2.3. Hahn–Banach and Convexity

2.3.1. The Hahn–Banach Theorem. The Hahn–Banach Theorem deals with bounded linear functionals on a subspace of a Banach space X and asserts that every such functional extends to a bounded linear functional on all of X. This theorem continues to hold in the more general setting where X is any real vector space and boundedness is replaced by a bound relative to a given quasi-seminorm on X.

Definition 2.3.1 (Quasi-Seminorm). Let X be a real vector space. A function $p : X \to \mathbb{R}$ is called a **quasi-seminorm** if it satisfies

$$p(x+y) \le p(x) + p(y), \qquad p(\lambda x) = \lambda p(x) \tag{2.3.1}$$

for all $x, y \in X$ and all $\lambda \ge 0$. It is called a **seminorm** if it is a quasi-seminorm and $p(\lambda x) = |\lambda| p(x)$ for all $x \in X$ and all $\lambda \in \mathbb{R}$. A seminorm has nonnegative values, because $2p(x) = p(x) + p(-x) \ge p(0) = 0$ for all $x \in X$. Thus a seminorm satisfies all the axioms of a norm except nondegeneracy (i.e. there may be nonzero elements $x \in X$ such that $p(x) = 0$).

Theorem 2.3.2 (Hahn–Banach). *Let X be a normed vector space and let $p : X \to \mathbb{R}$ be a quasi-seminorm. Let $Y \subset X$ be a linear subspace and let $\phi : Y \to \mathbb{R}$ be a linear functional such that $\phi(x) \le p(x)$ for all $x \in Y$. Then there exists a linear functional $\Phi : X \to \mathbb{R}$ such that*

$$\Phi|_Y = \phi, \qquad \Phi(x) \le p(x) \quad \text{for all } x \in X.$$

Proof. See page 66. □

Lemma 2.3.3. *Let X, p, Y, and ϕ be as in Theorem 2.3.2. Let $x_0 \in X \setminus Y$ and define $Y' := Y \oplus \mathbb{R}x_0$. Then there exists a linear functional $\phi' : Y' \to \mathbb{R}$ such that $\phi'|_Y = \phi$ and $\phi'(x) \le p(x)$ for all $x \in Y'$.*

Proof. An extension $\phi' : Y' \to \mathbb{R}$ of the linear functional $\phi : Y \to \mathbb{R}$ is uniquely determined by its value $a := \phi'(x_0) \in \mathbb{R}$ on x_0. This extension satisfies the required condition $\phi'(x) \le p(x)$ for all $x \in Y'$ if and only if

$$\phi(y) + \lambda a \le p(y + \lambda x_0) \qquad \text{for all } y \in Y \text{ and all } \lambda \in \mathbb{R}. \tag{2.3.2}$$

If this holds, then

$$\phi(y) \pm a \le p(y \pm x_0) \qquad \text{for all } y \in Y. \tag{2.3.3}$$

Conversely, if (2.3.3) holds and $\lambda > 0$, then

$$\begin{aligned} \phi(y) + \lambda a &= \lambda \left(\phi(\lambda^{-1}y) + a\right) \le \lambda p(\lambda^{-1}y + x_0) = p(y + \lambda x_0), \\ \phi(y) - \lambda a &= \lambda \left(\phi(\lambda^{-1}y) - a\right) \le \lambda p(\lambda^{-1}y - x_0) = p(y - \lambda x_0). \end{aligned}$$

This shows that (2.3.2) is equivalent to (2.3.3). Thus it remains to find a real number $a \in \mathbb{R}$ that satisfies (2.3.3). Equivalently, a must satisfy

$$\phi(y) - p(y - x_0) \le a \le p(y + x_0) - \phi(y) \qquad \text{for all } y \in Y. \tag{2.3.4}$$

To see that such a number exists, fix two vectors $y, y' \in Y$. Then

$$\begin{aligned}\phi(y) + \phi(y') &= \phi(y + y') \\ &\le p(y + y') \\ &= p(y + x_0 + y' - x_0) \\ &\le p(y + x_0) + p(y' - x_0).\end{aligned}$$

Thus

$$\phi(y') - p(y' - x_0) \le p(y + x_0) - \phi(y)$$

for all $y, y' \in Y$ and this implies

$$\sup_{y' \in Y} \big(\phi(y') - p(y' - x_0)\big) \le \inf_{y \in Y} \big(p(y + x_0) - \phi(y)\big).$$

Hence there exists a real number $a \in \mathbb{R}$ that satisfies (2.3.4) and this proves Lemma 2.3.3. □

Proof of Theorem 2.3.2. Define the set

$$\mathscr{P} := \left\{ (Z, \psi) \;\middle|\; \begin{array}{l} Z \text{ is a linear subspace of } X \text{ and} \\ \psi : Z \to \mathbb{R} \text{ is a linear functional such that} \\ Y \subset Z,\ \psi|_Y = \phi, \text{ and } \psi(x) \le p(x) \text{ for all } x \in Z \end{array} \right\}.$$

This set is partially ordered by the relation

$$(Z, \psi) \preccurlyeq (Z', \psi') \qquad \stackrel{\text{def}}{\Longleftrightarrow} \qquad Z \subset Z' \text{ and } \psi'|_Z = \psi$$

for $(Z, \psi), (Z', \psi') \in \mathscr{P}$. A chain in $\mathscr{P}$ is a totally ordered subset $\mathscr{C} \subset \mathscr{P}$. Every nonempty chain $\mathscr{C} \subset \mathscr{P}$ has a supremum (Z_0, ψ_0) given by

$$Z_0 := \bigcup_{(Z,\psi) \in \mathscr{C}} Z, \qquad \psi_0(x) := \psi(x) \quad \text{for all } (Z, \psi) \in \mathscr{C} \text{ and all } x \in Z.$$

Hence it follows from the Lemma of Zorn that $\mathscr{P}$ has a maximal element (Z, ψ). By Lemma 2.3.3 every such maximal element satisfies $Z = X$ and this proves Theorem 2.3.2. □

A special case of the Hahn–Banach Theorem is where the quasi-seminorm is actually a norm. In this situation the Hahn–Banach Theorem is an existence result for bounded linear functionals on real and complex normed vector spaces. It takes the following form.

Corollary 2.3.4 (Real Case). *Let X be a normed vector space over $\mathbb{R}$, let $Y \subset X$ be a linear subspace, let $\phi : Y \to \mathbb{R}$ be a linear functional, and let $c \geq 0$ such that $|\phi(x)| \leq c\,\|x\|$ for all $x \in Y$. Then there exists a bounded linear functional $\Phi : X \to \mathbb{R}$ such that*

$$\Phi|_Y = \phi, \qquad |\Phi(x)| \leq c\,\|x\| \quad \textit{for all } x \in X.$$

Proof. By Theorem 2.3.2 with $p(x) := c\,\|x\|$, there exists a linear functional $\Phi : X \to \mathbb{R}$ such that $\Phi|_Y = \phi$ and $\Phi(x) \leq c\,\|x\|$ for all $x \in X$. Since $\Phi(-x) = -\Phi(x)$ it follows that $|\Phi(x)| \leq c\,\|x\|$ for all $x \in X$ and this proves Corollary 2.3.4. □

Corollary 2.3.5 (Complex Case). *Let X be a normed vector space over $\mathbb{C}$, let $Y \subset X$ be a linear subspace, let $\psi : Y \to \mathbb{C}$ be a complex linear functional, and let $c \geq 0$ such that $|\psi(x)| \leq c\,\|x\|$ for all $x \in Y$. Then there exists a bounded complex linear functional $\Psi : X \to \mathbb{C}$ such that*

$$\Psi|_Y = \psi, \qquad |\Psi(x)| \leq c\,\|x\| \quad \textit{for all } x \in X.$$

Proof. By Corollary 2.3.4 there exists a real linear functional $\Phi : X \to \mathbb{R}$ such that

$$\Phi|_X = \mathrm{Re}\,\psi$$

and $|\Phi(x)| \leq c\,\|x\|$ for all $x \in X$. Define $\Psi : X \to \mathbb{C}$ by

$$\Psi(x) := \Phi(x) - \mathbf{i}\Phi(\mathbf{i}x) \qquad \text{for } x \in X.$$

Then $\Psi : X \to \mathbb{C}$ is complex linear and, for all $x \in Y$, we have

$$\begin{aligned} \Psi(x) &= \Phi(x) - \mathbf{i}\Phi(\mathbf{i}x) \\ &= \mathrm{Re}(\psi(x)) - \mathbf{i}\mathrm{Re}(\psi(\mathbf{i}x)) \\ &= \mathrm{Re}(\psi(x)) - \mathbf{i}\mathrm{Re}(\mathbf{i}\psi(x)) \\ &= \mathrm{Re}(\psi(x)) + \mathbf{i}\mathrm{Im}(\psi(x)) \\ &= \psi(x). \end{aligned}$$

To prove the estimate, fix a vector $x \in X$ such that $\Psi(x) \neq 0$ and choose a real number $\theta \in \mathbb{R}$ such that

$$e^{\mathbf{i}\theta} = |\Psi(x)|^{-1}\,\Psi(x).$$

Then

$$|\Psi(x)| = e^{-\mathbf{i}\theta}\Psi(x) = \Psi(e^{-\mathbf{i}\theta}x) = \Phi(e^{-\mathbf{i}\theta}x) \leq c\|e^{-\mathbf{i}\theta}x\| = c\|x\|.$$

Here the third equality follows from the fact that $\Psi(e^{-\mathbf{i}\theta}x)$ is real. This proves Corollary 2.3.5. □

2.3.2. Positive Linear Functionals. The Hahn–Banach Theorem has several important applications. The first is an extension theorem for positive linear functionals on ordered vector spaces. Recall that a partial order is a transitive, anti-symmetric, reflexive relation.

Definition 2.3.6 (Ordered Vector Space). An **ordered vector space** is a pair $(X, \preccurlyeq)$, where X is a real vector space and $\preccurlyeq$ is a partial order on X that satisfies the following two axioms for all $x, y, z \in X$ and all $\lambda \in \mathbb{R}$.

(O1) If $0 \preccurlyeq x$ and $0 \leq \lambda$, then $0 \preccurlyeq \lambda x$.

(O2) If $x \preccurlyeq y$, then $x + z \preccurlyeq y + z$.

In this situation the set $P := \{x \in X \,|\, 0 \preccurlyeq x\}$ is called the **positive cone**. A linear functional $\Phi : X \to \mathbb{R}$ is called **positive** if $\Phi(x) \geq 0$ for all $x \in P$.

Theorem 2.3.7 (Hahn–Banach for Positive Linear Functionals). *Let $(X, \preccurlyeq)$ be an ordered vector space and let $P \subset X$ be the positive cone. Let $Y \subset X$ be a linear subspace satisfying the following condition.*

(O3) *For each $x \in X$ there exists a $y \in Y$ such that $x \preccurlyeq y$.*

Let $\phi : Y \to \mathbb{R}$ be a positive linear functional, i.e. $\phi(y) \geq 0$ for all $y \in Y \cap P$. Then there is a positive linear functional $\Phi : X \to \mathbb{R}$ such that $\Phi|_Y = \phi$.

Proof. The proof has three steps.

Step 1. *For every $x \in X$ the set $\{y \in Y \,|\, x \preccurlyeq y\}$ is nonempty and the restriction of ϕ to this set is bounded below.*

Fix an element $x \in X$. Then the set $\{y \in Y \,|\, x \preccurlyeq y\}$ is nonempty by (O3). It follows also from (O3) that there exists a $y_0 \in Y$ such that $-x \preccurlyeq -y_0$. Thus we have $y_0 \preccurlyeq x$ by (O2). If $y \in Y$ satisfies $x \preccurlyeq y$, then $y_0 \preccurlyeq y$ and this implies $\phi(y_0) \leq \phi(y)$, because ϕ is positive. This proves Step 1.

Step 2. *By Step 1 the formula*

$$p(x) := \inf\{\phi(y) \,|\, y \in Y,\, x \preccurlyeq y\} \qquad \text{for } x \in X \tag{2.3.5}$$

defines a function $p : X \to \mathbb{R}$. This function is a quasi-seminorm and satisfies $p(y) = \phi(y)$ for all $y \in Y$.

Let $x_1, x_2 \in X$ and $\varepsilon > 0$. For $i = 1, 2$ choose $y_i \in Y$ such that $x_i \preccurlyeq y_i$ and $\phi(y_i) < p(x_i) + \varepsilon/2$. Then $x_1 + x_2 \preccurlyeq x_1 + y_2 \preccurlyeq y_1 + y_2$ by (O2), and so

$$p(x_1 + x_2) \leq \phi(y_1 + y_2) = \phi(y_1) + \phi(y_2) < p(x_1) + p(x_2) + \varepsilon.$$

This implies $p(x_1 + x_2) \leq p(x_1) + p(x_2)$ for all $x_1, x_2 \in X$.

Now let $x \in X$ and $\lambda > 0$. Then $\{y \in Y \mid \lambda x \preccurlyeq y\} = \{\lambda y \mid y \in Y,\, x \preccurlyeq y\}$ by (O1) and hence

$$p(\lambda x) = \inf_{\substack{y \in Y \\ \lambda x \preccurlyeq y}} \phi(y) = \inf_{\substack{y \in Y \\ x \preccurlyeq y}} \phi(\lambda y) = \inf_{\substack{y \in Y \\ x \preccurlyeq y}} \lambda \phi(y) = \lambda p(x).$$

Moreover $p(0) = 0$ by definition, and hence p is a quasi-seminorm. The formula $p(y) = \phi(y)$ for $y \in Y$ follows directly from the definition of p in (2.3.5) and this proves Step 2.

Step 3. *We prove Theorem* 2.3.7.

By Step 2 and the Hahn–Banach Theorem 2.3.2, there exists a linear functional $\Phi : X \to \mathbb{R}$ such that $\Phi|_Y = \phi$ and $\Phi(x) \le p(x)$ for all $x \in X$. If $x \in P$, then $-x \preccurlyeq 0 \in Y$, hence $\Phi(-x) \le p(-x) \le \phi(0) = 0$, and so $\Phi(x) \ge 0$. This proves Theorem 2.3.7. □

Exercise 2.3.8. Give a direct proof of Theorem 2.3.7 based on the Lemma of Zorn. **Hint:** *If $(X, \preccurlyeq)$ is an ordered vector space, $Y \subset X$ is a linear subspace satisfying (O3), $\phi : Y \to \mathbb{R}$ is a positive linear functional, and $x_0 \in X \setminus Y$, then there is a positive linear functional $\psi : Y \oplus \mathbb{R}x_0 \to \mathbb{R}$ such that $\psi|_Y = \phi$.* To see this, find a real number $a \in \mathbb{R}$ that satisfies the conditions

$$x_0 \preccurlyeq y \qquad \Longrightarrow \qquad a \le \phi(y)$$

and

$$y \preccurlyeq x_0 \qquad \Longrightarrow \qquad \phi(y) \le a$$

for all $y \in Y$.

Exercise 2.3.9. This exercise shows that the assumption (O3) cannot be removed in Theorem 2.3.7. The space $X := BC(\mathbb{R})$ of bounded continuous real valued functions on $\mathbb{R}$ is an ordered vector space with

$$f \preccurlyeq g \qquad \stackrel{\text{def}}{\Longleftrightarrow} \qquad f(t) \le g(t) \quad \text{for all } t \in \mathbb{R}.$$

The subspace $Y := C_c(\mathbb{R})$ of compactly supported continuous functions does not satisfy (O3) and the positive linear functional

$$C_c(\mathbb{R}) \to \mathbb{R} : f \mapsto \int_{-\infty}^{\infty} f(t)\, dt$$

does not extend to a positive linear functional on $BC(\mathbb{R})$. **Hint:** Every positive linear functional on $BC(\mathbb{R})$ is bounded with respect to the sup-norm.

2.3.3. Separation of Convex Sets. The second application of the Hahn–Banach Theorem concerns a pair of disjoint convex sets in a normed vector space. They can be separated by a hyperplane whenever one of them has nonempty interior (see Figure 2.3.1). The result and its proof carry over to general topological vector spaces (see Theorem 3.1.11 below).

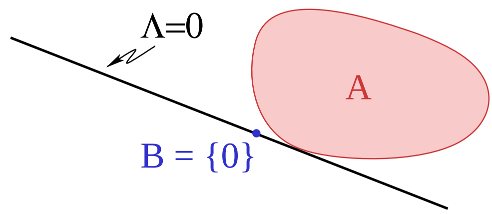

Figure 2.3.1. Two convex sets, separated by a hyperplane.

Theorem 2.3.10 (Separation of Convex Sets). *Let X be a real normed vector space and let $A, B \subset X$ be nonempty disjoint convex sets such that* $\mathrm{int}(A) \neq \emptyset$. *Then there exist a nonzero bounded linear functional $\Lambda : X \to \mathbb{R}$ and a constant $c \in \mathbb{R}$ such that $\Lambda(x) \geq c$ for all $x \in A$ and $\Lambda(x) \leq c$ for all $x \in B$. Moreover, every such bounded linear functional satisfies $\Lambda(x) > c$ for all $x \in \mathrm{int}(A)$.*

Proof. See page 71. □

Exercise 2.3.11. This exercise shows that the hypothesis that one of the convex sets has nonempty interior cannot be removed in Theorem 2.3.10. Consider the Hilbert space $H = \ell^2$ and define

$$A := \left\{ x \in \ell^2 \,\middle|\, \begin{array}{l} \exists n \in \mathbb{N}\, \forall i \in \mathbb{N} \\ i < n \implies x_i > 0 \\ i \geq n \implies x_i = 0 \end{array} \right\}, \quad B := \left\{ x \in \ell^2 \,\middle|\, \begin{array}{l} \exists n \in \mathbb{N}\, \forall i \in \mathbb{N} \\ i < n \implies x_i = 0 \\ i \geq n \implies x_i > 0 \end{array} \right\}.$$

Show that A, B are nonempty disjoint convex subsets of ℓ^2 with empty interior whose closures agree. If $\Lambda : \ell^2 \to \mathbb{R}$ is a bounded linear functional and c is a real number such that $\Lambda(x) \geq c$ for all $x \in A$ and $\Lambda(x) \leq c$ for all $x \in B$, show that $\Lambda = 0$ and $c = 0$.

Exercise 2.3.12. Define $A := \left\{ x \in \ell^2 \,|\, x_i = 0 \text{ for } i > 1 \right\}$ and

$$B := \left\{ x = (x_i)_{i=1}^\infty \in \mathbb{R}^{\mathbb{N}} \,\middle|\, |ix_i - i^{1/3}| \leq x_1 \text{ for all } i > 1 \right\} \subset \ell^2.$$

Show that A, B are nonempty disjoint closed convex subsets of ℓ^2 and $A - B$ is dense in ℓ^2. Deduce that A, B cannot be separated by an affine hyperplane.

Lemma 2.3.13. *Let X be a normed vector space and let $A \subset X$ be a convex set. Then* $\mathrm{int}(A)$ *and* $\overline{A}$ *are convex sets. Moreover, if* $\mathrm{int}(A) \neq \emptyset$*, then* $A \subset \overline{\mathrm{int}(A)}$.

Proof. The proof of convexity of $\mathrm{int}(A)$ and $\overline{A}$ is left as an exercise (see also Lemma 3.1.10). Let $x_0 \in \mathrm{int}(A)$ and choose $\delta > 0$ such that $B_\delta(x_0) \subset A$. If $x \in A$, then the set $U_x := \{tx + (1-t)y \,|\, y \in B_\delta(x_0),\, 0 < t < 1\} \subset A$ is open and hence $x \in \overline{U}_x \subset \overline{\mathrm{int}(A)}$. □

Lemma 2.3.14. *Let X be a normed vector space, let $A \subset X$ be a convex set with nonempty interior, let $\Lambda : X \to \mathbb{R}$ be a nonzero bounded linear functional, and let $c \in \mathbb{R}$ such that $\Lambda(x) \geq c$ for all $x \in$* $\mathrm{int}(A)$*. Then $\Lambda(x) \geq c$ for all $x \in A$ and $\Lambda(x) > c$ for all $x \in$* $\mathrm{int}(A)$.

Proof. Since A is convex and has nonempty interior, we have $A \subset \overline{\mathrm{int}(A)}$ by Lemma 2.3.13, and so $\Lambda(x) \geq c$ for all $x \in A$ because Λ is continuous. Now let $x \in \mathrm{int}(A)$, choose $x_0 \in X$ such that $\Lambda(x_0) = 1$, and choose $t > 0$ such that $x - tx_0 \in A$. Then $\Lambda(x) = t + \Lambda(x - tx_0) \geq t + c > c$. □

Proof of Theorem 2.3.10. The proof has three steps.

Step 1. *Let X be a real normed vector space, let $U \subset X$ be a nonempty open convex set such that $0 \notin U$, and define $P := \{tx \,|\, x \in U, t \in \mathbb{R}, t \geq 0\}$. Then P is a convex subset of X and satisfies the following.*

(P1) *If $x \in P$ and $\lambda \geq 0$, then $\lambda x \in P$.*

(P2) *If $x, y \in P$, then $x + y \in P$.*

(P3) *If $x \in P$ and $-x \in P$, then $x = 0$.*

If $x, y \in P \setminus \{0\}$, choose $x_0, x_1 \in U$ and $t_0, t_1 > 0$ such that $x = t_0 x_0$ and $y = t_1 x_1$; then $z := \frac{t_0}{t_0+t_1} x_0 + \frac{t_1}{t_0+t_1} x_1 \in U$ and hence $x + y = (t_0 + t_1) z \in P$. This proves (P2). That P satisfies (P1) is obvious and that it satisfies (P3) follows from the fact that $0 \notin U$. By (P1) and (P2) the set P is convex.

Step 2. *Let X and U be as in Step* 1*. Then there exists a bounded linear functional $\Lambda : X \to \mathbb{R}$ such that $\Lambda(x) > 0$ for all $x \in U$.*

Let P be as in Step 1. Then it follows from (P1), (P2), (P3) that the relation

$$x \preccurlyeq y \overset{\text{def}}{\iff} y - x \in P$$

defines a partial order $\preccurlyeq$ on X that satisfies (O1) and (O2).

Let $x_0 \in U$. Then the linear subspace $Y := \mathbb{R}x_0$ satisfies (O3). Namely, if $x \in X$, then $x_0 - tx \in U \subset P$ for $t > 0$ sufficiently small and so $x \preccurlyeq t^{-1}x_0$. Moreover, the linear functional $Y \to \mathbb{R} : tx_0 \mapsto t$ is positive by (P3). Hence, by Theorem 2.3.7, there is a linear functional $\Lambda : X \to \mathbb{R}$ such that $\Lambda(tx_0) = t$

for all $t \in \mathbb{R}$ and $\Lambda(x) \geq 0$ for all $x \in P$. We prove that this functional is bounded. Choose $\delta > 0$ such that $\overline{B}_\delta(x_0) \subset P$, and let $x \in X$ with $\|x\| \leq 1$. Then $x_0 - \delta x \in P$, hence $\Lambda(x_0 - \delta x) \geq 0$, and so $\Lambda(x) \leq \delta^{-1}\Lambda(x_0) = \delta^{-1}$. Thus $|\Lambda(x)| \leq \delta^{-1}\|x\|$ for all $x \in X$. Since $U \subset P$, we have $\Lambda(x) \geq 0$ for all $x \in U$, and so $\Lambda(x) > 0$ for all $x \in U$ by Lemma 2.3.14.

Step 3. *We prove Theorem* 2.3.10.

Let X, A, B be as in Theorem 2.3.10. Then $U := \mathrm{int}(A) - B$ is a nonempty open convex set and $0 \notin U$. Hence, by Step 2, there exists a bounded linear functional $\Lambda : X \to \mathbb{R}$ such that $\Lambda(x) > 0$ for all $x \in U$. Thus $\Lambda(x) > \Lambda(y)$ for all $x \in \mathrm{int}(A)$ and all $y \in B$. This implies $\Lambda(x) \geq c := \sup_{y \in B} \Lambda(y)$ for all $x \in \mathrm{int}(A)$. Hence $\Lambda(x) \geq c$ for all $x \in A$ and $\Lambda(x) > c$ for all $x \in \mathrm{int}(A)$ by Lemma 2.3.14. This proves Theorem 2.3.10. □

Definition 2.3.15 (Hyperplane). Let X be a real normed vector space. A **hyperplane** in X is a closed linear subspace of codimension one. An **affine hyperplane** is a translate of a hyperplane. An **open half-space** is a set of the form $\{x \in X \mid \Lambda(x) > c\}$ where $\Lambda : X \to \mathbb{R}$ is a nonzero bounded linear functional and $c \in \mathbb{R}$.

Exercise 2.3.16. Show that $H \subset X$ is an affine hyperplane if and only if there exist a nonzero bounded linear functional $\Lambda : X \to \mathbb{R}$ and a real number $c \in \mathbb{R}$ such that $H = \Lambda^{-1}(c)$.

Let X, A, B, Λ, c be as in Theorem 2.3.10. Then $H := \Lambda^{-1}(c)$ is an affine hyperplane that separates the convex sets A and B. It divides X into two connected components such that the interior of A is contained in one of them and B is contained in the closure of the other.

Corollary 2.3.17. *Let X be a real Banach space and let $A \subset X$ be an open convex set such that $0 \notin A$. Let $Y \subset X$ be a linear subspace such that $Y \cap A = \emptyset$. Then there is a hyperplane $H \subset X$ such that*

$$Y \subset H, \qquad H \cap A = \emptyset.$$

Proof. Assume without loss of generality that Y is closed, consider the quotient $X' := X/Y$, and denote by $\pi : X \to X'$ the obvious projection. Then π is open by Theorem 2.2.1, so $A' := \pi(A) \subset X'$ is an open convex set that does not contain the origin. Hence Theorem 2.3.10 asserts that there is a bounded linear functional $\Lambda' : X' \to \mathbb{R}$ such that $\Lambda'(x') > 0$ for all $x' \in A'$. Hence $\Lambda := \Lambda' \circ \pi : X \to \mathbb{R}$ is a bounded linear functional such that $Y \subset \ker(\Lambda)$ and $\Lambda(x) > 0$ for all $x \in A$. So $H := \ker(\Lambda)$ is the required hyperplane. □

Corollary 2.3.18. *Let X be a real normed vector space and let $A \subset X$ be a nonempty open convex set. Then A is the intersection of all open half-spaces containing A.*

Proof. Let $y \in X \setminus A$. Then, by Theorem 2.3.10 with $B = \{y\}$, there is a $\Lambda \in X^*$ and a $c \in \mathbb{R}$ such that $\Lambda(x) > c$ for all $x \in A$ and $\Lambda(y) \leq c$. Hence there is an open half-space containing A but not y. □

Corollary 2.3.19. *Let X be a real normed vector space and let $A, B \subset X$ be nonempty disjoint convex sets such that A is closed and B is compact. Then there exists a bounded linear functional $\Lambda : X \to \mathbb{R}$ such that*

$$\inf_{x \in A} \Lambda(x) > \sup_{y \in B} \Lambda(y).$$

Proof. We prove first that

$$\delta := \inf_{x \in A,\, y \in B} \|x - y\| > 0.$$

Choose sequences $x_n \in A$ and $y_n \in B$ such that

$$\lim_{n \to \infty} \|x_n - y_n\| = \delta.$$

Since B is compact, we may assume, by passing to a subsequence if necessary, that the sequence $(y_n)_{n \in \mathbb{N}}$ converges to an element $y \in B$. If $\delta = 0$ it would follow that the sequence $(x_n - y_n)_{n \in \mathbb{N}}$ converges to zero, so the sequence $x_n = y_n + (x_n - y_n)$ converges to y, and so $y \in A$, because A is closed, contradicting the fact that $A \cap B = \emptyset$. Thus $\delta > 0$ as claimed. Hence

$$U := \bigcup_{x \in A} B_\delta(x)$$

is an open convex set that contains A and is disjoint from B. Thus, by Theorem 2.3.10, there is a bounded linear functional $\Lambda : X \to \mathbb{R}$ such that

$$\Lambda(x) > c := \sup_{y \in B} \Lambda(y) \qquad \text{for all } x \in U.$$

Choose $\xi \in X$ such that $\|\xi\| < \delta$ and $\varepsilon := \Lambda(\xi) > 0$. Then every $x \in A$ satisfies $x - \xi \in U$ and hence

$$\Lambda(x) - \varepsilon = \Lambda(x - \xi) > c.$$

This proves Corollary 2.3.19. □

Exercise 2.3.20. Let X be a real normed vector space and let $A \subset X$ be a nonempty convex set. Prove that $\overline{A}$ is the intersection of all closed half-spaces of X containing A.

2.3.4. The Closure of a Linear Subspace. The third application of the Hahn–Banach Theorem is a characterization of the closure of a linear subspace of a real normed vector space X. Recall that the dual space of X is the space

$$X^* := \mathcal{L}(X, \mathbb{R})$$

of real valued bounded linear functionals on X. At this point it is convenient to introduce an alternative notation for the elements of the dual space. Denote a bounded linear functional on X by $x^* : X \to \mathbb{R}$ and denote the value of this linear functional on an element $x \in X$ by

$$\langle x^*, x \rangle := x^*(x).$$

This notation is reminiscent of the inner product on a Hilbert space and there are in fact many parallels between the pairing

$$X^* \times X \to \mathbb{R} : (x^*, x) \mapsto \langle x^*, x \rangle \tag{2.3.6}$$

and inner products on Hilbert spaces. Recall that X^* is a Banach space with respect to the norm

$$\|x^*\| := \sup_{x \in X \setminus \{0\}} \frac{|\langle x^*, x \rangle|}{\|x\|} \qquad \text{for } x^* \in X^* \tag{2.3.7}$$

(see Theorem 1.3.1). It follows directly from (2.3.7) that

$$|\langle x^*, x \rangle| \le \|x^*\| \, \|x\| \tag{2.3.8}$$

for all $x^* \in X^*$ and all $x \in X$, in analogy to the Cauchy–Schwarz inequality. Hence the pairing (2.3.6) is continuous by Corollary 2.1.7.

Definition 2.3.21 (Annihilator). Let X be a real normed vector space. For any subset $S \subset X$ define the **annihilator** of S as the space of bounded linear functionals on X that vanish on S and denote it by

$$S^\perp := \left\{ x^* \in X^* \,\middle|\, \langle x^*, x \rangle = 0 \text{ for all } x \in S \right\}. \tag{2.3.9}$$

Since the pairing (2.3.6) is continuous, the annihilator $S^\perp$ is a closed linear subspace of X^* for every subset $S \subset X$. As before, the closure of a subset $Y \subset X$ is denoted by $\overline{Y}$.

Theorem 2.3.22. *Let X be a real normed vector space, let $Y \subset X$ be a linear subspace, and let $x_0 \in X \setminus \overline{Y}$. Then*

$$\delta := d(x_0, Y) := \inf_{y \in Y} \|x_0 - y\| > 0 \tag{2.3.10}$$

and there exists a bounded linear functional $x^* \in Y^\perp$ such that

$$\|x^*\| = 1, \qquad \langle x^*, x_0 \rangle = \delta.$$

Proof. We prove first that the number δ in (2.3.10) is positive. Suppose by contradiction that $\delta = 0$. Then, by the axiom of countable choice, there exists a sequence $(y_n)_{n\in\mathbb{N}}$ in Y such that $\|x_0 - y_n\| < 1/n$ for all $n \in \mathbb{N}$. This implies that y_n converges to x_0 and hence $x_0 \in \overline{Y}$, in contradiction to our assumption. This shows that $\delta > 0$ as claimed.

Now define the subspace $Z \subset X$ by

$$Z := Y \oplus \mathbb{R}x_0 = \left\{ y + tx_0 \,\middle|\, y \in Y,\, t \in \mathbb{R} \right\}$$

and define the linear functional $\psi : Z \to \mathbb{R}$ by

$$\psi(y + tx_0) := \delta t \qquad \text{for } y \in Y \text{ and } t \in \mathbb{R}.$$

This functional is well defined because $x_0 \notin Y$. It satisfies $\psi(y) = 0$ for all $y \in Y$ and $\psi(x_0) = \delta$. Moreover, if $y \in Y$ and $t \in \mathbb{R} \setminus \{0\}$, then

$$\frac{|\psi(y + tx_0)|}{\|y + tx_0\|} = \frac{|t|\delta}{\|y + tx_0\|} = \frac{\delta}{\|t^{-1}y + x_0\|} \le 1.$$

Here the last inequality follows from the definition of δ. With this understood, it follows from Corollary 2.3.4 that there exists a bounded linear functional $x^* \in X^*$ such that

$$\|x^*\| \le 1$$

and

$$\langle x^*, x \rangle = \psi(x) \qquad \text{for all } x \in Z.$$

The norm of x^* is actually equal to one because

$$\|x^*\| \ge \sup_{y\in Y} \frac{|\psi(x_0 + y)|}{\|x_0 + y\|} = \sup_{y\in Y} \frac{|\delta|}{\|x_0 + y\|} = 1$$

by definition of δ. Moreover,

$$\langle x^*, x_0 \rangle = \psi(x_0) = \delta$$

and

$$\langle x^*, y \rangle = \psi(y) = 0 \quad \text{for all } y \in Y.$$

This proves Theorem 2.3.22. □

Corollary 2.3.23. *Let X be a real normed vector space and let $x_0 \in X$ be a nonzero vector. Then there exists a bounded linear functional $x^* \in X^*$ such that*

$$\|x^*\| = 1, \qquad \langle x^*, x_0 \rangle = \|x_0\|.$$

Proof. This follows directly from Theorem 2.3.22 with $Y := \{0\}$. □

The next corollary characterizes the closure of a linear subspace and gives rise to a criterion for a linear subspace to be dense.

Corollary 2.3.24 (Closure of a Subspace). *Let X be a real normed vector space, let $Y \subset X$ be a linear subspace, and let $x \in X$. Then*

$$x \in \overline{Y} \qquad \Longleftrightarrow \qquad \langle x^*, x\rangle = 0 \text{ for all } x^* \in Y^\perp.$$

Proof. If $x \in \overline{Y}$ and $x^* \in Y^\perp$, then there is a sequence $(y_n)_{n\in\mathbb{N}}$ in Y that converges to x and so $\langle x^*, x\rangle = \lim_{n\to\infty}\langle x^*, y_n\rangle = 0$. If $x \notin \overline{Y}$, then there is an element $x^* \in Y^\perp$ such that $\langle x^*, x\rangle > 0$ by Theorem 2.3.22. □

Corollary 2.3.25 (Dense Subspaces). *Let X be a real normed vector space and let $Y \subset X$ be a linear subspace. Then Y is dense in X if and only if $Y^\perp = \{0\}$.*

Proof. By Corollary 2.3.24 we have $\overline{Y} = X$ if and only if $\langle x^*, x\rangle = 0$ for all $x^* \in Y^\perp$ and all $x \in X$, and this is equivalent to $Y^\perp = \{0\}$. □

The next corollary asserts that the dual space of a quotient is a subspace of the dual space and vice versa.

Corollary 2.3.26 (Dual Spaces of Subspaces and Quotients). *Let X be a real normed vector space and let $Y \subset X$ be a linear subspace. Then the following hold.*

(i) *The linear map*

$$X^*/Y^\perp \to Y^* : [x^*] \mapsto x^*|_Y \tag{2.3.11}$$

is an isometric isomorphism.

(ii) *Assume Y is closed and let $\pi : X \to X/Y$ be the canonical projection, given by $\pi(x) := x + Y$ for $x \in X$. Then the linear map*

$$(X/Y)^* \to Y^\perp : \Lambda \mapsto \Lambda \circ \pi \tag{2.3.12}$$

is an isometric isomorphism.

Proof. We prove part (i). The linear map

$$X^* \to Y^* : x^* \mapsto x^*|_Y$$

vanishes on $Y^\perp$ and hence descends to the quotient $X^*/Y^\perp$. The resulting map (2.3.11) is injective by definition. Now fix any bounded linear functional $y^* \in Y^*$. Then Corollary 2.3.4 asserts that there is a bounded linear functional $x^* \in X^*$ such that

$$x^*|_Y = y^*, \qquad \|x^*\| = \|y^*\|.$$

Moreover, if $\xi^* \in X^*$ satisfies $\xi^*|_Y = y^*$, then $\|\xi^*\| \geq \|y^*\| = \|x^*\|$. Hence x^* minimizes the norm among all bounded linear functionals on X that restrict to y^* on Y. Thus $\|x^* + Y^\perp\|_{X^*/Y^\perp} = \|x^*\| = \|y^*\|$, and this shows that the map (2.3.11) is an isometric isomorphism.

We prove part (ii). Fix a bounded linear functional

$$\Lambda : X/Y \to \mathbb{R}$$

and define

$$x^* := \Lambda \circ \pi : X \to \mathbb{R}.$$

Then x^* is a bounded linear functional on X and $x^*|_Y = 0$. Thus

$$x^* \in Y^\perp.$$

Conversely, fix an element $x^* \in Y^\perp$. Then x^* vanishes on Y and hence descends to a unique linear map $\Lambda : X/Y \to \mathbb{R}$ such that

$$\Lambda \circ \pi = x^*.$$

To prove that Λ is bounded, observe that

$$\Lambda(x + Y) = \langle x^*, x\rangle = \langle x^*, x + y\rangle \le \|x^*\| \, \|x + y\|$$

for all $x \in X$ and all $y \in Y$, hence

$$|\Lambda(x + Y)| \le \|x^*\| \inf_{y\in Y} \|x + y\| = \|x^*\| \, \|x + Y\|_{X/Y}$$

for all $x \in X$, and hence

$$\|\Lambda\| \le \|x^*\| .$$

Conversely

$$\begin{aligned}\langle x^*, x\rangle &= \Lambda(x + Y) \\ &\le \|\Lambda\| \, \|x + Y\|_{X/Y} \\ &\le \|\Lambda\| \, \|x\|\end{aligned}$$

for all $x \in X$ and so

$$\|x^*\| \le \|\Lambda\| .$$

Hence the linear map (2.3.12) is an isometric isomorphism. This proves Corollary 2.3.26. □

Corollary 2.3.27. *Let X be a real normed vector space and let $Y \subset X$ be a closed linear subspace. Then*

$$\inf_{\xi^*\in Y^\perp} \|x^* + \xi^*\| = \sup_{y\in Y\setminus\{0\}} \frac{\langle x^*, y\rangle}{\|y\|} \qquad \text{for all } x^* \in X^* \tag{2.3.13}$$

and

$$\|x^*\| = \sup_{x\in X\setminus Y} \frac{\langle x^*, x\rangle}{\inf_{y\in Y} \|x + y\|} \qquad \text{for all } x^* \in Y^\perp. \tag{2.3.14}$$

Proof. This follows directly from Corollary 2.3.26. □

2.3.5. Complemented Subspaces. A familiar observation in linear algebra is that, for every subspace $Y \subset X$ of a finite-dimensional vector space X, there exists another subspace $Z \subset X$ such that $X = Y \oplus Z$. This continues to hold for infinite-dimensional vector spaces. However, it does not hold, in general, for closed subspaces of normed vector spaces. Here is the relevant definition.

Definition 2.3.28 (Complemented Subspace). Let X be a normed vector space. A closed subspace $Y \subset X$ is called **complemented** if there exists a closed subspace $Z \subset X$ such that $Y \cap Z = \{0\}$ and $X = Y \oplus Z$. A bounded linear operator $P : X \to X$ is called a **projection** if $P^2 = P$.

Exercise 2.3.29. Let X be a Banach space, let $Y \subset X$ be a closed linear subspace, and let $\pi : X \to X/Y$ be the canonical projection. (**Warning:** The term *projection* is used here with two different meanings.) Prove that the following are equivalent.

(i) Y is complemented.

(ii) There is a projection $P : X \to X$ such that $\mathrm{im}(P) = Y$.

(iii) There is a bounded linear operator $T : X/Y \to X$ such that $\pi \circ T = \mathrm{id}$. (The operator T, if it exists, is called a **right inverse** of π.)

Hint: For (i) $\Longrightarrow$ (ii) use Corollary 2.2.9. For (ii) $\Longrightarrow$ (i) define $Z := \ker(P)$. For (ii) $\Longrightarrow$ (iii) let $T[x] := x - Px$. For (iii) $\Longrightarrow$ (ii) let $P := \mathbb{1} - T \circ \pi$.

Lemma 2.3.30. *Let X be a normed vector space and let $Y \subset X$ be a closed linear subspace such that $\dim(Y) < \infty$ or $\dim(X/Y) < \infty$. Then Y is complemented.*

Proof. Assume $n := \dim(X/Y) < \infty$ and choose vectors $x_1, \dots, x_n \in X$ whose equivalence classes $[x_i] := x_i + Y$ form a basis of X/Y. Then the linear subspace $Z := \mathrm{span}\{x_1, \dots, x_n\}$ is closed by Corollary 1.2.7 and satisfies $X = Y \oplus Z$.

Now assume $n := \dim Y < \infty$ and choose a basis $x_1, \dots, x_n$ of Y. By the Hahn–Banach Theorem (Corollary 2.3.4) there exist bounded linear functionals $x_1^*, \dots, x_n^* \in X^*$ that satisfy $\langle x_i^*, x_j \rangle = \delta_{ij}$. Then the subspace

$$Z := \{x \in X \mid \langle x_i^*, x \rangle = 0 \text{ for } i = 1, \dots, n\}$$

is closed by Theorem 1.2.2. Moreover, $x - \sum_{i=1}^n \langle x_i^*, x \rangle x_i \in Z$ for all $x \in X$ and hence $X = Y \oplus Z$. This proves Lemma 2.3.30. □

There are examples of closed subspaces of infinite-dimensional Banach spaces that are not complemented. The simplest such example is the subspace $c_0 \subset \ell^\infty$. Phillips' Lemma asserts that it is not complemented. The proof is outlined in Exercise 2.5.1 below.

2.3.6. Orthonormal Bases.

Definition 2.3.31. Let H be an infinite-dimensional real Hilbert space. A sequence $(e_i)_{i\in\mathbb{N}}$ in H is called a **(countable) orthonormal basis** if

$$\langle e_i, e_j\rangle = \delta_{ij} := \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{if } i \neq j, \end{cases} \qquad \text{for all } i, j \in \mathbb{N}, \tag{2.3.15}$$

$$x \in H, \quad \langle e_i, x\rangle = 0 \text{ for all } i \in \mathbb{N} \qquad \Longrightarrow \qquad x = 0. \tag{2.3.16}$$

If $(e_i)_{i\in\mathbb{N}}$ is an orthonormal basis, then (2.3.15) implies that the e_i are linearly independent and (2.3.16) asserts that the set $E := \mathrm{span}(\{e_i \,|\, i \in \mathbb{N}\})$ is a dense linear subspace of H (Corollary 2.3.25).

Exercise 2.3.32. Show that an infinite-dimensional Hilbert space H admits a countable orthonormal basis if and only if it is separable. **Hint:** Assume H is separable. Choose a dense sequence, construct a linearly independent subsequence spanning a dense subspace, and use *Gram–Schmidt*.

Exercise 2.3.33. Let H be a separable Hilbert space and let $\{e_i\}_{i\in\mathbb{N}}$ be an orthonormal basis. Show that the map $\ell^2 \to H : x = (x_i)_{i\in\mathbb{N}} \mapsto \sum_{i=1}^\infty x_i e_i$ is well defined (i.e. $\xi_n := \sum_{i=1}^n x_i e_i$ is a Cauchy sequence in H for all $x \in \ell^2$) and defines a Hilbert space isometry. Deduce that

$$x = \sum_{i=1}^\infty \langle e_i, x\rangle e_i, \qquad \|x\|^2 = \sum_{i=1}^\infty \langle e_i, x\rangle^2 \qquad \text{for all } x \in H. \tag{2.3.17}$$

Example 2.3.34. The sequences $e_i := (\delta_{ij})_{j\in\mathbb{N}}$ for $i \in \mathbb{N}$ form an orthonormal basis of ℓ^2.

Example 2.3.35 (Fourier Series). The functions $e_k(t) := e^{2\pi \mathbf{i} kt}$, $k \in \mathbb{Z}$, form an orthonormal basis of the complex Hilbert space $L^2(\mathbb{R}/\mathbb{Z}, \mathbb{C})$. It is equipped with the complex valued **Hermitian inner product**

$$\langle f, g\rangle := \int_0^1 \overline{f(t)} g(t)\, dt \qquad \text{for } f, g \in L^2(\mathbb{R}/\mathbb{Z}, \mathbb{C}), \tag{2.3.18}$$

that is complex anti-linear in the first variable and complex linear in the second variable. To verify completeness, one can fix a continuous function $f : \mathbb{R}/\mathbb{Z} \to \mathbb{C}$, define $f_n := \sum_{k=-n}^n \langle e_k, f\rangle e_k$ for $n \in \mathbb{N}_0$, and prove that the sequence $n^{-1}(f_0 + f_1 + \cdots + f_{n-1})$ converges uniformly to f (Fejér's Theorem).

Example 2.3.36. The functions $s_n(t) := \sqrt{2}\sin(\pi nt)$ for $n \in \mathbb{N}$ form an orthonormal basis of the Hilbert space $L^2([0, 1])$ and so do the functions $c_0(t) := 1$ and $c_n(t) := \sqrt{2}\cos(\pi nt)$ for $n \in \mathbb{N}$. **Exercise:** Use completeness in Example 2.3.35 to verify the completeness axiom (2.3.16) for these two orthonormal bases.

2.4. Reflexive Banach Spaces

2.4.1. The Bidual Space. Let X be a real normed vector space. The **bidual space** of X is the dual space of the dual space and is denoted by

$$X^{**} := (X^*)^* = \mathcal{L}(X^*, \mathbb{R}).$$

There is a natural map $\iota = \iota_X : X \to X^{**}$ which assigns to every element $x \in X$ the linear functional $\iota(x) : X^* \to \mathbb{R}$ whose value at x^* is obtained by evaluating the bounded linear functional $x^* : X \to \mathbb{R}$ at the point $x \in X$. Thus the map $\iota : X \to X^{**}$ is defined by

$$\iota(x)(x^*) := \langle x^*, x\rangle \tag{2.4.1}$$

for $x \in X$ and $x^* \in X^*$. It is a consequence of the Hahn–Banach Theorem that the linear map $\iota : X \to X^{**}$ is an isometric embedding.

Lemma 2.4.1. *Let X be a real normed vector space. Then the linear map $\iota : X \to X^{**}$ is an isometric embedding. In particular,*

$$\|x\| = \sup_{x^* \in X^* \setminus \{0\}} \frac{|\langle x^*, x\rangle|}{\|x^*\|} \tag{2.4.2}$$

for all $x \in X$.

Proof. That the map $\iota : X \to X^{**}$ is linear follows directly from the definition. To prove that it preserves the norm, fix a nonzero vector $x_0 \in X$. Then, by Corollary 2.3.23, there exists a bounded linear functional $x_0^* \in X^*$ such that $\|x_0^*\| = 1$ and $\langle x_0^*, x_0\rangle = \|x_0\|$. Hence

$$\|x_0\| = \frac{|\langle x_0^*, x_0\rangle|}{\|x_0^*\|} \le \|\iota(x_0)\| = \sup_{x^* \in X^* \setminus \{0\}} \frac{|\langle x^*, x_0\rangle|}{\|x^*\|} \le \|x_0\|.$$

Here the last inequality follows from (2.3.8). This proves Lemma 2.4.1. □

Corollary 2.4.2. *Let X be a real normed vector space and let $Y \subset X$ be a closed linear subspace. Then, for every $x \in X$,*

$$\inf_{y \in Y} \|x + y\| = \sup_{x^* \in Y^\perp \setminus \{0\}} \frac{|\langle x^*, x\rangle|}{\|x^*\|}. \tag{2.4.3}$$

Proof. The left-hand side of equation (2.4.3) is the norm of the equivalence class $[x] = x + Y$ in the quotient space X/Y. The right-hand side is the norm of the bounded linear functional

$$\iota_{X/Y}(x + Y) : (X/Y)^* \cong Y^\perp \to \mathbb{R}$$

(see Corollary 2.3.26). Hence equation (2.4.3) follows from Lemma 2.4.1 with X replaced by X/Y. This proves Corollary 2.4.2. □

2.4.2. Reflexive Banach Spaces.

Definition 2.4.3 (**Reflexive Banach Space**). A real normed vector space X is called **reflexive** if the isometric embedding $\iota : X \to X^{**}$ in (2.4.1) is bijective. A reflexive normed vector space is necessarily complete by Theorem 1.3.1.

Theorem 2.4.4. *Let X be a Banach space. Then the following hold.*

(i) *X is reflexive if and only if X^* is reflexive.*

(ii) *If X is reflexive and $Y \subset X$ is a closed linear subspace, then the subspace Y and the quotient space X/Y are reflexive.*

Proof. We prove part (i). Assume X is reflexive and let $\Lambda : X^{**} \to \mathbb{R}$ be a bounded linear functional. Define

$$x^* := \Lambda \circ \iota : X \to \mathbb{R},$$

where $\iota = \iota_X : X \to X^{**}$ is the isometric embedding in (2.4.1). Since X is reflexive, this map ι is bijective. Fix an element $x^{**} \in X^{**}$ and define

$$x := \iota^{-1}(x^{**}) \in X.$$

Then

$$\Lambda(x^{**}) = \Lambda \circ \iota(x) = \langle x^*, x \rangle = \langle \iota(x), x^* \rangle = \langle x^{**}, x^* \rangle.$$

Here the first and last equation follow from the fact that $x^{**} = \iota(x)$, the second equation follows from the definition of $x^* = \Lambda \circ \iota$, and the third equation follows from the definition of the map ι in (2.4.1). This shows that

$$\Lambda = \iota_{X^*}(x^*),$$

where $\iota_{X^*} : X^* \to X^{***}$ is the isometric embedding in (2.4.1) with X replaced by X^*. This shows that the dual space X^* is reflexive.

Conversely, assume X^* is reflexive. The subspace $\iota(X)$ of X^{**} is complete by Lemma 2.4.1 and is therefore closed. We prove that $\iota(X)$ is a dense subspace of X^{**}. To see this, let $\Lambda : X^{**} \to \mathbb{R}$ be any bounded linear functional on X^{**} that vanishes on the image of ι, so that $\Lambda \circ \iota = 0$. Since X^* is reflexive, there exists an element $x^* \in X^*$ such that

$$\Lambda(x^{**}) = \langle x^{**}, x^* \rangle$$

for every $x^{**} \in X^{**}$. Since $\Lambda \circ \iota = 0$, this implies

$$\langle x^*, x \rangle = \langle \iota(x), x^* \rangle = \Lambda(\iota(x)) = 0$$

for all $x \in X$, hence $x^* = 0$, and hence $\Lambda = 0$. Thus the annihilator of the linear subspace $\iota(X) \subset X^{**}$ is zero, and so $\iota(X)$ is dense in X^{**} by Corollary 2.3.25. Hence $\iota(X) = X^{**}$ and this proves part (i).

We prove part (ii). Assume X is reflexive and let

$$Y \subset X$$

be a closed linear subspace. We prove first that Y is a reflexive Banach space. Define the linear operator

$$\pi : X^* \to Y^*$$

by

$$\pi(x^*) := x^*|_Y$$

for $x^* \in X^*$. Fix an element $y^{**} \in Y^{**}$ and define $x^{**} \in X^{**}$ by

$$x^{**} := y^{**} \circ \pi : X^* \to \mathbb{R}.$$

Since X is reflexive, there exists a unique element $y \in X$ such that

$$\iota_X(y) = x^{**}.$$

Every element $x^* \in Y^\perp$ satisfies $\pi(x^*) = 0$ and hence

$$\begin{aligned}
\langle x^*, y \rangle &= \langle \iota_X(y), x^* \rangle \\
&= \langle x^{**}, x^* \rangle \\
&= \langle y^{**} \circ \pi, x^* \rangle \\
&= \langle y^{**}, \pi(x^*) \rangle \\
&= 0.
\end{aligned}$$

In other words, $\langle x^*, y \rangle = 0$ for all $x^* \in Y^\perp$ and so

$$y \in \overline{Y} = Y$$

by Corollary 2.3.24. Now fix any element $y^* \in Y^*$. Then Corollary 2.3.4 asserts that there exists an element $x^* \in X^*$ such that

$$y^* = x^*|_Y = \pi(x^*)$$

and so

$$\begin{aligned}
\langle y^{**}, y^* \rangle &= \langle y^{**}, \pi(x^*) \rangle \\
&= \langle x^{**}, x^* \rangle \\
&= \langle \iota(y), x^* \rangle \\
&= \langle x^*, y \rangle \\
&= \langle y^*, y \rangle.
\end{aligned}$$

This shows that

$$\iota_Y(y) = y^{**}.$$

Since $y^{**} \in Y^{**}$ was chosen arbitrarily, this proves that the subspace Y is a reflexive Banach space.

Next we prove that the quotient

$$Z := X/Y$$

is reflexive. Let

$$\pi : X \to X/Y$$

be the canonical projection given by

$$\pi(x) := [x] = x + Y \qquad \text{for } x \in X$$

and define the linear operator $T : Z^* \to Y^\perp$ by

$$Tz^* := z^* \circ \pi : X \to \mathbb{R} \qquad \text{for } z^* \in Z^*.$$

Note that $Tz^* \in Y^\perp$ because $(Tz^*)(y) = z^*(\pi(y)) = 0$ for all $y \in Y$. Moreover, T is an isometric isomorphism by Corollary 2.3.26.

Now fix an element $z^{**} \in Z^{**}$. Then the map

$$z^{**} \circ T^{-1} : Y^\perp \to \mathbb{R}$$

is a bounded linear functional on a linear subspace of X^*. Hence, by Corollary 2.3.4, there exists a bounded linear functional $x^{**} : X^* \to \mathbb{R}$ such that

$$\langle x^{**}, x^* \rangle = \langle z^{**}, T^{-1}x^* \rangle \qquad \text{for all } x^* \in Y^\perp.$$

This condition on x^{**} can be expressed in the form

$$\langle x^{**}, z^* \circ \pi \rangle = \langle z^{**}, z^* \rangle \qquad \text{for all } z^* \in Z^*.$$

Since X is reflexive, there exists an element $x \in X$ such that

$$\iota_X(x) = x^{**}.$$

Define

$$z := [x] = \pi(x) \in Z.$$

Then, for all $z^* \in Z^*$, we have

$$\begin{aligned}
\langle z^{**}, z^* \rangle &= \langle x^{**}, z^* \circ \pi \rangle \\
&= \langle \iota(x), z^* \circ \pi \rangle \\
&= \langle z^* \circ \pi, x \rangle \\
&= \langle z^*, \pi(x) \rangle \\
&= \langle z^*, z \rangle.
\end{aligned}$$

This shows that

$$\iota_Z(z) = z^{**}.$$

Since $z^{**} \in Z^{**}$ was chosen arbitrarily, it follows that Z is reflexive. This proves Theorem 2.4.4. □

Example 2.4.5. **(i)** Every finite-dimensional normed vector space X is reflexive, because $\dim X = \dim X^* = \dim X^{**}$ (see Corollary 1.2.9).

(ii) Every Hilbert space H is reflexive by Theorem 1.4.4. **Exercise:** The composition of the isomorphisms $H \cong H^* \cong H^{**}$ is the map in (2.4.1).

(iii) Let $(M, \mathcal{A}, \mu)$ be a measure space and let $1 < p, q < \infty$ such that

$$\frac{1}{p} + \frac{1}{q} = 1.$$

Then $L^p(\mu)^* \cong L^q(\mu)$ (Example 1.3.3) and this implies that the Banach space $L^p(\mu)$ is reflexive. **Exercise:** Prove that the composition of the isomorphisms $L^p(\mu) \cong L^q(\mu)^* \cong L^p(\mu)^{**}$ is the map in (2.4.1).

(iv) Let $c_0 \subset \ell^\infty$ be the subspace of sequences $x = (x_i)_{i\in\mathbb{N}}$ of real numbers that converge to zero, equipped with the supremum norm. Then the map $\ell^1 \to c_0^* : y \mapsto \Lambda_y$, which assigns to every sequence $y = (y_i)_{i\in\mathbb{N}} \in \ell^1$ the bounded linear functional $\Lambda_y : c_0 \to \mathbb{R}$ defined by $\Lambda_y(x) := \sum_{i=1}^\infty x_i y_i$ for $x = (x_i)_{i\in\mathbb{N}} \in c_0$, is a Banach space isometry (see Example 1.3.7). This implies $c_0^{**} \cong (\ell^1)^* \cong \ell^\infty$ (see Example 1.3.6), and so c_0 is not reflexive. **Exercise:** The composition of the isometric embedding $\iota : c_0 \to c_0^{**}$ in (2.4.1) with the Banach space isometry $c_0^{**} \cong \ell^\infty$ is the canonical inclusion.

(v) The Banach space ℓ^1 is not reflexive. To see this, denote by $c \subset \ell^\infty$ the space of Cauchy sequences of real numbers and consider the bounded linear functional that assigns to each Cauchy sequence $x = (x_i)_{i\in\mathbb{N}} \in c$ its limit $\lim_{i\to\infty} x_i$. By the Hahn–Banach Theorem this functional extends to a bounded linear functional $\Lambda : \ell^\infty \to \mathbb{R}$ (see Corollary 2.3.4), which does not belong to the image of the inclusion $\iota : \ell^1 \to (\ell^1)^{**} \cong (\ell^\infty)^*$.

(vi) Let (M, d) be a compact metric space and let $X = C(M)$ be the Banach space of continuous real valued functions on M with the supremum norm (see part (v) of Example 1.1.3). Suppose M is an infinite set. Then $C(M)$ is not reflexive. To see this, let $A = \{a_1, a_2, \dots\} \subset M$ be a countably infinite subset such that $(a_i)_{i\in\mathbb{N}}$ is a Cauchy sequence and $a_i \neq a_j$ for $i \neq j$. Then $C_A(M) := \{f \in C(M) \,|\, f|_A = 0\}$ is a closed linear subspace of $C(M)$ and the quotient $C(M)/C_A(M)$ is isometrically isomorphic to the space c of Cauchy sequences of real numbers via $C(M)/C_A(M) \to c : [f] \mapsto (f(a_i))_{i=1}^\infty$. By Theorem 2.4.4 the Banach space c is not reflexive, because the closed subspace $c_0 \subset c$ is not reflexive by (iv) above. Hence $C(M)/C_A(M)$ is not reflexive, and so $C(M)$ is not reflexive by Theorem 2.4.4.

(vii) The dual space of the Banach space $C(M)$ in (vi) is isomorphic to the Banach space $\mathcal{M}(M)$ of signed Borel measures on M (see Example 1.3.8). Since $C(M)$ is not reflexive, neither is the space $\mathcal{M}(M)$ by Theorem 2.4.4.

2.4.3. Separable Banach Spaces. Recall that a normed vector space is called **separable** if it contains a countable dense subset (see Definition 1.1.6). Thus a Banach space X is separable if and only if there exists a sequence $e_1, e_2, e_3, \ldots$ in X such that the linear subspace of all (finite) linear combinations of the e_i is dense in X. If such a sequence exists, the required countable dense subset can be constructed as the set of all rational linear combinations of the e_i.

Theorem 2.4.6. *Let X be a normed vector space. The following hold.*

(i) *If X^* is separable, then X is separable.*

(ii) *If X is reflexive and separable, then X^* is separable.*

Proof. We prove part (i). Thus assume X^* is separable and choose a dense sequence $(x_i^*)_{i\in\mathbb{N}}$ in X^*. Choose a sequence $x_i \in X$ such that

$$\|x_i\| = 1, \qquad \langle x_i^*, x_i\rangle \geq \frac{1}{2}\|x_i^*\| \qquad \text{for all } i \in \mathbb{N}.$$

Let $Y \subset X$ be the linear subspace of all finite linear combinations of the x_i. We prove that Y is dense in X. To see this, fix any element $x^* \in Y^\perp$. Then there is a sequence $i_k \in \mathbb{N}$ such that $\lim_{k\to\infty}\|x^* - x_{i_k}^*\| = 0$. This implies

$$\begin{aligned} \|x_{i_k}^*\| &\leq 2|\langle x_{i_k}^*, x_{i_k}\rangle| = 2|\langle x_{i_k}^* - x^*, x_{i_k}\rangle| \\ &\leq 2\|x_{i_k}^* - x^*\|\|x_{i_k}\| = 2\|x_{i_k}^* - x^*\|. \end{aligned}$$

The last term on the right converges to zero as k tends to infinity, and hence $x^* = \lim_{k\to\infty} x_{i_k}^* = 0$. This shows that $Y^\perp = \{0\}$. Hence Y is dense in X by Corollary 2.3.25 and this proves part (i). If X is reflexive and separable, then X^{**} is separable, and so X^* is separable by (i). This proves part (ii) and Theorem 2.4.6. □

Example 2.4.7. (i) Finite-dimensional Banach spaces are separable.

(ii) The space ℓ^p is separable for $1 \leq p < \infty$, and $(\ell^1)^* \cong \ell^\infty$ is not separable. The subspace $c_0 \subset \ell^\infty$ of all sequences that converge to zero is separable.

(iii) Let M be a second countable locally compact Hausdorff space, denote by $\mathcal{B} \subset 2^M$ its Borel σ-algebra, and let $\mu : \mathcal{B} \to [0, \infty]$ be a locally finite Borel measure. Then the space $L^p(\mu)$ is separable for $1 \leq p < \infty$. (See for example [**75**, Thm. 4.13].)

(iv) Let (M, d) be a compact metric space. Then the Banach space $C(M)$ of continuous functions with the supremum norm is separable. Its dual space $\mathcal{M}(M)$ of signed Borel measures is in general not separable.

2.4.4. The James Space. In 1950 Robert C. James [**37**, **38**] discovered a remarkable example of a nonreflexive Banach space J that is isometrically isomorphic to its bidual space J^{**}. In this example the image of the canonical isometric embedding

$$\iota : J \to J^{**}$$

in (2.4.1) is a closed subspace of codimension one. Our exposition follows Megginson [**59**].

Recall that $c_0 \subset \ell^\infty$ is the Banach space of all sequences $(x_i)_{i\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$ that converge to zero, equipped with the supremum norm

$$\|x\|_\infty := \sup_{i\in\mathbb{N}} |x_i| \qquad \text{for } x = (x_i)_{i\in\mathbb{N}} \in c_0.$$

By Example 1.3.7 the dual space of c_0 is isomorphic to the space ℓ^1 of absolutely summable sequences of real numbers with the norm

$$\|x\|_1 := \sum_{i=1}^\infty |x_i| \qquad \text{for } x = (x_i)_{i\in\mathbb{N}} \in \ell^1.$$

Recall also that ℓ^2 is the Hilbert space of all square summable sequences of real numbers with the norm

$$\|x\|_2 := \left(\sum_{i=1}^\infty |x_i|^2\right)^{1/2} \qquad \text{for } x = (x_i)_{i\in\mathbb{N}} \in \ell^2.$$

Definition 2.4.8 (The James Space). Let $\mathcal{P} \subset 2^{\mathbb{N}}$ be the collection of all nonempty finite subsets of $\mathbb{N}$ and write the elements of $\mathcal{P}$ in the form $\mathbf{p} = (p_1, p_2, \ldots, p_k)$ with $1 \le p_1 < p_2 < \cdots < p_k$. For each $\mathbf{p} = (p_1, p_2, \ldots, p_k) \in \mathcal{P}$ and each sequence $x = (x_i)_{i\in\mathbb{N}}$ of real numbers define the number $\|x\|_{\mathbf{p}} \in [0, \infty)$ by $\|x\|_{\mathbf{p}} := 0$ when $k = 1$ and by

$$\|x\|_{\mathbf{p}} := \sqrt{\frac{1}{2}\left(\sum_{j=1}^{k-1} \left|x_{p_j} - x_{p_{j+1}}\right|^2 + \left|x_{p_k} - x_{p_1}\right|^2\right)} \tag{2.4.4}$$

when $k \ge 2$. The **James space** is the normed vector space defined by

$$J := \left\{ x \in c_0 \,\middle|\, \sup_{\mathbf{p}\in\mathcal{P}} \|x\|_{\mathbf{p}} < \infty \right\} \tag{2.4.5}$$

and

$$\|x\|_J := \sup_{\mathbf{p}\in\mathcal{P}} \|x\|_{\mathbf{p}} \tag{2.4.6}$$

for $x \in J$.

Before moving on to the main result of this section (Theorem 2.4.14) we explore some of the basic properties of the James space. This is the content of the next five lemmas.

Lemma 2.4.9. *The set J in* (2.4.5) *is a linear subspace of c_0 and $\|\cdot\|_J$ is a norm on J. With this norm J is a Banach space. Moreover,*

$$\|x\|_\infty \le \|x\|_J \le \sqrt{2}\,\|x\|_2 \qquad \text{for all } x \in c_0, \tag{2.4.7}$$

and thus $\ell^2 \subset J \subset c_0$.

Proof. By definition, $\|x+y\|_J \le \|x\|_J + \|y\|_J$ and $\|\lambda x\|_J = |\lambda|\|x\|_J$ for all $x, y \in c_0$ and all $\lambda \in \mathbb{R}$. Hence J is a linear subspace of c_0.

To prove the first inequality in (2.4.7), fix an element $\mathbf{p} = (i,j) \in \mathcal{P}$. Then $|x_i - x_j| = \|x\|_{\mathbf{p}} \le \|x\|_J$ for all $x \in c_0$ and all $i, j \in \mathbb{N}$ with $i < j$. Hence $|x_i| = \lim_{j\to\infty}|x_i - x_j| \le \|x\|_J$ for all $x \in c_0$ and all $i \in \mathbb{N}$. Now fix any element $\mathbf{p} = (p_1, p_2, \dots, p_k) \in \mathcal{P}$. Then

$$\begin{aligned}
\|x\|_{\mathbf{p}}^2 &= \frac{1}{2}\sum_{j=1}^{k-1}\left|x_{p_j} - x_{p_{j+1}}\right|^2 + \frac{1}{2}\left|x_{p_k} - x_{p_1}\right|^2 \\
&\le \sum_{j=1}^{k-1}\left|x_{p_j}\right|^2 + \sum_{j=1}^{k-1}\left|x_{p_{j+1}}\right|^2 + \left|x_{p_k}\right|^2 + \left|x_{p_1}\right|^2 \\
&= 2\sum_{j=1}^{k}\left|x_{p_j}\right|^2 \le 2\,\|x\|_2^2
\end{aligned}$$

for all $x \in c_0$. Take the supremum over all $\mathbf{p} \in \mathcal{P}$ to obtain $\|x\|_J \le \sqrt{2}\|x\|_2$. This proves (2.4.7). By (2.4.7) there are natural inclusions $\ell^2 \subset J \subset c_0$. Moreover, it follows from (2.4.7) that $\|x\|_J \ne 0$ for every $x \in J \setminus \{0\}$ and so $(J, \|\cdot\|_J)$ is a normed vector space.

We prove that J is complete. Let $(x_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in J. Then $(\|x_n\|_J)_{n\in\mathbb{N}}$ is a Cauchy sequence in $\mathbb{R}$, so the limit $C := \lim_{n\to\infty}\|x_n\|_J$ exists. Moreover, $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in c_0 by (2.4.7) and hence converges in the supremum norm to an element $x \in c_0$. Thus

$$\|x\|_{\mathbf{p}} = \lim_{n\to\infty}\|x_n\|_{\mathbf{p}} \le \lim_{n\to\infty}\|x_n\|_J = C$$

for all $\mathbf{p} \in \mathcal{P}$. Take the supremum over all $\mathbf{p} \in \mathcal{P}$ to obtain $x \in J$. We must prove that $\lim_{n\to\infty}\|x_n - x\|_J = 0$. To see this, fix a number $\varepsilon > 0$ and choose an integer $n_0 \in \mathbb{N}$ such that $\|x_n - x_m\|_J < \varepsilon/2$ for all integers $m, n \ge n_0$. Then $\|x_n - x\|_{\mathbf{p}} = \lim_{m\to\infty}\|x_n - x_m\|_{\mathbf{p}} \le \sup_{m\ge n_0}\|x_n - x_m\|_J \le \varepsilon/2$ for all $\mathbf{p} \in \mathcal{P}$ and all $n \ge n_0$. Thus $\|x_n - x\|_J = \sup_{\mathbf{p}\in\mathcal{P}}\|x_n - x\|_{\mathbf{p}} \le \varepsilon/2 < \varepsilon$ for every integer $n \ge n_0$. This shows that $\lim_{n\to\infty}\|x_n - x\|_J = 0$ and completes the proof of Lemma 2.4.9. □

The next goal is to prove that ℓ^2 is dense in J. For this it is convenient to introduce another norm on J. For every sequence $x = (x_i)_{i\in\mathbb{N}}$ of real numbers and every $\mathbf{p} = (p_1, p_2, \dots, p_k) \in \mathcal{P}$ define $|||x|||_{\mathbf{p}} := |x_{p_1}|$ in the case $k = 1$ and

$$|||x|||_{\mathbf{p}} := \sqrt{\frac{1}{2}\left(|x_{p_1}|^2 + \sum_{j=1}^{k-1} \left|x_{p_j} - x_{p_{j+1}}\right|^2 + |x_{p_k}|^2\right)} \tag{2.4.8}$$

in the case $k \geq 2$. Denote the supremum of these numbers over all $\mathbf{p} \in \mathcal{P}$ by

$$|||x|||_J := \sup_{\mathbf{p}\in\mathcal{P}} |||x|||_{\mathbf{p}}. \tag{2.4.9}$$

This is a norm on J that is equivalent to $\|\cdot\|_J$. Care must be taken. The second estimate in (2.4.10) below holds for $x \in c_0$ but not for all $x \in \ell^\infty$.

Lemma 2.4.10. *Every $x \in c_0$ satisfies the inequalities*

$$\frac{1}{\sqrt{2}} \|x\|_J \leq |||x|||_J \leq \|x\|_J\,. \tag{2.4.10}$$

Moreover, the function $J \to [0, \infty) : x \mapsto |||x|||_J$ is a norm.

Proof. Let $x \in c_0$ and $\mathbf{p} = (p_1, \dots, p_k) \in \mathcal{P}$. Then

$$\begin{aligned}
\|x\|_{\mathbf{p}}^2 &= \frac{1}{2}\left(\sum_{j=1}^{k} \left|x_{p_j} - x_{p_{j+1}}\right|^2 + |x_{p_k} - x_{p_1}|^2\right) \\
&\leq \sum_{j=1}^{k} \left|x_{p_j} - x_{p_{j+1}}\right|^2 + |x_{p_k}|^2 + \|x_{p_1}\|^2 \\
&= 2|||x|||_{\mathbf{p}}^2 \leq 2|||x|||_J^2.
\end{aligned}$$

Take the supremum over all $\mathbf{p} \in \mathcal{P}$ to obtain the inequality $\|x\|_J \leq \sqrt{2}|||x|||_J$. Now define $\mathbf{q}_n := (p_1, \dots, p_k, n)$ for every integer $n > p_k$. Then

$$\begin{aligned}
|||x|||_{\mathbf{p}}^2 &= \frac{1}{2}\left(\sum_{j=1}^{k} \left|x_{p_j} - x_{p_{j+1}}\right|^2 + |x_{p_k}|^2 + |x_{p_1}|^2\right) \\
&= \lim_{n\to\infty} \frac{1}{2}\left(\sum_{j=1}^{k} \left|x_{p_j} - x_{p_{j+1}}\right|^2 + |x_{p_k} - p_n|^2 + |p_n - x_{p_1}|^2\right) \\
&= \lim_{n\to\infty} \|x\|_{\mathbf{q}_n}^2 \leq \|x\|_J^2\,.
\end{aligned}$$

Take the supremum over all $\mathbf{p} \in \mathcal{P}$ to obtain the inequality $|||x|||_J \leq \|x\|_J$. This proves Lemma 2.4.10. □

Lemma 2.4.11. *The subspace ℓ^2 is dense in J.*

Proof. Fix a nonzero element $x \in J$ and a real number $\varepsilon > 0$, and choose a constant $0 < \delta < |||x|||_J$ such that

$$2\delta |||x|||_J < \varepsilon^2. \tag{2.4.11}$$

We claim that there are elements $n \in \mathbb{N}$ and $\mathbf{p} = (p_1, \dots, p_k) \in \mathcal{P}$ such that

$$\sup_{i \geq n} |x_i| < \delta, \qquad |||x|||_{\mathbf{p}} > |||x|||_J - \delta, \qquad p_k = n. \tag{2.4.12}$$

Namely, choose $n \in \mathbb{N}$ such that $\sup_{i \geq n} |x_i| < \delta$ and $\mathbf{p}_0 = (p_1, \dots, p_{k-1}) \in \mathcal{P}$ such that $|||x|||_{\mathbf{p}_0} > |||x|||_J - \delta$. Next choose $p_k > p_{k-1}$ so large that $p_k \geq n$ and the tuple $\mathbf{p} := (p_1, \dots, p_k)$ satisfies

$$|||x|||_{\mathbf{p}} = \sqrt{|||x|||_{\mathbf{p}_0}^2 - \tfrac{1}{2}|x_{p_{k-1}}|^2 + \tfrac{1}{2}|x_{p_{k-1}} - x_{p_k}|^2 + \tfrac{1}{2}|x_{p_k}|^2} > |||x|||_J - \delta.$$

Then increase n, if necessary, to obtain $p_k = n$.

Define $\xi := (x_1, \dots, x_n, 0, \dots)$. We prove that

$$|||x - \xi|||_J < \varepsilon. \tag{2.4.13}$$

To see this, let $\mathbf{q} = (q_1, \dots, q_\ell) \in \mathcal{P}$. If $q_\ell \leq n$, then $|||x - \xi|||_{\mathbf{q}} = 0$. Thus assume $q_\ell > n$, let $j \in \{1, \dots, \ell\}$ be the smallest element such that $q_j > n$, and define $\mathbf{q}' := (q_j, q_{j+1}, \dots, q_\ell) \in \mathcal{P}$. Then

$$|||x - \xi|||_{\mathbf{q}} = |||x|||_{\mathbf{q}'}. \tag{2.4.14}$$

Now consider the tuple $\mathbf{p}' := (p_1, \dots, p_k, q_j, q_{j+1}, \dots, q_\ell) \in \mathcal{P}$. By (2.4.12), it satisfies the inequality

$$\begin{aligned} |||x|||_J^2 &\geq |||x|||_{\mathbf{p}'}^2 \\ &= |||x|||_{\mathbf{p}}^2 + |||x|||_{\mathbf{q}'}^2 + \tfrac{1}{2}\left|x_{p_k} - x_{q_j}\right|^2 - \tfrac{1}{2}\left|x_{p_k}\right|^2 - \tfrac{1}{2}\left|x_{q_j}\right|^2 \\ &> (|||x|||_J - \delta)^2 - \delta^2 + |||x|||_{\mathbf{q}'}^2 \\ &= |||x|||_J^2 - 2\delta|||x|||_J + |||x|||_{\mathbf{q}'}^2. \end{aligned}$$

This implies $|||x|||_{\mathbf{q}'}^2 < 2\delta |||x|||_J$ and hence

$$|||x - \xi|||_{\mathbf{q}} = |||x|||_{\mathbf{q}'} < \sqrt{2\delta |||x|||_J} < \varepsilon$$

by (2.4.11) and (2.4.14). Take the supremum over all elements $\mathbf{q} \in \mathcal{P}$ to obtain the inequality (2.4.13). By (2.4.13) the set c_{00} of all finite sequences is dense in J and so is the subspace ℓ^2. This proves Lemma 2.4.11. □

The following lemma shows that the standard basis vectors $e_i := (\delta_{ij})_{j \in \mathbb{N}}$ form a Schauder basis of J (see Exercise 2.5.12).

Lemma 2.4.12. *For each $n \in \mathbb{N}$ define the projection $\Pi_n : J \to J$ by*

$$\Pi_n(x) := \sum_{i=1}^{n} x_i e_i \qquad \text{for } x = (x_i)_{i\in\mathbb{N}} \in J. \tag{2.4.15}$$

Then

$$\|\Pi_n(x)\|_J \le \|x\|_J, \qquad \|x - \Pi_n(x)\|_J \le \|x\|_J \tag{2.4.16}$$

for all $n \in \mathbb{N}$ and all $x \in J$, and

$$\lim_{n\to\infty} \|x - \Pi_n(x)\|_J = 0 \tag{2.4.17}$$

for all $x \in J$.

Proof. We prove (2.4.16). Fix an element $x \in J$, a positive integer n, and an element $\mathbf{p} = (p_1, \dots, p_k) \in \mathcal{P}$. If $p_k \le n$, then $\|\Pi_n(x)\|_{\mathbf{p}} = \|x\|_{\mathbf{p}}$ and, if $p_1 > n$, then $\|\Pi_n(x)\|_{\mathbf{p}} = 0$. Thus assume

$$p_1 \le n < p_k,$$

let $\ell \in \{1, \dots, k-1\}$ be the largest element such that $p_\ell \le n$, and define

$$\mathbf{q} := (p_1, \dots, p_\ell).$$

Then

$$\begin{aligned} \|\Pi_n(x)\|_{\mathbf{p}}^2 &= \|x\|_{\mathbf{q}}^2 - \tfrac{1}{2}\left|x_{p_\ell} - x_{p_1}\right|^2 + \tfrac{1}{2}\left|x_{p_\ell}\right|^2 + \tfrac{1}{2}\left|x_{p_1}\right|^2 \\ &= |||x|||_{\mathbf{q}}^2 \\ &\le \|x\|_J^2 \end{aligned}$$

by Lemma 2.4.10. Thus $\|\Pi_n(x)\|_{\mathbf{p}} \le \|x\|_J$ for all $\mathbf{p} \in \mathcal{P}$ and this proves the first inequality in (2.4.16). To prove the second inequality in (2.4.16), observe that $\|x - \Pi_n(x)\|_{\mathbf{p}} = \|x\|_{\mathbf{p}}$ whenever $p_1 > n$ and $\|x - \Pi_n(x)\|_{\mathbf{p}} = 0$ whenever $p_k \le n$. Thus assume $p_1 \le n < p_k$, let $\ell \in \{2, \dots, k\}$ be the smallest element such that $p_\ell > n$, and define $\mathbf{q} := (p_\ell, \dots, p_k)$. Then

$$\begin{aligned} \|x - \Pi_n(x)\|_{\mathbf{p}}^2 &= \|x\|_{\mathbf{q}}^2 - \tfrac{1}{2}\left|x_{p_\ell} - x_{p_k}\right|^2 + \tfrac{1}{2}\left|x_{p_\ell}\right|^2 + \tfrac{1}{2}\left|x_{p_k}\right|^2 \\ &= |||x|||_{\mathbf{q}}^2 \\ &\le \|x\|_J^2 \end{aligned}$$

by Lemma 2.4.10. Thus $\|x - \Pi_n(x)\|_{\mathbf{p}} \le \|x\|_J$ for all $\mathbf{p} \in \mathcal{P}$ and this proves the second inequality in (2.4.16).

We prove (2.4.17). When $x \in \ell^2$ this follows from (2.4.7). Since ℓ^2 is dense in J by Lemma 2.4.11, it follows from the estimate (2.4.16) and the Banach–Steinhaus Theorem 2.1.5 that (2.4.17) holds for all $x \in J$. This proves Lemma 2.4.12. □

With this preparation we are in a position to examine the dual space of the James space J. Fix a bounded linear functional $\Lambda : J \to \mathbb{R}$. By (2.4.7), the inclusion $\ell^2 \hookrightarrow J$ is a bounded linear operator and, by Lemma 2.4.11, it has a dense image. Thus the composition of Λ with this inclusion is a bounded linear functional $\Lambda|_{\ell^2} : \ell^2 \to \mathbb{R}$. Hence, by the Riesz Representation Theorem 1.4.4, there exists a unique sequence $y = (y_i)_{i\in\mathbb{N}} \in \ell^2$ such that

$$\Lambda(x) = \sum_{i=1}^{\infty} y_i x_i = \langle y, x\rangle \qquad \text{for all } x \in \ell^2 \subset J, \tag{2.4.18}$$

and, conversely, Λ is uniquely determined by this sequence $y \in \ell^2$. Thus the dual space of J can be identified with the space of all $y \in \ell^2$ such that

$$\|y\|_{J^*} := \sup_{0\neq x\in\ell^2} \frac{|\langle y, x\rangle|}{\|x\|_J} < \infty. \tag{2.4.19}$$

By (2.4.7) and (2.4.19), every $y \in J^*$ satisfies the inequalities

$$\frac{1}{\sqrt{2}} \|y\|_2 \le \|y\|_{J^*} \le \|y\|_1 . \tag{2.4.20}$$

Thus there are canonical inclusions

$$\ell^1 \subset J^* \subset \ell^2 \subset J \subset c_0.$$

At this point it is convenient to make use of two concepts that will only be introduced in Chapters 3 and 4. These are the dual operator $A^* : Y^* \to X^*$ of a bounded linear operator $A : X \to Y$ (Definition 4.1.1) and the weak* topology on the dual space of a Banach space (Example 3.1.9). A useful fact is that the dual operator has the same operator norm as the original operator (Lemma 4.1.2). Under our identification of J^* with a subspace of ℓ^2, the dual operator of the projection $\Pi_n : J \to J$ in (2.4.15) is the operator

$$\Pi_n : J^* \to J^*, \qquad \Pi_n(y) := \sum_{i=1}^{n} y_i e_i \quad \text{for } y = (y_i)_{i\in\mathbb{N}} \in J^*. \tag{2.4.21}$$

Thus it follows from the estimates in (2.4.16) that

$$\|\Pi_n(y)\|_{J^*} \le \|y\|_{J^*} , \qquad \|y - \Pi_n(y)\|_{J^*} \le \|y\|_{J^*} \tag{2.4.22}$$

for all $y \in J^*$ and all $n \in \mathbb{N}$. Moreover, the dual space of c_0 can be identified with ℓ^1 (Example 1.3.7) and the dual operator of the inclusion $J \hookrightarrow c_0$ is then the inclusion $\ell^1 \hookrightarrow J^*$. Hence it follows from general considerations that ℓ^1 is dense in J^* with respect to the weak* topology (Theorem 4.1.8). The next lemma shows that ℓ^1 is dense in J^* with respect to the norm topology.

Lemma 2.4.13. *Every* $y \in J^*$ *satisfies*

$$\lim_{n\to\infty} \|y - \Pi_n(y)\|_{J^*} = 0. \tag{2.4.23}$$

Proof. Fix an element $y \in J^*$. We prove that

$$\varepsilon_n := \|y - \Pi_n(y)\|_{J^*} = \sup_{\substack{0\neq x\in J\\ \Pi_n(x)=0}} \frac{|\langle y, x\rangle|}{\|x\|_J} \geq \varepsilon_{n+1} \tag{2.4.24}$$

for all $n \in \mathbb{N}$. To see this, fix an integer $n \in \mathbb{N}$ and recall from Lemma 2.4.12 that $\|x - \Pi_n(x)\|_J \leq \|x\|_J$ for all $x \in J$. Hence

$$\begin{aligned}
\|y - \Pi_n(y)\|_{J^*} &= \sup_{0\neq x\in J} \frac{|\langle y - \Pi_n(y), x\rangle|}{\|x\|_J} \\
&\leq \sup_{\substack{x\in J\\ \Pi_n(x)\neq x}} \frac{|\langle y, x - \Pi_n(x)\rangle|}{\|x - \Pi_n(x)\|_J} \\
&= \sup_{\substack{0\neq x\in J\\ \Pi_n(x)=0}} \frac{|\langle y, x\rangle|}{\|x\|_J} \\
&\leq \sup_{0\neq x\in J} \frac{|\langle y - \Pi_n(y), x\rangle|}{\|x\|_J} \\
&= \|y - \Pi_n(y)\|_{J^*}.
\end{aligned}$$

This proves the second equality in (2.4.24). This equality also shows that the sequence $(\varepsilon_n)_{n\in\mathbb{N}}$ is nonincreasing. Thus we have proved (2.4.24).

Now suppose, by contradiction, that $\lim_{n\to\infty} \varepsilon_n = \inf_{n\in\mathbb{N}} \varepsilon_n > 0$. Choose a constant $0 < \varepsilon < \inf_{n\in\mathbb{N}} \varepsilon_n$. Then, by (2.4.24) and the axiom of countable choice, there exists a sequence of sequences $x_n = (x_{n,i})_{i\in\mathbb{N}} \in J$ such that

$$\Pi_n(x_n) = 0, \qquad \|x_n\|_J = 1, \qquad \langle y, x_n\rangle > \varepsilon \tag{2.4.25}$$

for all $n \in \mathbb{N}$. Since c_{00} is dense in J by Lemma 2.4.12, the sequence can be chosen such that $x_n \in c_{00}$ for all $n \in \mathbb{N}$. By Lemma 2.4.9 each element x_n satisfies $\|x_n\|_\infty \leq \|x_n\|_J = 1$. Define the map $\kappa : \mathbb{N} \to \mathbb{N}$ by

$$\kappa(n) := \max\{i \in \mathbb{N} \mid x_{n,i} \neq 0\} \qquad \text{for } n \in \mathbb{N}. \tag{2.4.26}$$

Then $\kappa(n) > n$ for all $n \in \mathbb{N}$. Next define the sequence $n_j \in \mathbb{N}$ by $n_1 := 1$ and $n_{j+1} := \kappa(n_j) > n_j$ for $j \in \mathbb{N}$, and define the sequence $\xi = (\xi_i)_{i\in\mathbb{N}} \in c_0$ by $\xi_1 := 0$ and

$$\xi_i := \frac{x_{n_j,i}}{j} \qquad \text{for } j \in \mathbb{N} \text{ and } n_j + 1 \leq i \leq n_{j+1} = \kappa(n_j). \tag{2.4.27}$$

This sequence converges to zero because $|x_{n_j,i}| \le 1$ for all i and j. Moreover, it follows from (2.4.25), (2.4.26), and (2.4.27) that

$$\langle y, \Pi_{n_k}(\xi)\rangle = \sum_{i=1}^{n_k} y_i\xi_i = \sum_{j=1}^{k-1} \frac{\langle y, x_{n_j}\rangle}{j} \ge \sum_{j=1}^{k-1} \frac{\varepsilon}{j} \qquad \text{for all } k \in \mathbb{N}. \tag{2.4.28}$$

Now let $\mathbf{p} = (p_1, \dots, p_\ell) \in \mathcal{P}$. If $p_1 = 1$, then $|||\xi|||_{\mathbf{p}} = |||\xi|||_{(p_2,\dots,p_\ell)}$ in the case $\ell \ge 2$ and $|||\xi|||_{\mathbf{p}} = 0$ in the case $\ell = 1$. Thus assume $p_1 \ge 2$ and define

$$\mathcal{J} := \{ j \in \mathbb{N} \,|\, \text{ there exists an } i \in \{1, \dots, \ell\} \text{ such that } n_j < p_i \le n_{j+1} \}.$$

Then $\mathcal{J} \ne \emptyset$. Let $m := \max \mathcal{J}$ and define

$$\begin{aligned} k_j &:= \min \{ i \in \{1, \dots, \ell\} \,|\, n_j < p_i \le n_{j+1} \}, \\ \ell_j &:= \max \{ i \in \{1, \dots, \ell\} \,|\, n_j < p_i \le n_{j+1} \}, \\ \mathbf{p}_j &:= (p_{k_j}, \dots, p_{\ell_j}) \end{aligned}$$

for each $j \in \mathcal{J}$. Then $\{1, \dots, \ell\} = \bigcup_{j\in\mathcal{J}} \{k_j, \dots, \ell_j\}$ because $p_1 \ge 2$, and

$$|||\xi|||_{\mathbf{p}_j} = j^{-1} |||x_{n_j}|||_{\mathbf{p}_j} \le j^{-1} |||x_{n_j}|||_J \le j^{-1} \|x_{n_j}\|_J = j^{-1}$$

for all $j \in \mathcal{J}$ by (2.4.10) and (2.4.25). Hence

$$\begin{aligned} 2|||\xi|||_{\mathbf{p}}^2 &= |\xi_{p_1}|^2 + \sum_{m\ne j\in\mathcal{J}} \sum_{i=k_j}^{\ell_j} |\xi_{p_i} - \xi_{p_{i+1}}|^2 + \sum_{i=k_m}^{\ell_m - 1} |\xi_{p_i} - \xi_{p_{i+1}}|^2 + |\xi_{p_\ell}|^2 \\ &\le 2 \sum_{j\in\mathcal{J}} \left(|\xi_{p_{k_j}}|^2 + \sum_{i=k_j}^{\ell_j - 1} |\xi_{p_i} - \xi_{p_{i+1}}|^2 + |\xi_{p_{\ell_j}}|^2 \right) \\ &= 4 \sum_{j\in\mathcal{J}} |||\xi|||_{\mathbf{p}_j}^2 \\ &\le 4 \sum_{j\in\mathcal{J}} \frac{1}{j^2} \\ &\le \frac{2}{3}\pi^2. \end{aligned}$$

Take the supremum over all $\mathbf{p} \in \mathcal{P}$ and use Lemma 2.4.10 to obtain

$$\|\xi\|_J \le \sqrt{2}|||\xi|||_J = \sup_{\mathbf{p}\in\mathcal{P}} \sqrt{2}|||\xi|||_{\mathbf{p}} \le \sqrt{\frac{2}{3}}\pi < \infty$$

and so $\xi \in J$. It then follows from Lemma 2.4.12 that

$$\|\Pi_{n_k}(\xi)\|_J \le \|\xi\|_J \le \sqrt{\frac{2}{3}}\pi$$

for all $k \in \mathbb{N}$, in contradiction to the fact that the sequence $\langle y, \Pi_{n_k}(\xi)\rangle$ is unbounded by (2.4.28). This proves Lemma 2.4.13. □

We are now in a position to prove the main result of this subsection.

Theorem 2.4.14 (James). *The James space J is isometrically isomorphic to its bidual space J^{**} and the image of the canonical inclusion*

$$\iota : J \to J^{**}$$

has codimension one in J^{**}.

Proof. The proof has seven steps.

Step 1. *Let $\Lambda : J^* \to \mathbb{R}$ be a bounded linear functional and define*

$$z_i := \Lambda(e_i) \qquad \text{for } i \in \mathbb{N}.$$

Then $z := (z_i)_{i\in\mathbb{N}} \in \ell^\infty$,

$$\Lambda(\Pi_n(y)) = \langle y, \Pi_n(z)\rangle \tag{2.4.29}$$

for all $n \in \mathbb{N}$ and all $y \in J^*$, and

$$\Lambda(y) = \lim_{n\to\infty} \langle y, \Pi_n(z)\rangle \tag{2.4.30}$$

for all $y \in J^*$.

For every $i \in \mathbb{N}$ we have

$$|\langle e_i, e_i\rangle| = 1 = \|e_i\|_J$$

and thus

$$1 \le \|e_i\|_{J^*} = \sup_{0\ne x\in J} \frac{|\langle x, e_i\rangle|}{\|x\|_J} = \sup_{0\ne x\in J} \frac{|x_i|}{\|x\|_J} \le \sup_{0\ne x\in J} \frac{\|x\|_\infty}{\|x\|_J} \le 1$$

by Lemma 2.4.9. Hence

$$\|e_i\|_{J^*} = 1 \qquad \text{for all } i \in \mathbb{N}.$$

This implies

$$|z_i| = |\Lambda(e_i)| \le \|\Lambda\| \, \|e_i\|_{J^*} = \|\Lambda\|$$

for all $i \in \mathbb{N}$ and so $z \in \ell^\infty$. Now let $y = (y_i)_{i\in\mathbb{N}} \in J^*$. Then

$$\Lambda(\Pi_n(y)) = \sum_{i=1}^n y_i \Lambda(e_i) = \sum_{i=1}^n y_i z_i = \langle y, \Pi_n(z)\rangle \qquad \text{for all } n \in \mathbb{N}$$

and this proves (2.4.29). It follows from (2.4.23) and (2.4.29) that

$$\begin{aligned}
\lim_{n\to\infty} |\Lambda(y) - \langle y, \Pi_n(z)\rangle| &= \lim_{n\to\infty} |\Lambda(y - \Pi_n(y))| \\
&\le \lim_{n\to\infty} \|\Lambda\| \, \|y - \Pi_n(y)\|_{J^*} \\
&= 0.
\end{aligned}$$

This proves (2.4.30) and Step 1.

Step 2. *Let $\Lambda : J^* \to \mathbb{R}$ and $z \in \ell^\infty$ be as in Step 1. Then*

$$\sup_{p\in\mathcal{P}} \max\{\|z\|_{\mathbf{p}}, |||z|||_{\mathbf{p}}\} \leq \|\Lambda\|. \tag{2.4.31}$$

Fix an element $\mathbf{p} = (p_1, \dots, p_k) \in \mathcal{P}$ and choose an integer $n \geq p_k$. Then

$$\begin{aligned}
\max\{\|z\|_{\mathbf{p}}, |||z|||_{\mathbf{p}}\} &= \max\{\|\Pi_n(z)\|_{\mathbf{p}}, |||\Pi_n(z)|||_{\mathbf{p}}\} \\
&\leq \|\Pi_n(z)\|_J \\
&= \sup_{0\neq y\in J^*} \frac{|\langle y, \Pi_n(z)\rangle|}{\|y\|_{J^*}} \\
&= \sup_{0\neq y\in J^*} \frac{|\Lambda(\Pi_n(y))|}{\|y\|_{J^*}} \\
&\leq \sup_{0\neq y\in J^*} \frac{\|\Lambda\| \, \|\Pi_n(y)\|_{J^*}}{\|y\|_{J^*}} \\
&\leq \|\Lambda\|.
\end{aligned}$$

Here the second step follows from Lemma 2.4.10, the third step follows from Lemma 2.4.1, the fourth step follows from (2.4.29), and the last step follows from (2.4.22). This proves (2.4.31) and Step 2.

Step 3. *Let $z = (z_i)_{i\in\mathbb{N}} \in \ell^\infty$ be a bounded sequence such that*

$$\sup_{p\in\mathcal{P}} \max\{\|z\|_{\mathbf{p}}, |||z|||_{\mathbf{p}}\} < \infty. \tag{2.4.32}$$

Then z is a Cauchy sequence and the sequence $x := (x_i)_{i\in\mathbb{N}}$, defined by

$$\lambda := \lim_{j\to\infty} z_j, \qquad x_i := z_i - \lambda \quad \textit{for } i \in \mathbb{N}, \tag{2.4.33}$$

is an element of J.

Suppose, by contradiction, that z is not a Cauchy sequence. Then there exist two subsequences $(z_{p_i})_{i\in\mathbb{N}}$ and $(z_{q_i})_{i\in\mathbb{N}}$ converging to different limits. Passing to further subsequences we may assume that $p_i < q_i < p_{i+1}$ for all $i \in \mathbb{N}$ and that there exists a constant $\varepsilon > 0$ such that $|z_{p_i} - z_{q_j}| > \varepsilon$ for all $i, j \in \mathbb{N}$. For $n \in \mathbb{N}$ consider the tuple

$$\mathbf{p}_n := (p_1, q_1, p_2, q_2, \dots, p_n, q_n).$$

Then $\|z\|_{\mathbf{p}_n} > \sqrt{n}\varepsilon$ for all $n \in \mathbb{N}$, in contradiction to (2.4.32). This shows that z is a Cauchy sequence. Now the sequence x in (2.4.33) converges to zero, by definition, and satisfies

$$\|x\|_J = \sup_{\mathbf{p}\in\mathcal{P}} \|x\|_{\mathbf{p}} = \sup_{\mathbf{p}\in\mathcal{P}} \|z\|_{\mathbf{p}} < \infty$$

by (2.4.32). Hence $x \in J$ and this proves Step 3.

Step 4. *Let $z = (z_i)_{i\in\mathbb{N}} \in \ell^\infty$ be a bounded sequence that satisfies* (2.4.32) *and let $\lambda \in \mathbb{R}$ and $x \in J$ be given by* (2.4.33). *Then the limit*

$$\begin{aligned}\Lambda(y) &:= \lim_{n\to\infty} \langle y, \Pi_n(z)\rangle \\ &= \lim_{n\to\infty} \langle y, \Pi_n(x)\rangle + \lambda \lim_{n\to\infty} \sum_{i=1}^n y_i\end{aligned} \tag{2.4.34}$$

exists for every $y \in J^*$ and defines a linear functional $\Lambda : J^* \to \mathbb{R}$.

That the sequence $(\sum_{i=1}^n y_i)_{n\in\mathbb{N}}$ converges for $y \in \ell^1$ is obvious. Moreover, the subspace ℓ^1 is dense in J^* by Lemma 2.4.13, and

$$\begin{aligned}\left|\sum_{i=1}^n y_i\right| &= |\langle \mathbb{1}_n, y\rangle| \\ &\le \|\mathbb{1}_n\|_J \, \|y\|_{J^*} \\ &= \|y\|_{J^*}\end{aligned}$$

for all $n \in \mathbb{N}$. Here

$$\mathbb{1}_n := (1, \dots, 1, 0, \dots)$$

denotes the sequence whose first n entries are equal to one, followed by zeros. Hence the sequence of functionals

$$J^* \to \mathbb{R} : y \mapsto y_1 + \cdots + y_n$$

is uniformly bounded and converges for all y belonging to the dense subspace $\ell^1 \subset J^*$. Thus it follows from the Banach–Steinhaus Theorem 2.1.5 that the sequence $(\sum_{i=1}^n y_i)_{n\in\mathbb{N}}$ converges for all $y \in J^*$. Hence it follows from Step 3 and Lemma 2.4.12 that the limit in (2.4.34) exists for all $y \in J^*$ and this proves Step 4.

Step 5. *Let $z \in \ell^\infty$ be a sequence that satisfies* (2.4.32) *and let $\Lambda : J^* \to \mathbb{R}$ be the linear map defined by* (2.4.34) *in Step 4. Then Λ is a bounded linear functional on J^* and its norm is*

$$\|\Lambda\| = \sup_{p\in\mathcal{P}} \max\{\|z\|_{\mathbf{p}}, |||z|||_{\mathbf{p}}\}. \tag{2.4.35}$$

We prove that

$$\frac{|\Lambda(y)|}{\|y\|_{J^*}} \le \sup_{p\in\mathcal{P}} \max\{\|z\|_{\mathbf{p}}, |||z|||_{\mathbf{p}}\} \qquad \text{for all } y \in J^* \setminus \{0\}. \tag{2.4.36}$$

To see this, note first that

$$\|x\|_J = \sup_{p\in\mathcal{P}} \|x\|_{\mathbf{p}} = \sup_{p\in\mathcal{P}} \max\{\|x\|_{\mathbf{p}}, |||x|||_{\mathbf{p}}\} \tag{2.4.37}$$

for all $x \in J$ by Lemma 2.4.10.

Next we prove the inequality

$$\sup_{p\in\mathcal{P}} \max\{\|\Pi_n(z)\|_{\mathbf{p}}, |||\Pi_n(z)|||_{\mathbf{p}}\} \le \sup_{p\in\mathcal{P}} \max\{\|z\|_{\mathbf{p}}, |||z|||_{\mathbf{p}}\} \tag{2.4.38}$$

for all $n \in \mathbb{N}$. To see this, fix two elements

$$\mathbf{p} = (p_1, \dots, p_k) \in \mathcal{P}, \qquad n \in \mathbb{N}.$$

Then $\|\Pi_n(z)\|_{\mathbf{p}} = |||\Pi_n(z)|||_{\mathbf{p}} = 0$ whenever $p_1 > n$, and $|||\Pi_n(z)|||_{\mathbf{p}} = |||z|||_{\mathbf{p}}$ and $\|\Pi_n(z)\|_{\mathbf{p}} = \|z\|_{\mathbf{p}}$ whenever $p_k \le n$. Thus assume

$$p_1 \le n < p_k$$

and denote by $\ell \in \{1, \dots, k-1\}$ the largest number such that $p_\ell \le n$. Consider the element

$$\mathbf{q} := (p_1, \dots, p_\ell) \in \mathcal{P}.$$

It satisfies

$$\begin{aligned} 2\,\|\Pi_n(z)\|_{\mathbf{p}}^2 &= 2|||\Pi_n(z)|||_{\mathbf{p}}^2 \\ &= |z_{p_1}|^2 + \sum_{j=1}^{\ell-1} \left|z_{p_j} - z_{p_{j+1}}\right|^2 + |z_{p_\ell}|^2 \\ &= 2|||z|||_{\mathbf{q}}^2. \end{aligned}$$

This proves (2.4.38).

Now take $x = \Pi_n(z)$. Then, by (2.4.37) and (2.4.38),

$$\begin{aligned} \frac{|\langle y, \Pi_n(z)\rangle|}{\|y\|_{J^*}} &\le \|\Pi_n(z)\|_J \\ &= \sup_{p\in\mathcal{P}} \max\{\|\Pi_n(z)\|_{\mathbf{p}}, |||\Pi_n(z)|||_{\mathbf{p}}\} \\ &\le \sup_{p\in\mathcal{P}} \max\{\|z\|_{\mathbf{p}}, |||z|||_{\mathbf{p}}\} \end{aligned}$$

for all $y \in J^* \setminus \{0\}$ and all $n \in \mathbb{N}$. Take the limit $n \to \infty$. Then it follows from the definition of Λ in Step 4 via equation (2.4.34) that

$$\begin{aligned} \frac{|\Lambda(y)|}{\|y\|_{J^*}} &= \lim_{n\to\infty} \frac{|\langle y, \Pi_n(z)\rangle|}{\|y\|_{J^*}} \\ &\le \sup_{p\in\mathcal{P}} \max\{\|z\|_{\mathbf{p}}, |||z|||_{\mathbf{p}}\} \end{aligned}$$

for all $y \in J^* \setminus \{0\}$. This proves (2.4.36). Thus $\Lambda : J^* \to \mathbb{R}$ is a bounded linear functional. Now take the supremum over all $y \in J^* \setminus \{0\}$ to obtain

$$\begin{aligned} \|\Lambda\| &= \sup_{0\ne y\in J^*} \frac{|\Lambda(y)|}{\|y\|_{J^*}} \\ &\le \sup_{p\in\mathcal{P}} \max\{\|z\|_{\mathbf{p}}, |||z|||_{\mathbf{p}}\}. \end{aligned}$$

The converse inequality was established in Step 2 and this proves Step 5.

Step 6. *The canonical inclusion $\iota : J \to J^{**}$ has a codimension-one image.*

By Step 1, Step 2, Step 4, and Step 5, the bidual space of J is naturally isomorphic to the space

$$J^{**} := \left\{ z \in \ell^\infty \,\middle|\, \sup_{\mathbf{p}\in\mathcal{P}} \max\{\|z\|_{\mathbf{p}}, |||z|||_{\mathbf{p}}\} < \infty \right\}.$$

The correspondence assigns to a sequence $z \in J^{**}$ the bounded linear functional $\Lambda : J^* \to \mathbb{R}$ given by (2.4.34). That it is well defined for every $z \in J^{**}$ was proved in Step 4, that it is bounded was proved in Step 5, and that every bounded linear functional on J^* is of this form was proved in Steps 1 and 2. It was also proved in Step 5 that the identification of J^{**} with the dual space of J^* is an isometry with respect to the norm on J^{**}, defined by

$$\|z\|_{J^{**}} := \sup_{\mathbf{p}\in\mathcal{P}} \max\{\|z\|_{\mathbf{p}}, |||z|||_{\mathbf{p}}\} \qquad \text{for } x \in J^{**}.$$

Under this identification, the canonical inclusion $\iota : J \to J^{**}$ is the obvious inclusion of J into J^{**} as a subset. It is an isometric embedding by the general observation in Lemma 2.4.1 (see also Lemma 2.4.10 and equation (2.4.37)). Moreover, the constant sequence $\mathbb{1} := (1, 1, 1, \dots)$ is a unit vector in J^{**} and

$$J^{**} = J \oplus \mathbb{R}\mathbb{1}$$

by Step 3. This proves Step 6.

Step 7. *The map*

$$J \to J^{**} : x = (x_i)_{i\in\mathbb{N}} \mapsto (x_{i+1} - x_1)_{i\in\mathbb{N}}$$

is an isometric isomorphism.

The map is bijective by Step 3. If $x = (x_i)_{i\in\mathbb{N}} \in J$ and $z = (z_i)_{i\in\mathbb{N}} \in J^{**}$ are related by the conditions

$$x_1 = -\lim_{j\to\infty} z_j, \qquad x_{i+1} - x_1 = z_i \qquad \text{for } i \in \mathbb{N},$$

then

$$\|z\|_{(p_1,\dots,p_k)} = \|x\|_{(p_1+1,\dots,p_k+1)}, \qquad |||z|||_{(p_1,\dots,p_k)} = \|x\|_{(1,p_1+1,\dots,p_k+1)}$$

for all $(p_1, \dots, p_k) \in \mathcal{P}$, and hence

$$\|x\|_J = \sup_{\mathbf{p}\in\mathcal{P}} \|x\|_{\mathbf{p}} = \sup_{\mathbf{p}\in\mathcal{P}} \max\{\|z\|_{\mathbf{p}}, |||z|||_{\mathbf{p}}\} = \|z\|_{J^{**}}.$$

This proves Step 7 and Theorem 2.4.14. □

Remark 2.4.15. **(i)** Let X be a real Banach space. A **Schauder basis** of X is a sequence $(e_i)_{i\in\mathbb{N}}$ in X such that, for every $x \in X$, there exists a unique sequence $(x_i)_{i\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$ such that

$$\lim_{n\to\infty}\Big\|x - \sum_{i=1}^{n} x_i e_i\Big\| = 0. \tag{2.4.39}$$

Associated to every Schauder basis $(e_i)_{i\in\mathbb{N}}$ of X is a unique sequence of bounded linear functionals $e_i^* \in X^*$ such that $\langle e_i^*, e_j\rangle = \delta_{ij}$ for all $i, j \in \mathbb{N}$ (see Exercise 2.5.12). Thus the sequence $x_i = \langle e_i^*, x\rangle$ is characterized by the condition (2.4.39). A Schauder basis $(e_i)_{i\in\mathbb{N}}$ is called **normalized** if $\|e_i\| = 1$ for all $i \in \mathbb{N}$. Associated to every Schauder basis $(e_i)_{i\in\mathbb{N}}$ and every $n \in \mathbb{N}$ is a projection $\Pi_n : X \to X$ via

$$\Pi_n(x) := \sum_{i=1}^{n} \langle e_i^*, x\rangle e_i \qquad \text{for } x \in X. \tag{2.4.40}$$

The operator sequence $\Pi_n \in \mathcal{L}(X)$ is bounded by Exercise 2.5.12. A Schauder basis $(e_i)_{i\in\mathbb{N}}$ is called **monotone** if $\|\Pi_n\| \le 1$ for all $n \in \mathbb{N}$. It is called **shrinking** if $\lim_{n\to\infty}\|\Pi_n^*(x^*) - x^*\|_{X^*} = 0$ for every $x^* \in X^*$ and so the sequence $(e_i^*)_{i\in\mathbb{N}}$ is a Schauder basis of X^*. It is called **boundedly complete** if, for every sequence $(x_i)_{i\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$ such that $\sup_{n\in\mathbb{N}}\|\sum_{i=1}^{n} x_i e_i\| < \infty$, the sequence $\sum_{i=1}^{n} x_i e_i$ converges in X.

(ii) By Lemma 2.4.12 the standard basis $(e_i)_{i\in\mathbb{N}}$ of the James space J is a normalized monotone Schauder basis and, by Lemma 2.4.13, it is shrinking. It is not boundedly complete, because the constant sequence $x_i = 1$ satisfies $\|\sum_{i=1}^{n} e_i\|_J = 1$, however, the sequence $\sum_{i=1}^{n} e_i$ does not converge in J.

(iii) The standard basis $(e_i)_{i\in\mathbb{N}}$ of the dual space J^* is again normalized and monotone. One can deduce from Lemma 2.4.13 that this basis is boundedly complete. However, it is not shrinking, because the closure of the span of the dual sequence in J^{**} is the proper subspace $J \subset J^{**}$ by Theorem 2.4.14.

(iv) A theorem of Robert C. James asserts that a Banach space X with a Schauder basis $(e_i)_{i\in\mathbb{N}}$ is reflexive if and only if the basis is both shrinking and boundedly complete.

(v) A Schauder basis $(e_i)_{i\in\mathbb{N}}$ of a Banach space X is called **unconditional** if the sequence $(e_{\sigma(i)})_{i\in\mathbb{N}}$ is a Schauder basis for every bijection $\sigma : \mathbb{N} \to \mathbb{N}$. The James space J does not admit an unconditional Schauder basis.

(vi) There are many examples of Schauder bases, such as any orthonormal basis of a separable Hilbert space, which is always normalized, monotone, unconditional, shrinking, and boundedly complete.

(vii) The reader may verify that the standard basis of ℓ^p for $1 < p < \infty$ is normalized, monotone, unconditional, boundedly complete, and shrinking. For $p = 1$ it is still normalized, monotone, unconditional, and boundedly complete, but no longer shrinking. The Banach space ℓ^∞ does not admit a Schauder basis, because it is not separable.

(viii) There exist separable Banach spaces that do not admit Schauder bases. Examples are Banach spaces that do not have the approximation property (see Exercises 4.2.11 and 4.2.12).

Remark 2.4.16. **(i)** A **complex structure** on a real Banach space X is a bounded linear operator $I : X \to X$ such that

$$I^2 = -\mathbb{1}.$$

Such a complex structure induces a complex structure $I^{**} : X^{**} \to X^{**}$ on the bidual space such that the canonical inclusion $\iota : X \to X^{**}$ satisfies

$$\iota \circ I = I^{**} \circ \iota.$$

Thus the complex structure descends to the quotient space $X^{**}/\iota(X)$. In the case of the James space $X = J$, this quotient has one real dimension. Hence it does not admit a complex structure, and neither does the James space J.

(ii) Consider the product

$$X := J \times J^*$$

of the James space J with its dual, equipped with the norm

$$\|(x, y)\|_X := \sqrt{\|x\|_J^2 + \|y\|_{J^*}^2} \qquad \text{for } (x, y) \in J \times J^*.$$

By Theorem 2.4.14 the space X is isometrically isomorphic to its dual space. However, it is not reflexive.

(iii) The James space J is an example of a nonreflexive Banach space whose bidual space is separable.

(iv) Another question answered in the negative by the James space is whether a separable Banach space that is isometrically isomorphic to its bidual space must be reflexive. The James space satisfies both conditions, but is not reflexive.

(v) The James space J is an example of an infinite-dimensional Banach space that is not isomorphic to the product space

$$X := J \times J$$

(equipped with any product norm as in Subsection 1.2.3). This is because the canonical inclusion $\iota : X \to X^{**}$ has codimension two by Theorem 2.4.14. Moreover, X admits a complex structure and J does not.

3.5. The Kreĭn–Milman Theorem

The Kreĭn–Milman Theorem [**47**, **60**] is a general result about compact convex subsets of a locally convex Hausdorff topological vector space. It asserts that every such convex subset is the closed convex hull of its set of extremal points. In particular, the result applies to the dual space of a Banach space, equipped with the weak* topology. Here are the relevant definitions.

Definition 3.5.1 (Extremal Point and Face). Let X be a real vector space and let $K \subset X$ be a nonempty convex subset. A subset

$$F \subset K$$

is called a **face of** K if F is a nonempty convex subset of K and

$$\begin{array}{l} x_0, x_1 \in K,\ 0 < \lambda < 1, \\ (1-\lambda)x_0 + \lambda x_1 \in F \end{array} \quad \Longrightarrow \quad x_0, x_1 \in F. \tag{3.5.1}$$

An element $x \in K$ is called an **extremal point of** K if

$$\begin{array}{l} x_0, x_1 \in K,\ 0 < \lambda < 1, \\ (1-\lambda)x_0 + \lambda x_1 = x \end{array} \quad \Longrightarrow \quad x_0 = x_1 = x. \tag{3.5.2}$$

This means that the singleton $F := \{x\}$ is a face of K or, equivalently, that there is no open line segment in K that contains x (see Figure 3.5.1). Denote the set of extremal points of K by

$$\mathcal{E}(K) := \{x \in K \,|\, x \text{ satisfies } (3.5.2)\}\,.$$

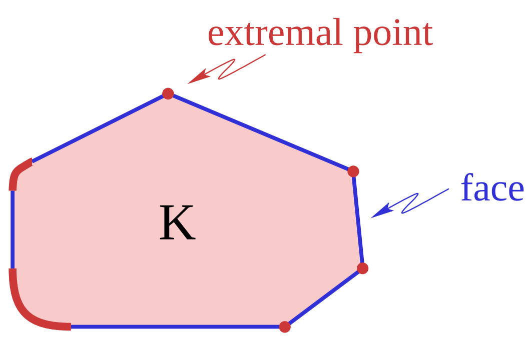

Figure 3.5.1. Extremal points and faces.

Recall that the convex hull of a set $E \subset X$ is denoted by $\mathrm{conv}(E)$ and that its closure, the closed convex hull of E, is denoted by $\overline{\mathrm{conv}}(E)$ whenever X is a topological vector space (see Definition 3.1.19).

Theorem 3.5.2 (Kreĭn–Milman). *Let X be a locally convex Hausdorff topological vector space and let $K \subset X$ be a nonempty compact convex set. Then K is the closed convex hull of its extremal points, i.e. $K = \overline{\mathrm{conv}}(\mathcal{E}(K))$. In particular, K admits an extremal point, i.e. $\mathcal{E}(K) \neq \emptyset$.*

Proof. The proof has five steps.

Step 1. *Let*

$$\mathscr{K} := \{K \subset X \,|\, K \text{ is a nonempty compact convex set}\}$$

and define the relation $\preccurlyeq$ *on* $\mathscr{K}$ *by*

$$F \preccurlyeq K \quad \overset{\text{def}}{\iff} \quad F \text{ is a face of } K \tag{3.5.3}$$

for $F, K \in \mathscr{K}$. *Then* $(\mathscr{K}, \preccurlyeq)$ *is a partially ordered set and every nonempty chain* $\mathscr{C} \subset \mathscr{K}$ *has an infimum.*

That the relation (3.5.3) is a partial order follows directly from the definition. Moreover, every element $K \in \mathscr{K}$ is a closed set because X is Hausdorff. This implies that every nonempty chain $\mathscr{C} \subset \mathscr{K}$ has an infimum

$$C_0 := \bigcap_{C \in \mathscr{C}} C.$$

This proves Step 1.

Step 2. *If* $K \in \mathscr{K}$ *and* $\Lambda : X \to \mathbb{R}$ *is a continuous linear functional, then*

$$F := K \cap \Lambda^{-1}(\sup_K \Lambda) \in \mathscr{K}$$

and $F \preccurlyeq K$.

Abbreviate $c := \sup_K \Lambda$. Since K is compact and Λ is continuous, the set $F = K \cap \Lambda^{-1}(c)$ is nonempty. Since K is closed and Λ is continuous, the set F is a closed subset of K and hence is compact. Since K is convex and Λ is linear, F is convex. Thus $F \in \mathscr{K}$.

To prove that F is a face of K, fix two elements $x_0, x_1 \in K$ and a real number $0 < \lambda < 1$ such that

$$x := (1 - \lambda)x_0 + \lambda x_1 \in F.$$

Then $(1 - \lambda)\Lambda(x_0) + \lambda\Lambda(x_1) = \Lambda(x) = c$ and hence

$$(1 - \lambda)(c - \Lambda(x_0)) + \lambda(c - \Lambda(x_1)) = 0.$$

Since $c - \Lambda(x_0) \geq 0$ and $c - \Lambda(x_1) \geq 0$, this implies

$$\Lambda(x_0) = \Lambda(x_1) = c$$

and hence $x_0, x_1 \in F$. Thus F is a face of K. This proves Step 2.

Step 3. *Every minimal element of $\mathscr{K}$ is a singleton.*

Fix an element $K \in \mathscr{K}$ which is not a singleton and choose two elements $x_0, x_1 \in K$ such that $x_0 \neq x_1$. Since X is a locally convex Hausdorff space, there exists a convex open set $U_1 \subset X$ such that $x_1 \in U_1$ and $x_0 \notin U_1$. Hence it follows from Theorem 3.1.11 that there exists a continuous linear functional $\Lambda : X \to \mathbb{R}$ such that $\Lambda(x_0) < \Lambda(x)$ for all $x \in U_1$ and so

$$\Lambda(x_0) < \Lambda(x_1).$$

By Step 2, the set $F := K \cap \Lambda^{-1}(\sup_K \Lambda)$ is a face of K and $x_0 \in K \setminus F$. Thus K is not a minimal element of $\mathscr{K}$.

Step 4. *Let $K \in \mathscr{K}$. Then $\mathcal{E}(K) \neq \emptyset$.*

By Step 1 and the Lemma of Zorn, there exists a minimal element $E \in \mathscr{K}$ such that $E \preccurlyeq K$. By Step 3,

$$E = \{x\}$$

is a singleton. Hence $x \in \mathcal{E}(K)$.

Step 5. *Let $K \in \mathscr{K}$. Then $K = \overline{\mathrm{conv}}(\mathcal{E}(K))$.*

It follows directly from the definitions that

$$\overline{\mathrm{conv}}(\mathcal{E}(K)) \subset K.$$

To prove the converse inclusion, assume, by contradiction, that there exists an element

$$x \in K \setminus \overline{\mathrm{conv}}(\mathcal{E}(K)).$$

Since X is a locally convex Hausdorff space, there exists an open convex set $U \subset X$ such that

$$x \in U, \qquad U \cap \overline{\mathrm{conv}}(\mathcal{E}(K)) = \emptyset.$$

Since $\mathcal{E}(K)$ is nonempty by Step 4, it follows from Theorem 3.1.11 that there exists a continuous linear functional $\Lambda : X \to \mathbb{R}$ such that

$$\Lambda(x) > \sup_{\overline{\mathrm{conv}}(\mathcal{E}(K))} \Lambda. \tag{3.5.4}$$

By Step 2, the set

$$F := K \cap \Lambda^{-1}(\sup_K \Lambda)$$

is a face of K and

$$F \cap \mathcal{E}(K) = \emptyset.$$

by (3.5.4). By Step 3, the set F has an extremal point x_0. Then x_0 is also an extremal point of K in contradiction to the fact that $F \cap \mathcal{E}(K) = \emptyset$. This proves Theorem 3.5.2. □

Example 3.5.3. This example shows that the extremal set of a compact convex set need not be compact. Let X be an infinite-dimensional reflexive Banach space. Assume X is **strictly convex**, i.e. for all $x, y \in X$,

$$\|x + y\| = 2\|x\| = 2\|y\| . \qquad \Longrightarrow \qquad x = y \tag{3.5.5}$$

Then the closed unit ball $B \subset X$ is weakly compact by Theorem 3.4.1 and its extremal set is the unit sphere $\mathcal{E}(B) = S$ (see Exercises 2.5.11 and 3.7.14). Thus the extremal set is not weakly compact and B is the weak closure of its extremal set by Lemma 3.1.21. **Exercise:** Prove that $B = \mathrm{conv}(S)$.

Example 3.5.4 (Infinite-Dimensional Simplex). The infinite product $\mathbb{R}^{\mathbb{N}}$ is a locally convex Hausdorff space with the product topology, induced by the metric

$$d(x, y) := \sum_{i=1}^{\infty} 2^{-i} \frac{|x_i - y_i|}{1 + |x_i - y_i|}$$

for $x = (x_i)_{i\in\mathbb{N}}$ and $y = (y_i)_{i\in\mathbb{N}}$ in $\mathbb{R}^{\mathbb{N}}$. The **infinite-dimensional simplex**

$$\Delta := \left\{ x = (x_i)_{i\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}} \,\middle|\, x_i \geq 0,\ \sum_{i=1}^{\infty} x_i \leq 1 \right\}$$

is a compact convex subset of $\mathbb{R}^{\mathbb{N}}$ by Tychonoff's Theorem A.2.1. Its set of extremal points is the compact set

$$\mathcal{E}(\Delta) = \{e_i \,|\, i \in \mathbb{N}\} \cup \{0\}, \qquad e_i := (\delta_{ij})_{j\in\mathbb{N}}.$$

The convex hull of $\mathcal{E}(\Delta)$ is strictly contained in Δ and hence is not compact. **Exercise:** The product topology on the infinite-dimensional simplex agrees with the weak* topology it inherits as a subset of $\ell^1 = c_0^*$ (see Example 1.3.7).

Example 3.5.5 (Hilbert Cube). The **Hilbert cube** is the set

$$Q := \left\{ x = (x_i)_{i\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}} \,\middle|\, 0 \leq x_i \leq 1/i \right\}.$$

This is a compact convex subset of $\mathbb{R}^{\mathbb{N}}$ with respect to the product topology. Its set of extremal points is the compact set

$$\mathcal{E}(Q) = \left\{ x = (x_i)_{i\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}} \,\middle|\, x_i \in \{0, 1/i\} \right\}.$$

The convex hull of any finite subset of $\mathcal{E}(Q)$ is nowhere dense in Q. Hence

$$\mathrm{conv}(\mathcal{E}(Q)) \subsetneq Q$$

by the Baire Category Theorem 1.6.4. **Exercise:** The product topology on the Hilbert cube agrees with the topology induced by the ℓ^2 norm.

3.6. Ergodic Theory

This section establishes the existence of an ergodic measure for any homeomorphism of a compact metric space. The proof is a fairly straightforward consequence of the Banach–Alaoglu Theorem 3.2.1 and the Kreĭn–Milman Theorem 3.5.2. We also show that the ergodic measures are precisely the extremal points of the convex set of all invariant measures (Theorem 3.6.3). The proof that every ergodic measure is extremal requires von Neumann's Mean Ergodic Theorem 3.6.5, the proof of which will in turn be based on an abstract ergodic theorem for operators on Banach spaces (Theorem 3.6.9).

3.6.1. Ergodic Measures. Let (M, d) be a compact metric space and let $\phi : M \to M$ be a homeomorphism. Denote by $\mathcal{B} \subset 2^M$ the Borel σ-algebra. Recall that the set $\mathcal{M}(\phi)$ of all ϕ-invariant Borel probability measures on M is a nonempty weak* compact convex subset of the space $\mathcal{M}(M) = C(M)^*$ of all signed Borel measures on M (see Subsection 3.2.2 and Corollary 3.2.6).

Definition 3.6.1 (**Ergodic Measure**). A ϕ-invariant Borel probability measure $\mu : \mathcal{B} \to [0, 1]$ is called **ϕ-ergodic** if, for every Borel set $B \subset M$,

$$\phi(B) = B \qquad \Longrightarrow \qquad \mu(B) \in \{0, 1\}. \tag{3.6.1}$$

The homeomorphism ϕ is called **μ-ergodic** if μ is an ergodic measure for ϕ.

Example 3.6.2. If $x \in M$ is a fixed point of ϕ, then the Dirac measure $\mu = \delta_x$ is ergodic for ϕ. If $\phi = \mathrm{id}$, then the Dirac measure at each point of M is ergodic for ϕ and there are no other ergodic measures.

Theorem 3.6.3 (**Ergodic Measures are Extremal**). *Let $\mu : \mathcal{B} \to [0, 1]$ be a ϕ-invariant Borel probability measure. Then the following are equivalent.*

(i) *μ is an ergodic measure for ϕ.*

(ii) *μ is an extremal point of $\mathcal{M}(\phi)$.*

Proof. We prove that (ii) implies (i) by an indirect argument. Assume that μ is not ergodic for ϕ. Then there exists a Borel set $\Lambda \subset M$ such that

$$\phi(\Lambda) = \Lambda, \qquad 0 < \mu(\Lambda) < 1.$$

Define $\mu_0, \mu_1 : \mathcal{B} \to [0, 1]$ by

$$\mu_0(B) := \frac{\mu(B \setminus \Lambda)}{1 - \mu(\Lambda)}, \qquad \mu_1(B) := \frac{\mu(B \cap \Lambda)}{\mu(\Lambda)}$$

for $B \in \mathcal{B}$. These are ϕ-invariant Borel probability measures and they are not equal because $\mu_0(\Lambda) = 0$ and $\mu_1(\Lambda) = 1$. Moreover, $\mu = (1 - \lambda)\mu_0 + \lambda\mu_1$ where $\lambda := \mu(\Lambda)$. Hence μ is not an extremal point of $\mathcal{M}(\phi)$. This shows that (ii) implies (i). The converse is proved on page 146. □

Corollary 3.6.4 (Existence of Ergodic Measures). *Every homeomorphism of a compact metric space admits an ergodic measure.*

Proof. The set $\mathcal{M}(\phi)$ of ϕ-invariant Borel probability measures on M is nonempty by Lemma 3.2.3 and is a weak* compact convex subset of $\mathcal{M}(M)$ by Corollary 3.2.6. Hence $\mathcal{M}(\phi)$ has an extremal point μ by Theorem 3.5.2. Thus μ is an ergodic measure by (ii) $\implies$ (i) in Theorem 3.6.3. □

3.6.2. Space and Time Averages. Given a homeomorphism

$$\phi : M \to M$$

of a compact metric space M, a ϕ-ergodic measure

$$\mu : \mathcal{B} \to [0, 1]$$

on the Borel σ-algebra $\mathcal{B} \subset 2^M$, a continuous function $f : M \to \mathbb{R}$, and an element $x \in M$, one can ask the question of whether the sequence of averages $\frac{1}{n}\sum_{k=0}^{n-1} f(\phi^k(x))$ converges. A theorem of Birkhoff [**13**] answers this question in the affirmative for almost every $x \in M$. This is **Birkhoff's Ergodic Theorem**. It asserts that, if μ is a ϕ-ergodic measure, then for every continuous function $f : M \to \mathbb{R}$, there exists a Borel set $\Lambda \subset M$ such that

$$\phi(\Lambda) = \Lambda, \qquad \mu(\Lambda) = 1, \tag{3.6.2}$$

and

$$\int_M f\, d\mu = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(\phi^k(x)) \qquad \text{for all } x \in \Lambda. \tag{3.6.3}$$

In other words, the *time average* of f agrees with the *space average* for almost every orbit of the dynamical system. If ϕ is **uniquely ergodic**, i.e. ϕ admits only one ergodic measure or, equivalently, only one ϕ-invariant Borel probability measure, then equation (3.6.3) actually holds for all $x \in M$. Birkhoff's Ergodic Theorem extends to μ-integrable functions and asserts that the sequence of measurable functions $\frac{1}{n}\sum_{k=0}^{n-1} f \circ \phi^k$ converges **pointwise** almost everywhere to the mean value of f. A particularly interesting case is where f is the characteristic function of a Borel set $B \subset M$. Then the integral of f is the measure of B and it follows from Birkhoff's Ergodic Theorem that

$$\mu(B) = \lim_{n\to\infty} \frac{\#\left\{k \in \{0, \dots, n-1\} \,\middle|\, \phi^k(x) \in B\right\}}{n} \tag{3.6.4}$$

for μ-almost all $x \in M$. A weaker result is von Neumann's Mean Ergodic Theorem [**62**]. It asserts that the sequence $\frac{1}{n}\sum_{k=0}^{n-1} f \circ \phi^k$ converges to the mean value of f in $L^p(\mu)$ for $1 < p < \infty$. This implies pointwise almost everywhere convergence for a suitable subsequence (see [**75**, Cor. 4.10]).

Theorem 3.6.5 (Von Neumann's Mean Ergodic Theorem). *Let (M, d) be a compact metric space, let $\phi : M \to M$ be a homeomorphism, let $\mu \in \mathcal{M}(\phi)$ be a ϕ-ergodic measure, let $1 < p < \infty$, and let $f \in L^p(\mu)$. Then*

$$\lim_{n\to\infty} \left\| \frac{1}{n} \sum_{k=0}^{n-1} f \circ \phi^k - \int_M f\,d\mu \right\|_{L^p} = 0. \tag{3.6.5}$$

Proof. See page 149. □

Theorem 3.6.5 implies Theorem 3.6.3. The proof has two steps.

Step 1. *Let $\mu_0, \mu_1 \in \mathcal{M}(\phi)$ be ergodic measures such that $\mu_0(\Lambda) = \mu_1(\Lambda)$ for every ϕ-invariant Borel set $\Lambda \subset M$. Then $\mu_0 = \mu_1$.*

Fix a continuous function $f : M \to \mathbb{R}$. Then it follows from Theorem 3.6.5 and [**75**, Cor. 4.10] that there exist Borel sets $B_0, B_1 \subset M$ and a sequence of integers $1 \le n_1 < n_2 < n_3 < \cdots$ such that $\mu_i(B_i) = 1$ and

$$\int_M f\,d\mu_i = \lim_{j\to\infty} \frac{1}{n_j} \sum_{k=0}^{n_j-1} f(\phi^k(x)) \qquad \text{for } x \in B_i \text{ and } i = 0, 1. \tag{3.6.6}$$

For $i = 0, 1$ define $\Lambda_i := \bigcap_{n\in\mathbb{Z}} \phi^n(B_i)$. So Λ_i is a ϕ-invariant Borel set such that $\mu_i(\Lambda_i) = 1$. Thus $\mu_1(\Lambda_0) = \mu_0(\Lambda_0) = 1$ and $\mu_0(\Lambda_1) = \mu_1(\Lambda_1) = 1$ by assumption. This implies that the ϕ-invariant Borel set $\Lambda := \Lambda_0 \cap \Lambda_1$ is nonempty. Since $\Lambda \subset B_0 \cap B_1$, it follows from (3.6.6) that

$$\int_M f\,d\mu_0 = \lim_{j\to\infty} \frac{1}{n_j} \sum_{k=0}^{n_j-1} f(\phi^k(x)) = \int_M f\,d\mu_1 \qquad \text{for all } x \in \Lambda.$$

Thus the integrals of f with respect to μ_0 and μ_1 agree for every continuous function $f : M \to \mathbb{R}$. Hence $\mu_0 = \mu_1$ by uniqueness in the Riesz Representation Theorem (see [**75**, Cor. 3.19]). This proves Step 1.

Step 2. *Let $\mu \in \mathcal{M}(\phi)$ be ergodic. Then μ is an extremal point of $\mathcal{M}(\phi)$.*

Let $\mu_0, \mu_1 \in \mathcal{M}(\phi)$ and $0 < \lambda < 1$ such that $\mu = (1-\lambda)\mu_0 + \lambda\mu_1$. If $B \subset M$ is a Borel set such that $\mu(B) = 0$, then $(1-\lambda)\mu_0(B) + \lambda\mu_1(B) = 0$, and hence $\mu_0(B) = \mu_1(B) = 0$ because $0 < \lambda < 1$. If $B \subset M$ is a Borel set such that $\mu(B) = 1$, then $\mu(M \setminus B) = 0$, hence $\mu_0(M \setminus B) = \mu_1(M \setminus B) = 0$, and therefore $\mu_0(B) = \mu_1(B) = 1$. Now let $\Lambda \subset M$ be a ϕ-invariant Borel set. Then $\mu(\Lambda) \in \{0, 1\}$ because μ is ϕ-ergodic, and hence $\mu_0(\Lambda) = \mu_1(\Lambda) = \mu(\Lambda)$. Thus μ_0 and μ_1 are ϕ-ergodic measures that agree on all ϕ-invariant Borel sets. Hence $\mu_0 = \mu_1 = \mu$ by Step 1 and this proves Step 2.

Step 2 shows that (i) implies (ii) in Theorem 3.6.3. The converse was proved on page 144. □

3.6.3. An Abstract Ergodic Theorem. Theorem 3.6.5 translates into a theorem about the iterates of a bounded linear operator from a Banach space to itself provided that these iterates are uniformly bounded. For an endomorphism

$$T : X \to X$$

of a vector space X and a positive integer n denote the nth iterate of T by

$$T^n := T \circ \cdots \circ T.$$

For $n = 0$ define

$$T^0 := \mathrm{id}.$$

The ergodic theorem in functional analysis asserts that, if $T : X \to X$ is a bounded linear operator on a reflexive Banach space whose iterates T^n form a bounded sequence of bounded linear operators, then its averages

$$S_n := \frac{1}{n} \sum_{k=1}^{n-1} T^k$$

form a sequence of bounded linear operators that converge strongly to a projection onto the kernel of the operator $\mathbb{1} - T$. Here is the relevant definition.

Definition 3.6.6 (Projection). Let X be a real normed vector space. A bounded linear operator $P : X \to X$ is called a **projection** if

$$P^2 = P.$$

Lemma 3.6.7. *Let X be a real normed vector space and let $P : X \to X$ be a bounded linear operator. Then the following are equivalent.*

(i) *P is a projection.*

(ii) *There exist closed linear subspaces $X_0, X_1 \subset X$ such that*

$$X_0 \cap X_1 = \{0\}, \qquad X_0 \oplus X_1 = X,$$

and

$$P(x_0 + x_1) = x_1$$

for all $x_0 \in X_0$ and all $x_1 \in X_1$.

Proof. If P is a projection, then $P^2 = P$ and hence the linear subspaces

$$X_0 := \ker(P), \qquad X_1 := \mathrm{im}(P) = \ker(\mathbb{1} - P)$$

satisfy the requirements of part (ii). If P is as in (ii), then $P^2 = P$ by definition and $P : X \to X$ is a bounded linear operator by Corollary 2.2.9. This proves Lemma 3.6.7. □

Example 3.6.8. The direct sum of two closed linear subspaces of a Banach space need not be closed. For example, let $X := C([0,1],\mathbb{R})$ be the Banach space of continuous functions $f : [0,1] \to \mathbb{R}$, equipped with the supremum norm. Then the linear subspaces

$$Y := \{(f,g) \in X \times X \mid f = 0\},$$
$$Z := \{(f,g) \in X \times X \mid f \in C^1([0,1]),\ f' = g\}$$

of $X \times X$ are closed, their intersection $Y \cap Z$ is trivial, and their direct sum $Y \oplus Z = \{(f,g) \in X \times X \mid f \in C^1([0,1])\}$ is not closed.

Theorem 3.6.9 (**Ergodic Theorem**). *Let $T : X \to X$ be a bounded linear operator on a Banach space X. Assume that there is a constant $c \geq 1$ such that*

$$\|T^n\| \leq c \qquad \text{for all } n \in \mathbb{N}. \tag{3.6.7}$$

For $n \in \mathbb{N}$ define the bounded linear operator $S_n : X \to X$ by

$$S_n := \frac{1}{n}\sum_{k=0}^{n-1} T^k. \tag{3.6.8}$$

Then the following hold.

(i) *Let $x \in X$. Then the sequence $(S_n x)_{n\in\mathbb{N}}$ converges if and only if it has a weakly convergent subsequence.*

(ii) *The set*

$$Z := \left\{x \in X \,\middle|\, \text{the sequence } (S_n x)_{n\in\mathbb{N}} \text{ converges}\right\} \tag{3.6.9}$$

is a closed T-invariant linear subspace of X and

$$Z = \ker(\mathbb{1} - T) \oplus \overline{\operatorname{im}(\mathbb{1} - T)}. \tag{3.6.10}$$

Moreover, if X is reflexive, then $Z = X$.

(iii) *Define the bounded linear operator*

$$S : Z \to Z$$

by

$$S(x + y) := x \qquad \text{for } x \in \ker(\mathbb{1} - T) \text{ and } y \in \overline{\operatorname{im}(\mathbb{1} - T)}. \tag{3.6.11}$$

Then

$$\lim_{n\to\infty} S_n z = Sz \tag{3.6.12}$$

for all $z \in Z$ and

$$ST = TS = S^2 = S, \qquad \|S\| \leq c. \tag{3.6.13}$$

Proof. See page 150. □

Theorem 3.6.9 implies Theorem 3.6.5. Let $\phi : M \to M$ be a homeomorphism of a compact metric space M and let $\mu \in \mathcal{M}(\phi)$ be an ergodic ϕ-invariant Borel probability measure on M. Define the bounded linear operator $T : L^p(\mu) \to L^p(\mu)$ by

$$Tf := f \circ \phi \qquad \text{for } f \in L^p(\mu).$$

Then $\|Tf\|_p = \|f\|_p$ for all $f \in L^p(\mu)$, by the ϕ-invariance of μ, and so

$$\|T\| = 1.$$

Thus T satisfies the requirement of Theorem 3.6.9. Let $f \in L^p(\mu)$. Since $L^p(\mu)$ is reflexive (Example 1.3.3), Theorem 3.6.9 asserts that the sequence

$$S_n f := \frac{1}{n}\sum_{k=0}^{n-1} T^k f = \frac{1}{n}\sum_{k=0}^{n-1} f \circ \phi^k$$

converges in $L^p(\mu)$ to a function $Sf \in \ker(\mathbb{1} - T)$. It remains to prove that Sf is equal to the constant $c := \int_M f\,d\mu$ almost everywhere. The key to the proof is the fact that every function in the kernel of the operator $\mathbb{1} - T$ is constant (almost everywhere). Once this is understood, it follows that there exists a constant $c \in \mathbb{R}$ such that $Sf = c$ almost everywhere, and hence

$$c = \int_M Sf\,d\mu = \lim_{n\to\infty}\int_M S_n f\,d\mu = \lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}\int_M (f\circ\phi^k)\,d\mu = \int_M f\,d\mu.$$

Thus it remains to prove that every function in the kernel of $\mathbb{1} - T$ is constant. Let $g \in L^p(\mu)$ and suppose that $Tg = g$. Choose a representative of the equivalence class of g, still denoted by $g \in \mathcal{L}^p(\mu)$. Then $g(x) = g(\phi(x))$ for almost all $x \in M$. Define

$$E_0 := \{x \in M \,|\, g(x) \neq g(\phi(x))\}, \qquad E := \bigcup_{k\in\mathbb{Z}} \phi^k(E_0).$$

Then $E \subset M$ is a Borel set with $\phi(E) = E$, $\mu(E) = 0$, and $g(\phi(x)) = g(x)$ for every $x \in M \setminus E$. Let $c := \int_M g\,d\mu$ and define $B_-, B_0, B_+ \subset M$ by

$$B_0 := \{x \in M\setminus E \,|\, g(x) = c\}, \qquad B_\pm := \{x \in M \setminus E \,|\, \pm g(x) > c\}.$$

Each of these three Borel sets is invariant under ϕ and hence has measure either zero or one. Moreover, $B_- \cup B_0 \cup B_+ = M \setminus E$ and this implies

$$\mu(B_-) + \mu(B_0) + \mu(B_+) = 1.$$

Hence one of the three sets has measure one and the other two have measure zero. This implies that $\mu(B_0) = 1$, because otherwise either $\int_M g\,d\mu < c$ or $\int_M g\,d\mu > c$. Thus g is equal to its mean value almost everywhere. This proves Theorem 3.6.5. □

Proof of Theorem 3.6.9. The proof has eight steps.

Step 1. *Let $n \in \mathbb{N}$. Then $\|S_n\| \le c$ and $\|S_n(\mathbb{1} - T)\| \le \frac{1+c}{n}$.*

By assumption, we have $\|S_n\| \le \frac{1}{n}\sum_{k=0}^{n-1}\|T^k\| \le c$ for all $n \in \mathbb{N}$. Moreover,

$$S_n(\mathbb{1} - T) = \frac{1}{n}\sum_{k=0}^{n-1} T^k - \frac{1}{n}\sum_{k=1}^{n} T^k = \frac{1}{n}(\mathbb{1} - T^n)$$

and so

$$\|S_n(\mathbb{1} - T)\| \le \frac{1}{n}\big(\|\mathbb{1}\| + \|T^n\|\big) \le \frac{1+c}{n}$$

for all $n \in \mathbb{N}$. This proves Step 1.

Step 2. *Let $x \in X$ such that $Tx = x$. Then $S_n x = x$ for all $n \in \mathbb{N}$ and*

$$\|x\| \le c\,\|x + \xi - T\xi\| \qquad \textit{for all } \xi \in X.$$

Since $Tx = x$ it follows by induction that $T^k x = x$ for all $k \in \mathbb{N}$ and hence

$$x = \frac{1}{n}\sum_{k=0}^{n-1} T^k x = S_n x \qquad \text{for all } n \in \mathbb{N}.$$

Moreover, $\lim_{n\to\infty}\|S_n(\xi - T\xi)\| = 0$ by Step 1 and hence

$$\|x\| = \lim_{n\to\infty}\|x + S_n(\xi - T\xi)\| = \lim_{n\to\infty}\|S_n(x + \xi - T\xi)\| \le c\,\|x + \xi - T\xi\|.$$

Here the inequality holds because $\|S_n\| \le c$ by Step 1. This proves Step 2.

Step 3. *If $x \in \ker(\mathbb{1} - T)$ and $y \in \overline{\mathrm{im}(\mathbb{1} - T)}$, then $\|x\| \le c\,\|x + y\|$.*

Choose a sequence $\xi_n \in X$ such that $y = \lim_{n\to\infty}(\xi_n - T\xi_n)$. Then, by Step 2, we have $\|x\| \le c\|x + \xi_n - T\xi_n\|$ for all $n \in \mathbb{N}$. Take the limit $n \to \infty$ to obtain $\|x\| \le c\|x + y\|$. This proves Step 3.

Step 4. $\ker(\mathbb{1} - T) \cap \overline{\mathrm{im}(\mathbb{1} - K)} = \{0\}$ *and the direct sum*

$$Z := \ker(\mathbb{1} - T) \oplus \overline{\mathrm{im}(\mathbb{1} - T)} \tag{3.6.14}$$

is a closed linear subspace of X.

Let $x \in \ker(\mathbb{1}-T)\cap\overline{\mathrm{im}(\mathbb{1} - T)}$ and define $y := -x$. Then $\|x\| \le c\|x + y\| = 0$ by Step 3 and hence $x = 0$. This shows that $\ker(\mathbb{1} - T) \cap \overline{\mathrm{im}(\mathbb{1} - T)} = \{0\}$. We prove that the subspace Z in (3.6.14) is closed. Let $x_n \in \ker(\mathbb{1} - T)$ and $y_n \in \overline{\mathrm{im}(\mathbb{1} - T)}$ be sequences whose sum $z_n := x_n + y_n$ converges to some element $z \in X$. Then $(z_n)_{n\in\mathbb{N}}$ is a Cauchy sequence and hence $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence by Step 3. This implies that $y_n = z_n - x_n$ is a Cauchy sequence and hence $z = x + y$, where $x := \lim_{n\to\infty} x_n \in \ker(\mathbb{1} - T)$ and $y := \lim_{n\to\infty} y_n \in \overline{\mathrm{im}(\mathbb{1} - T)}$. This proves Step 4.

Step 5. *If $z \in Z$, then $Tz \in Z$.*

Let $z \in Z$. Then

$$z = x + y, \qquad x \in \ker(\mathbb{1} - T), \qquad y \in \overline{\mathrm{im}(\mathbb{1} - T)}.$$

Choose a sequence $\xi_i \in X$ such that $y = \lim_{i\to\infty}(\xi_i - T\xi_i)$. Then

$$Ty = \lim_{i\to\infty} T(\xi_i - T\xi_i) = \lim_{i\to\infty} (\mathbb{1} - T)T\xi_i \in \overline{\mathrm{im}(\mathbb{1} - T)}.$$

Hence

$$Tz = Tx + Ty = x + Ty \in Z$$

and this proves Step 5.

Step 6. *Let $x \in \ker(\mathbb{1} - T)$ and $y \in \overline{\mathrm{im}(\mathbb{1} - T)}$. Then*

$$x = \lim_{n\to\infty} S_n(x + y).$$

By Step 1, the sequence

$$\|S_n(\mathbb{1} - T)\xi\| \leq \frac{1 + c}{n} \|\xi\|$$

converges to zero as n tends to infinity for every $\xi \in X$. Hence it follows from the estimate $\|S_n\| \leq c$ in Step 1 and the Banach–Steinhaus Theorem 2.1.5 that

$$\lim_{n\to\infty} S_n y = 0 \qquad \text{for all } y \in \overline{\mathrm{im}(\mathbb{1} - T)}.$$

Moreover,

$$S_n x = x \qquad \text{for all } n \in \mathbb{N}$$

by Step 2. Hence

$$x = \lim_{n\to\infty} S_n x = \lim_{n\to\infty} S_n(x + y).$$

This proves Step 6.

Step 7. *Let $x, z \in X$. Then the following are equivalent.*

(a) *$Tx = x$ and $z - x \in \overline{\mathrm{im}(\mathbb{1} - T)}$.*

(b) $\lim_{n\to\infty} \|S_n z - x\| = 0$.

(c) *There is a sequence of integers $1 \leq n_1 < n_2 < n_3 < \cdots$ such that*

$$\lim_{i\to\infty} \langle x^*, S_{n_i} z\rangle = \langle x^*, x\rangle \qquad \textit{for all } x^* \in X^*.$$

That (a) implies (b) follows immediately from Step 6 and that (b) implies (c) is obvious. We prove that (c) implies (a). Thus assume (c) and fix a bounded linear functional $x^* \in X^*$. Then

$$T^*x^* := x^* \circ T : X \to \mathbb{R}$$

is a bounded linear functional and

$$\begin{aligned}\langle x^*, x - Tx\rangle &= \langle x^* - T^*x^*, x\rangle \\ &= \lim_{i\to\infty} \langle x^* - T^*x^*, S_{n_i}z\rangle \\ &= \lim_{i\to\infty} \langle x^*, (\mathbb{1} - T)S_{n_i}z\rangle \\ &= 0.\end{aligned}$$

Here the last equation follows from Step 1. Hence

$$Tx = x$$

by the Hahn–Banach Theorem (Corollary 2.3.23). Next we prove that

$$z - x \in \overline{\mathrm{im}(\mathbb{1} - T)}.$$

Assume, by contradiction, that $z - x \in X \setminus \overline{\mathrm{im}(\mathbb{1} - T)}$. Then, by the Hahn–Banach Theorem 2.3.22, there exists an element $x^* \in X^*$ such that

$$\langle x^*, z - x\rangle = 1, \qquad \langle x^*, \xi - T\xi\rangle = 0 \quad \text{for all } \xi \in X. \tag{3.6.15}$$

This implies $\langle x^*, T^k\xi - T^{k+1}\xi\rangle = 0$ for all $k \in \mathbb{N}$ and all $\xi \in X$. Hence, by induction, $\langle x^*, \xi\rangle = \langle x^*, T^k\xi\rangle$ for every $\xi \in X$ and every integer $k \geq 0$. Thus

$$\langle x^*, S_n z\rangle = \frac{1}{n}\sum_{k=0}^{n-1}\langle x^*, T^k z\rangle = \langle x^*, z\rangle$$

for all $n \in \mathbb{N}$. Hence it follows from (c) that

$$\langle x^*, z - x\rangle = \lim_{i\to\infty}\langle x^*, S_{n_i}z - x\rangle = 0.$$

This contradicts (3.6.15). Thus $z - x \in \overline{\mathrm{im}(\mathbb{1} - T)}$ and this proves Step 7.

Step 8. *We prove Theorem* 3.6.9.

The subspace Z in (3.6.14) is closed by Step 4 and is T-invariant by Step 5. Moreover, Step 7 asserts that an element $z \in X$ belongs to Z if and only if the sequence $(S_n z)_{n\in\mathbb{N}}$ converges in the norm topology if and only if $(S_n z)_{n\in\mathbb{N}}$ has a weakly convergent subsequence. If X is reflexive, this holds for all $z \in X$ by Step 1 and Theorem 2.4.4. This proves (i) and (ii).

Define the operator $S : Z \to Z$ by (3.6.11). Then $\|S\| \leq c$ by Step 3, the equation $\lim_{n\to\infty} S_n z = Sz$ for $z \in Z$ follows from Step 6, and $S^2 = S$ by definition. The equation $ST = TS = S$ follows from the fact that S commutes with $T|_Z$ and vanishes on the image of the operator $\mathbb{1} - T$. This proves Theorem 3.6.9. □

6.3. Unbounded Operators on Hilbert Spaces

The dual operator of an unbounded operator between Banach spaces was introduced in Definition 6.2.1. For Hilbert spaces this leads to the notion of the adjoint of an unbounded densely defined operator which we explain next. As in Example 4.1.6 and Definition 5.3.7, the idea is to replace the dual space of a Hilbert space by the original Hilbert space via the isomorphism of Theorem 1.4.4, respectively Theorem 5.3.6 in the complex case.

6.3.1. The Adjoint of an Unbounded Operator.

Definition 6.3.1 (**Adjoint Operator**)**.** Let X and Y be complex Hilbert spaces and let $A : \mathrm{dom}(A) \to Y$ be an unbounded operator with a dense domain $\mathrm{dom}(A) \subset X$. The **adjoint operator**

$$A^* : \mathrm{dom}(A^*) \to X, \qquad \mathrm{dom}(A^*) \subset Y,$$

of A is defined as follows. Its domain is the linear subspace

$$\mathrm{dom}(A^*) := \left\{ y \in Y \;\middle|\; \begin{array}{l} \text{there exists a constant } c \geq 0 \text{ such that} \\ |\langle y, A\xi\rangle_Y| \leq c\,\|\xi\|_X \text{ for all } \xi \in \mathrm{dom}(A) \end{array} \right\}$$

and, for $y \in \mathrm{dom}(A^*)$, the element $A^*y \in X$ is the unique element of X that satisfies the equation

$$\langle A^*y, \xi\rangle_X = \langle y, A\xi\rangle_Y \qquad \text{for all } \xi \in \mathrm{dom}(A).$$

Thus the graph of the adjoint operator is characterized by the condition

$$\begin{array}{l} y \in \mathrm{dom}(A^*) \\ \text{and } x = A^*y \end{array} \quad \Longleftrightarrow \quad \begin{array}{l} \langle A^*y, \xi\rangle_X = \langle y, A\xi\rangle_Y \\ \text{for all } \xi \in \mathrm{dom}(A). \end{array} \tag{6.3.1}$$

The operator A is called **self-adjoint** if $X = Y$ and $A = A^*$.

Observe that an element $y \in Y$ belongs to the domain of A^* if and only if the complex linear functional $\mathrm{dom}(A) \to \mathbb{C} : \xi \mapsto \langle y, A\xi\rangle$ is bounded. In this case, the linear functional extends uniquely to a bounded complex linear functional on all of X, because $\mathrm{dom}(A)$ is a dense subspace of X, and Theorem 5.3.6 asserts that this extended complex linear functional is uniquely represented by an element of X. The reader may verify that $\mathrm{dom}(A^*)$ is a complex subspace of Y and that the operator $A^* : \mathrm{dom}(A^*) \to X$ is complex linear. Throughout the remainder of this chapter the symbol A^* will always denote the adjoint of an unbounded operator between Hilbert spaces as in Definition 6.3.1. The dual operator of Definition 6.2.1 is no longer used.

The next lemma summarizes the basic properties of the adjoint operator. Recall that in the Hilbert space setting the notation

$$S^\perp := \{y \in H \mid \langle x, y\rangle = 0 \text{ for all } x \in S\}$$

refers to the (complex) orthogonal complement of a subset $S \subset H$.

Lemma 6.3.2 (Properties of the Adjoint Operator). *Let X and Y be complex Hilbert spaces and let $A : \mathrm{dom}(A) \to Y$ be a linear operator with a dense domain $\mathrm{dom}(A) \subset X$. Then the following hold.*

(i) *If $P \in \mathcal{L}^c(X, Y)$ and $\lambda \in \mathbb{C}$, then $(A+P)^* = A^* + P^*$ and $(\lambda A)^* = \overline{\lambda} A^*$.*

(ii) *A is closeable if and only if $\mathrm{dom}(A^*)$ is a dense subspace of Y.*

(iii) *If A is closed, then $A^{**} = A$.*

(iv) *$\mathrm{im}(A)^\perp = \ker(A^*)$ and, if A is closed, then $\ker(A) = \mathrm{im}(A^*)^\perp$.*

(v) *A has a dense image if and only if A^* is injective.*

(vi) *Assume A is closed. Then A has a closed image if and only if A^* has a closed image if and only if $\mathrm{im}(A^*) = \ker(A)^\perp$.*

(vii) *If A is bijective, then so is A^* and $(A^{-1})^* = (A^*)^{-1}$.*

(viii) *If $X = Y = H$ and A is closed, then $\sigma(A^*) = \left\{\overline{\lambda} \,|\, \lambda \in \sigma(A)\right\}$ and*

$$\begin{aligned}
\mathrm{P}\sigma(A^*) &\subset \left\{\overline{\lambda} \,|\, \lambda \in \mathrm{P}\sigma(A) \cup \mathrm{R}\sigma(A)\right\}, \\
\mathrm{R}\sigma(A^*) &\subset \left\{\overline{\lambda} \,|\, \lambda \in \mathrm{P}\sigma(A)\right\}, \\
\mathrm{C}\sigma(A^*) &= \left\{\overline{\lambda} \,|\, \lambda \in \mathrm{C}\sigma(A)\right\}.
\end{aligned}$$

Proof. These assertions are proved by carrying over Theorem 6.2.2, Theorem 6.2.3, Corollary 6.2.4, and Lemma 6.2.6 to the Hilbert space setting. The details are left to the reader. □

6.3.2. Unbounded Self-Adjoint Operators. By definition, every self-adjoint operator on a Hilbert space $H = X = Y$ is symmetric, i.e. it satisfies

$$\langle x, Ay \rangle = \langle Ax, y \rangle \qquad \text{for all } x, y \in \mathrm{dom}(A).$$

However, the converse does not hold, even for operators with dense domains and closed graphs. (By Example 2.2.23 every symmetric operator is closeable.) Exercise 6.3.3 below illustrates the difference between symmetric and self-adjoint operators and shows how one can construct self-adjoint extensions of symmetric operators.

A skew-symmetric bilinear form

$$\omega : V \times V \to \mathbb{R}$$

on a real vector space V is called **symplectic** if it is nondegenerate, i.e. for every nonzero vector $v \in V$ there exists a vector $u \in V$ such that $\omega(u, v) \neq 0$. Assume $\omega : V \times V \to \mathbb{R}$ is a symplectic form. A linear subspace $\Lambda \subset V$ is called a **Lagrangian subspace** if $\omega(u, v) = 0$ for all $u, v \in \Lambda$ and if, for every $v \in V \setminus \Lambda$, there exists a vector $u \in \Lambda$ such that $\omega(u, v) \neq 0$.

Exercise 6.3.3 (Gelfand–Robbin Quotient). Let H be a real Hilbert space and let $A : \mathrm{dom}(A) \to H$ be a densely defined symmetric operator.

(i) Prove that $\mathrm{dom}(A) \subset \mathrm{dom}(A^*)$ and $A^*|_{\mathrm{dom}(A)} = A$.

(ii) Let $V := \mathrm{dom}(A^*)/\mathrm{dom}(A)$ and define the map $\omega : V \times V \to \mathbb{R}$ by

$$\omega(u, v) := \langle A^*x, y\rangle - \langle x, A^*y\rangle \tag{6.3.2}$$

for $x, y \in \mathrm{dom}(A^*)$, where $u := [x] \in V$ and $v := [y] \in V$. Prove that ω is a well-defined skew-symmetric bilinear form. Prove that ω is nondegenerate if and only if the operator A has a closed graph.

(iii) Assume A has a closed graph. For a subspace $\Lambda \subset V$ define the operator

$$A_\Lambda : \mathrm{dom}(A_\Lambda) \to H$$

by

$$\mathrm{dom}(A_\Lambda) := \{x \in \mathrm{dom}(A^*) \,|\, [x] \in \Lambda\}, \qquad A_\Lambda := A^*|_{\mathrm{dom}(A_\Lambda)}. \tag{6.3.3}$$

Prove that A_Λ is self-adjoint if and only if Λ is a Lagrangian subspace of V.

(iv) Prove that A admits a self-adjoint extension. **Hint:** The Lemma of Zorn.

(v) Prove that $\Lambda_0 := (\ker(A^*) + \mathrm{dom}(A))/\mathrm{dom}(A)$ is a Lagrangian subspace of V whenever A has a closed graph and a closed image.

Exercise 6.3.4. This example illustrates how the Gelfand–Robbin quotient gives rise to symplectic forms on the spaces of boundary data for symmetric differential operators. Let $n \in \mathbb{N}$ and consider the matrix

$$J := \begin{pmatrix} 0 & -\mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} \in \mathbb{R}^{2n\times 2n}.$$

Define the operator A on the Hilbert space $H := L^2([0,1], \mathbb{R}^{2n})$ by

$$\mathrm{dom}(A) := \left\{u \in W^{1,2}([0,1], \mathbb{R}^{2n}) \,|\, u(0) = u(1) = 0\right\}, \qquad Au := J\dot{u}.$$

Here $W^{1,2}([0,1], \mathbb{R}^{2n})$ denotes the space of all absolutely continuous functions $u : [0,1] \to \mathbb{R}^{2n}$ with square integrable derivatives. Prove the following.

(i) A is a symmetric operator with a closed graph.

(ii) $\mathrm{dom}(A^*) = W^{1,2}([0,1], \mathbb{R}^{2n})$ and $A^*u = J\dot{u}$ for all $u \in W^{1,2}([0,1], \mathbb{R}^{2n})$.

(iii) The map $W^{1,2}([0,1], \mathbb{R}^{2n}) \to \mathbb{R}^{2n} \times \mathbb{R}^{2n} : u \mapsto (u(0), u(1))$ descends to an isomorphism from the quotient space $V = \mathrm{dom}(A^*)/\mathrm{dom}(A)$ to $\mathbb{R}^{2n}\times\mathbb{R}^{2n}$. The resulting symplectic form determined by (6.3.2) on $\mathbb{R}^{2n} \times \mathbb{R}^{2n}$ is

$$\omega((u_0, u_1), (v_0, v_1)) = \langle Ju_1, v_1\rangle - \langle Ju_0, v_0\rangle$$

for $(u_0, u_1), (v_0, v_1) \in \mathbb{R}^{2n} \times \mathbb{R}^{2n}$. Here $\langle\cdot,\cdot\rangle$ denotes the standard inner product on $\mathbb{R}^{2n}$.

Exercise 6.3.5. Let H be a separable complex Hilbert space, let $(e_i)_{i\in\mathbb{N}}$ be a complex orthonormal basis, let $(\lambda_i)_{i\in\mathbb{N}}$ be a sequence of complex numbers, and let $A_\lambda : \mathrm{dom}(A_\lambda) \to H$ be the operator in Example 6.1.3. Prove that its adjoint is the operator $A_\lambda^* = A_{\overline{\lambda}}$ associated to the sequence $(\overline{\lambda}_i)_{i\in\mathbb{N}}$. Deduce that A_λ is self-adjoint if and only if $\lambda_i \in \mathbb{R}$ for all i.

Exercise 6.3.6. Prove that the operator A_f in Example 6.1.8 is self-adjoint for $p = 2$ and every measurable function $f : M \to \mathbb{R}$.

Another example of an unbounded self-adjoint operator is the Laplace operator on $\Delta : W^{2,2}(\mathbb{R}^n, \mathbb{C}) \to L^2(\mathbb{R}^n, \mathbb{C})$ in Example 6.1.6. The proof that this operator is self-adjoint requires elliptic regularity and goes beyond the scope of this book. However, this example can be recast as a special case of a general abstract setup, which is useful for many applications and which we now explain.

Definition 6.3.7 (Gelfand Triple). A **Gelfand triple** consists of a real Hilbert space $(H, \langle\cdot,\cdot\rangle_H)$ and a dense subspace $V \subset H$, equipped with an inner product $\langle\cdot,\cdot\rangle_V$ which renders V into a Hilbert space in its own right and the inclusion $V \hookrightarrow H$ into a bounded linear operator. Thus there exists a constant $\kappa > 0$ such that

$$\|v\|_H \le \kappa \|v\|_V \qquad \text{for all } v \in V. \tag{6.3.4}$$

We identify H with its dual space H^* via the isomorphism of Theorem 1.4.4. However, we do not identify V with its own dual space. Thus

$$V \subset H \subset V^*, \tag{6.3.5}$$

where the inclusion $H \cong H^* \hookrightarrow V^*$ assigns to each $u \in H$ the bounded linear functional $V \to \mathbb{R} : v \mapsto \langle u, v\rangle_H$. This is the dual operator of the inclusion $V \hookrightarrow H$ and so is injective and has a dense image by Theorem 4.1.8.

Theorem 6.3.8 (Gelfand Triples). *Let $V \subset H \subset V^*$ be a Gelfand triple and let $B : V \times V \to \mathbb{R}$ be a symmetric bilinear form. Suppose that there exist positive constants δ, c, and C such that*

$$\delta \|v\|_V^2 - c\|v\|_H^2 \le B(v,v) \le C\|v\|_V^2 \qquad \textit{for all } v \in V. \tag{6.3.6}$$

Then the linear subspace

$$\mathrm{dom}(A) := \left\{ u \in V \,\middle|\, \sup_{v\in V\setminus\{0\}} \frac{|B(u,v)|}{\|v\|_H} < \infty \right\} \tag{6.3.7}$$

is dense in V, there is a unique linear operator $A : \mathrm{dom}(A) \to H$ such that

$$\langle Au, v\rangle_H = B(u,v) \qquad \textit{for all } u \in \mathrm{dom}(A) \textit{ and all } v \in V, \tag{6.3.8}$$

and this operator A is self-adjoint. If H is a complex Hilbert space and V is a complex subspace of H such that the complex structure preserves the inner product on V and the bilinear form B, then A is complex linear.

Proof. The existence and uniqueness of an operator $A : \mathrm{dom}(A) \to H$ that satisfies (6.3.7) and (6.3.8) follow directly from the definitions and Theorem 1.4.4. Namely, if $u \in \mathrm{dom}(A)$, then, since V is dense in H, there exists a unique bounded linear functional $\Lambda_u : H \to \mathbb{R}$ such that

$$\Lambda_u(v) = B(u, v) \qquad \text{for all } v \in V,$$

and so, by Theorem 1.4.4, there exists a unique element $Au \in H$ such that

$$\langle Au, f\rangle_H = \Lambda_u(f) \qquad \text{for all } f \in H.$$

Then $A : \mathrm{dom}(A) \to H$ is a symmetric linear operator that satisfies (6.3.8). We prove in seven steps that A is self-adjoint.

Step 1. *If $u, v \in V$, then $|B(u, v)| \le C \|u\|_V \|v\|_V$.*

By Theorem 1.4.4 there exists a unique linear operator $\mathscr{B} : V \to V$ such that $\langle u, \mathscr{B}v\rangle_V = B(u, v)$ for all $u, v \in V$. Since B is symmetric, so is $\mathscr{B}$. Hence $\mathscr{B}$ is bounded by the Hellinger–Toeplitz Theorem (Corollary 2.2.16). Moreover, $|\langle v, \mathscr{B}v\rangle_V| = |B(v, v)| \le C$ for all $v \in V$ with $\|v\|_V = 1$ by (6.3.6). Hence $\|\mathscr{B}\|_{\mathcal{L}(V)} \le C$ by part (iv) of Theorem 5.3.16 and so

$$|B(u, v)| = |\langle u, \mathscr{B}v\rangle_V| \le \|u\|_V \|\mathscr{B}v\|_V \le C \|u\|_V \|v\|_V$$

for all $u, v \in V$. This proves Step 1.

Step 2. *If $u \in \mathrm{dom}(A)$, then $\|u\|_V \le \delta^{-1}\kappa \|cu + Au\|_H$.*

By (6.3.4), (6.3.6), and (6.3.8), every $u \in \mathrm{dom}(A)$ satisfies

$$\begin{aligned} \delta \|u\|_V^2 &\le c \|u\|_H^2 + B(u, u) \\ &= \langle cu + Au, u\rangle_H \\ &\le \|cu + Au\|_H \|u\|_H \\ &\le \kappa \|cu + Au\|_H \|u\|_V \end{aligned}$$

and this proves Step 2. (**Exercise:** Use Step 2 to show that A is closed.)

Step 3. *The formula*

$$\langle u, v\rangle_B := c \langle u, v\rangle_H + B(u, v) \qquad \textit{for } u, v \in V \tag{6.3.9}$$

defines an inner product on V whose norm $V \to \mathbb{R} : v \mapsto \|v\|_B := \sqrt{\langle v, v\rangle_B}$ is compatible with $\|\cdot\|_V$. Thus $(V, \langle\cdot,\cdot\rangle_B)$ is a Hilbert space.

The bilinear form (6.3.9) is symmetric because B is symmetric and satisfies the inequality $\delta \|v\|_V^2 \le B(v, v) \le c \|v\|_H^2 + C \|v\|_V^2 \le \big(c\kappa^2 + C\big) \|v\|_V^2$ for all $v \in V$ by (6.3.4) and (6.3.6). This proves Step 3.

The next step is the heart of the proof. It can be viewed as an abstract variant of the Dirichlet principle.

Step 4. *The operator* $c1\!\!1_H + A : \mathrm{dom}(A) \to H$ *is bijective.*

The operator is injective by Step 2. To prove that it is surjective, fix an element $f \in H$ and define the bounded linear functional $\Lambda : V \to \mathbb{R}$ by

$$\Lambda(v) := \langle f, v \rangle_H \qquad \text{for } v \in V.$$

Then, by Step 3 and Theorem 1.4.4, there exists an element $u \in V$ that satisfies $\langle u, v \rangle_B = \Lambda(v)$ for all $v \in V$. This implies

$$c \langle u, v \rangle_H + B(u, v) = \langle f, v \rangle_H$$

for all $v \in V$ and hence

$$|B(u, v)| = |\langle f - cu, v \rangle_H| \le \|f - cu\|_H \|v\|_H .$$

Thus $u \in \mathrm{dom}(A)$ and, for all $v \in V$, we have

$$\langle cu + Au - f, v \rangle_H = c \langle u, v \rangle_H + B(u, v) - \langle f, v \rangle_H = 0.$$

Since V is dense in H, it follows that $cu + Au = f$ and this proves Step 4.

Step 5. *The subspace* $\mathrm{dom}(A) \subset V$ *defined by* (6.3.7) *is dense in* V.

Let $\iota : V \to H$ denote the canonical inclusion and let $\iota^* : H \to V$ be its adjoint operator with respect to the inner products $\langle \cdot, \cdot \rangle_H$ on H and $\langle \cdot, \cdot \rangle_B$ on V (see Step 3). Then ι^* has a dense image by Theorem 4.1.8. Let $f \in H$ and define $u := (c1\!\!1_H + A)^{-1} f \in \mathrm{dom}(A)$ by Step 4. Then $cu + Au = f$ and

$$\langle \iota^*(f), v \rangle_B = \langle f, \iota(v) \rangle_H = \langle cu + Au, v \rangle_H = \langle u, v \rangle_B = \left\langle (c1\!\!1 + A)^{-1} f, v \right\rangle_B$$

for all $v \in V$. This shows that $\iota^* = (c1\!\!1_H - A)^{-1} : H \to V$ and hence the subspace $\mathrm{dom}(A) = \mathrm{im}(\iota^*)$ is dense in V. This proves Step 5.

Step 6. *Let* $v \in H$ *and suppose that there is a constant* $K \ge 0$ *such that*

$$|\langle v, Au \rangle_H| \le K \|u\|_V \qquad \textit{for all } u \in \mathrm{dom}(A). \tag{6.3.10}$$

Then $v \in V$.

By (6.3.4) and (6.3.10) we have $|\langle v, cu + Au \rangle_H| \le (c\kappa \|v\|_H + K) \|u\|_V$ for all $u \in \mathrm{dom}(A)$. Since $\mathrm{dom}(A)$ is dense in V by Step 5, this implies that there exists a unique bounded linear functional $\Lambda : V \to \mathbb{R}$ such that

$$\Lambda(u) = \langle v, cu + Au \rangle_H \qquad \text{for all } u \in \mathrm{dom}(A). \tag{6.3.11}$$

Hence, by Step 3 and Theorem 1.4.4, there exists a $w \in V$ such that

$$\Lambda(u) = \langle w, u \rangle_B = c \langle w, u \rangle_H + B(w, u) \qquad \text{for all } u \in V. \tag{6.3.12}$$

Take $u \in \mathrm{dom}(A)$ in (6.3.12) to obtain $\Lambda(u) = \langle w, cu + Au \rangle_H$. Hence it follows from (6.3.11) that $\langle v - w, cu + Au \rangle_H = 0$ for all $u \in \mathrm{dom}(A)$. Since the operator $c1\!\!1_H + A : \mathrm{dom}(A) \to H$ is surjective by Step 4, we have $v = w \in V$. This proves Step 6.

Step 7. *The operator* $A : \mathrm{dom}(A) \to H$ *is self-adjoint.*

The operator A is symmetric by definition. Hence $\mathrm{dom}(A) \subset \mathrm{dom}(A^*)$ and $A^*|_{\mathrm{dom}(A)} = A$. It remains to prove that $\mathrm{dom}(A^*) \subset \mathrm{dom}(A)$. To see this, fix an element $v \in \mathrm{dom}(A^*)$. Then

$$|\langle v, Au\rangle_H| = |\langle A^*v, u\rangle_H| \le \|A^*v\|_H \|u\|_H \le \kappa \|A^*v\|_H \|u\|_V$$

for all $u \in \mathrm{dom}(A)$ by (6.3.4). Hence $v \in V$ by Step 6. This implies

$$|B(v, u)| = |\langle v, Au\rangle_H| = |\langle A^*v, u\rangle_H| \le \|A^*v\|_H \|u\|_H$$

for all $u \in \mathrm{dom}(A)$. Since $\mathrm{dom}(A)$ is dense in V by Step 5, and the functions $V \to \mathbb{R} : u \mapsto \|u\|_H$ and $V \to \mathbb{R} : u \mapsto B(v, u)$ are continuous by (6.3.4) and Step 1, this implies $|B(v, u)| \le \|A^*v\|_H \|u\|_H$ for all $u \in V$ and therefore $v \in \mathrm{dom}(A)$. This proves Step 7 and Theorem 6.3.8. □

The next corollary explains how every closed densely defined unbounded operator gives rise to a self-adjoint operator by composition with its adjoint. The composition of two unbounded linear operators $A : \mathrm{dom}(A) \to Y$ with $\mathrm{dom}(A) \subset X$ and $B : \mathrm{dom}(B) \to Z$ with $\mathrm{dom}(B) \subset Y$ is the operator $BA : \mathrm{dom}(BA) \to Z$ defined by

$$\begin{aligned} \mathrm{dom}(BA) &:= \left\{x \in \mathrm{dom}(A) \,\middle|\, Ax \in \mathrm{dom}(B)\right\}, \\ BAx &:= B(Ax) \qquad \text{for } x \in \mathrm{dom}(BA). \end{aligned} \tag{6.3.13}$$

The domain of BA can be trivial even if A and B are densely defined. In the next theorem X and Y can either be real Hilbert spaces or complex Hilbert spaces with Hermitian inner products. In the latter case we assume that D is an unbounded complex linear operator and so D^*D is also complex linear.

Corollary 6.3.9 (**The Operator** D^*D). *Let X and Y be Hilbert spaces and let $D : \mathrm{dom}(D) \to Y$ be a closed unbounded operator with a dense domain* $\mathrm{dom}(D) \subset X$. *Then the operator $D^*D : \mathrm{dom}(D^*D) \to X$ is self-adjoint and its domain is dense in* $\mathrm{dom}(D)$ *with respect to the graph norm.*

Proof. This is a Gelfand triple with

$$H := X, \qquad V := \mathrm{dom}(D), \qquad \langle u, v\rangle_V := \langle u, v\rangle_X + \langle Du, Dv\rangle_Y$$

for $u, v \in \mathrm{dom}(D)$, and the bilinear form $B : V \times V \to \mathbb{R}$ is given by

$$B(u, v) := \langle Du, Dv\rangle_Y \qquad \text{for } u, v \in \mathrm{dom}(D) \subset X.$$

These data satisfy the hypotheses of Theorem 6.3.8 with $\delta = c = C = 1$. In particular, $\|v\|_V^2 = \|v\|_X^2 + \|Dv\|_Y^2 = \|v\|_X^2 + B(v, v) = \|v\|_B^2$ for all $v \in V$. The condition $\sup_{v \in V \setminus \{0\}} \|v\|_X^{-1} |\langle Du, Dv\rangle_Y| < \infty$ for $u \in V = \mathrm{dom}(D)$ in equation (6.3.7) is equivalent to $Du \in \mathrm{dom}(D^*)$, so the operator A in (6.3.8) agrees with D^*D. Hence the operator D^*D is self-adjoint by Theorem 6.3.8. This proves Corollary 6.3.9. □

Example 6.3.10 (Dirichlet Problem). The archetypal example of the situation in Theorem 6.3.8 and Corollary 6.3.9 is the operator

$$\nabla = \left(\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n}\right) : W_0^{1,2}(\Omega) \to L^2(\Omega, \mathbb{R}^n).$$

Here $\Omega \subset \mathbb{R}^n$ is a bounded open set with smooth boundary (i.e. $\partial\Omega$ is a smooth $(n-1)$-dimensional submanifold of $\mathbb{R}^n$ and $\Omega = \mathrm{int}(\overline{\Omega})$) and $W_0^{1,2}(\Omega)$ is the completion of the space $C_0^\infty(\Omega)$ of smooth functions $u : \Omega \to \mathbb{R}$ with compact support with respect to the norm

$$\|u\|_{W_0^{1,2}} := \sqrt{\int_\Omega \sum_{i=1}^n \left|\frac{\partial u}{\partial u_i}(x)\right|^2 dx} = \sqrt{\int_\Omega |\nabla u(x)|^2 \, dx}.$$

The Poincaré inequality asserts that this norm controls the L^2 norm of u. This example corresponds to the Gelfand triple with

$$H = X = L^2(\Omega), \qquad V = \mathrm{dom}(D) = W_0^{1,2}(\Omega),$$

where the bilinear form

$$B : W_0^{1,2}(\Omega) \times W_0^{1,2}(\Omega) \to \mathbb{R}$$

is given by

$$B(u,v) := \int_\Omega \langle \nabla u(x), \nabla v(x) \rangle \, dx$$

for $u, v \in W_0^{1,2}(\Omega)$. The operator $D = \nabla : \mathrm{dom}(D) \to Y$ takes values in the Hilbert space $Y = L^2(\Omega, \mathbb{R}^n)$, and $A = D^*D$ is the **Laplace operator**

$$\Delta = \sum_{i=1}^n \frac{\partial^2}{\partial x_i^2} : W^{2,2}(\Omega) \cap W_0^{1,2}(\Omega) \to L^2(\Omega). \tag{6.3.14}$$

Here $W^{2,2}(\Omega)$ denotes the space of equivalence classes, up to equality almost everywhere, of all L^2 functions $u : \Omega \to \mathbb{R}$ whose distributional derivatives up to order two can be represented by L^2 functions. The proof that

$$\mathrm{dom}(D^*D) = W^{2,2}(\Omega) \cap W_0^{1,2}(\Omega)$$

(for all domains $\Omega \subset \mathbb{R}^n$ with *"sufficiently nice boundary"*) requires elliptic regularity and goes beyond the scope of this book. Once this is established, Corollary 6.3.9 asserts that the Laplace operator (6.3.14) is self-adjoint. Moreover, this example satisfies condition (6.3.6) with $c = 0$. Hence it follows from Step 4 in the proof of Theorem 6.3.8 that the operator (6.3.14) is bijective. This translates into the observation that the Dirichlet problem

$$\begin{aligned} \Delta u &= f \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \partial\Omega \end{aligned} \tag{6.3.15}$$

has a unique solution $u \in W^{2,2}(\Omega) \cap W_0^{1,2}(\Omega)$ for every $f \in L^2(\Omega)$.

6.3.3. Unbounded Normal Operators. The next theorem introduces unbounded normal operators on Hilbert spaces.

Theorem 6.3.11 (Unbounded Normal Operator). *Let H be a complex Hilbert space and let $A : \mathrm{dom}(A) \to H$ be a closed unbounded complex linear operator with a dense domain $\mathrm{dom}(A) \subset H$. The following are equivalent.*

(i) $AA^* = A^*A$.

(ii) $\mathrm{dom}(A) = \mathrm{dom}(A^*)$ *and* $\|Ax\| = \|A^*x\|$ *for all* $x \in \mathrm{dom}(A)$.

(iii) *There exist complex linear self-adjoint operators* $A_i : \mathrm{dom}(A_i) \to H$ *for* $i = 1, 2$ *such that* $\mathrm{dom}(A) = \mathrm{dom}(A^*) = \mathrm{dom}(A_1) \cap \mathrm{dom}(A_2)$ *and*

$$Ax = A_1x + \mathbf{i}A_2x, \quad A^*x = A_1x - \mathbf{i}A_2x, \quad \|Ax\|^2 = \|A_1x\|^2 + \|A_2x\|^2$$

for all $x \in \mathrm{dom}(A)$.

Definition 6.3.12 (Unbounded Normal Operator). A closed unbounded complex linear operator $A : \mathrm{dom}(A) \to H$ on a Hilbert space H with a dense domain $\mathrm{dom}(A) \subset H$ is called **normal** if it satisfies the equivalent conditions of Theorem 6.3.11.

Proof. We prove that (i) implies (ii). Assume $AA^* = A^*A$. Then every element $x \in \mathrm{dom}(A^*A) = \mathrm{dom}(AA^*)$ satisfies $x \in \mathrm{dom}(A) \cap \mathrm{dom}(A^*)$ as well as $Ax \in \mathrm{dom}(A^*)$ and $A^*x \in \mathrm{dom}(A)$, and hence

$$\|Ax\|^2 = \langle Ax, Ax\rangle = \langle x, A^*Ax\rangle = \langle x, AA^*x\rangle = \langle A^*x, A^*x\rangle = \|A^*x\|^2 .$$

Next we prove that $\mathrm{dom}(A) \subset \mathrm{dom}(A^*)$. Let $x \in \mathrm{dom}(A)$. Then Corollary 6.3.9 asserts that there exists a sequence $x_i \in \mathrm{dom}(A^*A)$ such that

$$\lim_{i\to\infty} \|x - x_i\| = 0, \qquad \lim_{i\to\infty} \|Ax - Ax_i\| = 0.$$

Thus $(Ax_i)_{i\in\mathbb{N}}$ is a Cauchy sequence in H and so is the sequence $(A^*x_i)_{i\in\mathbb{N}}$ because $\|A^*x_i - A^*x_j\| = \|Ax_i - Ax_j\|$ for all $i, j \in \mathbb{N}$ by what we already proved. Hence $(A^*x_i)_{i\in\mathbb{N}}$ converges to some element $y := \lim_{i\to\infty} A^*x_i$. Since the sequence $(x_i)_{i\in\mathbb{N}}$ converges to x and $(A^*x_i)_{i\in\mathbb{N}}$ converges to y and A^* has a closed graph, it follows that $x \in \mathrm{dom}(A^*)$ and $A^*x = y$. Hence

$$\|A^*x\| = \|y\| = \lim_{i\to\infty} \|A^*x_i\| = \lim_{i\to\infty} \|Ax_i\| = \|Ax\| .$$

This shows that $\mathrm{dom}(A) \subset \mathrm{dom}(A^*)$ and $\|A^*x\| = \|Ax\|$ for all $x \in \mathrm{dom}(A)$. The converse inclusion $\mathrm{dom}(A^*) \subset \mathrm{dom}(A)$ follows by interchanging the roles of A and A^*. This shows that (i) implies (ii).

We prove that (ii) implies (i). Assume $\mathrm{dom}(A) = \mathrm{dom}(A^*)$ and

$$\|Ax\| = \|A^*x\| \qquad \text{for all } x \in \mathrm{dom}(A).$$

Then the same argument as in the proof of Lemma 5.3.14 shows that

$$\langle Ax, Ay\rangle = \langle A^*x, A^*y\rangle \qquad \text{for all } x, y \in \mathrm{dom}(A). \tag{6.3.16}$$

Now let $x \in \mathrm{dom}(A^*A)$. Then $x \in \mathrm{dom}(A)$ and $Ax \in \mathrm{dom}(A^*)$ and, by equation (6.3.16), we have

$$|\langle A^*x, A^*\xi\rangle| = |\langle Ax, A\xi\rangle| = |\langle A^*Ax, \xi\rangle| \le \|A^*Ax\| \, \|\xi\|$$

for all $\xi \in \mathrm{dom}(A^*)$. This implies $A^*x \in \mathrm{dom}(A)$ and hence $x \in \mathrm{dom}(AA^*)$. Thus we have proved that $\mathrm{dom}(A^*A) \subset \mathrm{dom}(AA^*)$. The same argument, with the roles of A and A^* reversed, shows that

$$\mathrm{dom}(A^*A) = \mathrm{dom}(AA^*).$$

Now let $x \in \mathrm{dom}(A^*A) = \mathrm{dom}(AA^*)$. Then, by equation (6.3.16), we have

$$\langle A^*Ax, \xi\rangle = \langle Ax, A\xi\rangle = \langle A^*x, A^*\xi\rangle = \langle AA^*x, \xi\rangle$$

for all $\xi \in \mathrm{dom}(A) = \mathrm{dom}(A^*)$. Since $\mathrm{dom}(A)$ is dense in H, this implies $A^*Ax = AA^*x$. Thus we have proved that (ii) implies (i).

We prove that (ii) implies (iii). Assume $\mathrm{dom}(A) = \mathrm{dom}(A^*)$ and

$$\|Ax\| = \|A^*x\| \qquad \text{for all } x \in \mathrm{dom}(A).$$

Define the operators $B_1, B_2 : \mathrm{dom}(A) \to H$ by

$$B_1x := \frac{1}{2}\big(Ax + A^*x\big), \qquad B_2x := \frac{1}{2\mathbf{i}}\big(Ax - A^*x\big)$$

for $x \in \mathrm{dom}(A)$. These operators are symmetric and hence closeable by Example 2.2.23. Thus they admit self-adjoint extensions $A_i : \mathrm{dom}(A_i) \to H$ for $i = 1, 2$ by Exercise 6.3.3. Moreover, $\mathrm{dom}(A) \subset \mathrm{dom}(A_1) \cap \mathrm{dom}(A_2)$ and every element $x \in \mathrm{dom}(A) = \mathrm{dom}(A^*)$ satisfies

$$Ax = A_1x + \mathbf{i}A_2x, \qquad A^*x = A_1x - \mathbf{i}A_2x,$$

and

$$\begin{aligned} \|Ax\|^2 &= \frac{1}{2}\Big(\|Ax\|^2 + \|A^*x\|^2\Big) \\ &= \frac{1}{4}\Big(\|Ax + A^*x\|^2 + \|Ax - A^*x\|^2\Big) \\ &= \|A_1x\|^2 + \|A_2x\|^2. \end{aligned}$$

Now let $x \in \mathrm{dom}(A_1) \cap \mathrm{dom}(A_2)$. Then

$$\begin{aligned} |\langle x, A\xi\rangle| &= |\langle x, A_1\xi + \mathbf{i}A_2\xi\rangle| \\ &= |\langle A_1x, \xi\rangle + \langle A_2x, \mathbf{i}\xi\rangle| \\ &\le \Big(\|A_1x\| + \|A_2x\|\Big) \|\xi\| \end{aligned}$$

for every $\xi \in \mathrm{dom}(A)$ and hence

$$x \in \mathrm{dom}(A^*) = \mathrm{dom}(A).$$

This shows that (ii) implies (iii).

We prove that (iii) implies (ii). Assume $A_i : \mathrm{dom}(A_i) \to H$ for $i = 1, 2$ are self-adjoint operators that satisfy the following four conditions.

(a) $\mathrm{dom}(A_1) \cap \mathrm{dom}(A_2)$ is a dense subspace of H.

(b) $\|A_1x + \mathbf{i}A_2x\|^2 = \|A_1x\|^2 + \|A_2x\|^2$ for all $x \in \mathrm{dom}(A_1) \cap \mathrm{dom}(A_2)$.

(c) Let $y \in H$ and $c > 0$ such that $|\langle y, A_1x + \mathbf{i}A_2x\rangle| \le c\,\|x\|$ for every element $x \in \mathrm{dom}(A_1) \cap \mathrm{dom}(A_2)$. Then $y \in \mathrm{dom}(A_1) \cap \mathrm{dom}(A_2)$.

(d) Let $x \in H$ and $c > 0$ such that $|\langle x, A_1y - \mathbf{i}A_2y\rangle| \le c\,\|y\|$ for every element $y \in \mathrm{dom}(A_1) \cap \mathrm{dom}(A_2)$. Then $x \in \mathrm{dom}(A_1) \cap \mathrm{dom}(A_2)$.

Define the operator $A : \mathrm{dom}(A) \to H$ by

$$\begin{aligned} \mathrm{dom}(A) &:= \mathrm{dom}(A_1) \cap \mathrm{dom}(A_2), \\ Ax &:= A_1x + \mathbf{i}A_2x \qquad \text{for } x \in \mathrm{dom}(A_1) \cap \mathrm{dom}(A_2). \end{aligned} \tag{6.3.17}$$

Its domain is dense by (a). We prove that its adjoint operator is given by

$$\begin{aligned} \mathrm{dom}(A^*) &= \mathrm{dom}(A_1) \cap \mathrm{dom}(A_2), \\ A^*y &= A_1y - \mathbf{i}A_2y \qquad \text{for } y \in \mathrm{dom}(A_1) \cap \mathrm{dom}(A_2). \end{aligned} \tag{6.3.18}$$

Let $y \in \mathrm{dom}(A^*)$. Then $\langle y, Ax\rangle = \langle A^*y, x\rangle$ for all $x \in \mathrm{dom}(A)$ and this implies $y \in \mathrm{dom}(A_1) \cap \mathrm{dom}(A_2)$ by (c). Hence

$$\langle A^*y, x\rangle = \langle y, A_1x + \mathbf{i}A_2x\rangle = \langle A_1y - \mathbf{i}A_2y, x\rangle$$

for all $x \in \mathrm{dom}(A_1) \cap \mathrm{dom}(A_2)$, and hence $A^*y = A_1y - \mathbf{i}A_2y$ by (a). The converse inclusion $\mathrm{dom}(A_1) \cap \mathrm{dom}(A_2) \subset \mathrm{dom}(A^*)$ follows directly from the assumptions. This shows that (6.3.18) is the adjoint operator of (6.3.17) and vice versa by the same argument, using (d) instead of (c). In particular, A has a closed graph. Moreover, it follows from (b) that $\|Ax\| = \|A^*x\|$ for all $x \in \mathrm{dom}(A) = \mathrm{dom}(A^*)$. This shows that (iii) implies (ii) and completes the proof of Theorem 6.3.11. □

Let H be a separable complex Hilbert space, equipped with an orthonormal basis $(e_i)_{i\in\mathbb{N}}$. Then the operator $A_\lambda : \mathrm{dom}(A_\lambda) \to H$ in Example 6.1.3 is normal for every sequence of complex numbers $(\lambda_i)_{i\in\mathbb{N}}$. The operator A_λ is bounded if and only if the sequence $(\lambda_i)_{i\in\mathbb{N}}$ is bounded, it is self-adjoint if and only if $\lambda_i \in \mathbb{R}$ for all i (Exercise 6.3.5), it is compact if and only if $\lim_{i\to\infty}|\lambda_i| = 0$ (Example 4.2.8), and it has a compact resolvent if and only if $\lim_{i\to\infty}|\lambda_i| = \infty$. This example shows that the domains of the self-adjoint operators $A_1 = A_{\mathrm{Re}\lambda}$ and $A_2 = A_{\mathrm{Im}\lambda}$ in Theorem 6.3.11 may differ dramatically from the domain of $A = A_\lambda$. It also shows that every nonempty closed subset of the complex plane can be the spectrum of an unbounded normal operator (Example 6.1.16). In particular, the resolvent set can be empty. The next theorem shows that every normal operator has a nonempty spectrum.

Theorem 6.3.13 (Spectrum of a Normal Operator). *Let H be a nonzero complex Hilbert space and let $A : \mathrm{dom}(A) \to H$ be an unbounded normal operator with $\mathrm{dom}(A) \subsetneq H$. Then the following hold.*

(i) *If $\lambda \in \mathbb{C}$, then $\lambda 1\!\!1 - A$ is normal and, if $\lambda \in \rho(A)$, then the resolvent operator $R_\lambda(A) = (\lambda 1\!\!1 - A)^{-1}$ is normal.*

(ii) $\sigma(A) \neq \emptyset$.

(iii) $\mathrm{R}\sigma(A) = \emptyset$ *and* $\mathrm{P}\sigma(A^*) = \left\{\overline{\lambda} \,|\, \lambda \in \mathrm{P}\sigma(A)\right\}$.

(iv) *If A has a compact resolvent, then the spectrum $\sigma(A) = \mathrm{P}\sigma(A)$ is discrete, for each $\lambda \in \mathrm{P}\sigma(A)$ the eigenspace $E_\lambda := \ker(\lambda 1\!\!1 - A)$ is finite-dimensional, and A admits an orthonormal basis of eigenvectors.*

(v) *If A is self-adjoint, then $\sigma(A) \subset \mathbb{R}$ and*

$$\begin{aligned} \sup \sigma(A) &= \sup \left\{ \langle x, Ax \rangle \,\middle|\, x \in \mathrm{dom}(A),\ \|x\| = 1 \right\}, \\ \inf \sigma(A) &= \inf \left\{ \langle x, Ax \rangle \,\middle|\, x \in \mathrm{dom}(A),\ \|x\| = 1 \right\}. \end{aligned} \tag{6.3.19}$$

Proof. We prove part (i). Let $\lambda \in \mathbb{C}$. Then $(\lambda 1\!\!1 - A)^* = \overline{\lambda} 1\!\!1 - A^*$ by part (i) of Lemma 6.3.2. Hence

$$\begin{aligned} \|\lambda x - Ax\|^2 &= |\lambda|^2 \|x\|^2 - 2\mathrm{Re}\langle \lambda x, Ax\rangle + \|Ax\|^2 \\ &= |\overline{\lambda}|^2 \|x\|^2 - 2\mathrm{Re}\langle A^* x, \overline{\lambda} x\rangle + \|A^* x\|^2 \\ &= \left\|\overline{\lambda} x - A^* x\right\|^2 \end{aligned}$$

for all $x \in \mathrm{dom}(A) = \mathrm{dom}(\lambda 1\!\!1 - A) = \mathrm{dom}(\overline{\lambda} 1\!\!1 - A^*)$. Thus $\lambda 1\!\!1 - A$ is normal. If A is invertible, then

$$\begin{aligned} A^{-1}(A^{-1})^* &= A^{-1}(A^*)^{-1} \\ &= (A^* A)^{-1} \\ &= (AA^*)^{-1} \\ &= (A^*)^{-1} A^{-1} \\ &= (A^{-1})^* A^{-1} \end{aligned}$$

by part (vii) of Lemma 6.3.2, and hence A^{-1} is normal. This proves part (i).

We prove part (ii). If $\rho(A) = \emptyset$, then $\sigma(A) = \mathbb{C} \neq \emptyset$. If $\rho(A) \neq \emptyset$ and $\mu \in \rho(A)$, then $R_\mu(A)$ is normal by part (i), hence

$$\sup_{z \in \sigma(R_\mu(A))} |z| = \|R_\mu(A)\| > 0$$

by Theorem 5.3.15, and hence

$$\sigma(A) = \left\{ \mu - z^{-1} \,|\, z \in \sigma(R_\mu(A)) \setminus \{0\} \right\} \neq \emptyset$$

by Lemma 6.1.12. This proves part (ii).

We prove part (iii). Fix an element $\lambda \in \mathbb{C} \setminus (\mathrm{P}\sigma(A) \cup \mathrm{C}\sigma(A))$. Then the operator $\lambda\mathbb{1} - A$ is normal by part (i) and is injective because $\lambda \notin \mathrm{P}\sigma(A)$. Hence the adjoint operator $(\lambda\mathbb{1} - A)^* = \overline{\lambda}\mathbb{1} - A^*$ is injective by definition of a normal operator in Theorem 6.3.11. Thus $\lambda\mathbb{1} - A$ has a dense image by part (v) of Lemma 6.3.2 and so $\lambda\mathbb{1} - A$ is surjective because $\lambda \notin \mathrm{C}\sigma(A)$. Thus $\lambda \in \rho(A)$ and this proves part (iii).

We prove part (iv). By assumption $\rho(A) \neq \emptyset$ and the resolvent operator $R_\mu(A)$ is compact for all $\mu \in \rho(A)$. Fix an element $\mu \in \rho(A)$. Then Theorem 5.2.8 asserts that $\sigma(R_\mu(A)) \setminus \{0\} = \mathrm{P}\sigma(R_\mu(A))$, that the spectrum of $R_\mu(A)$ can only accumulate at the origin, and that the eigenspaces of $R_\mu(A)$ are all finite-dimensional. Moreover, Theorem 5.3.15 asserts that the operator $R_\mu(A)$ admits an orthonormal basis of eigenvectors. Hence part (iv) follows from Lemma 6.1.12.

We prove part (v). Assume A is self-adjoint and let $\lambda \in \mathbb{C} \setminus \mathbb{R}$. Then

$$\|\lambda x - Ax\|^2 = (\mathrm{Im}\lambda)^2 \|x\|^2 + \|(\mathrm{Re}\lambda)x - Ax\|^2 \geq (\mathrm{Im}\lambda)^2 \|x\|^2$$

for all $x \in \mathrm{dom}(A)$ as in the proof of Theorem 5.3.16. Hence $\lambda\mathbb{1} - A$ is injective and has a closed image by Theorem 6.2.3. Replace λ by $\overline{\lambda}$ to deduce that the adjoint operator $\overline{\lambda}\mathbb{1} - A^* = \overline{\lambda}\mathbb{1} - A$ is also injective, hence $\lambda\mathbb{1} - A$ has a dense image by part (iv) of Lemma 6.3.2, so $\lambda\mathbb{1} - A$ is bijective and $\lambda \in \rho(A)$.

Now let $\lambda \in \mathbb{R}$ and assume

$$\lambda > \sup_{x \in \mathrm{dom}(A),\, \|x\|=1} \langle x, Ax \rangle =: c.$$

Then

$$\|x\|\|\lambda x - Ax\| \geq \langle x, \lambda x - Ax \rangle \geq (\lambda - c)\|x\|^2 \qquad \text{for all } x \in \mathrm{dom}(A).$$

Hence $\lambda\mathbb{1} - A$ is injective and has a closed image by Theorem 6.2.3 and so is bijective by Lemma 6.3.2. This shows that $\sigma(A) \subset (-\infty, c]$.

Conversely, assume

$$c := \sup \sigma(A) < \infty.$$

We must prove that $\langle x, Ax \rangle \leq c$ for all $x \in \mathrm{dom}(A)$ with $\|x\| = 1$. Suppose, by contradiction, that there exists an element $x \in \mathrm{dom}(A)$ such that $\|x\| = 1$ and $\langle x, Ax \rangle > c$. Choose a real number μ such that $c < \mu < \langle x, Ax \rangle$ and define $\xi := \mu x - Ax$. Then $\mu \in \rho(A)$ by assumption and

$$\langle \xi, R_\mu(A)\xi \rangle = \langle \mu x - Ax, x \rangle = \mu - \langle Ax, x \rangle < 0.$$

However, by Lemma 6.1.12, we have

$$\sigma(R_\mu(A)) = \{(\mu - \lambda)^{-1} \,|\, \lambda \in \sigma(A)\} \cup \{0\} \subset [0, \infty)$$

in contradiction to Theorem 5.3.16. This proves Theorem 6.3.13. $\square$

6.4. Functional Calculus and Spectral Measures

The purpose of the present section is to extend the measurable functional calculus and the spectral measure to unbounded self-adjoint operators.

6.4.1. Functional Calculus. For a topological space Σ let $B(\Sigma)$ denote the C* algebra of bounded Borel measurable functions $f : \Sigma \to \mathbb{C}$ with the supremum norm $\|f\| := \sup_{\lambda\in\Sigma}|f(\lambda)|$. Denote by $C_b(\Sigma) \subset B(\Sigma)$ the C* subalgebra of complex valued bounded continuous functions on Σ. The next theorem extends the functional calculus of Theorem 5.6.5 to unbounded self-adjoint operators.

Theorem 6.4.1 (**Functional Calculus**). *Let H be a nonzero complex Hilbert space, let $A : \mathrm{dom}(A) \to H$ be an unbounded self-adjoint operator, and let $\Sigma := \sigma(A) \subset \mathbb{R}$. Then there exists a C* algebra homomorphism*

$$B(\Sigma) \to \mathcal{L}^c(H) : f \mapsto f(A) =: \Psi_A(f) \tag{6.4.1}$$

that satisfies the following axioms.

(Normalization) *Let $f_i \in B(\Sigma)$ be a sequence such that $\sup_{i\in\mathbb{N}} |f_i(\lambda)| \le |\lambda|$ and $\lim_{i\to\infty} f_i(\lambda) = \lambda$ for all $\lambda \in \Sigma$. Then*

$$\lim_{i\to\infty} f_i(A)x = Ax \qquad \textit{for all } x \in \mathrm{dom}(A).$$

(Convergence) *Let $f_i \in B(\Sigma)$ be a sequence such that $\sup_{i\in\mathbb{N}} \|f_i\| < \infty$ and let $f \in B(\Sigma)$ such that $\lim_{i\to\infty} f_i(\lambda) = f(\lambda)$ for all $\lambda \in \Sigma$. Then*

$$\lim_{i\to\infty} f_i(A)x = f(A)x \qquad \textit{for all } x \in H.$$

(Positive) *If $f \in B(\Sigma, \mathbb{R})$ and $f \ge 0$, then $f(A) = f(A)^* \ge 0$.*

(Contraction) *$\|f(A)\| \le \|f\|$ for all $f \in B(\Sigma)$ and $\|f(A)\| = \|f\|$ for all $f \in C_b(\Sigma)$.*

(Commutative) *If $B \in \mathcal{L}^c(H)$ satisfies $AB = BA$, then $f(A)B = Bf(A)$ for all $f \in B(\Sigma)$.*

(Eigenvector) *If $\lambda \in \Sigma$ and $x \in \mathrm{dom}(A)$ satisfy $Ax = \lambda x$, then every function $f \in B(\Sigma)$ satisfies $f(A)x = f(\lambda)x$.*

(Spectrum) *If $f \in B(\Sigma)$, then $f(A)$ is normal and $\sigma(f(A)) \subset \overline{f(\Sigma)}$. Moreover, $\sigma(f(A)) = \overline{f(\Sigma)}$ for all $f \in C_b(\Sigma)$.*

(Composition) *If $f \in C_b(\Sigma)$ and $g \in B(\overline{f(\Sigma)})$, then $(g \circ f)(A) = g(f(A))$.*

The C algebra homomorphism* (6.4.1) *is uniquely determined by the (Normalization) and (Convergence) axioms.*

Proof. See page 329. □

Theorem 6.4.2 (Cayley Transform). *Let H be a complex Hilbert space.*

(i) *Let $A : \mathrm{dom}(A) \to H$ be a self-adjoint operator. Then the operator*

$$U := (A - \mathbf{i}\mathbb{1})(A + \mathbf{i}\mathbb{1})^{-1} : H \to H \tag{6.4.2}$$

is unitary, the operator $\mathbb{1} - U : H \to H$ is injective, and

$$\mathrm{dom}(A) = \mathrm{im}(\mathbb{1} - U), \qquad A = \mathbf{i}(\mathbb{1} + U)(\mathbb{1} - U)^{-1}. \tag{6.4.3}$$

The operator U is called the **Cayley transform** *of A.*

(ii) *Let $U \in \mathcal{L}^c(H)$ be a unitary operator such that $\mathbb{1} - U$ is injective. Then the operator*

$$A := \mathbf{i}(\mathbb{1} + U)(\mathbb{1} - U)^{-1} : \mathrm{dom}(A) \to H, \qquad \mathrm{dom}(A) := \mathrm{im}(\mathbb{1} - U),$$

is self-adjoint and U is the Cayley transform of A.

(iii) *Let $A : \mathrm{dom}(A) \to H$ be a self-adjoint operator and let $U \in \mathcal{L}^c(H)$ be its Cayley transform. Define the Möbius transformation $\phi : \mathbb{R} \to S^1 \setminus \{1\}$ by*

$$\phi(\lambda) := \frac{\lambda - \mathbf{i}}{\lambda + \mathbf{i}}, \qquad \phi^{-1}(\mu) = \mathbf{i}\frac{1 + \mu}{1 - \mu} \tag{6.4.4}$$

for $\lambda \in \mathbb{R}$ and $\mu \in S^1 \setminus \{1\}$. Then

$$\sigma(U) \setminus \{1\} = \phi(\sigma(A)), \qquad \mathrm{P}\sigma(U) = \phi(\mathrm{P}\sigma(A)), \tag{6.4.5}$$

and

$$\ker(\lambda\mathbb{1} - A) = \ker(\phi(\lambda)\mathbb{1} - U) \tag{6.4.6}$$

for all $\lambda \in \mathbb{R}$.

Proof. We prove (i). The operators

$$A \pm \mathbf{i}\mathbb{1} : \mathrm{dom}(A) \to H$$

are bijective and have bounded inverses by part (v) of Theorem 6.3.13 and they are normal by part (i) of Theorem 6.3.13. Hence

$$\|Ax - \mathbf{i}x\| = \|Ax + \mathbf{i}x\| \qquad \text{for all } x \in \mathrm{dom}(A)$$

and so the Cayley transform

$$U := (A - \mathbf{i}\mathbb{1})(A + \mathbf{i}\mathbb{1})^{-1}$$

in (6.4.2) is a unitary operator on H (Lemma 5.3.14). The operator U satisfies

$$\mathbb{1} - U = 2\mathbf{i}(A + \mathbf{i}\mathbb{1})^{-1}, \qquad \mathbb{1} + U = 2A(A + \mathbf{i}\mathbb{1})^{-1}.$$

Thus $\mathbb{1} - U$ is injective, $\mathrm{im}(\mathbb{1} - U) = \mathrm{dom}(A)$, and $\mathbf{i}^{-1}A(\mathbb{1} - U) = \mathbb{1} + U$, and hence A and U satisfy (6.4.3). This proves part (i).

We prove (ii). Assume $U \in \mathcal{L}^c(H)$ is a unitary operator such that $\mathbb{1} - U$ is injective. Then $1 \in \mathbb{C} \setminus \mathrm{P}\sigma(U)$ and hence the operator $\mathbb{1} - U$ has a dense image by Theorem 5.3.15. Define the operator $A : \mathrm{dom}(A) \to H$ by (6.4.3). We prove that A is self-adjoint. Thus let

$$x \in \mathrm{dom}(A^*), \qquad y := A^*x.$$

Then

$$\langle y, \zeta \rangle = \langle x, A\zeta \rangle = \langle x, \mathbf{i}(\mathbb{1} + U)(\mathbb{1} - U)^{-1}\zeta \rangle$$

for all $\zeta \in \mathrm{dom}(A) = \mathrm{im}(\mathbb{1} - U)$ and hence

$$\langle y, \xi - U\xi \rangle = \langle x, \mathbf{i}(\xi + U\xi) \rangle \qquad \text{for all } \xi \in H.$$

This implies $U^*y - y = \mathbf{i}(U^*x + x)$ and hence

$$y - Uy = \mathbf{i}(x + Ux). \tag{6.4.7}$$

Thus

$$x = \tfrac{1}{2}(x - Ux) + \tfrac{1}{2}(x + Ux) = \tfrac{1}{2}(\mathbb{1} - U)(x - \mathbf{i}y) \in \mathrm{im}(\mathbb{1} - U) = \mathrm{dom}(A),$$

hence

$$(\mathbb{1} - U)^{-1}x = \tfrac{1}{2}(x - \mathbf{i}y),$$

and therefore

$$Ax = \mathbf{i}(\mathbb{1} + U)(\mathbb{1} - U)^{-1}x = \tfrac{1}{2}(\mathbb{1} + U)(\mathbf{i}x + y) = y.$$

Here the last equation follows from (6.4.7). This shows that A is self-adjoint. Moreover, $A + \mathbf{i}\mathbb{1} = 2\mathbf{i}(\mathbb{1} - U)^{-1}$ and $\mathbb{A} - \mathbf{i}\mathbb{1} = 2\mathbf{i}U(\mathbb{1} - U)^{-1}$, and hence $U = (A - \mathbf{i}\mathbb{1})(A + \mathbf{i}\mathbb{1})^{-1}$ is the Cayley transform of A. This proves part (ii).

We prove (iii). Fix a real number λ. Then, by (6.4.2) and (6.4.4),

$$\begin{aligned}(\lambda + \mathbf{i})(\phi(\lambda)\mathbb{1} - U)(Ax + \mathbf{i}x) &= (\lambda - \mathbf{i})(Ax + \mathbf{i}x) - (\lambda + \mathbf{i})(Ax - \mathbf{i}x) \\ &= 2\mathbf{i}(\lambda x - Ax)\end{aligned}$$

for all $x \in \mathrm{dom}(A)$. Since the operator $A + \mathbf{i}\mathbb{1} : \mathrm{dom}(A) \to H$ is surjective, this implies that $\lambda\mathbb{1} - A : \mathrm{dom}(A) \to H$ is bijective if and only if $\phi(\lambda)\mathbb{1} - U$ is bijective. Moreover, if $x \in \mathrm{dom}(A)$ satisfies $Ax = \lambda x$, then

$$\begin{aligned}(\lambda + \mathbf{i})^2(\phi(\lambda)x - Ux) &= (\lambda + \mathbf{i})(\phi(\lambda)\mathbb{1} - U)(\lambda x + \mathbf{i}x) \\ &= (\lambda + \mathbf{i})(\phi(\lambda)\mathbb{1} - U)(Ax + \mathbf{i}x) \\ &= 2\mathbf{i}(\lambda x - Ax) \\ &= 0.\end{aligned}$$

Conversely, let $x \in H$ such that $Ux = \phi(\lambda)x$. Then $(1 - \phi(\lambda))x = x - Ux$ and so $x \in \mathrm{im}(\mathbb{1} - U) = \mathrm{dom}(A)$. Moreover, $\xi := (\mathbb{1} - U)^{-1}x = (1 - \phi(\lambda))^{-1}x$ and so

$$Ax = \mathbf{i}(\xi + U\xi) = \mathbf{i}\frac{x + Ux}{1 - \phi(\lambda)} = \mathbf{i}\frac{1 + \phi(\lambda)}{1 - \phi(\lambda)}x = \lambda x.$$

This proves part (iii) and Theorem 6.4.2. □

With these preparations we are now ready to establish the functional calculus for general unbounded self-adjoint operators. We give a proof of Theorem 6.4.1 which reduces the result to the functional calculus for bounded normal operators in Theorem 5.6.5 via the Cayley transform.

Proof of Theorem 6.4.1. Let $A : \mathrm{dom}(A) \to H$ be a self-adjoint operator with domain $\mathrm{dom}(A) \subsetneq H$ (so A is not bounded) and spectrum

$$\Sigma := \sigma(A) \subset \mathbb{R}.$$

Let

$$U := (A - \mathbf{i}\mathbb{1})(A + \mathbf{i}\mathbb{1})^{-1} \in \mathcal{L}^c(H)$$

be the Cayley transform of A. Then U is a unitary operator and $\mathbb{1} - U$ is injective and not surjective, because $\mathrm{im}(\mathbb{1} - U) = \mathrm{dom}(A) \neq H$, and so

$$1 \in \sigma(U).$$

Hence it follows from part (iii) of Theorem 6.4.2 that the spectrum of U is the (compact) set

$$\sigma(U) = \phi(\Sigma) \cup \{1\} \subset S^1. \tag{6.4.8}$$

Now denote by

$$B(\sigma(U)) \to \mathcal{L}^c(H) : g \mapsto g(U)$$

the C* algebra homomorphism in Theorem 5.6.5, and define the map

$$B(\Sigma) \to \mathcal{L}^c(H) : f \mapsto f(A)$$

by

$$f(A) := (f \circ \phi^{-1})(U) \qquad \text{for } f \in B(\Sigma). \tag{6.4.9}$$

Here the bounded measurable function $f \circ \phi^{-1} : S^1 \setminus \{1\} \to \mathbb{C}$ is extended to all of S^1 by setting $(f \circ \phi^{-1})(1) := 0$. We prove in seven steps that the map (6.4.9) satisfies the requirements of Theorem 6.4.1.

Step 1. *The map* (6.4.9) *is a C* algebra homomorphism. In particular, it satisfies* $1(A) = \mathbb{1}$.

Define $g_0 : \sigma(U) \to \mathbb{C}$ by $g_0(1) := 1$ and

$$g_0(\mu) := 0 \qquad \text{for } \mu \in \sigma(U) \setminus \{1\}.$$

Then the operator $g_0(U)$ is the orthogonal projection onto the kernel of the operator $\mathbb{1} - U$ by part (iii) of Theorem 5.6.11, and so $g_0(U) = 0$ because $\mathbb{1} - U$ is injective. This implies

$$1(A) = (1 \circ \phi^{-1})(U) = (1 - g_0)(U) = \mathbb{1}.$$

That the map (6.4.9) is linear and preserves multiplication follows directly from the definition. This proves Step 1.

Step 2. *The map* (6.4.9) *satisfies the (Normalization) axiom.*

Let $f_i : \Sigma \to \mathbb{C}$ be a sequence of bounded measurable functions such that

$$\sup_{i\in\mathbb{N}}|f_i(\lambda)| \le |\lambda|, \qquad \lim_{i\to\infty} f_i(\lambda) = \lambda \qquad \text{for all } \lambda \in \Sigma.$$

For $i \in \mathbb{N}$ define the function $h_i : \sigma(U) \to \mathbb{C}$ by

$$h_i(\mu) := (f_i \circ \phi^{-1})(\mu)(1-\mu) \qquad \text{for } \mu \in \sigma(U),$$

so $h_i : \sigma(U) \to \mathbb{C}$ is a bounded measurable function and

$$h_i(U) = f_i(A)(\mathbb{1} - U). \tag{6.4.10}$$

Moreover, $\phi^{-1}(\mu) = \mathbf{i}(1+\mu)(1-\mu)^{-1}$ for $\mu \in \sigma(U) \setminus \{1\}$ and hence

$$\begin{aligned}|h_i(\mu)| &= \left|f_i\left(\mathbf{i}\frac{1+\mu}{1-\mu}\right)\right||1-\mu| \\ &\le |1+\mu| \\ &\le 2\end{aligned}$$

for all $\mu \in \sigma(U) \setminus \{1\}$. Since $h_i(1) = 0$ for all i, this implies

$$\sup_{i\in\mathbb{N}}|h_i(\mu)| \le 2, \qquad \lim_{i\to\infty} h_i(\mu) = \mathbf{i}(1+\mu-2g_0(\mu)) \quad \text{for all } \mu \in \sigma(U), \tag{6.4.11}$$

where $g_0 : \sigma(U) \to \mathbb{C}$ is as in the proof of Step 1. Now let

$$x \in \mathrm{dom}(A) = \mathrm{im}(\mathbb{1} - U)$$

and define

$$\xi := (\mathbb{1} - U)^{-1}x.$$

Then it follows from (6.4.3), (6.4.10), (6.4.11), and the (Convergence) axiom in Theorem 5.6.5 that

$$\begin{aligned}\lim_{i\to\infty} f_i(A)x &= \lim_{i\to\infty} f_i(A)(\xi - U\xi) \\ &= \lim_{i\to\infty} h_i(U)\xi \\ &= \mathbf{i}(\xi + U\xi) \\ &= Ax.\end{aligned}$$

This proves Step 2.

Step 3. *The map* (6.4.9) *satisfies the (Convergence), (Positive), (Commutative), and (Eigenvector) axioms.*

The (Convergence) and (Positive) axioms follow directly from the definition and the corresponding axioms in Theorem 5.6.5. The (Commutative) axiom follows from the (Commutative) axiom in Theorem 5.6.5 and the fact that an operator $B \in \mathcal{L}^c(H)$ commutes with A if and only if it commutes with U (and hence also with $U^* = U^{-1}$). The (Eigenvector) axiom follows from equation (6.4.6) and the (Eigenvector) axiom in Theorem 5.6.5.

Step 4. *The map* (6.4.9) *satisfies the (Spectrum) axiom.*

Let $f \in B(\Sigma)$ and $\mu \in \mathbb{C} \setminus \overline{f(\Sigma)}$, and define the function $g : \Sigma \to \mathbb{C}$ by

$$g(\lambda) := \frac{1}{\mu - f(\lambda)} \qquad \text{for } \lambda \in \Sigma.$$

Then g is bounded and measurable and satisfies $g(\mu - f) = (\mu - f)g = 1$. Hence $g(A)(\mu\mathbb{1} - f(A)) = (\mu\mathbb{1} - f(A))g(A) = \mathbb{1}$ by Step 1, so $\mu\mathbb{1} - f(A)$ is invertible and thus $\mu \in \rho(f(A))$. This shows that $\sigma(f(A)) \subset \overline{f(\Sigma)}$.

Let $f \in C_b(\Sigma)$ and define the function $g : \sigma(U) \to \mathbb{C}$ by

$$g(z) := \begin{cases} f(\phi^{-1}(z)), & \text{for } z \in \sigma(U) \setminus \{1\}, \\ 0, & \text{for } z = 1. \end{cases}$$

Then g is continuous at every point $z \in \sigma(U) \setminus \{1\}$ and $f(A) = g(U)$. Hence

$$f(\lambda) = g(\phi(\lambda)) \in \sigma(g(U)) = \sigma(f(A)) \qquad \text{for all } \lambda \in \Sigma$$

by part (ii) of Theorem 5.6.11. Hence $\overline{f(\Sigma)} \subset \sigma(f(A))$ because the spectrum of $f(A)$ is a closed subset of $\mathbb{C}$. This proves Step 4.

Step 5. *The map* (6.4.9) *satisfies the (Contraction) axiom.*

This follows from Step 4 and the formula $\|f(A)\| = \sup_{\mu \in \sigma(f(A))} |\mu|$ in part (ii) of Theorem 5.3.15.

Step 6. *The map* (6.4.9) *satisfies the (Composition) axiom.*

Fix a function $f \in C_b(\Sigma)$ and define $A_f := f(A)$. Then $\Sigma_f := \sigma(A_f) = \overline{f(\Sigma)}$ by Step 4. Consider the map $B(\Sigma_f) \to \mathcal{L}^c(H) : g \mapsto g(A_f) := (g \circ f)(A)$. This map is a C* algebra homomorphism by Step 1, it is continuous by Step 5, it satisfies the (Normalization) axiom $\mathrm{id}(A_f) = A_f$ by definition, and it satisfies the (Convergence) axiom by Step 3. Hence Step 6 follows from uniqueness in Theorem 5.6.5.

Step 7. *The C* algebra homomorphism* (6.4.9) *is uniquely determined by the (Normalization) and (Convergence) axioms.*

Let $B(\Sigma) \to \mathcal{L}^c(H) : f \mapsto f(A)$ be any C* algebra homomorphism that satisfies the (Normalization) and (Convergence) axioms and define $U := \phi(A)$. Then $U(A + \mathbf{i}\mathbb{1}) = A - \mathbf{i}\mathbb{1}$ by the (Normalization) axiom, so U is the Cayley transform of A. Define the map $B(\sigma(U)) \to \mathcal{L}^c(H) : g \mapsto g(U)$ by $g(U) := (g \circ \phi)(A)$ for $g \in B(\sigma(U))$. By definition, this map is a C* algebra homomorphism that satisfies the (Convergence) axiom. Moreover, it satisfies $\mathrm{id}(U) = \phi(A) = U$. Hence the map $g \mapsto g(U)$ agrees with the functional calculus in Theorem 5.6.5. This proves Step 7 and Theorem 6.4.1. □

6.4.2. Spectral Measures. Let $\mathcal{B} \subset 2^{\mathbb{R}}$ be the Borel σ-algebra. Theorem 6.4.1 allows us to assign to every unbounded self-adjoint operator on a complex Hilbert space a projection valued measure (see Definition 5.6.1).

Definition 6.4.3 (Spectral Measure). Let H be a nonzero complex Hilbert space, let $A : \mathrm{dom}(A) \to H$ be an unbounded self-adjoint operator with spectrum $\Sigma := \sigma(A) \subset \mathbb{R}$, and let

$$\Psi_A : B(\Sigma) \to \mathcal{L}^c(H)$$

be the C* algebra homomorphism of Theorem 6.4.1. For every $\Omega \in \mathcal{B}$ define the operator $P_\Omega \in \mathcal{L}^c(H)$ by

$$P_\Omega := \Psi_A(\chi_\Omega|_\Sigma). \tag{6.4.12}$$

By Theorem 6.4.1 these operators are orthogonal projections and the map

$$\mathcal{B} \to \mathcal{L}^c(H) : \Omega \mapsto P_\Omega \tag{6.4.13}$$

is a projection valued measure. It is called the **spectral measure of** A.

Conversely, every projection valued measure (6.4.12) on the real axis gives rise to a family of self-adjoint operators $A_f : \mathrm{dom}(A_f) \to H$, one for every Borel measurable function $f : \mathbb{R} \to \mathbb{R}$. If f is bounded, then this operator is bounded, so $\mathrm{dom}(A_f) = H$, and it is given by the formula $A_f := \Psi(f)$ in Theorem 5.6.2. For unbounded functions f the operator A_f will in general be unbounded.

Theorem 6.4.4 (The Operator A_f). *Let H be a nonzero complex Hilbert space and fix any projection valued measure $\mathcal{B} \to \mathcal{L}^c(H) : \Omega \mapsto P_\Omega$ on the real axis. Define the signed Borel measures $\mu_{y,x} : \mathcal{B} \to \mathbb{R}$ by*

$$\mu_{y,x}(\Omega) := \mathrm{Re}\langle y, P_\Omega x\rangle \qquad \text{for } x, y \in H \text{ and } \Omega \in \mathcal{B}. \tag{6.4.14}$$

Let $f : \mathbb{R} \to \mathbb{R}$ be a Borel measurable function. Then the formula

$$\begin{aligned} \mathrm{dom}(A_f) &:= \left\{ x \in H \,\middle|\, \int_{\mathbb{R}} f^2 \, d\mu_{x,x} < \infty \right\}, \\ \mathrm{Re}\langle y, A_f x\rangle &:= \int_{\mathbb{R}} f \, d\mu_{y,x} \qquad \text{for } x \in \mathrm{dom}(A_f) \text{ and } y \in H, \end{aligned} \tag{6.4.15}$$

defines a self-adjoint operator $A_f : \mathrm{dom}(A_f) \to H$. This operator satisfies the equation

$$\|A_f x\|^2 = \int_{\mathbb{R}} f^2 \, d\mu_{x,x}$$

for all $x \in \mathrm{dom}(A_f)$.

Proof. For $x, y \in H$ the function $\mu_{y,x} : \mathcal{B} \to \mathbb{R}$ is a signed Borel measure. Its **total variation** is the Borel measure $|\mu_{y,x}| : \mathcal{B} \to [0, \infty)$, defined by

$$|\mu_{y,x}|(\Omega) := \sup\left\{\mu_{y,x}(\Omega') - \mu_{y,x}(\Omega \setminus \Omega') \,\middle|\, \Omega' \in \mathcal{B},\, \Omega' \subset \Omega\right\}$$

for every Borel set $\Omega \subset \mathbb{R}$ (see [**75**, Thm. 5.12]). By definition, the total variation satisfies $|\mu_{y,x}(\Omega)| \le |\mu_{y,x}|(\Omega)$ for all $\Omega \in \mathcal{B}$. The positive and negative parts of $\mu_{y,x}$ are the Borel measures $\mu^{\pm}_{y,x} : \mathcal{B} \to [0, \infty)$, defined by

$$\mu^{\pm}_{y,x}(\Omega) := \frac{|\mu_{y,x}|(\Omega) \pm \mu_{y,x}(\Omega)}{2} \qquad \text{for } \Omega \in \mathcal{B}.$$

They satisfy

$$\mu_{y,x} = \mu^{+}_{y,x} - \mu^{-}_{y,x}, \qquad |\mu_{y,x}| = \mu^{+}_{y,x} + \mu^{-}_{y,x}.$$

Let us now fix a Borel measurable function $f : \mathbb{R} \to \mathbb{R}$. Then two vectors $x, y \in H$ satisfy $\int_{\mathbb{R}} |f| d|\mu_{y,x}| < \infty$ if and only if $\int_{\mathbb{R}} |f| d\mu^{\pm}_{y,x} < \infty$, and if this holds, then the integral of f with respect to $\mu_{y,x}$ is defined by

$$\int_{\mathbb{R}} f\, d\mu_{y,x} := \int_{\mathbb{R}} f\, d\mu^{+}_{y,x} - \int_{\mathbb{R}} f\, d\mu^{-}_{y,x}.$$

With this understood, we prove in eight steps that the operator A_f is well defined and self-adjoint and satisfies $\|A_f x\|^2 = \int_{\mathbb{R}} f^2\, d\mu_{x,x}$ for all $x \in \mathrm{dom}(A_f)$.

Step 1. *The signed Borel measures $\mu_{y,x}$ in* (6.4.14) *satisfy the inequality*

$$|\mu_{y,x}|(\Omega) \le \sqrt{\mu_{x,x}(\Omega)}\sqrt{\mu_{y,y}(\Omega)} \tag{6.4.16}$$

for all $x, y \in H$ and all $\Omega \in \mathcal{B}$.

Fix two elements $x, y \in H$. If $\Omega_1, \Omega_2 \in \mathcal{B}$ are disjoint and $\Omega_1 \cup \Omega_2 =: \Omega$, then

$$\begin{aligned} \|P_{\Omega_1} x\|^2 + \|P_{\Omega_1} x\|^2 &= \langle x, P_{\Omega_1} x\rangle + \langle x, P_{\Omega_2} x\rangle \\ &= \langle x, P_{\Omega} x\rangle \\ &= \mu_{x,x}(\Omega). \end{aligned}$$

By the Cauchy–Schwarz inequality, this implies

$$\begin{aligned} \mu_{y,x}(\Omega') - \mu_{y,x}(\Omega \setminus \Omega') &= \mathrm{Re}\langle P_{\Omega'} y, P_{\Omega'} x\rangle - \mathrm{Re}\langle P_{\Omega\setminus\Omega'} y, P_{\Omega\setminus\Omega'} x\rangle \\ &\le \|P_{\Omega'} x\|\,\|P_{\Omega'} y\| + \left\|P_{\Omega\setminus\Omega'} x\right\|\,\left\|P_{\Omega\setminus\Omega'} y\right\| \\ &\le \sqrt{\|P_{\Omega'} x\|^2 + \left\|P_{\Omega\setminus\Omega'} x\right\|^2}\sqrt{\|P_{\Omega'} y\|^2 + \left\|P_{\Omega\setminus\Omega'} y\right\|^2} \\ &= \sqrt{\mu_{x,x}(\Omega)}\sqrt{\mu_{y,y}(\Omega)} \end{aligned}$$

for every pair of Borel sets $\Omega' \subset \Omega \subset \mathbb{R}$. Fix a Borel set $\Omega \subset \mathbb{R}$ and take the supremum over all Borel sets $\Omega' \subset \Omega$ to obtain (6.4.16). This proves Step 1.

Step 2. *Let $g : \mathbb{R} \to \mathbb{R}$ be a Borel measurable function. Then*

$$\int_{\mathbb{R}} |g|\, d|\mu_{y,x}| \le \|y\| \sqrt{\int_{\mathbb{R}} g^2 \, d\mu_{x,x}} \tag{6.4.17}$$

for all $x, y \in H$.

For every finite collection of pairwise disjoint Borel sets $\Omega_1, \dots, \Omega_n \subset \mathbb{R}$ and every finite collection of positive real numbers $a_1, \dots, a_n$, we have

$$\sum_{i=1}^n a_i |\mu_{y,x}|(\Omega_i) \le \left(\sum_{i=1}^n a_i^2 \mu_{x,x}(\Omega_i)\right)^{1/2} \left(\sum_{i=1}^n \mu_{y,y}(\Omega_i)\right)^{1/2}$$

by Step 1 and the Cauchy–Schwarz inequality. Moreover,

$$\sum_{i=1}^n \mu_{y,y}(\Omega_i) = \mu_{y,y}\left(\bigcup_{i=1}^n \Omega_i\right) \le \|y\|^2 .$$

This proves (6.4.17) for the Borel measurable step function $g := \sum_{i=1}^n a_i \chi_{\Omega_i}$. Since every nonnegative Borel measurable function can be approximated pointwise from below by a sequence of Borel measurable step functions (see for example [**75**, Thm. 1.26]), Step 2 follows from the Lebesgue Monotone Convergence Theorem.

Step 3. *The operator $A_f : \mathrm{dom}(A_f) \to H$ in (6.4.15) is well defined. More precisely, fix an element $x \in \mathrm{dom}(A_f)$. Then the function $|f| : \mathbb{R} \to [0, \infty)$ is integrable with respect to the Borel measure $|\mu_{y,x}|$ for every $y \in H$, and there exists a unique element $A_f x \in H$ such that*

$$\mathrm{Re}\langle y, A_f x\rangle = \int_{\mathbb{R}} f \, d\mu_{y,x}$$

for all $y \in H$. Moreover, $\|A_f x\|^2 \le \int_{\mathbb{R}} f^2 \, d\mu_{x,x}$.

Fix an element $x \in \mathrm{dom}(A_f)$ and define $c := \left(\int_{\mathbb{R}} f^2 \, d\mu_{x,x}\right)^{1/2} < \infty$. Then Step 2 asserts that $\int_{\mathbb{R}} |f| d|\mu_{y,x}| \le c\,\|y\| < \infty$ and so the integral $\int_{\mathbb{R}} f d\mu_{y,x}$ is well defined for all $y \in H$. Now define the map $\Lambda_x : H \to \mathbb{R}$ by

$$\Lambda_x(y) := \int_{\mathbb{R}} f \, d\mu_{y,x} \qquad \text{for } y \in H.$$

This map is real linear and satisfies the inequality

$$|\Lambda_x(y)| \le \int_{\mathbb{R}} |f|\, d|\mu_{y,x}| \le \left(\int_{\mathbb{R}} f^2 \, d\mu_{x,x}\right)^{1/2} \|y\| = c\,\|y\|$$

for all $y \in H$ by Step 2. Hence, by Theorem 1.4.4, there exists a unique element $A_f x \in H$ such that $\mathrm{Re}\langle y, A_f x\rangle = \int_{\mathbb{R}} f \, d\mu_{y,x}$ for all $y \in H$. Moreover $\|A_f x\| = \|\Lambda_x\| \le c$ and this proves Step 3.

Step 4. *The set* $\mathrm{dom}(A_f) \subset H$ *is a complex linear subspace and the operator* A_f *in* (6.4.15) *is complex linear and symmetric.*

Let $x, x' \in \mathrm{dom}(A_f)$. Then

$$\begin{aligned} \mu_{x+x',x+x'}(\Omega) &= \langle x + x', P_\Omega x + P_\Omega x' \rangle \\ &= \|P_\Omega x\|^2 + 2\mathrm{Re}\langle P_\Omega x', P_\Omega x\rangle + \|P_\Omega x'\|^2 \\ &\le 2\|P_\Omega x\|^2 + 2\|P_\Omega x'\|^2 \\ &= 2\mu_{x,x}(\Omega) + 2\mu_{x',x'}(\Omega) \end{aligned}$$

for all $\Omega \in \mathcal{B}$ and this implies $x+x' \in \mathrm{dom}(A_f)$. Moreover, $\mu_{\lambda x,\lambda x} = |\lambda|^2\mu_{x,x}$, so $\lambda x \in \mathrm{dom}(A_f)$ for all $\lambda \in \mathbb{C}$. Thus $\mathrm{dom}(A_f)$ is a complex subspace of H. Since $\mu_{y,x+x'} = \mu_{y,x} + \mu_{y,x'}$ and $\mu_{y,\lambda x} = \lambda\mu_{y,x}$ for all $x, x' \in \mathrm{dom}(A_f)$ and all $\lambda \in \mathbb{R}$, the operator A_f is real linear. To prove that it is complex linear, let $x \in \mathrm{dom}(A_f)$ and $y \in H$. Then $\mu_{y,\mathbf{i}x} = -\mu_{\mathbf{i}y,x}$ and hence

$$\mathrm{Re}\langle y, A_f\mathbf{i}x\rangle = \int_{\mathbb{R}} f\,d\mu_{y,\mathbf{i}x} = -\int_{\mathbb{R}} f\,d\mu_{\mathbf{i}y,x} = -\mathrm{Re}\langle \mathbf{i}y, A_f x\rangle = \mathrm{Re}\langle y, \mathbf{i}A_f x\rangle.$$

This shows that $A_f\mathbf{i}x = \mathbf{i}A_f x$ for all $x \in \mathrm{dom}(A_f)$, so A_f is complex linear. Moreover, A_f is symmetric because the bilinear map

$$\mathrm{dom}(A_f) \times \mathrm{dom}(A_f) \to \mathcal{M}(\mathbb{R}) : (x, y) \mapsto \mu_{x,y}$$

is symmetric. This proves Step 4.

Step 5. *The operator* $A_f : \mathrm{dom}(A_f) \to H$ *in* (6.4.15) *has a dense domain.*

For $n \in \mathbb{N}$ define $\Omega_n := \{\lambda \in \mathbb{R} \,|\, |f(\lambda)| \le n\}$. Then $\mathbb{R} = \bigcup_{n=1}^\infty \Omega_n$. Hence it follows from the (σ-Additive) and (Normalization) axioms in Definition 5.6.1 that $\lim_{n\to\infty} P_{\Omega_n} x = x$ for all $x \in H$. Now let $x \in H$ and define $x_n := P_{\Omega_n} x$. Then $\mu_{x_n,x_n}(\Omega) = \mu_{x,x}(\Omega \cap \Omega_n)$ for all $\Omega \in \mathcal{B}$ by the (Intersection) axiom in Definition 5.6.1. Hence

$$\int_{\mathbb{R}} f^2\,d\mu_{x_n,x_n} = \int_{\Omega_n} f^2\,d\mu_{x,x} \le n^2\|x\|^2$$

and so $x_n \in \mathrm{dom}(A_f)$ for all $n \in \mathbb{N}$. This proves Step 5.

Step 6. *Let* $x \in \mathrm{dom}(A_f)$ *and* $\Omega \in \mathcal{B}$. *Then* $P_\Omega x \in \mathrm{dom}(A_f)$ *and*

$$A_f P_\Omega x = P_\Omega A_f x.$$

The estimate $\int_{\mathbb{R}} f^2\,d\mu_{P_\Omega x, P_\Omega x} = \int_\Omega f^2\,d\mu_{x,x} < \infty$ implies $P_\Omega x \in \mathrm{dom}(A_f)$. Moreover,

$$\mathrm{Re}\langle y, A_f P_\Omega x\rangle = \int_{\mathbb{R}} f\,d\mu_{y,P_\Omega x} = \int_{\mathbb{R}} f\,d\mu_{P_\Omega y,x} = \mathrm{Re}\langle P_\Omega y, A_f x\rangle$$

for all $y \in H$ and this proves Step 6.

Step 7. *Let* $x \in \mathrm{dom}(A_f)$. *Then* f *is integrable with respect to the Borel measure* $|\mu_{x,A_f x}|$ *and*

$$\int_{\mathbb{R}} f^2\, d\mu_{x,x} = \int_{\mathbb{R}} f\, d\mu_{x,A_f x} = \|A_f x\|^2. \tag{6.4.18}$$

That f is integrable with respect to $|\mu_{x,A_f x}| = |\mu_{A_f x,x}|$ was proved in Step 3. Next we observe that

$$\int_{\mathbb{R}} \chi_\Omega\, d\mu_{x,A_f x} = \mu_{x,A_f x}(\Omega) = \mathrm{Re}\langle P_\Omega x, A_f x\rangle = \int_{\mathbb{R}} f\, d\mu_{P_\Omega x,x} = \int_{\mathbb{R}} \chi_\Omega f\, d\mu_{x,x}$$

for every Borel set $\Omega \subset \mathbb{R}$. This shows that $\int_{\mathbb{R}} g\, d\mu_{x,A_f x} = \int_{\mathbb{R}} gf\, d\mu_{x,x}$ for every Borel measurable step function $g : \mathbb{R} \to \mathbb{R}$. Now approximate f pointwise by a sequence of Borel measurable step functions $g_n : \mathbb{R} \to \mathbb{R}$ such that $|g_n(\lambda)| \le |f(\lambda)|$ for all $n \in \mathbb{N}$ and all $\lambda \in \mathbb{R}$ (see [**75**, Thm. 1.26]). Then the Lebesgue Dominated Convergence Theorem asserts that

$$\int_{\mathbb{R}} f\, d\mu_{x,A_f x} = \lim_{n\to\infty} \int_{\mathbb{R}} g_n\, d\mu_{x,A_f x} = \lim_{n\to\infty} \int_{\mathbb{R}} g_n f\, d\mu_{x,x} = \int_{\mathbb{R}} f^2 d\mu_{x,x}.$$

This proves the first equality in (6.4.18). The second equality follows from Step 3 and this proves Step 7.

Step 8. *The operator* $A_f : \mathrm{dom}(A_f) \to H$ *in* (6.4.15) *is self-adjoint.*

By Step 4 it suffices to prove that $\mathrm{dom}(A_f^*) \subset \mathrm{dom}(A_f)$. Let $x \in \mathrm{dom}(A_f^*)$ and define $y := A_f^* x$. Then, for all $\xi \in \mathrm{dom}(A_f)$, we have

$$\int_{\mathbb{R}} f d\mu_{x,\xi} = \mathrm{Re}\langle x, A_f \xi\rangle = \mathrm{Re}\langle A_f^* x, \xi\rangle = \mathrm{Re}\langle y, \xi\rangle \tag{6.4.19}$$

by Step 3. For $n \in \mathbb{N}$ let $\Omega_n := \{\lambda \in \mathbb{R} \mid |f(\lambda)| \le n\}$ and $x_n := P_{\Omega_n} x$ as in the proof of Step 5. Then

$$\int_{\mathbb{R}} f^2\, d\mu_{x,x} = \lim_{n\to\infty} \int_{\Omega_n} f^2\, d\mu_{x,x} = \lim_{n\to\infty} \int_{\mathbb{R}} f^2\, d\mu_{x_n,x_n}$$

by the Lebesgue Monotone Convergence Theorem. Moreover, it follows from Steps 5 and 6 that $A_f x_n = A_f P_{\Omega_n} x_n = P_{\Omega_n} A_f x_n \in \mathrm{dom}(A_f)$. Hence

$$\int_{\mathbb{R}} f^2\, d\mu_{x_n,x_n} = \int_{\mathbb{R}} f\, d\mu_{x,A_f x_n} = \mathrm{Re}\langle y, A_f x_n\rangle \le \|y\| \sqrt{\int_{\mathbb{R}} f^2\, d\mu_{x_n,x_n}}.$$

Here the first equality uses Step 7 and the fact that the signed Borel measures $\mu_{x_n,A_f x_n}$ and $\mu_{x,A_f x_n}$ agree, the second equality follows from (6.4.19) with $\xi := A_f x_n \in \mathrm{dom}(A_f)$, and the inequality follows from Step 3. Thus

$$\int_{\mathbb{R}} f^2\, d\mu_{x,x} = \lim_{n\to\infty} \int_{\mathbb{R}} f^2\, d\mu_{x_n,x_n} \le \|y\|^2$$

and so $x \in \mathrm{dom}(A_f)$. This proves Step 8 and Theorem 6.4.4. □

Remark 6.4.5. **(i)** Theorem 6.4.4 can be used to extend the functional calculus for self-adjoint operators to unbounded functions $f : \mathbb{R} \to \mathbb{R}$, starting from a projection valued measure as in Theorem 5.6.2. This functional calculus can then be used to prove that the operator $A_f + \mathbf{i}\mathbb{1}$ is invertible and thus gives rise to an alternative proof that A_f is self-adjoint. This approach is used in Kato [**44**, p. 355]. Steps 6 and 7 in the above proof of Theorem 6.4.4 can be understood as a special case of this functional calculus, using one unbounded function f and the bounded functions χ_Ω for $\Omega \in \mathcal{B}$.

(ii) There is an entirely different approach to the measurable functional calculus for unbounded self-adjoint operators. One can start by assigning to an unbounded self-adjoint operator A its spectral measure and use Theorem 5.6.2 to construct the C* algebra homomorphism $\Psi_A : B(\Sigma) \to \mathcal{L}^c(H)$. For the construction of the spectral measure one can proceed as follows. First show that every self-adjoint operator $A : \mathrm{dom}(A) \to H$ can be written as a difference $A = A^+ - A^-$ of two positive semidefinite self-adjoint operators $A^\pm : \mathrm{dom}(A^\pm) \to H$ with $\mathrm{dom}(A^+) \cap \mathrm{dom}(A^-) = \mathrm{dom}(A)$. Then the operators $\mathbb{1} + A^\pm$ are invertible by Theorem 6.3.13 and one can use the spectral measures of their inverses in Theorem 5.6.3 to find the spectral measure for A. This approach is taken in Kato [**44**, pp. 353–361]. It does not require the functional calculus for normal operators in Section 5.5.

(iii) Suppose the projection valued measure is supported on a closed subset $\Sigma \subset \mathbb{R}$. Then the functional calculus for unbounded functions can be used as in Step 5 of the proof of Theorem 5.6.2 to show that $\sigma(A_f) \subset \overline{f(\Sigma)}$.

(iv) The functional calculus extends to unbounded normal operators. It can be reduced to the self-adjoint case by writing an unbounded normal operator as $A = A_1 + \mathbf{i}A_2$ where A_1 and A_2 are self-adjoint (Theorem 6.3.11). For bounded normal operators this approach is outlined in [**72**, pp. 245–247].

The next theorem shows that (6.4.13), (6.4.14), (6.4.15) give rise to a one-to-one correspondence between projection valued measures on the real axis with values in $\mathcal{L}^c(H)$ and unbounded self-adjoint operators on H.

Theorem 6.4.6 (Spectral Measures). *Let H be a nonzero complex Hilbert space and let $\mathcal{B} \subset 2^{\mathbb{R}}$ be the Borel σ-algebra.*

(i) *Let $A : \mathrm{dom}(A) \to H$ be a self-adjoint operator and let $\{P_\Omega\}_{\Omega \in \mathcal{B}}$ be the spectral measure of A in Definition 6.4.3. Then $A = A_{\mathrm{id}}$ is the operator in Theorem 6.4.4 with $f = \mathrm{id} : \mathbb{R} \to \mathbb{R}$.*

(ii) *Let $\mathcal{B} \to \mathcal{L}^c(H) : \Omega \mapsto P_\Omega$ be a projection valued measure and let A_{id} be the operator in Theorem 6.4.4 with $f = \mathrm{id} : \mathbb{R} \to \mathbb{R}$. Then $\{P_\Omega\}_{\Omega \in \mathcal{B}}$ is the spectral measure of A_{id} in Definition 6.4.3.*

Proof. See page 338. □

Corollary 6.4.7 (Characterization of the Spectral Measure). *Let $A : \mathrm{dom}(A) \to H$ be a self-adjoint operator on a nonzero complex Hilbert space H. Then there exists a unique projection valued measure $\{P_\Omega\}_{\Omega\in\mathcal{B}}$ on the real axis such that*

$$(6.4.20)\qquad \begin{aligned} \mathrm{dom}(A) &= \left\{ x \in H \,\middle|\, \int_{\mathbb{R}} \lambda^2\, d\mu_{x,x}(\lambda) < \infty \right\}, \\ \mathrm{Re}\langle y, Ax\rangle &= \int_{\mathbb{R}} \lambda\, d\mu_{y,x}(\lambda) \qquad \text{for } x \in \mathrm{dom}(A) \text{ and } y \in H, \end{aligned}$$

where $\{\mu_{y,x}\}_{x,y\in H}$ is the collection of signed Borel measures on the real axis defined by $\mu_{y,x}(\Omega) := \mathrm{Re}\,\langle y, P_\Omega x\rangle$ for all $x, y \in H$ and all Borel sets $\Omega \subset \mathbb{R}$. It agrees with the spectral measure of Definition 6.4.3.

Proof. Uniqueness follows from part (ii) of Theorem 6.4.6, and existence follows from Theorem 6.4.4 and part (i) of Theorem 6.4.6. □

Proof of Theorem 6.4.6. We prove part (i). Let $A : \mathrm{dom}(A) \to H$ be an unbounded self-adjoint operator with spectrum

$$\Sigma := \sigma(A)$$

and take $\{P_\Omega\}_{\Omega\in\mathcal{B}}$ to be the projection valued measure in Definition 6.4.3, associated to the C* algebra homomorphism $\Psi_A : B(\Sigma) \to \mathcal{L}^c(H)$ in Theorem 6.4.1. For $i \in \mathbb{N}$ define the function $f_i : \mathbb{R} \to \mathbb{R}$ by

$$f_i(\lambda) := \begin{cases} \lambda, & \text{if } |\lambda| \le i, \\ 0, & \text{if } |\lambda| > i. \end{cases}$$

Then the (Normalization) axiom in Theorem 6.4.1 asserts that

$$(6.4.21)\qquad \lim_{i\to\infty} \Psi_A(f_i|_\Sigma)x = Ax \qquad \text{for all } x \in \mathrm{dom}(A).$$

Moreover, by definition of P_Ω in (6.4.12) and of $\mu_{y,x}$ in (6.4.14), we have

$$\mu_{y,x}(\Omega) = \mathrm{Re}\langle y, P_\Omega x\rangle = \mathrm{Re}\langle y, \Psi_A(\chi_{\Sigma\cap\Omega})x\rangle$$

for all $x, y \in H$ and all $\Omega \in \mathcal{B}$. Hence the (Convergence) axiom implies

$$\int_{\mathbb{R}} f\, d\mu_{y,x} = \mathrm{Re}\langle y, \Psi_A(f|_\Sigma)x\rangle$$

for all $x, y \in H$ and all bounded Borel measurable functions $f : \mathbb{R} \to \mathbb{R}$. In particular,

$$(6.4.22)\qquad \int_{\mathbb{R}} f_i\, d\mu_{y,x} = \mathrm{Re}\langle y, \Psi_A(f_i|_\Sigma)x\rangle, \qquad \int_{\mathbb{R}} f_i^2\, d\mu_{x,x} = \|\Psi_A(f_i|_\Sigma)x\|^2$$

for all $i \in \mathbb{N}$ and all $x, y \in H$.

Now let $x \in \mathrm{dom}(A)$. Then, by equations (6.4.21) and (6.4.22) and the Lebesgue Monotone Convergence Theorem, we have

$$\int_{\mathbb{R}} \lambda^2 \, d\mu_{x,x}(\lambda) = \lim_{i\to\infty} \int_{\mathbb{R}} f_i^2 \, d\mu_{x,x} = \lim_{i\to\infty} \|\Psi_A(f_i|_\Sigma)x\|^2 = \|Ax\|^2 .$$

This implies $x \in \mathrm{dom}(A_{\mathrm{id}})$ and hence, by equations (6.4.21) and (6.4.22) and the Lebesgue Dominated Convergence Theorem,

$$\begin{aligned} \mathrm{Re}\langle y, A_{\mathrm{id}}x\rangle &= \int_{\mathbb{R}} \lambda \, d\mu_{y,x} \\ &= \lim_{i\to\infty} \int_{\mathbb{R}} f_i \, d\mu_{y,x} \\ &= \lim_{i\to\infty} \mathrm{Re}\langle y, \Psi_A(f_i|_\Sigma)x\rangle \\ &= \mathrm{Re}\langle y, Ax\rangle \end{aligned}$$

for all $y \in H$. Thus $\mathrm{dom}(A) \subset \mathrm{dom}(A_{\mathrm{id}})$ and $A_{\mathrm{id}}|_{\mathrm{dom}(A)} = A$. This implies $A_{\mathrm{id}} = A$ by Exercise 6.5.4 and proves part (i).

We prove part (ii). Thus let $\mathcal{B} \to \mathcal{L}^c(H) : \Omega \mapsto P_\Omega$ be a projection valued measure on the real axis, let

$$A := A_{\mathrm{id}}$$

be the operator in Theorem 6.4.4 with $f = \mathrm{id}$, and let $\Psi : B(\mathbb{R}) \to \mathcal{L}^c(H)$ be the C* algebra homomorphism in Theorem 5.6.2 associated to $\{P_\Omega\}_{\Omega\in\mathcal{B}}$. Then Ψ satisfies the (Convergence) axiom in Theorem 6.4.1 by definition. We prove that

$$P_{\mathbb{R}\setminus\Sigma} = 0, \qquad \Sigma := \sigma(A_{\mathrm{id}}). \tag{6.4.23}$$

Suppose, by contradiction, that $P_{\mathbb{R}\setminus\Sigma} \neq 0$, choose a vector $x \in X$ such that $P_{\mathbb{R}\setminus\Sigma}x \neq 0$, and consider the Borel measure $\mu_x : \mathcal{B} \to [0,\infty)$ defined by $\mu_x(\Omega) := \langle x, P_\Omega x\rangle$ for $\Omega \in \mathcal{B}$. Then $\mu_x(\mathbb{R}\setminus\Sigma) > 0$ and so, since every Borel measure on $\mathbb{R}$ is inner regular by [**75**, Thm. 3.18], there exists a compact set $K \subset \mathbb{R} \setminus \Sigma$ such that $\mu_x(K) > 0$. Hence $P_K \neq 0$ and so

$$E_K := \mathrm{im}(P_K)$$

is a nonzero closed subspace of H. Since the identity function $f = \mathrm{id} : \mathbb{R} \to \mathbb{R}$ is bounded on K, it follows from the definition of the operator $A = A_{\mathrm{id}}$ in (6.4.15) that $E_K \subset \mathrm{dom}(A)$ and E_K is invariant under A. Since $E_K \neq \{0\}$ and the operator $A_K := A|_{E_K} : E_K \to E_K$ is self-adjoint, its spectrum is nonempty. Since $\mu_{y,x}(\Omega) = \mu_{y,x}(\Omega \cap K)$ for $x, y \in E_K$ and $\Omega \in \mathcal{B}$, we have

$$\mathrm{Re}\langle y, A_K x\rangle = \int_{\mathbb{R}} \lambda \, d\mu_{y,x}(\lambda) = \int_K \lambda \, d\mu_{y,x}(\lambda) \qquad \text{for all } x, y \in E_K.$$

Hence $\sigma(A_K) \subset K$ by Theorem 5.6.2 and so $\emptyset \neq \sigma(A_K) \subset \sigma(A) \cap K = \emptyset$, a contradiction. This proves (6.4.23).

Since $P_{\mathbb{R}\setminus\Sigma} = 0$, the C* algebra homomorphism Ψ of Theorem 5.6.2 descends to a unique C* algebra homomorphism

$$\Psi_\Sigma : B(\Sigma) \to \mathcal{L}^c(H)$$

such that $\Psi(f) = \Psi_\Sigma(f|_\Sigma)$ for all $f \in B(\mathbb{R})$. We prove that Ψ_Σ satisfies the (Normalization) axiom in Theorem 6.4.1 with $A = A_{\mathrm{id}}$. To see this, choose a sequence of bounded Borel measurable functions $f_i : \mathbb{R} \to \mathbb{R}$ that satisfy

$$\sup_{i\in\mathbb{N}}|f_i(\lambda)| \le |\lambda|, \qquad \lim_{i\to\infty} f_i(\lambda) = \lambda \qquad \text{for all } \lambda \in \mathbb{R}.$$

Fix an element $x \in \mathrm{dom}(A_{\mathrm{id}})$. Then the identity function $\mathrm{id} : \mathbb{R} \to \mathbb{R}$ is square integrable with respect to the Borel measure $\mu_{x,x}$, and is integrable with respect to the Borel measure $|\mu_{y,x}|$ for every $y \in H$. Hence it follows from the Lebesgue Dominated Convergence Theorem and the Hahn Decomposition Theorem that

$$\begin{aligned}\langle y, A_{\mathrm{id}}x\rangle &= \int_{\mathbb{R}} \lambda \, d\mu_{y,x}(\lambda) \\ &= \lim_{i\to\infty} \int_{\mathbb{R}} f_i \, d\mu_{y,x} \\ &= \lim_{i\to\infty} \langle y, \Psi_\Sigma(f_i|_\Sigma)x\rangle\end{aligned}$$

for all $y \in H$ and

$$\begin{aligned}\|A_{\mathrm{id}}x\|^2 &= \int_{\mathbb{R}} \lambda^2 \, d\mu_{x,x}(\lambda) \\ &= \lim_{i\to\infty} \int_{\mathbb{R}} f_i^2 \, d\mu_{x,x} \\ &= \lim_{i\to\infty} \|\Psi_\Sigma(f_i|_\Sigma)x\|^2 .\end{aligned}$$

Hence the sequence $\Psi_\Sigma(f_i|_\Sigma)x$ converges weakly to $A_{\mathrm{id}}x$ and its norm converges to that of $A_{\mathrm{id}}x$. By Exercise 3.7.1 this implies

$$\lim_{i\to\infty} \|A_{\mathrm{id}}x - \Psi_\Sigma(f_i|_\Sigma)x\| = 0.$$

Thus the reduced C* algebra homomorphism $\Psi_\Sigma : B(\Sigma) \to \mathcal{L}^c(H)$ satisfies the (Normalization) axiom in Theorem 6.4.1 with $A = A_{\mathrm{id}}$. Hence it follows from uniqueness in Theorem 6.4.1 that $\Psi_\Sigma = \Psi_{A_{\mathrm{id}}}$ is the functional calculus associated to the self-adjoint operator A_{id}. Hence

$$P_\Omega = \Psi(\chi_\Omega) = \Psi_\Sigma(\chi_\Omega|_\Sigma) = \Psi_{A_{\mathrm{id}}}(\chi_\Omega|_\Sigma)$$

for every Borel set $\Omega \subset \mathbb{R}$. Here the first equality holds by definition of the C* algebra homomorphism $\Psi : B(\mathbb{R}) \to \mathcal{L}^c(H)$ in Theorem 5.6.2. Hence the projection valued measure $\{P_\Omega\}_{\Omega\in\mathcal{B}}$ is the spectral measure of A_{id} as introduced in Definition 6.4.3. This proves part (ii) and Theorem 6.4.6. □

Example 6.4.8. Let $A : \mathrm{dom}(A) \to H$ be a self-adjoint operator on a nonzero complex Hilbert space H.

(i) Consider the operator family

$$\mathbb{R} \to \mathcal{L}(H) : t \mapsto U(t)$$

associated to the functions $\lambda \mapsto e^{\mathbf{i}\lambda t}$ via the functional calculus of Theorem 6.4.1. In terms of the spectral measure the operators $U(t)$ are determined by the formula

$$\langle y, U(t)x\rangle := \int_{-\infty}^{\infty} e^{\mathbf{i}\lambda t}\, d\langle y, P_\lambda x\rangle \qquad \text{for all } x, y \in H \text{ and all } t \in \mathbb{R}.$$

Here the expression $\int_{\mathbb{R}} f(\lambda)\, d\langle y, P_\lambda x\rangle$ denotes the integral of a Borel measurable function $f : \mathbb{R} \to \mathbb{C}$ with respect to the complex valued Borel measure

$$\mathcal{B} \to \mathbb{C} : \Omega \mapsto \langle y, P_\Omega x\rangle$$

on the real axis. The operator family $\mathbb{R} \to \mathcal{L}^c(H) : t \mapsto U(t)$ is strongly continuous, by the (Convergence) axiom, and satisfies

$$U(s+t) = U(s)U(t), \qquad U(0) = \mathbb{1}$$

for all $s, t \in \mathbb{R}$. This means that U is a *strongly continuous group* of (unitary) operators. Such groups play an important role in quantum mechanics. For example, they appear as solutions of the Schrödinger equation.

(ii) Assume, in addition, that

$$\langle x, Ax\rangle \le 0 \qquad \text{for all } x \in \mathrm{dom}(A).$$

Then $\sigma(A) \subset (-\infty, 0]$ and a similar construction leads to an operator family

$$[0, \infty) \to \mathcal{L}^c(H) : t \mapsto S(t)$$

associated to the functions $\lambda \mapsto e^{\lambda t}$ on the negative real axis. In terms of the spectral measure the operators $S(t)$ are determined by the formula

$$\langle y, S(t)x\rangle := \int_{-\infty}^{0} e^{\lambda t}\, d\langle y, P_\lambda x\rangle \qquad \text{for all } x, y \in H \text{ and all } t \ge 0.$$

The restriction $t \ge 0$ is needed to obtain bounded functions on the negative real axis and bounded linear operators $S(t)$. These operators form a *strongly continuous semigroup* of operators on H. For example, the solutions of the heat equation on $\mathbb{R}^n$ can be expressed in this form with A the Laplace operator. The study of strongly continuous semigroups is the subject of the next and final chapter of this book.

of strongly continuous semigroups. The dual semigroup is the subject of Section 7.3 and analytic semigroups are discussed in Section 7.4. A preparatory Section 7.5 is devoted to Banach space valued measurable functions, and inhomogeneous equations are examined in Section 7.6.

7.1. Strongly Continuous Semigroups

7.1.1. Definition and Examples. The existence and uniqueness theorem for solutions of a time-independent ordinary differential equation implies that the solutions define a flow. This means that the value of the solution with initial condition x_0 at time $s+t$ agrees with the value at time s of the solution whose initial condition is taken to be the value of the original solution at time t. For linear differential equations on Banach spaces this translates into a semigroup condition on the family of linear operators, parametrized by a nonnegative real variable t, that assign to a given initial condition the solution of the respective linear differential equation at time t. Continuous dependence on time translates into strong continuity of the semigroup of operators and continuous dependence on the initial condition translates into boundedness of the operators.

Definition 7.1.1 (Strongly Continuous Semigroup). Let X be a real Banach space. A **one-parameter semigroup (of operators on X)** is a map $S:[0,\infty)\to\mathcal{L}(X)$ that satisfies

$$S(0)=\mathbb{1},\qquad S(s+t)=S(s)S(t) \tag{7.1.1}$$

for all $s,t\ge 0$. A **one-parameter group (of operators on X)** is a map $S:\mathbb{R}\to\mathcal{L}(X)$ that satisfies (7.1.1) for all $s,t\in\mathbb{R}$. A **strongly continuous semigroup (of operators on X)** is a map $S:[0,\infty)\to\mathcal{L}(X)$ that satisfies (7.1.1) for all $s,t\ge 0$ and satisfies

$$\lim_{t\to 0}\|S(t)x-x\|=0 \tag{7.1.2}$$

for all $x\in X$. A **strongly continuous group (of operators on X)** is a map $S:\mathbb{R}\to\mathcal{L}(X)$ that satisfies (7.1.1) for all $s,t\in\mathbb{R}$ and satisfies (7.1.2) for all $x\in X$.

Example 7.1.2 (Groups Generated by Bounded Operators). Let X be a real Banach space and let $A:X\to X$ be a bounded linear operator. Then the operators

$$S(t):=e^{tA}=\sum_{k=0}^{\infty}\frac{t^kA^k}{k!} \tag{7.1.3}$$

for $t\in\mathbb{R}$ form a strongly continuous group of operators on X. In this example the map $\mathbb{R}\to\mathcal{L}(X):t\mapsto S(t)$ is continuous with respect to the norm topology on $\mathcal{L}(X)$ (see Exercise 5.2.13).

Example 7.1.3 (Semigroups and Orthonormal Bases). Let H be a separable complex Hilbert space, let $(e_i)_{i\in\mathbb{N}}$ be a complex orthonormal basis, and let $(\lambda_i)_{i\in\mathbb{N}}$ be a sequence of complex numbers such that

$$\sup_{i\in\mathbb{N}} \mathrm{Re}\lambda_i < \infty.$$

Define the map $S : [0,\infty) \to \mathcal{L}^c(H)$ by

$$S(t)x := \sum_{i=1}^{\infty} e^{\lambda_i t}\langle e_i, x\rangle e_i \tag{7.1.4}$$

for $x \in H$ and $t \geq 0$. **Exercise:** Show that this is a strongly continuous semigroup of operators on H. Show that it extends to a strongly continuous group $S : \mathbb{R} \to \mathcal{L}^c(H)$ if and only if

$$\sup_{i\in\mathbb{N}} |\mathrm{Re}\lambda_i| < \infty.$$

Example 7.1.4 (Shift Semigroups). Fix a constant $1 \leq p < \infty$ and let $X = L^p([0,\infty))$ be the Banach space of real valued L^p-functions on $[0,\infty)$ with respect to the Lebesgue measure.

(i) Define the map $L : [0,\infty) \to \mathcal{L}(X)$ by

$$(L(t)f)(s) := f(s+t) \tag{7.1.5}$$

for $f \in L^p([0,\infty))$ and $s,t \geq 0$. **Exercise:** Show that this is a strongly continuous semigroup of operators. Show that this example extends to the space of continuous functions on $[0,\infty)$ that converge to zero at infinity. Show that strong continuity fails when $L^p([0,\infty))$ is replaced by $L^\infty([0,\infty))$ or by the space of bounded continuous real valued functions on $[0,\infty)$. Show that the formula (7.1.5) defines a group on $L^p(\mathbb{R})$ for $1 \leq p < \infty$.

(ii) Define the map $R : [0,\infty) \to \mathcal{L}(X)$ by

$$(R(t)f)(s) := \begin{cases} 0, & \text{if } s < t, \\ f(s-t), & \text{if } s \geq t, \end{cases} \tag{7.1.6}$$

for $f \in L^p([0,\infty))$ and $s,t \geq 0$. **Exercise:** Show that this is a strongly continuous semigroup of isometric embeddings. Show that this example extends to the space of continuous functions $f : [0,\infty) \to \mathbb{R}$ that vanish at the origin and converge to zero at infinity.

(iii) Define the map $S : [0,\infty) \to \mathcal{L}(L^p([0,1]))$ by

$$(S(t)f)(s) := \begin{cases} f(s+t), & \text{if } s+t \leq 1, \\ 0, & \text{if } s+t > 1, \end{cases} \tag{7.1.7}$$

for $f \in L^p([0,1])$, $s \in [0,1]$, and $t \geq 0$. **Exercise:** Show that this is a strongly continuous semigroup of operators such that $S(t) = 0$ for $t \geq 1$.

Example 7.1.5 (Flows). Let (M,d) be a compact metric space and suppose that the map

$$\mathbb{R} \times M \to M : (t,p) \mapsto \phi_t(p)$$

is a **continuous flow**, i.e. it is continuous and satisfies

$$\phi_0 = \mathrm{id}, \qquad \phi_{s+t} = \phi_s \circ \phi_t$$

for all $s,t \in \mathbb{R}$. Let $X := C(M)$ be the Banach space of continuous real valued functions on M equipped with the supremum norm. Define

$$S(t)f := f \circ \phi_t \qquad \text{for } t \in \mathbb{R} \text{ and } f \in C(M). \tag{7.1.8}$$

Then $S : \mathbb{R} \to \mathcal{L}(C(M))$ is a strongly continuous group of operators.

Example 7.1.6 (Heat Equation). Fix a positive integer n and a real number $1 \le p < \infty$. Define the **heat kernel** $K_t : \mathbb{R}^n \to \mathbb{R}$ by

$$K_t(x) := \frac{1}{(4\pi t)^{n/2}} e^{-|x|^2/4t} \qquad \text{for } x \in \mathbb{R}^n \text{ and } t > 0. \tag{7.1.9}$$

Here $|x| := \sqrt{\sum_{i=1}^n x_i^2}$ denotes the Euclidean norm of $x = (x_1, \dots, x_n) \in \mathbb{R}^n$. These functions are nonnegative and Lebesgue integrable and satisfy

$$\int_{\mathbb{R}^n} K_t(\xi)\, d\xi = 1, \qquad K_{s+t} = K_s * K_t \tag{7.1.10}$$

for all $s,t > 0$, where $(f*g)(x) := \int_{\mathbb{R}^n} f(x-\xi)g(\xi)\,d\xi$ denotes the convolution of two Lebesgue integrable functions $f,g : \mathbb{R}^n \to \mathbb{R}$. Equation (7.1.10) implies that the operators $S(t) : L^p(\mathbb{R}^n) \to L^p(\mathbb{R}^n)$, defined by

$$S(t)f := \begin{cases} K_t * f, & \text{for } t > 0, \\ f, & \text{for } t = 0, \end{cases} \tag{7.1.11}$$

define a semigroup of operators. Since $\lim_{t\to 0} \sup_{|x|\ge\delta} K_t(x) = 0$ for all $\delta > 0$ and $\int_{\mathbb{R}^n} K_t = 1$ for all $t > 0$, the functions $S(t)f = K_t * f$ converge uniformly to f for every continuous function $f : \mathbb{R}^n \to \mathbb{R}$ with compact support. The convergence is also in $L^p(\mathbb{R}^n)$. Since $C_c(\mathbb{R}^n)$ is dense in $L^p(\mathbb{R}^n)$ by [**75**, Thm. 4.15] and $\|S(t)\| \le 1$ for all $t \ge 0$ by Young's inequality, it follows from Theorem 2.1.5 that $\lim_{t\to 0} \|S(t)f - f\|_{L^p} = 0$ for all $f \in L^p(\mathbb{R}^n)$. Thus the semigroup (7.1.11) is strongly continuous. Moreover, for each $f \in L^p(\mathbb{R}^n)$, the function $u : (0,\infty) \times \mathbb{R}^n \to \mathbb{R}$, defined by $u(t,x) := (K_t * f)(x)$ for $t > 0$ and $x \in \mathbb{R}^n$, is smooth and satisfies the **heat equation**

$$\frac{\partial u}{\partial t} = \sum_{i=1}^n \frac{\partial^2 u}{\partial x_i^2}, \qquad \lim_{t\to 0} \int_{\mathbb{R}^n} |u(t,x) - f(x)|^p\, dx = 0. \tag{7.1.12}$$

Exercise: Fill in the details.

Example 7.1.7 (Wave Equation). Let $L^2(\mathbb{R})$ be the space of square integrable real valued functions on $\mathbb{R}$ with respect to the Lebesgue measure, modulo equality almost everywhere, and let $W^{1,2}(\mathbb{R})$ denote the space of absolutely continuous functions $f : \mathbb{R} \to \mathbb{R}$ such that f and f' are square integrable. Then $H := W^{1,2}(\mathbb{R}) \times L^2(\mathbb{R})$ is a Hilbert space with the norm

$$\|(f,g)\|_H := \sqrt{\int_{-\infty}^{\infty} \left(|f(x)|^2 + \left|\frac{df}{dx}(x)\right|^2 + |g(x)|^2 \right) dx}$$

for $f \in W^{1,2}(\mathbb{R})$ and $g \in L^2(\mathbb{R})$. Given a pair $(f,g) \in H$, define the function $u : \mathbb{R}^2 \to \mathbb{R}$ by

$$u(t,x) = \frac{f(x+t) + f(x-t)}{2} + \frac{1}{2}\int_{x-t}^{x+t} g(s)\, ds \tag{7.1.13}$$

for $t, x \in \mathbb{R}$. Then $u(t,\cdot) \in W^{1,2}(\mathbb{R})$ and $\partial_t u(t,\cdot) \in L^2(\mathbb{R})$ for all $t \in \mathbb{R}$, and the linear operators $S(t) : H \to H$ given by $S(t)(f,g) := (u(t,\cdot), \partial_t u(t,\cdot))$ for $(f,g) \in H$ and $t \in \mathbb{R}$ define a strongly continuous group of operators on H. If f and g are smooth, then the function (7.1.13) is the unique solution of the one-dimensional **wave equation**

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2}, \qquad u(0,x) = f(x), \qquad \frac{\partial u}{\partial t}(0,x) = g(x). \tag{7.1.14}$$

The energy identity asserts that the function

$$E(t) := \frac{1}{2}\int_{-\infty}^{\infty} \left(\left|\frac{\partial u}{\partial x}(t,x)\right|^2 + \left|\frac{\partial u}{\partial t}(t,x)\right|^2 \right) dx$$

is constant for every solution of (7.1.14). Thus the operators $S(t) \in \mathcal{L}(H)$ extend to isometries of the completion $\mathcal{H}$ of H with respect to the norm

$$\|(f,g)\|_{\mathcal{H}} := \sqrt{\int_{-\infty}^{\infty} \left(\left|\frac{df}{dx}(x)\right|^2 + |g(x)|^2 \right) dx}.$$

The completion can be identified with the quotient of the space of all pairs (f,g), where $g \in L^2(\mathbb{R})$ and $f : \mathbb{R} \to \mathbb{R}$ is absolutely continuous with square integrable derivative, under the equivalence relation $(f_1,g_1) \sim (f_2,g_2)$ iff $g_1 = g_2$ and $f_1 - f_2$ is constant (Exercise 7.7.5). If one identifies $\mathcal{H}$ with $\mathscr{H} := L^2(\mathbb{R},\mathbb{R}^2)$ via the isomorphism $\mathcal{H} \to \mathscr{H} : (f,g) \mapsto (f',g)$, one obtains the strongly continuous group $\mathscr{S} : \mathbb{R} \to \mathcal{L}(\mathscr{H})$ of isometries, given by $\mathscr{S}(t)(f,g) := (u(t,\cdot), v(t,\cdot))$ for $t \in \mathbb{R}$ and $f, g \in L^2(\mathbb{R})$, where

$$\begin{aligned} u(t,x) &:= \frac{f(x+t) + f(x-t)}{2} + \frac{g(x+t) - g(x-t)}{2}, \\ v(t,x) &:= \frac{f(x+t) - f(x-t)}{2} + \frac{g(x+t) + g(x-t)}{2}. \end{aligned} \tag{7.1.15}$$

7.1.2. Basic Properties. The next two lemmas examine some of the elementary properties of strongly continuous semigroups.

Lemma 7.1.8. *Let X be a real Banach space and let $S : [0, \infty) \to \mathcal{L}(X)$ be a strongly continuous semigroup. Then the following hold.*

(i) $\sup_{0 \le t \le T} \|S(t)\| < \infty$ *for all* $T > 0$.

(ii) *The function* $[0, \infty) \to X : t \mapsto S(t)x$ *is continuous for all* $x \in X$.

(iii) *The function* $t^{-1} \log \|S(t)\|$ *converges in* $\mathbb{R} \cup \{-\infty\}$ *as* t *tends to infinity and*

$$\lim_{t\to\infty} t^{-1} \log\|S(t)\| = \inf_{t>0} t^{-1} \log\|S(t)\| =: \omega_0. \tag{7.1.16}$$

(iv) *Let* ω_0 *be as in (iii) and fix a real number* $\omega > \omega_0$. *Then there exists a constant* $M \ge 1$ *such that*

$$\|S(t)\| \le Me^{\omega t} \qquad \text{for all } t \ge 0. \tag{7.1.17}$$

Proof. To prove (i) we show first that there exist constants $\delta > 0$ and $M \ge 1$ such that, for all $t \in \mathbb{R}$,

$$0 \le t \le \delta \qquad \Longrightarrow \qquad \|S(t)\| \le M. \tag{7.1.18}$$

Suppose by contradiction that there do not exist such constants. Then

$$\sup_{0\le t\le\delta} \|S(t)\| = \infty$$

for all $\delta > 0$. Hence there exists a sequence of real numbers $t_n > 0$ such that $\lim_{n\to\infty} t_n = 0$ and the sequence $\|S(t_n)\|$ is unbounded. By the Uniform Boundedness Theorem 2.1.1 this implies that there exists an element $x \in X$ such that the sequence $\|S(t_n)x\|$ is unbounded. This contradicts the fact that $\lim_{n\to\infty} \|S(t_n)x - x\| = 0$. Thus we have proved (7.1.18).

Now fix a number $T > 0$ and choose $N \in \mathbb{N}$ such that $N\delta > T$. Fix an element $t \in [0, T]$. Then there exists a unique integer $k \in \{0, 1, \dots, N-1\}$ such that $k\delta \le t < (k + 1)\delta$ and hence, by (7.1.18),

$$\|S(t)\| = \|S(\delta)^k S(t - k\delta)\| \le \|S(\delta)\|^k \|S(t - k\delta)\| \le M^{k+1} \le M^N.$$

This proves part (i).

Part (ii) follows from part (i) and the inequalities

$$\|S(t + h)x - S(t)x\| \le \|S(t)\| \|S(h)x - x\|$$

and

$$\|S(t - h)x - S(t)x\| \le \|S(t - h)\| \|x - S(h)x\|$$

for $0 \le h \le t$.

We prove part (iii). Equation (7.1.16) holds obviously with $\omega_0 = -\infty$ whenever $S(t) = 0$ for some $t > 0$. Hence assume $S(t) \neq 0$ for all $t > 0$. Then for every $t > 0$ there is a constant $c \geq 1$ such that $c^{-1} \leq \|S(s)\| \leq c$ for $0 \leq s \leq t$. Define the function $g : [0,\infty) \to \mathbb{R}$ by

$$g(t) := \log \|S(t)\| \qquad \text{for } t \geq 0.$$

Then it follows from the semigroup property and part (i) that

$$g(0) = 0, \qquad g(s+t) \leq g(s) + g(t), \qquad M(t) := \sup_{0\leq s\leq t} |g(s)| < \infty$$

for all $s, t \geq 0$. Fix a real number $t_0 > 0$ and let $t > 0$ be any positive real number. Then there exists an integer $k \geq 0$ and a real number s such that

$$t = kt_0 + s, \qquad 0 \leq s < t_0.$$

Hence

$$\frac{g(t)}{t} \leq \frac{kg(t_0) + g(s)}{t} = \frac{g(t_0)}{t_0} - \frac{sg(t_0)}{t_0 t} + \frac{g(s)}{t} \leq \frac{g(t_0)}{t_0} + \frac{2M(t_0)}{t}.$$

This implies

$$\limsup_{t\to\infty} \frac{g(t)}{t} \leq \frac{g(t_0)}{t_0}.$$

Since this holds for all $t_0 > 0$, we have $\limsup_{t\to\infty} t^{-1}g(t) \leq \inf_{t>0} t^{-1}g(t)$ and this proves part (iii).

We prove part (iv). Fix a real number $\omega > \omega_0$. By part (iii) there exists a constant $T > 0$ such that

$$\frac{\log \|S(t)\|}{t} \leq \omega \qquad \text{for all } t \geq T.$$

Thus $\log \|S(t)\| \leq \omega t$ and so $\|S(t)\| \leq e^{\omega t}$ for all $t \geq T$. Define

$$M := \sup_{0\leq t\leq T} \|S(t)\| \, e^{-\omega t}.$$

Then $\|S(t)\| \leq Me^{\omega t}$ for all $t \geq 0$ and this proves Lemma 7.1.8. □

Lemma 7.1.9. *Let X be a real Banach space and let $S : [0,\infty) \to \mathcal{L}(X)$ be a strongly continuous semigroup. Then the following hold.*

(i) *The operator $S(t)$ is injective for some $t > 0$ if and only if it is injective for all $t > 0$.*

(ii) *The operator $S(t)$ is surjective for some $t > 0$ if and only if it is surjective for all $t > 0$.*

(iii) *The operator $S(t)$ has a dense image for some $t > 0$ if and only if it has a dense image for all $t > 0$.*

(iv) *Assume $S(t)$ is injective for all $t > 0$. Then $S(t)$ has a closed image for some $t > 0$ if and only if it has a closed image for all $t > 0$.*

Proof. We prove part (i). Assume that there exists a real number $t_0 > 0$ such that $S(t_0)$ is injective. Let $t > 0$ and choose an integer $k > 0$ such that $kt_0 \geq t$. If $x \in X$ satisfies $S(t)x = 0$, then $S(t_0)^k x = S(kt_0 - t)S(t)x = 0$ and hence $x = 0$. Thus $S(t)$ is injective for all $t > 0$.

We prove part (ii). Assume that there exists a real number $t_0 > 0$ such that $S(t_0)$ is surjective. Let $t > 0$ and choose an integer $k > 0$ such that $kt_0 \geq t$. Then $S(kt_0) = S(t_0)^k$ is surjective and this implies that $\mathrm{im}(S(t)) \supset \mathrm{im}(S(t)S(kt_0 - t)) = \mathrm{im}(S(kt_0)) = X$. Thus $S(t)$ is surjective for all $t > 0$.

We prove part (iii). Assume that there exists a real number $t_0 > 0$ such that $S(t_0)$ has a dense image. Let $t > 0$ and choose an integer $k > 0$ such that $kt_0 \geq t$. Then the operator $S(kt_0) = S(t_0)^k$ has a dense image. Since $\mathrm{im}(S(t)) \supset \mathrm{im}(S(t)S(kt_0 - t)) = \mathrm{im}(S(kt_0))$ this implies that $S(t)$ has a dense image.

We prove part (iv). Thus assume $S(t)$ is injective for all $t > 0$ and that there exists a real number $t_0 > 0$ such that $S(t_0)$ has a closed image. Then it follows from part (ii) of Corollary 4.1.17 that there exists a constant $\delta > 0$ such that $\delta \|x\| \leq \|S(t_0)x\|$ for all $x \in X$. By induction this implies $\delta^k \|x\| \leq \|S(kt_0)x\|$ for all $x \in X$ and all $k \in \mathbb{N}$. Let $t > 0$ and choose an integer $k > 0$ such that $kt_0 \geq t$. Then

$$\|S(kt_0 - t)\| \, \|S(t)x\| \geq \|S(kt_0)x\| \geq \delta^k \|x\|$$

and so $\|S(t)x\| \geq \|S(kt_0 - t)\|^{-1} \delta^k \|x\|$ for all $x \in X$. Hence $S(t)$ has a closed image by Theorem 4.1.16 and this proves Lemma 7.1.9. □

Example 7.1.10. This example shows that the hypothesis that $S(t)$ is injective for all $t > 0$ cannot be removed in part (iv) of Lemma 7.1.9. Consider the real Banach space

$$X := \left\{ f \in \mathcal{L}^2([0,1]) \,\middle|\, f \text{ is continuous on } [0, \tfrac{1}{2}] \text{ and } f(0) = 0 \right\} \Big/ \sim .$$

Here the equivalence relation is defined by $f \sim g$ if and only if $f - g$ vanishes almost everywhere on the interval $[\frac{1}{2}, 1]$, and the norm is defined by

$$\|f\|_X := \sup_{0 \leq s \leq \frac{1}{2}} |f(s)| + \sqrt{\int_{\frac{1}{2}}^1 f(s)^2 \, ds}$$

for $f \in X$. Then the formula

$$(S(t)f)(s) := \begin{cases} f(s-t), & \text{if } s \geq t, \\ 0, & \text{if } s < t, \end{cases}$$

for $f \in X$, $t \geq 0$, and $0 \leq s \leq 1$ defines a strongly continuous semigroup on X. The operator $S(t)$ has a nontrivial kernel for all $t > 0$, does not have a closed image for $0 < t < 1$, and vanishes for all $t \geq 1$.

7.1.3. The Infinitesimal Generator. The starting point of the present section was to introduce strongly continuous semigroups of operators as a generalization of the space of solutions of a linear differential equation. Given such a space of *"solutions"* it is then a natural question to ask whether there is actually a linear differential equation that a given strongly continuous semigroup provides the solutions of. The quest for such an equation leads to the following definition.

Definition 7.1.11 (Infinitesimal Generator). Let X be a real Banach space and let $S : [0, \infty) \to \mathcal{L}(X)$ be a strongly continuous semigroup. The **infinitesimal generator of** S is the linear operator $A : \mathrm{dom}(A) \to X$, whose domain is the linear subspace $\mathrm{dom}(A) \subset X$ defined by

$$\mathrm{dom}(A) := \left\{ x \in X \,\middle|\, \text{the limit } \lim_{h \searrow 0} \frac{S(h)x - x}{h} \text{ exists} \right\}, \tag{7.1.19}$$

and which is given by

$$Ax := \lim_{h \searrow 0} \frac{S(h)x - x}{h} \qquad \text{for } x \in \mathrm{dom}(A). \tag{7.1.20}$$

Example 7.1.12. Let H be a separable complex Hilbert space, let $(e_i)_{i\in\mathbb{N}}$ be a complex orthonormal basis, and let $(\lambda_i)_{i\in\mathbb{N}}$ be a sequence of complex numbers such that

$$\sup_{i\in\mathbb{N}} \mathrm{Re}\lambda_i < \infty.$$

Let $S : [0, \infty) \to \mathcal{L}^c(H)$ be the strongly continuous semigroup in Example 7.1.3, i.e.

$$S(t)x = \sum_{i=1}^{\infty} e^{\lambda_i t} \langle e_i, x \rangle e_i$$

for $x \in H$ and $t \geq 0$. Then the infinitesimal generator of S is the linear operator

$$A : \mathrm{dom}(A) \to H$$

in Example 6.1.3, given by

$$\mathrm{dom}(A) = \left\{ x \in H \,\middle|\, \sum_{i=1}^{\infty} |\lambda_i \langle e_i, x \rangle|^2 < \infty \right\} \tag{7.1.21}$$

and

$$Ax = \sum_{i=1}^{\infty} \lambda_i \langle e_i, x \rangle e_i \qquad \text{for } x \in \mathrm{dom}(A). \tag{7.1.22}$$

Exercise: Prove this.

Lemma 7.1.13. *Let X be a real Banach space and let $S : [0, \infty) \to \mathcal{L}(X)$ be a strongly continuous semigroup with infinitesimal generator*

$$A : \mathrm{dom}(A) \to X.$$

Let $x \in X$. Then the following are equivalent.

(i) $x \in \mathrm{dom}(A)$.

(ii) *The function $[0, \infty) \to X : t \mapsto S(t)x$ is continuously differentiable, takes values in the domain of A, and satisfies the differential equation*

$$\frac{d}{dt}S(t)x = AS(t)x = S(t)Ax \qquad \text{for all } t \geq 0. \tag{7.1.23}$$

Proof. That (ii) implies (i) follows directly from the definitions. To prove the converse, fix an element $x \in \mathrm{dom}(A)$. Then, for $t \geq 0$, we have

$$S(t)Ax = \lim_{h \searrow 0} S(t)\frac{S(h)x - x}{h} = \lim_{h \searrow 0} \frac{S(t+h)x - S(t)x}{h}$$

and, for $t > 0$,

$$S(t)Ax = \lim_{h \searrow 0} S(t-h)\frac{S(h)x - x}{h} = \lim_{h \searrow 0} \frac{S(t-h)x - S(t)x}{-h}.$$

This shows that the function $[0, \infty) \to X : t \mapsto S(t)x$ is continuously differentiable and that its derivative at $t \geq 0$ is $S(t)Ax$. Moreover,

$$\lim_{h \searrow 0} \frac{S(h)S(t)x - S(t)x}{h} = \lim_{h \searrow 0} S(t)\frac{S(h)x - x}{h} = S(t)Ax.$$

Thus $S(t)x \in \mathrm{dom}(A)$ and

$$AS(t)x = S(t)Ax.$$

This proves Lemma 7.1.13. □

Lemma 7.1.14 (Variation of Constants). *Let X be a real Banach space and let $S : [0, \infty) \to \mathcal{L}(X)$ be a strongly continuous semigroup with infinitesimal generator $A : \mathrm{dom}(A) \to X$. Let $f : [0, \infty) \to X$ be a continuously differentiable function and define the function $x : [0, \infty) \to X$ by*

$$x(t) := \int_0^t S(t-s)f(s)\,ds \qquad \text{for } t \geq 0. \tag{7.1.24}$$

Then x is continuously differentiable, $x(t) \in \mathrm{dom}(A)$ for all $t \geq 0$, and

$$\dot{x}(t) = Ax(t) + f(t) = S(t)f(0) + \int_0^t S(t-s)\dot{f}(s)\,ds \tag{7.1.25}$$

for all $t \geq 0$.

Proof. Fix a constant $t \geq 0$ and let $h > 0$. Then

$$\begin{aligned}
\frac{S(h)x(t) - x(t)}{h} &= \frac{S(h) - \mathbb{1}}{h} \int_0^t S(s)f(t-s)\,ds \\
&= \frac{1}{h}\int_0^t S(s+h)f(t-s)\,ds - \frac{1}{h}\int_0^t S(s)f(t-s)\,ds \\
&= \frac{1}{h}\int_h^{t+h} S(s)f(t+h-s)\,ds - \frac{1}{h}\int_0^t S(s)f(t-s)\,ds \\
&= \frac{1}{h}\int_t^{t+h} S(s)f(t+h-s)\,ds - \frac{1}{h}\int_0^h S(s)f(t+h-s)\,ds \\
&\quad + \int_0^t S(s)\frac{f(t+h-s) - f(t-s)}{h}\,ds.
\end{aligned}$$

Take the limit $h \to 0$ to obtain $x(t) \in \mathrm{dom}(A)$ and

$$Ax(t) = S(t)f(0) - f(t) + \int_0^t S(t-s)\dot{f}(s)\,ds. \tag{7.1.26}$$

This proves the second equation in (7.1.25) and shows that Ax is continuous. Next observe that

$$\begin{aligned}
\frac{x(t+h) - x(t)}{h} &= \frac{1}{h}\int_0^{t+h} S(t+h-s)f(s)\,ds - \frac{1}{h}\int_0^t S(t-s)f(s)\,ds \\
&= \frac{S(h)x(t) - x(t)}{h} + \frac{1}{h}\int_t^{t+h} S(t+h-s)f(s)\,ds
\end{aligned}$$

for all $h > 0$. Take the limit $h \to 0$ to obtain that x is right differentiable and $\frac{d}{dt^+}x(t) = Ax(t) + f(t)$. Third, observe that

$$\begin{aligned}
\frac{x(t) - x(t-h)}{h} &= \frac{1}{h}\int_0^t S(t-s)f(s)\,ds - \frac{1}{h}\int_0^{t-h} S(t-h-s)f(s)\,ds \\
&= \frac{1}{h}\int_0^t S(t-s)f(s)\,ds - \frac{1}{h}\int_h^t S(t-s)f(s-h)\,ds \\
&= \frac{1}{h}\int_0^h S(t-s)f(s)\,ds + \int_h^t S(t-s)\frac{f(s) - f(s-h)}{h}\,ds
\end{aligned}$$

for $0 < h < t$. Take the limit $h \to 0$ to obtain that x is left differentiable and $\frac{d}{dt^-}x(t) = S(t)f(0) + \int_0^t S(t-s)\dot{f}(s)\,ds = Ax(t) + f(t)$. Here the last equation follows from (7.1.26). This proves Lemma 7.1.14. □

Example 7.1.15. Let $x \in X$ and take $f(t) = x$ in Lemma 7.1.14. Then

$$\int_0^t S(s)x\,ds \in \mathrm{dom}(A), \qquad A\int_0^t S(s)x\,ds = S(t)x - x$$

for all $t > 0$.

Lemma 7.1.16. *Let X be a real Banach space and let $S:[0,\infty)\to\mathcal{L}(X)$ be a strongly continuous semigroup with infinitesimal generator*

$$A:\mathrm{dom}(A)\to X.$$

For $n\in\mathbb{N}$ define the linear subspaces $\mathrm{dom}(A^n)\subset X$ recursively by

$$\mathrm{dom}(A^1):=\mathrm{dom}(A),\qquad \mathrm{dom}(A^n):=\left\{x\in\mathrm{dom}(A)\,|\,Ax\in\mathrm{dom}(A^{n-1})\right\}$$

for $n\ge 2$. Then the linear subspace $\mathrm{dom}(A^\infty):=\bigcap_{n\in\mathbb{N}}\mathrm{dom}(A^n)$ is dense in X and A has a closed graph.

Proof. The proof has three steps.

Step 1. *Let $x\in X$ and let $\phi:\mathbb{R}\to X$ be a smooth function with compact support contained in the interval $[\delta,\delta^{-1}]$ for some constant $0<\delta<1$. Then, for every $n\in\mathbb{N}$, we have $\int_0^\infty\phi(t)S(t)x\,dt\in\mathrm{dom}(A^n)$ and*

$$A^n\int_0^\infty\phi(t)S(t)x\,dt=(-1)^n\int_0^\infty\phi^{(n)}(t)S(t)x\,dt.$$

For $n=1$ this follows from Lemma 7.1.14 with $t>\delta^{-1}$ and $f(s):=\phi(t-s)x$ for $s\ge 0$. For $n\ge 2$ the assertion follows by induction.

Step 2. $\mathrm{dom}(A^\infty)$ *is dense in X.*

Let $x\in X$ and choose a smooth function $\phi:\mathbb{R}\to[0,\infty)$, with compact support in the interval $[1/2,1]$, such that $\int_0^1\phi(t)\,dt=1$. Define

$$x_n:=n\int_0^\infty\phi(nt)S(t)x\,dt\qquad\text{for }n\in\mathbb{N}.$$

Then $x_n\in\mathrm{dom}(A^\infty)$ by Step 1 and

$$\|x_n-x\|=\left\|n\int_0^{1/n}\phi(nt)(S(t)x-x)\,dt\right\|\le\sup_{0\le t\le 1/n}\|S(t)x-x\|.$$

Hence $\lim_{n\to\infty}\|x_n-x\|=0$ and this proves Step 2.

Step 3. *A has a closed graph.*

Choose a sequence $x_n\in\mathrm{dom}(A)$ and $x,y\in X$ such that

$$\lim_{n\to\infty}\|x_n-x\|=0,\qquad\lim_{n\to\infty}\|Ax_n-y\|=0.$$

Then, by Lemma 7.1.13,

$$S(t)x-x=\lim_{n\to\infty}\big(S(t)x_n-x_n\big)=\lim_{n\to\infty}\int_0^t S(s)Ax_n\,ds=\int_0^t S(s)y\,ds$$

for all $t>0$. Hence $y=\lim_{t\searrow 0}t^{-1}(S(t)x-x)$ and this implies $x\in\mathrm{dom}(A)$ and $Ax=y$. This proves Step 3 and Lemma 7.1.16. □

Recall from Exercise 2.2.12 that the domain of a closed densely defined operator $A : \mathrm{dom}(A) \to X$ is a Banach space with the **graph norm**

$$\|x\|_A := \|x\|_X + \|Ax\|_X \qquad \text{for } x \in \mathrm{dom}(A).$$

Moreover, the operator A can be viewed as a bounded operator from $\mathrm{dom}(A)$ to X rather than as an unbounded densely defined operator from X to itself.

Lemma 7.1.17. *Let X be a real Banach space and let $S : [0,\infty) \to \mathcal{L}(X)$ be a strongly continuous semigroup. Let $A : \mathrm{dom}(A) \to X$ be a linear operator with a dense domain $\mathrm{dom}(A) \subset X$ and a closed graph. Then the following are equivalent.*

(i) *The operator A is the infinitesimal generator of the semigroup S.*

(ii) *If $x \in \mathrm{dom}(A)$ and $t > 0$, then $S(t)x \in \mathrm{dom}(A)$, $AS(t)x = S(t)Ax$, and $S(t)x - x = \int_0^t S(s)Ax\,ds$.*

(iii) *If $x_0 \in \mathrm{dom}(A)$, then the function $[0,\infty) \to X : t \mapsto x(t) := S(t)x_0$ is continuously differentiable, takes values in $\mathrm{dom}(A)$, and satisfies the differential equation $\dot{x}(t) = Ax(t)$ for all $t \geq 0$.*

Proof. That (i) implies (ii) follows directly from Lemma 7.1.13. That (ii) implies (iii) follows directly from part (vii) of Lemma 5.1.10. We prove in three steps that (iii) implies (i). Assume A satisfies (iii).

Step 1. *Let $x \in \mathrm{dom}(A)$ and $t > 0$. Then*

$$\int_0^t S(s)x\,ds \in \mathrm{dom}(A), \qquad A\int_0^t S(s)x\,ds = S(t)x - x. \tag{7.1.27}$$

By part (iii) the function $\xi : [0,t] \to X$ defined by $\xi(s) := S(s)x$ for $0 \leq s \leq t$ takes values in $\mathrm{dom}(A)$ and the function $A\xi = \dot{\xi} : [0,t] \to X$ is continuous. Hence the function $\xi : [0,t] \to \mathrm{dom}(A)$ is continuous with respect to the graph norm. Thus it follows from part (iii) of Lemma 5.1.10 that

$$\int_0^t \xi(s)\,ds \in \mathrm{dom}(A)$$

and

$$A\int_0^t \xi(s)\,ds = \int_0^t A\xi(s)\,ds = \xi(t) - \xi(0) = S(t)x - x.$$

This proves Step 1.

Step 2. *If $x \in X$ and $t > 0$, then* (7.1.27) *holds.*

Let $x \in X$ and $t > 0$. Choose a sequence $x_i \in \mathrm{dom}(A)$ that converges to x. Then $\xi_i := \int_0^t S(s)x_i\,ds \in \mathrm{dom}(A)$ and $A\xi_i = S(t)x_i - x_i$ by Step 1. Since A has a closed graph, ξ_i converges to $\int_0^t S(s)x\,ds$, and $A\xi_i$ converges to $S(t)x - x$, it follows that x and t satisfy (7.1.27). This proves Step 2.

Step 3. *Let $x, y \in X$. Then*

$$\lim_{h\to 0}\frac{S(h)x - x}{h} = y \qquad \Longleftrightarrow \qquad x \in \mathrm{dom}(A), \quad Ax = y. \tag{7.1.28}$$

If $x \in \mathrm{dom}(A)$ and $y = Ax$, then $\lim_{h\to 0} h^{-1}(S(h)x - x) = y$ by part (iii). Conversely, suppose that $\lim_{h\to 0} h^{-1}(S(h)x - x) = y$. For each $h > 0$ define $x_h := h^{-1}\int_0^h S(s)x\,ds$. Then $\lim_{h\to 0} x_h = x$ and by Step 2 $x_h \in \mathrm{dom}(A)$ and $Ax_h = h^{-1}(S(h)x - x)$. Hence $\lim_{h\to 0} Ax_h = y$. Since A has a closed graph this implies $x \in \mathrm{dom}(A)$ and $Ax = y$. This proves Lemma 7.1.17. □

Lemma 7.1.18. *Let X be a real Banach space and let $S : [0,\infty) \to \mathcal{L}(X)$ be a strongly continuous semigroup with infinitesimal generator A. Then the following are equivalent.*

(i) $\mathrm{dom}(A) = X$.

(ii) *A is bounded.*

(iii) *The semigroup S is continuous in the norm topology on $\mathcal{L}(X)$.*

Proof. The Closed Graph Theorem 2.2.13 asserts that (i) and (ii) are equivalent. That (ii) implies (iii) follows from Exercise 1.5.4 and Corollary 7.2.3 below. We prove that (iii) implies (i), following [**26**, p. 615]. Assume that $S : [0,\infty) \to \mathcal{L}(X)$ is continuous with respect to the norm topology on $\mathcal{L}(X)$. Then $\lim_{t\to 0}\|S(t) - 1\!\!1\| = 0$. Hence there exists a $\delta > 0$ such that $\sup_{0\le t\le\delta}\|S(t) - 1\!\!1\| < 1$. For $0 \le t \le \delta$ define

$$B(t) := \sum_{n=1}^{\infty}\frac{(-1)^{n-1}}{n}\big(S(t) - 1\!\!1\big)^n.$$

Then the following hold.

(I) The function $B : [0,\delta] \to \mathcal{L}(X)$ is norm-continuous.

(II) $e^{B(t)} = S(t)$ for $0 \le t \le \delta$.

(III) If $k \in \mathbb{N}$ and $0 \le t \le \delta/k$, then $B(kt) = kB(t)$.

Part (II) uses the fact that the power series $f(z) := \sum_{n=1}^{\infty}(-1)^{n-1}(z-1)^n/n$ satisfies $\exp(f(z)) = z$ for all $z \in \mathbb{C}$ with $|z-1| < 1$. Part (III) follows from the fact that $f(z^k) = kf(z)$ whenever $|z^j - 1| < 1$ for $j = 1, 2, \dots, k$.

By (III), $B(\delta) = \ell B(\delta/\ell)$ and so $B(k\delta/\ell) = kB(\delta/\ell) = (k/\ell)B(\delta)$ for all integers $0 \le k \le \ell$. Since B is continuous by (I), this implies

$$B(t) = t\delta^{-1}B(\delta) \qquad \text{for } 0 \le t \le \delta.$$

(Approximate $t\delta^{-1}$ by a sequence of rational numbers in $[0,1]$.) Now define the operator $A := \delta^{-1}B(\delta) \in \mathcal{L}(X)$. Then by (II) we have $S(t) = e^{B(t)} = e^{tA}$ for $0 \le t \le \delta$. So $S(t) = e^{tA}$ for all $t \ge 0$ and this proves Lemma 7.1.18. □

7.2. The Hille–Yosida–Phillips Theorem

7.2.1. Well-Posed Cauchy Problems. Let us now change the point of view and suppose that $A : \mathrm{dom}(A) \to X$ is a linear operator on a Banach space X whose domain is a linear subspace $\mathrm{dom}(A) \subset X$. Consider the **Cauchy problem**

$$\dot{x} = Ax, \qquad x(0) = x_0. \tag{7.2.1}$$

Definition 7.2.1. (i) Let $I \subset [0, \infty)$ be a closed interval with $0 \in I$. A continuously differentiable function $x : I \to X$ is called a **solution** of (7.2.1) if it takes values in $\mathrm{dom}(A)$ and $x(0) = x_0$ and $\dot{x}(t) = Ax(t)$ for all $t \in I$.

(ii) The Cauchy problem (7.2.1) is called **well-posed** if it satisfies the following axioms.

(Existence) For each $x_0 \in \mathrm{dom}(A)$ there is a solution of (7.2.1) on $[0, \infty)$.

(Uniqueness) Let $x_0 \in \mathrm{dom}(A)$ and $T > 0$. If $x, y : [0, T] \to X$ are solutions of (7.2.1), then $x(t) = y(t)$ for all $t \in [0, T]$.

(Continuous Dependence) Define the map $\phi : [0, \infty) \times \mathrm{dom}(A) \to X$ by $\phi(t, x_0) := x(t)$ for $t \geq 0$ and $x_0 \in \mathrm{dom}(A)$, where $x : [0, \infty) \to X$ is the unique solution of (7.2.1). Then, for every $T > 0$, there is a constant $M \geq 1$ such that $\|\phi(t, x_0)\| \leq M\|x_0\|$ for all $t \in [0, T]$ and all $x_0 \in \mathrm{dom}(A)$.

The next theorem characterizes well-posed Cauchy problems and was proved by Ralph S. Phillips [**68**] in 1954.

Theorem 7.2.2 (Phillips). *Let $A : \mathrm{dom}(A) \to X$ be a linear operator with a dense domain $\mathrm{dom}(A) \subset X$ and a closed graph. The following are equivalent.*

(i) *A is the infinitesimal generator of a strongly continuous semigroup.*

(ii) *The Cauchy problem* (7.2.1) *is well-posed.*

Proof. We prove that (i) implies (ii). Thus assume that A is the infinitesimal generator of a strongly continuous semigroup $S : [0, \infty) \to \mathcal{L}(X)$ and fix an element $x_0 \in \mathrm{dom}(A)$. Then the function $[0, \infty) \to X : t \mapsto S(t)x_0$ is a solution of equation (7.2.1) by Lemma 7.1.13. To prove uniqueness, assume that $x : [0, \infty) \to X$ is any solution of (7.2.1). Fix a constant $t > 0$. We will prove that the function $[0, t] \to X : s \mapsto S(t - s)x(s)$ is constant. To see this, note that $x(s) \in \mathrm{dom}(A)$ and so

$$\lim_{\substack{h \to 0 \\ h \leq t-s}} \frac{S(t - s - h)x(s) - S(t - s)x(s)}{-h} = S(t - s)Ax(s) \qquad \text{for } 0 \leq s \leq t.$$

This implies

$$\begin{aligned}
&\lim_{h\to 0} \frac{S(t-s-h)x(s+h) - S(t-s)x(s)}{h} \\
&= \lim_{h\to 0} S(t-s-h)\left(\frac{x(s+h)-x(s)}{h} - Ax(s)\right) \\
&\quad + \lim_{h\to 0}\left(\frac{S(t-s-h)x(s) - S(t-s)x(s)}{h} + S(t-s)Ax(s)\right) \\
&\quad + \lim_{h\to 0}\big(S(t-s-h)Ax(s) - S(t-s)Ax(s)\big) \\
&= 0.
\end{aligned}$$

Hence the function $[0,t] \to X : s \mapsto S(t-s)x(s)$ is everywhere differentiable and its derivative vanishes. Thus it is constant and hence $x(t) = S(t)x_0$. Since $t > 0$ was chosen arbitrarily this proves uniqueness. Continuous dependence follows from the estimate $\|S(t)\| \le Me^{\omega t}$ in Lemma 7.1.8. This shows that (i) implies (ii).

We prove that (ii) implies (i). Assume the Cauchy problem (7.2.1) is well-posed and let

$$\phi : [0,\infty) \times \mathrm{dom}(A) \to \mathrm{dom}(A)$$

be the map that assigns to each element $x_0 \in \mathrm{dom}(A)$ the unique solution $[0,\infty) \to X : t \mapsto \phi(t,x_0)$ of (7.2.1). We claim that, for each $t \ge 0$, there is a unique bounded linear operator $S(t) : X \to X$ such that

$$S(t)x_0 = \phi(t,x_0) \qquad \text{for all } x_0 \in \mathrm{dom}(A). \tag{7.2.2}$$

To see this, note first that the space of solutions $x : [0,\infty) \to X$ of (7.2.1) is a linear subspace of the space of all functions from $[0,\infty)$ to X. Hence it follows from uniqueness that the map $\mathrm{dom}(A) \to X : x_0 \mapsto \phi(t,x_0)$ is linear. Second, it follows from continuous dependence that the linear operator $\mathrm{dom}(A) \to X : x_0 \mapsto \phi(t,x_0)$ is bounded. Since $\mathrm{dom}(A)$ is a dense linear subspace of X it follows that this operator extends uniquely to a bounded linear operator $S(t) \in \mathcal{L}(X)$. More precisely, fix an element $x \in X$. Then there exists a sequence $x_n \in \mathrm{dom}(A)$ that converges to x. Hence $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in X and so is the sequence $(\phi(t,x_n))_{n\in\mathbb{N}}$ by continuous dependence. Hence it converges and the limit

$$S(t)x := \lim_{n\to\infty} \phi(t,x_n)$$

is independent of the choice of the sequence $x_n \in \mathrm{dom}(A)$ used to define it. This proves the existence of a bounded linear operator $S(t)$ that satisfies (7.2.2).

We prove that these operators form a one-parameter semigroup. Fix a real number $t \geq 0$ and an element $x_0 \in \mathrm{dom}(A)$. Then

$$S(t)x_0 = \phi(t, x_0) \in \mathrm{dom}(A)$$

and the function $[0,\infty) \to X : s \mapsto S(s+t)x_0 = \phi(s+t, x_0)$ is a solution of the Cauchy problem (7.2.1) with x_0 replaced by $S(t)x_0 = \phi(t, x_0)$. Hence

$$S(s+t, x_0) = \phi(s, S(t)x_0) = S(s)S(t)x_0.$$

Since this holds for all $x_0 \in \mathrm{dom}(A)$, the set $\mathrm{dom}(A)$ is dense in X, and the operators $S(s+t)$ and $S(s)S(t)$ are both continuous maps, it follows that $S(s+t) = S(s)S(t)$ for all $s \geq 0$. This shows that $S : [0,\infty) \to \mathcal{L}(X)$ is a one-parameter semigroup.

We prove that S is strongly continuous. To see this, fix an element $x \in X$ and a constant $\varepsilon > 0$. By continuous dependence there exists an $M \geq 1$ such that $\sup_{0 \leq t \leq 1} \|\phi(t, x_0)\| \leq M \|x_0\|$ for all $x_0 \in \mathrm{dom}(A)$. This shows that $\sup_{0 \leq t \leq 1} \|S(t)\| \leq M$. Choose an element $y \in \mathrm{dom}(A)$ such that

$$\|x - y\| \leq \frac{\varepsilon}{2(M+1)}.$$

Next choose a constant $0 < \delta < 1$ such that, for all $t \in \mathbb{R}$,

$$0 \leq t < \delta \qquad \Longrightarrow \qquad \|\phi(t, y) - y\| < \frac{\varepsilon}{2}.$$

Fix a real number $0 \leq t < \delta$. Then

$$\begin{aligned} \|S(t)x - x\| &\leq \|S(t)x - S(t)y\| + \|S(t)y - y\| + \|y - x\| \\ &\leq (M+1)\|x - y\| + \|\phi(t, y) - y\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

This shows that S is strongly continuous.

We prove that A is the infinitesimal generator of S. Let $x_0 \in \mathrm{dom}(A)$ and define the function $x : [0,\infty) \to X$ by $x(t) := S(t)x_0 = \phi(t, x_0)$. It is continuously differentiable, takes values in $\mathrm{dom}(A)$, and satisfies the equation $\dot{x}(t) = Ax(t)$ for all $t \geq 0$. Thus A and S satisfy condition (iii) in Lemma 7.1.17, so A is the infinitesimal generator of S. This proves Theorem 7.2.2. □

Corollary 7.2.3 (Uniqueness). *A linear operator on a Banach space is the infinitesimal generator of at most one strongly continuous semigroup.*

Proof. Let A be the infinitesimal generator of two strongly continuous semigroups $S, T : [0,\infty) \to \mathcal{L}(X)$. Let $x_0 \in \mathrm{dom}(A)$. Then the functions $x(t) := S(t)x_0$ and $y(t) := T(t)x_0$ both satisfy (7.2.1) and hence agree by Theorem 7.2.2. Since $\mathrm{dom}(A)$ is dense in X by Lemma 7.1.16, it follows that $S(t)x = T(t)x$ for all $x \in X$ and all $t \geq 0$. □

Theorem 7.2.4 (Strongly Continuous Groups). *Let X be a real Banach space, let $S : [0,\infty) \to \mathcal{L}(X)$ be a strongly continuous semigroup, and let $A : \mathrm{dom}(A) \to X$ be the infinitesimal generator of S. Then the following are equivalent.*

(i) *The semigroup S extends to a strongly continuous group $S : \mathbb{R} \to \mathcal{L}(X)$.*

(ii) *$-A$ is the infinitesimal generator of a strongly continuous semigroup.*

(iii) *The operator $S(t)$ is bijective for all $t > 0$.*

Proof. We prove that (i) implies (ii). Thus assume that S extends to a strongly continuous group $S : \mathbb{R} \to \mathcal{L}(X)$. Then

$$S(t)S(-t) = S(-t)S(t) = \mathbb{1}$$

for all $t > 0$ by definition of a one-parameter group of operators. This implies that $S(t)$ is bijective and

$$S(t)^{-1} = S(-t)$$

for all $t > 0$. Define the map $T : [0,\infty) \to \mathcal{L}(X)$ by

$$T(t) := S(-t) = S(t)^{-1} \qquad \text{for } t \geq 0.$$

Then T is a strongly continuous semigroup by definition. Denote its infinitesimal generator by $B : \mathrm{dom}(B) \to X$. We must prove that $B = -A$. To see this, choose a constant $M \geq 1$ such that

$$\|S(t)\| \leq M \quad \text{and} \quad \|T(t)\| \leq M \qquad \text{for } 0 \leq t \leq 1.$$

Now let $x \in \mathrm{dom}(A)$. Then

$$\begin{aligned} \left\| \frac{T(h)x - x}{h} + Ax \right\| &\leq \left\| T(h) \left(\frac{x - S(h)x}{h} + Ax \right) \right\| + \|Ax - T(h)Ax\| \\ &\leq M \left\| \frac{x - S(h)x}{h} + Ax \right\| + \|Ax - T(h)Ax\| \end{aligned}$$

for $0 < h < 1$. Since the right-hand side converges to zero it follows that

$$x \in \mathrm{dom}(B), \qquad Bx = -Ax.$$

Thus we have proved that

$$\mathrm{dom}(A) \subset \mathrm{dom}(B), \qquad B|_{\mathrm{dom}(A)} = -A.$$

Interchange the roles of S and T to obtain

$$\mathrm{dom}(B) = \mathrm{dom}(A), \qquad B = -A.$$

This shows that (i) implies (ii).

We prove that (ii) implies (iii). Let $T : [0, \infty) \to \mathcal{L}(X)$ be the strongly continuous semigroup generated by $-A$. We prove that $S(t)$ is bijective and $T(t) = S(t)^{-1}$ for all $t > 0$. To see this, fix an element $x \in \mathrm{dom}(A)$ and a real number $t > 0$. Define the functions $y, z : [0, t] \to X$ by

$$y(s) := S(t-s)x, \qquad z(s) := T(t-s)x \qquad \text{for } 0 \le s \le t.$$

Then y and z are continuously differentiable, take values in the domain of A, and satisfy the Cauchy problems

$$\dot{y}(s) = -Ay(s) \quad \text{for } 0 \le s \le t, \qquad y(0) = S(t)x,$$

and

$$\dot{z}(s) = Az(s) \quad \text{for } 0 \le s \le t, \qquad z(0) = T(t)x.$$

By Theorem 7.2.2 this implies

$$y(s) = T(s)S(t)x, \qquad z(s) = S(s)T(t)x \qquad \text{for } 0 \le s \le t.$$

Take $s = t$ to obtain $T(t)S(t)x = y(t) = x$ and $S(t)T(t)x = z(t) = x$. Thus we have proved that $S(t)T(t)x = T(t)S(t)x = x$ for all $t > 0$ and all $x \in \mathrm{dom}(A)$. Since the domain of A is dense in X this implies

$$S(t)T(t) = T(t)S(t) = \mathbb{1} \qquad \text{for all } t > 0.$$

Hence $S(t)$ is bijective for all $t > 0$. This shows that (ii) implies (iii).

We prove that (iii) implies (i). Thus assume that $S(t)$ is bijective for all $t > 0$. Then $S(t)^{-1} : X \to X$ is a bounded linear operator for every $t > 0$ by the Open Mapping Theorem 2.2.1. Define

$$S(-t) := S(t)^{-1} \qquad \text{for } t > 0.$$

We prove that the extended function $S : \mathbb{R} \to \mathcal{L}(X)$ is a one-parameter group. The formula $S(t+s) = S(t)S(s)$ holds by definition whenever $s, t \ge 0$ or $s, t \le 0$. Moreover, if $0 \le s < t$, then $S(t-s)S(s) = S(t)$ and hence

$$S(t-s) = S(t)S(s)^{-1} = S(t)S(-s).$$

This implies that, for $0 \le t < s$, we have $S(s-t) = S(s)S(-t)$ and hence

$$S(t-s) = S(s-t)^{-1} = S(-t)^{-1}S(s)^{-1} = S(t)S(-s).$$

This shows that S is a one-parameter group. Strong continuity at $t = 0$ follows from the equation

$$S(-h)x - x = S(h)^{-1}(x - S(h)x)$$

for $h > 0$. Strong continuity at $-t < 0$ follows from the equation

$$S(-t+h)x - S(-t)x = S(t)^{-1}\big(S(h)x - x\big)$$

for $h \in \mathbb{R}$. This proves Theorem 7.2.4. □

7.2.2. The Hille–Yosida–Phillips Theorem. The following theorem is the main result of this chapter. For the special case $M = 1$ it was discovered by Hille [**35**] and Yosida [**87**] independently in 1948. It was extended to the case $M > 1$ by Phillips [**67**] in 1952.

Theorem 7.2.5 (Hille–Yosida–Phillips). *Let X be a real Banach space and $A : \mathrm{dom}(A) \to X$ be a linear operator with a dense domain $\mathrm{dom}(A) \subset X$. Fix real numbers ω and $M \geq 1$. Then the following are equivalent.*

(i) *The operator A is the infinitesimal generator of a strongly continuous semigroup $S : [0, \infty) \to \mathcal{L}(X)$ that satisfies*

$$\|S(t)\| \leq M e^{\omega t} \qquad \text{for all } t \geq 0. \tag{7.2.3}$$

(ii) *For every real number $\lambda > \omega$ the operator $\lambda 1\!\!1 - A : \mathrm{dom}(A) \to X$ is invertible and*

$$\|(\lambda 1\!\!1 - A)^{-k}\| \leq \frac{M}{(\lambda - \omega)^k} \qquad \text{for all } \lambda > \omega \text{ and all } k \in \mathbb{N}. \tag{7.2.4}$$

Proof. See page 371. □

The necessity of the condition (7.2.4) is a straightforward consequence of Lemma 7.2.6 below which expresses the resolvent operator $(\lambda 1\!\!1 - A)^{-1}$ in terms of the semigroup. At this point it is convenient to allow for λ to be a complex number and therefore to extend the discussion to complex Banach spaces. When X is a real Banach space we will tacitly assume that X has been complexified so as to make sense of the operator $\lambda 1\!\!1 - A : \mathrm{dom}(A) \to X$ for complex numbers λ (see Exercise 5.1.5).

Lemma 7.2.6 (Resolvent Identity for Semigroups). *Let X be a complex Banach space and let*

$$A : \mathrm{dom}(A) \to X$$

be the infinitesimal generator of a strongly continuous semigroup

$$S : [0, \infty) \to \mathcal{L}^c(X).$$

Let $\lambda \in \mathbb{C}$ such that

$$\mathrm{Re}\lambda > \omega_0 := \lim_{t \to \infty} \frac{\log\|S(t)\|}{t}. \tag{7.2.5}$$

Then $\lambda \in \rho(A)$ and

$$(\lambda 1\!\!1 - A)^{-k} x = \frac{1}{(k-1)!} \int_0^\infty t^{k-1} e^{-\lambda t} S(t) x \, dt \tag{7.2.6}$$

for all $x \in X$ and all $k \in \mathbb{N}$.

Proof. We first prove the assertion for $k = 1$. Fix a complex number λ such that $\mathrm{Re}\lambda > \omega_0$ and choose a real number ω such that $\omega_0 < \omega < \mathrm{Re}\lambda$. By Lemma 7.1.8, there exists a constant $M \geq 1$ such that $\|S(t)\| \leq Me^{\omega t}$ for all $t \geq 0$. Hence $\|e^{-\lambda t}S(t)x\| \leq Me^{(\omega - \mathrm{Re}\lambda)t}\|x\|$ for all $x \in X$ and all $t \geq 0$. This implies that the formula

$$R_\lambda x := \int_0^\infty e^{-\lambda t}S(t)x\,dt = \lim_{T\to\infty}\int_0^T e^{-\lambda t}S(t)x\,dt \quad \text{for } x \in X$$

defines a bounded linear operator $R_\lambda \in \mathcal{L}^c(X)$. We prove the following.

Claim 1. *If $x \in X$ and $T > 0$, then $\xi_T := \int_0^T e^{-\lambda t}S(t)x\,dt \in \mathrm{dom}(A)$ and*

$$A\xi_T = e^{-\lambda T}S(T)x - x + \lambda\int_0^T e^{-\lambda t}S(t)x\,dt =: \eta_T.$$

Claim 2. *If $x \in \mathrm{dom}(A)$ and $T > 0$, then $\int_0^T e^{-\lambda t}S(t)Ax\,ds = \eta_T$.*

Claim 1 follows from Lemma 7.1.14 with $t = T$ and $f(t) := e^{-\lambda(T-t)}x$. Claim 2 follows from integration by parts with $\frac{d}{dt}S(t)x = S(t)Ax$. Now

$$A\xi_T = \eta_T, \qquad \lim_{T\to\infty}\xi_T = R_\lambda x, \qquad \lim_{T\to\infty}\eta_T = \lambda R_\lambda x - x$$

by Claim 1. Since A has a closed graph this implies

$$R_\lambda x \in \mathrm{dom}(A), \qquad AR_\lambda x = \lambda R_\lambda x - x \qquad \text{for all } x \in X.$$

If $x \in \mathrm{dom}(A)$ it follows from Claim 2 that

$$R_\lambda Ax = \lim_{T\to\infty}\int_0^T e^{-\lambda t}S(t)Ax\,dt = \lambda R_\lambda x - x.$$

Thus $(\lambda\mathbb{1} - A)R_\lambda x = x$ for all $x \in X$ and $R_\lambda(\lambda\mathbb{1} - A)x = x$ for all $x \in \mathrm{dom}(A)$. Hence $\lambda\mathbb{1} - A$ is bijective and $(\lambda\mathbb{1} - A)^{-1} = R_\lambda$. This proves (7.2.6) for $k = 1$. To prove the equation for $k \geq 2$ observe that the function

$$\rho(A) \to X : \lambda \mapsto (\lambda\mathbb{1} - A)^{-1}x$$

is holomorphic by Lemma 6.1.10 and satisfies

$$\begin{aligned}
(\lambda\mathbb{1} - A)^{-k}x &= \frac{(-1)^{k-1}}{(k-1)!}\frac{d^{k-1}}{d\lambda^{k-1}}(\lambda\mathbb{1} - A)^{-1}x \\
&= \frac{(-1)^{k-1}}{(k-1)!}\frac{d^{k-1}}{d\lambda^{k-1}}\int_0^\infty e^{-\lambda t}S(t)x\,dt \\
&= \frac{1}{(k-1)!}\int_0^\infty t^{k-1}e^{-\lambda t}S(t)x\,dt
\end{aligned}$$

for all $x \in X$ and all $\lambda \in \mathbb{C}$ with $\mathrm{Re}\lambda > \omega_0$. This proves Lemma 7.2.6. □

It follows from Lemma 7.2.6 that

$$\sup_{\lambda\in\sigma(A)} \mathrm{Re}\lambda \le \omega_0 = \lim_{t\to\infty} \frac{\log\|S(t)\|}{t} \tag{7.2.7}$$

for every strongly continuous semigroup S with infinitesimal generator A. The following example by Einar Hille shows that the inequality in (7.2.7) can be strict.

Example 7.2.7. Fix a real number $\omega > 0$ and consider the Banach space

$$X := \left\{ f : [0,\infty) \to \mathbb{C} \;\middle|\; \begin{array}{l} f \text{ is continuous and bounded} \\ \text{and } \int_0^\infty e^{\omega s}|f(s)|\,ds < \infty \end{array} \right\},$$

equipped with the norm

$$\|f\| := \sup_{s\ge 0}|f(s)| + \int_0^\infty e^{\omega s}|f(s)|\,ds \qquad \text{for } f \in X.$$

The formula

$$(S(t)f)(s) := f(s+t) \qquad \text{for } f \in X \text{ and } s,t \ge 0$$

defines a strongly continuous semigroup on X and its infinitesimal generator is the operator $A : \mathrm{dom}(A) \to X$ given by

$$\mathrm{dom}(A) = \left\{ u : [0,\infty) \to \mathbb{C} \;\middle|\; \begin{array}{l} u \text{ is continuously differentiable} \\ \text{and } u, \dot{u} \in X \end{array} \right\},$$

$$Au = \dot{u}.$$

The operator $S(t)$ satisfies $\|S(t)\| = 1$ for all $t \ge 0$ and so

$$\omega_0 = \lim_{t\to\infty} \frac{\log\|S(t)\|}{t} = 0.$$

Now let $\lambda \in \mathbb{C}$ with $\mathrm{Re}\lambda > -\omega$ and let $f \in X$. Then, for $u \in \mathrm{dom}(A)$,

$$\lambda u - Au = f \qquad \iff \qquad \dot{u} = \lambda u - f.$$

This equation has a unique solution $u \in \mathrm{dom}(A)$ given by

$$u(s) = \int_s^\infty e^{\lambda(s-t)} f(t)\,dt \qquad \text{for } s \ge 0.$$

Thus the operator $\lambda\mathbb{1} - A$ is bijective for all $\lambda \in \mathbb{C}$ with $\mathrm{Re}\lambda > -\omega$. It has a one-dimensional kernel for all $\lambda \in \mathbb{C}$ with $\mathrm{Re}\lambda < -\omega$. Thus

$$\sup_{\lambda\in\sigma(A)} \mathrm{Re}\lambda = -\omega < 0 = \lim_{t\to\infty} \frac{\log\|S(t)\|}{t}.$$

Exercise: For $t > 0$ the spectrum of $S(t)$ is the closed unit disc and the point spectrum of $S(t)$ is the open disc of radius $e^{-\omega t}$ centered at the origin.

Proof of Theorem 7.2.5. We prove that (i) implies (ii). Thus assume that $A : \mathrm{dom}(A) \to X$ is the infinitesimal generator of a strongly continuous semigroup $S : [0, \infty) \to \mathcal{L}(X)$ that satisfies (7.2.3). Fix a real number

$$\lambda > \omega$$

and a positive integer k. Then

$$(\lambda 1\!\!1 - A)^{-k} x = \frac{1}{(k-1)!} \int_0^\infty t^{k-1} e^{-\lambda t} S(t) x \, dt$$

for all $x \in X$ by Lemma 7.2.6 and hence

$$\begin{aligned} \|(\lambda 1\!\!1 - A)^{-k} x\| &\leq \frac{1}{(k-1)!} \int_0^\infty t^{k-1} e^{-\lambda t} \|S(t)x\| \, dt \\ &\leq \frac{M\|x\|}{(k-1)!} \int_0^\infty t^{k-1} e^{-(\lambda-\omega)t} \, dt \\ &= \frac{M\|x\|}{(\lambda - \omega)^k}. \end{aligned}$$

Hence the operator A satisfies (ii).

We prove that (ii) implies (i). Thus assume that $A : \mathrm{dom}(A) \to X$ is a linear operator with a dense domain such that

$$\lambda 1\!\!1 - A : \mathrm{dom}(A) \to X$$

is bijective and satisfies the estimate (7.2.4) for $\lambda > \omega$. We prove in five steps that A is the infinitesimal generator of a strongly continuous semigroup that satisfies the estimate (7.2.3).

Step 1. $x = \lim_{\lambda \to \infty} \lambda(\lambda 1\!\!1 - A)^{-1} x$ *for all* $x \in X$.

If $x \in \mathrm{dom}(A)$, then

$$\lambda(\lambda 1\!\!1 - A)^{-1} x - x = A(\lambda 1\!\!1 - A)^{-1} x = (\lambda 1\!\!1 - A)^{-1} A x$$

for all $\lambda > \omega$ and so it follows from (7.2.4) that

$$\left\| \lambda(\lambda 1\!\!1 - A)^{-1} x - x \right\| \leq \frac{M}{\lambda - \omega} \|Ax\|.$$

Thus

$$x = \lim_{\lambda \to \infty} \lambda(\lambda 1\!\!1 - A)^{-1} x$$

for all $x \in \mathrm{dom}(A)$. Moreover

$$\left\| \lambda(\lambda 1\!\!1 - A)^{-1} \right\| \leq \frac{M\lambda}{\lambda - \omega} \leq 2M \qquad \text{for all } \lambda > 2\omega.$$

Hence Step 1 follows from Theorem 2.1.5.

Step 2. *For $\lambda > \omega$ and $t \ge 0$ define*

$$A_\lambda := \lambda A(\lambda \mathbb{1} - A)^{-1}, \qquad S_\lambda(t) := e^{tA_\lambda} = \sum_{k=0}^\infty \frac{t^k A_\lambda^k}{k!}.$$

Then

$$\|S_\lambda(t)\| \le M e^{\frac{\lambda\omega t}{\lambda-\omega}}$$

for all $\lambda > \omega$ and all $t \ge 0$.

The operator A_λ can be written as

$$A_\lambda = \lambda^2(\lambda\mathbb{1} - A)^{-1} - \lambda\mathbb{1}.$$

Hence

$$\begin{aligned}
\|S_\lambda(t)\| &= e^{-\lambda t}\left\|e^{t\lambda^2(\lambda\mathbb{1}-A)^{-1}}\right\| \\
&\le e^{-\lambda t}\sum_{k=0}^\infty \frac{t^k\lambda^{2k}}{k!}\left\|(\lambda\mathbb{1}-A)^{-k}\right\| \\
&\le e^{-\lambda t}\sum_{k=0}^\infty \frac{t^k\lambda^{2k}}{k!}\frac{M}{(\lambda-\omega)^k} \\
&= Me^{-\lambda t}e^{\frac{\lambda^2 t}{\lambda-\omega}} = Me^{\frac{\lambda\omega t}{\lambda-\omega}}
\end{aligned}$$

for all $\lambda > \omega$ and all $t \ge 0$. This proves Step 2.

Step 3. *Fix real numbers $\lambda > \mu > \omega$. Then*

$$\|S_\lambda(t)x - S_\mu(t)x\| \le M^2 e^{\frac{\mu\omega t}{\mu-\omega}} t \|A_\lambda x - A_\mu x\|$$

for all $x \in X$ and all $t \ge 0$.

Since $A_\lambda A_\mu = A_\mu A_\lambda$, we have $A_\lambda S_\mu(t) = S_\mu(t)A_\lambda$ and so

$$\begin{aligned}
S_\lambda(t)x - S_\mu(t)x &= \int_0^t \frac{d}{ds}S_\mu(t-s)S_\lambda(s)x\,ds \\
&= \int_0^t S_\mu(t-s)S_\lambda(s)(A_\lambda x - A_\mu x)\,ds
\end{aligned}$$

for all $x \in X$ and all $t \ge 0$. Hence

$$\begin{aligned}
\|S_\lambda(t)x - S_\mu(t)x\| &\le \int_0^t \|S_\mu(t-s)\|\,\|S_\lambda(s)\|\,ds\,\|A_\lambda x - A_\mu x\| \\
&\le M^2 e^{\frac{\mu\omega t}{\mu-\omega}} \int_0^t e^{-\frac{\mu\omega s}{\mu-\omega}} e^{\frac{\lambda\omega s}{\lambda-\omega}}\,ds\,\|A_\lambda x - A_\mu x\| \\
&\le M^2 e^{\frac{\mu\omega t}{\mu-\omega}} t\,\|A_\lambda x - A_\mu x\|.
\end{aligned}$$

Here the last step uses the inequality $\frac{\lambda\omega}{\lambda-\omega} \le \frac{\mu\omega}{\mu-\omega}$. This proves Step 3.

Step 4. *The limit*

$$S(t)x := \lim_{\lambda\to\infty} S_\lambda(t)x \tag{7.2.8}$$

exists for all $x \in X$ and all $t \geq 0$. The resulting map $S : [0,\infty) \to \mathcal{L}(X)$ is a strongly continuous semigroup that satisfies (7.2.3).

Assume first that $x \in \mathrm{dom}(A)$. Then $\lim_{\lambda\to\infty} A_\lambda x = Ax$ by Step 1. Hence the limit (7.2.8) exists for all $t \geq 0$ by Step 3 and the convergence is uniform on every compact interval $[0,T]$. Since the operator family $\{S_\lambda(t)\}_{\lambda\geq 2\omega}$ is bounded by Step 2 it follows from Theorem 2.1.5 that the limit (7.2.8) exists for all $x \in X$ and that $S(t) \in \mathcal{L}(X)$ for all $t \geq 0$. Apply Theorem 2.1.5 to the operator family $X \to C([0,T],X) : x \mapsto S_\lambda(\cdot)x$ to deduce that the map $[0,T] \to X : t \mapsto S(t)x$ is continuous for all $x \in X$ and all $T > 0$. Moreover,

$$S(s)S(t)x = \lim_{\lambda\to\infty} S_\lambda(s)S_\lambda(t)x = \lim_{\lambda\to\infty} S_\lambda(s+t)x = S(s+t)x$$

for all $s,t \geq 0$ and all $x \in X$ and $S(0)x = \lim_{\lambda\to\infty} S_\lambda(t)x = x$ for all $x \in X$. Thus S is a strongly continuous semigroup. By Step 2 it satisfies the estimate

$$\|S(t)x\| = \lim_{\lambda\to\infty} \|S_\lambda(t)x\| \leq \lim_{\lambda\to\infty} Me^{\frac{\lambda\omega t}{\lambda-\omega}} \|x\| = Me^{\omega t} \|x\|$$

and this proves Step 4.

Step 5. *The operator A is the infinitesimal generator of S.*

Let B be the infinitesimal generator of S and let $x \in \mathrm{dom}(A)$. Then

$$\|S_\lambda(t)A_\lambda x - S(t)Ax\| \leq \|S_\lambda(t)\| \|A_\lambda x - Ax\| + \|S_\lambda(t)Ax - S(t)Ax\|$$

for all $t \geq 0$. Hence it follows from Step 1 and Step 2 that the functions $S_\lambda(\cdot)A_\lambda x : [0,h] \to X$ converge uniformly to $S(\cdot)Ax$ as λ tends to infinity. This implies

$$\int_0^h S(t)Ax\,dt = \lim_{\lambda\to\infty} \int_0^h S_\lambda(t)A_\lambda x\,dt = \lim_{\lambda\to\infty} S_\lambda(h)x - x = S(h)x - x$$

for all $h > 0$ and so

$$\lim_{h\to 0} \frac{S(h)x - x}{h} = \lim_{h\to 0} \frac{1}{h} \int_0^h S(t)Ax\,dt = Ax.$$

This shows that $\mathrm{dom}(A) \subset \mathrm{dom}(B)$ and $B|_{\mathrm{dom}(A)} = A$. Now let $y \in \mathrm{dom}(B)$ and $\lambda > \omega$. Define $x := (\lambda\mathbb{1} - A)^{-1}(\lambda y - By)$. Then $x \in \mathrm{dom}(A) \subset \mathrm{dom}(B)$ and $\lambda x - Bx = \lambda x - Ax = \lambda y - By$. Since $\lambda\mathbb{1} - B : \mathrm{dom}(B) \to X$ is injective by Lemma 7.2.6, this implies $y = x \in \mathrm{dom}(A)$. Thus $\mathrm{dom}(B) \subset \mathrm{dom}(A)$ and so $\mathrm{dom}(B) = \mathrm{dom}(A)$. This proves Step 5 and Theorem 7.2.5. □

Corollary 7.2.8. *Let X be a complex Banach space and let*

$$A : \mathrm{dom}(A) \to X$$

be a complex linear operator with a dense domain $\mathrm{dom}(A) \subset X$. *Fix two real numbers $M \geq 1$ and ω. Then the following are equivalent.*

(i) *The operator A is the infinitesimal generator of a strongly continuous semigroup $S : [0, \infty) \to \mathcal{L}^c(X)$ that satisfies the estimate* (7.2.3).

(ii) *For every real number $\lambda > \omega$ the operator $\lambda\mathbb{1} - A : \mathrm{dom}(A) \to X$ is bijective and satisfies the estimate* (7.2.4).

(iii) *For every $\lambda \in \mathbb{C}$ with $\mathrm{Re}\lambda > \omega$ the operator $\lambda\mathbb{1} - A : \mathrm{dom}(A) \to X$ is bijective and satisfies the estimate*

$$\|(\lambda\mathbb{1} - A)^{-k}\| \leq \frac{M}{(\mathrm{Re}\lambda - \omega)^k} \qquad \text{for all } k \in \mathbb{N}. \tag{7.2.9}$$

Proof. That (i) implies (iii) follows from Lemma 7.2.6 by the same argument that was used in the proof of Theorem 7.2.5. That (iii) implies (ii) is obvious and that (ii) implies (i) follows from Theorem 7.2.5 and the fact that the operators $S_\lambda(t)$ in the proof of Theorem 7.2.5 are complex linear whenever A is complex linear. This proves Corollary 7.2.8. □

7.2.3. Contraction Semigroups. The archetypal example of a contraction semigroup is the heat flow in Example 7.1.6. Here is the general definition.

Definition 7.2.9 (Contraction Semigroup). Let X be a real Banach space. A **contraction semigroup on** X is a strongly continuous semigroup $S : [0, \infty) \to \mathcal{L}(X)$ that satisfies the inequality

$$\|S(t)\| \leq 1 \tag{7.2.10}$$

for all $t \geq 0$.

Definition 7.2.10 (Dissipative Operator). Let X be a complex Banach space. A complex linear operator $A : \mathrm{dom}(A) \to X$ with a dense domain $\mathrm{dom}(A) \subset X$ is called **dissipative** if, for every $x \in \mathrm{dom}(A)$, there exists an element $x^* \in X^*$ such that

$$\|x^*\|^2 = \|x\|^2 = \langle x^*, x\rangle, \qquad \mathrm{Re}\langle x^*, Ax\rangle \leq 0. \tag{7.2.11}$$

When $X = H$ is a complex Hilbert space, a linear operator $A : \mathrm{dom}(A) \to H$ with a dense domain $\mathrm{dom}(A) \subset H$ is dissipative if and only if

$$\mathrm{Re}\langle x, Ax\rangle \leq 0 \tag{7.2.12}$$

for all $x \in \mathrm{dom}(A)$.

The next theorem characterizes the infinitesimal generators of contraction semigroups. It was proved by Lumer–Phillips [**58**] in 1961. They also introduced the notion of a dissipative operator.

Theorem 7.2.11 (**Lumer–Phillips**). *Let X be a complex Banach space and let $A : \mathrm{dom}(A) \to X$ be a complex linear operator with a dense domain $\mathrm{dom}(A) \subset X$. Then the following are equivalent.*

(i) *The operator A is the infinitesimal generator of a contraction semigroup.*

(ii) *For every real number $\lambda > 0$ the operator $\lambda \mathbb{1} - A : \mathrm{dom}(A) \to X$ is bijective and satisfies the estimate*

$$\left\|(\lambda \mathbb{1} - A)^{-1}\right\| \le \frac{1}{\lambda}. \tag{7.2.13}$$

(iii) *For every $\lambda \in \mathbb{C}$ with $\mathrm{Re}\lambda > 0$ the operator $\lambda \mathbb{1} - A : \mathrm{dom}(A) \to X$ is bijective and satisfies the estimate*

$$\left\|(\lambda \mathbb{1} - A)^{-1}\right\| \le \frac{1}{\mathrm{Re}\lambda}. \tag{7.2.14}$$

(iv) *The operator A is dissipative and there exists a $\lambda > 0$ such that the operator $\lambda \mathbb{1} - A : \mathrm{dom}(A) \to X$ has a dense image.*

Proof. The equivalence of (i), (ii), and (iii) follows from Corollary 7.2.8 with $M = 1$ and $\omega = 0$. We prove the remaining implications in three steps.

Step 1. *If A is dissipative, then*

$$\|\lambda x - Ax\| \ge \mathrm{Re}\lambda \, \|x\| \tag{7.2.15}$$

for all $x \in \mathrm{dom}(A)$ and all $\lambda \in \mathbb{C}$ with $\mathrm{Re}\lambda > 0$.

Let $x \in \mathrm{dom}(A)$ and $\lambda \in \mathbb{C}$ such that $\mathrm{Re}\lambda > 0$. Since A is dissipative, there exists an element $x^* \in X^*$ such that (7.2.11) holds. This implies

$$\begin{aligned}
\|x\| \, \|\lambda x - Ax\| &= \|x^*\| \, \|\lambda x - Ax\| \\
&\ge \mathrm{Re}\langle x^*, \lambda x - Ax\rangle \\
&= \mathrm{Re}\lambda \langle x^*, x\rangle - \mathrm{Re}\langle x^*, Ax\rangle \\
&\ge \mathrm{Re}\lambda \, \|x\|^2 .
\end{aligned}$$

Hence

$$\|\lambda x - Ax\| \ge \mathrm{Re}\lambda \, \|x\|$$

and this proves Step 1.

Step 2. *We prove that (iv) implies (iii).*

Assume A satisfies (iv) and define the set

$$\Omega = \{\lambda \in \mathbb{C} \,|\, \mathrm{Re}\lambda > 0 \text{ and } \lambda\mathbb{1} - A \text{ has a dense image}\}.$$

This set is nonempty by (iv). Moreover, it follows from Step 1 that the operator $\lambda\mathbb{1} - A : \mathrm{dom}(A) \to X$ is injective and has a closed image for every $\lambda \in \mathbb{C}$ with $\mathrm{Re}\lambda > 0$. Hence $\Omega \subset \rho(A)$ and

$$\left\|(\lambda\mathbb{1} - A)^{-1}\right\| \le \frac{1}{\mathrm{Re}\lambda} \qquad \text{for all } \lambda \in \Omega \subset \rho(A). \tag{7.2.16}$$

If $\lambda \in \Omega$ and $|\mu - \lambda| < \mathrm{Re}\lambda$, then $\mathrm{Re}\mu > 0$ and $|\mu - \lambda|\|(\lambda\mathbb{1} - A)^{-1}\| < 1$, hence $\mu \in \rho(A)$ by Lemma 6.1.10, and hence $\mu \in \Omega$. Thus

$$\lambda \in \Omega \text{ and } |\mu - \lambda| < \mathrm{Re}\lambda \qquad \Longrightarrow \qquad \mu \in \Omega. \tag{7.2.17}$$

Fix an element $\lambda \in \Omega$. Then it follows from (7.2.17) that

$$\{\mu \in \mathbb{C} \,|\, \mathrm{Im}\mu = \mathrm{Im}\lambda,\, 0 < \mathrm{Re}\mu < 2\mathrm{Re}\lambda\} \subset \Omega.$$

Thus an induction argument shows that

$$\{\mu \in \mathbb{C} \,|\, \mathrm{Im}\mu = \mathrm{Im}\lambda,\, \mathrm{Re}\mu > 0\} \subset \Omega.$$

Hence it follows from (7.2.17) that $B_{\mathrm{Re}\mu}(\mu) \subset \Omega$ for every $\mu \in \mathbb{C}$ such that $\mathrm{Im}\mu = \mathrm{Im}\lambda$ and $\mathrm{Re}\mu > 0$. The union of these open discs is the entire positive half-plane in $\mathbb{C}$. Thus $\{z \in \mathbb{C} \,|\, \mathrm{Re}z > 0\} = \Omega \subset \rho(A)$ and hence it follows from (7.2.16) that A satisfies (iii). This proves Step 2.

Step 3. *We prove that (i) implies (iv).*

Assume that $A : \mathrm{dom}(A) \to X$ is the infinitesimal generator of a contraction semigroup $S : [0, \infty) \to \mathcal{L}^c(X)$. Let $x \in \mathrm{dom}(A)$. By the Hahn–Banach Theorem (Corollary 2.3.23) there exists an element $x^* \in X^*$ such that

$$\|x^*\|^2 = \|x\|^2 = \langle x^*, x\rangle.$$

Since S is a contraction semigroup this implies

$$\mathrm{Re}\langle x^*, S(h)x - x\rangle \le \|x^*\|\|S(h)x\| - \|x\|^2 \le 0$$

for all $h > 0$ and hence

$$\mathrm{Re}\langle x^*, Ax\rangle = \lim_{h\to 0} \frac{\mathrm{Re}\langle x^*, S(h)x - x\rangle}{h} \le 0.$$

This proves Step 3 and Theorem 7.2.11. □

7.3. The Dual Semigroup

When $S : [0, \infty) \to \mathcal{L}(X)$ is a strongly continuous semigroup on a real Banach space X the dual operators define a semigroup

$$S^* : [0, \infty) \to \mathcal{L}(X^*),$$

called the **dual semigroup**. One might expect that the dual semigroup is again strongly continuous, however, an elementary example shows that this need not always be the case (see Example 7.3.3 below). The failure of strong continuity of the dual semigroup is related to the fact that the Banach space X in Example 7.3.3 is not reflexive. On a reflexive Banach space it turns out that the dual semigroup is always strongly continuous and this is the content of Corollary 7.3.2 below, which will be derived as a consequence of the main theorem about the dual semigroup. The other subsections deal with self-adjoint semigroups and with unitary groups.

7.3.1. The Dual Semigroup and its Infinitesimal Generator. The following theorem is the main result of the present section. It was proved in 1955 by R. S. Phillips [**69**].

Theorem 7.3.1 (Phillips). *Let $S : [0, \infty) \to \mathcal{L}(X)$ be a strongly continuous semigroup on a real Banach space X and let $A : \mathrm{dom}(A) \to X$ be its infinitesimal generator. Denote by*

$$[0, \infty) \to \mathcal{L}(X^*) : t \mapsto S^*(t) := S(t)^*$$

the dual semigroup and by

$$E := \left\{ x^* \in X^* \;\middle|\; \begin{array}{l} \text{there exists a sequence } x_i^* \in \mathrm{dom}(A^*) \\ \text{such that } \lim_{i\to\infty} \|x_i^* - x^*\| = 0 \end{array} \right\} \tag{7.3.1}$$

the strong closure of the domain of the dual operator $A^* : \mathrm{dom}(A^*) \to X^*$. Then the following hold.

(i) *Let $x^* \in X^*$. Then $x^* \in E$ if and only if $\lim_{t\to 0} \|S^*(t)x^* - x^*\| = 0$.*

(ii) *The closed subspace $E \subset X^*$ is invariant under the operator $S^*(t)$ for every $t \geq 0$ and the map $T : [0, \infty) \to \mathcal{L}(E)$, defined by*

$$T(t) := S^*(t)|_E \qquad \text{for } t \geq 0,$$

is a strongly continuous semigroup.

(iii) *The infinitesimal generator of the strongly continuous semigroup T in part (ii) is the operator $B : \mathrm{dom}(B) \to E$ with*

$$\mathrm{dom}(B) = \{x^* \in \mathrm{dom}(A^*) \mid A^*x^* \in E\}$$

and $Bx^* = A^*x^*$ for $x^* \in \mathrm{dom}(B)$.

Proof. It follows directly from Lemma 4.1.3 that S^* is a one-parameter semigroup. The remaining assertions are proved in eight steps.

Step 1. *Let $x^* \in X^*$ and $h > 0$ and define the element $x_h^* \in X^*$ by*

$$\langle x_h^*, x\rangle = \frac{1}{h}\int_0^h \langle x^*, S(t)x\rangle\, dt \tag{7.3.2}$$

for $x \in X$. Then $x_h^* \in \mathrm{dom}(A^*)$ and $A^*x_h^* = h^{-1}(S^*(h)x^* - x^*)$.

Let $M := \sup_{0\le t\le h}\|S(t)\|$. The functional $X \to \mathbb{R} : x \mapsto \frac{1}{h}\int_0^h \langle x^*, S(t)x\rangle dt$ is linear and satisfies the inequality

$$\begin{aligned}\left|\frac{1}{h}\int_0^h \langle x^*, S(t)x\rangle\, dt\right| &\le \frac{1}{h}\int_0^h |\langle x^*, S(t)x\rangle|\, dt \\ &\le \frac{1}{h}\int_0^h \|x^*\|\,\|S(t)x\|\, dt \\ &\le M\,\|x^*\|\,\|x\|\end{aligned}$$

for all $x \in X$. Hence (7.3.2) defines an element $x_h^* \in X^*$. For $x \in \mathrm{dom}(A)$ this element satisfies the equation

$$\begin{aligned}\langle x_h^*, Ax\rangle &= \left\langle x^*, \int_0^h \frac{S(t)Ax}{h} dt\right\rangle \\ &= \left\langle x^*, \frac{S(h)x - x}{h}\right\rangle \\ &= \left\langle \frac{S^*(h)x^* - x^*}{h}, x\right\rangle.\end{aligned}$$

Here the second step follows from Lemma 7.1.13. This implies $x_h^* \in \mathrm{dom}(A^*)$ and $A^*x_h^* = h^{-1}(S^*(h)x^* - x^*)$. This proves Step 1.

Step 2. *Let $x^* \in \mathrm{dom}(A^*)$ and $t > 0$. Then $S^*(t)x^* \in \mathrm{dom}(A^*)$ and*

$$A^*S^*(t)x^* = S^*(t)A^*x^*.$$

If $x \in \mathrm{dom}(A)$, then $S(t)x \in \mathrm{dom}(A)$ and $S(t)Ax = AS(t)x$ by Lemma 7.1.13, and hence $\langle S^*(t)A^*x^*, x\rangle = \langle A^*x^*, S(t)x\rangle = \langle x^*, AS(t)x\rangle = \langle S^*(t)x^*, Ax\rangle$. By definition of the dual operator, this implies that $S^*(t)x^* \in \mathrm{dom}(A^*)$ and $A^*S^*(t)x^* = S^*(t)A^*x^*$. This proves Step 2.

Step 3. *Let $x^* \in E$ and $t \ge 0$. Then $S^*(t)x^* \in E$.*

Choose a sequence $x_i^* \in \mathrm{dom}(A^*)$ such that $\lim_{i\to\infty}\|x_i^* - x^*\| = 0$. Then it follows from Step 2 that $S^*(t)x_i^* \in \mathrm{dom}(A^*)$. Since $S^*(t) : X^* \to X^*$ is a bounded linear operator, we also have $\lim_{i\to\infty}\|S^*(t)x_i^* - S^*(t)x^*\| = 0$, and hence $S^*(t)x^* \in E$. This proves Step 3.

Step 4. *Let $x^* \in \mathrm{dom}(A^*)$ and $x \in X$. Then*

$$\langle S^*(t)x^* - x^*, x\rangle = \int_0^t \langle S^*(s)A^*x^*, x\rangle\, ds.$$

By Example 7.1.15 we have

$$\int_0^t S(s)x\, ds \in \mathrm{dom}(A), \qquad A\int_0^t S(s)x\, ds = S(t)x - x,$$

and hence

$$\begin{aligned}\langle S^*(t)x^* - x^*, x\rangle &= \langle x^*, S(t)x - x\rangle \\ &= \langle x^*, A\int_0^t S(s)x\, ds\rangle \\ &= \langle A^*x^*, \int_0^t S(s)x\, ds\rangle \\ &= \int_0^t \langle A^*x^*, S(s)x\rangle\, ds \\ &= \int_0^t \langle S^*(s)A^*x^*, x\rangle\, ds.\end{aligned}$$

Here the fourth equality follows from Lemma 5.1.8. This proves Step 4.

Step 5. *If $x^* \in E$, then $\lim_{t\to 0} \|S^*(t)x^* - x^*\| = 0$.*

Define $M := \sup_{0\le t\le 1} \|S(t)\|$ and let $x^* \in \mathrm{dom}(A^*)$. Then, by Step 4,

$$\begin{aligned}\langle S^*(t)x^* - x^*, x\rangle &= \int_0^t \langle A^*x^*, S(s)x\rangle\, ds \\ &\le \|A^*x^*\| \int_0^t \|S(s)x\|\, ds \\ &\le tM\,\|A^*x^*\|\,\|x\|\end{aligned}$$

for $0 \le t \le 1$. This implies

$$\|S^*(t)x^* - x^*\| = \sup_{x\in X\setminus\{0\}} \frac{\langle S^*(t)x^* - x^*, x\rangle}{\|x\|} \le tM\,\|A^*x^*\|$$

for $0 \le t \le 1$ and so $\lim_{t\to 0} \|S^*(t)x^* - x^*\| = 0$. Since $\mathrm{dom}(A^*)$ is dense in E and $\|S^*(t)\| = \|S(t)\| \le M$ for $0 \le t \le 1$, it follows from the Banach–Steinhaus Theorem 2.1.5 that

$$\lim_{t\to 0} \|S^*(t)x^* - x^*\| = 0 \qquad \text{for all } x^* \in E.$$

This proves Step 5.

Step 6. *Let* $x^* \in X^*$ *such that* $\lim_{t\to 0}\|S^*(t)x^* - x^*\| = 0$. *Then* $x^* \in E$.

For $h > 0$ let $x_h^* \in X^*$ be as in Step 1. Then $x_h^* \in \mathrm{dom}(A^*)$ and

$$\langle x_h^* - x^*, x\rangle = \frac{1}{h}\int_0^h \langle x^*, S(t)x - x\rangle\, dt.$$

Now fix a constant $\varepsilon > 0$ and choose $\delta > 0$ such that

$$0 \le t < \delta \qquad \Longrightarrow \qquad \|S^*(t)x^* - x^*\| < \varepsilon.$$

Let $0 < h < \delta$. Then

$$\begin{aligned}\frac{\langle x^*, S(t)x - x\rangle}{\|x\|} &= \frac{\langle S^*(t)x^* - x^*, x\rangle}{\|x\|}\\ &\le \|S^*(t)x^* - x^*\|\\ &\le \varepsilon\end{aligned}$$

for $0 \le t \le h$ and $x \in X \setminus \{0\}$, and hence

$$\frac{\langle x_h^* - x^*, x\rangle}{\|x\|} = \frac{1}{h}\int_0^h \frac{\langle x^*, S(t)x - x\rangle}{\|x\|}\, dt \le \varepsilon.$$

Take the supremum over all $x \in X \setminus \{0\}$ to obtain the inequality

$$\|x_h^* - x^*\| = \sup_{x\in X\setminus\{0\}} \frac{\langle x_h^* - x^*, x\rangle}{\|x\|} \le \varepsilon$$

for $0 < h < \delta$. Thus we have proved that

$$\lim_{h\to 0}\|x_h^* - x^*\| = 0,$$

and hence $x^* \in E$. This proves Step 6.

Step 7. *Let* $x^* \in \mathrm{dom}(A^*)$ *such that* $y^* := A^*x^* \in E$. *Then*

$$\lim_{t\to 0}\left\|\frac{S^*(t)x^* - x^*}{t} - y^*\right\| = 0.$$

By Step 3 and Step 5, S^* restricts to a strongly continuous semigroup on the subspace E. Thus the function $[0,\infty) \to E : t \mapsto S^*(t)y^* = S^*(t)A^*x^*$ is continuous and so

$$S^*(t)x^* - x^* = \int_0^t S^*(s)y^*\, ds$$

for all $t > 0$ by Step 4. Hence

$$\left\|\frac{S^*(t)x^* - x^*}{t} - y^*\right\| = \left\|\frac{1}{t}\int_0^t \Big(S^*(s)y^* - y^*\Big)\, ds\right\| \le \sup_{0\le s\le t}\|S^*(s)y^* - y^*\|$$

and this proves Step 7.

Step 8. *Let* $x^*, y^* \in X^*$ *such that* $\lim_{h\to 0}\left\|\frac{S^*(t)x^*-x^*}{t} - y^*\right\| = 0$. *Then*

$$x^* \in \mathrm{dom}(A^*), \qquad y^* = A^*x^* \in E.$$

It follows from the assumptions of Step 8 that $\lim_{t\to 0}\|S^*(t)x^* - x^*\| = 0$ and hence $x^* \in E$ by Step 6. This implies $t^{-1}(S^*(t)x^* - x^*) \in E$ by Step 3, and so $y^* \in E$ because E is a closed subspace of X^*. Since the function $[0,h] \to E : t \mapsto S^*(t)x^*$ is continuous by Step 3 and Step 5, the element $x_h^* \in X^*$ in Step 1 is given by

$$x_h^* = \frac{1}{h}\int_0^h S^*(t)x^*\,dt$$

and converges to x^* as h tends to zero. Moreover, by Step 1, we have that $x_h^* \in \mathrm{dom}(A^*)$ and $A^*x_h^* = h^{-1}(S^*(h)x^* - x^*)$ converges to y^* as h tends to zero. Since A^* is a closed operator, this implies $x^* \in \mathrm{dom}(A^*)$ and $A^*x^* = y^* \in E$. This proves Step 8.

Part (i) follows from Steps 5 and 6, part (ii) from Steps 3 and 5, and part (iii) from Steps 7 and 8. This proves Theorem 7.3.1. □

Corollary 7.3.2. *Let X be a real reflexive Banach space and let S be a strongly continuous semigroup on X with the infinitesimal generator A. Then the dual semigroup $S^* : [0,\infty) \to \mathcal{L}(X^*)$ is strongly continuous and its infinitesimal generator is the dual operator $A^* : \mathrm{dom}(A^*) \to X^*$.*

Proof. The domain of the dual operator A^* is weak* dense in X^* by part (iii) of Theorem 6.2.2, and so it is dense because X is reflexive. Hence the result follows from Theorem 7.3.1 with $E = X^*$. □

The shift group in the following example shows that Corollary 7.3.2 does not extend to nonreflexive Banach spaces. In Example 7.3.3 the subspace E is not invariant under A^* although it is invariant under $S^*(t)$ for all t.

Example 7.3.3. Let $X := L^1(\mathbb{R})$ and, for $t \in \mathbb{R}$, define the linear operator $S(t) : L^1(\mathbb{R}) \to L^1(\mathbb{R})$ for $t \in \mathbb{R}$ by

$$(S(t)f)(s) := f(s+t) \qquad \text{for } f \in L^1(\mathbb{R}) \text{ and } s,t \in \mathbb{R}.$$

Then $X^* \cong L^\infty(\mathbb{R})$ and under this identification the dual group is given by

$$(S^*(t)g)(s) := g(s-t) \qquad \text{for } g \in L^\infty(\mathbb{R}) \text{ and } s,t \in \mathbb{R}.$$

For a general element $g \in L^\infty(\mathbb{R})$ the function $\mathbb{R} \to L^\infty(\mathbb{R}) : t \mapsto S^*(t)g$ is weak* continuous but not continuous. In this example the domain of A^* is the space of bounded Lipschitz continuous functions on $\mathbb{R}$. This space is weak* dense in $L^\infty(\mathbb{R})$ but not dense. Its closure is the space $E \subset L^\infty(\mathbb{R})$ of bounded uniformly continuous functions on $\mathbb{R}$.

7.3.2. Self-Adjoint Semigroups. The next theorem characterizes the infinitesimal generators of self-adjoint semigroups.

Theorem 7.3.4 (Self-Adjoint Semigroups). *Let H be a real Hilbert space and let $A : \mathrm{dom}(A) \to H$ be a linear operator with a dense domain $\mathrm{dom}(A) \subset H$. Then the following are equivalent.*

(i) *The operator A is the infinitesimal generator of a strongly continuous semigroup $S : [0, \infty) \to \mathcal{L}(H)$ such that $S(t) = S(t)^*$ for all $t \geq 0$.*

(ii) *The operator A is self-adjoint and*

$$\sup_{x \in \mathrm{dom}(A) \setminus \{0\}} \frac{\langle x, Ax \rangle}{\|x\|^2} < \infty.$$

If these equivalent conditions are satisfied, then

$$\frac{\log\|S(t)\|}{t} = \sup_{\substack{x \in \mathrm{dom}(A) \\ x \neq 0}} \frac{\langle x, Ax \rangle}{\|x\|^2} \tag{7.3.3}$$

for all $t > 0$.

Proof. We prove that (i) implies (ii) and

$$\sup_{\substack{x \in \mathrm{dom}(A) \\ x \neq 0}} \frac{\langle x, Ax \rangle}{\|x\|^2} \leq \frac{\log\|S(t)\|}{t} = \lim_{s \to \infty} \frac{\log\|S(s)\|}{s} \qquad \text{for all } t > 0. \tag{7.3.4}$$

For Hilbert spaces Theorem 7.3.1 asserts that the adjoint A^* of the infinitesimal generator A of a semigroup S is the infinitesimal generator of the adjoint semigroup S^*. Since $S(t)^* = S(t)$ for all $t \geq 0$ in the case at hand, it follows that the infinitesimal generator A is self-adjoint. Moreover,

$$\|S(t)\| = \|S(t)^n\|^{1/n} = \|S(nt)\|^{1/n}$$

by part (i) of Theorem 5.3.15 and hence

$$\frac{\log\|S(t)\|}{t} = \frac{\log\|S(nt)\|}{nt} \qquad \text{for all } t > 0 \text{ and all } n \in \mathbb{N}.$$

Take the limit $n \to \infty$ and use Lemma 7.1.8 to obtain

$$\frac{\log\|S(t)\|}{t} = \omega_0 := \lim_{s \to \infty} \frac{\log\|S(s)\|}{s} \qquad \text{for all } t > 0.$$

This implies $\log\|S(t)\| = t\omega_0$ and so $\|S(t)\| = e^{t\omega_0}$ for all $t > 0$. Thus

$$\langle x, S(t)x \rangle \leq e^{t\omega_0}\|x\|^2 \qquad \text{for all } x \in H \text{ and all } t \geq 0.$$

Differentiate this inequality at $t = 0$ to obtain $\langle x, Ax \rangle \leq \omega_0\|x\|^2$ for every $x \in \mathrm{dom}(A)$. This shows that (i) implies (ii) and (7.3.4).

We prove that (ii) implies (i). Thus assume A is self-adjoint and

$$\omega := \sup_{\substack{x\in\mathrm{dom}(A)\\ x\neq 0}} \frac{\langle x, Ax\rangle}{\|x\|^2} < \infty.$$

We prove in five steps that A generates a self-adjoint semigroup.

Step 1. *If $\lambda > \omega$ and $x \in \mathrm{dom}(A)$, then $\|\lambda x - Ax\| \geq (\lambda - \omega)\|x\|$.*

Let $x \in \mathrm{dom}(A)$ and $\lambda > \omega$. Then $\langle x, Ax\rangle \leq \omega \|x\|^2$ and so

$$\|x\|\|\lambda x - Ax\| \geq \langle x, \lambda x - Ax\rangle \geq (\lambda - \omega)\|x\|^2.$$

This proves Step 1.

Step 2. *If $\lambda > \omega$, then $\lambda\mathbb{1} - A$ is injective and has a closed image.*

Let $\lambda > \omega$. Assume x_n is a sequence in $\mathrm{dom}(A)$ such that $y_n := \lambda x_n - Ax_n$ converges to y. Then x_n is a Cauchy sequence by Step 1 and so converges to some element $x \in H$. Hence $Ax_n = \lambda x_n - y_n$ converges to $\lambda x - y$. Since A has a closed graph by Theorem 6.2.2, this implies $x \in \mathrm{dom}(A)$ and $Ax = \lambda x - y$. Thus $y = \lambda x - Ax \in \mathrm{im}(\lambda\mathbb{1} - A)$, and so $\lambda\mathbb{1} - A$ has a closed image. That it is injective follows directly from the estimate in Step 1. This proves Step 2.

Step 3. *If $\lambda > \omega$, then $\lambda\mathbb{1} - A$ is surjective.*

Let $\lambda > \omega$ and suppose $y \in H$ is orthogonal to the image of $\lambda\mathbb{1} - A$. Then $\langle y, \lambda x\rangle = \langle y, Ax\rangle$ for all $x \in \mathrm{dom}(A)$. Hence $y \in \mathrm{dom}(A^*) = \mathrm{dom}(A)$ and $Ay = A^*y = \lambda y$. Thus $y = 0$ by Step 2. This shows that $\lambda\mathbb{1} - A$ has a dense image. Hence it is surjective by Step 2. This proves Step 3.

Step 4. *The operator A is the infinitesimal generator of a strongly continuous semigroup $S : [0,\infty) \to \mathcal{L}(H)$ such that $\|S(t)\| \leq e^{\omega t}$ for all $t \geq 0$.*

Let $\lambda > \omega$. Then $\lambda\mathbb{1} - A : \mathrm{dom}(A) \to H$ is bijective by Step 2 and Step 3 and $\|(\lambda\mathbb{1} - A)^{-1}\| \leq (\lambda - \omega)^{-1}$ by Step 1. Hence Step 4 follows from the Hille–Yosida–Phillips Theorem 7.2.5 with $M = 1$.

Step 5. *The semigroup S in Step 4 is self-adjoint and satisfies* (7.3.3)*.*

The operator $A = A^*$ is the infinitesimal generator of S by Step 4 and of the adjoint semigroup S^* by Theorem 7.3.1. Hence Corollary 7.2.3 asserts that $S(t) = S^*(t)$ for all $t \geq 0$. This implies that A and S satisfy (7.3.4). By (7.3.4), we have $\omega \leq t^{-1}\log\|S(t)\|$ and by Step 4 we have $\|S(t)\| \leq e^{\omega t}$ and hence $t^{-1}\log\|S(t)\| \leq \omega$ for all $t > 0$. Thus equality holds in (7.3.4). This proves (7.3.3) and Theorem 7.3.4. □

7.3.3. Unitary Groups. On complex Hilbert spaces it is interesting to examine the infinitesimal generators of strongly continuous unitary groups. This is the content of Theorem 7.3.6 below which was proved in 1932 by M. H. Stone [**81**].

Definition 7.3.5. Let H be a complex Hilbert space. A strongly continuous group $S : \mathbb{R} \to \mathcal{L}^c(H)$ is called **unitary** if $\|S(t)x\| = \|x\|$ for all $t \in \mathbb{R}$ and all $x \in H$ or, equivalently,

$$S^*(t) = S(t)^{-1} = S(-t)$$

for all $t \in \mathbb{R}$, where $S^*(t) = S(t)^*$ denotes the adjoint operator of $S(t)$.

Theorem 7.3.6 (**Stone**). *Let H be a complex Hilbert space and suppose that $A : \mathrm{dom}(A) \to H$ is a linear operator with a dense domain $\mathrm{dom}(A) \subset H$. Then the following are equivalent.*

(i) *A is the infinitesimal generator of a unitary group.*

(ii) *The operator $\mathbf{i}A : \mathrm{dom}(A) \to H$ is self-adjoint.*

Proof. We prove that (i) implies (ii). Thus assume that A is the infinitesimal generator of a unitary group $S : \mathbb{R} \to \mathcal{L}^c(H)$. Then

$$S^*(t) = S(t)^{-1} = S(-t) \qquad \text{for all } t \in \mathbb{R}.$$

The operator $-A : \mathrm{dom}(A) \to H$ is the infinitesimal generator of the group $\mathbb{R} \to \mathcal{L}^c(H) : t \mapsto S(-t)$ by Theorem 7.2.4 and $A^* : \mathrm{dom}(A^*) \to H$ is the infinitesimal generator of the group $\mathbb{R} \to \mathcal{L}^c(H) : t \mapsto S^*(t)$ by Theorem 7.3.1. Hence

$$A^* = -A$$

and so

$$(\mathbf{i}A)^* = -\mathbf{i}A^* = \mathbf{i}A.$$

Thus $\mathbf{i}A$ is self-adjoint.

We prove that (ii) implies (i). Suppose that

$$A = \mathbf{i}B,$$

where $B : \mathrm{dom}(B) \to H$ is a complex linear self-adjoint operator. Then A has a dense domain $\mathrm{dom}(A) = \mathrm{dom}(B)$ and a closed graph. Moreover,

$$A^* = (\mathbf{i}B)^* = -\mathbf{i}B^* = -\mathbf{i}B = -A.$$

This implies

$$\mathrm{Re}\langle x, Ax\rangle = \frac{\langle x, Ax\rangle + \langle Ax, x\rangle}{2} = \frac{\langle x, (A + A^*)x\rangle}{2} = 0 \tag{7.3.5}$$

for all $x \in \mathrm{dom}(A)$.

We prove that the operator $\mathbb{1} - A : \mathrm{dom}(A) \to H$ has a dense image. Assume that $y \in H$ is orthogonal to the image of $\mathbb{1} - A$. Then

$$0 = \langle y, x - Ax \rangle = \langle y, x \rangle - \langle y, Ax \rangle \qquad \text{for all } x \in \mathrm{dom}(A).$$

Hence it follows from the definition of the adjoint operator that

$$y \in \mathrm{dom}(A^*) = \mathrm{dom}(A), \qquad y = A^* y = -Ay.$$

This implies $\|y\|^2 = -\langle y, Ay \rangle = -\langle A^* y, y \rangle = -\|y\|^2$ and so $y = 0$. Hence the operator $\mathbb{1} - A$ has a dense image by the Hahn–Banach Theorem (Corollary 2.3.25).

Since $\mathbb{1} - A$ has a dense image it follows from (7.3.5) and the Lumer–Phillips Theorem 7.2.11 that A is the infinitesimal generator of a contraction semigroup $S : [0, \infty) \to \mathcal{L}^c(H)$. The adjoint semigroup $S^* : [0, \infty) \to \mathcal{L}^c(H)$ is also a contraction semigroup and is generated by the operator A^* by Theorem 7.3.1. Hence $-A = A^*$ is the infinitesimal generator of the semigroup S^* and so S extends to a strongly continuous group $S : \mathbb{R} \to \mathcal{L}^c(H)$ by Theorem 7.2.4. Since S^* is the group generated by $-A = A^*$ it follows that $S(t)^{-1} = S(-t) = S^*(t)$ for all $t \in \mathbb{R}$ and this proves Theorem 7.3.6. □

Example 7.3.7 (Shift Group). Consider the Hilbert space

$$H := L^2(\mathbb{R}, \mathbb{C})$$

and define the operator $A : \mathrm{dom}(A) \to H$ by

$$\begin{aligned} \mathrm{dom}(A) &:= W^{1,2}(\mathbb{R}, \mathbb{C}) \\ &:= \left\{ f : \mathbb{R} \to \mathbb{C} \,\middle|\, \begin{array}{l} f \text{ is absolutely continuous} \\ \text{and } f, \frac{df}{ds} \in L^2(\mathbb{R}, \mathbb{C}) \end{array} \right\}, \\ Af &:= \frac{df}{ds} \qquad \text{for } f \in W^{1,2}(\mathbb{R}, \mathbb{C}). \end{aligned} \tag{7.3.6}$$

Here s is the variable in $\mathbb{R}$. Recall that an absolutely continuous function is almost everywhere differentiable, that its derivative is locally integrable, and that it can be written as the integral of its derivative, i.e. the fundamental theorem of calculus holds in this setting (see [**75**, Thm. 6.19]). The operator

$$\mathbf{i}A = \mathbf{i}\frac{d}{ds} : W^{1,2}(\mathbb{R}, \mathbb{C}) \to L^2(\mathbb{R}, \mathbb{C})$$

is self-adjoint and hence A generates a unitary group $U : \mathbb{R} \to \mathcal{L}^c(L^2(\mathbb{R}, \mathbb{C}))$. This group is in fact the shift group in Example 7.1.4 given by

$$(U(t)f)(s) = f(s + t) \qquad \text{for } f \in L^2(\mathbb{R}, \mathbb{C}) \text{ and } s, t \in \mathbb{R}.$$

(See also Example 7.3.3 and Exercise 7.7.3.) **Exercise:** Verify the details.

Example 7.3.8 (**Schrödinger Equation**)**.** **(i)** Define the unbounded linear operator A on the Hilbert space $H := L^2(\mathbb{R}, \mathbb{C})$ by

$$\begin{aligned} \mathrm{dom}(A) &:= W^{2,2}(\mathbb{R}, \mathbb{C}) \\ &:= \left\{ f : \mathbb{R} \to \mathbb{C} \;\middle|\; \begin{array}{l} f \text{ is absolutely continuous and} \\ \frac{df}{dx} \text{ is absolutely continuous and} \\ \int_{-\infty}^{\infty} (|f|^2 + |\frac{df}{dx}|^2 + |\frac{d^2 f}{dx^2}|^2)\, dx < \infty \end{array} \right\}, \\ Af &:= \mathbf{i}\hbar \frac{d^2 f}{dx^2} \qquad \text{for } f \in W^{2,2}(\mathbb{R}, \mathbb{C}). \end{aligned} \tag{7.3.7}$$

(See Example 6.1.7.) Here $\hbar$ is a positive real number (**Planck's constant**) and x is the variable in $\mathbb{R}$. The operator

$$\mathbf{i}A = -\hbar \frac{d^2}{dx^2} : W^{2,2}(\mathbb{R}, \mathbb{C}) \to L^2(\mathbb{R}, \mathbb{C})$$

is self-adjoint and hence A generates a unitary group $U : \mathbb{R} \to \mathcal{L}^c(L^2(\mathbb{R}, \mathbb{C}))$. If $f : \mathbb{R} \to \mathbb{C}$ is a smooth function with compact support and $u : \mathbb{R}^2 \to \mathbb{C}$ is defined by $u(t, x) := (U(t)f)(x)$, then u satisfies the **Schrödinger equation**

$$\mathbf{i}\hbar \frac{\partial u}{\partial t} = -\hbar^2 \frac{\partial^2 u}{\partial x^2} \tag{7.3.8}$$

with the initial condition $u(0, \cdot) = f$. **Exercise:** Prove that the operator $\mathbf{i}A$ is self-adjoint.

(ii) Another variant of the Schrödinger equation is associated to the operator $A : \mathrm{dom}(A) \to L^2(\mathbb{R}, \mathbb{C})$, defined by

(7.3.9)

$$\begin{aligned} \mathrm{dom}(A) &:= \left\{ f : \mathbb{R} \to \mathbb{C} \;\middle|\; \begin{array}{l} f \text{ is absolutely continuous and} \\ \frac{df}{dx} \text{ is absolutely continuous and} \\ \int_{-\infty}^{\infty} (|f|^2 + |-\hbar^2 \frac{d^2 f}{dx^2} + x^2 f|^2)\, dx < \infty \end{array} \right\}, \\ (Af)(x) &:= \mathbf{i}\hbar \frac{d^2 f}{dx^2}(x) + \frac{x^2}{\mathbf{i}\hbar} f(x) \quad \text{for } f \in W^{2,2}(\mathbb{R}, \mathbb{C}) \text{ and } x \in \mathbb{R}. \end{aligned}$$

The operator $\mathbf{i}A$ is again self-adjoint and hence the operator A generates a unitary group $U : \mathbb{R} \to \mathcal{L}^c(L^2(\mathbb{R}, \mathbb{C}))$. If $f : \mathbb{R} \to \mathbb{C}$ is a smooth function with compact support and $u : \mathbb{R}^2 \to \mathbb{C}$ is defined by $u(t, x) := (U(t)f)(x)$, then u satisfies the **Schrödinger equation with quadratic potential**

$$\mathbf{i}\hbar \frac{\partial u}{\partial t}(t, x) = -\hbar^2 \frac{\partial^2 u}{\partial x^2}(t, x) + x^2 u(t, x) \tag{7.3.10}$$

with the initial condition $u(0, \cdot) = f$. **Exercise:** Prove that the operator $\mathbf{i}A$ is self-adjoint.

Corollary 7.3.9 (**Groups of Isometries**). *Let H be a real Hilbert space and suppose that $A : \mathrm{dom}(A) \to H$ is a linear operator with a dense domain $\mathrm{dom}(A) \subset H$. Then the following are equivalent.*

(i) *A is the infinitesimal generator of a group of isometries.*

(ii) *If $\lambda \in \mathbb{R} \setminus \{0\}$, then $\lambda \mathbb{1} - A$ is bijective and $\left\|(\lambda \mathbb{1} - A)^{-1}\right\| \le |\lambda|^{-1}$.*

(iii) $\mathrm{dom}(A^*) = \mathrm{dom}(A)$ *and $A^* x + Ax = 0$ for all $x \in \mathrm{dom}(A)$.*

Proof. By Theorem 7.2.4, a map $S : \mathbb{R} \to \mathcal{L}(H)$ is a strongly continuous group of isometries if and only if both $[0, \infty) \to \mathcal{L}(H) : t \mapsto S(t)$ and $[0, \infty) \to \mathcal{L}(H) : t \mapsto S(-t)$ are contraction semigroups. Hence the equivalence of (i) and (ii) follows from the Lumer–Phillips Theorem 7.2.11. The equivalence of (i) and (iii) follows from Theorem 7.3.6 for the complexified operator $A^c : \mathrm{dom}(A^c) := \mathrm{dom}(A)^c \to H^c$. □

Example 7.3.10 (**Shift Group**). **(i)** The formula $(L(t)f)(s) := f(s + t)$ for $s, t \in \mathbb{R}$ and $f \in H := L^2(\mathbb{R})$ defines a shift group $L : \mathbb{R} \to \mathcal{L}(H)$ of isometries. Its infinitesimal generator $A : \mathrm{dom}(A) = W^{1,2}(\mathbb{R}) \to L^2(\mathbb{R})$ is given by $Af = f'$ for $f \in W^{1,2}(\mathbb{R})$ and satisfies $A^* = -A$. (See Example 7.1.4.)

(ii) The formulas $(R(t)f)(s) := f(s - t)$ for $s \ge t \ge 0$ and $(R(t)f)(s) := 0$ for $t > s \ge 0$ and $f \in H := L^2([0, \infty))$ define a semigroup $R : [0, \infty) \to \mathcal{L}(H)$ of isometric embeddings. The infinitesimal generator $B : \mathrm{dom}(B) \to H$ has the domain $\mathrm{dom}(B) = W_0^{1,2}([0, \infty)) := \{f \in W^{1,2}([0, \infty)) \,|\, f(0) = 0\}$ and is given by $Bf = -f'$. Its adjoint has the domain $\mathrm{dom}(B^*) = W^{1,2}([0, \infty))$ and satisfies $Bf + B^* f = 0$ for $f \in \mathrm{dom}(B) \subsetneq \mathrm{dom}(B^*)$.

Example 7.3.11 (**Wave Equation**). **(i)** The group $\mathscr{S} : \mathbb{R} \to \mathcal{L}(\mathscr{H})$ on the Hilbert space $\mathscr{H} = L^2(\mathbb{R}, \mathbb{R}^2)$, given by (7.1.15) in Example 7.1.7, consists of isometries and has the infinitesimal generator $\mathscr{A} = -\mathscr{A}^*$ on $\mathscr{H}$, given by $\mathrm{dom}(\mathscr{A}) = W^{1,2}(\mathbb{R}, \mathbb{R}^2)$ and $\mathscr{A}(f, g) = (g', f')$.

(ii) Fix real numbers $a < b$ and consider the wave equation

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2}, \qquad u(t, a) = u(t, b) = 0, \tag{7.3.11}$$

on the compact interval $I := [a, b]$. Equation (7.3.11) gives rise to a strongly continuous group of isometries on the Hilbert space $H := W_0^{1,2}(I) \times L^2(I)$, where $W_0^{1,2}(I) := \left\{f \in W^{1,2}(I) \,|\, f(a) = f(b) = 0\right\}$ and

$$\|(f, g)\|_H := \sqrt{\int_a^b \Big(|f'(x)|^2 + |g(x)|^2\Big)\, dx}$$

for $f \in W_0^{1,2}(I)$ and $g \in L^2(I)$. Its infinitesimal generator is the operator

$$\mathrm{dom}(A) = (W^{2,2}(I) \cap W_0^{1,2}(I)) \times W_0^{1,2}(I), \qquad A(f, g) = (g, f'').$$

1.7. Problems

Exercise 1.7.1 (Precompact Sets). Let X and Y be topological spaces such that Y is Hausdorff. Let $f : X \to Y$ be a continuous map and let $A \subset X$ be a precompact subset of X (i.e. its closure $\overline{A}$ is compact). Prove that $B := f(A)$ is a precompact subset of Y. **Hint:** Show that $f(\overline{A}) \subset \overline{B}$. If $\overline{A}$ is compact and Y is Hausdorff show that $f(\overline{A}) = \overline{B}$.

Exercise 1.7.2 (Totally Bounded Sets). Let A be a subset of a metric space. Show that A is totally bounded if and only if $\overline{A}$ is totally bounded.

Exercise 1.7.3 (Complete and Closed Subspaces). Let (X, d_X) be a metric space, let $Y \subset X$ be a subset, and denote by $d_Y := d_X|_{Y \times Y}$ the induced distance function on Y. Prove the following.

(a) If (Y, d_Y) is complete, then Y is a closed subset of X.

(b) If (X, d_X) is complete and $Y \subset X$ is closed, then (Y, d_Y) is complete.

Exercise 1.7.4 (Completion of a Metric Space). Let (X, d) be a metric space. A **completion** of (X, d) is a triple $(\overline{X}, \overline{d}, \iota)$, consisting of a complete metric space $(\overline{X}, \overline{d})$ and an isometric embedding $\iota : X \to \overline{X}$ with a dense image.

(a) Every completion $(\overline{X}, \overline{d}, \iota)$ of (X, d) has the following **universality property**: If (Y, d_Y) is a complete metric space and $\phi : X \to Y$ is a **1-Lipschitz map** (i.e. a Lipschitz continuous map with Lipschitz constant one), then there exists a unique 1-Lipschitz map $\overline{\phi} : \overline{X} \to Y$ such that

$$\phi = \overline{\phi} \circ \iota.$$

(b) If $(\overline{X}_1, \overline{d}_1, \iota_1)$ and $(\overline{X}_2, \overline{d}_2, \iota_2)$ are completions of (X, d), then there exists a unique isometry $\psi : \overline{X}_1 \to \overline{X}_2$ such that $\psi \circ \iota_1 = \iota_2$.

(c) (X, d) admits a completion. **Hint:** The space $C_b(X)$ of bounded continuous functions $f : X \to \mathbb{R}$ is a Banach space with the supremum norm. Let $x_0 \in X$ and, for $x \in X$, define $f_x \in C_b(X)$ by

$$f_x(y) := d(y, x) - d(y, x_0) \qquad \text{for } y \in X.$$

Prove that the map $X \to C_b(X) : x \mapsto f_x$ is an isometric embedding and that the closure of its image is a completion of (X, d).

(d) Let $(\overline{X}, \overline{d})$ be a complete metric space and let $\iota : X \to \overline{X}$ be a 1-Lipschitz map that satisfies the universality property in (a). Prove that $(\overline{X}, \overline{d}, \iota)$ is a completion of (X, d).

Exercise 1.7.5 (Completion of a Normed Vector Space). The completion of a normed vector space is a Banach space.

Exercise 1.7.6 (Operator Norm). This exercise shows that the supremum in the definition of the operator norm need not be a maximum (see Definition 1.2.1). Consider the Banach space $X := C([-1,1])$ of continuous functions $f : [-1,1] \to \mathbb{R}$ equipped with the supremum norm and define the bounded linear functional

$$\Lambda : C([-1,1]) \to \mathbb{R}$$

by

$$\Lambda(f) := \int_0^1 f(t)\,dt - \int_{-1}^0 f(t)\,dt \qquad \text{for } f \in C([-1,1]).$$

Prove that there does not exist a function $f \in C([-1,1])$ such that $\|f\|_\infty = 1$ and $|\Lambda(f)| = \|\Lambda\| = 2$.

Exercise 1.7.7 (Continuously Differentiable Functions). Let $I := [0,1]$ be the unit interval and denote by $C^1(I)$ the space of continuously differentiable functions $f : I \to \mathbb{R}$ (with one-sided derivatives at $t = 0$ and $t = 1$). Define

$$\|f\|_{C^1} := \sup_{0\le t\le 1} |f(t)| + \sup_{0\le t\le 1} \left|f'(t)\right| \qquad \text{for } f \in C^1(I). \tag{1.7.1}$$

(a) Prove that $C^1(I)$ is a Banach space with the norm (1.7.1).

(b) Show that the inclusion $\iota : C^1(I) \to C(I)$ is a bounded linear operator.

(c) Let $B \subset C^1(I)$ be the unit ball. Show that $\iota(B)$ has compact closure.

(d) Is $\iota(B)$ a closed subset of $C(I)$?

(e) Does the linear operator $\iota : C^1(I) \to C(I)$ have a dense image?

Exercise 1.7.8 (Integration Against a Kernel). Let $I := [0,1]$, let $K : I \times I \to \mathbb{R}$ be a continuous function, and define the linear operator $T_K : C(I) \to C(I)$ by

$$(T_K f)(t) := \int_0^1 K(t,s) f(s)\,ds \qquad \text{for } f \in C(I) \text{ and } 0 \le t \le 1.$$

Prove that T_K is continuous. Let $B \subset C(I)$ be the unit ball and prove that its image $T_K(B)$ has a compact closure in $C(I)$.

Exercise 1.7.9 (Fekete's Lemma). Let $(\alpha_n)_{n\in\mathbb{N}}$ be a sequence of real numbers and suppose that there exists a constant $c \ge 0$ such that

$$\alpha_{n+m} \le \alpha_n + \alpha_m + c \qquad \text{for all } n,m \in \mathbb{N}.$$

Prove that $\lim_{n\to\infty} \alpha_n/n = \inf_{n\in\mathbb{N}} \alpha_n/n$. Here both sides of the equation may be minus infinity. Compare this with part (i) of Theorem 1.5.5 by taking $\alpha_n := \log \|a^n\|$.

Exercise 1.7.10 (The Inverse in a Unital Banach Algebra). Let $\mathcal{A}$ be a unital Banach algebra and let $a, b \in \mathcal{A}$ such that $\mathbb{1} - ab$ is invertible. Prove that $\mathbb{1} - ba$ is invertible. **Hint:** An explicit formula for the inverse of $\mathbb{1} - ba$ in terms of the inverse of $\mathbb{1} - ab$ can be guessed by expanding $(\mathbb{1} - ab)^{-1}$ and $(\mathbb{1} - ba)^{-1}$ formally as geometric series (Theorem 1.5.5).

Exercise 1.7.11 (Cantor's Intersection Theorem). The **diameter** of a nonempty subset A of a metric space (X, d) is defined by

$$\mathrm{diam}(A) := \sup_{x,y \in A} d(x, y). \tag{1.7.2}$$

(a) Prove that a metric space (X, d) is complete if and only if every nested sequence $A_1 \supset A_2 \supset A_3 \supset \cdots$ of nonempty closed subsets $A_n \subset X$ satisfying $\lim_{n\to\infty} \mathrm{diam}(A_n) = 0$ has a nonempty intersection (consisting of a single point).

(b) Find an example of a complete metric space and a nested sequence of nonempty closed bounded sets whose intersection is empty. **Hint:** Consider the unit sphere in an infinite-dimensional Hilbert space.

Exercise 1.7.12 (Convergence Along Arithmetic Sequences). Let $f : [0, \infty) \to \mathbb{R}$ be a continuous function such that

$$\lim_{n\to\infty} f(nt) = 0 \qquad \text{for all } t > 0.$$

Prove that

$$\lim_{x\to\infty} f(x) = 0.$$

Hint: Fix a constant $\varepsilon > 0$ and show that the set

$$A_n := \{t > 0 \,|\, |f(mt)| \le \varepsilon \text{ for every integer } m \ge n\}$$

has a nonempty interior for some $n \in \mathbb{N}$ (using the Baire Category Theorem 1.6.4). Assume without loss of generality that $[a, b] \subset A_n$ for $0 < a < b$ with $n(b - a) \ge a$. Deduce that $|f(x)| \le \varepsilon$ for all $x \ge na$.

Exercise 1.7.13 (Nowhere Differentiable Continuous Functions). Prove that the set

$$\mathcal{R} := \big\{ f : [0,1] \to \mathbb{R} \,\big|\, f \text{ is continuous and nowhere differentiable} \big\}$$

is residual in the Banach space $C([0, 1])$ and hence is dense. (This result is due to Stefan Banach and was proved in 1931.) **Hint:** Prove that the set

$$\mathcal{U}_n := \left\{ f \in C([0,1]) \,\middle|\, \sup_{\substack{0 \le s \le 1 \\ s \ne t}} \left| \frac{f(s) - f(t)}{s - t} \right| > n \text{ for all } t \in [0, 1] \right\}$$

is open and dense in $C([0, 1])$ for every $n \in \mathbb{N}$ and that $\bigcap_{n=1}^{\infty} \mathcal{U}_n \subset \mathcal{R}$.

The proof of the Baire Category Theorem uses the axiom of dependent choice. A theorem of Blair asserts that the Baire Category Theorem is equivalent to the axiom of dependent choice. That the axiom of dependent choice follows from the Baire Category Theorem is the content of the next exercise.

Exercise 1.7.14 (**Baire Category and Dependent Choice**). Let $\mathbf{X}$ be a nonempty set and let $\mathbf{A} : \mathbf{X} \to 2^{\mathbf{X}}$ be a map which assigns to each $\mathbf{x} \in \mathbf{X}$ a nonempty subset $\mathbf{A}(\mathbf{x}) \subset \mathbf{X}$. Use Theorem 1.6.4 to prove that there is a sequence $(\mathbf{x}_n)_{n\in\mathbb{N}}$ in $\mathbf{X}$ such that $\mathbf{x}_{n+1} \in \mathbf{A}(\mathbf{x}_n)$ for all $n \in \mathbb{N}$.

Hint: Denote by $\mathcal{X} := \mathbf{X}^{\mathbb{N}}$ the set of all sequences $\xi = (\mathbf{x}_n)_{n\in\mathbb{N}}$ in $\mathbf{X}$ and define the function $d : \mathcal{X} \times \mathcal{X} \to [0, \infty)$ by $d(\xi, \xi) := 0$ and

$$d(\xi, \eta) := 2^{-n}, \qquad n := \min\{k \in \mathbb{N} \,|\, \mathbf{x}_k \neq \mathbf{y}_k\},$$

for every pair of distinct sequences $\xi = (\mathbf{x}_n)_{n\in\mathbb{N}}, \eta = (\mathbf{y}_n)_{n\in\mathbb{N}} \in \mathcal{X}$. Prove that $(\mathcal{X}, d)$ is a complete metric space. For $k \in \mathbb{N}$ define

$$\mathcal{U}_k := \left\{ \xi = (\mathbf{x}_n)_{n\in\mathbb{N}} \in \mathbf{X}^{\mathbb{N}} \,\middle|\, \begin{array}{l} \text{there is an integer } \ell > k \\ \text{such that } \mathbf{x}_\ell \in \mathbf{A}(\mathbf{x}_k) \end{array} \right\}.$$

Prove that $\mathcal{U}_k$ is a dense open subset of $\mathcal{X}$ for every $k \in \mathbb{N}$ and deduce that the set $\mathcal{R} := \bigcap_{k\in\mathbb{N}} \mathcal{U}_k$ is nonempty. Construct the desired sequence as a suitable subsequence of an element $\xi = (\mathbf{x}_n)_{n\in\mathbb{N}} \in \mathcal{R}$.

Exercise 1.7.15 (**Borel Measurable Linear Operators**).

(a) Sets with the Baire property. A subset B of a topological space is said to have the **Baire property** if there exists an open set U such that the symmetric difference $B \Delta U := (B \setminus U) \cup (U \setminus B)$ is meagre, i.e. B and U differ by a meagre set (see Definition 1.6.1). Prove that the collection of all sets with the Baire property is the smallest σ-algebra containing the Borel sets and the meagre sets.

(b) Pettis' Lemma. Let X be a Banach space and let $B \subset X$ be a nonmeagre subset that has the Baire property. Prove that the set $B - B$ is a neighborhood of the origin. In particular, if B is a linear subspace of X, then $B = X$. **Hint:** Let U be an open subset of X such that $B\Delta U$ is meagre. Show that $U \neq \emptyset$, fix an element $x \in U$, and find an open neighborhood V of the origin such that $x + V - V \subset U$. For every $v \in V$ show that $U \cap (v + U) \neq \emptyset$ and deduce that $B \cap (v + B) \neq \emptyset$.

(c) Borel measurable linear operators. Let $f : X \to Y$ be a Borel measurable linear operator from a Banach space X to a separable normed vector space Y. Prove that f is continuous. **Hint:** $B := \{x \in X \mid \|f(x)\|_Y < 1/2\}$ is a nonmeagre Borel set.

2.5. Problems

Exercise 2.5.1 (Phillips' Lemma). Prove that the subspace

$$c_0 \subset \ell^\infty$$

of all sequences of real numbers that converge to zero is not complemented. This result is due to Phillips [**65**]. The hints are based on [**3**, p. 45].

Hint 1: *There exists an uncountable collection $\{A_i\}_{i\in I}$ of infinite subsets $A_i \subset \mathbb{N}$ such that $A_i \cap A_{i'}$ is a finite set for all $i, i' \in I$ such that $i \neq i'$.*

For example, take

$$I := \mathbb{R} \setminus \mathbb{Q},$$

choose a bijection $\mathbb{N} \to \mathbb{Q} : n \mapsto a_n$, choose sequences $(n_{i,k})_{k\in\mathbb{N}}$ in $\mathbb{N}$, one for each $i \in I$, such that $\lim_{k\to\infty} a_{n_{i,k}} = i$ for all $i \in I = \mathbb{R} \setminus \mathbb{Q}$, and define

$$A_i := \{n_{i,k} \,|\, k \in \mathbb{N}\} \subset \mathbb{N} \qquad \text{for } i \in I.$$

Hint 2: *Let $Q : \ell^\infty \to \ell^\infty$ be a bounded linear operator with $c_0 \subset \ker(Q)$. Then there exists an infinite subset $A \subset \mathbb{N}$ such that $Q(x) = 0$ for every sequence $x = (x_j)_{j\in\mathbb{N}} \in \ell^\infty$ that satisfies $x_j = 0$ for all $j \in \mathbb{N} \setminus A$.*

The set A can be taken as one of the sets A_i in Hint 1. Argue by contradiction and suppose that, for each $i \in I$, there exists a sequence

$$x_i = (x_{ij})_{j\in\mathbb{N}} \in \ell^\infty$$

such that

$$Q(x_i) \neq 0, \qquad \|x_i\|_\infty = 1, \qquad x_{ij} = 0 \text{ for all } j \in \mathbb{N} \setminus A_i.$$

Define the maps $Q_n : \ell^\infty \to \mathbb{R}$ by $Q(x) =: (Q_n(x))_{n\in\mathbb{N}}$ for $x \in \ell^\infty$. For each pair of integers $n, k \in \mathbb{N}$ define the set

$$I_{n,k} := \{i \in I \,|\, |Q_n(x_i)| \geq 1/k\}.$$

Fix a finite set $I' \subset I_{n,k}$ and consider the value of the operator Q on the element

$$x := \sum_{i\in I'} \varepsilon_i x_i, \qquad \varepsilon_i := \operatorname{sign}(Q_n(x_i)).$$

Use the fact that the set

$$B := \{j \in \mathbb{N} \,|\, \exists\, i, i' \in I' \text{ such that } i \neq i' \text{ and } x_{ij} \neq 0 \neq x_{i'j}\}$$

is finite to deduce that $|Q_n(x)| \leq \|Q(x)\| \leq \|Q\|$ and so

$$\#I_{n,k} \leq k\,\|Q\| \qquad \text{for all } n, k \in \mathbb{N}.$$

This contradicts the fact that the set $I = \bigcup_{n,k\in\mathbb{N}} I_{n,k}$ is uncountable.

Hint 3: *There is no bounded linear operator $Q : \ell^\infty \to \ell^\infty$ with $\ker(Q) = c_0$.*

Exercise 2.5.2 (Uniform Boundedness and Open Mappings). The Uniform Boundedness Principle, the Open Mapping Theorem, and the Closed Graph Theorem do not extend to normed vector spaces that are not complete. Let $X = \mathbb{R}^\infty$ be the vector space of sequences $x = (x_i)_{i\in\mathbb{N}}$ of real numbers with only finitely many nonzero terms. For $x \in X$ define

$$\|x\|_1 := \sum_{i=1}^{\infty} |x_i|\,, \qquad \|x\|_\infty := \sup_{i\in\mathbb{N}} |x_i|_\infty\,.$$

Prove the following.

(a) For $n \in \mathbb{N}$ define the linear functional $\Lambda_n : X \to \mathbb{R}$ by $\Lambda_n(x) := nx_n$. Then Λ_n is bounded for all $n \in \mathbb{N}$ and $\sup_{n\in\mathbb{N}} |\Lambda_n(x)| < \infty$ for all $x \in X$, however, $\sup_{n\in\mathbb{N}} \|\Lambda_n\|_{X^*} = \infty$ (for either norm on X).

(b) The identity operator $\mathrm{id} : (X, \|\cdot\|_1) \to (X, \|\cdot\|_\infty)$ is bounded but does not have a bounded inverse.

(c) The identity operator $\mathrm{id} : (X, \|\cdot\|_\infty) \to (X, \|\cdot\|_1)$ has a closed graph but is not bounded.

Exercise 2.5.3 (Zabreĭko's Lemma).
(a) Prove **Zabreĭko's Lemma**. *Let X be a Banach space and let $p : X \to \mathbb{R}$ be a seminorm. Then the following are equivalent.*

(i) *p is continuous.*

(ii) *There exists a constant $c > 0$ such that $p(x) \le c\,\|x\|$ for all $x \in X$.*

(iii) *The seminorm p is countably subadditive, i.e.*

$$p\left(\sum_{i=1}^{\infty} x_i\right) \le \sum_{i=1}^{\infty} p(x_i)$$

for every absolutely convergent series $x = \sum_{i=1}^{\infty} x_i$ in X.

Hint: See Definition 2.3.1 for seminorms and Lemma 1.5.1 for absolutely convergent series. To prove that (iii) implies (ii), define the sets

$$A_n := \{x \in X \,|\, p(x) \le n\}\,, \qquad F_n := \overline{\{x \in X \,|\, p(x) \le n\}}$$

for $n \in \mathbb{N}$. Show that F_n is convex and symmetric for each n. Use the Baire Category Theorem 1.6.4 to prove that there exists an $n \in \mathbb{N}$ such that F_n contains the open unit ball $B := \{x \in X \mid \|x\| < 1\}$. Prove that $B \subset A_n$ by mimicking the proof of the open mapping theorem (Lemma 2.2.3).

(b) Deduce the Uniform Boundedness Principle, the Open Mapping Theorem, and the Closed Graph Theorem from Zabreĭko's Lemma.

Exercise 2.5.4 (Complex Hahn–Banach). The dual space of a complex normed vector space X is the space of bounded complex linear functionals $x^* : X \to \mathbb{C}$. Adapt Corollaries 2.3.23–2.3.26 of the Hahn–Banach Theorem and their proofs to complex normed vector spaces.

Exercise 2.5.5 (Fourier Series of Continuous Functions). This exercise shows that there exist continuous functions whose Fourier series do not converge uniformly. Denote by $C(\mathbb{R}/2\pi\mathbb{Z}, \mathbb{C})$ the space of continuous 2π-periodic complex valued functions $f : \mathbb{R} \to \mathbb{C}$ equipped with the supremum norm.

(a) For $n \in \mathbb{N}$ the **Dirichlet kernel** $D_n \in C(\mathbb{R}/2\pi\mathbb{Z}, \mathbb{C})$ is defined by

$$D_n(t) := \sum_{k=-n}^{n} e^{\mathbf{i}kt} = \frac{\sin((n+\frac{1}{2})t)}{\sin(\frac{1}{2}t)} \qquad \text{for } t \in \mathbb{R}. \tag{2.5.1}$$

Prove that $\|D_n\|_{L^1} \geq \frac{8}{\pi}\sum_{k=1}^{n}\frac{1}{k}$.

(b) The nth Fourier expansion of a function $f \in C(\mathbb{R}/2\pi\mathbb{Z}, \mathbb{C})$ is defined by

$$(\mathcal{F}_n(f))(x) := (D_n * f)(x) = \sum_{k=-n}^{n} \int_0^{2\pi} f(t)e^{\mathbf{i}k(x-t)}\,dt \tag{2.5.2}$$

for $x \in \mathbb{R}$. Prove that the operator $\mathcal{F}_n : C(\mathbb{R}/2\pi\mathbb{Z}, \mathbb{C}) \to C(\mathbb{R}/2\pi\mathbb{Z}, \mathbb{C})$ has the operator norm $\|\mathcal{F}_n\| = \|D_n\|_{L^1}$.

(c) Deduce from the Uniform Boundedness Principle (Theorem 2.1.1) that there exists a function $f \in C(\mathbb{R}/2\pi\mathbb{Z}, \mathbb{C})$ such that the sequence $\mathcal{F}_n(f)$ does not converge uniformly.

Exercise 2.5.6 (Fourier Series of Integrable Functions). The **Fourier coefficients** of a function $f \in L^1([0, 2\pi], \mathbb{C})$ are given by

$$\widehat{f}(k) := \int_0^{2\pi} e^{-\mathbf{i}t} f(t)\,dt \qquad \text{for } k \in \mathbb{Z}, \tag{2.5.3}$$

and the **Fourier series of** f is $\mathscr{F}(f) := \big(\widehat{f}(k)\big)_{k\in\mathbb{Z}}$.

(a) Prove the **Riemann–Lebesgue Lemma**, which asserts that

$$\lim_{|k|\to\infty} \left|\widehat{f}(k)\right| = 0$$

for all $f \in L^1([0, 2\pi], \mathbb{C})$.

(b) Denote by $c_0(\mathbb{Z}, \mathbb{C}) \subset \ell^\infty(\mathbb{Z}, \mathbb{C})$ the closed subspace of all bi-infinite sequences of complex numbers that converge to zero as $|k|$ tends to infinity. Prove that the bounded linear operator $\mathscr{F} : L^1([0, 2\pi], \mathbb{C}) \to c_0(\mathbb{Z}, \mathbb{C})$ has a dense image but is not surjective. **Hint:** Investigate the Fourier coefficients of the Dirichlet kernels in Exercise 2.5.5.

Exercise 2.5.7 (**Banach Limits**). Let ℓ^∞ be the Banach space of bounded sequences of real numbers with the supremum norm as in part (ii) of Example 1.1.3 and define the shift operator $T : \ell^\infty \to \ell^\infty$ by

$$Tx := (x_{n+1})_{n\in\mathbb{N}} \qquad \text{for } x = (x_n)_{n\in\mathbb{N}} \in \ell^\infty.$$

Consider the subspace

$$Y := \mathrm{im}(\mathrm{id} - T) = \{x - Tx \,|\, x \in \ell^\infty\}.$$

Prove the following.

(a) The subspace $c_0 \subset \ell^\infty$ of all sequences that converge to zero is contained in the closure of Y.

(b) Let $\mathbb{1} = (1, 1, 1, \dots) \in \ell^\infty$ be the constant sequence with entries 1. Prove that $\sup_{n\in\mathbb{N}} |1 + x_{n+1} - x_n| \geq 1$ for all $x \in \ell^\infty$ and deduce that

$$d(\mathbb{1}, Y) = \inf_{y\in Y} \|\mathbb{1} - y\|_\infty = 1.$$

(c) By the Hahn–Banach Theorem 2.3.22 there exists a bounded linear functional $\Lambda : \ell^\infty \to \mathbb{R}$ such that

$$(2.5.4) \qquad \Lambda(\mathbb{1}) = 1, \qquad \|\Lambda\| = 1, \qquad \Lambda(x - Tx) = 0 \quad \text{for all } x \in \ell^\infty.$$

Prove that any such functional has the following properties.

(i) $\Lambda(Tx) = \Lambda(x)$ for all $x \in \ell^\infty$.

(ii) If $x \in \ell^\infty$ satisfies $x_n \geq 0$ for all $n \in \mathbb{N}$, then $\Lambda(x) \geq 0$.

(iii) $\liminf_{n\to\infty} x_n \leq \Lambda(x) \leq \limsup_{n\to\infty} x_n$ for all $x \in \ell^\infty$.

(iv) If $x \in \ell^\infty$ converges, then $\Lambda(x) = \lim_{n\to\infty} x_n$.

(d) Let Λ be as in (c). Find $x, y \in \ell^\infty$ such that $\Lambda(xy) \neq \Lambda(x)\Lambda(y)$. **Hint:** Consider the sequence $x_n := (-1)^n$ and show that $\Lambda(x) = 0$.

(e) Let Λ be as in (c). Prove that there does not exist a sequence $y \in \ell^1$ such that $\Lambda(x) = \sum_{n=1}^\infty x_n y_n$ for all $x \in \ell^\infty$. **Hint:** Any such sequence would have nonnegative entries $y_n \geq 0$ by part (ii) in (c) and satisfy $\sum_{n=1}^\infty y_n = 1$. Hence $\sum_{n=1}^N y_n > 0$ for some $N \in \mathbb{N}$ in contradiction to part (iv) in (c).

Exercise 2.5.8 (**Minkowski Functionals**). Let X be a normed vector space and let $C \subset X$ be a convex subset such that $0 \in C$. The **Minkowski functional** of C is the function

$$p : X \to [0, \infty]$$

defined by

$$(2.5.5) \qquad p(x) := \inf\left\{\lambda > 0 \,|\, \lambda^{-1}x \in C\right\} \qquad \text{for } x \in X.$$

The convex set C is called **absorbing** if, for every $x \in X$, there is a $\lambda > 0$ such that $\lambda^{-1}x \in C$. Let p be the Minkowski functional of C.

(a) Prove that $p(x+y) \le p(x) + p(y)$ and $p(\lambda x) = \lambda p(x)$ for all $x, y \in X$ and all $\lambda > 0$.

(b) Prove that C is absorbing if and only if p takes values in $[0, \infty)$ and hence is a quasi-seminorm (see Definition 2.3.1).

(c) Suppose C is absorbing. Find conditions on C which ensure that p is a seminorm or a norm. Do this both for real scalars and complex scalars.

(d) Prove that p is continuous if and only if zero is an interior point of C. In this case, show that $\mathrm{int}(C) = p^{-1}([0,1))$ and $\overline{C} = p^{-1}([0,1])$.

Exercise 2.5.9 (Reflexive Banach Spaces). Let X be a normed vector space and let $Y \subset X$ be a closed subspace. Assume Y and X/Y are reflexive. Prove that X is reflexive.

Exercise 2.5.10 (Schatten's Projective Tensor Product). Let X and Y be real normed vector spaces.

(a) For every normed vector space Z, the space $\mathcal{B}(X, Y; Z)$ of bounded bilinear maps $B : X \times Y \to Z$ is a normed vector space with the norm

$$\|B\| := \sup_{\substack{x \in X \setminus \{0\} \\ y \in Y \setminus \{0\}}} \frac{\|B(x,y)\|_Z}{\|x\|_X \, \|y\|_Y} \qquad \text{for } B \in \mathcal{B}(X, Y; Z).$$

(b) The map

$$\mathcal{B}(X, Y; Z) \to \mathcal{L}(X, \mathcal{L}(Y, Z)) : B \mapsto (x \mapsto B(x, \cdot))$$

is an isometric isomorphism.

(c) Associated to every pair $(x, y) \in X \times Y$ is a linear functional

$$x \otimes y \in \mathcal{B}(X, Y; \mathbb{R})^*$$

defined by $\langle x \otimes y, B \rangle := B(x, y)$ for $B \in \mathcal{B}(X, Y; \mathbb{R})$. It satisfies

$$\|x \otimes y\| = \|x\|_X \, \|y\|_Y \, .$$

Hint: Use the Hahn–Banach Theorem to prove the inequality $\|x \otimes y\| \ge \|x\|_X \, \|y\|_Y$. Namely, consider the bilinear functional $B : X \times Y \to \mathbb{R}$, defined by $B(x, y) := \langle x^*, x \rangle \langle y^*, y \rangle$ for suitable elements $x^* \in X^*$ and $y^* \in Y^*$ of the dual spaces.

(d) Let $X \otimes Y \subset \mathcal{B}(X, Y; \mathbb{R})^*$ be the smallest closed subspace containing the image of the bilinear map $X \times Y \to \mathcal{B}(X, Y; \mathbb{R})^* : (x, y) \mapsto x \otimes y$ in (c). Then, for every normed vector space Z, the map

$$\mathcal{L}(X \otimes Y, Z) \to \mathcal{B}(X, Y; Z) : A \mapsto B_A$$

defined by $B_A(x, y) := A(x \otimes y)$ for $x, y \in X$ and $A \in \mathcal{L}(X \otimes Y, Z)$ is an isometric isomorphism.

Exercise 2.5.11 (Strict Convexity and Hahn–Banach).

(a) Prove **Ruston's Theorem:** *The following properties of a normed vector space X are equivalent.*

(i) *If $x, y \in X$ satisfy $x \neq y$ and $\|x\| = \|y\| = 1$, then $\|x + y\| < 2$.*

(ii) *If $x, y \in X$ satisfy $x \neq 0 \neq y$ and $\|x + y\| = \|x\| + \|y\|$, then $x = \lambda y$ for some $\lambda > 0$.*

(iii) *If $x^* \in X^*$ is a nonzero bounded linear functional then there exists at most one element $x \in X$ such that $\|x\| = 1$ and $\langle x^*, x\rangle = \|x^*\|$.*

The normed vector space X is called **strictly convex** *if it satisfies these equivalent conditions.*

Condition (i) says that the unit sphere contains no nontrivial line segment. Condition (ii) says that equality in the triangle inequality occurs only in the trivial situation. Condition (iii) says that the support hyperplane $H_{x^*} := \{x \in X \,|\, \langle x^*, x\rangle = \|x^*\|\}$ meets the unit sphere in at most one point. (Note that $\inf_{x \in H_{x^*}} \|x\| = 1$.)

(b) For which p is $L^p([0,1])$ strictly convex? Is $C([0,1])$ strictly convex?

(c) If X is a normed vector space such that X^* is strictly convex, $Y \subset X$ is a linear subspace, and $y^* : Y \to \mathbb{R}$ is a bounded linear functional, then there is a unique $x^* \in X^*$ such that $x^*|_Y = y^*$ and $\|x^*\| = \|y^*\|$.

Exercise 2.5.12 (Schauder Bases). Let X be a separable real Banach space and let $(e_i)_{i\in\mathbb{N}}$ be a **Schauder basis** of X. This means that, for each element $x \in X$, there exists a unique sequence $(x_i)_{i\in\mathbb{N}}$ of real numbers such that

$$\lim_{n\to\infty} \Big\| x - \sum_{i=1}^{n} x_i e_i \Big\| = 0. \tag{2.5.6}$$

Let $n \in \mathbb{N}$ and define the map $\Pi_n : X \to X$ by

$$\Pi_n(x) := \sum_{i=1}^{n} x_i e_i \tag{2.5.7}$$

for $x \in X$, where $(x_i)_{i\in\mathbb{N}}$ is the unique sequence that satisfies (2.5.6).

(a) Prove that the operators $\Pi_n : X \to X$ are linear and satisfy

$$\Pi_n \circ \Pi_m = \Pi_m \circ \Pi_n = \Pi_m \tag{2.5.8}$$

for all integers $n \geq m \geq 1$. In particular, they are projections.

(b) Define a map $X \to [0,\infty) : x \mapsto |||x|||$ by the formula

$$|||x||| := \sup_{n\in\mathbb{N}} \|\Pi_n(x)\| \qquad \text{for } x \in X. \tag{2.5.9}$$

Prove that this is a norm and that $\|x\| \le |||x|||$ for all $x \in X$.

(c) Prove that $(X, |||\cdot|||)$ is a Banach space. **Hint:** Let $(x_k)_{k\in\mathbb{N}}$ be a Cauchy sequence in $(X, |||\cdot|||)$. Then $(x_k)_{k\in\mathbb{N}}$ is a Cauchy sequence in $(X, \|\cdot\|)$ by (b). Hence there is an $x \in X$ such that $\lim_{k\to\infty} \|x - x_k\| = 0$. Also, $(\Pi_n(x_k))_{k\in\mathbb{N}}$ is a Cauchy sequence in $(X, \|\cdot\|)$ for all n. Thus there is a sequence $(\xi_n)_{n\in\mathbb{N}}$ in X such that $\lim_{k\to\infty} \|\xi_n - \Pi_n(x_k)\| = 0$ for all $n \in \mathbb{N}$. Prove that

$$\Pi_m(\xi_n) = \xi_m \qquad \text{for all integers } n \ge m \ge 1. \tag{2.5.10}$$

(The restriction of Π_m to every finite-dimensional subspace is continuous.) Let $\varepsilon > 0$. Choose $k_0 \in \mathbb{N}$ such that $|||x_k - x_\ell||| < \varepsilon/3$ for all $k, \ell \ge k_0$. Then choose $n_0 \in \mathbb{N}$ such that $\|x_{k_0} - \Pi_n(x_{k_0})\| < \varepsilon/3$ for all $n \ge n_0$. Then

$$\begin{aligned} \|x - \xi_n\| &= \lim_{k\to\infty} \|x_k - \Pi_n(x_k)\| \\ &\le \lim_{k\to\infty} \Big(2|||x_k - x_{k_0}||| + \|x_{k_0} - \Pi_n(x_{k_0})\| \Big) \\ &< \varepsilon \end{aligned}$$

for $n \ge n_0$. Deduce that $\xi_n = \Pi_n(x)$ for all n and $\lim_{k\to\infty} |||x - x_k||| = 0$.

(d) Prove that there exists a constant $c > 0$ such that

$$\sup_{n\in\mathbb{N}} \|\Pi_n(x)\| \le c\,\|x\| \qquad \text{for all } x \in X. \tag{2.5.11}$$

Hint: Use parts (b) and (c) and the Open Mapping Theorem 2.2.1.

Exercise 2.5.13 (The Canonical Inclusion). Let X be a normed vector space and let $\iota_X : X \to X^{**}$ be the canonical inclusion defined by (2.4.1).

(a) Show that $(\iota_X)^*\iota_{X^*} = \mathrm{id}_{X^*}$ and determine the kernel of the projection

$$P := \iota_{X^*}(\iota_X)^* : X^{***} \to X^{***}.$$

(b) Assume X is complete. Show that X is reflexive if and only if

$$\iota_{X^*}(\iota_X)^* = \mathrm{id}_{X^{***}}.$$

(c) Linton's Pullback. *Let $Y \subset X$ be a closed subspace and let $j : Y \to X$ be the obvious inclusion. Then $\iota_X \circ j = j^{**} \circ \iota_Y : Y \to X^{**}$. This map is an isometric embedding of Y into X^{**} whose image is*

$$\iota_X \circ j(Y) = j^{**} \circ \iota_Y(Y) = \iota_X(X) \cap j^{**}(Y^{**}) \subset X^{**}.$$

(d) Deduce from Linton's Pullback that Y is reflexive whenever X is reflexive.

(e) Show that X is reflexive if and only if $\iota_{X^{**}} = (\iota_X)^{**}$.

Note. This exercise requires the notion of the dual operator, introduced in Definition 4.1.1 below.

3.7. Problems

Exercise 3.7.1 (Weak and Strong Convergence). Let H be a real Hilbert space and let $(x_i)_{i\in\mathbb{N}}$ be a sequence in H that converges weakly to $x \in H$. Assume also that

$$\|x\| = \lim_{i\to\infty} \|x_i\|.$$

Prove that $(x_i)_{i\in\mathbb{N}}$ converges strongly to x, i.e.

$$\lim_{i\to\infty} \|x_i - x\| = 0.$$

Exercise 3.7.2 (Weak Convergence and Weak Closure). Let H be an infinite-dimensional separable real Hilbert space and let $(e_n)_{n\in\mathbb{N}}$ be an orthonormal basis of H. Prove the following.

(a) The sequence $(e_n)_{n\in\mathbb{N}}$ converges weakly to zero.

(b) The set

$$A := \left\{\sqrt{n}e_n \,|\, n \in \mathbb{N}\right\}$$

is sequentially weakly closed, but the weak closure of A contains zero. **Hint:** Let $U \subset H$ be a weakly open neighborhood of the origin. Show that there are vectors $y_1, \dots, y_m \in H$ and a number $\varepsilon > 0$ such that

$$V := \{x \in H \mid \max_{i=1,\dots,m} |\langle x, y_i\rangle| < \varepsilon\} \subset U.$$

Show that the sequence

$$z_n := \max_{i=1,\dots,m} |\langle e_n, y_i\rangle|$$

is square summable and deduce that $V \cap A \neq \emptyset$.

Exercise 3.7.3 (The Weak Topology of ℓ^1). Prove the following.

(a) The standard basis e_n of ℓ^1 does not converge weakly to zero.

(b) View ℓ^1 as the dual space of c_0 (see Example 1.3.7). Then the standard basis converges to zero in the weak* topology.

(c) Schur's Theorem. *A sequence in ℓ^1 converges (to zero) in the weak topology if and only if it converges (to zero) in the norm topology.*

Exercise 3.7.4 (Weak* Topology and Distance Function). Let X be a separable normed vector space and let $(x_n)_{n\in\mathbb{N}}$ be a dense sequence in the unit ball of X. Prove that the map

$$d(x^*, y^*) := \sum_{n=1}^{\infty} 2^{-n} |\langle x^* - y^*, x_n\rangle| \qquad \text{for } x^*, y^* \in B^* \tag{3.7.1}$$

defines a distance function on the closed unit ball $B^* \subset X^*$. Prove that the topology induced by this distance function is the weak* topology on B^*.

Exercise 3.7.5 (Compact-Open Topology). Let X be a topological space, let Y be a metric space, and let $C(X,Y)$ be the space of continuous functions $f: X \to Y$. The **compact-open topology** on $C(X,Y)$ is the smallest topology such that the set

$$\mathscr{S}(K,V) := \{f \in C(X,Y) \,|\, f(K) \subset V\}$$

is open for every compact set $K \subset X$ and every open set $V \subset Y$. Thus a set $\mathcal{U} \subset C(X,Y)$ is open with respect to the compact-open topology if and only if, for each $f \in \mathcal{U}$, there are finitely many compact sets $K_1, \dots, K_m \subset X$ and open sets $V_1, \dots, V_m \subset Y$ such that $f \in \bigcap_{i=1}^m \mathscr{S}(K_i, V_i) \subset \mathcal{U}$.

(a) If X is compact, prove that the compact-open topology on $C(X,Y)$ agrees with the topology induced by the metric

$$d(f,g) := \sup_{x \in X} d_Y(f(x), g(x)) \qquad \text{for } f, g \in C(X,Y). \tag{3.7.2}$$

Hint 1: Let $f \in C(X,Y)$ and suppose that $K_1, \dots, K_m \subset X$ are compact sets and $V_1, \dots, V_m \subset Y$ are open sets such that $f(K_i) \subset V_i$ for $i = 1, \dots, m$. Prove that there is a constant $\varepsilon > 0$ such that $B_\varepsilon(f_i(x)) \subset V_i$ for all $x \in K_i$ and all $i \in \{1, \dots, m\}$. Deduce that every $g \in C(X,Y)$ with $d(f,g) < \varepsilon$ satisfies $g(K_i) \subset V_i$ for $i = 1, \dots, m$.

Hint 2: Let $f \in C(X,Y)$ and $\varepsilon > 0$. Find elements $x_1, \dots, x_m \in X$ such that $X = \bigcup_{i=1}^m K_i$, where $K_i := \{x \in X \,|\, d_Y(f(x_i), f(x)) \le \varepsilon/4\}$. Define

$$\mathcal{U} := \{g \in C(X,Y) \,|\, g(K_i) \subset V_i \text{ for } i = 1, \dots, m\}, \qquad V_i := B_{\varepsilon/2}(f(x_i)).$$

Show that $f \in \mathcal{U}$ and $d(f,g) < \varepsilon$ for all $g \in \mathcal{U}$.

(b) For each compact subset $K \subset X$ define the seminorm $p_K : C(X,\mathbb{R}) \to \mathbb{R}$ by

$$p_K(f) := \sup_K |f| \qquad \text{for } f \in C(X,\mathbb{R}).$$

Prove that these seminorms generate the compact-open topology, i.e. the compact-open topology on $C(X,\mathbb{R})$ is the smallest topology such that p_K is continuous for every compact set $K \subset X$.

(c) Prove that $C(X,\mathbb{R})$ is a locally convex topological vector space with the compact-open topology.

(d) Prove that a subset $\mathscr{F} \subset C(X,Y)$ is precompact with respect to the compact-open topology if and only if, for every compact set $K \subset X$, the set

$$\mathscr{F}_K := \{f|_K \,|\, f \in \mathscr{F}\} \subset C(K,Y) \tag{3.7.3}$$

is precompact. **Hint:** Let $\mathscr{K} \subset 2^X$ be the collection of compact subsets. Prove that the map $C(X,Y) \to \prod_{K \in \mathscr{K}} C(K,Y) : f \mapsto (f|_K)_{K \in \mathscr{K}}$ is a homeomorphism onto its image and use Tychonoff's Theorem A.2.1.

(e) Prove the following variant of the Arzelà–Ascoli Theorem.

Arzelà–Ascoli. *Let X be a topological space and let Y be a metric space. A subset $\mathscr{F} \subset C(X,Y)$ is precompact with respect to the compact-open topology if and only if it is pointwise precompact and the set $\mathscr{F}_K \subset C(K,Y)$ in* (3.7.3) *is equi-continuous for every compact set $K \subset X$.*

Hint: Use part (d) and Exercise 1.1.15.

Exercise 3.7.6 (Banach–Alaoglu). Let X be a normed vector space. Deduce the Banach–Alaoglu Theorem 3.2.4 from the Arzelà–Ascoli Theorem in part (e) of Exercise 3.7.5. **Hint:** The closed unit ball in X^* is equi-continuous as a subset of $C(X,\mathbb{R})$. Prove that the compact-open topology on X^* is finer than the weak* topology, i.e. every weak* open subset of X^* is also open with respect to the compact-open topology.

Exercise 3.7.7 (Functions Vanishing at Infinity). Let M be a locally compact Hausdorff space. A continuous real valued function $f : M \to \mathbb{R}$ is said to **vanish at infinity** if, for every $\varepsilon > 0$, there exists a compact set $K \subset M$ such that

$$\sup_{x \in M \setminus K} |f(x)| < \varepsilon.$$

Denote by $C_0(M)$ the space of all continuous functions $f : M \to \mathbb{R}$ that vanish at infinity (see Exercise 3.2.10).

(a) Prove that $C_0(M)$ is a Banach space with the supremum norm.

(b) The dual space $C_0(M)^*$ can be identified with the space $\mathcal{M}(M)$ of signed Radon measures on M with the norm (1.1.4), by the Riesz Representation Theorem (see [**75**, Thm. 3.15 & Ex. 3.35]). Here a **signed Radon measure** on M is a signed Borel measure μ with the property that, for each Borel set $B \subset M$ and each $\varepsilon > 0$, there exists a compact set $K \subset B$ such that $|\mu(A) - \mu(A \cap K)| < \varepsilon$ for every Borel set $A \subset B$.

(c) Prove that the map $\delta : M \to C_0(M)^*$, which assigns to each $x \in M$ the bounded linear functional $\delta_x : C_0(M) \to \mathbb{R}$ given by

$$\delta_x(f) := f(x) \qquad \text{for } f \in C_0(M),$$

is a homeomorphism onto its image $\delta(M) \subset C_0(M)^*$, equipped with the weak* topology. Under the identification in (b) this image is contained in the set

$$P(M) := \{\mu \in \mathcal{M}(M) \,|\, \mu \geq 0,\ \|\mu\| = \mu(M) = 1\}$$

of Radon probability measures. Determine the weak* closure of the set

$$\delta(M) = \{\delta_x \,|\, x \in M\} \subset P(M).$$

Exercise 3.7.8 (Alaoglu–Bourbaki Theorem). Let X and Y be real vector spaces and let

$$Y \times X \to \mathbb{R} : (y, x) \mapsto \langle y, x \rangle \tag{3.7.4}$$

be a nondegenerate pairing. For two subsets $A \subset X$ and $B \subset Y$ define the **polar sets** $A^\circ \subset Y$ and $B_\circ \subset X$ by

$$\begin{aligned} A^\circ &:= \{y \in Y \mid \langle y, a \rangle \le 1 \text{ for all } a \in A\}, \\ B_\circ &:= \{x \in X \mid \langle b, x \rangle \le 1 \text{ for all } b \in B\}. \end{aligned} \tag{3.7.5}$$

Thus A° and $B_\circ$ are intersections of half-spaces.

(a) Suppose X is a real normed vector space, $Y = X^*$ is its dual space, and (3.7.4) is the standard pairing. Let $S \subset X$ and $S^* \subset X^*$ denote the unit spheres and $B \subset X$ and $B^* \subset X^*$ the closed unit balls. Verify that

$$S^0 = B^*, \qquad (S^*)_0 = B.$$

(b) Bipolar Theorem. *Equip X with the topology induced by the linear maps $X \to \mathbb{R} : x \mapsto \langle y, x \rangle$ for $y \in Y$. Then*

$$(A^0)_0 = \overline{\mathrm{conv}}(A \cup \{0\}).$$

(c) Goldstine's Theorem. *If X is a normed vector space and B is the closed unit ball, then the weak* closure of $\iota(B)$ is the closed unit ball in X^{**}.* (See also Corollary 3.1.29.)

(d) Alaoglu–Bourbaki Theorem. *Suppose $(X, \mathscr{U})$ is a locally convex topological vector space over the reals, Y is the space of $\mathscr{U}$-continuous linear functionals $\Lambda : X \to \mathbb{R}$, and (3.7.4) is the standard pairing. Equip Y with the topology $\mathscr{V} \subset 2^Y$ induced by the linear maps $Y \to \mathbb{R} : y \mapsto \langle y, x \rangle$ for $x \in X$. If $A \subset X$ is a $\mathscr{U}$-neighborhood of the origin, then $A^\circ \subset Y$ is $\mathscr{V}$-compact.*

Exercise 3.7.9 (Milman–Pettis Theorem). A normed vector space X over the reals is called **uniformly convex** if, for every $\varepsilon > 0$, there exists a constant $\delta > 0$ such that, for all $x, y \in X$,

$$\|x\| = \|y\| = 1, \quad \|x + y\| > 2 - \delta \qquad \Longrightarrow \qquad \|x - y\| < \varepsilon.$$

The Milman–Pettis Theorem asserts that every uniformly convex Banach space is reflexive. This can be proved as follows.

The proof requires the concept of a net, which generalizes the concept of a sequence. A **directed set** is a nonempty set A, equipped with a reflexive and transitive relation $\preccurlyeq$, such that, for all $\alpha, \beta \in A$, there exists a $\gamma \in A$ with $\alpha \preccurlyeq \gamma$ and $\beta \preccurlyeq \gamma$. Anti-symmetry is not required, so a directed set need not be partially ordered. An example of a directed set is the collection of open neighborhoods of a point x_0 in a topological space X, equipped with the relation $U \preccurlyeq V \iff V \subset U$.

A **net** in a space X is a map

$$A \to X : \alpha \mapsto x_\alpha,$$

defined on a directed set A. A net $(x_\alpha)_{\alpha \in A}$ in a topological space X is said to **converge to** $x \in X$ if, for every open neighborhood $U \subset X$ of x, there exists an element $\alpha_0 \in A$, such that $x_\alpha \in U$ for all $\alpha \in A$ with $\alpha_0 \preccurlyeq \alpha$.

If X and Y are topological spaces, then a map $f : X \to Y$ is continuous if and only if, for every net $(x_\alpha)_{\alpha \in A}$ in X that converges to $x \in X$, the net $(f(x_\alpha))_{\alpha \in A}$ in Y converges to $f(x)$.

Let $(x_\alpha)_{\alpha \in A}$ be a net in X. A **subnet** of $(x_\alpha)_{\alpha \in A}$ is a net of the form $(x_{h(\beta)})_{\beta \in B}$ where $h : B \to A$ is a monotone final map between directed sets. Here the map $h : B \to A$ is called **monotone** if

$$\beta_1 \preccurlyeq \beta_2 \qquad \Longrightarrow \qquad h(\beta_1) \preccurlyeq h(\beta_2)$$

for all $\beta_1, \beta_2 \in B$, and it is called **final** if, for every $\alpha \in A$, there exists an element $\beta \in B$ such that

$$h(\alpha) \preccurlyeq \beta.$$

With this understood, a topological space X is compact if and only if every net in X has a convergent subnet.

A net $(x_\alpha)_{\alpha \in A}$ in a normed vector space X is called a **Cauchy net** if the net $(\|x_\alpha - x_\beta\|)_{(\alpha,\beta) \in A \times A}$ (product order on $A \times A$) converges to zero. If X is a Banach space, then every Cauchy net in X converges.

(a) Let X be a uniformly convex normed vector space. Let $(x_\alpha)_{\alpha \in A}$ be a net in the unit sphere of X such that the net $(\|x_\alpha + x_\beta\|)_{(\alpha,\beta) \in A \times A}$ converges to 2. Prove that $(x_\alpha)_{\alpha \in A}$ is a Cauchy net.

(b) Let X be a normed vector space and let $x^{**} \in X^{**}$ with $\|x^{**}\| = 1$. Prove that there exists a net $(x_\alpha)_{\alpha \in A}$ in the unit sphere of X such that the net $(\iota(x_\alpha))_{\alpha \in A}$ in X^{**} converges to x^{**} with respect to the weak* topology.

(c) Let X be a normed vector space and let $(x_\alpha)_{\alpha \in A}$ be a net in the unit sphere of X such that the net $(\iota(x_\alpha))_{\alpha \in A}$ in X^{**} converges to x^{**} with respect to the weak* topology, where $\|x^{**}\| = 1$. Prove that the net

$$(\iota(x_\alpha + x_\beta))_{(\alpha,\beta) \in A \times A}$$

converges to $2x^{**}$ in the weak* topology. If X is uniformly convex, use (a) to prove that $(x_\alpha)_{\alpha \in A}$ is a Cauchy net.

(d) Assume X is a uniformly convex Banach space, let $x^{**} \in X^{**}$ such that $\|x^{**}\| = 1$, and choose a net $(x_\alpha)_{\alpha \in A}$ as in (b). Use (c) to prove that the net $(x_\alpha)_{\alpha \in A}$ converges to some element $x \in X$. Deduce that $\iota(x) = x^{**}$.

Exercise 3.7.10 (Banach–Mazur Theorem). Let X be a Banach space and let $B^* \subset X^*$ be the closed unit ball in the dual space, equipped with the weak* topology.

(a) Prove that the map $X \to C(B^*) : x \mapsto f_x$, defined by $f_x(x^*) := \langle x^*, x\rangle$ for $x \in X$ and $x^* \in B^*$, is a linear isometric embedding.

(b) If K is a compact metric space, then there is a continuous surjective map $\pi : F \to K$, defined on a closed subset $F \subset \{0,1\}^{\mathbb{N}}$ of the Cantor set. Deduce that there exists a linear isometric embedding $\pi^* : C(K) \to C(F)$. **Hint:** The **Cantor function** is a continuous surjection $\{0,1\}^{\mathbb{N}} \to [0,1]$. Use it to construct a continuous surjection $\{0,1\}^{\mathbb{N}} \to [0,1]^{\mathbb{N}}$ and then find an embedding $K \hookrightarrow [0,1]^{\mathbb{N}}$.

(c) For every closed subset $F \subset [0,1]$ of the unit interval find a linear isometric embedding $\iota_F : C(F) \to C([0,1])$. **Hint:** The complement of F is a countable union of intervals.

(d) Banach–Mazur Theorem. *Every separable Banach space is isometrically isomorphic to a closed subspace of* $C([0,1])$.

Exercise 3.7.11 (Helly's Theorem). (Another proof of Lemma 3.4.3.)

(a) Let X be a normed vector space, let $x_1^*, \dots, x_n^* \in X^*$, and let $c_1, \dots, c_n$ be scalars. Prove that there exists an element $x \in X$ such that

$$\langle x_i^*, x\rangle = c_i \qquad \text{for } i = 1, \dots, n \tag{3.7.6}$$

if and only if there is a constant $M > 0$ such that, for all scalars $\lambda_1, \dots, \lambda_n$,

$$\left|\sum_{i=1}^n \lambda_i c_i\right| \le M \left\|\sum_{i=1}^n \lambda_i x_i^*\right\|. \tag{3.7.7}$$

Hint: Assume $x_1^*, \dots, x_m^*$ are linearly independent and span the same space as $x_1^*, \dots, x_n^*$. Define the map $T : X \to \mathbb{R}^m$ by $Tx := (\langle x_1^*, x\rangle, \dots, \langle x_m^*, x\rangle)$ for $x \in X$. Then T is surjective by Lemma 3.1.13. Use the inequality (3.7.7) to show that every element $x \in T^{-1}(c_1, \dots, c_m)$ satisfies (3.7.6).

(b) Assume (3.7.7) and let $\varepsilon > 0$. Prove that there exists an element $x \in X$ that satisfies (3.7.6) and $\|x\| < M + \varepsilon$. **Hint:** By (a) there exists some element $y \in X$ such that $\langle x_i^*, y\rangle = c_i$ for $i = 1, \dots, n$. Define $Z := \bigcap_{i=1}^n \ker(x_i^*)$. If $y \notin Z$, then, by Theorem 2.3.22, there is an element $x^* \in X^*$ such that

$$\|x^*\| = 1, \qquad x^*|_Z = 0, \qquad \langle x^*, y\rangle = d(y, Z) = \inf_{z\in Z} \|y - z\|.$$

By Lemma 3.1.14 the element x^* is a linear combination of the x_i^*. Use this to deduce from (3.7.7) that $d(y, Z) \le M$. Find $z \in Z$ with $\|y + z\| < M + \varepsilon$. (If $\dim X = \infty$, then $Z \ne \{0\}$, so the norm of $y + z$ can be chosen equal to any number bigger than M.)

Exercise 3.7.12 (Šmulyan–James Theorem). *Let X be a normed vector space. Then the following are equivalent.*

(i) *X is reflexive.*

(ii) *Every bounded sequence in X has a weakly convergent subsequence.*

(iii) *If $C_1 \supset C_2 \supset C_3 \supset \cdots$ is a nested sequence of nonempty bounded closed convex subsets of X, then their intersection is nonempty.*

The implication $(iii) \implies (i)$ of the Šmulyan–James Theorem strengthens the Eberlein–Šmulyan Theorem 3.4.1.

(a) Prove that (i) implies (ii) and (ii) implies (iii).

(b) Prove that (iii) implies that X is complete.

(c) Let X be a nonreflexive Banach space. Prove that there exists a constant $0 < \alpha < 1$ and an element $x^{**} \in X^{**}$ such that

$$\alpha < d(x^{**}, \iota(X)) \le \|x^{**}\| < 1. \tag{3.7.8}$$

Hint: Use the Riesz Lemma 1.2.12.

(d) Let $0 < \alpha < 1$ and $x^{**} \in X^{**}$ be as in (c). Find sequences of unit vectors $(x_n^*)_{n\in\mathbb{N}}$ in X^* and $(x_k)_{k\in\mathbb{N}}$ in X such that

$$\langle x_n^*, x_k\rangle = \begin{cases} 0, & \text{if } k < n, \\ \alpha, & \text{if } k \ge n, \end{cases} \qquad \langle x^{**}, x_n^*\rangle = \alpha \quad \text{for all } k, n \in \mathbb{N}. \tag{3.7.9}$$

Hint: Argue by induction. First find unit vectors $x_1^* \in X^*$ and $x_1 \in X$ such that $\langle x^{**}, x_1^*\rangle = \alpha$ and $\langle x_1^*, x_1\rangle = \alpha$. Now let $N > 1$ and assume by induction that unit vectors $x_1, \dots, x_{N-1} \in X$ and $x_1^*, \dots, x_{N-1}^* \in X^*$ have been found that satisfy (3.7.9) for $k, n = 1, \dots, N-1$. With $M := \alpha d(x^{**}, \iota(X)) < 1$ we have $|\lambda_0 \alpha| \le M\|\lambda_0 x^{**} + \sum_{k=1}^{N-1} \lambda_k \iota(x_k)\|$ for all $\lambda_0, \dots, \lambda_{N-1} \in \mathbb{R}$. Hence, by Helly's Theorem, there exists a unit vector $x_N^* \in X^*$ such that

$$\langle x^{**}, x_N^*\rangle = \alpha, \qquad \langle x_N^*, x_k\rangle = \langle \iota(x_k), x_N^*\rangle = 0 \quad \text{for } k = 1, \dots, N-1.$$

Moreover,

$$\alpha \left|\sum_{n=1}^N \lambda_n\right| = \left|\left\langle x^{**}, \sum_{n=1}^N \lambda_n x_n^*\right\rangle\right| \le \|x^{**}\| \left\|\sum_{n=1}^N \lambda_n x_n^*\right\|$$

for all $\lambda_1, \dots, \lambda_N \in \mathbb{R}$. Since $\|x^{**}\| < 1$ it follows again from Helly's Theorem that there is a unit vector $x_N \in X$ such that $\langle x_n^*, x_N\rangle = \alpha$ for $n = 1, \dots, N$. This completes the induction step for the proof of (3.7.9).

(e) Let x_k, x_n^* be as in (d) and define $C_N := \overline{\mathrm{conv}}(\{x_k \,|\, k \ge N\})$ for $N \in \mathbb{N}$. Prove that $\langle x_N^*, x\rangle = \alpha$ and $\lim_{n\to\infty}\langle x_n^*, x\rangle = 0$ for all $x \in C_N$. Deduce that the C_N have an empty intersection.

Exercise 3.7.13 (Birkhoff–von Neumann Theorem). An $n \times n$-matrix $M = (m_{ij})$ with nonnegative coefficients $m_{ij} \geq 0$ is called **doubly stochastic** if its row sums and column sums are all equal to one. The Birkhoff–von Neumann Theorem asserts the following.

Every doubly stochastic matrix
is a convex combination of permutation matrices.

Thus the doubly stochastic matrices form a convex set whose extremal points are the permutation matrices. This can be proved as follows.

Let M be a doubly stochastic matrix and denote by $\nu(M)$ the number of positive entries. If $\nu(M) > n$ find a permutation matrix P and a constant $0 < \lambda < 1$ such that the matrix $N := M - \lambda P_1$ has nonnegative entries and strictly fewer positive entries than M. In the case $N \neq 0$ the matrix $M_1 := (1-\lambda)^{-1} N$ is again doubly stochastic with $\nu(M_1) < \nu(M)$, and $M = \lambda P_1 + (1-\lambda) M_1$. Continue by induction until $\nu(M_k) = n$ and so M_k is a permutation matrix. Here is a method to find P_1 and λ.

Hall's Marriage Theorem. *Let X and Y be finite sets and let $\Gamma \subset X \times Y$. Then the following are equivalent.*

(i) *There is an injective map $f : X \to Y$ whose graph is contained in Γ.*

(ii) *For every $A \subset X$ the set*

$$\Gamma(A) := \{y \in Y \,|\, \textit{there is an } x \in A \textit{ such that } (x, y) \in \Gamma\}$$

satisfies $\#\Gamma(A) \geq \#A$.

Take $X = Y = \{1, \ldots, n\}$ and $\Gamma := \{(i, j) \,|\, m_{ij} > 0\}$. Use the fact that M is doubly stochastic to verify that Γ satisfies (ii). Use the injective map f in (i) to determine the permutation matrix P_1 and take $\lambda := \min_{j=f(i)} m_{ij}$.

Exercise 3.7.14 (Strict Convexity and Extremal Points). A normed vector space is strictly convex (see Example 3.5.3 and Exercise 2.5.11) if and only if the unit sphere is equal to the set of extremal points of the closed unit ball.

Exercise 3.7.15 (A Noncompact Set of Extremal Points). Let $C \subset \mathbb{R}^3$ be the closed convex hull of the set

$$S := \big\{(1 + \cos(\theta), \sin(\theta), 0) \,\big|\, \theta \in \mathbb{R}\big\} \cup \big\{(0, 0, 1), (0, 0, -1)\big\}.$$

Determine the extremal points of C.

Exercise 3.7.16 (Extremal Points of Unit Balls). Determine the extremal points of the closed unit balls in the Banach spaces

$$c_0,\ c,\ C([0,1]),\ \ell^1,\ \ell^p,\ \ell^\infty,\ L^1([0,1]),\ L^p([0,1]),\ L^\infty([0,1])$$

for $1 < p < \infty$.

Exercise 3.7.17 (Hilbert Cube). (See Example 3.5.5.)

(a) Show that the Hilbert cube $Q := \{x = (x_i)_{i\in\mathbb{N}} \in \ell^2 \,|\, 0 \le x_i \le 1/i\}$ is a compact subset of ℓ^2 with respect to the norm topology.

(b) Is the set $R := \{x = (x_i)_{i\in\mathbb{N}} \in \ell^2 \,|\, 0 \le x_i \le 1/\sqrt{i}\}$ compact in ℓ^2 with respect to either the norm topology or the weak topology?

Exercise 3.7.18. Let X be a real normed vector space, let $B^* \subset X^*$ be the closed unit ball in the dual space, and let $\Lambda : X^* \to \mathbb{R}$ be a linear functional such that the restriction $\Lambda|_{B^*} : B^* \to \mathbb{R}$ is weak* continuous. Then there exists an element $x \in X$ such that $\Lambda = \iota(x)$.

Exercise 3.7.19 (Markov–Kakutani Fixed Point Theorem). Let X be a locally convex Hausdorff topological vector space and let $\mathcal{A}$ be a collection of pairwise commuting continuous linear operators $A : X \to X$. Let $C \subset X$ be a nonempty $\mathcal{A}$-invariant compact convex subset of X, so that

$$A(C) \subset C \qquad \text{for all } A \in \mathcal{A}.$$

Then there exists an element $x \in C$ such that $Ax = x$ for all $A \in \mathcal{A}$.

(a) For $A \in \mathcal{A}$ and $k \in \mathbb{N}$ define

$$A_k := \frac{1}{k}\big(\mathbb{1} + A + A^2 + \cdots + A^{k-1}\big).$$

Then $A_k(C)$ is a nonempty compact convex subset of C.

(b) If $A, B \in \mathcal{A}$ and $k, \ell \in \mathbb{N}$, then

$$A_k(B_\ell(C)) \subset A_k(C) \cap B_\ell(C).$$

Use this to prove that the set

$$F := \bigcap_{k\in\mathbb{N}} \bigcap_{A\in\mathcal{A}} A_k(C)$$

is nonempty.

(c) Prove that every element $x \in F$ is a fixed point of $\mathcal{A}$. **Hint:** Fix an element $A \in \mathcal{A}$. If $Ax \ne x$ find a continuous linear functional $\Lambda : X \to \mathbb{R}$ such that $\Lambda(x - Ax) = 1$. Prove that, for every $k \in \mathbb{N}$, there exists an element $y \in C$ such that

$$A_k y = x.$$

Now observe that

$$y - A^k y = k(x - Ax)$$

and deduce that the functional Λ is unbounded on the compact set $C - C$, contradicting continuity.

Exercise 3.7.20 (Bell–Fremlin Theorem). *The axiom of choice is equivalent to the assertion that the closed unit ball in the dual space of every nonzero Banach space has an extremal point.*

(a) Let X be any nonzero Banach space. Use the Banach–Alaoglu Theorem 3.2.4, the Hahn–Banach Theorem 2.3.2, and the Kreĭn–Milman Theorem 3.5.2 to prove that the closed unit ball in X^* has an extremal point.

(b) Let I be any index set and, for each $i \in I$, let X_i be a nonzero Banach space. Define the Banach spaces

$$\bigoplus_{i\in I} X_i := \left\{ x = (x_i)_{i\in I} \,\middle|\, x_i \in X_i \text{ and } \|x\|_1 := \sum_{i\in I} \|x_i\|_{X_i} < \infty \right\} \tag{3.7.10}$$

and

$$\prod_{i\in I} X_i := \left\{ x = (x_i)_{i\in I} \,\middle|\, x_i \in X_i \text{ and } \|x\|_\infty := \sup_{i\in I} \|x_i\|_{X_i} < \infty \right\}. \tag{3.7.11}$$

Prove that $\prod_{i\in I} X_i^*$ is isomorphic to the dual space of $\bigoplus_{i\in I} X_i$.

(c) Let S be a nonempty set. Define $c_0(S)$ to be the space of all functions $f : S \to \mathbb{R}$ that satisfy $\#\{s \in S \,|\, |f(s)| > \varepsilon\} < \infty$ for all $\varepsilon > 0$, equipped with the supremum norm $\|f\|_\infty := \sup_{s\in S} |f(s)|$. Define $\ell^1(S)$ to be the space of all functions $g : S \to \mathbb{R}$ such that $\|g\|_1 := \sum_{s\in S} |g(s)| < \infty$. Prove that $\ell^1(S)$ is isomorphic to the dual space of $c_0(S)$.

(d) Let $(S_i)_{i\in I}$ be a family of pairwise disjoint nonempty sets. Then the Banach space $\prod_{i\in I} \ell^1(S_i)$ is isomorphic to the dual space of $\bigoplus_{i\in I} c_0(S_i)$ by (b) and (c). Suppose the closed unit ball in $\prod_{i\in I} \ell^1(S_i)$ has an extremal point $g = (g_i)_{i\in I}$. Prove that $g_i \neq 0$ for all $i \in I$. Show that, for each $i \in I$, there is a unique element $s_i \in S_i$ such that $g_i(s_i) \neq 0$.

4.5. Problems

Exercise 4.5.1 (Injections and Surjections). Let X and Y be Banach spaces. Prove the following.

(a) The set of all surjective bounded linear operators $A : X \to Y$ is an open subset of $\mathcal{L}(X, Y)$ with respect to the norm topology.

(b) The set of all injective bounded linear operators $A : X \to Y$ is not necessarily an open subset of $\mathcal{L}(X, Y)$ with respect to the norm topology.

(c) The set of all injective bounded linear operators $A : X \to Y$ with closed image is an open subset of $\mathcal{L}(X, Y)$ with respect to the norm topology.

Exercise 4.5.2 (The Image of a Compact Operator). Let X and Y be Banach spaces and let $K : X \to Y$ be a compact operator. Prove the following.

(a) If K has a closed image, then $\dim \mathrm{im}(K) < \infty$.

(b) The image of K is a separable subspace of Y.

(c) If Y is separable, then there exist a Banach space X and a compact operator $K : X \to Y$ with a dense image.

Exercise 4.5.3 (Compact Subsets of Banach Spaces). *Let X be a Banach space and let $C \subset X$ be a closed subset. Then the following are equivalent.*

(i) *C is compact.*

(ii) *There exists a sequence $x_n \in C$ such that*

$$\lim_{n\to\infty} \|x_n\| = 0, \qquad C \subset \overline{\mathrm{conv}}(\{x_n \,|\, n \in \mathbb{N}\}). \tag{4.5.1}$$

Hint 1: To prove that (ii) implies (i) observe that

$$\overline{\mathrm{conv}}(\{x_n \,|\, n \in \mathbb{N}\}) = \left\{ \sum_{n=1}^{\infty} \lambda_n x_n \,\middle|\, \lambda_n \geq 0, \sum_{n=1}^{\infty} \lambda_n = 1 \right\} \tag{4.5.2}$$

whenever $\lim_{n\to\infty} \|x_n\| = 0$.

Hint 2: To prove that (i) implies (ii), choose a sequence of compact sets $C_k \subset X$ and a sequence of finite subsets $A_k \subset C_k$ such that $C_1 = C$ and

$$2C_k \subset \bigcup_{x \in A_k} \overline{B}_{4^{-k}}(x), \qquad C_{k+1} := \bigcup_{x \in A_k} \Big(\big(2C \cap \overline{B}_{4^{-k}}(x)\big) - x \Big)$$

for $k \in \mathbb{N}$. Prove that, for every $c \in C$, there is a sequence $x_k \in A_k$ such that $x = \sum_{k=1}^{\infty} 2^{-k} x_k$. Note that $\|x\| \leq 4^{-k}$ for all $x \in A_{k+1}$ and all $k \in \mathbb{N}$.

Exercise 4.5.4 (Continuity). Let X and Y be normed vector spaces.

(a) A linear operator $A : X \to Y$ is bounded if and only if it is continuous with respect to the weak topologies on X and Y.

(b) A linear operator $B : Y^* \to X^*$ is continuous with respect to the weak* topologies on Y^* and X^* if and only if there exists a bounded linear operator $A : X \to Y$ such that $B = A^*$.

(c) A linear operator $A : X \to Y$ is continuous with respect to the weak topology on X and the norm topology on Y if and only if it is bounded and has finite rank.

(d) Suppose X and Y are Banach spaces and denote by $B^* \subset Y^*$ the closed unit ball. Then a bounded linear operator $A : X \to Y$ is compact if and only if $A^*|_{B^*} : B^* \to X^*$ is continuous with respect to the weak* topology on B^* and the norm topology on X^*.

(e) Suppose X and Y are reflexive Banach spaces and denote by $B \subset X$ the closed unit ball. Then a bounded linear operator $A : X \to Y$ is compact if and only if $A|_B : B \to X$ is continuous with respect to the weak topology on B and the norm topology on Y.

Exercise 4.5.5 (Gantmacher's Theorem). Let X and Y be Banach spaces and let $A : X \to Y$ be a bounded linear operator. Then the following are equivalent.

(i) A is **weakly compact**, i.e. if $B \subset X$ is a bounded set, then the weak closure of $A(B)$ is a weakly compact subset of Y.

(ii) If $(x_n)_{n\in\mathbb{N}}$ is a bounded sequence in X, then the sequence $(Ax_n)_{n\in\mathbb{N}}$ in Y has a weakly convergent subsequence.

(iii) $A^{**}(X^{**}) \subset \iota_Y(Y)$.

(iv) $A^* : Y^* \to X^*$ *is continuous with respect to the weak* topology on* Y^* *and the weak topology on* X^*.

(v) *The dual operator* $A^* : Y^* \to X^*$ *is weakly compact.*

Hint: To prove that (i) implies (iii) denote by

$$B \subset X, \qquad B^{**} \subset X^{**}$$

the closed unit balls and denote by $C \subset Y$ the weak closure of $A(B)$. If (i) holds, then $\iota_Y(C)$ is a weak* compact subset of Y^{**}. Use Goldstine's Theorem (Corollary 3.1.29) to prove that

$$A^{**}(B^{**}) \subset \iota_Y(C).$$

(See Exercise 3.7.8.)

Exercise 4.5.6 (Pitt's Theorem). Let $1 \le p < q < \infty$. Then every bounded linear operator $A : \ell^q \to \ell^p$ is compact.

(a) Fix a bounded linear operator $A : \ell^q \to \ell^p$ such that $\|A\| = 1$ and a sequence $(x_n)_{n\in\mathbb{N}}$ in ℓ^q that converges weakly to zero. It suffices to prove

$$\lim_{n\to\infty} \|Ax_n\|_p = 0.$$

Hint: Use Theorem 3.4.1 and part (e) of Exercise 4.5.4.

(b) If $(y_n)_{n\in\mathbb{N}}$ is a sequence in ℓ^p that converges weakly to zero, then

$$\limsup_{n\to\infty} \|y + y_n\|_p^p = \|y\|_p^p + \limsup_{n\to\infty} \|y_n\|_p^p \tag{4.5.3}$$

for every $y \in \ell^p$. **Hint:** Assume first that y has finite support.

(c) Let x_n be as in (a), fix a constant $\varepsilon > 0$, and choose $x \in \ell^q$ such that

$$\|x\|_q = 1, \qquad 1 - \varepsilon < \|Ax\|_p < 1. \tag{4.5.4}$$

Then

$$\left(\|Ax\|_p^p + \lambda^p \limsup_{n\to\infty} \|Ax_n\|_p^p\right)^{1/p} \le \left(\|x\|_q^q + \lambda^q \limsup_{n\to\infty} \|x_n\|_q^q\right)^{1/q} \tag{4.5.5}$$

for all $\lambda > 0$. **Hint:** Use the equation (4.5.3) in part (b) with

$$y_n := \lambda A x_n$$

and the inequality $\|Ax + \lambda Ax_n\|_p \le \|x + \lambda x_n\|_q$.

(d) There exists a constant $C > 0$ such that

$$\limsup_{n\to\infty} \|Ax_n\|_p^p \le \frac{(1 + \lambda^q C^q)^{p/q} - (1-\varepsilon)^p}{\lambda^p} \tag{4.5.6}$$

for all $\lambda > 0$ and all $\varepsilon > 0$. **Hint:** Take $C \ge \sup_{n\in\mathbb{N}} \|x_n\|_q$ and use the inequalities (4.5.4) and (4.5.5) in part (c).

(e) Choose $\lambda := C^{-1}\varepsilon^{1/q}$ in (4.5.6) to obtain

$$\limsup_{n\to\infty} \|Ax_n\|_p^p \le C^p \varepsilon^{1-p/q} \left(\frac{(1+\varepsilon)^{p/q} - 1}{\varepsilon} + \frac{1 - (1-\varepsilon)^p}{\varepsilon} \right) \tag{4.5.7}$$

for all $\varepsilon > 0$. Take the limit $\varepsilon \to 0$ in (4.5.7) to obtain $\lim_{n\to\infty} \|Ax_n\|_p = 0$.

Exercise 4.5.7 (Existence of Fredholm Operators). Let X and Y be Banach spaces and suppose that there exists a Fredholm operator from X to Y. Prove the following.

(a) X is reflexive if and only if Y is reflexive.

(b) X is separable if and only if Y is separable.

Exercise 4.5.8 (Codimension One Subspaces). Let X be a real Banach space. Prove that any two closed codimension one subspaces of X are isomorphic to one another. **Hint:** If Y and Z are distinct closed codimension one subspaces of X, then each of them is isomorphic to $(Y \cap Z) \times \mathbb{R}$.

Exercise 4.5.9 (Existence of Index One Fredholm Operators). Let X be an infinite-dimensional real Banach space. Prove that the following are equivalent.

(i) X is isomorphic to $X \times \mathbb{R}$.

(ii) There exists a codimension one subspace of X that is isomorphic to X.

(iii) Every closed codimension one subspace of X is isomorphic to X.

(iv) There exists a Fredholm operator $A : X \to X$ of index one.

(v) The homomorphism (4.4.4) is surjective.

Exercise 4.5.10 (Existence of Index Zero Fredholm Operators).

(a) Let X and Y be Banach spaces and suppose that there exists an index zero Fredholm operator from X to Y. Prove that X and Y are isomorphic.

(b) Let X be a Banach space and let $Y \subset X$ be a closed codimension one subspace. Prove that there is an index one Fredholm operator $A : X \to Y$. If X is not isomorphic to any proper closed subspace of X, prove that every Fredholm operator from X to Y has index one.

Exercise 4.5.11 (Fredholm Operators Between ℓ^p Spaces).

(a) Let $1 \le p \le \infty$. For every integer $n \in \mathbb{Z}$ construct a Fredholm operator $A : \ell^p \to \ell^p$ of index n.

(b) Construct a family of examples in (a) that are neither injective nor surjective.

(c) Let $1 \le p, q \le \infty$ and $p \ne q$. Does there exist a Fredholm operator from ℓ^p to ℓ^q?

Exercise 4.5.12 (Fredholm Operators and Vector Bundles). Let H be a separable infinite-dimensional Hilbert space and, for $k \in \mathbb{Z}$, denote by $\mathcal{F}_k(H)$ the space of Fredholm operators $A : H \to H$ of index k. Find a continuous map

$$A : S^1 \to \mathcal{F}_1(H)$$

such that the Fredholm operator $A(z) : H \to H$ is surjective for all $z \in S^1$, and the vector bundle

$$E := \left\{ (z, \xi) \in S^1 \times H \,|\, A(z)\xi = 0 \right\}$$

over S^1 is a Möbius band.

Exercise 4.5.13 (Fredholm Alternative). Fix an interval $I := [a, b]$ with $a < b$, let $f_i, g_i \in \mathcal{L}^2(I)$ for $i = 1, \dots, n$, and define

$$K(x, y) := \sum_{i=1}^{n} f_i(x) g_i(y) \qquad \text{for } a \leq x, y \leq b.$$

For $h \in L^2(I)$ consider the equation

$$u(x) + \int_a^b K(x, y) u(y)\, dy = h(x) \qquad \text{for } a \leq x \leq b. \tag{4.5.8}$$

Prove that equation (4.5.8) either has a unique solution $u \in L^2(I)$ for every h, or the homogeneous equation with $h = 0$ has a nonzero solution u.

Exercise 4.5.14 (Hilbert Spheres).

(a) The unit sphere

$$S := \left\{ x \in \ell^2 \mid \|x\|_2 = 1 \right\}$$

is contractible, i.e. there exists a continuous map $f : [0, 1] \times S \to S$ and an element $e \in S$ such that

$$f(0, x) = e, \qquad f(1, x) = x$$

for all $x \in S$.

Hint: Let $e_1, e_2, e_3, \dots$ be the standard orthonormal basis of ℓ^2 and define the shift operator $T : \ell^2 \to \ell^2$ by

$$T(x_1, x_2, x_3, \dots) := (0, x_1, x_2, x_3, \dots) \qquad \text{for } x = (x_i)_{i \in \mathbb{N}} \in \ell^2.$$

Then $Te_n = e_{n+1}$ for all $n \in \mathbb{N}$. Consider the maps $g : [0, 1] \times \ell^2 \to \ell^2$ and $h : [0, 1] \times \ell^2 \to \ell^2$ defined by

$$g(t, x) := (1 - t)e_1 + tTx, \qquad h(t, x) := (1 - t)Tx + tx$$

for $0 \leq t \leq 1$ and $x \in \ell^2$. Use these maps to show that $\ell^2 \setminus \{0\}$ is contractible and then normalize to deduce that S is contractible.

(b) Refine the construction in (a) to obtain a map $f : [0, 1] \times S \to S$ that satisfies

$$f(0, x) = e, \qquad f(1, x) = x, \qquad f(t, e) = e$$

for all $x \in S$ and all $t \in [0, 1]$. This means that the singleton $\{e\}$ is a **deformation retract of** S.

(c) Prove that the unit sphere in any infinite-dimensional Hilbert space is contractible.

Exercise 4.5.15 (Fredholm Intersection Theory). Let X be a Banach space and let $X_1, X_2 \subset X$ be closed subspaces. The triple (X, X_1, X_2) is called a **Fredholm triple** if the subspace $X_1 + X_2$ is closed, and the spaces $X_1 \cap X_2$ and $X/(X_1 + X_2)$ are finite-dimensional. The **Fredholm index** of a Fredhom triple (X, X_1, X_2) is defined by

$$\mathrm{index}(X, X_1, X_2) := \dim(X_1 \cap X_2) - \dim(X/(X_1 + X_2)). \tag{4.5.9}$$

(a) Prove that (X, X_1, X_2) is a Fredholm triple if and only if the operator

$$X_1 \times X_2 \to X : (x_1, x_2) \mapsto x_1 + x_2$$

is Fredholm. Show that the Fredholm indices agree. **Hint:** Corollary 2.2.9.

(b) Assume $X_1 + X_2$ has finite codimension in X. Prove that $X_1 + X_2$ is a closed subspace of X. **Hint:** Lemma 4.3.2.

(c) Assume (X, X_1, X_2) is a Fredholm triple. Prove that the subspaces X_1 and X_2 are complemented.

(d) Define the notion of a *small deformation* of a complemented subspace.

(e) Prove that the Fredholm property and the Fredholm index of a Fredholm triple (X, X_1, X_2) are stable under small deformations of the subspaces X_1 and X_2. **Hint:** Theorem 4.4.2.

Exercise 4.5.16 (Rellich's Theorem). Let $I := [0, 1] \subset \mathbb{R}$ be the unit interval and fix a real number $p \geq 1$. Denote by

$$W^{1,p}(I) := \left\{ f : I \to \mathbb{R} \,\middle|\, \begin{array}{l} f \text{ is absolutely continuous} \\ \text{and } \int_0^1 |f'(t)|^p \, dt < \infty \end{array} \right\} \tag{4.5.10}$$

the Sobolev space of $W^{1,p}$-functions on I with the norm

$$\|f\|_{W^{1,p}} := \left(\int_0^1 \left(|f(t)|^p + |f'(t)|^p \right) dt \right)^{1/p} \tag{4.5.11}$$

for $f \in W^{1,p}(I, \mathbb{R})$. In particular, $W^{1,1}(I)$ is the Banach space of absolutely continuous functions.

(a) Prove that $W^{1,p}(I)$ is a Banach space with the norm (4.5.11). **Hint:** Use [**75**, Thm. 6.19] or Theorem 7.5.18 with $X = \mathbb{R}$.

(b) Prove that the inclusion of $W^{1,p}(I)$ into the Banach space $C(I)$ of continuous functions $f : I \to \mathbb{R}$, equipped with the supremum norm, is a bounded linear operator.

(c) Prove that the inclusion $W^{1,p}(I) \hookrightarrow C(I)$ is a compact operator for $p > 1$ but not for $p = 1$. **Hint:** Show that the unit ball in $W^{1,p}(I)$ is equicontinuous for $p > 1$ and use the Arzelà–Ascoli Theorem (Corollary 1.1.13). For $p = 1$ consider the functions $f_n(t) := t^n$.

Exercise 4.5.17 (Fredholm Theory and Homological Algebra).

(a) Exact Sequences. A finite sequence

$$0 \longrightarrow V_0 \xrightarrow{d_0} V_1 \xrightarrow{d_1} V_2 \xrightarrow{d_2} \cdots \longrightarrow V_{n-1} \xrightarrow{d_{n-1}} V_n \longrightarrow 0$$

of vector spaces and linear maps is called **exact** if d_0 is injective, d_{n-1} is surjective, and $\ker(d_k) = \mathrm{im}(d_{k-1})$ for $k = 1, \dots, n-1$. If the sequence is exact and the vector spaces V_k are all finite-dimensional, then its **Euler characteristic** vanishes, i.e. $\sum_{k=0}^n (-1)^k \dim V_k = 0$.

(b) Two linear operators $A : X \to Y$ and $B : Y \to Z$ between vector spaces determine a natural **long exact sequence**

$$0 \longrightarrow \ker(A) \longrightarrow \ker(BA) \longrightarrow \ker(B) \xrightarrow{\delta}$$
$$\xrightarrow{\delta} \mathrm{coker}(A) \longrightarrow \mathrm{coker}(BA) \longrightarrow \mathrm{coker}(B) \longrightarrow 0,$$

where the map $\delta : \ker(B) \to \mathrm{coker}(A)$ assigns to an element $y \in \ker(B)$ the equivalence class of y in the quotient space $Y/\mathrm{im}(A) = \mathrm{coker}(A)$.

(c) If the vector spaces X, Y, Z in (b) are Banach spaces and two out of the three operators A, B, BA are Fredholm operators, then so is the third and $\mathrm{index}(BA) = \mathrm{index}(A) + \mathrm{index}(B)$. (See also Theorem 4.4.1.)

(d) The Snake Lemma. Consider a commutative diagram

$$\begin{array}{ccccccccc} 0 & \longrightarrow & U & \longrightarrow & V & \longrightarrow & W & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle A} & & \downarrow{\scriptstyle B} & & \downarrow{\scriptstyle C} & & \\ 0 & \longrightarrow & X & \longrightarrow & Y & \longrightarrow & Z & \longrightarrow & 0 \end{array}$$

of vector spaces and linear operators such that the horizontal rows are short exact sequences. Then there is a natural long exact sequence

$$0 \longrightarrow \ker(A) \longrightarrow \ker(B) \longrightarrow \ker(C) \xrightarrow{\delta}$$
$$\xrightarrow{\delta} \mathrm{coker}(A) \longrightarrow \mathrm{coker}(B) \longrightarrow \mathrm{coker}(C) \longrightarrow 0,$$

where the boundary map

$$\delta : \ker(C) \to \mathrm{coker}(A)$$

is defined as follows. Let $w \in \ker(C)$ and choose an element $v \in V$ that maps to w under the surjection $V \to W$. Then $Bv \in Y$ belongs to the kernel of the map $Y \to Z$; so there is a unique element $x \in X$ that maps to Bv under the injection $X \to Y$ and $\delta w := [x] \in X/\mathrm{im}(A) = \mathrm{coker}(A)$ is independent of the choice of v.

(e) Deduce from the Snake Lemma that, if U, V, W, X, Y, Z are Banach spaces and two out of the three operators A, B, C are Fredholm operators, then so is the third and $\mathrm{index}(B) = \mathrm{index}(A) + \mathrm{index}(C)$.