1

Elementary operator theory

1.1 Banach spaces

In this chapter we collect together material which should be covered in an introductory course of functional analysis and operator theory. We do not always include proofs, since there are many excellent textbooks on the subject.[1] The theorems provide a list of results which we use throughout the book.

We start at the obvious point. A normed space is a vector space $\mathcal{B}$ (assumed to be over the complex number field $\mathbf{C}$) provided with a norm $\|\cdot\|$ satisfying

$$\begin{aligned} \|f\| &\geq 0, \\ \|f\| &= 0 \;\text{ implies }\; f = 0, \\ \|\alpha f\| &= |\alpha|\,\|f\|, \\ \|f+g\| &\leq \|f\| + \|g\|, \end{aligned}$$

for all $\alpha \in \mathbf{C}$ and all $f, g \in \mathcal{B}$. Many of our definitions and theorems also apply to real normed spaces, but we will not keep pointing this out. We say that $\|\cdot\|$ is a seminorm if it satisfies all of the axioms except the second.

A Banach space is defined to be a normed space $\mathcal{B}$ which is complete in the sense that every Cauchy sequence in $\mathcal{B}$ converges to a limit in $\mathcal{B}$. Every normed space $\mathcal{B}$ has a completion $\overline{\mathcal{B}}$, which is a Banach space in which $\mathcal{B}$ is embedded isometrically and densely. (An isometric embedding is a linear, norm-preserving (and hence one-one) map of one normed space into another in which every element of the first space is identified with its image in the second.)

[1] One of the most systematic is [Dunford and Schwartz 1966].

Problem 1.1.1 Prove that the following conditions on a normed space $\mathcal{B}$ are equivalent:

(i) $\mathcal{B}$ is complete.
(ii) Every series $\sum_{n=1}^{\infty} f_n$ in $\mathcal{B}$ such that $\sum_{n=1}^{\infty} \|f_n\| < \infty$ is norm convergent.
(iii) Every series $\sum_{n=1}^{\infty} f_n$ in $\mathcal{B}$ such that $\|f_n\| \leq 2^{-n}$ for all n is norm convergent.

Prove also that any two completions of a normed space $\mathcal{B}$ are isometrically isomorphic. □

The following results from point set topology are rarely used below, but they provide worthwhile background knowledge. We say that a topological space X is normal if given any pair of disjoint closed subsets A, B of X there exists a pair of disjoint open sets U, V such that $A \subseteq U$ and $B \subseteq V$. All metric spaces and all compact Hausdorff spaces are normal. The size of the space of continuous functions on a normal space is revealed by Urysohn's lemma.

Lemma 1.1.2 *(Urysohn)*[2] *If A, B are disjoint closed sets in the normal topological space X, then there exists a continuous function $f : X \to [0, 1]$ such that $f(x) = 0$ for all $x \in A$ and $f(x) = 1$ for all $x \in B$.*

Problem 1.1.3 Use the continuity of the distance function $x \to \operatorname{dist}(x, A)$ to provide a direct proof of Urysohn's lemma when X is a metric space. □

Theorem 1.1.4 *(Tietze) Let S be a closed subset of the normal topological space X and let $f : S \to [0, 1]$ be a continuous function. Then there exists a continuous extension of f to X, i.e. a continuous function $g : X \to [0, 1]$ which coincides with f on S.*[3]

Problem 1.1.5 Prove the Tietze extension theorem by using Urysohn's lemma to construct a sequence of functions $g_n : X \to [0, 1]$ which converge uniformly on X and also uniformly on S to f. □

If K is a compact Hausdorff space then $C(K)$ stands for the space of all continuous complex-valued functions on K with the supremum norm

$$\|f\|_\infty := \sup\{|f(x)| : x \in K\}.$$

$C(K)$ is a Banach space with this norm, and the supremum is actually a maximum. We also use the notation $C_{\mathbf{R}}(K)$ to stand for the real Banach space of all continuous, real-valued functions on K.

[2] See [Bollobas 1999], [Simmons 1963, p. 135] or [Kelley 1955, p. 115].
[3] See [Bollobas 1999].

The following theorem is of interest in spite of the fact that it is rarely useful: in most applications it is equally evident that all four statements are true (or false).

Theorem 1.1.6 *(Urysohn) If K is a compact Hausdorff space then the following statements are equivalent.*

(i) K is metrizable;
(ii) the topology of K has a countable base;
(iii) K can be homeomorphically embedded in the unit cube $\Omega := \prod_{n=1}^{\infty}[0, 1]$ of countable dimension;
(iv) the space $C_{\mathbf{R}}(K)$ is separable in the sense that it contains a countable norm dense subset.

The equivalence of the first three statements uses methods of point-set topology, for which we refer to [Kelley 1955, p. 125]. The equivalence of the fourth statement uses the Stone-Weierstrass theorem 2.3.17.

Problem 1.1.7 Without using Theorem 1.1.6, prove that the topological product of a countable number of compact metrizable spaces is also compact metrizable. □

We say that $\mathcal{H}$ is a Hilbert space if it is a Banach space with respect to a norm associated with an inner product $f, g \to \langle f, g\rangle$ according to the formula

$$\|f\| := \sqrt{\langle f, f\rangle}.$$

We always assume that an inner product is linear in the first variable and conjugate linear in the second variable. We assume familiarity with the basic theory of Hilbert spaces. Although we do not restrict the statements of many theorems in the book to separable Hilbert spaces, we frequently only give the proof in that case. The proof in the non-separable context can usually be obtained by either of two devices: one may replace the word sequence by generalized sequence, or one may show that if the result is true on every separable subspace then it is true in general.

Example 1.1.8 If X is a finite or countable set then $l^2(X)$ is defined to be the space of all functions $f : X \to \mathbf{C}$ such that

$$\|f\|_2 := \sqrt{\sum_{x\in X} |f(x)|^2} < \infty.$$

This is the norm associated with the inner product

$$\langle f, g\rangle := \sum_{x\in X} f(x)\overline{g(x)},$$

the sum being absolutely convergent for all $f, g \in l^2(X)$. □

A sequence $\{\phi_n\}_{n=1}^{\infty}$ in a Hilbert space $\mathcal{H}$ is said to be an orthonormal sequence if

$$\langle \phi_m, \phi_n\rangle = \begin{cases} 1 & \text{if } m = n,\\ 0 & \text{otherwise.}\end{cases}$$

It is said to be a complete orthonormal sequence or an orthonormal basis, if it satisfies the conditions of the following theorem.

Theorem 1.1.9 *The following conditions on an orthonormal sequence $\{\phi_n\}_{n=1}^{\infty}$ in a Hilbert space $\mathcal{H}$ are equivalent.*

(i) The linear span of $\{\phi_n\}_{n=1}^{\infty}$ is a dense linear subspace of $\mathcal{H}$.
(ii) The identity

$$f = \sum_{n=1}^{\infty} \langle f, \phi_n\rangle \phi_n \tag{1.1}$$

holds for all $f \in \mathcal{H}$.
(iii) The identity

$$\|f\|^2 = \sum_{n=1}^{\infty} |\langle f, \phi_n\rangle|^2$$

holds for all $f \in \mathcal{H}$.
(iv) The identity

$$\langle f, g\rangle = \sum_{n=1}^{\infty} \langle f, \phi_n\rangle \langle \phi_n, g\rangle$$

holds for all $f, g \in \mathcal{H}$, the series being absolutely convergent.

The formula (1.1) is sometimes called a generalized Fourier expansion and $\langle f, \phi_n\rangle$ are then called the Fourier coefficients of f. The rate of convergence in (1.1) depends on f, and is discussed further in Theorem 5.4.12.

Problem 1.1.10 (Haar) Let $\{v_n\}_{n=0}^{\infty}$ be a dense sequence of distinct numbers in $[0, 1]$ such that $v_0 = 0$ and $v_1 = 1$. Put $e_1(x) := 1$ for all $x \in (0, 1)$ and define $e_n \in L^2(0,1)$ for $n = 2, 3, \ldots$ by

$$e_n(x) := \begin{cases} 0 & \text{if } x < u_n, \\ \alpha_n & \text{if } u_n < x < v_n, \\ -\beta_n & \text{if } v_n < x < w_n, \\ 0 & \text{if } x > w_n, \end{cases}$$

where

$$u_n := \max\{v_r : r < n \text{ and } v_r < v_n\},$$

$$w_n := \min\{v_r : r < n \text{ and } v_r > v_n\},$$

and $\alpha_n > 0$, $\beta_n > 0$ are the solutions of

$$\alpha_n(v_n - u_n) - \beta_n(w_n - v_n) = 0,$$

$$(v_n - u_n)\alpha_n^2 + (w_n - v_n)\beta_n^2 = 1.$$

Prove that $\{e_n\}_{n=1}^\infty$ is an orthonormal basis in $L^2(0,1)$. If $\{v_n\}_{n=0}^\infty$ is the sequence $\{0, 1, 1/2, 1/4, 3/4, 1/8, 3/8, 5/8, 7/8, 1/16, \ldots\}$ one obtains the standard Haar basis of $L^2(0,1)$, discussed in all texts on wavelets and of importance in image processing. If $\{m_r\}_{r=1}^\infty$ is a sequence of integers such that $m_1 \geq 2$ and m_r is a proper factor of m_{r+1} for all r, then one may define a generalized Haar basis of $L^2(0,1)$ by concatenating 0, 1, $\{r/m_1\}_{r=1}^{m_1}$, $\{r/m_2\}_{r=1}^{m_2}$, $\{r/m_3\}_{r=1}^{m_3}, \ldots$ and removing duplicated numbers as they arise. □

If X is a set with a σ-algebra Σ of subsets, and $\mathrm{d}x$ is a countably additive σ-finite measure on Σ, then the formula

$$\|f\|_2 := \sqrt{\int_X |f(x)|^2 \,\mathrm{d}x}$$

defines a norm on the space $L^2(X, \mathrm{d}x)$ of all functions f for which the integral is finite. The norm is associated with the inner product

$$\langle f, g \rangle := \int_X f(x)\overline{g(x)} \,\mathrm{d}x.$$

Strictly speaking one only gets a norm by identifying two functions which are equal almost everywhere. If the integral used is that of Lebesgue, then $L^2(X, \mathrm{d}x)$ is complete.[4]

Notation If $\mathcal{B}$ is a Banach space of functions on a locally compact, Hausdorff space X, then we will always use the notation $\mathcal{B}_c$ to stand for all those

[4] See [Lieb and Loss 1997] for one among many more complete accounts of Lebesgue integration. See also Section 2.1.

functions in $\mathcal{B}$ which have compact support, and $\mathcal{B}_0$ to stand for the closure of $\mathcal{B}_c$ in $\mathcal{B}$. Also $C_0(X)$ stands for the closure of $C_c(X)$ with respect to the supremum norm; equivalently $C_0(X)$ is the space of continuous functions on X that vanish at infinity. If X is a region in $\mathbf{R}^N$ then $C^n(X)$ will stand for the space of n times continuously differentiable functions on X.

Problem 1.1.11 The space $L^1(a, b)$ may be defined as the abstract completion of the space $\mathcal{P}$ of piecewise continuous functions on $[a, b]$, with respect to the norm

$$\|f\|_1 := \int_a^b |f(x)|\,\mathrm{d}x.$$

Without using any properties of Lebesgue integration prove that $C^k[a, b]$ is dense in $L^1(a, b)$ for every $k \geq 0$. $\square$

Lemma 1.1.12 *A finite-dimensional normed space V is necessarily complete. Any two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ on V are equivalent in the sense that there exist positive constants a and b such that*

$$a\|f\|_1 \leq \|f\|_2 \leq b\|f\|_1 \tag{1.2}$$

for all $f \in V$.

Problem 1.1.13 Find the optimal values of the constants a and b in (1.2) for the norms on $\mathbf{C}^n$ given by

$$\|f\|_1 := \sum_{r=1}^n |f_r|, \qquad \|f\|_2 := \left\{\sum_{r=1}^n |f_r|^2\right\}^{1/2}. \qquad \square$$

A bounded linear functional $\phi : \mathcal{B} \to \mathbf{C}$ is a linear map for which

$$\|\phi\| := \sup\{|\phi(f)| : \|f\| \leq 1\}$$

is finite. The dual space $\mathcal{B}^*$ of $\mathcal{B}$ is defined to be the space of all bounded linear functionals on $\mathcal{B}$, and is itself a Banach space for the norm given above. The Hahn-Banach theorem states that if L is any linear subspace of $\mathcal{B}$, then any bounded linear functional ϕ on L has a linear extension ψ to $\mathcal{B}$ which has the same norm:

$$\sup\{|\phi(f)|/\|f\| : 0 \neq f \in L\} = \sup\{|\psi(f)|/\|f\| : 0 \neq f \in \mathcal{B}\}.$$

It is not always easy to find a useful representation of the dual space of a Banach space, but the Hilbert space is particularly simple.

Theorem 1.1.14 *(Fréchet-Riesz)*[5] *If $\mathcal{H}$ is a Hilbert space then the formula*

$$\phi(f) := \langle f, g\rangle$$

defines a one-one correspondence between all $g \in \mathcal{H}$ and all $\phi \in \mathcal{H}^*$. Moreover $\|\phi\| = \|g\|$.

Note The correspondence $\phi \leftrightarrow g$ is conjugate linear rather than linear, and this can cause some confusion if forgotten.

Problem 1.1.15 Prove that if ϕ is a bounded linear functional on the closed linear subspace $\mathcal{L}$ of a Hilbert space $\mathcal{H}$, then there is only one linear extension of ϕ to $\mathcal{H}$ with the same norm. □

The following theorem is not elementary, and we will not use it until Chapter 13.1. The notation $C_{\mathbf{R}}(K)$ refers to the real Banach space of continuous functions $f : K \to \mathbf{R}$ with the supremum norm.[6]

Theorem 1.1.16 *(Riesz-Kakutani) Let K be a compact Hausdorff space and let $\phi \in C_{\mathbf{R}}(K)^*$. If ϕ is non-negative in the sense that $\phi(f) \geq 0$ for all non-negative $f \in C_{\mathbf{R}}(K)$ then there exists a non-negative countably additive measure μ on K such that*

$$\phi(f) = \int_X f(x)\, \mu(\mathrm{d}x)$$

for all $f \in C_{\mathbf{R}}(K)$. Moreover $\|\phi\| = \phi(1) = \mu(K)$.

One may reduce the representation of more general bounded linear functionals to the above special case by means of the following theorem. Given $\phi, \psi \in C_{\mathbf{R}}(K)^*$, we write $\phi \geq \psi$ if $\phi(f) \geq \psi(f)$ for all non-negative $f \in C_{\mathbf{R}}(K)$.

Theorem 1.1.17 *If K is a compact Hausdorff space and $\phi \in C_{\mathbf{R}}(K)^*$ then one may write $\phi := \phi_+ - \phi_-$ where $\phi_\pm$ are canonically defined, non-negative, bounded linear functionals. If $|\phi| := \phi_+ + \phi_-$ then $|\phi| \geq \pm\phi$. If $\psi \geq \pm\phi \in C_{\mathbf{R}}(K)^*$ then $\psi \geq |\phi|$. Finally $\|\,|\phi|\,\| = \|\phi\|$.*

[5] See [Dunford and Schwartz 1966, Theorem IV.4.5] for the proof.

[6] A combination of the next two theorems is usually called the Riesz representation theorem. According to [Dunford and Schwartz 1966, p. 380] Riesz provided an explicit representation of $C[0, 1]^*$. The corresponding theorem for $C_{\mathbf{R}}(K)^*$, where K is a general compact Hausdorff space, was obtained some years later by Kakutani. The formula $\phi := \phi_+ - \phi_-$ is called the Jordan decomposition. For the proof of the theorem see [Dunford and Schwartz 1966, Theorem IV.6.3]. A more abstract formulation, in terms of Banach lattices and AM-spaces, is given in [Schaefer 1974, Proposition II.5.5 and Section II.7].

Proof. The proof is straightforward but lengthy. Let $\mathcal{B} := C_{\mathbf{R}}(K)$, let $\mathcal{B}_+$ denote the convex cone of all non-negative continuous functions on K, and let $\mathcal{B}_+^*$ denote the convex cone of all non-negative functionals $\psi \in \mathcal{B}^*$.

Given $\phi \in \mathcal{B}^*$, we define $\phi_+ : \mathcal{B}_+ \to \mathbf{R}_+$ by

$$\phi_+(f) := \sup\{\phi(f_0) : 0 \le f_0 \le f\}.$$

If $0 \le f_0 \le f$ and $0 \le g_0 \le g$ then

$$\phi(f_0) + \phi(g_0) = \phi(f_0 + g_0) \le \phi_+(f+g).$$

Letting f_0 and g_0 vary subject to the stated constraints, we deduce that

$$\phi_+(f) + \phi_+(g) \le \phi_+(f+g)$$

for all $f,\ g \in \mathcal{B}_+$.

The reverse inequality is harder to prove. If $f,\ g \in \mathcal{B}_+$ and $0 \le h \le f+g$ then one puts $f_0 := \min\{h, f\}$ and $g_0 := h - f_0$. By considering each point $x \in K$ separately one sees that $0 \le f_0 \le f$ and $0 \le g_0 \le g$. hence

$$\phi(h) = \phi(f_0) + \phi(g_0) \le \phi_+(f) + \phi_+(g).$$

Since h is arbitrary subject to the stated constraints one obtains

$$\phi_+(f+g) \le \phi_+(f) + \phi_+(g)$$

for all $f,\ g \in \mathcal{B}_+$.

We are now in a position to extend ϕ_+ to the whole of $\mathcal{B}$. If $f \in \mathcal{B}$ we put

$$\phi_+(f) := \phi_+(f + \alpha 1) - \alpha\phi_+(1)$$

where $\alpha \in \mathbf{R}$ is chosen so that $f + \alpha 1 \ge 0$. The linearity of ϕ_+ on $\mathcal{B}_+$ implies that the particular choice of α does not matter subject to the stated constraint.

Our next task is to prove that the extended ϕ_+ is a linear functional on $\mathcal{B}_+$. If $f,\ g \in \mathcal{B}$, $f + \alpha 1 \ge 0$ and $g + \beta 1 \ge 0$, then

$$\begin{aligned}\phi_+(f+g) &= \phi_+(f + g + \alpha 1 + \beta 1) - (\alpha+\beta)\phi_+(1)\\ &= \phi_+(f+\alpha 1) + \phi_+(g + \beta 1) - (\alpha+\beta)\phi_+(1)\\ &= \phi_+(f) + \phi_+(g).\end{aligned}$$

It follows immediately from the definition that $\phi_+(\lambda h) = \lambda\phi_+(h)$ for all $\lambda \ge 0$ and $h \in \mathcal{B}_+$. Hence $f \in \mathcal{B}$ implies

$$\phi_+(\lambda f) = \phi(\lambda f + \lambda\alpha 1) - \lambda\alpha\phi_+(1) = \lambda\phi(f + \alpha 1) - \lambda\alpha\phi_+(1) = \lambda\phi_+(f).$$

If $\lambda < 0$ then

$$0 = \phi_+(\lambda f + |\lambda| f) = \phi_+(\lambda f) + \phi_+(|\lambda| f) = \phi_+(\lambda f) + |\lambda|\phi_+(f).$$

Therefore

$$\phi_+(\lambda f) = -|\lambda|\phi_+(f) = \lambda\phi_+(f).$$

Therefore ϕ_+ is a linear functional on $\mathcal{B}$. It is non-negative in the sense defined above.

We define ϕ_- by $\phi_- := \phi_+ - \phi$, and deduce immediately that it is linear. Since $f \in \mathcal{B}_+$ implies that $\phi_+(f) \geq \phi(f)$, we see that ϕ_- is non-negative. The boundedness of $\phi_\pm$ will be a consequence of the boundedness of $|\phi|$ and the formulae

$$\phi_+ = \tfrac{1}{2}(|\phi| + \phi), \qquad \phi_- = \tfrac{1}{2}(|\phi| - \phi).$$

We will need the following formula for $|\phi|$. If $f \in \mathcal{B}_+$ then the identity $|\phi| = 2\phi_+ - \phi$ implies

$$\begin{aligned} |\phi|(f) &= 2\sup\{\phi(f_0) : 0 \leq f_0 \leq f\} - \phi(f) \\ &= \sup\{\phi(2f_0 - f) : 0 \leq f_0 \leq f\} \\ &= \sup\{\phi(f_1) : -f \leq f_1 \leq f\}. \end{aligned} \tag{1.3}$$

The inequality $|\phi| \geq \pm\phi$ of the theorem follows from

$$\begin{aligned} |\phi| = \phi + 2\phi_- &\geq \phi \\ |\phi| = 2\phi_+ - \phi &\geq -\phi. \end{aligned}$$

If $\psi \geq \pm\phi$, $f \geq 0$ and $-f \leq f_1 \leq f$ then adding the two inequalities $(\psi + \phi)(f - f_1) \geq 0$ and $(\psi - \phi)(f + f_1) \geq 0$ yields $\psi(f) \geq \phi(f_1)$. Letting f_1 vary subject to the stated constraint we obtain $\psi(f) \geq |\phi|(f)$ by using (1.3). Therefore $\psi \geq |\phi|$.

We finally have to evaluate $\| \, |\phi| \, \|$. If $f \in \mathcal{B}$ and $\phi \in \mathcal{B}^*$ then

$$\begin{aligned} |\phi(f)| &= |\phi_+(f_+) - \phi_+(f_-) - \phi_-(f_+) + \phi_-(f_-)| \\ &\leq \phi_+(f_+) + \phi_+(f_-) + \phi_-(f_+) + \phi_-(f_-) \\ &= |\phi|(|f|) \\ &\leq \| \, |\phi| \, \| \, \| \, |f| \, \| \\ &= \| \, |\phi| \, \| \, \|f\|. \end{aligned}$$

Since f is arbitrary we deduce that $\|\phi\| \leq \| \, |\phi| \, \|$.

Conversely suppose that $f \in \mathcal{B}$. The inequality $-|f| \le f \le |f|$ implies

$$-|\phi|(|f|) \le |\phi|(f) \le |\phi|(|f|).$$

Therefore

$$\begin{aligned} \big| |\phi|(f) \big| &\le |\phi|(|f|) \\ &= \sup\{\phi(f_1) : -|f| \le f_1 \le |f|\} \\ &\le \|\phi\| \sup\{\|f_1\| : -|f| \le f_1 \le |f|\} \\ &= \|\phi\| \, \|f\|. \end{aligned}$$

Hence $\| \, |\phi| \, \| \le \|\phi\|$. □

If L is a closed linear subspace of the normed space $\mathcal{B}$, then the quotient space $\mathcal{B}/L$ is defined to be the algebraic quotient, provided with the quotient norm

$$\|f + L\| := \inf\{\|f + g\| : g \in L\}.$$

It is known that if $\mathcal{B}$ is a Banach space then so is $\mathcal{B}/L$.

Problem 1.1.18 If $\mathcal{B} = C[a, b]$ and L is the subspace of all functions in $\mathcal{B}$ which vanish on the closed subset K of $[a, b]$, find an explicit representation of $\mathcal{B}/L$ and of its norm. □

The Hahn-Banach theorem implies immediately that there is a canonical and isometric embedding j from $\mathcal{B}$ into the second dual space $\mathcal{B}^{**} = (\mathcal{B}^*)^*$, given by

$$(jx)(\phi) := \phi(x)$$

for all $x \in \mathcal{B}$ and all $\phi \in \mathcal{B}^*$. The space $\mathcal{B}$ is said to be reflexive if j maps $\mathcal{B}$ one-one onto $\mathcal{B}^{**}$.

We will often use the more symmetrical notation $\langle x, \phi \rangle$ in place of $\phi(x)$, and regard $\mathcal{B}$ as a subset of $\mathcal{B}^{**}$, suppressing mention of its natural embedding.

Problem 1.1.19 Prove that $\mathcal{B}$ is reflexive if and only if $\mathcal{B}^*$ is reflexive. □

Example 1.1.20 The dual $\mathcal{B}^*$ of a Banach space $\mathcal{B}$ is usually not isometrically isomorphic to $\mathcal{B}$ even if $\mathcal{B}$ is reflexive. The following provides a large number of spaces for which they are isometrically isomorphic. We simply