \to \|y\|^2$ and we need $\|y\| = |\mu|$. In fact, $\|Tx_n - \mu x_n\|^2 = \|Tx_n\|^2 - 2\mu \langle Tx_n, x_n \rangle + \mu^2 \leq \mu^2 - 2\mu \langle Tx_n, x_n \rangle + \mu^2 \to 2\mu^2 - 2\mu^2 = 0$. Hence $Tx_n - \mu x_n \to 0$, so $\mu x_n \to y$, giving $x_n \to e := y/\mu$ with $\|e\| = 1$ and $Te = \mu e$. $\qquad \square$

Proof of Theorem 9.14. We construct the eigenvectors iteratively.

Step 1. By Lemma 9.15, if $T \neq 0$ then either $\|T\|$ or $-\|T\|$ is an eigenvalue λ_1 with eigenvector e_1 ($\|e_1\| = 1$, $|\lambda_1| = \|T\|$).

Step 2. Set $H_1 = \{e_1\}^\perp$. Since $T = T^*$, H_1 is T-invariant: for $x \perp e_1$, $\langle Tx, e_1 \rangle = \langle x, Te_1 \rangle = \lambda_1 \langle x, e_1 \rangle = 0$. The restriction $T_1 = T|_{H_1}$ is compact and self-adjoint on H_1 with $\|T_1\| \leq \|T\|$. Apply the lemma to get λ_2, e_2 with $|\lambda_2| \leq |\lambda_1|$.

Step 3. Iterate: $H_n = \{e_1, \ldots, e_n\}^\perp$, and $|\lambda_{n+1}| = \|T|_{H_n}\|$.

Step 4. $\lambda_n \to 0$: if $|\lambda_n| \geq \delta > 0$ for all n, then (e_n) is orthonormal with $\|Te_m - Te_n\|^2 = \lambda_m^2 + \lambda_n^2 \geq 2\delta^2$ for $m \neq n$, contradicting compactness.

Step 5 (Completeness). Let $H_\infty = \bigcap_n H_n$. Then $T|_{H_\infty}$ is compact self-adjoint with $\|T|_{H_\infty}\| = \lim_n |\lambda_{n+1}| = 0$, so $T|_{H_\infty} = 0$. Every vector in H_∞ is an eigenvector for eigenvalue 0. Choose an orthonormal basis of H_∞ and adjoin it to (e_n) to obtain a complete orthonormal basis of H, each element being an eigenvector of T. $\qquad\square$

9.6 The Hilbert–Schmidt Theorem

> **Definition 9.16** (Hilbert–Schmidt operator)**.** Let H be a separable Hilbert space with orthonormal basis (e_n). An operator $T \in \mathcal{L}(H)$ is *Hilbert–Schmidt* if
> $$\|T\|_{\mathrm{HS}}^2 := \sum_{n=1}^\infty \|Te_n\|^2 < \infty.$$
> This quantity is independent of the choice of basis.

> **Theorem 9.17** (Hilbert–Schmidt)**.** *Every Hilbert–Schmidt operator is compact. The space $\mathcal{S}_2(H)$ of Hilbert–Schmidt operators is a two-sided $*$-ideal in $\mathcal{L}(H)$, and $(\mathcal{S}_2(H), \|\cdot\|_{\mathrm{HS}})$ is a Hilbert space with inner product $\langle S, T \rangle_{\mathrm{HS}} = \sum_n \langle Se_n, Te_n \rangle$.*

Proof. Define $T_N x = \sum_{n=1}^N \langle x, e_n \rangle Te_n$, a finite-rank operator. Then

$$\|(T - T_N)x\|^2 = \|[\![\Big] \sum_{n>N} \langle x, e_n \rangle Te_n \Big.^2 \leq \left(\sum_{n>N} \|Te_n\|^2 \right) \|x\|^2$$

by Cauchy–Schwarz. Hence $\|T - T_N\| \leq (\sum_{n>N} \|Te_n\|^2)^{1/2} \to 0$. Since $\mathcal{K}(H)$ is closed, T is compact.

The ideal and Hilbert space properties are verified by direct computation using Parseval's identity. □

Example 9.18 (Hilbert–Schmidt integral operator). If $K \in L^2(\Omega \times \Omega)$, the integral operator T_K of Example 9.3 is Hilbert–Schmidt with $\|T_K\|_{\mathrm{HS}} = \|K\|_{L^2}$. Indeed, for any ONB (e_n) of $L^2(\Omega)$:

$$\sum_n \|T_K e_n\|^2 = \int_\Omega \sum_n |\langle K(x, \cdot), e_n \rangle|^2 \, \mathrm{d}x = \int_\Omega \|K(x, \cdot)\|_{L^2}^2 \, \mathrm{d}x = \|K\|_{L^2(\Omega \times \Omega)}^2.$$

9.7 Application: Fredholm Integral Equations

Theorem 9.19 (Solvability of Fredholm integral equations). *Let $K \in L^2(\Omega \times \Omega)$ be Hermitian ($K(x, y) = \overline{K(y, x)}$ a.e.) and (λ_n, e_n) the eigenpairs from Theorem 9.14. Given $g \in L^2(\Omega)$ and $\mu \neq 0$, consider the equation $f - \mu T_K f = g$.*

(i) If $1/\mu \notin \{\lambda_n\}$, the unique solution is

$$f = g + \sum_{n \geq 1} \frac{\mu \lambda_n}{1 - \mu \lambda_n} \langle g, e_n \rangle \, e_n.$$

(ii) If $1/\mu = \lambda_k$ for some k, a solution exists if and only if $g \perp \mathrm{Ker}(\mathrm{Id} - \mu T_K)$.

Proof. Expand $f = \sum_n c_n e_n$ and $g = \sum_n g_n e_n$. The equation yields $c_n(1 - \mu \lambda_n) = g_n$. If $1 - \mu \lambda_n \neq 0$ for all n, then $c_n = g_n/(1 - \mu \lambda_n)$ and

$$f = \sum_n \frac{g_n}{1 - \mu \lambda_n} e_n = g + \sum_n \frac{\mu \lambda_n}{1 - \mu \lambda_n} g_n e_n.$$

Convergence holds since $\lambda_n \to 0$. When $1 - \mu \lambda_k = 0$, solvability requires $g_k = 0$, i.e., $g \perp e_k$. □

9.8 Exercises

Exercise 9.1 ($\star$). Let H be a Hilbert space and $T \in \mathcal{K}(H)$. Show that if $x_n \rightharpoonup x$, then $Tx_n \to Tx$ in norm.

Exercise 9.2 ($\star$). Let $T \in \mathcal{L}(E, F)$ and $S \in \mathcal{K}(F, G)$. Show ST is compact. Give an example where neither S nor T is compact but ST is.

Exercise 9.3 ($\star\star$). Let $(Vf)(x) = \int_0^x f(t)\,dt$ on $L^2([0, 1])$. Show V is compact with $\sigma(V) = \{0\}$ and $\sigma_p(V) = \emptyset$.

Exercise 9.4 ($\star\star$). Show that every trace class operator is Hilbert–Schmidt. If T is trace class and self-adjoint, show $\mathrm{Tr}(T) = \sum_n \langle Te_n, e_n \rangle$ is basis-independent.

Exercise 9.5 ($\star\star$). (Mercer's theorem.) Let $K \colon [0, 1]^2 \to \mathbb{R}$ be continuous, symmetric, with $T_K \geq 0$. Show $K(x, y) = \sum_n \lambda_n e_n(x) e_n(y)$ uniformly.

Exercise 9.6 ($\star\star\star$). (Schauder.) Let C be closed, bounded, convex in a Banach space E and $T \colon C \to C$ compact. Prove T has a fixed point.

Exercise 9.7 ($\star\star$). Let $T \in \mathcal{K}(H)$ be positive self-adjoint. Prove the min-max principle:

$$\lambda_n = \min_{\mathrm{codim}\, V = n-1} \max_{\substack{x \in V \\ \|x\| = 1}} \langle Tx, x \rangle = \max_{\dim W = n} \min_{\substack{x \in W \\ \|x\| = 1}} \langle Tx, x \rangle.$$

Exercise 9.8 ($\star$). Give an example of a bounded operator on ℓ^2 whose spectrum is $[0, 1]$.

Exercise 9.9 ($\star\star\star$)**.** Show: (a) $\mathrm{Id} - K$ is Fredholm of index zero for all $K \in \mathcal{K}(E)$; (b) the set of Fredholm operators is open and the index is locally constant.

Chapter 10

The Spectral Theorem for Self-Adjoint Operators

Contents

10.1 Continuous Functional Calculus

The spectral theorem for compact self-adjoint operators expresses T as a (norm-convergent) sum involving its eigenvalues and eigenprojections. For a general bounded self-adjoint operator, eigenvalues need not exist (consider multiplication by x on $L^2([0,1])$), and we must develop a more sophisticated framework.

Notation 10.1. For a bounded self-adjoint operator T on a Hilbert space H, we write $\sigma(T) \subset \mathbb{R}$ for its spectrum and set $m = \inf_{\|x\|=1} \langle Tx, x \rangle$, $M = \sup_{\|x\|=1} \langle Tx, x \rangle$, so that $\sigma(T) \subset [m, M]$.

> **Theorem 10.2** (Continuous functional calculus)**.** *Let* $T \in \mathcal{L}(H)$ *be self-adjoint. There exists a unique isometric $*$-algebra homomorphism*
>
> $$\Phi_T \colon C(\sigma(T)) \longrightarrow \mathcal{L}(H), \qquad f \longmapsto f(T),$$
>
> *satisfying:*
>
> (i) $\Phi_T(1) = \mathrm{Id}$ *and* $\Phi_T(\mathrm{id}) = T$, *where* $\mathrm{id}(\lambda) = \lambda$.
>
> (ii) $\Phi_T(\bar{f}) = \Phi_T(f)^*$.
>
> (iii) $\|f(T)\| = \|f\|_{C(\sigma(T))} := \sup_{\lambda \in \sigma(T)} |f(\lambda)|$.
>
> (iv) $\sigma(f(T)) = f(\sigma(T))$ *(spectral mapping theorem).*

Proof. **Step 1 (Polynomials).** For a polynomial $p(\lambda) = \sum_{k=0}^n a_k \lambda^k$, define $p(T) = \sum_{k=0}^n a_k T^k$. We must show that $\|p(T)\| = \sup_{\lambda \in \sigma(T)} |p(\lambda)|$.

Since T is self-adjoint, $p(T)$ is normal (it commutes with its adjoint). For a self-adjoint operator S, $\|S\| = r(S)$ (the spectral radius). For a normal operator N, $\|N^2\| = \|N^*N\| = \|N\|^2$, hence by induction $\left\|N^{2^k}\right\| = \|N\|^{2^k}$, giving $\|N\| = r(N) = \sup_{\mu \in \sigma(N)} |\mu|$. Since $\sigma(p(T)) = p(\sigma(T))$ by the polynomial spectral mapping theorem,

$$\|p(T)\| = r(p(T)) = \sup_{\mu \in \sigma(p(T))} |\mu| = \sup_{\lambda \in \sigma(T)} |p(\lambda)|.$$

Step 2 (Extension). The map $p \mapsto p(T)$ is an isometry from $(C_{\mathrm{poly}}(\sigma(T)), \|\cdot\|_\infty)$ to $\mathcal{L}(H)$. By the Weierstrass approximation theorem, polynomials are dense in $C(\sigma(T))$ (since $\sigma(T)$ is a compact subset of $\mathbb{R}$). The isometry extends uniquely to a map $\Phi_T \colon C(\sigma(T)) \to \mathcal{L}(H)$.

Step 3 (Properties). The $*$-homomorphism properties follow by continuity from the polynomial case. The spectral mapping theorem: $\mu \in \sigma(f(T))$ iff $f(T) - \mu\,\mathrm{Id}$ is not invertible. By the isometry, this happens iff $f - \mu$ has no inverse in $C(\sigma(T))$, i.e., iff $f(\lambda) = \mu$ for some $\lambda \in \sigma(T)$. $\qquad\square$

10.2 Spectral Measures and Resolution of the Identity

Definition 10.3 (Spectral measure)**.** Let H be a Hilbert space and $(\Omega, \mathcal{B})$ a measurable space. A *spectral measure* (or *projection-valued measure*) on $(\Omega, \mathcal{B})$ with values in $\mathcal{L}(H)$ is a map $E \colon \mathcal{B} \to \mathcal{L}(H)$ satisfying:

   (i) $E(\Delta)$ is an orthogonal projection for each $\Delta \in \mathcal{B}$.

   (ii) $E(\emptyset) = 0$ and $E(\Omega) = \mathrm{Id}$.

   (iii) $E(\Delta_1 \cap \Delta_2) = E(\Delta_1)E(\Delta_2)$ for all $\Delta_1, \Delta_2 \in \mathcal{B}$.

   (iv) If (Δ_n) are pairwise disjoint, then $E(\bigcup_n \Delta_n) = \sum_n E(\Delta_n)$ (strong operator convergence).

Remark 10.4. For any $x, y \in H$, the map $\Delta \mapsto \langle E(\Delta)x, y \rangle$ defines a complex Borel measure $\mu_{x,y}$ on Ω. When $x = y$, the measure $\mu_{x,x}(\Delta) = \|E(\Delta)x\|^2$ is positive. These are called the *scalar spectral measures.*

Definition 10.5 (Resolution of the identity)**.** For a bounded self-adjoint operator T with $\sigma(T) \subset [m, M]$, a *resolution of the identity* is a family of projections $(E_\lambda)_{\lambda \in \mathbb{R}}$ satisfying:

   (i) $E_\lambda \le E_\mu$ (i.e., $\mathrm{im}\, E_\lambda \subset \mathrm{im}\, E_\mu$) for $\lambda \le \mu$.

   (ii) $E_\lambda = 0$ for $\lambda < m$ and $E_\lambda = \mathrm{Id}$ for $\lambda \ge M$.

   (iii) E_λ is strongly right-continuous: $E_{\lambda+} = E_\lambda$.

10.3 The Spectral Theorem (Spectral Measure Version)

> **Theorem 10.6** (Spectral theorem for bounded self-adjoint operators).
> *Let $T \in \mathcal{L}(H)$ be self-adjoint. There exists a unique spectral measure E on $(\sigma(T), \mathcal{B}(\sigma(T)))$ such that*
>
> $$T = \int_{\sigma(T)} \lambda \, dE(\lambda),$$
>
> *meaning $\langle Tx, y \rangle = \int_{\sigma(T)} \lambda \, d\mu_{x,y}(\lambda)$ for all $x, y \in H$. More generally, for every bounded Borel function $f \colon \sigma(T) \to \mathbb{C}$:*
>
> $$f(T) = \int_{\sigma(T)} f(\lambda) \, dE(\lambda) \in \mathcal{L}(H).$$

Proof. **Step 1 (From continuous to Borel functions).** By Theorem 10.2, we have a $*$-homomorphism $\Phi_T \colon C(\sigma(T)) \to \mathcal{L}(H)$. For each $x \in H$, the map $f \mapsto \langle f(T)x, x \rangle$ is a positive linear functional on $C(\sigma(T))$ (positive because $f \geq 0$ implies $f(T) \geq 0$).

By the Riesz representation theorem, there exists a unique positive Borel measure μ_x on $\sigma(T)$ such that

$$\langle f(T)x, x \rangle = \int_{\sigma(T)} f(\lambda) \, d\mu_x(\lambda) \quad \text{for all } f \in C(\sigma(T)).$$

Note $\mu_x(\sigma(T)) = \langle \text{Id } x, x \rangle = \|x\|^2$.

Step 2 (Complex measures). By polarization, for $x, y \in H$ define the complex measure $\mu_{x,y}$ via

$$\mu_{x,y} = \tfrac{1}{4}\left(\mu_{x+y} - \mu_{x-y} + i\mu_{x+iy} - i\mu_{x-iy} \right).$$

Then $\langle f(T)x, y \rangle = \int f \, d\mu_{x,y}$ for all $f \in C(\sigma(T))$.

Step 3 (Extension to bounded Borel functions). For a bounded Borel function f, the map $(x, y) \mapsto \int f \, d\mu_{x,y}$ is a bounded sesquilinear form. By the Riesz representation theorem for bounded sesquilinear forms, there exists a unique $f(T) \in \mathcal{L}(H)$ with $\langle f(T)x, y \rangle = \int f \, d\mu_{x,y}$.

The map $f \mapsto f(T)$ is a $*$-homomorphism from $B(\sigma(T))$ (bounded Borel functions) to $\mathcal{L}(H)$. The multiplicativity $f(T)g(T) = (fg)(T)$ follows from the continuous case by a monotone class argument.

Step 4 (Spectral measure). For $\Delta \in \mathcal{B}(\sigma(T))$, set $E(\Delta) = \mathbf{1}_\Delta(T)$. Since $\mathbf{1}_\Delta^2 = \mathbf{1}_\Delta = \overline{\mathbf{1}_\Delta}$, we have $E(\Delta)^2 = E(\Delta) = E(\Delta)^*$, so $E(\Delta)$ is an orthogonal projection. The spectral measure axioms follow from the properties of the $*$-homomorphism:

- $E(\sigma(T)) = \mathbf{1}_{\sigma(T)}(T) = \mathrm{Id}$.

- $E(\Delta_1 \cap \Delta_2) = \mathbf{1}_{\Delta_1 \cap \Delta_2}(T) = \mathbf{1}_{\Delta_1}(T)\mathbf{1}_{\Delta_2}(T) = E(\Delta_1)E(\Delta_2)$.

- Countable additivity in the strong operator topology follows from the dominated convergence theorem applied to $\mu_{x,x}$.

Uniqueness. If E' is another spectral measure with $T = \int \lambda \, dE'(\lambda)$, then $p(T) = \int p \, dE'$ for every polynomial p. By density of polynomials and uniqueness in the Riesz representation theorem, $\mu'_{x,x} = \mu_{x,x}$ for all x, hence $E' = E$. $\qquad\square$

10.4 Borel Functional Calculus

The spectral theorem gives us a powerful extension of the continuous functional calculus.

Definition 10.7 (Borel functional calculus)**.** For T self-adjoint and $f\colon \sigma(T) \to \mathbb{C}$ a bounded Borel function, we define

$$f(T) = \int_{\sigma(T)} f(\lambda) \, dE(\lambda),$$

where E is the spectral measure of T.

Proposition 10.8. *Let T be bounded self-adjoint with spectral measure E, and let $f, g \in B(\sigma(T))$.*

(i) $(\alpha f + \beta g)(T) = \alpha f(T) + \beta g(T)$.

(ii) $(fg)(T) = f(T)\, g(T)$.

(iii) $\bar{f}(T) = f(T)^*$.

(iv) $\|f(T)\| \le \|f\|_\infty$.

(v) *If $f_n \to f$ pointwise with $\sup_n \|f_n\|_\infty < \infty$, then $f_n(T) \to f(T)$ in the strong operator topology.*

> (vi) $\sigma(f(T)) \subset \overline{f(\sigma(T))}$.

Proof. Properties (i)–(iv) follow from the construction in Theorem 10.6. For (v), let $x \in H$; then $\|f_n(T)x - f(T)x\|^2 = \int |f_n - f|^2 \, d\mu_{x,x}$. By dominated convergence (the integrand is bounded by $(2 \sup_n \|f_n\|_\infty)^2$ and $\mu_{x,x}$ is finite), this tends to 0. Property (vi) follows from (ii): if $\mu \notin \overline{f(\sigma(T))}$, then $h = 1/(f - \mu)$ is bounded and Borel on $\sigma(T)$, so $h(T)(f(T) - \mu \, \mathrm{Id}) = \mathrm{Id}$. $\square$

10.5 Positive Operators and Square Root

> **Definition 10.9.** A self-adjoint operator $T \in \mathcal{L}(H)$ is *positive* (written $T \geq 0$) if $\langle Tx, x \rangle \geq 0$ for all $x \in H$. Equivalently, $\sigma(T) \subset [0, \infty)$.

> **Theorem 10.10** (Square root). *If $T \geq 0$, there exists a unique $S \geq 0$ with $S^2 = T$. We write $S = T^{1/2}$ or $S = \sqrt{T}$. Moreover, S commutes with every operator that commutes with T.*

Proof. **Existence.** Apply the continuous functional calculus with $f(\lambda) = \sqrt{\lambda}$ on $\sigma(T) \subset [0, \|T\|]$. Set $S = f(T) = \sqrt{\cdot}(T)$. Then $S^2 = (\sqrt{\cdot})^2(T) = \mathrm{id}(T) = T$, $S^* = \overline{\sqrt{\cdot}}(T) = \sqrt{\cdot}(T) = S$, and $\sigma(S) = f(\sigma(T)) \subset [0, \infty)$, so $S \geq 0$.

Uniqueness. Let $R \geq 0$ with $R^2 = T$. Then R commutes with $T = R^2$, hence with any polynomial in T, and by continuity with $S = \sqrt{T}$. Let $N = \mathrm{Ker}(R - S)^\perp$ and decompose $H = \mathrm{Ker}(R - S) \oplus N$. On N, $R - S$ is injective. Now $(R - S)(R + S) = R^2 - S^2 = 0$ (using $RS = SR$), so for $x \in N$, $(R + S)x \in \mathrm{Ker}(R - S)$, hence $\langle (R + S)x, x \rangle = 0$ for $x \in N$ (since $N \perp \mathrm{Ker}(R - S)$). But $R + S \geq 0$, so $\langle Rx, x \rangle + \langle Sx, x \rangle = 0$ with both terms non-negative forces $Rx = Sx = 0$ for $x \in N$, hence $(R - S)x = 0$. Thus $N \subset \mathrm{Ker}(R - S)$, so $N = \{0\}$ and $R = S$.

Commuting property. If A commutes with T, then A commutes with $p(T)$ for every polynomial p, hence with $\sqrt{T}$ by continuity of Φ_T. $\square$

10.6 Polar Decomposition

Definition 10.11. For $T \in \mathcal{L}(H)$, define $|T| = (T^*T)^{1/2}$. Note $T^*T \geq 0$, so $|T|$ exists by Theorem 10.10.

Theorem 10.12 (Polar decomposition). *For every $T \in \mathcal{L}(H)$, there exists a unique partial isometry $U \in \mathcal{L}(H)$ such that $T = U\,|T|$ and $\operatorname{Ker} U = \operatorname{Ker} T$. Moreover:*

(i) U is an isometry from $(\operatorname{Ker} T)^\perp = \overline{\operatorname{im} |T|}$ onto $\overline{\operatorname{im} T}$.

(ii) $|T| = U^*T$.

(iii) $T^* = |T|\,U^*$ and $|T^*| = U\,|T|\,U^*$.

Proof. **Construction of U.** For $x \in H$, observe $\|Tx\|^2 = \langle T^*Tx, x \rangle = \langle |T|^2 x, x \rangle = \||T|\,x\|^2$. Hence the map $|T|\,x \mapsto Tx$ is a well-defined isometry from $\operatorname{im} |T|$ to $\operatorname{im} T$. Extend by continuity to an isometry $U_0 \colon \overline{\operatorname{im} |T|} \to \overline{\operatorname{im} T}$, and set $U = 0$ on $(\overline{\operatorname{im} |T|})^\perp = \operatorname{Ker} |T| = \operatorname{Ker} T$ (the last equality because $\||T|\,x\| = \|Tx\|$). Then U is a partial isometry with $T = U\,|T|$ and $\operatorname{Ker} U = \operatorname{Ker} T$.

Uniqueness. If $T = V\,|T|$ with V a partial isometry and $\operatorname{Ker} V = \operatorname{Ker} T$, then V agrees with U on $\operatorname{im} |T|$ (since $V\,|T|\,x = Tx = U\,|T|\,x$) and on $\operatorname{Ker} |T| = \operatorname{Ker} T$ (since both vanish). By density, $V = U$.

Property (ii): $U^*T = U^*U\,|T| = P_{\overline{\operatorname{im} |T|}}\,|T| = |T|$ since $\operatorname{im} |T| \subset \overline{\operatorname{im} |T|}$. $\square$

10.7 Unbounded Operators

Many important operators in analysis and physics (differential operators, for instance) are not bounded. We develop the basic framework.

Definition 10.13. An *unbounded operator* on a Hilbert space H is a linear map $T \colon \operatorname{dom}(T) \to H$ where $\operatorname{dom}(T)$ is a linear subspace of H (the *domain* of T), not necessarily all of H.

Definition 10.14 (Graph, closed operator). The *graph* of T is $\Gamma(T) = \{(x, Tx) : x \in \mathrm{dom}(T)\} \subset H \times H$. The operator T is *closed* if $\Gamma(T)$ is closed in $H \times H$, i.e., whenever $x_n \in \mathrm{dom}(T)$, $x_n \to x$, and $Tx_n \to y$, we have $x \in \mathrm{dom}(T)$ and $Tx = y$.

Definition 10.15 (Closable operator). T is *closable* if the closure $\overline{\Gamma(T)}$ is the graph of an operator (equivalently, if $x_n \in \mathrm{dom}(T)$, $x_n \to 0$, $Tx_n \to y$ implies $y = 0$). The resulting operator $\bar{T}$ is the *closure* of T.

Definition 10.16 (Symmetric, self-adjoint). Let T be densely defined $(\overline{\mathrm{dom}(T)} = H)$.

- T is *symmetric* if $\langle Tx, y \rangle = \langle x, Ty \rangle$ for all $x, y \in \mathrm{dom}(T)$.

- The *adjoint* T^* has domain $\mathrm{dom}(T^*) = \{y \in H : x \mapsto \langle Tx, y \rangle$ is bounded on $\mathrm{dom}(T)\}$, and T^*y is the unique element such that $\langle Tx, y \rangle = \langle x, T^*y \rangle$.

- T is *self-adjoint* if $T = T^*$ (meaning $\mathrm{dom}(T) = \mathrm{dom}(T^*)$ and $Tx = T^*x$ for all $x \in \mathrm{dom}(T)$).

A symmetric operator satisfies $T \subset T^*$ (i.e., $\mathrm{dom}(T) \subset \mathrm{dom}(T^*)$ and $T^*|_{\mathrm{dom}(T)} = T$). Self-adjointness is the stronger condition $\mathrm{dom}(T) = \mathrm{dom}(T^*)$.

Remark 10.17. The distinction between symmetric and self-adjoint is crucial. A symmetric operator may have many self-adjoint extensions, exactly one, or none. The spectral theorem requires genuine self-adjointness.

Example 10.18 (The Laplacian). On $H = L^2(\mathbb{R}^d)$, define $T = -\Delta$ with $\mathrm{dom}(T) = H^2(\mathbb{R}^d)$ (the Sobolev space). This is self-adjoint. Indeed, for $f, g \in H^2(\mathbb{R}^d)$:

$$\langle -\Delta f, g \rangle_{L^2} = \int_{\mathbb{R}^d} \nabla f \cdot \overline{\nabla g} \, \mathrm{d}x = \langle f, -\Delta g \rangle_{L^2}$$

by integration by parts. Self-adjointness (not just symmetry) can be proved via the Fourier transform: under $\mathcal{F}$, $-\Delta$ becomes multiplication

by $|\xi|^2$, which is self-adjoint on dom $= \{f \in L^2 : |\xi|^2 \hat{f} \in L^2\} = H^2(\mathbb{R}^d)$.

Example 10.19 (Momentum operator). On $L^2(\mathbb{R})$, define $P = -i\frac{\mathrm{d}}{\mathrm{d}x}$ with $\mathrm{dom}(P) = H^1(\mathbb{R})$. Then P is self-adjoint. Via Fourier transform, P becomes multiplication by ξ.

Proposition 10.20 (Criteria for self-adjointness). *Let T be symmetric and densely defined. The following are equivalent:*

(i) T is self-adjoint.

(ii) T is closed and $\mathrm{Ker}(T^* \pm i) = \{0\}$.

(iii) $\mathrm{ran}(T \pm i) = H$.

Proof. (i)$\Rightarrow$(ii): If $T = T^*$, then T is closed (since T^* is always closed). If $T^*y = iy$, then $\langle Ty, y \rangle = \langle y, T^*y \rangle = \langle y, iy \rangle = -i\|y\|^2$ but also $\langle Ty, y \rangle = \langle T^*y, y \rangle = \langle iy, y \rangle = i\|y\|^2$, giving $y = 0$.

(ii)$\Rightarrow$(iii): We show $\mathrm{ran}(T - i)$ is closed and dense. Closedness: if $x \in \mathrm{dom}(T)$, $\|(T - i)x\|^2 = \|Tx\|^2 + \|x\|^2$ (since $\mathrm{Re}\langle Tx, ix \rangle = 0$ for symmetric T), so $T - i$ is bounded below, forcing $\mathrm{ran}(T - i)$ closed. Density: if $y \perp \mathrm{ran}(T - i)$, then $y \in \mathrm{dom}(T^*)$ and $T^*y = iy$, hence $y = 0$ by hypothesis. Similarly for $\mathrm{ran}(T + i)$.

(iii)$\Rightarrow$(i): Let $y \in \mathrm{dom}(T^*)$. We show $y \in \mathrm{dom}(T)$. Since $\mathrm{ran}(T+i) = H$, there exists $x \in \mathrm{dom}(T)$ with $(T + i)x = (T^* + i)y$. Since $T \subset T^*$, we have $T^*(x - y) = -i(x - y)$, i.e., $(x - y) \in \mathrm{Ker}(T^* + i)$. But $\mathrm{ran}(T - i) = H$ implies $\mathrm{Ker}(T^* + i) = \mathrm{ran}(T - i)^\perp = \{0\}$. Hence $x = y$ and $y \in \mathrm{dom}(T)$. $\square$

10.8 Spectral Theorem for Unbounded Self-Adjoint Operators

Theorem 10.21 (Spectral theorem, unbounded version). *Let T be a self-adjoint operator on H (possibly unbounded). There exists a unique*

spectral measure E on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ with $E(\mathbb{R} \setminus \sigma(T)) = 0$ and

$$T = \int_{-\infty}^{\infty} \lambda \, dE(\lambda),$$

in the sense that

$$\operatorname{dom}(T) = \left\{ x \in H : \int_{-\infty}^{\infty} \lambda^2 \, d\mu_{x,x}(\lambda) < \infty \right\}$$

and $\langle Tx, y \rangle = \int \lambda \, d\mu_{x,y}(\lambda)$ for $x \in \operatorname{dom}(T)$, $y \in H$.
For every Borel measurable function $f \colon \mathbb{R} \to \mathbb{C}$, the operator

$$f(T) = \int f(\lambda) \, dE(\lambda)$$

is defined on $\operatorname{dom}(f(T)) = \{ x \in H : \int |f|^2 \, d\mu_{x,x} < \infty \}$.

The proof uses the Cayley transform $U = (T - i)(T + i)^{-1}$, which is a unitary operator when T is self-adjoint (by Proposition 10.20). One applies the spectral theorem for unitary operators (a variant of the bounded case) to U, then transfers back via $T = i(\operatorname{Id} + U)(\operatorname{Id} - U)^{-1}$.

10.9   Application: Quantum Mechanics

The spectral theorem provides the mathematical foundation for quantum mechanics.

Remark 10.22 (Observables as self-adjoint operators). In the quantum-mechanical formalism:

- The *state space* is a separable Hilbert space H (typically $L^2(\mathbb{R}^3)$).

- *Observables* (position, momentum, energy, etc.) are represented by self-adjoint operators on H.

- If T is an observable with spectral measure E and the system is in state ψ ($\|\psi\| = 1$), the probability of measuring T in a Borel set $\Delta \subset \mathbb{R}$ is $\Pr(T \in \Delta) = \langle E(\Delta)\psi, \psi \rangle = \mu_{\psi,\psi}(\Delta)$.

- The expected value is $\langle T \rangle_\psi = \langle T\psi, \psi \rangle = \int \lambda \, d\mu_{\psi,\psi}(\lambda)$.

- The Heisenberg uncertainty principle $\Delta T \cdot \Delta S \geq \frac{1}{2} |\langle [T,S]\psi, \psi \rangle|$ (where $[T,S] = TS - ST$) follows from the Cauchy–Schwarz inequality.

10.10 Exercises

Exercise 10.1 ($\star$). Let T be bounded self-adjoint with spectral measure E. Show that T has an eigenvalue λ if and only if $E(\{\lambda\}) \neq 0$, and in that case $E(\{\lambda\})$ is the orthogonal projection onto $\mathrm{Ker}(T - \lambda)$.

Exercise 10.2 ($\star\star$). Let T be self-adjoint, f continuous, and g bounded Borel. Show $g(f(T)) = (g \circ f)(T)$.

Exercise 10.3 ($\star$). Let $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ on $\mathbb{R}^2$. Compute $A^{1/2}$.

Exercise 10.4 ($\star\star$). Let $g \in L^\infty(\mathbb{R})$ be real-valued and M_g the multiplication operator on $L^2(\mathbb{R})$. Describe the spectral measure of M_g explicitly.

Exercise 10.5 ($\star\star$). Let $T = -\frac{d^2}{dx^2}$ on $\mathrm{dom}(T) = C_c^\infty((0,1))$ in $L^2((0,1))$. Show T is symmetric but not self-adjoint. Find its self-adjoint extensions (parametrized by boundary conditions).

Exercise 10.6 ($\star\star\star$). (Stone's theorem.) Let $(U(t))_{t \in \mathbb{R}}$ be a strongly continuous one-parameter unitary group on H. Show there exists a self-adjoint operator A such that $U(t) = e^{itA}$ for all t.

Exercise 10.7 ($\star\star$). Show that T is normal if and only if $|T^*| = |T|$ (equivalently, U in the polar decomposition is unitary on $(\mathrm{Ker}\, T)^\perp$).

Exercise 10.8 ($\star\star\star$)**.** (Kato–Rellich.) Let A be self-adjoint and B symmetric with $\mathrm{dom}(A) \subset \mathrm{dom}(B)$. Suppose there exist $a < 1$ and $b \geq 0$ with $\|Bx\| \leq a\|Ax\| + b\|x\|$ for all $x \in \mathrm{dom}(A)$. Prove that $A + B$ is self-adjoint on $\mathrm{dom}(A)$.

Exercise 10.9 ($\star$)**.** Describe the Hamiltonian of the hydrogen atom $H = -\Delta - 1/|x|$ on $L^2(\mathbb{R}^3)$ and explain why the Kato–Rellich theorem guarantees it is self-adjoint on $H^2(\mathbb{R}^3)$.

Chapter 11

L^p Spaces and Duality

Contents

11.1   Hölder and Minkowski Inequalities

Throughout this chapter, $(\Omega, \mathcal{A}, \mu)$ denotes a σ-finite measure space and $1 \leq p \leq \infty$. We write p' (or q) for the conjugate exponent: $\frac{1}{p} + \frac{1}{p'} = 1$ (with $1' = \infty$ and $\infty' = 1$).

> **Lemma 11.1** (Young's inequality for products). *For $a, b \geq 0$ and $1 < p < \infty$ with conjugate q:*
> $$ab \leq \frac{a^p}{p} + \frac{b^q}{q}.$$

> *Equality holds if and only if $a^p = b^q$.*

Proof. The function $t \mapsto e^t$ is convex, so $e^{\alpha s + \beta t} \leq \alpha e^s + \beta e^t$ for $\alpha, \beta > 0$, $\alpha + \beta = 1$. Take $\alpha = 1/p$, $\beta = 1/q$, $s = p \ln a$, $t = q \ln b$ (for $a, b > 0$; the case $a = 0$ or $b = 0$ is trivial). $\qquad\square$

Theorem 11.2 (Hölder's inequality)**.** *Let $1 \leq p \leq \infty$ and $q = p'$. If $f \in L^p(\mu)$ and $g \in L^q(\mu)$, then $fg \in L^1(\mu)$ and*

$$\|fg\|_{L^1} \leq \|f\|_{L^p} \|g\|_{L^q}.$$

Proof. The cases $p = 1$ (or $p = \infty$) are immediate. For $1 < p < \infty$, we may assume $\|f\|_p, \|g\|_q > 0$. Set $\tilde{f} = f/\|f\|_p$ and $\tilde{g} = g/\|g\|_q$. By Young's inequality:

$$\left|\tilde{f}(x)\tilde{g}(x)\right| \leq \frac{\left|\tilde{f}(x)\right|^p}{p} + \frac{|\tilde{g}(x)|^q}{q}.$$

Integrating: $\int \left|\tilde{f}\tilde{g}\right| \, d\mu \leq \frac{1}{p} + \frac{1}{q} = 1$. Hence $\|fg\|_1 \leq \|f\|_p \|g\|_q$. $\qquad\square$

Theorem 11.3 (Minkowski's inequality)**.** *For $1 \leq p \leq \infty$ and $f, g \in L^p(\mu)$:*

$$\|f + g\|_{L^p} \leq \|f\|_{L^p} + \|g\|_{L^p}.$$

Proof. For $p = 1$ and $p = \infty$ this is the triangle inequality. For $1 < p < \infty$:

$$\|f + g\|_p^p = \int |f + g|^p \, d\mu \leq \int |f + g|^{p-1} |f| \, d\mu + \int |f + g|^{p-1} |g| \, d\mu.$$

Since $(p - 1)q = p$, the function $|f + g|^{p-1} \in L^q$ with $\left\||f + g|^{p-1}\right\|_q = \|f + g\|_p^{p/q}$. Apply Hölder to each term:

$$\|f + g\|_p^p \leq \|f + g\|_p^{p/q} (\|f\|_p + \|g\|_p).$$

Divide by $\|f + g\|_p^{p/q}$ (assuming it is nonzero) and use $p - p/q = 1$. $\qquad\square$

11.2　Completeness of L^p: The Riesz–Fischer Theorem

Theorem 11.4 (Riesz–Fischer). *For $1 \leq p \leq \infty$, the space $L^p(\Omega, \mu)$ is a Banach space.*

Proof for $1 \leq p < \infty$. Let (f_n) be Cauchy in L^p. We show it converges by extracting a subsequence that converges both in L^p and almost everywhere.

Step 1 (Subsequence). Choose $n_1 < n_2 < \cdots$ such that $\left\| f_{n_{k+1}} - f_{n_k} \right\|_p < 2^{-k}$. Set $g_k = f_{n_k}$ and

$$S_N(x) = \sum_{k=1}^{N} |g_{k+1}(x) - g_k(x)|, \qquad S(x) = \sum_{k=1}^{\infty} |g_{k+1}(x) - g_k(x)|.$$

By Minkowski's inequality, $\|S_N\|_p \leq \sum_{k=1}^{N} 2^{-k} < 1$. By the monotone convergence theorem, $\|S\|_p \leq 1 < \infty$, so $S(x) < \infty$ a.e.

Step 2 (Pointwise limit). Where $S(x) < \infty$, the telescoping series $g_1(x) + \sum_{k=1}^{\infty}(g_{k+1}(x) - g_k(x))$ converges absolutely. Define $f(x) = \lim_{k \to \infty} g_k(x)$ where this limit exists, and $f(x) = 0$ elsewhere.

Step 3 (L^p convergence). We have $|g_k(x) - f(x)|^p \to 0$ a.e. and $|g_k - f|^p \leq (2S)^p \in L^1$. By dominated convergence, $\|g_k - f\|_p \to 0$.

Step 4 (Full sequence). Since (f_n) is Cauchy and the subsequence $(g_k) = (f_{n_k})$ converges to f in L^p, the full sequence $f_n \to f$ in L^p. $\qquad\square$

Proof for $p = \infty$. If (f_n) is Cauchy in L^∞, then for each pair m, n, $|f_m(x) - f_n(x)| \leq \|f_m - f_n\|_\infty$ outside a μ-null set $N_{m,n}$. The union $N = \bigcup_{m,n} N_{m,n}$ is still μ-null, and on $\Omega \setminus N$ the sequence $(f_n(x))$ is uniformly Cauchy. Define $f(x) = \lim f_n(x)$ on $\Omega \setminus N$, $f = 0$ on N. Then $\|f_n - f\|_\infty \to 0$. $\qquad\square$

11.3 Density of C_c^∞ in L^p

Theorem 11.5. *Let $\Omega \subset \mathbb{R}^d$ be open and $1 \leq p < \infty$. Then $C_c^\infty(\Omega)$ is dense in $L^p(\Omega)$.*

The proof uses convolution with an approximation to the identity.

Definition 11.6 (Mollifier). Let $\rho \in C_c^\infty(\mathbb{R}^d)$ with $\rho \geq 0$, $\operatorname{supp} \rho \subset B(0,1)$, and $\int \rho = 1$. For $\varepsilon > 0$, set $\rho_\varepsilon(x) = \varepsilon^{-d} \rho(x/\varepsilon)$, so that $\operatorname{supp} \rho_\varepsilon \subset B(0,\varepsilon)$ and $\int \rho_\varepsilon = 1$. The family (ρ_ε) is called an *approximation to the identity* (or *mollifier*).

11.4   Convolution and Approximation

Definition 11.7. For $f \in L^p(\mathbb{R}^d)$ and $g \in L^1(\mathbb{R}^d)$, the *convolution* is

$$(f * g)(x) = \int_{\mathbb{R}^d} f(x - y) \, g(y) \, \mathrm{d}y.$$

Proposition 11.8 (Young's convolution inequality). *If $f \in L^p(\mathbb{R}^d)$ and $g \in L^1(\mathbb{R}^d)$, then $f * g \in L^p(\mathbb{R}^d)$ with $\|f * g\|_p \le \|f\|_p \|g\|_1$.*

Proof. For $p = 1$ this is Fubini–Tonelli. For $1 < p < \infty$: write $|g| = |g|^{1/q} |g|^{1/p}$ (with $1/p + 1/q = 1$) and apply Hölder:

$$|(f * g)(x)| \le \int |f(x - y)| \, |g(y)| \, \mathrm{d}y$$
$$\le \left(\int |g(y)| \, \mathrm{d}y \right)^{1/q} \left(\int |f(x - y)|^p \, |g(y)| \, \mathrm{d}y \right)^{1/p}.$$

Raising to the p-th power and integrating in x: $\|f * g\|_p^p \le \|g\|_1^{p/q} \int\int |f(x - y)|^p \, |g(y)| \, \mathrm{d}y \, \mathrm{d}x = \|g\|_1^{p/q} \|f\|_p^p \|g\|_1 = \|f\|_p^p \|g\|_1^p.$ $\qquad\square$

Proposition 11.9 (Approximation by mollifiers). *If $f \in L^p(\mathbb{R}^d)$ with $1 \le p < \infty$, then $f * \rho_\varepsilon \in C^\infty(\mathbb{R}^d)$ and $f * \rho_\varepsilon \to f$ in L^p as $\varepsilon \to 0$.*

Proof. Smoothness: differentiating under the integral sign (justified by dominated convergence and the compactness of $\operatorname{supp}\rho_\varepsilon$) gives $D^\alpha(f * \rho_\varepsilon) = f * D^\alpha \rho_\varepsilon \in C^\infty$.

L^p convergence: by Minkowski's integral inequality,

$$\|f * \rho_\varepsilon - f\|_p = \left\| \int [\tau_y f - f] \, \rho_\varepsilon(y) \, \mathrm{d}y \right\|_p \le \int \|\tau_y f - f\|_p \, \rho_\varepsilon(y) \, \mathrm{d}y,$$

where $(\tau_y f)(x) = f(x - y)$. Since translations are continuous in L^p for $p < \infty$, given $\delta > 0$ there exists $\eta > 0$ with $\|\tau_y f - f\|_p < \delta$ for $|y| < \eta$. For $\varepsilon < \eta$, $\operatorname{supp}\rho_\varepsilon \subset B(0, \eta)$, so the integral above is $< \delta$. $\qquad\square$

Proof of Theorem 11.5. **Step 1.** Approximate $f \in L^p(\Omega)$ by $f_n = f\mathbf{1}_{K_n}$ where $K_n \nearrow \Omega$ are compact. Then $f_n \to f$ in L^p by dominated convergence.

Step 2. Each f_n (extended by zero to $\mathbb{R}^d$) lies in $L^p(\mathbb{R}^d)$. By Proposition 11.9, $g_\varepsilon = f_n * \rho_\varepsilon \in C^\infty$ and $g_\varepsilon \to f_n$ in L^p. For $\varepsilon < d(K_n, \Omega^c)$, $\operatorname{supp} g_\varepsilon \subset \Omega$ and is compact, so $g_\varepsilon \in C_c^\infty(\Omega)$. $\qquad\square$

11.5 The Dual of L^p

The following is one of the central results in the theory of L^p spaces.

> **Theorem 11.10** (Dual of L^p). *Let $(\Omega, \mathcal{A}, \mu)$ be σ-finite and $1 \leq p < \infty$. The map*
>
> $$\Phi \colon L^q(\mu) \longrightarrow (L^p(\mu))', \qquad \Phi(g)(f) = \int_\Omega f\,\bar g\,\mathrm{d}\mu,$$
>
> *where $q = p'$, is an isometric isomorphism. That is, $(L^p(\mu))' \cong L^q(\mu)$.*

Proof. **Step 1 (Φ is an isometry).** By Hölder, $|\Phi(g)(f)| \leq \|f\|_p \|g\|_q$, so $\|\Phi(g)\| \leq \|g\|_q$. For the reverse, take $f = |g|^{q-2}\,\bar g$ (when $g \neq 0$; here $|g|^{q-2}\,\bar g$ has $|f|^p = |g|^{(q-1)p} = |g|^q$, so $f \in L^p$). Then $\Phi(g)(f) = \int |g|^q\,\mathrm{d}\mu = \|g\|_q^q$ and $\|f\|_p = \|g\|_q^{q/p}$, giving $\|\Phi(g)\| \geq \|g\|_q^q / \|g\|_q^{q/p} = \|g\|_q^{q-q/p} = \|g\|_q$.

Step 2 (Φ is surjective). Let $\varphi \in (L^p(\mu))'$.

Case $p = 1$ ($q = \infty$): For measurable A with $\mu(A) < \infty$, define $\nu(A) = \varphi(\mathbf{1}_A)$. Then ν is a signed (or complex) measure absolutely continuous with respect to μ (since $\mu(A) = 0$ implies $\mathbf{1}_A = 0$ in L^1). By the Radon–Nikodym theorem, $\mathrm{d}\nu = g\,\mathrm{d}\mu$ for some measurable g. One checks $g \in L^\infty$ with $\|g\|_\infty \leq \|\varphi\|$.

Case $1 < p < \infty$: Define ν as above. The key estimate is: for any measurable set A with $\mu(A) > 0$,

$$|\nu(A)| = |\varphi(\mathbf{1}_A)| \leq \|\varphi\|\, \|\mathbf{1}_A\|_p = \|\varphi\|\, \mu(A)^{1/p},$$

so $\nu \ll \mu$. By Radon–Nikodym, $\mathrm{d}\nu = g\,\mathrm{d}\mu$. We need $g \in L^q$.

For simple functions $s = \sum_k c_k \mathbf{1}_{A_k}$: $\varphi(s) = \int s\,g\,\mathrm{d}\mu$. By density of simple functions in L^p, $\varphi(f) = \int f\,g\,\mathrm{d}\mu$ for all $f \in L^p$ (once we verify integrability).

To show $g \in L^q$: let $g_n = g\,\mathbf{1}_{\{|g| \leq n\}}$. Set $f_n = |g_n|^{q-2}\,\overline{g_n}$, so $|f_n|^p = |g_n|^{(q-1)p} = |g_n|^q$. Then

$$\int |g_n|^q\,\mathrm{d}\mu = \int f_n\,g\,\mathrm{d}\mu = \varphi(f_n) \leq \|\varphi\|\,\|f_n\|_p = \|\varphi\| \left(\int |g_n|^q\,\mathrm{d}\mu \right)^{1/p}.$$

Hence $\|g_n\|_q^{q/q} = \|g_n\|_q^q / \|g_n\|_q^{q/p} \leq \|\varphi\|$, i.e., $\|g_n\|_q \leq \|\varphi\|$. By monotone convergence, $\|g\|_q \leq \|\varphi\| < \infty$. $\qquad\square$

Remark 11.11. The dual of L^∞ is *not* L^1 in general; $(L^\infty(\mu))'$ is strictly larger, consisting of all finitely additive signed measures absolutely continuous with respect to μ. One has $L^1 \hookrightarrow (L^\infty)'$ isometrically, but the inclusion is proper when μ is not purely atomic.

Remark 11.12. Since $(L^p)' \cong L^q$ and $(L^q)' \cong L^p$ for $1 < p < \infty$, the spaces L^p are *reflexive* for $1 < p < \infty$. The spaces L^1 and L^∞ are not reflexive (except in trivial cases).

11.6 L^2 as a Hilbert Space

Proposition 11.13. $L^2(\Omega, \mu)$ *is a Hilbert space with inner product* $\langle f, g \rangle = \int_\Omega f \, \bar{g} \, \mathrm{d}\mu$. *The Riesz representation theorem for Hilbert spaces gives* $(L^2)' \cong L^2$ *directly, consistent with Theorem 11.10 at* $p = q = 2$.

11.7   The Radon–Nikodym Theorem

We state the theorem used in the proof of Theorem 11.10 and give a proof using Hilbert space methods (following von Neumann).

Theorem 11.14 (Radon–Nikodym)**.** *Let* μ *and* ν *be* σ-*finite measures on* $(\Omega, \mathcal{A})$ *with* $\nu \ll \mu$ *(i.e.,* $\mu(A) = 0 \Rightarrow \nu(A) = 0$*). Then there exists a measurable function* $g \geq 0$ *(unique* μ-*a.e.) such that* $\mathrm{d}\nu = g \, \mathrm{d}\mu$, *i.e.,* $\nu(A) = \int_A g \, \mathrm{d}\mu$ *for all* $A \in \mathcal{A}$. *The function* $g = \frac{\mathrm{d}\nu}{\mathrm{d}\mu}$ *is the* Radon–Nikodym *derivative.*

Proof (von Neumann). Set $\lambda = \mu + \nu$. Then $L^2(\lambda) \to \mathbb{R}$, $f \mapsto \int f \, \mathrm{d}\nu$, is a bounded linear functional (since $\left| \int f \, \mathrm{d}\nu \right| \leq \int |f| \, \mathrm{d}\nu \leq \int |f| \, \mathrm{d}\lambda$ and by Cauchy–Schwarz $\leq \lambda(\Omega)^{1/2} \|f\|_{L^2(\lambda)}$ when $\lambda(\Omega) < \infty$; the σ-finite case requires a standard exhaustion argument).

By the Riesz representation theorem for $L^2(\lambda)$, there exists $h \in L^2(\lambda)$ with $\int f \, \mathrm{d}\nu = \int f h \, \mathrm{d}\lambda$ for all $f \in L^2(\lambda)$. Taking $f = \mathbf{1}_A$: $\nu(A) = \int_A h \, \mathrm{d}\lambda$. Since $0 \leq \nu(A) \leq \lambda(A)$, we get $0 \leq h \leq 1$ λ-a.e.

Now $\int f \, \mathrm{d}\nu = \int f h \, \mathrm{d}\mu + \int f h \, \mathrm{d}\nu$, so $\int f(1 - h) \, \mathrm{d}\nu = \int f h \, \mathrm{d}\mu$. On $\{h = 1\}$, this gives $\mu(\{h = 1\}) = 0$ (take $f = \mathbf{1}_{\{h=1\}}$; the left side is 0). Set

$g = h/(1 - h)$ on $\{h < 1\}$ and $g = 0$ on $\{h = 1\}$. Then for any measurable A, taking $f = \mathbf{1}_A/(1 - h)$ (which is integrable) yields $\nu(A) = \int_A g \, d\mu$. $\qquad \square$

11.8 Interpolation: The Riesz–Thorin Theorem

Theorem 11.15 (Riesz–Thorin). *Let (Ω_1, μ_1) and (Ω_2, μ_2) be σ-finite measure spaces. Let $1 \leq p_0, p_1, q_0, q_1 \leq \infty$ and let T be a linear operator defined on $L^{p_0}(\mu_1) + L^{p_1}(\mu_1)$ satisfying*

$$\|Tf\|_{q_0} \leq M_0 \|f\|_{p_0}, \qquad \|Tf\|_{q_1} \leq M_1 \|f\|_{p_1}.$$

Then for every $0 < \theta < 1$, with

$$\frac{1}{p} = \frac{1 - \theta}{p_0} + \frac{\theta}{p_1}, \qquad \frac{1}{q} = \frac{1 - \theta}{q_0} + \frac{\theta}{q_1},$$

we have $\|Tf\|_q \leq M_0^{1-\theta} M_1^\theta \|f\|_p$ for all $f \in L^p(\mu_1)$.

Proof. We use the Hadamard three-lines lemma. Consider the strip $S = \{z \in \mathbb{C} : 0 \leq \operatorname{Re} z \leq 1\}$.

Step 1 (Reduction). By duality, it suffices to show that for simple functions $f = \sum_j a_j \mathbf{1}_{A_j}$ and $g = \sum_k b_k \mathbf{1}_{B_k}$ with $\|f\|_p = \|g\|_{q'} = 1$:

$$\left| \int (Tf) \, \bar{g} \, d\mu_2 \right| \leq M_0^{1-\theta} M_1^\theta.$$

Step 2 (Analytic family). Define, for $z \in S$:

$$f_z(x) = |f(x)|^{p\left(\frac{1-z}{p_0} + \frac{z}{p_1}\right)} \frac{f(x)}{|f(x)|}, \qquad g_z(y) = |g(y)|^{q'\left(\frac{1-z}{q_0'} + \frac{z}{q_1'}\right)} \frac{g(y)}{|g(y)|}.$$

(Set $f_z = 0$ where $f = 0$, similarly for g_z.) At $z = \theta$: $f_\theta = f$, $g_\theta = g$.

Define $F(z) = \int (Tf_z) \, \overline{g_z} \, d\mu_2$. Then F is continuous on S, analytic on the interior (since f and g are simple functions), and bounded on S.

Step 3 (Boundary estimates). On $\operatorname{Re} z = 0$: $\|f_{it}\|_{p_0} = 1$ and $\|g_{it}\|_{q_0'} = 1$, so $|F(it)| \leq \|Tf_{it}\|_{q_0} \|g_{it}\|_{q_0'} \leq M_0$. Similarly, $|F(1 + it)| \leq M_1$ on $\operatorname{Re} z = 1$.

Step 4 (Three-lines lemma). By the Hadamard three-lines lemma (if F is bounded and continuous on S, analytic inside, with $|F| \leq M_j$ on the line $\operatorname{Re} z = j$ for $j = 0, 1$, then $|F(z)| \leq M_0^{1-\operatorname{Re} z} M_1^{\operatorname{Re} z}$):

$$\left| \int (Tf) \bar{g} \, d\mu_2 \right| = |F(\theta)| \leq M_0^{1-\theta} M_1^\theta. \qquad \square$$

11.9 Application: Fourier Transform on L^1 and L^2

Definition 11.16. For $f \in L^1(\mathbb{R}^d)$, the *Fourier transform* is

$$\hat{f}(\xi) = \int_{\mathbb{R}^d} f(x) \, e^{-2\pi i x \cdot \xi} \, \mathrm{d}x, \qquad \xi \in \mathbb{R}^d.$$

Proposition 11.17. *The Fourier transform $\mathcal{F} \colon L^1(\mathbb{R}^d) \to C_0(\mathbb{R}^d)$ is a bounded linear map with $\left\| \hat{f} \right\|_\infty \leq \|f\|_1$.*

Theorem 11.18 (Plancherel)**.** *The Fourier transform extends uniquely to a unitary operator $\mathcal{F} \colon L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$:*

$$\left\| \hat{f} \right\|_{L^2} = \|f\|_{L^2} \qquad \text{for all } f \in L^2(\mathbb{R}^d).$$

Proof. For $f \in L^1 \cap L^2$ (which is dense in L^2), one verifies $\left\| \hat{f} \right\|_2 = \|f\|_2$ by computing $\left\langle \hat{f}, \hat{g} \right\rangle_2 = \langle f, g \rangle_2$ for Schwartz functions and extending by density.

Alternatively, by the Riesz–Thorin theorem: $\mathcal{F}$ maps $L^1 \to L^\infty$ with norm 1 and (once established on a dense subspace) $L^2 \to L^2$ with norm 1. By interpolation at $\theta = 1/2$ (giving $p = 4/3$, $q = 4$), one obtains the Hausdorff–Young inequality $\left\| \hat{f} \right\|_q \leq \|f\|_p$ for $1 \leq p \leq 2$, $q = p'$. $\qquad \square$

11.10 Exercises

Exercise 11.1 ($\star$)**.** Characterize equality in Hölder's inequality.

Exercise 11.2 ($\star$)**.** Let $\mu(\Omega) < \infty$. Show that $L^q(\Omega) \subset L^p(\Omega)$ for $1 \leq p \leq q \leq \infty$ with $\|f\|_p \leq \mu(\Omega)^{1/p - 1/q} \|f\|_q$. Give a counterexample when $\mu(\Omega) = \infty$.

Exercise 11.3 ($\star\star$). Show that $L^p(\mathbb{R}^d)$ is separable for $1 \leq p < \infty$ but $L^\infty(\mathbb{R}^d)$ is not.

Exercise 11.4 ($\star\star$). Let $1 < p < \infty$ and $f_n \rightharpoonup f$ in L^p. Show $\|f\|_p \leq \liminf \|f_n\|_p$. If additionally $\|f_n\|_p \to \|f\|_p$, show $f_n \to f$ in L^p.

Exercise 11.5 ($\star\star$). Prove directly that $(\ell^1)' \cong \ell^\infty$.

Exercise 11.6 ($\star\star\star$). (Dunford–Pettis.)   A subset $K \subset L^1(\mu)$ is weakly sequentially compact if and only if it is bounded and uniformly integrable. Prove this.

Exercise 11.7 ($\star$). Show $f * g * h = f * (g * h)$ for appropriate f, g, h in L^p spaces.

Exercise 11.8 ($\star\star$). Show the Fourier transform is a bijection on the Schwartz space $\mathcal{S}(\mathbb{R}^d)$ and satisfies $\widehat{D^\alpha f}(\xi) = (2\pi i \xi)^\alpha \hat{f}(\xi)$.

Exercise 11.9 ($\star\star\star$). (Marcinkiewicz interpolation.) Prove the weak-type version of Riesz–Thorin: if T is weak-(p_0, q_0) and weak-(p_1, q_1), then T is strong-(p, q) for $p_0 < p < p_1$.

Chapter 12

Banach Algebras and C^*-Algebras — Introduction

Contents

12.1 Banach Algebras

Definition 12.1 (Banach algebra)**.** A *Banach algebra* is a Banach space $(A, \|\cdot\|)$ equipped with a bilinear multiplication $A \times A \to A$ satisfying:

(i) $\|ab\| \leq \|a\| \, \|b\|$ for all $a, b \in A$ (submultiplicativity).

(ii) Associativity: $(ab)c = a(bc)$.

If there exists a unit $e \in A$ with $ea = ae = a$ for all a and $\|e\| = 1$, we say A is *unital*.

Example 12.2. (a) $\mathcal{L}(E)$ with operator composition, for any Banach space E.

(b) $C(K)$ with pointwise multiplication, for a compact Hausdorff space K. This is a commutative unital Banach algebra.

(c) $\ell^1(\mathbb{Z})$ with convolution: $(a * b)_n = \sum_{k \in \mathbb{Z}} a_k b_{n-k}$. The unit is δ_0. This is a commutative unital Banach algebra.

(d) $L^1(\mathbb{R})$ with convolution. This is a commutative Banach algebra *without* unit.

12.2 Spectrum and Spectral Radius

Definition 12.3. Let A be a unital Banach algebra and $a \in A$. The *spectrum* of a is

$$\mathrm{Spec}(a) = \{\lambda \in \mathbb{C} : \lambda e - a \text{ is not invertible in } A\}.$$

The *spectral radius* is $r(a) = \sup\{|\lambda| : \lambda \in \mathrm{Spec}(a)\}$.

Proposition 12.4. *If $\|a\| < 1$ in a unital Banach algebra, then $e - a$ is invertible with $(e - a)^{-1} = \sum_{n=0}^{\infty} a^n$ and $\|(e - a)^{-1}\| \leq (1 - \|a\|)^{-1}$.*

Proof. The series converges absolutely since $\sum \|a^n\| \le \sum \|a\|^n < \infty$. Then $(e - a) \sum_{n=0}^{N} a^n = e - a^{N+1} \to e$. $\square$

> **Theorem 12.5** (Spectral radius formula)*. For a in a unital Banach algebra:*
> $$r(a) = \lim_{n \to \infty} \|a^n\|^{1/n} = \inf_{n \ge 1} \|a^n\|^{1/n}.$$
>
> *Moreover,* $\mathrm{Spec}(a)$ *is a nonempty compact subset of the disk* $\{|\lambda| \le \|a\|\}$.

Proof. **Step 1** ($\mathrm{Spec}(a)$ **is compact and nonempty**). If $|\lambda| > \|a\|$, then $\lambda e - a = \lambda(e - a/\lambda)$ is invertible by Proposition 12.4, so $\lambda \notin \mathrm{Spec}(a)$. Hence $\mathrm{Spec}(a) \subset \overline{B}(0, \|a\|)$.

The map $\lambda \mapsto \lambda e - a$ is continuous $\mathbb{C} \to A$, and the set of invertible elements $A^\times$ is open (if b is invertible and $\|c - b\| < \|b^{-1}\|^{-1}$, then $c = b(e - b^{-1}(b - c))$ is invertible). Hence $\mathbb{C} \setminus \mathrm{Spec}(a)$ is open, so $\mathrm{Spec}(a)$ is closed, thus compact.

Nonemptiness: for any $\varphi \in A'$, the function $f(\lambda) = \varphi((\lambda e - a)^{-1})$ is analytic on $\mathbb{C} \setminus \mathrm{Spec}(a)$. If $\mathrm{Spec}(a) = \emptyset$, f is entire. For $|\lambda| > \|a\|$: $|f(\lambda)| \le \|\varphi\| \, \|(\lambda e - a)^{-1}\| \le \|\varphi\| / (|\lambda| - \|a\|) \to 0$ as $|\lambda| \to \infty$. By Liouville's theorem, $f \equiv 0$. Since this holds for all $\varphi \in A'$, Hahn–Banach gives $(\lambda e - a)^{-1} = 0$, a contradiction. Hence $\mathrm{Spec}(a) \neq \emptyset$.

Step 2 ($r(a) = \lim \|a^n\|^{1/n}$). If $\lambda \in \mathrm{Spec}(a)$, then $\lambda^n \in \mathrm{Spec}(a^n)$ (since $\lambda^n e - a^n$ factors), so $|\lambda|^n \le \|a^n\|$, giving $r(a) \le \inf_n \|a^n\|^{1/n}$.

For the reverse, let $R = \limsup_n \|a^n\|^{1/n}$. The resolvent $R(\lambda) = (\lambda e - a)^{-1} = \sum_{n=0}^{\infty} a^n / \lambda^{n+1}$ converges for $|\lambda| > R$ (by the root test). Since this Laurent series converges on $\mathbb{C} \setminus \mathrm{Spec}(a)$ and diverges at points of $\mathrm{Spec}(a)$, we have $r(a) \ge R$.

To see $R = \lim$ (not just $\limsup$): let $\alpha = \inf_n \|a^n\|^{1/n}$. For any n, writing $m = nq + r$ with $0 \le r < n$: $\|a^m\|^{1/m} \le \|a^n\|^{q/m} \|a\|^{r/m}$. Taking $\limsup$ as $m \to \infty$ (so $q/m \to 1/n$, $r/m \to 0$): $R \le \|a^n\|^{1/n}$ for all n, hence $R \le \alpha$. Combined with $\alpha \le R$, we get $\alpha = R = \lim_n \|a^n\|^{1/n}$. $\square$

12.3 Ideals, Quotients, and Morphisms

> **Definition 12.6.** A *(two-sided) ideal* in a Banach algebra A is a closed subspace I such that $aI \subset I$ and $Ia \subset I$ for all $a \in A$. An ideal is *maximal* if it is proper ($I \neq A$) and not contained in any larger proper ideal.

> **Proposition 12.7.** *If I is a closed ideal in A, then A/I is a Banach algebra with $\|a + I\| = \inf_{x \in I} \|a + x\|$, the quotient norm. If A is unital, so is A/I (with unit $e + I$).*

> **Definition 12.8.** A *morphism* of Banach algebras is a bounded linear map $\phi \colon A \to B$ satisfying $\phi(ab) = \phi(a)\phi(b)$. If A and B are unital and $\phi(e_A) = e_B$, we say ϕ is *unital*.

12.4 Commutative Banach Algebras: Characters and Gelfand Space

> **Definition 12.9.** Let A be a commutative unital Banach algebra. A *character* (or *multiplicative linear functional*) is a nonzero algebra homomorphism $\chi \colon A \to \mathbb{C}$, i.e., $\chi(ab) = \chi(a)\chi(b)$ and $\chi \neq 0$.

> **Proposition 12.10.** *Let χ be a character of a commutative unital Banach algebra A.*
>
> *(i) χ is continuous with $\|\chi\| = 1$.*
>
> *(ii) $\chi(e) = 1$.*
>
> *(iii) $\operatorname{Ker} \chi$ is a maximal ideal.*
>
> *(iv) $\chi(a) \in \operatorname{Spec}(a)$ for all $a \in A$.*

Proof. (ii) $\chi(e) = \chi(e \cdot e) = \chi(e)^2$ and $\chi \neq 0$, so $\chi(e) = 1$.

(iv) $\chi(a)e - a \in \operatorname{Ker} \chi$ which is a proper ideal, hence $\chi(a)e - a$ is not invertible.

(i) From (iv), $|\chi(a)| \leq r(a) \leq \|a\|$, so $\|\chi\| \leq 1$. Since $\chi(e) = 1$, $\|\chi\| = 1$.

(iii) Ker χ has codimension 1 (since χ is surjective onto $\mathbb{C}$) and is an ideal, hence maximal. $\qquad\square$

> **Definition 12.11** (Gelfand space)**.** The *Gelfand space* (or *maximal ideal space*, or *character space*) of A is the set $\hat{A}$ of all characters of A, equipped with the weak-$*$ topology inherited from A'.

> **Proposition 12.12.** $\hat{A}$ *is a compact Hausdorff space (in the weak-$*$ topology).*

Proof. By Proposition 12.10, $\hat{A} \subset \{\varphi \in A' : \|\varphi\| \leq 1\}$, which is weak-$*$ compact by Banach–Alaoglu. It suffices to show $\hat{A} \cup \{0\}$ is weak-$*$ closed. If $\chi_\alpha \to \varphi$ weak-$*$ with $\chi_\alpha \in \hat{A}$, then $\varphi(ab) = \lim \chi_\alpha(ab) = \lim \chi_\alpha(a)\chi_\alpha(b) = \varphi(a)\varphi(b)$. So φ is multiplicative, hence $\varphi \in \hat{A} \cup \{0\}$. Since $\chi(e) = 1$ for all $\chi \in \hat{A}$ and $\varphi(e) = \lim \chi_\alpha(e) = 1$, $\varphi \neq 0$, so actually $\hat{A}$ is itself closed, hence compact. $\qquad\square$

12.5 The Gelfand Transform

> **Definition 12.13.** The *Gelfand transform* is the map $\Gamma \colon A \to C(\hat{A})$ defined by
> $$\hat{a}(\chi) = \chi(a), \qquad a \in A, \ \chi \in \hat{A}.$$

> **Theorem 12.14.** *The Gelfand transform* $\Gamma \colon a \mapsto \hat{a}$ *is a unital algebra homomorphism from* A *to* $C(\hat{A})$ *with:*
>
> *(i)* $\widehat{ab} = \hat{a}\hat{b}$ *and* $\widehat{\alpha a + \beta b} = \alpha\hat{a} + \beta\hat{b}$.
>
> *(ii)* $\|\hat{a}\|_\infty = r(a)$ *for all* $a \in A$.
>
> *(iii)* $\hat{a}(\hat{A}) = \operatorname{Spec}(a)$ *for all* a.
>
> *(iv)* $\operatorname{Ker}\Gamma = \{a \in A : r(a) = 0\}$ *(the* radical *of A).*

Proof. (i) is immediate: $\widehat{ab}(\chi) = \chi(ab) = \chi(a)\chi(b) = \hat{a}(\chi)\hat{b}(\chi)$.

(iii) $\lambda \in \operatorname{Spec}(a)$ iff $\lambda e - a$ is not invertible. In a commutative unital Banach algebra, a noninvertible element is contained in a maximal ideal, and every maximal ideal is Ker χ for some $\chi \in \hat{A}$ (by Zorn's lemma and the

correspondence between maximal ideals and characters). Hence $\lambda \in \mathrm{Spec}(a)$ iff $\chi(\lambda e - a) = 0$ for some χ, iff $\lambda = \chi(a)$ for some χ.

(ii) $\|\hat{a}\|_\infty = \sup_\chi |\chi(a)| = \sup_{\lambda \in \mathrm{Spec}(a)} |\lambda| = r(a)$.

(iv) $\hat{a} = 0$ iff $\|\hat{a}\|_\infty = 0$ iff $r(a) = 0$. $\qquad\square$

12.6 C^*-Algebras

Definition 12.15 (C^*-algebra). A C^*-*algebra* is a Banach algebra A equipped with an *involution* $*\colon A \to A$ (an antilinear map with $(a^*)^* = a$ and $(ab)^* = b^*a^*$) satisfying the C^*-*identity*:

$$\|a^*a\| = \|a\|^2 \qquad \text{for all } a \in A.$$

Example 12.16. (a) $\mathcal{L}(H)$ with $T^* = $ adjoint, for a Hilbert space H.

(b) $C(K)$ with $f^* = \bar{f}$, for compact Hausdorff K.

(c) Any norm-closed $*$-subalgebra of $\mathcal{L}(H)$.

Proposition 12.17. *In a C^*-algebra: $\|a\| = r(a)$ for every normal element ($a^*a = aa^*$). In particular, $\|a\| = r(a)$ for self-adjoint a.*

Proof. If a is normal, $\|a^2\|^2 = \|(a^2)^*(a^2)\| = \|(a^*a)^2\| = \|a^*a\|^2 = \|a\|^4$, so $\|a^2\| = \|a\|^2$. By induction, $\|a^{2^n}\| = \|a\|^{2^n}$, hence $r(a) = \lim \|a^n\|^{1/n} = \|a\|$. $\qquad\square$

12.7 The Commutative Gelfand–Naimark Theorem

Theorem 12.18 (Commutative Gelfand–Naimark). *Let A be a commutative unital C^*-algebra. The Gelfand transform $\Gamma\colon A \to C(\hat{A})$ is an isometric $*$-isomorphism.*

Proof. **Step 1 (Γ is a $*$-homomorphism).** We know Γ is an algebra homomorphism. We must show $\widehat{a^*}(\chi) = \overline{\hat{a}(\chi)}$, i.e., $\chi(a^*) = \overline{\chi(a)}$.

Every element of a C^*-algebra can be written $a = h + ik$ with $h = (a + a^*)/2$ and $k = (a - a^*)/(2i)$ self-adjoint. It suffices to show $\chi(h) \in \mathbb{R}$ for self-adjoint h.

For self-adjoint h and $t \in \mathbb{R}$, set $u_t = e^{ith}$ (defined by the power series, which converges). One checks $u_t^* = e^{-ith} = u_t^{-1}$, so u_t is unitary: $\|u_t\|^2 = \|u_t^* u_t\| = \|e\| = 1$. Hence $|\chi(u_t)| \leq \|u_t\| = 1$ and $|\chi(u_t)^{-1}| = |\chi(u_{-t})| \leq 1$, so $|\chi(u_t)| = 1$.

Now $\chi(u_t) = e^{it\chi(h)}$ (since χ is an algebra homomorphism preserving the exponential series). Hence $\left| e^{it\chi(h)} \right| = 1$ for all $t \in \mathbb{R}$, which forces $\chi(h) \in \mathbb{R}$.

Step 2 (Γ is isometric). Since A is commutative, every element is normal (as $a^* a = a a^*$ when multiplication commutes). By Proposition 12.17, $\|a\| = r(a) = \|\hat{a}\|_\infty$.

Step 3 (Γ is surjective). By the isometry, $\Gamma(A)$ is a closed $*$-subalgebra of $C(\hat{A})$ containing the constants (since $\hat{e} = 1$) and separating points of $\hat{A}$ (if $\chi_1 \neq \chi_2$, there exists a with $\chi_1(a) \neq \chi_2(a)$, i.e., $\hat{a}(\chi_1) \neq \hat{a}(\chi_2)$). Since Γ is a $*$-map, $\Gamma(A)$ is closed under conjugation. By the Stone–Weierstrass theorem, $\Gamma(A) = C(\hat{A})$. $\qquad\square$

12.8 Continuous Functional Calculus in C^*-Algebras

> **Theorem 12.19.** *Let A be a unital C^*-algebra and $a \in A$ normal. There exists a unique isometric $*$-isomorphism $\Phi_a \colon C(\mathrm{Spec}(a)) \to C^*(a, e)$ (the C^*-subalgebra generated by a and e) with $\Phi_a(\mathrm{id}) = a$ and $\Phi_a(1) = e$.*

Proof. $C^*(a, e)$ is a commutative C^*-algebra (since a is normal). By the Gelfand–Naimark theorem, $C^*(a, e) \cong C(\widehat{C^*(a, e)})$. The evaluation map $\chi \mapsto \chi(a)$ is a homeomorphism $\widehat{C^*(a, e)} \to \mathrm{Spec}(a)$ (it is continuous and injective—since χ is determined by $\chi(a)$—and surjective by Theorem 12.14(iii)). The composition gives the desired isomorphism $C(\mathrm{Spec}(a)) \cong C^*(a, e)$. $\quad\square$

12.9 States and Positive Forms

> **Definition 12.20.** A linear functional $\omega \colon A \to \mathbb{C}$ on a C^*-algebra is *positive* if $\omega(a^* a) \geq 0$ for all $a \in A$.

Definition 12.21. A *state* on a unital C^*-algebra A is a positive linear functional ω with $\omega(e) = 1$. The set of all states is the *state space* $S(A)$.

Proposition 12.22. *Let ω be a state on a unital C^*-algebra A.*

   (i) $\omega(a^*) = \overline{\omega(a)}$.

  (ii) $|\omega(a)|^2 \leq \omega(a^*a)$ (Cauchy–Schwarz).

 (iii) $\|\omega\| = \omega(e) = 1$.

 (iv) $S(A)$ is convex and weak-$*$ compact.

Proof. (i)–(ii): The map $(a, b) \mapsto \omega(b^*a)$ is a positive semidefinite sesquilinear form, so (i) and (ii) follow from standard properties.

(iii): $|\omega(a)|^2 \leq \omega(a^*a) \leq \|\omega\| \|a^*a\| = \|\omega\| \|a\|^2$, hence $\|\omega\| \leq$ itself (circular unless we note $\omega(e) = 1$ gives $\|\omega\| \geq 1$, and the Cauchy–Schwarz bound gives $|\omega(a)| \leq \omega(e)^{1/2}\omega(a^*a)^{1/2} \leq \omega(e)^{1/2} \|a^*a\|^{1/2} \|\omega\|^{1/2}$... more cleanly: $|\omega(a)| \leq \|a\|$ follows from $|\omega(a)|^2 \leq \omega(a^*a) \leq \|a^*a\| = \|a\|^2$ once we know ω is bounded, which follows from positivity on a C^*-algebra). Thus $\|\omega\| = 1$.

(iv): Convexity: if ω_1, ω_2 are states and $0 \leq t \leq 1$, then $t\omega_1 + (1 - t)\omega_2$ is positive with value 1 at e. Weak-$*$ compactness: $S(A) \subset \{\varphi \in A' : \|\varphi\| \leq 1\}$ is weak-$*$ closed (being defined by the conditions $\omega(e) = 1$ and $\omega(a^*a) \geq 0$), hence compact by Banach–Alaoglu. $\square$

12.10 The GNS Construction

The Gelfand–Naimark–Segal (GNS) construction produces a Hilbert space representation from a state.

Theorem 12.23 (GNS construction). *Let A be a unital C^*-algebra and ω a state on A. There exist a Hilbert space H_ω, a unit vector $\xi_\omega \in H_\omega$, and a $*$-representation $\pi_\omega : A \to \mathcal{L}(H_\omega)$ such that:*

  (i) $\omega(a) = \langle \pi_\omega(a)\xi_\omega, \xi_\omega \rangle$ for all $a \in A$.

 (ii) ξ_ω is cyclic: $\overline{\pi_\omega(A)\xi_\omega} = H_\omega$.

The triple $(H_\omega, \pi_\omega, \xi_\omega)$ is unique up to unitary equivalence.

Proof. **Step 1 (Inner product on A).** Define the sesquilinear form on A: $\langle a, b \rangle_\omega = \omega(b^*a)$. This is positive semidefinite: $\langle a, a \rangle_\omega = \omega(a^*a) \geq 0$.

Step 2 (Quotient by null space). Let $N_\omega = \{a \in A : \omega(a^*a) = 0\}$. By Cauchy–Schwarz ($|\omega(b^*a)|^2 \leq \omega(a^*a)\omega(b^*b)$), N_ω is a left ideal of A (if $a \in N_\omega$, then for any $b \in A$: $\omega((ba)^*ba) = \omega(a^*b^*ba) \leq \|b^*b\|\,\omega(a^*a) = 0$ by positivity and C^*-identity reasoning). The quotient A/N_ω inherits an inner product: $\langle a + N_\omega, b + N_\omega \rangle = \omega(b^*a)$.

Step 3 (Completion). Let H_ω be the completion of A/N_ω.

Step 4 (Representation). For $a \in A$, define $\pi_\omega(a)(b + N_\omega) = ab + N_\omega$. This is well-defined since N_ω is a left ideal. We check boundedness:

$$\|\pi_\omega(a)(b + N_\omega)\|^2 = \omega(b^*a^*ab) \leq \|a^*a\|\,\omega(b^*b) = \|a\|^2\,\|b + N_\omega\|^2,$$

where the inequality uses $a^*a \leq \|a\|^2\, e$ (in the order of self-adjoint elements) and positivity of ω. Hence $\|\pi_\omega(a)\| \leq \|a\|$ and $\pi_\omega(a)$ extends to H_ω.

One verifies $\pi_\omega(ab) = \pi_\omega(a)\pi_\omega(b)$ and $\pi_\omega(a^*) = \pi_\omega(a)^*$ (from the definition of the inner product).

Step 5 (Cyclic vector). Set $\xi_\omega = e + N_\omega \in H_\omega$. Then $\pi_\omega(a)\xi_\omega = a + N_\omega$, so $\pi_\omega(A)\xi_\omega = A/N_\omega$, which is dense in H_ω by construction.

Also, $\langle \pi_\omega(a)\xi_\omega, \xi_\omega \rangle = \omega(e^* \cdot a \cdot e) = \omega(a)$.

Uniqueness. If (H', π', ξ') is another GNS triple, define $U \colon \pi_\omega(a)\xi_\omega \mapsto \pi'(a)\xi'$. This is isometric (since both inner products equal $\omega(a^*a)$) and extends to a unitary $U \colon H_\omega \to H'$ with $U\pi_\omega(a) = \pi'(a)U$ and $U\xi_\omega = \xi'$. $\square$

> **Corollary 12.24** (Gelfand–Naimark embedding). *Every C^*-algebra admits an isometric $*$-representation on a Hilbert space.*

Proof. For each $a \in A$ with $a \neq 0$, there exists a state ω with $\omega(a^*a) = \|a\|^2$ (by the Hahn–Banach theorem applied to the C^*-subalgebra generated by a^*a). The direct sum $\pi = \bigoplus_{\omega \in S(A)} \pi_\omega$ on $H = \bigoplus_\omega H_\omega$ gives a faithful $*$-representation. $\square$

12.11 Application: Representation Theory and Quantum Mechanics

Remark 12.25 (Algebraic quantum mechanics). In the algebraic approach to quantum mechanics:

- Observables form a C^*-algebra A (self-adjoint elements represent physical quantities).

- States are positive normalized linear functionals $\omega \in S(A)$.

- The GNS construction recovers the Hilbert space formulation: given a state ω, the observables act on H_ω via π_ω.

- Different states may give rise to unitarily inequivalent representations, which is related to the phenomenon of *spontaneous symmetry breaking* and *superselection sectors* in quantum field theory.

12.12 Exercises

Exercise 12.1 ($\star$). Compute $\mathrm{Spec}(a)$ for: (a) the identity e; (b) an idempotent $p = p^2$; (c) a nilpotent element $a^n = 0$.

Exercise 12.2 ($\star\star$). Identify the Gelfand space of $\ell^1(\mathbb{Z})$ (with convolution product) with the unit circle $\mathbb{T}$, and the Gelfand transform with the Fourier series.

Exercise 12.3 ($\star$). Show that a C^*-algebra has a unique C^*-norm: if $\|\cdot\|_1$ and $\|\cdot\|_2$ are two norms making A a C^*-algebra, then $\|a\|_1 = \|a\|_2$ for all a.

Exercise 12.4 ($\star\star$). Let A be a C^*-algebra and $a \in A$ self-adjoint. Show $a \geq 0$ if and only if $\mathrm{Spec}(a) \subset [0, \infty)$.

Exercise 12.5 ($\star\star$). A state ω is *pure* if it is an extreme point of $S(A)$. Show that ω is pure if and only if π_ω is irreducible.

Exercise 12.6 ($\star \star \star$). Using the GNS construction and Corollary 12.24, prove that every C^*-algebra A is isometrically $*$-isomorphic to a norm-closed $*$-subalgebra of $\mathcal{L}(H)$ for some Hilbert space H.

Exercise 12.7 ($\star\star$). (Wiener's lemma.) Let $f \in \ell^1(\mathbb{Z})$ have absolutely convergent Fourier series $\hat{f}(\theta) = \sum_n a_n e^{in\theta}$. Show that if $\hat{f}$ never vanishes on $\mathbb{T}$, then $1/\hat{f}$ also has absolutely convergent Fourier series. *Hint:* use the Gelfand theory of $\ell^1(\mathbb{Z})$.

Exercise 12.8 ($\star\star$). (Kadison's inequality.) Let $\pi \colon A \to \mathcal{L}(H)$ be a unital $*$-homomorphism. Show that $\pi(a)^*\pi(a) \leq \pi(a^*a)$ for all a.

Exercise 12.9 ($\star \star \star$). (Spectral permanence.) Let B be a C^*-subalgebra of a C^*-algebra A with the same unit. Show that $\mathrm{Spec}_B(b) = \mathrm{Spec}_A(b)$ for all $b \in B$.

Appendix A

Review of Analysis and Topology

Contents

This appendix collects fundamental results from analysis and topology that are used throughout the text. Proofs are included for the reader's convenience.

A.1 Baire's Category Theorem

Theorem A.1 (Baire category theorem)**.** *Let X be a complete metric space (or, more generally, a locally compact Hausdorff space). If $(U_n)_{n \geq 1}$ is a sequence of dense open subsets of X, then $\bigcap_{n=1}^{\infty} U_n$ is dense in X.*

Equivalently, X cannot be written as a countable union of closed sets with empty interior: if $X = \bigcup_n F_n$ with F_n closed, then at least one F_n has nonempty interior.

Proof. Let $B_0 \subset X$ be a nonempty open ball. Since U_1 is dense, $B_0 \cap U_1 \neq \emptyset$; choose a closed ball $\overline{B}_1 \subset B_0 \cap U_1$ with radius $r_1 < 1$. Since U_2 is dense, $B_1 \cap U_2 \neq \emptyset$; choose $\overline{B}_2 \subset B_1 \cap U_2$ with $r_2 < 1/2$. Continue inductively: $\overline{B}_{n+1} \subset B_n \cap U_{n+1}$, $r_n < 1/n$.

The centers (x_n) form a Cauchy sequence (since $x_m, x_n \in B_N$ for $m, n \geq N$, with diam $B_N < 2/N$). By completeness, $x_n \to x$. Since $x \in \overline{B}_n$ for all n (as the nested intersection of closed sets), $x \in \overline{B}_n \subset U_n$ for all n. Hence $x \in B_0 \cap \bigcap_n U_n$, proving density. $\qquad\square$

> *Remark* A.2. The three pillars of functional analysis—the uniform boundedness principle, the open mapping theorem, and the closed graph theorem—all rely on Baire's theorem. In this text, it is invoked in Chapters 2, 3, and 4.

A.2   Zorn's Lemma

> **Theorem A.3** (Zorn's lemma)**.** *Let $(P, \leq)$ be a nonempty partially ordered set in which every totally ordered subset (chain) has an upper bound. Then P has at least one maximal element.*

Zorn's lemma is equivalent to the axiom of choice and is used throughout functional analysis, notably in proving the Hahn–Banach theorem, the existence of bases (Hamel bases), and the existence of maximal ideals.

A.3   Tychonoff's Theorem

> **Theorem A.4** (Tychonoff)**.** *An arbitrary product of compact topological spaces is compact (in the product topology).*

Proof sketch using Alexander's subbasis theorem. By Alexander's subbasis theorem, it suffices to show that every cover by subbasis elements has a finite subcover. The subbasis for the product topology $\prod_\alpha X_\alpha$ consists of sets $\pi_\alpha^{-1}(U_\alpha)$ where $U_\alpha \subset X_\alpha$ is open. Suppose $\prod X_\alpha \subset \bigcup_{i \in I} \pi_{\alpha_i}^{-1}(U_{\alpha_i})$ has no finite subcover. For each index α, consider the open sets U_{α_i} with $\alpha_i = \alpha$. If these cover X_α, a finite subcover of X_α (by compactness) gives a finite subcover of the product restricted to that coordinate—one checks this yields a contradiction. Hence for each α, X_α is not covered, and one can choose $x_\alpha \notin \bigcup_i U_{\alpha_i}$ (for those i with $\alpha_i = \alpha$). The point $(x_\alpha)_\alpha$ then lies in no set of the cover, contradicting the assumption that it was a cover. $\qquad\square$

> *Remark* A.5. Tychonoff's theorem is equivalent to the axiom of choice. It is used in the proof of the Banach–Alaoglu theorem (weak-∗ compactness of the unit ball in a dual space).

A.4   Lebesgue Measure and Integration

We recall the essential facts. Let $(\Omega, \mathcal{A}, \mu)$ be a measure space.

> **Definition A.6.** A function $f \colon \Omega \to \overline{\mathbb{R}}$ (or $\mathbb{C}$) is *measurable* if $f^{-1}(B) \in \mathcal{A}$ for every Borel set B.

> **Theorem A.7** (Monotone convergence)**.** *If $0 \le f_1 \le f_2 \le \cdots$ are measurable with $f_n \nearrow f$ pointwise, then $\int f_n \, d\mu \nearrow \int f \, d\mu$.*

> **Theorem A.8** (Dominated convergence)**.** *If $f_n \to f$ a.e., $|f_n| \le g$ a.e. for some $g \in L^1(\mu)$, then $f \in L^1(\mu)$ and $\int f_n \, d\mu \to \int f \, d\mu$.*

Proof. By Fatou's lemma applied to $g + f_n \ge 0$ and $g - f_n \ge 0$:

$$\int g + \int f \le \liminf \int (g + f_n) = \int g + \liminf \int f_n,$$

$$\int g - \int f \le \liminf \int (g - f_n) = \int g - \limsup \int f_n.$$

Hence $\int f \le \liminf \int f_n \le \limsup \int f_n \le \int f$. □

> **Theorem A.9** (Fubini–Tonelli)**.** *Let (X, μ) and (Y, ν) be σ-finite measure spaces.*
>
> (i) *(Tonelli) If $f \ge 0$ is measurable on $X \times Y$, then*
>
> $$\int_{X \times Y} f \, d(\mu \otimes \nu) = \int_X \left(\int_Y f(x, y) \, d\nu(y) \right) d\mu(x) = \int_Y \left(\int_X f(x, y) \, d\mu(x) \right) d\nu(y).$$
>
> (ii) *(Fubini) If $f \in L^1(\mu \otimes \nu)$, the same iterated integral equalities hold and the inner integrals define integrable functions a.e.*

Remark A.10. Fubini's theorem is used extensively in Chapter 11 (for convolution and duality proofs) and in Chapter 9 (for Hilbert–Schmidt operators).

A.5 Additional Useful Results

Theorem A.11 (Urysohn's lemma)**.** *If X is a normal topological space and $A, B \subset X$ are disjoint closed sets, there exists a continuous function $f \colon X \to [0, 1]$ with $f|_A = 0$ and $f|_B = 1$.*

Theorem A.12 (Partition of unity)**.** *Let X be a paracompact Hausdorff space and (U_α) an open cover. There exists a partition of unity (φ_α) subordinate to (U_α): $\varphi_\alpha \geq 0$, $\operatorname{supp} \varphi_\alpha \subset U_\alpha$, $\sum_\alpha \varphi_\alpha = 1$ (locally finite sum).*

Theorem A.13 (Arzelà–Ascoli)**.** *Let K be a compact metric space. A subset $\mathcal{F} \subset C(K)$ is relatively compact if and only if it is bounded and equicontinuous.*

Theorem A.14 (Stone–Weierstrass)**.** *Let K be a compact Hausdorff space and $A \subset C(K, \mathbb{R})$ a subalgebra that separates points and contains the constants. Then A is dense in $C(K, \mathbb{R})$. The complex version requires A to be closed under conjugation.*

Bibliography

[1] H. G. Dales, *Banach Algebras and Automatic Continuity*, London Mathematical Society Monographs 24, Clarendon Press, 2000.

[2] R. C. James, Weakly compact sets, *Transactions of the AMS* **113** (1964), 129–140.

[3] N. Dunford and J. T. Schwartz, *Linear Operators, Part I: General Theory*, Interscience Publishers, 1958.

[4] R. E. Megginson, *An Introduction to Banach Space Theory*, Graduate Texts in Mathematics 183, Springer-Verlag, 1998.

Index