2

Functional Analysis

In this chapter we survey the main theorems of functional analysis that deal with Banach spaces, including the Hahn–Banach, Baire Category, Uniform Boundedness, Open Mapping, and Closed Graph Theorems. References for additional information on this material (and sources for many of the proofs that we give) include the texts by Conway [Con90], Folland [Fol99], and Rudin [Rud91].

2.1 The Hahn–Banach Theorem and Its Implications

Orthogonality played an essential role in many of the proofs for Hilbert spaces that appeared in Sections 1.5 and 1.6. The analysis of general Banach spaces is much more difficult because there need not be any notion of orthogonality in a Banach space. The Hahn–Banach Theorem is a fundamental result for Banach spaces that allows us to do some things in Banach spaces that at first glance seem to be impossible without having the tools that orthogonality provides.

The abstract form of the Hahn–Banach Theorem is a statement about extension of linear functionals. We state a form that applies to both real and complex vector spaces.

Theorem 2.1 (Hahn–Banach Theorem). *Let X be a vector space over $\mathbf{F}$ and let ρ be a seminorm on X. If M is a subspace of X and $\lambda\colon M \to \mathbf{F}$ is a linear functional on M satisfying*

$$|\langle x, \lambda\rangle| \;\le\; \rho(x), \qquad x \in M,$$

then there exists a linear functional $\Lambda\colon X \to \mathbf{F}$ such that

$$\Lambda|_M \;=\; \lambda \qquad \text{and} \qquad |\langle x, \Lambda\rangle| \;\le\; \rho(x), \quad x \in X. \qquad \diamondsuit$$

C. Heil, *A Basis Theory Primer: Expanded Edition*, Applied and Numerical Harmonic Analysis,
DOI 10.1007/978-0-8176-4687-5_2,

The proof of the Hahn–Banach Theorem takes some preparation, and therefore we will omit it (see [Con90] for a proof). The most important point to note is that the extension Λ obeys the same bound that is satisfied by λ, but does so on the *entire* space X and not just on the subspace M.

In practice, it is usually not the Hahn–Banach Theorem itself but rather one of its many corollaries that is applied. Therefore we will concentrate in this section on these implications. Since these corollaries are so important, when invoking any one of them it is customary to write "by the Hahn–Banach Theorem" instead of "by a corollary to the Hahn–Banach Theorem."

Our first corollary states that any bounded linear functional on a subspace M of a normed space X has an extension to the entire space whose operator norm on X equals the operator norm on M. This is easy to prove when the space is a Hilbert space (see Exercise 2.1), but it is far from obvious that such an extension should be possible on non-inner product spaces.

Corollary 2.2 (Hahn–Banach). *Let X be a normed linear space and M a subspace of X. If $\lambda \in M^*$, then there exists $\Lambda \in X^*$ such that*

$$\Lambda|_M = \lambda \qquad \textit{and} \qquad \|\Lambda\|_{X^*} = \|\lambda\|_{M^*}.$$

Proof. Set $\rho(x) = \|\lambda\|_{M^*}\,\|x\|_X$ for $x \in X$. Note that ρ is defined on all of X, and is a seminorm on X (in fact, it is a norm if $\lambda \neq 0$). Further,

$$\forall\, x \in M, \quad |\langle x, \lambda\rangle| \;\le\; \|x\|_X\,\|\lambda\|_{M^*} \;=\; \rho(x).$$

Hence Theorem 2.1 implies that there exists a linear functional $\Lambda\colon X \to \mathbf{F}$ such that $\Lambda|_M = \lambda$ (which implies $\|\Lambda\|_{X^*} \ge \|\lambda\|_{M^*}$) and

$$\forall\, x \in X, \quad |\langle x, \Lambda\rangle| \;\le\; \rho(x) \;=\; \|\lambda\|_{M^*}\,\|x\|_X,$$

which implies that $\|\Lambda\|_{X^*} \le \|\lambda\|_{M^*}$. □

Given a normed space X and given $x^* \in X^*$, the operator norm of x^* is

$$\|x^*\|_{X^*} \;=\; \sup_{x \in X,\, \|x\|_X = 1} |\langle x, x^*\rangle|.$$

Thus, we obtain the operator norm of x^* on X^* by "looking back" at its action on X. The next corollary provides a complementary viewpoint: The norm of $x \in X$ can be obtained by "looking forward" to its action on X^*. Again, this is easy to prove directly for Hilbert spaces (see Theorem 1.37), but is a much more subtle fact for generic Banach spaces.

Corollary 2.3 (Hahn–Banach). *Let X be a Banach space. Then for each $x \in X$ we have*

$$\|x\|_X \;=\; \sup_{x^* \in X^*,\, \|x^*\|_{X^*} = 1} |\langle x, x^*\rangle|. \tag{2.1}$$

Further, the supremum is achieved.

Proof. Fix $x \in X$, and let α denote the supremum on the right-hand side of equation (2.1). Since $|\langle x, x^* \rangle| \le \|x\|_X \, \|x^*\|_{X^*}$, we have $\alpha \le \|x\|_X$.

Let $M = \operatorname{span}\{x\}$, and define $\lambda \colon M \to \mathbf{F}$ by $\langle cx, \lambda \rangle = c\,\|x\|_X$. Then $\lambda \in M^*$ and $\|\lambda\|_{M^*} = 1$. Corollary 2.2 therefore implies that there exists some $\Lambda \in X^*$ with $\Lambda|_M = \lambda$ and $\|\Lambda\|_{X^*} = \|\lambda\|_{M^*} = 1$. In particular, since $x \in M$, we have $\alpha \ge |\langle x, \Lambda \rangle| = |\langle x, \lambda \rangle| = \|x\|_X$, and therefore the supremum in equation (2.1) is achieved. □

Now we can give one of the most powerful and often-used implications of the Hahn–Banach Theorem. It states that we can find a bounded linear functional that separates a point from a closed subspace of a normed space. This is easy to prove constructively for the case of a Hilbert space (see Exercise 2.2), but it is quite amazing that we can do this in arbitrary normed spaces.

Corollary 2.4 (Hahn–Banach). *Let X be a normed linear space. Suppose that:*

(a) *M is a closed subspace of X,*

(b) *$x_0 \in X \backslash M$, and*

(c) *$d = \operatorname{dist}(x_0, M) = \inf\{\|x_0 - m\| : m \in M\}$.*

Then there exists $\Lambda \in X^*$ such that

$$\langle x_0, \Lambda \rangle = 1, \qquad \Lambda|_M = 0, \qquad \textit{and} \qquad \|\Lambda\|_{X^*} = \frac{1}{d}.$$

Proof. Note that $d > 0$ since M is closed. Define $M_1 = \operatorname{span}\{M, x_0\}$. Then each $x \in M_1$ can be written as $x = m_x + t_x x_0$ for some $m_x \in M$ and $t_x \in \mathbf{F}$, and since $x_0 \notin M$, this representation is unique (verify!). Define $\lambda \colon M_1 \to \mathbf{F}$ by $\langle x, \lambda \rangle = t_x$. Then λ is linear, $\lambda|_M = 0$, and $\langle x_0, \lambda \rangle = 1$.

If $x \in M_1$ and $t_x \neq 0$, then we have $m_x/t_x \in M$, so

$$\|x\| \;=\; \|t_x x_0 + m_x\|_X \;=\; |t_x| \left\| x_0 - \left(\frac{-m_x}{t_x}\right) \right\|_X \;\ge\; |t_x|\, d.$$

If $t_x = 0$ (so $x \in M$), this is still true. Hence, $|\langle x, \lambda \rangle| = |t_x| \le \|x\|_X / d$ for all $x \in M_1$. Therefore λ is continuous on M_1, and $\|\lambda\|_{M_1^*} \le 1/d$.

On the other hand, there exist vectors $m_n \in M$ such that $\|x_0 - m_n\|_X \to d$. Since λ vanishes on M, we therefore have

$$1 \;=\; \langle x_0, \lambda \rangle \;=\; \langle x_0 - m_n, \lambda \rangle \;\le\; \|x_0 - m_n\|_X \, \|\lambda\|_{M_1^*} \;\to\; d\, \|\lambda\|_{M_1^*}.$$

Therefore $\|\lambda\|_{M_1^*} \ge 1/d$.

Applying Corollary 2.2, there exists a $\Lambda \in X^*$ such that $\Lambda|_{M_1} = \lambda$ and $\|\Lambda\|_{X^*} = \|\lambda\|_{M_1^*}$. This functional Λ has all of the required properties. □

Unlike the preceding corollaries, the next corollary is usually not given a special name, but we will have occasion to use it often (compare Lemma 1.44 for the case of Hilbert spaces).

Corollary 2.5. *Let X be a Banach space. Then $\{x_n\} \subseteq X$ is complete if and only if the following statement holds:*

$$x^* \in X^* \text{ and } \langle x_n, x^* \rangle = 0 \text{ for every } n \quad \Longrightarrow \quad x^* = 0.$$

Proof. $\Rightarrow$. Suppose that $\{x_n\}$ is complete, i.e., $\overline{\text{span}}\{x_n\} = X$. Suppose that $x^* \in X^*$ satisfies $\langle x_n, x^* \rangle = 0$ for every n. Since x^* is linear, we therefore have $\langle x, x^* \rangle = 0$ for every $x = \sum_{n=1}^{N} c_n x_n \in \text{span}\{x_n\}$. However, x^* is continuous, so this implies $\langle x, x^* \rangle = 0$ for every $x \in \overline{\text{span}}\{x_n\} = X$. Hence x^* is the zero functional.

$\Leftarrow$. Suppose that the only $x^* \in X^*$ satisfying $\langle x_n, x^* \rangle = 0$ for every n is $x^* = 0$. Define $Z = \overline{\text{span}}\{x_n\}$, and suppose that $Z \neq X$. Then we can find an element $y \in X$ such that $y \notin Z$. Since Z is a closed subset of X, we therefore have $d = \text{dist}(y, Z) > 0$. By the Hahn–Banach Theorem (Corollary 2.4), there exists a functional $\Lambda \in X^*$ satisfying $\langle y, \Lambda \rangle = 1 \neq 0$ and $\langle z, \Lambda \rangle = 0$ for every $z \in Z$. However, this implies that $\langle x_n, \Lambda \rangle = 0$ for every n. By hypothesis, Λ must then be the zero functional, contradicting the fact that $\langle y, \Lambda \rangle \neq 0$. Hence, we must have $Z = X$, so $\{x_n\}$ is complete in X. □

Exercises

2.1. Let M be a subspace of a Hilbert space H and fix $\lambda \in M$. Show directly that there exists some $\Lambda \in H$ such that $\langle x, \Lambda \rangle = \langle x, \lambda \rangle$ for all $x \in M$ and $\|\Lambda\| = \|\lambda\|$.

2.2. Suppose that M is a closed subspace of a Hilbert space H, $x_0 \in H \backslash M$, and $d = \text{dist}(x_0, M)$. Show directly that there exists a $\mu \in H$ such that $\langle x_0, \mu \rangle = 1$, $\langle x, \mu \rangle = 0$ for all $x \in M$, and $\|\mu\| = 1/d$.

2.3. Let X be a normed space. Show that if X^* is separable then X is separable, but the converse can fail.

2.4. The Weierstrass Approximation Theorem implies that $\{x^k\}_{k \geq 0}$ is complete in $C[0,1]$. Show that $\{x^{2k}\}_{k \geq 0}$ is also complete in $C[0,1]$.

2.5. Given a subset A of a normed space X, define its *orthogonal complement* $A^\perp \subseteq X^*$ by

$$A^\perp = \bigl\{\mu \in X^* : \langle x, \mu \rangle = 0 \text{ for all } x \in A\bigr\}.$$

Prove that $A^\perp$ is a closed subspace of X^*, and explain how this relates to Corollary 2.5.

2.6. Let S be a subspace of a normed space X, and show that its closure $\overline{S}$ is given by

$$\overline{S} = \bigcap \bigl\{\ker(\mu) : \mu \in X^* \text{ and } S \subseteq \ker(\mu)\bigr\}.$$

2.2 Reflexivity

Given a normed space X, its dual space X^* is a Banach space, so we can consider the dual of the dual space, which we denote by X^{**}. The next result shows that there is a natural isometry that maps X into X^{**}.

Theorem 2.6. *Let X be a normed linear space. Given $x \in X$, define $\pi(x) : X^* \to \mathbf{F}$ by*

$$\big\langle x^*, \pi(x) \big\rangle = \langle x, x^* \rangle, \qquad x^* \in X^*.$$

Then $\pi(x)$ is a bounded linear functional on X^*, and has operator norm

$$\|\pi(x)\|_{X^{**}} = \|x\|_X.$$

Consequently, the mapping

$$\begin{aligned} \pi\colon X &\to X^{**} \\ x &\mapsto \pi(x), \end{aligned}$$

is a linear isometry of X into X^{**}.

Proof. By definition of the operator norm,

$$\|\pi(x)\|_{X^{**}} = \sup_{x^* \in X^*,\, \|x^*\|_{X^*}=1} |\langle x^*, \pi(x) \rangle|.$$

On the other hand, by the Hahn–Banach Theorem in the form of Corollary 2.3,

$$\|x\| = \sup_{x^* \in X^*,\, \|x^*\|_{X^*}=1} |\langle x, x^* \rangle|.$$

Since $\langle x, x^* \rangle = \langle x^*, \pi(x) \rangle$, the result follows. □

Definition 2.7 (Natural Embedding of X into X^{**}). Let X be a normed space.

(a) The mapping $\pi\colon X \to X^{**}$ defined in Theorem 2.6 is called the *natural embedding* or the *canonical embedding* of X into X^{**}.

(b) If the natural embedding of X into X^{**} is surjective, then we say that X is *reflexive*. ◊

Note that in order for X to be called reflexive, the natural embedding must be a surjective isometry. There exist Banach spaces X such that X is isometrically isomorphic to X^{**} even though X is not reflexive [Jam51].

By the Riesz Representation Theorem (Theorem 1.75), every Hilbert space is reflexive.

Exercise 2.7 asks for a proof that ℓ^p is reflexive for each $1 < p < \infty$. However, ℓ^1 and ℓ^∞ are not reflexive. Another nonreflexive example is the space c_0, since by Exercise 1.75 we have $c_0{}^{**} \cong (\ell^1)^* \cong \ell^\infty$. The space c_0 is one of the few easily exhibited nonreflexive separable spaces whose dual is separable.

It is likewise true that $L^p(E)$ is reflexive when $1 < p < \infty$, but not for $p = 1$ or $p = \infty$.

Exercises

2.7. Show that ℓ^p is reflexive for each $1 < p < \infty$.

2.8. Let X be a Banach space. Show that if X is separable but X^* is not, then X is not reflexive. Use this to show that ℓ^1 is a proper subspace of $(\ell^\infty)^*$.

2.3 Adjoints of Operators on Banach Spaces

The duality between Banach spaces and their dual spaces allows us to define the "dual" of a bounded linear operator on Banach spaces.

Let X and Y be Banach spaces, and let $T\colon X \to Y$ be a bounded linear operator. Fix $\nu \in Y^*$, and define a functional $\mu\colon X \to \mathbf{F}$ by

$$\langle x, \mu \rangle \;=\; \langle Tx, \nu \rangle, \qquad x \in X.$$

That is, $\mu = \nu \circ T$. Then μ is linear since T and ν are linear. Further,

$$|\langle x, \mu \rangle| \;=\; |\langle Tx, \nu \rangle| \;\le\; \|Tx\|_Y \, \|\nu\|_{Y^*} \;\le\; \|T\| \, \|x\|_X \, \|\nu\|_{Y^*},$$

so

$$\|\mu\|_{X^*} \;=\; \sup_{\|x\|_X = 1} |\langle x, \mu \rangle| \;\le\; \|T\| \, \|\nu\|_{Y^*} \;<\; \infty. \tag{2.2}$$

Hence μ is bounded, so $\mu \in X^*$. Thus, for each $\nu \in Y^*$ we have defined a functional $\mu \in X^*$, so we can define an operator $T^*\colon Y^* \to X^*$ by setting $T^*\nu = \mu$. This mapping T^* is linear, and by equation (2.2) we have

$$\|T^*\nu\|_{X^*} \;=\; \|\mu\|_{X^*} \;\le\; \|T\| \, \|\nu\|_{Y^*}.$$

Taking the supremum over all unit vectors $\nu \in Y^*$, we conclude that T^* is bounded and $\|T^*\| \le \|T\|$.

We can use the Hahn–Banach Theorem to show that $\|T^*\| = \|T\|$. Choose any $x \in X$ with $\|x\|_X = 1$. By Corollary 2.3,

$$\|Tx\|_Y \;=\; \sup_{\|\nu\|_{Y^*} = 1} |\langle Tx, \nu \rangle|, \tag{2.3}$$

and this supremum is achieved. Let $\nu \in Y^*$ be any particular functional with unit norm that achieves the supremum in equation (2.3). Then we have

$$\begin{aligned}
\|Tx\|_Y \;=\; |\langle Tx, \nu \rangle| \;&=\; |\langle x, T^*\nu \rangle| \\
&\le\; \|x\|_X \, \|T^*\nu\|_{X^*} \\
&\le\; \|x\|_X \, \|T^*\| \, \|\nu\|_{Y^*} \\
&=\; \|x\|_X \, \|T^*\|.
\end{aligned}$$

Since this is true for every unit vector $x \in X$, we conclude that $\|T\| \le \|T^*\|$.

In summary, given $T \in \mathcal{B}(X,Y)$, we have constructed an operator $T^* \in \mathcal{B}(Y^*,X^*)$ that satisfies

$$\forall\, x \in X, \quad \forall\, \nu \in Y^*, \quad \big\langle Tx,\, \nu \big\rangle \;=\; \big\langle x,\, T^*\nu \big\rangle. \tag{2.4}$$

According to Exercise 2.9, there is a unique such operator, and we call it the adjoint of T.

Definition 2.8 (Adjoint). Given $T \in \mathcal{B}(X,Y)$, the unique operator $T^* \in \mathcal{B}(Y^*,X^*)$ satisfying equation (2.4) is called the *adjoint* of T. $\diamondsuit$

Example 2.9. Let $E \subseteq \mathbf{R}$ be Lebesgue measurable, choose $1 \le p < \infty$, and fix $m \in L^\infty(\mathbf{R})$. Let $T_m \colon L^p(\mathbf{R}) \to L^p(\mathbf{R})$ be the operation of pointwise multiplication of f by m, i.e., $T_m f = fm$ for $f \in L^p(\mathbf{R})$. Exercise 1.67 shows that T_m is bounded and has operator norm $\|T_m\| = \|m\|_{L^\infty}$. Therefore, $T_m^* \colon L^{p'}(E) \to L^{p'}(E)$ is the unique operator that satisfies

$$\langle f, T_m^* g\rangle \;=\; \langle T_m f, g\rangle \;=\; \langle fm, g\rangle \;=\; \int_E f(t)\, m(t)\, g(t)\, dt \;=\; \langle f, gm\rangle$$

for $f \in L^p(E)$ and $g \in L^p(E)^* = L^{p'}(E)$. Therefore $T_m^* g = gm$, so T_m^* is also multiplication by the function m. Technically, however, T_m and T_m^* are not the same operator, since T_m maps $L^p(E)$ into itself, while T_m^* maps $L^{p'}(E)$ into itself. $\diamondsuit$

Exercises

2.9. Let X, Y be Banach spaces. Given $T \in \mathcal{B}(X,Y)$, show that there is a unique operator $T^* \in \mathcal{B}(Y^*,X^*)$ that satisfies equation (2.4).

2.10. Let X be a Banach space. Given $\mu \in X^* = \mathcal{B}(X,\mathbf{F})$, explicitly describe its adjoint μ^*.

2.11. Let M be a closed subspace M of a normed space X, and fix $L \in \mathcal{B}(X)$. We say that M is *invariant* under L if $L(M) \subseteq M$. Show that if $M \subseteq X$ is invariant under L, then $M^\perp$ is invariant under L^*, where $M^\perp$ is the orthogonal complement defined in Exercise 2.5.

2.12. Suppose that M is a closed subspace of a Banach space X. Let $\epsilon \colon M \to X$ be the embedding map, i.e., $\epsilon(x) = x$ for $x \in M$. Show that $\epsilon^* \colon X^* \to M^*$ is the restriction map, i.e., if $\mu \in X^*$, then $\epsilon^*\mu = \mu|_M$.

2.4 Adjoints of Operators on Hilbert Spaces

Since Hilbert spaces are Banach spaces, if H, K are Hilbert spaces and $T \in \mathcal{B}(H,K)$, then there exists a unique adjoint operator $T^* \in \mathcal{B}(K^*, H^*)$. However, since Hilbert spaces are self-dual, we can regard the adjoint as belonging to $\mathcal{B}(K,H)$. In particular, if $K = H$ then T and T^* both belong to $\mathcal{B}(H)$. This makes adjoints of operators on Hilbert spaces quite special, and so we study them in more detail in this section.

Because of the conflict between our bilinear form notation $\langle x, x^* \rangle$ for functionals x^* acting on elements x and the inner product $\langle x, y \rangle$, which is antilinear as a function of y, the definition of adjoints on Hilbert spaces differs slightly from the definition on Banach spaces. We defined the adjoint using the bilinear form notation, but when dealing with a space that we know is a Hilbert space, it is usually more convenient to employ that space's inner product. Therefore, we define the adjoint of an operator on a Hilbert space as follows.

Definition 2.10 (Adjoint). Let H and K be Hilbert spaces. Let $\langle \cdot, \cdot \rangle_H$ denote the inner product on H, and $\langle \cdot, \cdot \rangle_K$ the inner product on K. If $A \in \mathcal{B}(H,K)$, then the *adjoint* of A is the unique operator $A^* \in \mathcal{B}(K,H)$ satisfying

$$\forall\, x \in H, \quad \forall\, y \in K, \quad \langle Ax, y \rangle_K \;=\; \langle x, A^*y \rangle_H. \qquad \diamondsuit$$

Comparing Definitions 2.8 and 2.10, we see that there is an ambiguity in the definition of an adjoint. We use the convention that if X, Y are Banach spaces then the adjoint of $T \in \mathcal{B}(X,Y)$ is defined by Definition 2.8, while if we know that H, K are Hilbert spaces then the adjoint of $A \in \mathcal{B}(H,K)$ is defined by Definition 2.10.

Example 2.11. Consider again the mapping T_m discussed in Example 2.9, but now consider the particular case $p = 2$. Since $L^2(E)$ is a Hilbert space, we define T_m^* to be the unique operator that, for f, $g \in L^2(E)$, satisfies

$$\begin{aligned} \langle f, T_m^* g \rangle \;=\; \langle T_m f, g \rangle \;=\; \langle fm, g \rangle \;&=\; \int_E f(t)\, m(t)\, \overline{g(t)}\, dt \\ &=\; \int_E f(t)\, \overline{g(t)\, \overline{m(t)}}\, dt \;=\; \langle f, g\overline{m} \rangle. \end{aligned}$$

Therefore $T_m^* g = g\overline{m}$, i.e., T_m^* is multiplication by the function $\overline{m}$. $\diamondsuit$

Thus, we see that Definitions 2.8 and 2.10 differ in how they define the adjoint. Fortunately, this is not a significant problem in practice.

Example 2.12. Consider the finite-dimensional Hilbert spaces $H = \mathbf{C}^n$ and $K = \mathbf{C}^m$. A linear operator $A\colon \mathbf{C}^n \to \mathbf{C}^m$ is given by multiplication by an $m \times n$ matrix A, which we identify with the operator A. The Hilbert space adjoint of A corresponds to multiplication by the *conjugate transpose* or *Hermitian* matrix $A^* = \overline{A^{\mathrm{T}}}$, while the Banach space adjoint corresponds to multiplication by the transpose matrix A^{T} (see Exercise 2.13). $\diamondsuit$

The next result summarizes some of the properties of adjoints on Hilbert spaces (see Exercise 2.16).

Theorem 2.13. *Let H, K, L be Hilbert spaces, and fix $A \in \mathcal{B}(H,K)$ and $B \in \mathcal{B}(K,L)$.*

(a) $(A^*)^* = A$.

(b) $(BA)^* = A^*B^*$.

(c) $\ker(A) = \mathrm{range}(A^*)^\perp$.

(d) $\ker(A)^\perp = \overline{\mathrm{range}(A^*)}$.

(e) *A is injective if and only if* $\mathrm{range}(A^*)$ *is dense in H.*

(f) $\|A\| = \|A^*\| = \|A^*A\|^{1/2} = \|AA^*\|^{1/2}$. ♢

We now make some definitions specifically for the case of adjoints of operators that map a Hilbert space into itself.

Definition 2.14. Let H be a Hilbert space, and let $A, B\colon H \to H$ be bounded linear operators.

(a) A is *self-adjoint* or *Hermitian* if $A = A^*$. By definition,

$$A \text{ is self-adjoint} \quad \Longleftrightarrow \quad \forall\, x, y \in H, \quad \langle Ax, y\rangle \;=\; \langle x, Ay\rangle.$$

(b) A is *positive*, denoted $A \ge 0$, if A is self-adjoint and $\langle Ax, x\rangle$ is real with $\langle Ax, x\rangle \ge 0$ for every $x \in H$.

(c) A is *positive definite* or *strictly positive*, denoted $A > 0$, if A is self-adjoint and $\langle Ax, x\rangle$ is real with $\langle Ax, x\rangle > 0$ for every $x \ne 0$.

(d) We write $A \ge B$ if $A - B \ge 0$, and $A > B$ if $A - B > 0$. ♢

We will need the following results for self-adjoint and positive operators.

Theorem 2.15. *If $A \in \mathcal{B}(H)$ is self-adjoint, then*

$$\|A\| \;=\; \sup_{\|x\|=1} |\langle Ax, x\rangle|.$$

Proof. Let us take $\mathbf{F} = \mathbf{C}$; the proof for real scalars is similar. Set $M = \sup_{\|x\|=1} |\langle Ax, x\rangle|$. By the Cauchy–Bunyakovski–Schwarz Inequality and the definition of operator norm, we have $M \le \|A\|$.

Choose any unit vectors $x, y \in H$. Then, by expanding the inner products, canceling terms, and using the fact that $A = A^*$, we see that

$$\begin{aligned}\big\langle A(x+y),\, x+y\big\rangle - \big\langle A(x-y),\, x-y\big\rangle \;&=\; 2\,\langle Ax, y\rangle + 2\,\langle Ay, x\rangle\\ &=\; 2\,\langle Ax, y\rangle + 2\,\langle y, Ax\rangle\\ &=\; 4\,\mathrm{Re}(\langle Ax, y\rangle).\end{aligned}$$

Applying the definition of M and using the Parallelogram Law, it follows that

$$\begin{aligned} 4\,\mathrm{Re}(\langle Ax,y\rangle) &\le |\langle A(x+y),x+y\rangle| + |\langle A(x-y),x-y\rangle| \\ &\le M\,\|x+y\|^2 + M\,\|x-y\|^2 \\ &= 2M\left(\|x\|^2+\|y\|^2\right) = 4M. \end{aligned}$$

That is, $\mathrm{Re}(\langle Ax,y\rangle) \le M$ for every choice of unit vectors x and y. Write $|\langle Ax,y\rangle| = \alpha\,\langle Ax,y\rangle$ where $\alpha \in \mathbf{C}$ satisfies $|\alpha|=1$. Then $\bar{\alpha}y$ is another unit vector, so

$$|\langle Ax,y\rangle| = \alpha\langle Ax,y\rangle = \langle Ax,\bar{\alpha}y\rangle \le M.$$

Using Lemma 1.36(c), we therefore have

$$\|Ax\| = \sup_{\|y\|=1} |\langle Ax,y\rangle| \le M.$$

Since this is true for every unit vector x, we conclude that $\|A\| \le M$. □

As a corollary, we obtain the following useful fact for self-adjoint operators.

Corollary 2.16. *Let H be a Hilbert space. If $A \in \mathcal{B}(H)$ is self-adjoint and $\langle Ax,x\rangle = 0$ for every $x \in H$, then $A = 0$.* ♢

Although we will not prove it, it can be shown that if H is a *complex* Hilbert space, then $A \in \mathcal{B}(H)$ is self-adjoint if and only if $\langle Ax,x\rangle$ is real for every $x \in H$. Hence for complex Hilbert spaces, the hypothesis in Corollary 2.16 that A is self-adjoint is redundant.

We end this section by proving that every positive operator A on a Hilbert space has a square root. That is, there exists a positive operator S such that $S^2 = A$. The idea of the proof is that if a is a real number with $0 < a < 1$ and if $(1-t)^2 = a$, then $t = \frac{1}{2}(1-a) + \frac{1}{2}t^2$ and the iteration $t_{n+1} = \frac{1}{2}(1-a) + \frac{1}{2}t_n^2$ converges to t. We make an operator analogue of this recursion. To prove convergence, we need the following lemma, which will be useful to us again in Chapter 8.

Lemma 2.17. *If $T\colon H \to H$ is a positive operator on a Hilbert space H, then*

$$\forall\, x,y \in H, \quad |\langle Tx,y\rangle|^2 \le \langle Tx,x\rangle\,\langle Ty,y\rangle.$$

Proof. By definition of a positive operator, $\langle Tx,x\rangle \ge 0$ for every $x \in H$. Therefore $(x,y) = \langle Tx,y\rangle$ defines a semi-inner product on H, and $|||x||| = (x,x)^{1/2}$ is a seminorm on H. In general, $(\cdot,\cdot)$ need not be an inner product (this happens if and only if T is positive definite). Still, the Cauchy–Bunyakovski–Schwarz Inequality holds for semi-inner products by Exercise 1.34, so we have

$$|\langle Tx,y\rangle|^2 = |(x,y)|^2 \le |||x|||^2\,|||y|||^2 = (x,x)\,(y,y) = \langle Tx,x\rangle\,\langle Ty,y\rangle. \quad □$$

Theorem 2.18. *If $A \in \mathcal{B}(H)$ is a positive operator on a Hilbert space H, then there exists a positive operator $A^{1/2} \in \mathcal{B}(H)$ such that $A^{1/2}A^{1/2} = A$. Moreover, $A^{1/2}$ commutes with A and with all operators that commute with A.*

Proof. We present some parts of the proof and assign the remainder as Exercise 2.23.

Suppose that $A \geq 0$. The result is trivial if A is the zero operator, so assume $A \neq 0$. Let $c = \|A\|^{-1}$. Then for every x we have

$$\langle cAx, x\rangle \;\leq\; |c|\, \|Ax\|\, \|x\| \;\leq\; |c|\, \|A\|\, \|x\|^2 \;=\; \|x\|^2 \;=\; \langle Ix, x\rangle,$$

which in operator notation says that $cA \leq I$. Since A has a square root if and only if cA has a square root, we can simply replace A by cA. That is, it suffices to prove the result under the assumptions that $A \geq 0$, $A \leq I$, and $\|A\| = 1$.

Let $B = I - A$. Set $T_0 = 0$, $T_1 = \frac{1}{2}B$, and

$$T_{n+1} \;=\; \frac{1}{2}\,(B + T_n^2), \qquad n \geq 2.$$

Each T_n is a polynomial in B, and therefore commutes with T_m and with every operator that commutes with B. The polynomial defining T_n has only nonnegative coefficients, so $T_n \geq 0$. Further, $T_{n+1} - T_n$ is also a polynomial in B *with all nonnegative coefficients*. Consequently, $T_n - T_m \geq 0$ for all $n \geq m \geq 0$.

By induction, $\|T_n\| \leq 1$ for every n. Therefore, if we fix $x \in H$ then the sequence $\big(\langle T_n x, x\rangle\big)$ is a bounded, increasing sequence of nonnegative real scalars. Hence this sequence must converge, and so is Cauchy. Now, if $n \geq m$ then by using Theorem 1.37(c) and Lemma 2.17 we compute that

$$\begin{aligned}
\|T_n x - T_m x\|^2 \;&=\; \sup_{\|y\|=1} |\langle (T_n - T_m)x, y\rangle|^2 \\
&\leq\; \sup_{\|y\|=1} |\langle (T_n - T_m)x, x\rangle|\; |\langle (T_n - T_m)y, y\rangle| \\
&\leq\; \sup_{\|y\|=1} |\langle T_n x, x\rangle - \langle T_m x, x\rangle|\; \|T_n - T_m\|\; \|y\|^2 \\
&\leq\; 2\, |\langle T_n x, x\rangle - \langle T_m x, x\rangle|.
\end{aligned}$$

Since $\big(\langle T_n x, x\rangle\big)$ is a Cauchy sequence of scalars, we conclude that $\{T_n x\}$ is a Cauchy sequence of vectors in H. Therefore $\{T_n x\}_{n\in\mathbf{N}}$ converges in H, and we define Tx to be the limit of this sequence. This operator T is bounded, linear, and positive, and it commutes with B and with every operator that commutes with B. Further, $T = \frac{1}{2}\,(B+T^2)$. Consequently, the operator $S = I-T$ satisfies $S^2 = A$, and S is positive since $\|T\| \leq 1$. □

In fact, the square root $A^{1/2}$ is unique; see Exercise 2.24.

Exercises

2.13. Let $A\colon \mathbf{C}^n \to \mathbf{C}^m$ be a linear operator, which we identify with its $m \times n$ matrix representation. Show that the adjoint of A in the Hilbert space sense (Definition 2.10) is the conjugate transpose matrix $A^* = \overline{A^{\mathrm{T}}}$, while the adjoint of A in the Banach space sense (Definition 2.8) is the transpose matrix A^{T}.

2.14. Let L, R be the left- and right-shift operators on ℓ^2 defined in Exercise 1.64. Show that $R = L^*$.

2.15. Fix $\lambda \in \ell^\infty$, and let M_λ be the multiplication operator defined in Exercise 1.66. Find M_λ^*, and determine when M_λ is self-adjoint, positive, or positive definite.

2.16. Prove Theorem 2.13.

2.17. Let M be a closed subspace of a Hilbert space H, and let $P \in \mathcal{B}(H)$ be given. Show that P is the orthogonal projection of H onto M if and only if $P^2 = P$, $P^* = P$, and $\mathrm{range}(P) = M$.

2.18. Let H be a Hilbert space and suppose that $A, B \in \mathcal{B}(H)$ are self-adjoint. Show that ABA, and BAB are self-adjoint, but AB is self-adjoint if and only if $AB = BA$. Exhibit self-adjoint operators A, B that do not commute.

2.19. Let H be a Hilbert space and let $A \in \mathcal{B}(H)$ be fixed.

(a) Show that if A is self-adjoint then all eigenvalues of A are real, and eigenvectors of A corresponding to distinct eigenvalues are orthogonal.

(b) Show that if A is a positive operator then all eigenvalues of A are real and nonnegative.

(c) Show that if A is a positive definite operator then all eigenvalues of A are real and strictly positive.

2.20. Let H, K be Hilbert spaces. Show that if $A \in \mathcal{B}(H,K)$, then $A^*A \in \mathcal{B}(H)$ and $AA^* \in \mathcal{B}(K)$ are positive operators.

2.21. Let H be a Hilbert space. Given $A \in \mathcal{B}(H)$, show that $\ker(A) = \ker(A^*A)$ and $\overline{\mathrm{range}(A^*A)} = \overline{\mathrm{range}(A^*)}$.

2.22. Let H, K be Hilbert spaces, and fix $U \in \mathcal{B}(H,K)$. Show that U is unitary if and only if U is a bijection and $U^{-1} = U^*$.

2.23. Fill in the details in the proof of Theorem 2.18.

2.24. Let A be a positive operator on a Hilbert space H.

(a) Show that $\langle Ax, x\rangle = 0$ if and only if $Ax = 0$.

(b) Show that the operator $A^{1/2}$ constructed in Theorem 2.18 is unique, i.e., there is only one positive operator S satisfying $S^2 = A$.

2.5 The Baire Category Theorem

Just as it is not possible to write the Euclidean plane $\mathbf{R}^2$ as the union of *countably many* straight lines, the Baire Category Theorem states that a complete metric space cannot be written as a countable union of "nowhere dense" sets. Since we are mainly interested in Banach spaces in this volume, we will prove this theorem in the setting of complete normed spaces, but the proof carries over without change to complete metric spaces.

Definition 2.19 (Nowhere Dense Sets). Let X be a Banach space, and let $E \subseteq X$ be given.

(a) E is *nowhere dense* or *rare* if $X \backslash \overline{E}$ is dense in X.

(b) E is *meager* or *first category* if it can be written as a countable union of nowhere dense sets.

(c) E is *nonmeager* or *second category* if it is not meager. ♢

We can restate the meaning of nowhere dense sets as follows (see Exercise 2.25).

Lemma 2.20. *Let E be a nonempty subset of a Banach space X. Then E is nowhere dense if and only if $\overline{E}$ contains no nonempty open subsets.* ♢

The set of rationals $\mathbf{Q}$ is not a nowhere dense subset of $\mathbf{R}$, but it is meager in $\mathbf{R}$. Although it is not a real vector space and hence not a normed space, $\mathbf{Q}$ under the metric $\mathrm{d}(x,y) = |x-y|$ is an example of an incomplete metric space that is a meager subset of itself.

Now we prove the Baire Category Theorem.

Theorem 2.21 (Baire Category Theorem). *Every Banach space X is a nonmeager subset of itself. Consequently, if*

$$X = \bigcup_{n=1}^{\infty} E_n$$

where each E_n is a closed subset of X, then at least one E_n contains a nonempty open subset.

Proof. Suppose that $X = \cup E_n$ where each E_n is nowhere dense. Then, by definition, $U_n = X \backslash \overline{E_n}$ is dense, and it is open since $\overline{E_n}$ is closed.

Choose $x_1 \in U_1$ and let $r_1 > 0$ be such that $B_1 = B_{r_1}(x_1) \subseteq U_1$. Then since U_2 is dense, there exists a point $x_2 \in U_2 \cap B_1$. Since U_2 and B_1 are both open, there exists some $r_2 > 0$ such that $B_2 = B_{r_2}(x_2) \subseteq U_2 \cap B_1$. Without loss of generality, we can take r_2 small enough that we have both $r_2 < r_1/2$ and $\overline{B_2} \subseteq B_1$. Continuing in this way we obtain points $x_n \in U_n$ and open balls $B_n = B_{r_n}(x_n) \subseteq U_n$ such that

$$r_n < \frac{r_{n-1}}{2} \qquad \text{and} \qquad \overline{B_n} \subseteq B_{n-1}.$$

In particular, $r_n \to 0$ and the balls B_n are nested.

Fix $\varepsilon > 0$, and let N be large enough so that $r_N < \varepsilon/2$. If $m, n > N$, then we have $x_m, x_n \in B_N$. Hence $\|x_m - x_n\| < 2r_N < \varepsilon$. Thus $\{x_n\}_{n\in\mathbf{N}}$ is Cauchy, and therefore there exists some $x \in X$ such that $x_n \to x$.

Now fix any $N > 0$. Then, since the B_n are nested, we have $x_n \in B_{N+1}$ for all $n > N$. As $x_n \to x$, this implies that $x \in \overline{B_{N+1}} \subseteq B_N$. This is true for every N, so

$$x \in \bigcap_{n=1}^{\infty} B_n \subseteq \bigcap_{n=1}^{\infty} U_n = \bigcap_{n=1}^{\infty} (X \backslash \overline{E_n}).$$

But then $x \notin \cup E_n$, which is a contradiction. □

Exercises

2.25. Prove Lemma 2.20.

2.26. Show that $C_c(\mathbf{R})$ is a meager subset of $C_0(\mathbf{R})$.

2.27. Suppose that f is an infinitely differentiable function on $\mathbf{R}$ such that for each $t \in \mathbf{R}$ there exists some integer $n_t \geq 0$ so that $f^{(n_t)}(t) = 0$. Prove that there exists some open interval (a, b) and some polynomial p such that $f(t) = p(t)$ for all $t \in (a, b)$.

2.28. Let D be the subset of $C[0,1]$ consisting of all functions $f \in C[0,1]$ that have a right-hand derivative at at least one point in $[0,1]$. Show that D is meager in $C[0,1]$, and conclude that there are functions in $C[0,1]$ that are not differentiable at any point.

2.6 The Uniform Boundedness Principle

The Uniform Boundedness Principle states that a family of bounded linear operators on a Banach space that are uniformly bounded *at each individual point* must actually be *uniformly bounded in operator norm.*

Theorem 2.22 (Uniform Boundedness Principle). *Let X be a Banach space and Y a normed linear space. If $\{A_i\}_{i\in I}$ is any collection of operators in $\mathcal{B}(X,Y)$ such that*

$$\forall\, x \in X, \quad \sup_{i\in I} \|A_i x\| < \infty,$$

then

$$\sup_{i\in I} \|A_i\| < \infty.$$

Proof. Set

$$E_n = \Big\{x \in X : \sup_{i \in I} \|A_i x\| \le n\Big\}.$$

Then $X = \cup E_n$ by hypothesis, and since each A_i is continuous it follows that E_n is closed. Consequently, the Baire Category Theorem implies that some E_n must contain an open ball, say $B_r(x_0) \subseteq E_n$.

Given any nonzero $x \in X$, if we set $y = x_0 + sx$ with $s = \frac{r}{2\|x\|}$ then we have $y \in B_r(x_0) \subseteq E_n$, and therefore

$$\|A_i x\| = \Big\|A_i\Big(\frac{y - x_0}{s}\Big)\Big\| \le \frac{1}{s}\big(\|A_i y\| + \|A_i x_0\|\big) \le \frac{2\,\|x\|}{r}\,2n = \frac{4n}{r}\,\|x\|.$$

Consequently, $\|A_i\| \le 4n/r$, which is a constant independent of i. □

The following special case of the Uniform Boundedness Principle is often useful (sometimes the names "Uniform Boundedness Principle" and "Banach–Steinhaus Theorem" are used interchangeably). The proof of Theorem 2.23 is assigned as Exercise 2.29.

Theorem 2.23 (Banach–Steinhaus Theorem). *Let X and Y be Banach spaces. If $A_n \in \mathcal{B}(X,Y)$ for $n \in \mathbf{N}$ and $Ax = \lim_{n\to\infty} A_n x$ exists for each $x \in X$, then $A \in \mathcal{B}(X,Y)$ and $\|A\| \le \sup_n \|A_n\| < \infty$.* ♢

Note that the hypotheses of the Banach–Steinhaus Theorem do *not* imply that $A_n \to A$ in operator norm. A counterexample is given in Exercise 2.30.

As an application of the Banach–Steinhaus Theorem, we prove a fact that was used earlier to show that the dual space ℓ^p is (isomorphic to) $\ell^{p'}$ when p is finite (see Theorem 1.73).

Theorem 2.24. *Fix $1 \le p \le \infty$ and any sequence of scalars $y = (y_k)$. Then $\sum x_k y_k$ converges for all $x \in \ell^p$ if and only if $y \in \ell^{p'}$. Furthermore, in this case $T_y x = (x_k y_k)$ defines a bounded linear map of ℓ^p into ℓ^1, and*

$$\sum_k |x_k y_k| = \|T_y x\|_{\ell^1} \le \|x\|_{\ell^p}\,\|y\|_{\ell^{p'}}, \qquad x \in \ell^p.$$

Proof. We will prove the case $1 < p < \infty$ (the cases $p = 1$ and $p = \infty$ are Exercise 2.32). Assume that $\sum x_k y_k$ converges for all $x \in \ell^p$. Define functionals T_N, $T\colon \ell^p \to \mathbf{F}$ by

$$Tx = \sum_{k=1}^{\infty} x_k y_k \qquad \text{and} \qquad T_N x = \sum_{k=1}^{N} x_k y_k.$$

Clearly T_N is linear, and for $x \in \ell^p$ we have

$$|T_N x| \le \Big(\sum_{k=1}^{N} |x_k|^p\Big)^{1/p} \Big(\sum_{k=1}^{N} |y_k|^{p'}\Big)^{1/p'} \le C_N\,\|x\|_{\ell^p},$$

where $C_N = \left(\sum_{k=1}^N |y_k|^{p'}\right)^{1/p'}$ is a finite constant independent of x (though not independent of N). Therefore $T_N \in \mathcal{B}(\ell^p, \mathbf{F}) = (\ell^p)^*$ for each N.

By hypothesis, $T_N x \to Tx$ as $N \to \infty$ for each $x \in \ell^p$. The Banach–Steinhaus Theorem therefore implies that $T \in \mathcal{B}(\ell^p, \mathbf{F}) = (\ell^p)^*$ and $\|T\| \le C = \sup \|T_N\| < \infty$.

At this point, if we accept the fact that $(\ell^p)^* = \ell^{p'}$ then we can argue as follows. Since $T \in (\ell^p)^*$ there must exist some $z \in \ell^{p'}$ such that $Tx = \langle x, z\rangle = \sum x_k z_k$ for all $x \in \ell^p$. Letting $\{\delta_k\}$ denote the standard basis vectors on ℓ^p, we have $y_k = T\delta_k = z_k$ for every k, so $y = z \in \ell^{p'}$.

However, since the current theorem was used in the proof that $(\ell^p)^* = \ell^{p'}$, in order to avoid circularity we need to give a direct proof that y belongs to $\ell^{p'}$. To do this, set

$$x_N \;=\; \left(\alpha_1 |y_1|^{p'-1}, \dots, \alpha_N |y_N|^{p'-1}, 0, 0, \dots\right) \;\in\; \ell^p,$$

where α_k is a scalar of unit modulus such that $\alpha_k y_k = |y_k|$. Then we have from the definition of T that

$$|Tx_N| \;=\; \sum_{k=1}^N \alpha_k |y_k|^{p'-1} y_k \;=\; \sum_{k=1}^N |y_k|^{p'},$$

while from $\|T\| \le C$ we obtain

$$|Tx_N| \;\le\; C\,\|x_N\|_{\ell^p} \;=\; C\left(\sum_{k=1}^N |y_k|^{(p'-1)p}\right)^{1/p} \;=\; C\left(\sum_{k=1}^N |y_k|^{p'}\right)^{1/p}.$$

Combining the two preceding equations, dividing through by $\left(\sum_{k=1}^N |y_k|^{p'}\right)^{1/p}$, and noting that $1 - \frac{1}{p} = \frac{1}{p'}$, this implies that

$$\left(\sum_{k=1}^N |y_k|^{p'}\right)^{1/p'} \;=\; \left(\sum_{k=1}^N |y_k|^{p'}\right)^{1-\frac{1}{p}} \;\le\; C.$$

Letting $N \to \infty$, we see that $\|y\|_{\ell^{p'}} \le C$. □

Exercises

2.29. Prove Theorem 2.23.

2.30. Let $\{e_n\}$ be an orthonormal basis for a Hilbert space H, and let P_N be the orthogonal projection of H onto $\mathrm{span}\{e_1, \dots, e_N\}$. Show that $P_N x \to x$ for every $x \in H$, but $\|I - P_N\| \not\to 0$ as $N \to \infty$.

2.31. Let X, Y be Banach spaces. Suppose $A_n \in \mathcal{B}(X,Y)$ for $n \in \mathbf{N}$ and $Ax = \lim_{n\to\infty} A_n x$ exists for each x in a dense subspace S of X.

(a) Show that if $\sup_n \|A_n\| < \infty$ then A extends to a bounded map on X, and $Ax = \lim_{n\to\infty} A_n x$ for all $x \in X$.

(b) Give an example that shows that the hypothesis $\sup_n \|A_n\| < \infty$ in part (a) is necessary.

2.32. Prove Theorem 2.24 for the cases $p = 1$ and $p = \infty$.

2.33. (a) Let X be a Banach space. Show that $S \subseteq X^*$ is bounded if and only if $\sup\bigl\{|\langle x, x^*\rangle| : x^* \in S\bigr\} < \infty$ for each $x \in X$.

(b) Let X be a normed linear space. Show that $S \subseteq X$ is bounded if and only if $\sup\bigl\{|\langle x, x^*\rangle| : x \in S\bigr\} < \infty$ for each $x^* \in X^*$.

2.34. Fix $1 \le p,\ q \le \infty$. Let $A = [a_{ij}]_{i,j\in\mathbf{N}}$ be an infinite matrix and set $a_i = (a_{ij})_{j\in\mathbf{N}}$ for each $i \in \mathbf{N}$. Suppose that

(a) $(Ax)_i = \langle x, a_i\rangle = \sum_j a_{ij}x_j$ converges for each $x \in \ell^p$ and $i \in \mathbf{N}$, and

(b) $Ax = \bigl((Ax)_i\bigr)_{i\in\mathbf{N}} = \bigl(\langle x, a_i\rangle\bigr)_{i\in\mathbf{N}} \in \ell^q$ for each $x \in \ell^p$.

Identifying the matrix A with the map $x \mapsto Ax$, prove that $A \in \mathcal{B}(\ell^p, \ell^q)$.

2.7 The Open Mapping Theorem

By Theorem 1.59, a function $f\colon X \to Y$ is continuous if the inverse image under f of any open subset of Y is open in X. It is often important to consider direct images of open sets as well.

Definition 2.25 (Open Mapping). Let X, Y be normed linear spaces. A function $A\colon X \to Y$ is an *open mapping* if

$$U \text{ is open in } X \implies A(U) \text{ is open in } Y. \qquad \diamondsuit$$

In general, a continuous function need not be an open mapping. For example, $f(x) = \sin x$ is a continuous mapping of the real line into itself, but f maps the open interval $(0, 2\pi)$ onto the closed interval $[-1, 1]$.

The Open Mapping Theorem asserts that any continuous linear surjection of one Banach space onto another must be an open mapping. The key to the proof is the following lemma. For clarity, we will write $B_r^X(x)$ and $B_r^Y(y)$ to distinguish open balls in X from open balls in Y.

Lemma 2.26. *Let X, Y be Banach spaces and fix $A \in \mathcal{B}(X,Y)$. If $\overline{A(B_1^X(0))}$ contains an open ball in Y, then $A(B_1^X(0))$ contains an open ball $B_r^Y(0)$ for some $r > 0$.*

Proof. Suppose that $\overline{A(B_1^X(0))}$ contains some open ball $B_s^Y(z)$. We claim that if we set $r = s/2$, then

$$B_r^Y(0) \subseteq \overline{A(B_1^X(0))}. \tag{2.5}$$

To see this, fix $x \in B_r^Y(0)$, i.e., $\|x\|_X < r = s/2$. Then $2x + z \in B_s^Y(z) \subseteq \overline{A(B_1^X(0))}$. Hence there exist vectors $y_n \in X$ with $\|y_n\|_X < 1$ such that $Ay_n \to 2x + z$. Also, $z \in B_s^Y(z) \subseteq \overline{A(B_1^X(0))}$, so there exist vectors $z_n \in X$ with $\|z_n\|_X < 1$ such that $Az_n \to z$. Then $w_n = (y_n - z_n)/2 \in B_1^X(0)$, and

$$Aw_n = \frac{Ay_n - Az_n}{2} \to \frac{(2x+z) - z}{2} = x \quad \text{as } n \to \infty.$$

Hence $x \in \overline{A(B_1^X(0))}$, so equation (2.5) holds.

Now we will show that we actually have $B_{r/2}^Y(0) \subseteq A(B_1^X(0))$. To see this, suppose that $y \in B_{r/2}^Y(0)$. Rescaling equation (2.5), we have $y \in \overline{A(B_{1/2}^X(0))}$, so there exists some $x_1 \in X$ with $\|x_1\| < 1/2$ such that $\|y - Ax_1\| < r/4$. Then $y - Ax_1 \in B_{r/4}^Y(0) \subseteq \overline{A(B_{1/4}^X(0))}$, so there exists some $x_2 \in X$ with $\|x_2\| < 1/4$ such that $\|(y - Ax_1) - Ax_2\| < r/8$. Continuing in this way, we obtain vectors $x_n \in X$ with $\|x_n\| < 2^{-n}$ such that

$$\|y - Az_n\| < \frac{r}{2^{n+1}},$$

where $z_n = \sum_{k=1}^n x_k$. Hence $Az_n \to y$. However, $\{z_n\}_{n\in\mathbf{N}}$ is Cauchy in X, so $z_n \to z$ for some $z \in X$. Since A is continuous, it follows that $y = Az$. Since $\|y\| < 1$, we therefore have $y \in A(B_1^X(0))$. □

Theorem 2.27 (Open Mapping Theorem). *If X, Y are Banach spaces and $A\colon X \to Y$ is a continuous linear surjection, then A is an open mapping.*

Proof. Since A is surjective, we have

$$Y = \bigcup_{k=1}^{\infty} \overline{A(B_k^X(0))}.$$

The Baire Category Theorem implies that some set $\overline{A(B_k^X(0))}$ must contain an open ball. Therefore, by Lemma 2.26, there is some $r > 0$ such that

$$B_r^Y(0) \subseteq A(B_1^X(0)). \tag{2.6}$$

Now suppose that $U \subseteq X$ is open and $y \in A(U)$. Then $y = Ax$ for some $x \in U$, so $B_s^X(x) \subseteq U$ for some $s > 0$. Rescaling equation (2.6), we have $B_t^Y(0) \subseteq A(B_s^X(0))$ for some $t > 0$. Therefore

$$B_t^Y(y) = B_t^Y(0) + Ax \subseteq A(B_s^X(0) + x) = A(B_s^X(x)) \subseteq A(U),$$

so $A(U)$ is open. □

The hypotheses in the Open Mapping Theorem that X and Y are both complete is necessary; see [Con90].

Exercises

2.35. Let X and Y be Banach spaces. Show that $A \in \mathcal{B}(X,Y)$ is surjective if and only if range(A) is not meager in Y.

2.8 Topological Isomorphisms

Topological isomorphisms will play an important role in the remainder of this volume.

Definition 2.28. Let X, Y be normed linear spaces.

(a) A linear operator $T\colon X \to Y$ is a *topological isomorphism* if T is a bijection and both T and T^{-1} are continuous.

(b) We say that X and Y are *topologically isomorphic* if there exists a topological isomorphism $T\colon X \to Y$. ♢

Every isometric isomorphism is a topological isomorphism, but the converse need not hold (see Exercise 2.39).

For the case of linear operators on Banach spaces, we have the following useful consequence of the Open Mapping Theorem.

Theorem 2.29 (Inverse Mapping Theorem). *If X, Y are Banach spaces and $T\colon X \to Y$ is a continuous linear bijection, then $T^{-1}\colon Y \to X$ is continuous. Consequently T is a topological isomorphism.*

Proof. The Open Mapping Theorem implies that T is an open mapping, so if $U \subseteq X$ is open then $T(U)$ is an open subset of Y. However, since T is a bijection we have $(T^{-1})^{-1}(U) = T(U)$. Hence the inverse image under T^{-1} of any open set is open, which implies by Theorem 1.59 that T^{-1} is continuous. □

The next result is a typical application of the Inverse Mapping Theorem.

Theorem 2.30. *Suppose X is a vector space that is complete with respect to each of two norms $\|\cdot\|$ and $|||\cdot|||$. If there exists $C > 0$ such that $\|x\| \le C\,|||x|||$ for all $x \in X$, then $\|\cdot\|$ and $|||\cdot|||$ are equivalent norms on X.*

Proof. The hypotheses imply that the identity map $I\colon (X, |||\cdot|||) \to (X, \|\cdot\|)$ is a bounded bijection, so, by the Inverse Mapping Theorem, the inverse map $I^{-1}\colon (X, \|\cdot\|) \to (X, |||\cdot|||)$ is a bounded bijection. Hence there is some $c > 0$ such that

$$|||x||| \;=\; |||I^{-1}(x)||| \;\le\; c\,\|x\|,$$

so the two norms are equivalent. □

The next theorem, whose proof is Exercise 2.42, states that the adjoint of a topological isomorphism is itself a topological isomorphism.

Theorem 2.31. *Let X, Y be Banach spaces. If $T\colon X \to Y$ is a topological isomorphism, then its adjoint $T^*\colon Y^* \to X^*$ is a topological isomorphism, and if T is an isometric isomorphism then so is T^*.* $\diamondsuit$

We will use the Inverse Mapping Theorem to derive two results for operators on Hilbert spaces. The following theorem shows that a bounded linear operator has closed range if and only if its adjoint has closed range (this also holds for operators on Banach space, see [Rud91, Thm. 4.14]).

Theorem 2.32. *Fix $A \in \mathcal{B}(H,K)$, where H and K are Hilbert spaces. Then*

$$\operatorname{range}(A) \text{ is closed} \quad \Longleftrightarrow \quad \operatorname{range}(A^*) \text{ is closed.}$$

Proof. $\Leftarrow$. Suppose that $\operatorname{range}(A^*)$ is closed, and let $M = \overline{\operatorname{range}(A)}$. Define $T \in \mathcal{B}(H,M)$ by $Tx = Ax$ for $x \in H$. Since $\operatorname{range}(T)$ is dense in M, Theorem 2.13 implies that $T^*\colon M \to H$ is injective. Given $y \in K$, write $y = m+e$ where $m \in M$ and $e \in M^\perp$. Since $\ker(A^*) = \operatorname{range}(A)^\perp = M^\perp$, for any $x \in H$ we have

$$\langle x, A^*y\rangle \;=\; \langle x, A^*m\rangle \;=\; \langle Ax, m\rangle \;=\; \langle Tx, m\rangle \;=\; \langle x, T^*m\rangle.$$

Hence $A^*y = A^*m = T^*m$, and it follows from this that $\operatorname{range}(T^*) = \operatorname{range}(A^*)$, which is closed. Now set $N = \operatorname{range}(T^*)$ and define $U \in \mathcal{B}(M,N)$ by $Uy = T^*y$ for $y \in M$. Then U is a continuous bijection, so it is a topological isomorphism by the Inverse Mapping Theorem. Theorem 2.31 therefore implies that $U^* \in \mathcal{B}(N,M)$ is a topological isomorphism. In particular, $\operatorname{range}(U^*) = M$ is closed.

Fix $y \in M$, so $y = U^*x$ for some $x \in N$. Let z be any vector in K, and let p be its orthogonal projection onto M. Then, since Ax, U^*x, and y all belong to M,

$$\begin{aligned}
\langle y, z\rangle &= \langle U^*x, z\rangle \\
&= \langle U^*x, p\rangle \\
&= \langle x, Up\rangle \\
&= \langle x, T^*p\rangle \\
&= \langle Tx, p\rangle \\
&= \langle Ax, p\rangle \;=\; \langle Ax, z\rangle.
\end{aligned}$$

Therefore $y = Ax$, so $M \subseteq \operatorname{range}(A)$ and hence $\operatorname{range}(A) = M$ is closed.

$\Rightarrow$. Since $(A^*)^* = A$, this follows from the previous case. $\square$

Our next application of the Inverse Mapping Theorem constructs a "pseudoinverse" of a bounded operator A that has closed range. Although A need not be injective, the pseudoinverse $A^\dagger$ acts as a right-inverse of A, at least when we restrict the domain of $A^\dagger$ to $\operatorname{range}(A)$.

Theorem 2.33. *Let H and K be Hilbert spaces. Assume that $A \in \mathcal{B}(H, K)$ has closed range, and let P be the orthogonal projection of K onto* range(A). *Then the mapping $B\colon \ker(A)^\perp \to \mathrm{range}(A)$ defined by $Bx = Ax$ for $x \in \ker(A)^\perp$ is a topological isomorphism, and $A^\dagger = B^{-1}P \in \mathcal{B}(K, H)$ satisfies the following:*

(a) $AA^\dagger y = y$ *for every* $y \in \mathrm{range}(A)$,

(b) $AA^\dagger$ *is the orthogonal projection of K onto* range(A), *and*

(c) $A^\dagger A$ *is the orthogonal projection of H onto* range(A^*).

Proof. The mapping B is bounded and linear since it is a restriction of the bounded mapping A. Further, the fact that $H = \ker(A) \oplus \ker(A)^\perp$ implies that B is a bijection of $\ker(A)^\perp$ onto range(A). Applying the Inverse Mapping Theorem, we conclude that $B\colon \ker(A)^\perp \to \mathrm{range}(A)$ is a topological isomorphism. Hence $B^{-1}\colon \mathrm{range}(A) \to \ker(A)^\perp$ is a topological isomorphism, and therefore $A^\dagger = B^{-1}P$ is bounded. We assign the proof of statements (a)–(c) as Exercise 2.43. □

Definition 2.34 (Pseudoinverse). Given $A \in \mathcal{B}(H, K)$, the operator $A^\dagger$ constructed in Theorem 2.33 is called the *Moore–Penrose pseudoinverse*, or simply the *pseudoinverse*, of A. ◇

Exercise 2.44 gives an equivalent characterization of the pseudoinverse.

Exercises

2.36. Show that if $T\colon X \to Y$ is a topological isomorphism of a normed space X onto a normed space Y, then a sequence $\{x_n\}$ is complete in X if and only if $\{Tx_n\}$ is complete in Y.

2.37. Let X and Y be normed linear spaces. Show that if $T\colon X \to Y$ is a topological isomorphism, then $\|T^{-1}\|^{-1}\,\|x\| \le \|Tx\| \le \|T\|\,\|x\|$ for all $x \in X$.

2.38. Let X be a Banach space and Y a normed linear space. Suppose that $L\colon X \to Y$ is bounded and linear. Prove that the following two statements are equivalent.

(a) There exists $c > 0$ such that $\|Lx\| \ge c\|x\|$ for all $x \in X$.

(b) L is injective and range(L) is closed.

Show further that, in case these hold, $L\colon X \to \mathrm{range}(L)$ is a topological isomorphism.

2.39. Given a sequence of scalars $\lambda = (\lambda_k)$, define a mapping T_λ on sequences $x = (x_k)$ by $T_\lambda x = (\lambda_k x_k)$. Prove the following statements.

(a) T_λ is a bounded map of ℓ^2 into itself if and only if $\lambda \in \ell^\infty$.

(b) T_λ is a topological isomorphism of ℓ^2 onto itself if and only if $0 < \inf |\lambda_k| \le \sup |\lambda_k| < \infty$.

(c) T_λ is an isometric isomorphism of ℓ^2 onto itself if and only if $|\lambda_k| = 1$ for every n.

2.40. Let X be a Banach space. Given $T \in \mathcal{B}(X)$, define $T^0 = I$. Show that if $\|T\| < 1$, then $I - T$ is a topological isomorphism of X onto itself and $(I-T)^{-1} = \sum_{n=0}^{\infty} T^n$, where the series converges in operator norm (this is called a *Neumann series* for $(I-T)^{-1}$).

2.41. Show that if X is a Banach space, Y is a normed linear space, and $T\colon X \to Y$ is a topological isomorphism, then Y is a Banach space.

2.42. Prove Theorem 2.31.

2.43. Show that the operator $A^\dagger = B^{-1}P$ defined in Theorem 2.33 satisfies statements (a)–(c) of that theorem.

2.44. Assume that $A \in \mathcal{B}(H,K)$ has closed range, and let $A^\dagger$ be its pseudoinverse. Prove the following statements.

(a) $\ker(A^\dagger) = \mathrm{range}(A)^\perp$.

(b) $\mathrm{range}(A^\dagger) = \ker(A)^\perp$.

(c) $AA^\dagger y = y$ for all $y \in \mathrm{range}(A)$.

(d) $A^\dagger$ is the unique operator in $\mathcal{B}(K,H)$ that satisfies statements (a)–(c) above.

2.45. Let H be a Hilbert space. Given a positive definite operator $A \in \mathcal{B}(H)$, prove the following statements.

(a) A is injective and has dense range.

(b) A is a topological isomorphism of H onto itself if and only if it is surjective. Show by example that a positive definite operator need not be a topological isomorphism.

(c) If A is a surjective positive definite operator, then $(x,y) = \langle Ax, y\rangle$ defines an inner product that is equivalent to the original inner product $\langle\cdot,\cdot\rangle$ on H.

2.9 The Closed Graph Theorem

The Closed Graph Theorem provides a convenient means of testing whether a linear operator on Banach spaces is continuous.

Theorem 2.35 (Closed Graph Theorem). *Let X and Y be Banach spaces. If $T\colon X \to Y$ is linear, then the following statements are equivalent.*

(a) *T is continuous.*

(b) *If $x_n \to x$ in X and $Tx_n \to y$ in Y, then $y = Tx$.*

Proof. (a) ⇒ (b). This follows immediately from the definition of continuity.

(b) ⇒ (a). Assume that statement (b) holds. Define

$$|||x||| \;=\; \|x\|_X + \|Tx\|_Y, \qquad x \in X.$$

Now we appeal to Exercise 2.46, which states that $|||\cdot|||$ is a norm on X and X is complete with respect to this norm.

Since $\|x\|_X \le |||x|||$ for $x \in X$ and X is complete with respect to both norms, it follows from Theorem 2.30 that there exists a constant $C > 0$ such that $|||x||| \le C\,\|x\|_X$ for $x \in X$. Consequently, $\|Tx\|_Y \le |||x||| \le C\,\|x\|_X$, so T is bounded. □

The name of the Closed Graph Theorem comes from the fact that hypothesis (b) in Theorem 2.35 can be equivalently formulated as follows: The graph of T, $\mathrm{graph}(T) = \{(f, Tf) : f \in X\}$, is a closed subset of the product space $X \times Y$.

Exercise 2.49 shows that the hypothesis in the Closed Graph Theorem that X is complete is necessary, and it can be shown that it is also necessary that Y be complete.

Exercises

2.46. Prove the claim in Theorem 2.35 that $|||\cdot|||$ is a norm on X and X is complete with respect to this norm.

2.47. Use the Closed Graph Theorem to give another proof of Theorem 2.24.

2.48. Use the Closed Graph Theorem to give another proof of Exercise 2.34.

2.49. Let $C_b(\mathbf{R})$ and $C_b^1(\mathbf{R})$ be as in Exercise 1.22, and assume that the norm on both of these spaces is the uniform norm. In this case $C_b(\mathbf{R})$ is complete, but $C_b^1(\mathbf{R})$ is not. Show that the differentiation operator $D\colon C_b^1(\mathbf{R}) \to C_b(\mathbf{R})$ given by $Df = f'$ is unbounded, but has a closed graph, i.e., if $f_n \to f$ uniformly and $f_n' \to g$ uniformly then $f' = g$.

2.10 Weak Convergence

In this section we discuss some types of "weak convergence" that we will occasionally make use of (see especially Section 4.7). Part (a) of the following definition recalls the usual notion of convergence as given in Definition 1.2, and parts (b) and (c) introduce some new types of convergence.

Definition 2.36. Let X be a Banach space.

(a) We say that a sequence $\{x_n\}$ of elements of X *converges* to $x \in X$ if $\lim_{n\to\infty} \|x - x_n\| = 0$. For emphasis, we sometimes refer to this type of convergence as *strong convergence* or *norm convergence.* We denote norm convergence by $x_n \to x$ or $\lim_{n\to\infty} x_n = x$.

(b) A sequence $\{x_n\}$ of elements of X *converges weakly* to $x \in X$ if

$$\forall\, x^* \in X^*, \qquad \lim_{n\to\infty} \langle x_n, x^* \rangle \;=\; \langle x, x^* \rangle.$$

We denote weak convergence by $x_n \xrightarrow{\text{w}} x$.

(c) A sequence $\{x_n^*\}$ of functionals in X^* *converges weak** to $x^* \in X^*$ if

$$\forall\, x \in X, \qquad \lim_{n\to\infty} \langle x, x_n^* \rangle \;=\; \langle x, x^* \rangle.$$

We denote weak* convergence by $x_n^* \xrightarrow{\text{w}^*} x^*$. ♢

Note that weak* convergence only applies to convergence of functionals in a dual space X^*. However, since X^* is a Banach space, we can consider strong or weak convergence of functionals in X^* as well as weak* convergence. By definition, strong (norm), weak, and weak* convergence of a sequence $\{x_n^*\}$ in X^* mean:

$$\begin{aligned}
x_n^* \to x^* \quad &\iff \quad \lim_{n\to\infty} \|x^* - x_n^*\| \;=\; 0,\\
x_n^* \xrightarrow{\text{w}} x^* \quad &\iff \quad \forall\, x^{**} \in X^{**}, \quad \lim_{n\to\infty} \langle x_n^*, x^{**} \rangle \;=\; \langle x^*, x^{**} \rangle,\\
x_n^* \xrightarrow{\text{w}^*} x^* \quad &\iff \quad \forall\, x \in X, \quad \lim_{n\to\infty} \langle x, x_n^* \rangle \;=\; \langle x, x^* \rangle.
\end{aligned}$$

If X is reflexive then $X = X^{**}$, and therefore $x_n^* \xrightarrow{\text{w}} x^*$ if and only if $x_n^* \xrightarrow{\text{w}^*} x^*$. For general Banach spaces, we have the following implications.

Lemma 2.37. *Let X be a Banach space, and let $x_n, x \in X$ and $x_n^*, x^* \in X^*$ be given.*

(a) *Strong convergence in X implies weak convergence in X:*

$$x_n \to x \qquad \Longrightarrow \qquad x_n \xrightarrow{\text{w}} x.$$

(b) *Weak convergence in X^* implies weak* convergence in X^*:*

$$x_n^* \xrightarrow{\text{w}} x^* \qquad \Longrightarrow \qquad x_n^* \xrightarrow{\text{w}^*} x^*.$$

Proof. (a) Suppose that $x_n \to x$ strongly, and fix any $x^* \in X^*$. Since x^* is continuous we have $\lim_{n\to\infty} \langle x_n, x^* \rangle = \langle x, x^* \rangle$, so $x_n \xrightarrow{\text{w}} x$.

(b) Suppose that x_n^*, $x^* \in X^*$ and $x_n^* \xrightarrow{\text{w}} x^*$. Given $x \in X$ we have $\pi(x) \in X^{**}$, where $\pi\colon X \to X^{**}$ is the natural embedding of X into X^{**}. By definition of weak convergence, $\lim_{n\to\infty} \langle x_n^*, x^{**}\rangle = \langle x^*, x^{**}\rangle$ for every $x^{**} \in X^{**}$. Taking $x^{**} = \pi(x)$ in particular, we have

$$\lim_{n\to\infty} \langle x, x_n^*\rangle \;=\; \lim_{n\to\infty} \big\langle x_n^*,\, \pi(x)\big\rangle \;=\; \big\langle x^*,\, \pi(x)\big\rangle \;=\; \langle x, x^*\rangle.$$

Thus $x_n^* \xrightarrow{\text{w}^*} x^*$. □

It is easy to see that strongly convergent sequences are norm-bounded above. It is a more subtle fact that the same is true of weakly convergent sequences.

Theorem 2.38. *Let X be a Banach space.*

(a) *If $\{x_n\} \subseteq X$ and $x_n \xrightarrow{\text{w}} x$ in X, then x is unique and* $\sup \|x_n\|_X < \infty$.

(b) *If $\{x_n^*\} \subseteq X^*$ and $x_n^* \xrightarrow{\text{w}^*} x^*$ in X^*, then x^* is unique and* $\sup \|x_n^*\|_{X^*} < \infty$.

Proof. We prove statement (a) and assign statement (b) as Exercise 2.50.

Suppose that $x_n \xrightarrow{\text{w}} x$. If we also had $x_n \xrightarrow{\text{w}} y$, then for each $x^* \in X^*$ we would have

$$\langle x-y, x^*\rangle \;=\; \langle x, x^*\rangle - \langle y, x^*\rangle \;=\; \lim_{n\to\infty} \langle x_n, x^*\rangle - \lim_{n\to\infty} \langle x_n, x^*\rangle \;=\; 0.$$

The Hahn–Banach Theorem (Corollary 2.3) therefore implies that $x = y$.

For each $x \in X$, let $\pi(x)$ be the image of x in X^{**} under the natural embedding of X into X^{**}. Then for each $x^* \in X^*$,

$$\lim_{n\to\infty} \big\langle x^*,\, \pi(x_n)\big\rangle \;=\; \lim_{n\to\infty} \langle x_n, x^*\rangle \;=\; \langle x, x^*\rangle.$$

Since convergent sequences of scalars are bounded, we therefore have

$$\forall\, x^* \in X^*, \quad \sup_n \big|\big\langle x^*,\, \pi(x_n)\big\rangle\big| < \infty.$$

Hence, by the Uniform Boundedness Principle, $\sup \|\pi(x_n)\|_{X^{**}} < \infty$. Since $\|\pi(x_n)\|_{X^{**}} = \|x_n\|_X$ (Theorem 2.6), we conclude that $\{x_n\}$ is bounded in X. □

Strong, weak, and weak* convergence can all be defined in terms of topologies on X or X^*. For example, the strong topology is induced from the norm $\|\cdot\|$ on X. The weak topology on X is induced from the family of seminorms $\rho_{x^*}(x) = |\langle x, x^*\rangle|$ with x^* ranging through X^*. The weak* topology on X^* is induced from the family of seminorms $\rho_x(x^*) = |\langle x, x^*\rangle|$ with x ranging through X. One difference between these latter two topologies and the strong topology is that, because the weak and weak* topologies are not defined by a norm, in order to rigorously relate topological concepts to limit concepts we must use *nets* instead of ordinary sequences indexed by the natural numbers. For example, a set $E \subseteq X$ is weakly closed if its complement is an open set in the weak topology, and this is equivalent to the requirement that E contains all of its weak limit points (compare Lemma 1.16). However, the definition of a weak limit is a point $x \in X$ for which there exists a *net* $\{x_i\}_{i \in I}$ such that x_i converges to x in the appropriate net sense (see the discussion in Section 3.2).

We will not pursue the connection between weak or weak* convergence and topologies in this volume, but we sketch the proof of one result in order to give a brief (albeit incomplete) illustration of these ideas.

Theorem 2.39. *Let M be a subspace of a normed space X. If M is strongly closed (i.e., closed with respect to the norm topology), then it is weakly closed (i.e., closed with respect to the weak topology).*

Proof. If $M = X$ then we are done, so suppose that M is strongly closed and there exists some vector $x \notin M$. Then, by the Hahn–Banach Theorem (Corollary 2.4), there exists an $x^* \in X^*$ such that $x^*|_M = 0$ and $\langle x, x^* \rangle = 1$.

By definition, X^* is the set of all *strongly continuous* linear functionals on X, so we know that the functional x^* is strongly continuous. On the other hand, if $x_n \xrightarrow{\text{w}} x$ then, by definition of weak convergence, $\langle x_n, x^* \rangle \to \langle x, x^* \rangle$. Hence, simply by definition, each element of X^* is *weakly continuous* (technically, we should justify this by using nets instead of sequences, but the idea is the same).

Just as in Theorem 1.59, weak continuity of x^* is equivalent to the fact that the inverse image of any open set in the codomain of x^* (which is $\mathbf{F}$) is weakly open in X. Therefore, since $\mathbf{F} \backslash \{0\}$ is an open subset of $\mathbf{F}$, the set $U = (x^*)^{-1}(\mathbf{F} \backslash \{0\}) \subseteq X$ is open in the weak topology. Since x^* maps every element of M to zero, no element of M is contained in U, i.e., $U \subseteq X \backslash M$. Further, $x \in U$ since $\langle x, x^* \rangle \neq 0$. Thus, given an arbitrary element $x \in X \backslash M$, we have found a weakly open set U such that $x \in U \subseteq X \backslash M$. Therefore $X \backslash M$ is open in the weak topology, which says that M is closed in the weak topology. $\square$

The converse of Theorem 2.39 is true as well, i.e., every weakly closed subspace is strongly closed. In fact, since strong convergence always implies weak convergence, every strong limit point of an arbitrary set is a weak limit point. Therefore, if a set is weakly closed then it contains all of its weak limit points and hence contains all of its strong limit points. Thus every weakly closed set is strongly closed. By taking complements, every weakly open set is strongly open, so the weak topology is a subset of the strong topology. However, the strong and weak topologies are distinct in infinite-dimensional spaces, so in general it is not true that every strongly closed *set* is weakly closed—this is why Theorem 2.39 is interesting!

The strong, weak, and weak* topologies are only three specific examples of topologies on a Banach space X or X^*. There are many other topologies that are useful in specific applications. Additionally, there are many other useful vector spaces that are not Banach spaces, but for which topologies can still be defined. We shall not deal with such *topological vector spaces*, but instead refer to texts such as [Con90] for details.

Exercises

2.50. Prove part (b) of Theorem 2.38.

2.51. In this exercise we will denote the components of $x \in \ell^p$ by $x = \big(x(k)\big)$.

(a) Given $1 < p < \infty$ and $x_n,\, y \in \ell^p$, show that $x_n \xrightarrow{\text{w}} y$ in ℓ^p if and only if $\sup \|x_n\|_{\ell^p} < \infty$ and x_n *converges componentwise* to y, i.e., $\lim_{n\to\infty} x_n(k) = y(k)$ for each $k \in \mathbf{N}$. Does either implication remain valid if $p = 1$?

(b) Given $1 \le p \le \infty$ and $x_n,\, y \in \ell^p$, show that $x_n \xrightarrow{\text{w*}} y$ in ℓ^p if and only if x_n converges componentwise to y and $\sup \|x_n\|_{\ell^p} < \infty$ (recall that $\ell^1 \cong {c_0}^*$ and $\ell^{p'} \cong (\ell^p)^*$ for $1 \le p < \infty$).

2.52. Show that if $\{x_n\}$ is an orthonormal sequence in a Hilbert space H, then $x_n \xrightarrow{\text{w}} 0$.