Chapter 5

Banach Spaces

Many linear equations may be formulated in terms of a suitable linear operator acting on a Banach space. In this chapter, we study Banach spaces and linear operators acting on Banach spaces in greater detail. We give the definition of a Banach space and illustrate it with a number of examples. We show that a linear operator is continuous if and only if it is bounded, define the norm of a bounded linear operator, and study some properties of bounded linear operators. Unbounded linear operators are also important in applications: for example, differential operators are typically unbounded. We will study them in later chapters, in the simpler context of Hilbert spaces.

5.1 Banach spaces

A normed linear space is a metric space with respect to the metric d derived from its norm, where $d(x, y) = \|x - y\|$.

Definition 5.1 A *Banach space* is a normed linear space that is a complete metric space with respect to the metric derived from its norm.

The following examples illustrate the definition. We will study many of these examples in greater detail later on, so we do not present proofs here.

Example 5.2 For $1 \leq p < \infty$, we define the p-norm on $\mathbb{R}^n$ (or $\mathbb{C}^n$) by

$$\|(x_1, x_2, \ldots, x_n)\|_p = \left(|x_1|^p + |x_2|^p + \ldots + |x_n|^p\right)^{1/p}.$$

For $p = \infty$, we define the ∞, or maximum, norm by

$$\|(x_1, x_2, \ldots, x_n)\|_\infty = \max\left\{|x_1|, |x_2|, \ldots, |x_n|\right\}.$$

Then $\mathbb{R}^n$ equipped with the p-norm is a finite-dimensional Banach space for $1 \leq p \leq \infty$.

Example 5.3 The space $C([a,b])$ of continuous, real-valued (or complex-valued) functions on $[a,b]$ with the sup-norm is a Banach space. More generally, the space $C(K)$ of continuous functions on a compact metric space K equipped with the sup-norm is a Banach space.

Example 5.4 The space $C^k([a,b])$ of k-times continuously differentiable functions on $[a,b]$ is not a Banach space with respect to the sup-norm $\|\cdot\|_\infty$ for $k \geq 1$, since the uniform limit of continuously differentiable functions need not be differentiable. We define the C^k-norm by

$$\|f\|_{C^k} = \|f\|_\infty + \|f'\|_\infty + \ldots + \|f^{(k)}\|_\infty.$$

Then $C^k([a,b])$ is a Banach space with respect to the C^k-norm. Convergence with respect to the C^k-norm is uniform convergence of functions and their first k derivatives.

Example 5.5 For $1 \leq p < \infty$, the *sequence space* $\ell^p(\mathbb{N})$ consists of all infinite sequences $x = (x_n)_{n=1}^\infty$ such that

$$\sum_{n=1}^{\infty} |x_n|^p < \infty,$$

with the p-norm,

$$\|x\|_p = \left(\sum_{n=1}^{\infty} |x_n|^p \right)^{1/p}.$$

For $p = \infty$, the sequence space $\ell^\infty(\mathbb{N})$ consists of all bounded sequences, with

$$\|x\|_\infty = \sup\{|x_n| \mid n = 1, 2, \ldots\}.$$

Then $\ell^p(\mathbb{N})$ is an infinite-dimensional Banach space for $1 \leq p \leq \infty$. The sequence space $\ell^p(\mathbb{Z})$ of bi-infinite sequences $x = (x_n)_{n=-\infty}^\infty$ is defined in an analogous way.

Example 5.6 Suppose that $1 \leq p < \infty$, and $[a,b]$ is an interval in $\mathbb{R}$. We denote by $L^p([a,b])$ the set of Lebesgue measurable functions $f : [a,b] \to \mathbb{R}$ (or $\mathbb{C}$) such that

$$\int_a^b |f(x)|^p \, dx < \infty,$$

where the integral is a Lebesgue integral, and we identify functions that differ on a set of measure zero (see Chapter 12). We define the L^p-norm of f by

$$\|f\|_p = \left(\int_a^b |f(x)|^p \, dx \right)^{1/p}.$$

For $p = \infty$, the space $L^{\infty}([a, b])$ consists of the Lebesgue measurable functions $f : [a, b] \to \mathbb{R}$ (or $\mathbb{C}$) that are *essentially bounded* on $[a, b]$, meaning that f is bounded on a subset of $[a, b]$ whose complement has measure zero. The norm on $L^{\infty}([a, b])$ is the *essential supremum*

$$\|f\|_{\infty} = \inf\{M \mid |f(x)| \leq M \text{ a.e. in } [a, b]\}.$$

More generally, if Ω is a measurable subset of $\mathbb{R}^n$, which could be equal to $\mathbb{R}^n$ itself, then $L^p(\Omega)$ is the set of Lebesgue measurable functions $f : \Omega \to \mathbb{R}$ (or $\mathbb{C}$) whose pth power is Lebesgue integrable, with the norm

$$\|f\|_p = \left(\int_{\Omega} |f(x)|^p \, dx\right)^{1/p}.$$

We identify functions that differ on a set of measure zero. For $p = \infty$, the space $L^{\infty}(\Omega)$ is the space of essentially bounded Lebesgue measurable functions on Ω with the essential supremum as the norm. The spaces $L^p(\Omega)$ are Banach spaces for $1 \leq p \leq \infty$.

Example 5.7 The *Sobolev spaces*, $W^{k,p}$, consist of functions whose derivatives satisfy an integrability condition. If (a, b) is an open interval in $\mathbb{R}$, then we define $W^{k,p}((a, b))$ to be the space of functions $f : (a, b) \to \mathbb{R}$ (or $\mathbb{C}$) whose derivatives of order less than or equal to k belong to $L^p((a, b))$, with the norm

$$\|f\|_{W^{k,p}} = \left(\sum_{j=0}^{k} \int_a^b \left|f^{(j)}(x)\right|^p \, dx\right)^{1/p}.$$

The derivatives $f^{(j)}$ are defined in a weak, or distributional, sense as we explain later on. More generally, if Ω is an open subset of $\mathbb{R}^n$, then $W^{k,p}(\Omega)$ is the set of functions whose partial derivatives of order less than or equal to k belong to $L^p(\Omega)$. Sobolev spaces are Banach spaces. We will give more detailed definitions of these spaces, and state some of their main properties, in Chapter 12.

A closed linear subspace of a Banach space is a Banach space, since a closed subset of a complete space is complete. Infinite-dimensional subspaces need not be closed, however. For example, infinite-dimensional Banach spaces have proper dense subspaces, something which is difficult to visualize from our intuition of finite-dimensional spaces.

Example 5.8 The space of polynomial functions is a linear subspace of $C([0, 1])$, since a linear combination of polynomials is a polynomial. It is not closed, and Theorem 2.9 implies that it is dense in $C([0, 1])$. The set $\{f \in C([0, 1]) \mid f(0) = 0\}$ is a closed linear subspace of $C([0, 1])$, and is a Banach space equipped with the sup-norm.

Example 5.9 The set $\ell_c(\mathbb{N})$ of all sequences of the form $(x_1, x_2, \ldots, x_n, 0, 0, \ldots)$ whose terms vanish from some point onwards is an infinite-dimensional linear subspace of $\ell^p(\mathbb{N})$ for any $1 \le p \le \infty$. The subspace $\ell_c(\mathbb{N})$ is not closed, so it is not a Banach space. It is dense in $\ell^p(\mathbb{N})$ for $1 \le p < \infty$. Its closure in $\ell^\infty(\mathbb{N})$ is the space $c_0(\mathbb{N})$ of sequences that converge to zero.

A *Hamel basis*, or algebraic basis, of a linear space is a maximal linearly independent set of vectors. Each element of a linear space may be expressed as a unique *finite* linear combination of elements in a Hamel basis. Every linear space has a Hamel basis, and any linearly independent set of vectors may be extended to a Hamel basis by the repeated addition of linearly independent vectors to the set until none are left (a procedure which is formalized by the axiom of choice, or Zorn's lemma, in the case of infinite-dimensional spaces). A Hamel basis of an infinite-dimensional space is frequently very large. In a normed space, we have a notion of convergence, and we may therefore consider various types of topological bases in which infinite sums are allowed.

Definition 5.10 Let X be a separable Banach space. A sequence (x_n) is a *Schauder basis* of X if for every $x \in X$ there is a unique sequence of scalars (c_n) such that $x = \sum_{n=1}^{\infty} c_n x_n$.

The concept of a Schauder basis is not as straightforward as it may appear. The Banach spaces that arise in applications typically have Schauder bases, but Enflo showed in 1973 that there exist separable Banach spaces that do not have any Schauder bases. As we will see, this problem does not arise in Hilbert spaces, which always have an orthonormal basis.

Example 5.11 A Schauder basis $(f_n)_{n=0}^{\infty}$ of $C([0,1])$ may be constructed from "tent" functions. For $n = 0, 1$, we define

$$f_0(x) = 1, \qquad f_1(x) = x.$$

For $2^{k-1} < n \le 2^k$, where $k \ge 1$, we define

$$f_n(x) = \begin{cases} 2^k \left[x - \left(2^{-k}(2n-2) - 1\right)\right] & \text{if } x \in I_n, \\ 1 - 2^k \left[x - \left(2^{-k}(2n-1) - 1\right)\right] & \text{if } x \in J_n, \\ 0 & \text{otherwise,} \end{cases}$$

where

$$\begin{aligned} I_n &= [2^{-k}(2n-2), 2^{-k}(2n-1)), \\ J_n &= [2^{-k}(2n-1), 2^{-k}2n). \end{aligned}$$

The graphs of these functions form a sequence of "tents" of height one and width 2^{-k+1} that sweep across the interval $[0,1]$. If $f \in C([0,1])$, then we may compute the coefficients c_n in the expansion

$$f(x) = \sum_{n=0}^{\infty} c_n f_n(x)$$

by equating the values of f and the series at the points $x = 2^{-k}m$ for $k \in \mathbb{N}$ and $m = 0, 1, \ldots, 2^k$. The uniform continuity of f implies that the resulting series converges uniformly to f.

5.2 Bounded linear maps

A *linear map* or *linear operator* T between real (or complex) linear spaces X, Y is a function $T : X \to Y$ such that

$$T(\lambda x + \mu y) = \lambda Tx + \mu Ty \qquad \text{for all } \lambda, \mu \in \mathbb{R} \text{ (or } \mathbb{C}\text{) and } x, y \in X.$$

A linear map $T : X \to X$ is called a *linear transformation of* X, or a *linear operator on* X. If $T : X \to Y$ is one-to-one and onto, then we say that T is *nonsingular* or *invertible*, and define the inverse map $T^{-1} : Y \to X$ by $T^{-1}y = x$ if and only if $Tx = y$, so that $TT^{-1} = I$, $T^{-1}T = I$. The linearity of T implies the linearity of T^{-1}.

If X, Y are normed spaces, then we can define the notion of a bounded linear map. As we will see, the boundedness of a linear map is equivalent to its continuity.

Definition 5.12 Let X and Y be two normed linear spaces. We denote both the X and Y norms by $\|\cdot\|$. A linear map $T : X \to Y$ is *bounded* if there is a constant $M \geq 0$ such that

$$\|Tx\| \leq M\|x\| \qquad \text{for all } x \in X. \tag{5.1}$$

If no such constant exists, then we say that T is *unbounded*. If $T : X \to Y$ is a bounded linear map, then we define the *operator norm* or *uniform norm* $\|T\|$ of T by

$$\|T\| = \inf\{M \mid \|Tx\| \leq M\|x\| \text{ for all } x \in X\}. \tag{5.2}$$

We denote the set of all linear maps $T : X \to Y$ by $\mathcal{L}(X, Y)$, and the set of all bounded linear maps $T : X \to Y$ by $\mathcal{B}(X, Y)$. When the domain and range spaces are the same, we write $\mathcal{L}(X, X) = \mathcal{L}(X)$ and $\mathcal{B}(X, X) = \mathcal{B}(X)$.

Equivalent expressions for $\|T\|$ are:

$$\|T\| = \sup_{x \neq 0} \frac{\|Tx\|}{\|x\|}; \quad \|T\| = \sup_{\|x\| \leq 1} \|Tx\|; \quad \|T\| = \sup_{\|x\| = 1} \|Tx\|. \tag{5.3}$$

We also use the notation $\mathbb{R}^{m \times n}$, or $\mathbb{C}^{m \times n}$, to denote the space of linear maps from $\mathbb{R}^n$ to $\mathbb{R}^m$, or $\mathbb{C}^n$ to $\mathbb{C}^m$, respectively.

Example 5.13 The linear map $A : \mathbb{R} \to \mathbb{R}$ defined by $Ax = ax$, where $a \in \mathbb{R}$, is bounded, and has norm $\|A\| = |a|$.

Example 5.14 The identity map $I : X \to X$ is bounded on any normed space X, and has norm one. If a map has norm zero, then it is the zero map $0x = 0$.

Linear maps on infinite-dimensional normed spaces need not be bounded.

Example 5.15 Let $X = C^\infty([0,1])$ consist of the smooth functions on $[0,1]$ that have continuous derivatives of all orders, equipped with the maximum norm. The space X is a normed space, but it is not a Banach space, since it is incomplete. The differentiation operator $Du = u'$ is an unbounded linear map $D : X \to X$. For example, the function $u(x) = e^{\lambda x}$ is an eigenfunction of D for any $\lambda \in \mathbb{R}$, meaning that $Du = \lambda u$. Thus $\|Du\|/\|u\| = |\lambda|$ may be arbitrarily large. The unboundedness of differential operators is a fundamental difficulty in their study.

Suppose that $A : X \to Y$ is a linear map between finite-dimensional real linear spaces X, Y with $\dim X = n$, $\dim Y = m$. We choose bases $\{e_1, e_2, \ldots, e_n\}$ of X and $\{f_1, f_2, \ldots, f_m\}$ of Y. Then

$$A(e_j) = \sum_{i=1}^{m} a_{ij} f_i,$$

for a suitable $m \times n$ matrix (a_{ij}) with real entries. We expand $x \in X$ as

$$x = \sum_{i=1}^{n} x_i e_i, \tag{5.4}$$

where $x_i \in \mathbb{R}$ is the ith component of x. It follows from the linearity of A that

$$A\left(\sum_{j=1}^{n} x_j e_j\right) = \sum_{i=1}^{m} y_i f_i,$$

where

$$y_i = \sum_{j=1}^{n} a_{ij} x_j.$$

Thus, given a choice of bases for X, Y we may represent A as a linear map $A : \mathbb{R}^n \to \mathbb{R}^m$ with matrix $A = (a_{ij})$, where

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}. \tag{5.5}$$

We will often use the same notation A to denote a linear map on a finite-dimensional space and its associated matrix, but it is important not to confuse the geometrical notion of a linear map with the matrix of numbers that represents it.

Each pair of norms on $\mathbb{R}^n$ and $\mathbb{R}^m$ induces a corresponding operator, or matrix, norm on A. We first consider the *Euclidean norm*, or 2-norm, $\|A\|_2$ of A. The Euclidean norm of a vector x is given by $\|x\|_2^2 = (x, x)$, where $(x, y) = x^T y$. From (5.3), we may compute the Euclidean norm of A by maximizing the function $\|Ax\|_2^2$ on the unit sphere $\|x\|_2^2 = 1$. The maximizer x is a critical point of the function

$$f(x, \lambda) = (Ax, Ax) - \lambda \{(x, x) - 1\},$$

where λ is a Lagrange multiplier. Computing ∇f and setting it equal to zero, we find that x satisfies

$$A^T A x = \lambda x. \tag{5.6}$$

Hence, x is an eigenvector of the matrix $A^T A$ and λ is an eigenvalue. The matrix $A^T A$ is an $n \times n$ symmetric matrix, with real, nonnegative eigenvalues. At an eigenvector x of $A^T A$ that satisfies (5.6), normalized so that $\|x\|_2 = 1$, we have $(Ax, Ax) = \lambda$. Thus, the maximum value of $\|Ax\|_2^2$ on the unit sphere is the maximum eigenvalue of $A^T A$.

We define the *spectral radius* $r(B)$ of a matrix B to be the maximum absolute value of its eigenvalues. It follows that the Euclidean norm of A is given by

$$\|A\|_2 = \sqrt{r\left(A^T A\right)}. \tag{5.7}$$

In the case of linear maps $A : \mathbb{C}^n \to \mathbb{C}^m$ on finite dimensional complex linear spaces, equation (5.7) holds with $A^T A$ replaced by $A^* A$, where A^* is the Hermitian conjugate of A. Proposition 9.7 gives a formula for the spectral radius of a bounded operator in terms of the norms of its powers.

To compute the maximum norm of A, we observe from (5.5) that

$$\begin{aligned} |y_i| &\leq |a_{i1}||x_1| + |a_{i2}||x_2| + \ldots + |a_{in}||x_n| \\ &\leq (|a_{i1}| + |a_{i2}| + \ldots + |a_{in}|)\, \|x\|_\infty. \end{aligned}$$

Taking the maximum of this equation with respect to i and comparing the result with the definition of the operator norm, we conclude that

$$\|A\|_\infty \leq \max_{1 \leq i \leq m} (|a_{i1}| + |a_{i2}| + \ldots + |a_{in}|).$$

Conversely, suppose that the maximum on the right-hand side of this equation is attained at $i = i_0$. Let x be the vector with components $x_j = \operatorname{sgn} a_{i_0 j}$, where sgn is the *sign function*,

$$\operatorname{sgn} x = \begin{cases} 1 & \text{if } x > 0, \\ 0 & \text{if } x = 0, \\ -1 & \text{if } x < 0. \end{cases} \tag{5.8}$$

Then, if A is nonzero, we have $\|x\|_\infty = 1$, and

$$\|Ax\|_\infty = |a_{i_0 1}| + |a_{i_0 2}| + \ldots + |a_{i_0 n}|.$$

Since $\|A\|_\infty \geq \|Ax\|_\infty$, we obtain that

$$\|A\|_\infty \geq \max_{1\leq i\leq m} \left(|a_{i1}| + |a_{i2}| + \ldots + |a_{in}|\right).$$

Therefore, we have equality, and the maximum norm of A is given by the maximum row sum,

$$\|A\|_\infty = \max_{1\leq i\leq m} \left\{ \sum_{j=1}^{n} |a_{ij}| \right\}. \tag{5.9}$$

A similar argument shows that the sum norm of A is given by the maximum column sum

$$\|A\|_1 = \max_{1\leq j\leq n} \left\{ \sum_{i=1}^{m} |a_{ij}| \right\}.$$

For $1 < p < \infty$, one can show (see Kato [26]) that the p-matrix norm satisfies

$$\|A\|_p \leq \|A\|_1^{1/p} \|A\|_\infty^{1-1/p}.$$

There are norms on the space $\mathcal{B}(\mathbb{R}^n, \mathbb{R}^m) = \mathbb{R}^{m\times n}$of $m \times n$ matrices that are not associated with any vector norms on $\mathbb{R}^n$ and $\mathbb{R}^m$. An example is the *Hilbert-Schmidt* norm

$$\|A\| = \left(\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2 \right)^{1/2}.$$

Next, we give some examples of linear operators on infinite-dimensional spaces.

Example 5.16 Let $X = \ell^\infty(\mathbb{N})$ be the space of bounded sequences $\{(x_1, x_2, \ldots)\}$ with the norm

$$\|(x_1, x_2, \ldots)\|_\infty = \sup_{i\in\mathbb{N}} |x_i|.$$

A linear map $A : X \to X$ is represented by an infinite matrix $(a_{ij})_{i,j=1}^\infty$, where

$$(Ax)_i = \sum_{j=1}^{\infty} a_{ij} x_j.$$

In order for this sum to converge for any $x \in \ell^\infty(\mathbb{N})$, we require that

$$\sum_{j=1}^{\infty} |a_{ij}| < \infty$$

for each $i \in \mathbb{N}$, and in order for Ax to belong to $\ell^\infty(\mathbb{N})$, we require that

$$\sup_{i\in\mathbb{N}} \left\{ \sum_{j=1}^{\infty} |a_{ij}| \right\} < \infty.$$

Then A is a bounded linear operator on $\ell^\infty(\mathbb{N})$, and its norm is the maximum row sum,

$$\|A\|_\infty = \sup_{i\in\mathbb{N}} \left\{ \sum_{j=1}^{\infty} |a_{ij}| \right\}.$$

Example 5.17 Let $X = C([0,1])$ with the maximum norm, and

$$k : [0,1] \times [0,1] \to \mathbb{R}$$

be a continuous function. We define the linear Fredholm integral operator $K : X \to X$ by

$$Kf(x) = \int_0^1 k(x,y) f(y)\, dy.$$

Then K is bounded and

$$\|K\| = \max_{0\le x\le 1} \left\{ \int_0^1 |k(x,y)|\, dy \right\}.$$

This expression is the "continuous" analog of the maximum row sum for the ∞-norm of a matrix.

For linear maps, boundedness is equivalent to continuity.

Theorem 5.18 A linear map is bounded if and only if it is continuous.

Proof. First, suppose that $T : X \to Y$ is bounded. Then, for all $x, y \in X$, we have

$$\|Tx - Ty\| = \|T(x-y)\| \le M\|x-y\|,$$

where M is a constant for which (5.1) holds. Therefore, we can take $\delta = \epsilon/M$ in the definition of continuity, and T is continuous.

Second, suppose that T is continuous at 0. Since T is linear, we have $T(0) = 0$. Choosing $\epsilon = 1$ in the definition of continuity, we conclude that there is a $\delta > 0$ such that $\|Tx\| \le 1$ whenever $\|x\| \le \delta$. For any $x \in X$, with $x \neq 0$, we define $\tilde{x}$ by

$$\tilde{x} = \delta \frac{x}{\|x\|}.$$

Then $\|\tilde{x}\| \leq \delta$, so $\|T\tilde{x}\| \leq 1$. It follows from the linearity of T that

$$\|Tx\| = \frac{\|x\|}{\delta}\|T\tilde{x}\| \leq M\|x\|,$$

where $M = 1/\delta$. Thus T is bounded. □

The proof shows that if a linear map is continuous at zero, then it is continuous at every point. A nonlinear map may be bounded but discontinuous, or continuous at zero but discontinuous at other points.

The following theorem, sometimes called the BLT theorem for "bounded linear transformation" has many applications in defining and studying linear maps.

Theorem 5.19 (Bounded linear transformation) Let X be a normed linear space and Y a Banach space. If M is a dense linear subspace of X and

$$T : M \subset X \to Y$$

is a bounded linear map, then there is a unique bounded linear map $\overline{T} : X \to Y$ such that $\overline{T}x = Tx$ for all $x \in M$. Moreover, $\left\|\overline{T}\right\| = \|T\|$.

Proof. For every $x \in X$, there is a sequence (x_n) in M that converges to x. We define

$$\overline{T}x = \lim_{n\to\infty} Tx_n.$$

This limit exists because (Tx_n) is Cauchy, since T is bounded and (x_n) Cauchy, and Y is complete. We claim that the value of the limit does not depend on the sequence in M that is used to approximate x. Suppose that (x_n) and (x'_n) are any two sequences in M that converge to x. Then

$$\|x_n - x'_n\| \leq \|x_n - x\| + \|x - x'_n\|,$$

and, taking the limit of this equation as $n \to \infty$, we see that

$$\lim_{n\to\infty} \|x_n - x'_n\| = 0.$$

It follows that

$$\|Tx_n - Tx'_n\| \leq \|T\|\,\|x_n - x'_n\| \to 0 \qquad \text{as } n \to \infty.$$

Hence, (Tx_n) and (Tx'_n) converge to the same limit.

The map $\overline{T}$ is an extension of T, meaning that $\overline{T}x = Tx$, for all $x \in M$, because if $x \in M$, we can use the constant sequence with $x_n = x$ for all n to define $\overline{T}x$. The linearity of $\overline{T}$ follows from the linearity of T.

The fact that $\overline{T}$ is bounded follows from the inequality

$$\left\|\overline{T}x\right\| = \lim_{n\to\infty} \|Tx_n\| \leq \lim_{n\to\infty} \|T\|\,\|x_n\| = \|T\|\,\|x\|.$$

It also follows that $\left\|\overline{T}\right\| \leq \|T\|$. Since $\overline{T}x = Tx$ for $x \in M$, we have $\left\|\overline{T}\right\| = \|T\|$.

Finally, we show that $\overline{T}$ is the unique bounded linear map from X to Y that coincides with T on M. Suppose that $\widetilde{T}$ is another such map, and let x be any point in X, We choose a sequence (x_n) in M that converges to x. Then, using the continuity of $\widetilde{T}$, the fact that $\widetilde{T}$ is an extension of T, and the definition of $\overline{T}$, we see that

$$\widetilde{T}x = \lim_{n\to\infty} \widetilde{T}x_n = \lim_{n\to\infty} Tx_n = \overline{T}x.$$

□

We can use linear maps to define various notions of equivalence between normed linear spaces.

Definition 5.20 Two linear spaces X, Y are *linearly isomorphic* if there is a one-to-one, onto linear map $T : X \to Y$. If X and Y are normed linear spaces and T, T^{-1} are bounded linear maps, then X and Y are *topologically isomorphic*. If T also preserves norms, meaning that $\|Tx\| = \|x\|$ for all $x \in X$, then X, Y are *isometrically isomorphic*.

When we say that two normed linear spaces are "isomorphic" we will usually mean that they are topologically isomorphic. We are often interested in the case when we have two different norms defined on the same space, and we would like to know if the norms define the same topologies.

Definition 5.21 Let X be a linear space. Two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ on X are *equivalent* if there are constants $c > 0$ and $C > 0$ such that

$$c\|x\|_1 \le \|x\|_2 \le C\|x\|_1 \qquad \text{for all } x \in X. \tag{5.10}$$

Theorem 5.22 Two norms on a linear space generate the same topology if and only if they are equivalent.

Proof. Let $\|\cdot\|_1$ and $\|\cdot\|_2$ be two norms on a linear space X. We consider the identity map

$$I : (X, \|\cdot\|_1) \to (X, \|\cdot\|_2).$$

From Corollary 4.20, the topologies generated by the two norms are the same if and only if I and I^{-1} are continuous. Since I is linear, it is continuous if and only if it is bounded. The boundedness of the identity map and its inverse is equivalent to the existence of constants c and C such that (5.10) holds. □

Geometrically, two norms are equivalent if the unit ball of either one of the norms is contained in a ball of finite radius of the other norm.

We end this section by stating, without proof, a fundamental fact concerning linear operators on Banach spaces.

Theorem 5.23 (Open mapping) Suppose that $T : X \to Y$ is a one-to-one, onto bounded linear map between Banach spaces X, Y. Then $T^{-1} : Y \to X$ is bounded.

This theorem states that the existence of the inverse of a continuous linear map between Banach spaces implies its continuity. Contrast this result with Example 4.9.

5.3 The kernel and range of a linear map

The kernel and range are two important linear subspaces associated with a linear map.

Definition 5.24 Let $T : X \to Y$ be a linear map between linear spaces X, Y. The *null space* or *kernel* of T, denoted by $\ker T$, is the subset of X defined by

$$\ker T = \{x \in X \mid Tx = 0\} .$$

The *range* of T, denoted by $\operatorname{ran} T$, is the subset of Y defined by

$$\operatorname{ran} T = \{y \in Y \mid \text{there exists } x \in X \text{ such that } Tx = y\} .$$

The word "kernel" is also used in a completely different sense to refer to the kernel of an integral operator. A map $T : X \to Y$ is one-to-one if and only if $\ker T = \{0\}$, and it is onto if and only if $\operatorname{ran} T = Y$.

Theorem 5.25 Suppose that $T : X \to Y$ is a linear map between linear spaces X, Y. The kernel of T is a linear subspace of X, and the range of T is a linear subspace of Y. If X and Y are normed linear spaces and T is bounded, then the kernel of T is a closed linear subspace.

Proof. If $x_1, x_2 \in \ker T$ and $\lambda_1, \lambda_2 \in \mathbb{R}$ (or $\mathbb{C}$), then the linearity of T implies that

$$T\left(\lambda_1 x_1 + \lambda_2 x_2\right) = \lambda_1 T x_1 + \lambda_2 T x_2 = 0,$$

so $\lambda_1 x_1 + \lambda_2 x_2 \in \ker T$. Therefore, $\ker T$ is a linear subspace. If $y_1, y_2 \in \operatorname{ran} T$, then there are $x_1, x_2 \in X$ such that $Tx_1 = y_1$ and $Tx_2 = y_2$. Hence

$$T(\lambda_1 x_1 + \lambda_2 x_2) = \lambda_1 T x_1 + \lambda_2 T x_2 = \lambda_1 y_1 + \lambda_2 y_2,$$

so $\lambda_1 y_1 + \lambda_2 y_2 \in \operatorname{ran} T$. Therefore, $\operatorname{ran} T$ is a linear subspace.

Now suppose that X and Y are normed spaces and T is bounded. If (x_n) is a sequence of elements in $\ker T$ with $x_n \to x$ in X, then the continuity of T implies that

$$Tx = T\left(\lim_{n\to\infty} x_n\right) = \lim_{n\to\infty} Tx_n = 0,$$

so $x \in \ker T$, and $\ker T$ is closed. □

The *nullity* of T is the dimension of the kernel of T, and the *rank* of T is the dimension of the range of T. We now consider some examples.

Example 5.26 The right shift operator S on $\ell^\infty(\mathbb{N})$ is defined by

$$S(x_1, x_2, x_3, \ldots) = (0, x_1, x_2, \ldots),$$

and the left shift operator T by

$$T(x_1, x_2, x_3, \ldots) = (x_2, x_3, x_4, \ldots).$$

These maps have norm one. Their matrices are the infinite-dimensional Jordan blocks,

$$[S] = \begin{pmatrix} 0 & 0 & 0 & \ldots \\ 1 & 0 & 0 & \ldots \\ 0 & 1 & 0 & \ldots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \quad [T] = \begin{pmatrix} 0 & 1 & 0 & \ldots \\ 0 & 0 & 1 & \ldots \\ 0 & 0 & 0 & \ldots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

The kernel of S is $\{0\}$ and the range of S is the subspace

$$\operatorname{ran} S = \{(0, x_2, x_3, \ldots) \in \ell^\infty(\mathbb{N})\}.$$

The range of T is the whole space $\ell^\infty(\mathbb{N})$, and the kernel of T is the one-dimensional subspace

$$\ker T = \{(x_1, 0, 0, \ldots) \mid x_1 \in \mathbb{R}\}.$$

The operator S is one-to-one but not onto, and T is onto but not one-to-one. This cannot happen for linear maps $T : X \to X$ on a finite-dimensional space X, such as $X = \mathbb{R}^n$. In that case, $\ker T = \{0\}$ if and only if $\operatorname{ran} T = X$.

Example 5.27 An integral operator $K : C([0,1]) \to C([0,1])$

$$Kf(x) = \int_0^1 k(x,y) f(y)\, dy$$

is said to be *degenerate* if $k(x,y)$ is a finite sum of separated terms of the form

$$k(x,y) = \sum_{i=1}^n \varphi_i(x)\psi_i(y),$$

where $\varphi_i, \psi_i : [0,1] \to \mathbb{R}$ are continuous functions. We may assume without loss of generality that $\{\varphi_1, \ldots, \varphi_n\}$ and $\{\psi_1, \ldots, \psi_n\}$ are linearly independent. The range of K is the finite-dimensional subspace spanned by $\{\varphi_1, \varphi_2, \ldots, \varphi_n\}$, and the kernel of K is the subspace of functions $f \in C([0,1])$ such that

$$\int_0^1 f(y)\psi_i(y)\, dy = 0 \qquad \text{for } i = 1, \ldots, n.$$

Both the range and kernel are closed linear subspaces of $C([0,1])$.

Example 5.28 Let $X = C([0,1])$ with the maximum norm. We define the integral operator $K : X \to X$ by

$$Kf(x) = \int_0^x f(y)\,dy. \tag{5.11}$$

An integral operator like this one, with a variable range of integration, is called a *Volterra integral operator*. Then K is bounded, with $\|K\| \le 1$, since

$$\|Kf\| \le \sup_{0\le x\le 1} \int_0^x |f(y)|\,dy \le \int_0^1 |f(y)|\,dy \le \|f\|.$$

In fact, $\|K\| = 1$, since $K(1) = x$ and $\|x\| = \|1\|$. The range of K is the set of continuously differentiable functions on $[0,1]$ that vanish at $x = 0$. This is a linear subspace of $C([0,1])$ but it is not closed. The lack of closure of the range of K is due to the "smoothing" effect of K, which maps continuous functions to differentiable functions. The problem of inverting integral operators with similar properties arises in a number of inverse problems, where one wants to reconstruct a source distribution from remotely sensed data. Such problems are ill-posed and require special treatment.

Example 5.29 Consider the operator $T = I + K$ on $C([0,1])$, where K is defined in (5.11), which is a perturbation of the identity operator by K. The range of T is the whole space $C([0,1])$, and is therefore closed. To prove this statement, we observe that $g = Tf$ if and only if

$$f(x) + \int_0^x f(y)\,dy = g(x).$$

Writing $F(x) = \int_0^x f(y)\,dy$, we have $F' = f$ and

$$F' + F = g, \qquad F(0) = 0.$$

The solution of this initial value problem is

$$F(x) = \int_0^x e^{-(x-y)} g(y)\,dy.$$

Differentiating this expression with respect to x, we find that f is given by

$$f(x) = g(x) - \int_0^x e^{-(x-y)} g(y)\,dy.$$

Thus, the operator $T = I + K$ is invertible on $C([0,1])$ and

$$(I+K)^{-1} = I - L,$$

where L is the Volterra integral operator

$$Lg(x) = \int_0^x e^{-(x-y)} g(y)\,dy.$$

The following result provides a useful way to show that an operator T has closed range. It states that T has closed range if one can estimate the norm of the solution x of the equation $Tx = y$ in terms of the norm of the right-hand side y. In that case, it is often possible to deduce the existence of solutions (see Theorem 8.18).

Proposition 5.30 Let $T : X \to Y$ be a bounded linear map between Banach spaces X, Y. The following statements are equivalent:

(a) there is a constant $c > 0$ such that

$$c\|x\| \leq \|Tx\| \qquad \text{for all } x \in X;$$

(b) T has closed range, and the only solution of the equation $Tx = 0$ is $x = 0$.

Proof. First, suppose that T satisfies (a). Then $Tx = 0$ implies that $\|x\| = 0$, so $x = 0$. To show that $\operatorname{ran} T$ is closed, suppose that (y_n) is a convergent sequence in $\operatorname{ran} T$, with $y_n \to y \in Y$. Then there is a sequence (x_n) in X such that $Tx_n = y_n$. The sequence (x_n) is Cauchy, since (y_n) is Cauchy and

$$\|x_n - x_m\| \leq \frac{1}{c}\|T(x_n - x_m)\| = \frac{1}{c}\|y_n - y_m\|.$$

Hence, since X is complete, we have $x_n \to x$ for some $x \in X$. Since T is bounded, we have

$$Tx = \lim_{n\to\infty} Tx_n = \lim_{n\to\infty} y_n = y,$$

so $y \in \operatorname{ran} T$, and $\operatorname{ran} T$ is closed.

Conversely, suppose that T satisfies (b). Since $\operatorname{ran} T$ is closed, it is a Banach space. Since $T : X \to Y$ is one-to-one, the operator $T : X \to \operatorname{ran} T$ is a one-to-one, onto map between Banach spaces. The open mapping theorem, Theorem 5.23, implies that $T^{-1} : \operatorname{ran} T \to X$ is bounded, and hence that there is a constant $C > 0$ such that

$$\left\|T^{-1}y\right\| \leq C\|y\| \qquad \text{for all } y \in \operatorname{ran} T.$$

Setting $y = Tx$, we see that $c\|x\| \leq \|Tx\|$ for all $x \in X$, where $c = 1/C$. □

Example 5.31 Consider the Volterra integral operator $K : C([0, 1]) \to C([0, 1])$ defined in (5.11). Then

$$K[\cos n\pi x] = \int_0^x \cos n\pi y \, dy = \frac{\sin n\pi x}{n\pi}.$$

We have $\|\cos n\pi x\| = 1$ for every $n \in \mathbb{N}$, but $\|K[\cos n\pi x]\| \to 0$ as $n \to \infty$. Thus, it is not possible to estimate $\|f\|$ in terms of $\|Kf\|$, consistent with the fact that the range of K is not closed.

5.4 Finite-dimensional Banach spaces

In this section, we prove that every finite-dimensional (real or complex) normed linear space is a Banach space, that every linear operator on a finite-dimensional space is continuous, and that all norms on a finite-dimensional space are equivalent. None of these statements is true for infinite-dimensional linear spaces. As a result, topological considerations can often be neglected when dealing with finite-dimensional spaces but are of crucial importance when dealing with infinite-dimensional spaces.

We begin by proving that the components of a vector with respect to any basis of a finite-dimensional space can be bounded by the norm of the vector.

Lemma 5.32 Let X be a finite-dimensional normed linear space with norm $\|\cdot\|$, and $\{e_1, e_2, \ldots, e_n\}$ any basis of X. There are constants $m > 0$ and $M > 0$ such that if $x = \sum_{i=1}^n x_i e_i$, then

$$m \sum_{i=1}^{n} |x_i| \leq \|x\| \leq M \sum_{i=1}^{n} |x_i| . \tag{5.12}$$

Proof. By the homogeneity of the norm, it suffices to prove (5.12) for $x \in X$ such that $\sum_{i=1}^n |x_i| = 1$. The "cube"

$$C = \left\{ (x_1, \ldots, x_n) \in \mathbb{R}^n \,\middle|\, \sum_{i=1}^{n} |x_i| = 1 \right\}$$

is a closed, bounded subset of $\mathbb{R}^n$, and is therefore compact by the Heine-Borel theorem. We define a function $f : C \to X$ by

$$f\left((x_1, \ldots, x_n)\right) = \sum_{i=1}^{n} x_i e_i.$$

For $(x_1, \ldots, x_n) \in \mathbb{R}^n$ and $(y_1, \ldots, y_n) \in \mathbb{R}^n$, we have

$$\|f\left((x_1, \ldots, x_n)\right) - f\left((y_1, \ldots, y_n)\right)\| \leq \sum_{i=1}^{n} |x_i - y_i| \|e_i\|,$$

so f is continuous. Therefore, since $\|\cdot\| : X \to \mathbb{R}$ is continuous, the map

$$(x_1, \ldots, x_n) \mapsto \|f\left((x_1, \ldots, x_n)\right)\|$$

is continuous. Theorem 1.68 implies that $\|f\|$ is bounded on C and attains its infimum and supremum. Denoting the minimum by $m \geq 0$ and the maximum by $M \geq m$, we obtain (5.12). Let $(\overline{x}_1, \ldots, \overline{x}_n)$ be a point in C where $\|f\|$ attains its minimum, meaning that

$$\|\overline{x}_1 e_1 + \ldots + \overline{x}_n e_n\| = m.$$

The linear independence of the basis vectors $\{e_1, \ldots, e_n\}$ implies that $m \neq 0$, so $m > 0$. □

This result is not true in an infinite-dimensional space because, if a basis consists of vectors that become "almost" parallel, then the cancellation in linear combinations of basis vectors may lead to a vector having large components but small norm.

Theorem 5.33 Every finite-dimensional normed linear space is a Banach space.

Proof. Suppose that $(x_k)_{k=1}^{\infty}$ is a Cauchy sequence in a finite-dimensional normed linear space X. Let $\{e_1, \ldots, e_n\}$ be a basis of X. We expand x_k as

$$x_k = \sum_{i=1}^{n} x_{i,k} e_i,$$

where $x_{i,k} \in \mathbb{R}$. For $1 \leq i \leq n$, we consider the real sequence of ith components, $(x_{i,k})_{k=1}^{\infty}$. Equation (5.12) implies that

$$|x_{i,j} - x_{i,k}| \leq \frac{1}{m} \|x_j - x_k\|,$$

so $(x_{i,k})_{k=1}^{\infty}$ is Cauchy. Since $\mathbb{R}$ is complete, there is a $y_i \in \mathbb{R}$, such that

$$\lim_{k \to \infty} x_{i,k} = y_i.$$

We define $y \in X$ by

$$y = \sum_{i=1}^{k} y_i e_i.$$

Then, from (5.12),

$$\|x_k - y\| \leq M \sum_{i=1}^{n} |x_{i,k} - y_i| \, \|e_i\|,$$

and hence $x_k \to y$ as $k \to \infty$. Thus, every Cauchy sequence in X converges, and X is complete. □

Since a complete space is closed, we have the following corollary.

Corollary 5.34 Every finite-dimensional linear subspace of a normed linear space is closed.

In Section 5.2, we proved explicitly the boundedness of linear maps on finite-dimensional linear spaces with respect to certain norms. In fact, linear maps on finite-dimensional spaces are always bounded.

Theorem 5.35 Every linear operator on a finite-dimensional linear space is bounded.

Proof. Suppose that $A : X \to Y$ is a linear map and X is finite dimensional. Let $\{e_1, \dots, e_n\}$ be a basis of X. If $x = \sum_{i=1}^n x_i e_i \in X$, then (5.12) implies that

$$\|Ax\| \le \sum_{i=1}^n |x_i| \, \|Ae_i\| \le \max_{1\le i\le n} \{\|Ae_i\|\} \sum_{i=1}^n |x_i| \le \frac{1}{m} \max_{1\le i\le n} \{\|Ae_i\|\} \, \|x\|,$$

so A is bounded. □

Finally, we show that although there are many different norms on a finite-dimensional linear space they all lead to the same topology and the same notion of convergence. This fact follows from Theorem 5.22 and the next result.

Theorem 5.36 Any two norms on a finite-dimensional space are equivalent.

Proof. Let $\|\cdot\|_1$ and $\|\cdot\|_2$ be two norms on a finite-dimensional space X. We choose a basis $\{e_1, e_2, \dots, e_n\}$ of X. Then Lemma 5.32 implies that there are strictly positive constants m_1, m_2, M_1, M_2 such that if $x = \sum_{i=1}^n x_i e_i$, then

$$m_1 \sum_{i=1}^n |x_i| \le \|x\|_1 \le M_1 \sum_{i=1}^n |x_i| \,,$$
$$m_2 \sum_{i=1}^n |x_i| \le \|x\|_2 \le M_2 \sum_{i=1}^n |x_i| \,.$$

Equation (5.10) then follows with $c = m_2/M_1$ and $C = M_2/m_1$. □

5.5 Convergence of bounded operators

The set $\mathcal{B}(X,Y)$ of bounded linear maps from a normed linear space X to a normed linear space Y is a linear space with respect to the natural pointwise definitions of vector addition and scalar multiplication:

$$(S+T)x = Sx + Tx, \qquad (\lambda T)x = \lambda(Tx).$$

It is straightforward to check that the operator norm in Definition 5.12,

$$\|T\| = \sup_{x\ne 0} \frac{\|Tx\|}{\|x\|},$$

defines a norm on $\mathcal{B}(X,Y)$, so that $\mathcal{B}(X,Y)$ is a normed linear space.

The composition of two linear maps is linear, and the following theorem states that the composition of two bounded linear maps is bounded.

Theorem 5.37 Let X, Y, and Z be normed linear spaces. If $T \in \mathcal{B}(X,Y)$ and $S \in \mathcal{B}(Y,Z)$, then $ST \in \mathcal{B}(X,Z)$, and

$$\|ST\| \le \|S\| \, \|T\|. \tag{5.13}$$

Proof. For all $x \in X$ we have

$$\|STx\| \leq \|S\| \, \|Tx\| \leq \|S\| \, \|T\| \, \|x\|.$$

□

For example, if $T \in \mathcal{B}(X)$, then $T^n \in \mathcal{B}(X)$ and $\|T^n\| \leq \|T\|^n$. It may well happen that we have strict inequality in (5.13).

Example 5.38 Consider the linear maps A, B on $\mathbb{R}^2$ with matrices

$$A = \begin{pmatrix} \lambda & 0 \\ 0 & 0 \end{pmatrix}, \qquad B = \begin{pmatrix} 0 & 0 \\ 0 & \mu \end{pmatrix}.$$

These matrices have the Euclidean (or sum, or maximum) norms $\|A\| = |\lambda|$ and $\|B\| = |\mu|$, but $\|AB\| = 0$.

A linear space with a product defined on it is called an *algebra*. The composition of maps defines a product on the space $\mathcal{B}(X)$ of bounded linear maps on X into itself, so $\mathcal{B}(X)$ is an algebra. The algebra is associative, meaning that $(RS)T = R(ST)$, but is not commutative, since in general ST is not equal to TS. If $S, T \in \mathcal{B}(X)$, we define the *commutator* $[S, T] \in \mathcal{B}(X)$ of S and T by

$$[S, T] = ST - TS.$$

If $ST = TS$, or equivalently if $[S, T] = 0$, then we say that S and T *commute*.

The convergence of operators in $\mathcal{B}(X, Y)$ with respect to the operator norm is called uniform convergence.

Definition 5.39 If (T_n) is a sequence of operators in $\mathcal{B}(X, Y)$ and

$$\lim_{n \to \infty} \|T_n - T\| = 0$$

for some $T \in \mathcal{B}(X, Y)$, then we say that T_n *converges uniformly* to T, or that T_n converges to T in the uniform, or operator norm, topology on $\mathcal{B}(X, Y)$.

Example 5.40 Let $X = C([0, 1])$ equipped with the supremum norm. For $k_n(x, y)$ is a real-valued continuous function on $[0, 1] \times [0, 1]$, we define $K_n \in \mathcal{B}(X)$ by

$$K_n f(x) = \int_0^1 k_n(x, y) f(y) \, dy. \tag{5.14}$$

Then $K_n \to 0$ uniformly as $n \to \infty$ if

$$\|K_n\| = \max_{x \in [0,1]} \left\{ \int_0^1 |k_n(x, y)| \, dy \right\} \to 0 \qquad \text{as } n \to \infty. \tag{5.15}$$

An example of functions k_n satisfying (5.15) is $k_n(x, y) = xy^n$.

A basic fact about a space of bounded linear operators that take values in a Banach space is that it is itself a Banach space.

Theorem 5.41 If X is a normed linear space and Y is a Banach space, then $\mathcal{B}(X,Y)$ is a Banach space with respect to the operator norm.

Proof. We have to prove that $\mathcal{B}(X,Y)$ is complete. Let (T_n) be a Cauchy sequence in $\mathcal{B}(X,Y)$. For each $x \in X$, we have

$$\|T_n x - T_m x\| \leq \|T_n - T_m\| \, \|x\|,$$

which shows that $(T_n x)$ is a Cauchy sequence in Y. Since Y is complete, there is a $y \in Y$ such that $T_n x \to y$. It is straightforward to check that $Tx = y$ defines a linear map $T : X \to Y$. We show that T is bounded. For any $\epsilon > 0$, let N_ϵ be such that $\|T_n - T_m\| < \epsilon/2$ for all $n, m \geq N_\epsilon$. Take $n \geq N_\epsilon$. Then for each $x \in X$, there is an $m(x) \geq N_\epsilon$ such that $\|T_{m(x)}x - Tx\| \leq \epsilon/2$. If $\|x\| = 1$, we have

$$\|T_n x - Tx\| \leq \|T_n x - T_{m(x)}x\| + \|T_{m(x)}x - Tx\| \leq \epsilon. \tag{5.16}$$

It follows that if $n \geq N_\epsilon$, then

$$\|Tx\| \leq \|T_n x\| + \|Tx - T_n x\| \leq \|T_n\| + \epsilon$$

for all x with $\|x\| = 1$, so T is bounded. Finally, from (5.16) it follows that $\lim_{n\to\infty} \|T_n - T\| = 0$. Hence, $T_n \to T$ in the uniform norm. □

A particularly important class of bounded operators is the class of compact operators.

Definition 5.42 A linear operator $T : X \to Y$ is *compact* if $T(B)$ is a precompact subset of Y for every bounded subset B of X.

An equivalent formulation is that T is compact if and only if every bounded sequence (x_n) in X has a subsequence (x_{n_k}) such that (Tx_{n_k}) converges in Y. We do not require that the range of T be closed, so $T(B)$ need not be compact even if B is a closed bounded set. We leave the proof of the following properties of compact operators as an exercise.

Proposition 5.43 Let X, Y, Z be Banach spaces. (a) If $S, T \in \mathcal{B}(X,Y)$ are compact, then any linear combination of S and T is compact. (b) If (T_n) is a sequence of compact operators in $\mathcal{B}(X,Y)$ converging uniformly to T, then T is compact. (c) If $T \in \mathcal{B}(X,Y)$ has finite-dimensional range, then T is compact. (d) Let $S \in \mathcal{B}(X,Y)$, $T \in \mathcal{B}(Y,Z)$. If S is bounded and T is compact, or S is compact and T is bounded, then $TS \in \mathcal{B}(X,Z)$ is compact.

It follows from parts (a)–(b) of this proposition that the space $\mathcal{K}(X,Y)$ of compact linear operators from X to Y is a closed linear subspace of $\mathcal{B}(X,Y)$. Part (d) implies that $\mathcal{K}(X)$ is a two-sided ideal of $\mathcal{B}(X)$, meaning that if $K \in \mathcal{K}(X)$, then $AK \in \mathcal{K}(X)$ and $KA \in \mathcal{K}(X)$ for all $A \in \mathcal{B}(X)$.

From parts (b)–(c), an operator that is the uniform limit of operators with finite rank, that is with finite-dimensional range, is compact. The converse is also true for compact operators on many Banach spaces, including all Hilbert spaces, although there exist separable Banach spaces on which some compact operators cannot be approximated by finite-rank operators. As a result, compact operators on infinite-dimensional spaces behave in many respects like operators on finite-dimensional spaces. We will discuss compact operators on a Hilbert space in greater detail in Chapter 9.

Another type of convergence of linear maps is called strong convergence.

Definition 5.44 A sequence (T_n) in $\mathcal{B}(X, Y)$ *converges strongly* if

$$\lim_{n\to\infty} T_n x = Tx \qquad \text{for every } x \in X.$$

Thus, strong convergence of linear maps is convergence of their pointwise values with respect to the norm on Y. The terminology here is a little inconsistent: strong and norm convergence mean the same thing for vectors in a Banach space, but different things for operators on a Banach space. The associated strong topology on $\mathcal{B}(X, Y)$ is distinct from the uniform norm topology whenever X is infinite-dimensional, and is not derived from a norm. We leave the proof of the following theorem as an exercise.

Theorem 5.45 If $T_n \to T$ uniformly, then $T_n \to T$ strongly.

The following examples show that strong convergence does not imply uniform convergence.

Example 5.46 Let $X = \ell^2(\mathbb{N})$, and define the projection $P_n : X \to X$ by

$$P_n(x_1, x_2, \ldots, x_n, x_{n+1}, x_{n+2}, \ldots) = (x_1, x_2, \ldots, x_n, 0, 0, \ldots).$$

Then $\|P_n - P_m\| = 1$ for $n \neq m$, so (P_n) does not converge uniformly. Nevertheless, if $x \in \ell^2(\mathbb{N})$ is any fixed vector, we have $P_n x \to x$ as $n \to \infty$. Thus, $P_n \to I$ strongly.

Example 5.47 Let $X = C([0, 1])$, and consider the sequence of continuous linear functionals $K_n : X \to \mathbb{R}$, given by

$$K_n f = \int_0^1 \sin(n\pi x)\, f(x)\, dx.$$

If p is a polynomial, then an integration by parts implies that

$$K_n p = \frac{p(0) - \cos(n\pi)p(1)}{n\pi} + \frac{1}{n\pi} \int_0^1 \cos(n\pi x)\, p'(x)\, dx.$$

Hence, $K_n p \to 0$ as $n \to \infty$. If $f \in C([0,1])$, then by Theorem 2.9 for any $\epsilon > 0$ there is a polynomial p such that $\|f - p\| < \epsilon/2$, and there is an N such that $|K_n p| < \epsilon/2$ for $n \geq N$. Since $\|K_n\| \leq 1$ for all n, it follows that

$$|K_n f| \leq \|K_n\| \, \|f - p\| + |K_n p| < \epsilon$$

when $n \geq N$. Thus, $K_n f \to 0$ as $n \to \infty$ for every $f \in C([0,1])$. This result is a special case of the Riemann-Lebesgue lemma, which we prove in Theorem 11.34. On the other hand, if $f_n(x) = \sin(n\pi x)$, then $\|f_n\| = 1$ and $\|K_n f_n\| = 1/2$, which implies that $\|K_n\| \geq 1/2$. (In fact, $\|K_n\| = 2/\pi$ for each n.) Hence, $K_n \to 0$ strongly, but not uniformly.

A third type of convergence of operators, *weak convergence*, may be defined using the notion of weak convergence in a Banach space, given in Definition 5.59 below. We say that T_n *converges weakly* to T in $\mathcal{B}(X, Y)$ if the pointwise values $T_n x$ converge weakly to Tx in Y. We will not consider the weak convergence of operators in this book.

We end this section with two applications of operator convergence. First we define the *exponential* of an operator, and use it to solve a linear evolution equation. If $A : X \to X$ is a bounded linear operator on a Banach space X, then, by analogy with the power series expansion of e^a, we define

$$e^A = I + A + \frac{1}{2!}A^2 + \frac{1}{3!}A^3 + \ldots + \frac{1}{n!}A^n + \ldots. \tag{5.17}$$

A comparison with the convergent real series

$$e^{\|A\|} = 1 + \|A\| + \frac{1}{2!}\|A\|^2 + \frac{1}{3!}\|A\|^3 + \ldots + \frac{1}{n!}\|A\|^n + \ldots,$$

implies that the series on the right hand side of (5.17) is absolutely convergent in $\mathcal{B}(X)$, and hence norm convergent. It also follows that

$$\left\|e^A\right\| \leq e^{\|A\|}.$$

If A and B commute, then multiplication and rearrangement of the series for the exponentials implies that

$$e^A e^B = e^{A+B}.$$

The solution of the initial value problem for the linear, scalar ODE $x_t = ax$ with $x(0) = x_0$ is given by $x(t) = x_0 e^{at}$. This result generalizes to a linear system,

$$x_t = Ax, \qquad x(0) = x_0, \tag{5.18}$$

where $x : \mathbb{R} \to X$, with X a Banach space, and $A : X \to X$ is a bounded linear operator on X. The solution of (5.18) is given by

$$x(t) = e^{tA} x_0.$$

This is a solution because

$$\frac{d}{dt}e^{tA} = Ae^{tA},$$

where the derivative is given by the uniformly convergent limit,

$$\begin{aligned}
\frac{d}{dt}e^{tA} &= \lim_{h\to 0}\left(\frac{e^{A(t+h)} - e^{tA}}{h}\right)\\
&= e^{tA}\lim_{h\to 0}\left(\frac{e^{Ah} - I}{h}\right)\\
&= Ae^{tA}\lim_{h\to 0}\sum_{n=0}^{\infty}\frac{1}{(n+1)!}A^n h^n\\
&= Ae^{tA}.
\end{aligned}$$

An important application of this result is to linear systems of ODEs when $x(t) \in \mathbb{R}^n$ and A is an $n \times n$ matrix, but it also applies to linear equations on infinite-dimensional spaces.

Example 5.48 Suppose that $k : [0,1] \times [0,1] \to \mathbb{R}$ is a continuous function, and $K : C([0,1]) \to C([0,1])$ is the integral operator

$$Ku(x) = \int_0^1 k(x,y)u(y)\,dy.$$

The solution of the initial value problem

$$u_t(x,t) + \lambda u(x,t) = \int_0^1 k(x,y)u(y,t)\,dy, \qquad u(x,0) = u_0(x),$$

with $u(\cdot,t) \in C([0,1])$, is $u = e^{(K-\lambda I)t}u_0$.

The one-parameter family of operators $T(t) = e^{tA}$ is called the *flow* of the evolution equation (5.18). The operator $T(t)$ maps the solution at time 0 to the solution at time t. We leave the proof of the following properties of the flow as an exercise.

Theorem 5.49 If $A : X \to X$ is a bounded linear operator and $T(t) = e^{tA}$ for $t \in \mathbb{R}$, then:

(a) $T(0) = I$;
(b) $T(s)T(t) = T(s+t)$ for $s, t \in \mathbb{R}$;
(c) $T(t) \to I$ uniformly as $t \to 0$.

A family of bounded linear operators $\{T(t) \mid t \in \mathbb{R}\}$ that satisfies the properties (a)–(c) in this theorem is called a one-parameter *uniformly continuous group*. Properties (a)–(b) imply that the operators form a commutative group under composition, while (c) states that $T : \mathbb{R} \to \mathcal{B}(X)$ is continuous with respect to the uniform, or norm, topology on $\mathcal{B}(X)$ at $t = 0$. The group property implies that T is uniformly continuous on $\mathbb{R}$, meaning that $\|T(t) - T(t_0)\| \to 0$ as $t \to t_0$ for any $t_0 \in \mathbb{R}$.

Any one-parameter uniformly continuous group of operators can be written as $T(t) = e^{tA}$ for a suitable operator A, called the *generator* of the group. The generator A may be recovered from the operators $T(t)$ by

$$A = \lim_{t \to 0} \left(\frac{T(t) - I}{t} \right). \tag{5.19}$$

Many linear partial differential equations can be written as evolution equations of the form (5.18) in which A is an unbounded operator. Under suitable conditions on A, there exist solution operators $T(t)$, which may be defined only for $t \geq 0$, and which are strongly continuous functions of t, rather than uniformly continuous. The solution operators are then said to form a C_0-*semigroup*. For an example, see the discussion of the heat equation in Section 7.3.

As a second application of operator convergence, we consider the convergence of approximation schemes. Suppose we want to solve an equation of the form

$$Au = f, \tag{5.20}$$

where $A : X \to Y$ is a nonsingular linear operator between Banach spaces and $f \in Y$ is given. Suppose we can approximate (5.20) by an equation

$$A_\epsilon u_\epsilon = f_\epsilon, \tag{5.21}$$

whose solution u_ϵ can be computed more easily. We assume that $A_\epsilon : X \to Y$ is a nonsingular linear operator with a bounded inverse. We call the family of equations (5.21) an *approximation scheme* for (5.20). For instance, if (5.20) is a differential equation, then (5.21) may be obtained by a finite difference or finite element approximation, where ϵ is a grid spacing. One complication is that a numerical approximation A_ϵ may act on a different space X_ϵ than the space X. For simplicity, we suppose that the approximations A_ϵ may be defined on the same space as A. The primary requirement of an approximation scheme is that it is convergent.

Definition 5.50 The approximation scheme (5.21) is *convergent* to (5.20) if $u_\epsilon \to u$ as $\epsilon \to 0$ whenever $f_\epsilon \to f$.

We make precise the idea that A_ϵ approximates A in the following definition of consistency.

Definition 5.51 The approximation scheme (5.21) is *consistent* with (5.20) if $A_\epsilon v \to Av$ as $\epsilon \to 0$ for each $v \in X$.

In other words, the approximation scheme is consistent if A_ϵ converges strongly to A as $\epsilon \to 0$. Consistency on its own is not sufficient to guarantee convergence. We also need a second property called *stability*.

Definition 5.52 The approximation scheme (5.21) is *stable* if there is a constant M, independent of ϵ, such that

$$\|A_\epsilon^{-1}\| \leq M.$$

Consistency and convergence relate the operators A_ϵ to A, while stability is a property of the approximate operators A_ϵ alone. Stability plays a crucial role in convergence, because it prevents the amplification of errors in the approximate solutions as $\epsilon \to 0$.

Theorem 5.53 (Lax equivalence) An consistent approximation scheme is convergent if and only if it is stable.

Proof. First, we prove that a stable scheme is convergent. If $Au = f$ and $A_\epsilon u_\epsilon = f_\epsilon$, then

$$u - u_\epsilon = A_\epsilon^{-1}\left(A_\epsilon u - Au + f - f_\epsilon\right).$$

Taking the norm of this equation, using the definition of the operator norm, and the triangle inequality, we find that

$$\|u - u_\epsilon\| \leq \|A_\epsilon^{-1}\|\left(\|A_\epsilon u - Au\| + \|f - f_\epsilon\|\right). \tag{5.22}$$

If the scheme is stable, then

$$\|u - u_\epsilon\| \leq M\left(\|A_\epsilon u - Au\| + \|f - f_\epsilon\|\right),$$

and if the scheme is consistent, then $A_\epsilon u \to Au$ as $\epsilon \to 0$. It follows that $u_\epsilon \to u$ if $f_\epsilon \to f$, and the scheme is convergent.

Conversely, we prove that a convergent scheme is stable. For any $f \in Y$, let $u_\epsilon = A_\epsilon^{-1} f$. Then, since the scheme is convergent, we have $u_\epsilon \to u$ as $\epsilon \to 0$, where $u = A^{-1} f$, so that u_ϵ is bounded. Thus, there exists a constant M_f, independent of ϵ, such that $\|A_\epsilon^{-1} f\| \leq M_f$. The uniform boundedness theorem, which we do not prove here, then implies that there exists a constant M such that $\|A_\epsilon^{-1}\| \leq M$, so the scheme is stable. □

An analogous result holds for linear evolution equations of the form (5.18) (see Strikwerder [53], for example). There is, however, no general criterion for the convergence of approximation schemes for nonlinear equations.

5.6 Dual spaces

The dual space of a linear space consists of the scalar-valued linear maps on the space. Duality methods play a crucial role in many parts of analysis. In this section, we consider real linear spaces for definiteness, but all the results hold for complex linear spaces.

Definition 5.54 A scalar-valued linear map from a linear space X to $\mathbb{R}$ is called a *linear functional* or *linear form* on X. The space of linear functionals on X is called the *algebraic dual space* of X, and the space of continuous linear functionals on X is called the *topological dual space* of X.

In terms of the notation in Definition 5.12, the algebraic dual space of X is $\mathcal{L}(X, \mathbb{R})$, and the topological dual space is $\mathcal{B}(X, \mathbb{R})$. A linear functional $\varphi : X \to \mathbb{R}$ is bounded if there is a constant M such that

$$|\varphi(x)| \le M\|x\| \qquad \text{for all } x \in X,$$

and then we define $\|\varphi\|$ by

$$\|\varphi\| = \sup_{x\neq 0} \frac{|\varphi(x)|}{\|x\|}. \tag{5.23}$$

If X is infinite dimensional, then $\mathcal{L}(X, \mathbb{R})$ is much larger than $\mathcal{B}(X, \mathbb{R})$, as we illustrate in Example 5.57 below. Somewhat confusingly, both dual spaces are commonly denoted by X^*. We will use X^* to denote the topological dual space of X. Either dual space is itself a linear space under the operations of pointwise addition and scalar multiplication of maps, and the topological dual is a Banach space, since $\mathbb{R}$ is complete.

If X is finite dimensional, then $\mathcal{L}(X, \mathbb{R}) = \mathcal{B}(X, \mathbb{R})$, so there is no need to distinguish between the algebraic and topological dual spaces. Moreover, the dual space X^* of a finite-dimensional space X is linearly isomorphic to X. To show this, we pick a basis $\{e_1, e_2, \ldots, e_n\}$ of X. The map $\omega_i : X \to \mathbb{R}$ defined by

$$\omega_i \left(\sum_{j=1}^{n} x_j e_j \right) = x_i \tag{5.24}$$

is an element of the algebraic dual space X^*. The linearity of ω_i is obvious.

For example, if $X = \mathbb{R}^n$ and

$$e_1 = (1, 0, \ldots, 0), \quad e_2 = (0, 1, \ldots, 0), \ldots, e_n = (0, 0, \ldots, 1),$$

are the coordinate basis vectors, then

$$\omega_i : (x_1, x_2, \ldots, x_n) \mapsto x_i$$

is the map that takes a vector to its ith coordinate.

The action of a general element φ of the dual space $\varphi : X \to \mathbb{R}$ on a vector $x \in X$ is given by a linear combination of the components of x, since

$$\varphi\left(\sum_{i=1}^{n} x_i e_i\right) = \sum_{i=1}^{n} \varphi_i x_i,$$

where $\varphi_i = \varphi(e_i) \in \mathbb{R}$. It follows that, as a map,

$$\varphi = \sum_{i=1}^{n} \varphi_i \omega_i.$$

Thus, $\{\omega_1, \omega_2, \ldots, \omega_n\}$ is a basis of X^*, called the *dual basis* of $\{e_1, e_2, \ldots, e_n\}$, and both X and X^* are linearly isomorphic to $\mathbb{R}^n$. The dual basis has the property that

$$\omega_i(e_j) = \delta_{ij},$$

where δ_{ij} is the *Kronecker delta function*, defined by

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases} \tag{5.25}$$

Although a finite-dimensional space is linearly isomorphic with its dual space, there is no canonical way to identify the space with its dual; there are many isomorphisms, depending on an arbitrary choice of a basis. In the following chapters, we will study Hilbert spaces, and show that the topological dual space of a Hilbert space can be identified with the original space in a natural way through the inner product (see the Riesz representation theorem, Theorem 8.12). The dual of an infinite-dimensional Banach space is, in general, different from the original space.

Example 5.55 In Section 12.8, we will see that for $1 \leq p < \infty$ the dual of $L^p(\Omega)$ is $L^{p'}(\Omega)$, where $1/p + 1/p' = 1$. The Hilbert space $L^2(\Omega)$ is self-dual.

Example 5.56 Consider $X = C([a, b])$. For any $\rho \in L^1([a, b])$, the following formula defines a continuous linear functional φ on X:

$$\varphi(f) = \int_a^b f(x)\rho(x)\, dx. \tag{5.26}$$

Not all continuous linear functionals are of the form (5.26). For example, if $x_0 \in [a, b]$, then the evaluation of f at x_0 is a continuous linear functional. That is, if we define $\delta_{x_0} : C([a, b]) \to \mathbb{R}$ by

$$\delta_{x_0}(f) = f(x_0),$$

then δ_{x_0} is a continuous linear functional on $C([a, b])$. A full description of the dual space of $C([a, b])$ is not so simple: it may be identified with the space of Radon measures on $[a, b]$ (see [12], for example).

One way to obtain a linear functional on a linear space is to start with a linear functional defined on a subspace, extend a Hamel basis of the subspace to a Hamel basis of the whole space and extend the functional to the whole space, by use of linearity and an arbitrary definition of the functional on the additional basis elements. The next example uses this procedure to obtain a discontinuous linear functional on $C([0,1])$.

Example 5.57 Let $M = \{x^n \mid n = 0, 1, 2, \ldots\}$ be the set of monomials in $C([0,1])$. The set M is linearly independent, so it may be extended to a Hamel basis H. Each $f \in C([0,1])$ can be written uniquely as

$$f = c_1 h_1 + \cdots + c_N h_N, \tag{5.27}$$

for suitable basis functions $h_i \in H$ and nonzero scalar coefficients c_i. For each $n = 0, 1, 2, \ldots$, we define $\varphi_n(f)$ by

$$\varphi_n(f) = \begin{cases} c_i & \text{if } h_i = x^n, \\ 0 & \text{otherwise.} \end{cases}$$

Due to the uniqueness of the decomposition in (5.27), the functional φ_n is well-defined. We define a linear functional φ on $C([0,1])$ by

$$\varphi(f) = \sum_{n=1}^{\infty} n\varphi_n(f).$$

For each f, only a finite number of terms in this sum are nonzero, so φ is a well-defined linear functional on $C([0,1])$. The functional is unbounded, since for each $n = 0, 1, 2, \ldots$ we have $\|x^n\| = 1$ and $|\varphi(x^n)| = n$.

A similar construction shows that every infinite-dimensional linear space has discontinuous linear functionals defined on it. On the other hand, Theorem 5.35 implies that all linear functionals on a finite-dimensional linear space are bounded.

It is not obvious that this extension procedure can be used to obtain bounded linear functionals on an infinite-dimensional linear space, or even that there are any nonzero bounded linear functionals at all, because the extension need not be bounded. In fact, it is possible to maintain boundedness of an extension by a suitable choice of its values off the original subspace, as stated in the following version of the Hahn-Banach theorem.

Theorem 5.58 (Hahn-Banach) If Y is a linear subspace of a normed linear space X and $\psi : Y \to \mathbb{R}$ is a bounded linear functional on Y with $\|\psi\| = M$, then there is a bounded linear functional $\varphi : X \to \mathbb{R}$ on X such that φ restricted to Y is equal to ψ and $\|\varphi\| = M$.

We omit the proof here. One consequence of this theorem is that there are enough bounded linear functionals to separate X, meaning that if $\varphi(x) = \varphi(y)$ for all $\varphi \in X^*$, then $x = y$ (see Exercise 5.6).

Since X^* is a Banach space, we can form its dual space X^{**}, called the *bidual* of X. There is no natural way to identify an element of X with an element of the dual X^*, but we can naturally identify an element of X with an element of the bidual X^{**}. If $x \in X$, then we define $F_x \in X^{**}$ by evaluation at x:

$$F_x(\varphi) = \varphi(x) \qquad \text{for every } \varphi \in X^*. \tag{5.28}$$

Thus, we may regard X as a subspace of X^{**}. If all continuous linear functionals on X^* are of the form (5.28), then $X = X^{**}$ under the identification $x \mapsto F_x$, and we say that X is *reflexive*.

Linear functionals may be used to define a notion of convergence that is weaker than norm, or strong, convergence on an infinite-dimensional Banach space.

Definition 5.59 A sequence (x_n) in a Banach space X *converges weakly* to x, denoted by $x_n \rightharpoonup x$ as $n \to \infty$, if

$$\varphi(x_n) \to \varphi(x) \qquad \text{as } n \to \infty,$$

for every bounded linear functional φ in X^*.

If we think of a linear functional $\varphi : X \to \mathbb{R}$ as defining a coordinate $\varphi(x)$ of x, then weak convergence corresponds to coordinate-wise convergence. Strong convergence implies weak convergence: if $x_n \to x$ in norm and φ is a bounded linear functional, then

$$|\varphi(x_n) - \varphi(x)| = |\varphi(x_n - x)| \leq \|\varphi\|\|x_n - x\| \to 0.$$

Weak convergence does not imply strong convergence on an infinite-dimensional space, as we will see in Section 8.6.

If X^* is the dual of a Banach space X, then we can define another type of weak convergence on X^*, called weak-$*$ convergence, pronounced "weak star."

Definition 5.60 Let X^* be the dual of a Banach space X. We say $\varphi \in X^*$ is the *weak-$*$* limit of a sequence (φ_n) in X^* if

$$\varphi_n(x) \to \varphi(x) \qquad \text{as } n \to \infty,$$

for every $x \in X$. We denote weak-$*$ convergence by

$$\varphi_n \overset{*}{\rightharpoonup} \varphi.$$

By contrast, weak convergence of (φ_n) in X^* means that

$$F(\varphi_n) \to F(\varphi) \qquad \text{as } n \to \infty,$$

for every $F \in X^{**}$. If X is reflexive, then weak and weak-$*$ convergence in X^* are equivalent because every bounded linear functional on X^* is of the form (5.28). If X^* is the dual space of a nonreflexive space X, then weak and weak-$*$ convergence are different, and it is preferable to use weak-$*$ convergence in X^* instead of weak convergence.

One reason for the importance of weak-$*$ convergence is the following compactness result, called the Banach-Alaoglu theorem.

Theorem 5.61 (Banach-Alaoglu) Let X^* be the dual space of a Banach space X. The closed unit ball in X^* is weak-$*$ compact.

We will not prove this result here, but we prove a special case of it in Theorem 8.45.

5.7 References

For more on linear operators in Banach spaces, see Kato [26]. For proofs of the Hahn-Banach, open mapping, and Banach-Alaoglu theorems, see Folland [12], Reed and Simon [45], or Rudin [48]. The use of linear and Banach spaces in optimization theory is discussed in [34]. Applied functional analysis is discussed in Lusternik and Sobolev [33]. For an introduction to semigroups associated with evolution equations, see [4]. For more on matrices, see [24]. An introduction to the numerical aspects of matrices and linear algebra is in [54]. For more on the stability, consistency, and convergence of finite difference schemes for partial differential equations, see Strikwerder [53].

5.8 Exercises

Exercise 5.1 Prove that the expressions in (5.2) and (5.3) for the norm of a bounded linear operator are equivalent.

Exercise 5.2 Suppose that $\{e_1, e_2, \ldots, e_n\}$ and $\{\widetilde{e}_1, \widetilde{e}_2, \ldots, \widetilde{e}_n\}$ are two bases of the n-dimensional linear space X, with

$$\widetilde{e}_i = \sum_{j=1}^{n} L_{ij} e_j, \qquad e_i = \sum_{j=1}^{n} \widetilde{L}_{ij} \widetilde{e}_j,$$

where (L_{ij}) is an invertible matrix with inverse $\left(\widetilde{L}_{ij}\right)$, meaning that

$$\sum_{j=1}^{n} L_{ij} \widetilde{L}_{jk} = \delta_{ik}.$$

Let $\{\omega_1, \omega_2, \ldots, \omega_n\}$ and $\{\widetilde{\omega}_1, \widetilde{\omega}_2, \ldots, \widetilde{\omega}_n\}$ be the associated dual bases of X^*.

(a) If $x = \sum x_i e_i = \sum \widetilde{x}_i \widetilde{e}_i \in X$, then prove that the components of x transform under a change of basis according to

$$\widetilde{x}_i = \widetilde{L}_{ij} x_j. \tag{5.29}$$

(b) If $\varphi = \sum \varphi_i \omega_i = \sum \widetilde{\varphi}_i \widetilde{\omega}_i \in X^*$, then prove that the components of φ transform under a change of basis according to

$$\widetilde{\varphi}_i = L_{ji} \varphi_j. \tag{5.30}$$

Exercise 5.3 Let $\delta : C([0,1]) \to \mathbb{R}$ be the linear functional that evaluates a function at the origin: $\delta(f) = f(0)$. If $C([0,1])$ is equipped with the sup-norm,

$$\|f\|_\infty = \sup_{0 \le x \le 1} |f(x)|,$$

show that δ is bounded and compute its norm. If $C([0,1])$ is equipped with the one-norm,

$$\|f\|_1 = \int_0^1 |f(x)| dx,$$

show that δ is unbounded.

Exercise 5.4 Consider the 2×2 matrix

$$A = \begin{pmatrix} 0 & a^2 \\ b^2 & 0 \end{pmatrix},$$

where $a > b > 0$. Compute the spectral radius $r(A)$ of A. Show that the Euclidean norms of powers of the matrix are given by

$$\left\|A^{2n}\right\| = a^{2n} b^{2n}, \qquad \left\|A^{2n+1}\right\| = a^{2n+2} b^{2n}.$$

Verify that $r(A) = \lim_{n \to \infty} \|A^n\|^{1/n}$.

Exercise 5.5 Define $K : C([0,1]) \to C([0,1])$ by

$$Kf(x) = \int_0^1 k(x,y) f(y)\, dy,$$

where $k : [0,1] \times [0,1] \to \mathbb{R}$ is continuous. Prove that K is bounded and

$$\|K\| = \max_{0 \le x \le 1} \left\{ \int_0^1 |k(x,y)|\, dy \right\}.$$

Exercise 5.6 Let X be a normed linear space. Use the Hahn-Banach theorem to prove the following statements.

(a) For any nonzero $x \in X$, there is a bounded linear functional $\varphi \in X^*$ such that $\|\varphi\| = 1$ and $\varphi(x) = \|x\|$.

(b) If $x, y \in X$ and $\varphi(x) = \varphi(y)$ for all $\varphi \in X^*$, then $x = y$.

Exercise 5.7 Find the kernel and range of the linear operator $K : C([0,1]) \to C([0,1])$ defined by

$$Kf(x) = \int_0^1 \sin \pi(x - y) f(y)\, dy.$$

Exercise 5.8 Prove that equivalent norms on a normed linear space X lead to equivalent norms on the space $\mathcal{B}(X)$ of bounded linear operators on X.

Exercise 5.9 Prove Proposition 5.43.

Exercise 5.10 Suppose that $k : [0,1] \times [0,1] \to \mathbb{R}$ is a continuous function. Prove that the integral operator $K : C([0,1]) \to C([0,1])$ defined by

$$Kf(x) = \int_0^1 k(x, y) f(y)\, dy$$

is compact.

Exercise 5.11 Prove that if $T_n \to T$ uniformly, then $\|T_n\| \to \|T\|$.

Exercise 5.12 Prove that if T_n converges to T uniformly, then T_n converges to T strongly.

Exercise 5.13 Suppose that Λ is the diagonal $n \times n$ matrix and N is the $n \times n$ *nilpotent matrix* (meaning that $N^k = 0$ for some k)

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}, \qquad N = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}.$$

(a) Compute the two-norms and spectral radii of Λ and N.
(b) Compute $e^{\Lambda t}$ and e^{Nt}.

Exercise 5.14 Suppose that A is an $n \times n$ matrix. For $t \in \mathbb{R}$ we define $f(t) = \det e^{tA}$.

(a) Show that

$$\lim_{t \to 0} \frac{f(t) - 1}{t} = \operatorname{tr} A,$$

where $\operatorname{tr} A$ is the trace of the matrix A, that is the sum of its diagonal elements.

(b) Deduce that $f : \mathbb{R} \to \mathbb{R}$ is differentiable, and is a solution of the ODE $\dot{f} = (\operatorname{tr} A) f$.

(c) Show that

$$\det e^A = e^{\operatorname{tr} A}.$$

Exercise 5.15 Suppose that A and B are bounded linear operators on a Banach space.

(a) If A and B commute, then prove that $e^A e^B = e^{A+B}$.

(b) If $[A, [A, B]] = [B, [A, B]] = 0$, then prove that

$$e^A e^B = e^{A+B+[A,B]/2}.$$

This result is called the *Baker-Campbell-Hausdorff formula*.

Exercise 5.16 Suppose that A and B are, possibly noncommuting, bounded operators on a Banach space. Show that

$$\lim_{t \to 0} \frac{e^{t(A+B)} - e^{tA} e^{tB}}{t^2} = -\frac{1}{2}[A, B],$$
$$\lim_{t \to 0} \frac{e^{t(A+B)} - e^{tA/2} e^{tB} e^{tA/2}}{t^2} = 0.$$

Show that for small t the function $e^{tA/2} e^{tB} e^{tA/2} x(0)$ provides a better approximation to the solution of the equation $x_t = (A + B)x$ than the function $e^{tA} e^{tB} x(0)$. The approximation $e^{t(A+B)} \approx e^{tA/2} e^{tB} e^{tA/2}$, called *Strang splitting*, is useful in the numerical solution of evolution equations by *fractional step methods*.

Exercise 5.17 Suppose that $K : X \to X$ is a bounded linear operator on a Banach space X with $\|K\| < 1$. Prove that $I - K$ is invertible and

$$(I - K)^{-1} = I + K + K^2 + K^3 + \dots,$$

where the series on the right hand side converges uniformly in $\mathcal{B}(X)$.