Christopher Heil

Introduction to Real Analysis

Alternative Chapter 1

An Introduction to Norms and Banach Spaces

Last Updated: January 13, 2021

Chapter 1′

An Introduction to Norms and Banach Spaces

Overview. This online chapter is an alternative, expanded version of Chapter 1 from the text "An Introduction to Real Analysis" by C. Heil (hereafter called the "main text").

The main goal of both Chapter 1 and Chapter 1′ is to provide a review of normed spaces, which appear in various contexts in the main text. Metric spaces also appear in the main text, although more rarely; in fact, metric spaces only appear in problems or exercises in the main text.

Chapter 1 provides a very brief review of both metric and normed spaces, while this alternative chapter provides a much more comprehensive review of normed spaces (with metric spaces defined in the exercises). This Chapter 1′ gives detailed coverage of the background topics on normed spaces that are most important for the main text, and it does so with discussion, motivation, and examples (whereas Chapter 1 just gives a brief review of the definitions and main theorems).

Chapter 1′ presents many examples that appear in the later chapters of the main text. In particular, the ℓ^p spaces are introduced and discussed in detail in Chapter 1′, while in the main text they do not appear until Chapter 7. Likewise we see an introduction to the L^p-norms in Chapter 1′, although these are not studied in detail until Chapter 7 in the main text.

Introduction. Much of what we do in analysis centers on issues of *convergence* or *approximation.* What does it mean for one object to be close to (or to approximate) another object? How can we define the *limit* of a sequence of objects that appear to be converging in some sense? If our objects are points in $\mathbb{R}^d$ then we can simply grab a ruler, embed ourselves into d-dimensional space, and measure the physical (or *Euclidean*) distance between the points. Two points are close if the distance between them is small, and a sequence of points x_n converges to a point x if the distance between x_n and x shrinks to zero in the limit as $n \to \infty$. However, the objects we work with are often not points in $\mathbb{R}^d$ but instead are elements of some other set X (perhaps a set of sequences, or a set of functions, or some other abstract set). Even so, if we can find a way to measure the distance between elements of X then we can still think about approximation or convergence. We simply say that two elements are close if the distance between them is small, and a sequence of elements x_n converges to an element x if the distance from x_n to x shrinks to zero. If the *properties* of the distance function on X are similar to those of Euclidean distance, then we will be able to prove useful theorems about X and its elements. We will make these ideas precise in this chapter.

Notation. Recall from the Preliminaries that we let the symbol $\overline{\mathbf{F}}$ denote a choice of the extended real line $[-\infty, \infty]$ or the complex plane $\mathbb{C}$. Further, in this context:

- if $\overline{\mathbf{F}} = [-\infty, \infty]$, then the word *scalar* means a *finite real number* $c \in \mathbb{R}$;
- if $\overline{\mathbf{F}} = \mathbb{C}$, then the word *scalar* means a *complex number* $c \in \mathbb{C}$.

Consequently, if we say that "$x = (x_k)_{k\in\mathbb{N}}$ is a series of scalars," then it is a sequence of real numbers if we have chosen $\overline{\mathbf{F}} = [-\infty, \infty]$, while it is a sequence of complex numbers if we have chosen $\overline{\mathbf{F}} = \mathbb{C}$. Likewise, "a scalar-valued function f on X" is a real-valued function $f\colon X \to \mathbb{R}$ if $\overline{\mathbf{F}} = [-\infty, \infty]$, and a complex-valued function $f\colon X \to \mathbb{C}$ if $\overline{\mathbf{F}} = \mathbb{C}$. ◊

1′.1 The Definition of a Norm

In a vector space we can add vectors and multiply vectors by scalars. A *norm* assigns to each vector x in X a *length* $\|x\|$ in a way that respects the structure of X. Specifically, a norm must be *homogeneous* in the sense that $\|cx\| = |c|\,\|x\|$ for all scalars c and all vectors x, and a norm must satisfy the Triangle Inequality, which in this setting takes the form $\|x+y\| \le \|x\| + \|y\|$. When we have a norm we also have a way of measuring the *distance between points*; the distance between x and y is *length of their difference*, i.e., $\|x-y\|$.

Definition 1′.1.1 (Seminorms and Norms). Let X be a vector space. A *seminorm* on X is a function $\|\cdot\|\colon X \to \mathbb{R}$ such that for all vectors $x, y \in X$ and all scalars c we have:

(a) Nonnegativity: $0 \le \|x\| < \infty$,

(b) Homogeneity: $\|cx\| = |c|\,\|x\|$, and

(c) The Triangle Inequality: $\|x + y\| \le \|x\| + \|y\|$.

A seminorm is a *norm* if we also have:

(d) Uniqueness: $\|x\| = 0$ if and only if $x = 0$.

A vector space X together with a norm $\|\cdot\|$ is called a *normed vector space*, a *normed linear space*, or simply a *normed space*. ◊

We often refer to the elements of a normed space X as "points" or "vectors," and we mostly use letters such as x, y, z to denote elements of the space. If the elements of our set X are functions (which is the case for many of the examples in this text), then we may refer to them as "points," "vectors," or "functions." Further, if we know that the elements of X are functions, then we usually denote them by letters such as f, g, h (instead of x, y, z).

We refer to the number $\|x\|$ as the *length* of a vector x, and we say that $\|x-y\|$ is the *distance* between the vectors x and y. A vector x that has length 1 is called a *unit vector*, or is said to be *normalized*. If y is any nonzero vector, then

$$x = \frac{1}{\|y\|}\,y = \frac{y}{\|y\|}$$

is a unit vector.

Here are some examples of norms on $\mathbb{R}^d$.

Exercise 1′.1.2. Prove that each of the following is a norm on $\mathbb{R}^d$, where $x = (x_1, \dots, x_d)$ denotes a vector in $\mathbb{R}^d$.

(a) The ℓ^1-*norm*: $\|x\|_1 = |x_1| + \cdots + |x_d|$.

(b) The *Euclidean norm* or ℓ^2-*norm*: $\|x\|_2 = \bigl(|x_1|^2 + \cdots + |x_d|^2\bigr)^{1/2}$.

(c) The ℓ^∞-*norm*: $\|x\|_\infty = \max\bigl\{|x_1|, \dots, |x_d|\bigr\}$. ♢

The Euclidean norm of a vector in $\mathbb{R}^d$ is the ordinary physical length of the vector (as measured by a ruler in d-dimensional space). The Triangle Inequality for the Euclidean norm on $\mathbb{R}^2$ is illustrated in Figure 1′.1.

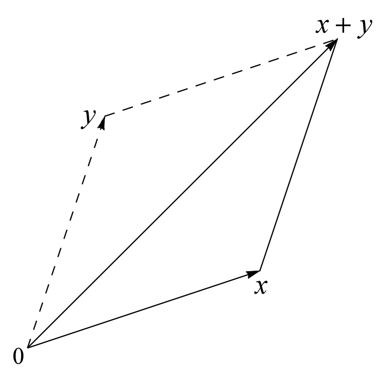

Fig. 1′.1 The Triangle Inequality for the Euclidean norm on $\mathbb{R}^2$. Two vectors x and y and their sum $x+y$ are pictured. The lengths of the three edges of the triangle drawn with solid lines are $\|x\|_2$, $\|y\|_2$, and $\|x+y\|_2$, and we can see that $\|x\|_2 + \|y\|_2 \geq \|x+y\|_2$.

1′.1.1 The Sequence Space ℓ^1

Now we will introduce an infinite-dimensional vector space known as "ℓ^1" (pronounced "little ell one"). The elements of this vector space are infinite sequences of scalars that satisfy a certain "summability" property, and we will define a norm on ℓ^1 that is related to summability.

Given a sequence of scalars $x = (x_k)_{k\in\mathbb{N}} = (x_1, x_2, \dots)$, we define the ℓ^1-*norm of* x to be

$$\|x\|_1 \;=\; \|(x_k)_{k\in\mathbb{N}}\|_1 \;=\; \sum_{k=1}^{\infty} |x_k|. \tag{1′.1}$$

In some sense, the quantity $\|x\|_1$ measures the "size" of x. This size is finite for some sequences but infinite for others. For example, if $x = (1, 1, 1, \dots)$ then $\|x\|_1 = \infty$, but for the sequence $y = (1, \frac{1}{4}, \frac{1}{9}, \dots)$ we have (by *Euler's Formula*) that $\|y\|_1 = \sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6} < \infty$.

We say that a sequence $x = (x_k)_{k\in\mathbb{N}}$ is *absolutely summable*, or just *summable* for short, if $\|x\|_1 < \infty$. We let ℓ^1 denote the space of all summable sequences, i.e.,

$$\ell^1 \;=\; \biggl\{x = (x_k)_{k\in\mathbb{N}} \;:\; \|x\|_1 \;=\; \sum_{k=1}^{\infty} |x_k| \;<\; \infty\biggr\}.$$

That is, ℓ^1 contains exactly those sequences $x = (x_k)_{k\in\mathbb{N}}$ for which $\|x\|_1$ is finite.

Here are some examples (the reader should verify these, and construct other examples of sequences that do or do not belong to ℓ^1).

- $x = (1, 1, 1, \dots) \notin \ell^1$;
- $y = \bigl(\frac{1}{k^2}\bigr)_{k\in\mathbb{N}} = (1, \frac{1}{4}, \frac{1}{9}, \frac{1}{16}, \dots) \in \ell^1$;
- $s = (1, 0, -1, 0, 0, 1, 0, 0, 0, -1, , 0, 0, 0, 0, 1, \dots) \notin \ell^1$;
- $t = (-1, 0, \frac{1}{2}, 0, 0, -\frac{1}{3}, 0, 0, 0, \frac{1}{4}, 0, 0, 0, 0, -\frac{1}{5}, \dots) \notin \ell^1$;
- if $p > 1$, then $u = \Bigl(\frac{(-1)^k}{k^p}\Bigr)_{k\in\mathbb{N}} \in \ell^1$;
- $v = (2^{-k})_{k\in\mathbb{N}} = \bigl(\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \dots\bigr) \in \ell^1$;
- if z is a scalar with absolute value $|z| < 1$, then $w = \bigl(z^k\bigr)_{k\in\mathbb{N}} \in \ell^1$.

Observe that a sequence $x = (x_k)_{k\in\mathbb{N}}$ belongs to ℓ^1 if and only if the sequence of absolute values $y = (|x_k|)_{k\in\mathbb{N}}$ belongs to ℓ^1.

If $x = (x_1, x_2, \dots)$ and $y = (y_1, y_2, \dots)$ are two elements of ℓ^1, then

$$x + y \;=\; (x_k + y_k)_{k\in\mathbb{N}} \;=\; (x_1 + y_1,\, x_2 + y_2,\, \dots)$$

also belongs to ℓ^1 (why?). Likewise, if c is scalar then $cx = (cx_1, cx_2, \dots)$ is an element of ℓ^1. Thus ℓ^1 is closed under addition of vectors and under scalar multiplication, and the reader should verify that it follows from this that ℓ^1 is a vector space. Therefore we often call a sequence $x = (x_k)_{k\in\mathbb{N}}$ in ℓ^1 a "vector." Although x is a sequence of infinitely many numbers x_k, it is just one element of ℓ^1, and so we also often say that x is one "point" in ℓ^1.

Remark 1′.1.3. The quantity $\|x\|_1$ is defined for every sequence x, although it could be ∞. For example, if $x = (1,1,1,\dots)$ then $\|x\|_1 = \infty$, while if $x = (\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \dots)$ then $\|x\|_1 = 1$. The space ℓ^1 consists of all those x for which $\|x\|_1$ is finite. In this example, it is not that $\|x\|_1$ does not exist when $x \notin \ell^1$, but rather that it is not finite when $x \notin \ell^1$. This need not be the case in other circumstances. If $\|\cdot\|$ is a norm on a vector space x, then $\|x\|$ need only be defined for vectors in x; it could simply be undefined (rather than being infinite) for vectors not in X. ◻

Lemma 1′.1.4. *$\|\cdot\|_1$ is a norm on ℓ^1.*

Proof. The nonnegativity and finiteness condition $0 \le \|x\|_1 < \infty$ is certainly satisfied for each $x \in \ell^1$, and the homogeneity condition is straightforward. The Triangle Inequality follows from the calculation

$$\begin{aligned} \|x+y\|_1 \;=\; \sum_{k=1}^{\infty} |x_k + y_k| \;&\le\; \sum_{k=1}^{\infty} \bigl(|x_k| + |y_k|\bigr) \\ &=\; \sum_{k=1}^{\infty} |x_k| + \sum_{k=1}^{\infty} |y_k| \\ &=\; \|x\|_1 + \|y\|_1. \end{aligned}$$

Finally, if $\|x\|_1 = 0$ then we must have $x_k = 0$ for every k, so $x = (0,0,\dots)$, which is the zero vector in ℓ^1. This shows that $\|\cdot\|_1$ is a norm on ℓ^1. ◻

There are many norms on ℓ^1 other than $\|\cdot\|_1$. However, unless we explicitly state otherwise, whenever we deal with ℓ^1 we always assume that $\|\cdot\|_1$ is the norm we are using.

The following particular elements of ℓ^1 appear so often that we introduce a name for them.

Notation 1′.1.5 (The Standard Basis Vectors). Given an integer $n \in \mathbb{N}$, we let δ_n denote the sequence

$$\delta_n \;=\; (0, \dots, 0, 1, 0, 0, \dots),$$

where the 1 is in the nth component and the other components are all zero. We often denote the kth component of δ_n by $\delta_n(k)$, so $\delta_n = \bigl(\delta_n(k)\bigr)_{k\in\mathbb{N}}$ where $\delta_n(k) = 0$ if $k \ne n$, and $\delta_n(n) = 1$.

We call δ_n the *nth standard basis vector*, and we refer to the family $\{\delta_n\}_{n\in\mathbb{N}}$ as the *sequence of standard basis vectors*, or simply the *standard basis*. $\diamondsuit$

The use of the word "basis" in Notation 1′.1.5 should just be regarded as a name for now. Problem 7.4.11 considers the issue of in what sense $\{\delta_n\}_{n\in\mathbb{N}}$ is or is not a "basis" for ℓ^1.

1′.1.2 The Sequence Space ℓ^∞

Now we define another infinite-dimensional vector space whose elements are infinite sequences of scalars.

Example 1′.1.6. Given a sequence of scalars $x = (x_k)_{k\in\mathbb{N}} = (x_1, x_2, \dots)$, we define the *sup-norm* or ℓ^∞*-norm of* x to be

$$\|x\|_\infty \;=\; \|(x_k)_{k\in\mathbb{N}}\|_\infty \;=\; \sup_{k\in\mathbb{N}} |x_k|. \tag{1′.2}$$

The quantity $\|x\|_\infty$ can be infinite; indeed, x is a bounded sequence if and only if $\|x\|_\infty < \infty$. For example, if $x = (1, 1, 1, \dots)$ then $\|x\|_\infty = 1$, while if $x = (1, 2, 3, \dots)$ then $\|x\|_\infty = \infty$.

We let ℓ^∞ denote the space of all bounded sequences, i.e.,

$$\ell^\infty \;=\; \Bigl\{x = (x_k)_{k\in\mathbb{N}} \;:\; \|x\|_\infty \;=\; \sup_{k\in\mathbb{N}} |x_k| \;<\; \infty\Bigr\}.$$

This is a vector space, and the reader should show that $\|\cdot\|_\infty$ is a norm on ℓ^∞. Unless we state otherwise, we always assume that this ℓ^∞*-norm* is the norm on ℓ^∞.

The two sets ℓ^1 and ℓ^∞ do not contain exactly the same vectors. For example, some of the vectors x, y, s, t, u, v, w discussed in Section 1′.1.1 belong to ℓ^∞ but do not belong to ℓ^1 (which ones?), and therefore $\ell^\infty \neq \ell^1$. On the other hand, the reader should prove that every vector in ℓ^1 also belongs to ℓ^∞, and therefore ℓ^1 is a *proper subspace* of ℓ^∞. $\diamondsuit$

We have to use a supremum rather than a maximum in equation (1′.2) because it is not true that every bounded sequence has a maximum. For example, if $x = (x_k)_{k\in\mathbb{N}}$ is the sequence whose components are $x_k = k/(k+1)$, then x is a bounded sequence and hence is a vector in ℓ^∞, but there is no maximum component x_k.

1′.1.3 The Uniform Norm

Next we give an example of an infinite-dimensional normed space whose elements are functions rather than sequences. For convenience, we will take the interval $[0,1]$ to be the domain of the functions in this example, but we could just as well have chosen other sets to be the domain.

Example 1′.1.7. We define $\mathcal{F}[0,1]$ to be the set of all scalar-valued functions whose domain is the closed interval $[0,1]$. A "point" or "vector" in $\mathcal{F}[0,1]$ is an element of $\mathcal{F}[0,1]$, i.e., it is a function that maps $[0,1]$ to scalars. The zero vector in $\mathcal{F}[0,1]$ is the *zero function*, i.e., the function that takes the value zero at every point. We denote this function by the symbol 0. That is, 0 is the function defined by the rule $0(t) = 0$ for $t \in [0,1]$.

Consider the subspace of $\mathcal{F}[0,1]$ that only contains bounded functions:

$$\mathcal{F}_b[0,1] \;=\; \bigl\{f \in \mathcal{F}[0,1] \,:\, f \text{ is bounded}\bigr\}.$$

By definition, a function f is bounded if and only if there is some finite number M such that $|f(t)| \le M$ for all t. Given any function f (not necessarily bounded), we call

$$\|f\|_{\mathrm{u}} \;=\; \sup_{t\in[0,1]} |f(t)| \tag{1′.3}$$

the *uniform norm* of f. The bounded functions are precisely the functions whose uniform norm is finite. Thus $\mathcal{F}_b[0,1]$ consists of those functions f on $[0,1]$ for which $\|f\|_{\mathrm{u}}$ is finite. Problem 1′.1.11 asks for a proof that $\|\cdot\|_{\mathrm{u}}$ is a norm on $\mathcal{F}_b[0,1]$.

The *uniform distance* between two bounded functions f and g is

$$\|f-g\|_{\mathrm{u}} \;=\; \sup_{t\in[0,1]} |f(t)-g(t)|. \tag{1′.4}$$

In some sense, $\|f-g\|_{\mathrm{u}}$ is the "maximum deviation" between $f(t)$ and $g(t)$ over all t. However, it is important to note that the supremum in equation (1′.4) need not be achieved, so there need not be an actual maximum to the values of $|f(t)-g(t)|$ (if f and g are continuous then it is true that there is a maximum to $|f(t)-g(t)|$ on the finite closed interval $[a,b]$, but this need not be the case if f or g is discontinuous or if the domain is not compact).

To illustrate, consider the two continuous functions $f(t) = 5t-1$ and $g(t) = 9t^3 - 18t^2 + 11t - 1$, whose graphs over the domain $[0,1]$ are shown in Figure 1′.2. The distance between f and g, as measured by the uniform norm, is the supremum of $|f(t)-g(t)|$ over all values of $t \in [0,1]$. For these specific functions f and g this supremum is achieved at the point $t = 1$, so

$$\|f-g\|_{\mathrm{u}} \;=\; \sup_{0\le t\le 1} |f(t)-g(t)| \;=\; |f(1)-g(1)| \;=\; 4-1 \;=\; 3.$$

Thus, using this norm, the points f and g are 3 units apart in $\mathcal{F}_b[0,1]$. ◇

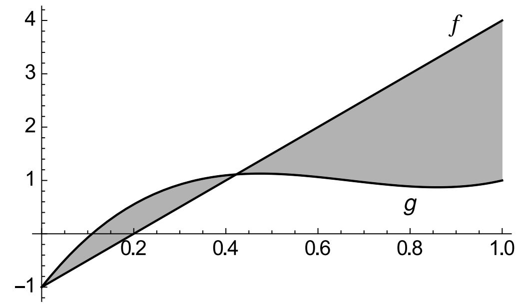

Fig. 1′.2 Graphs of the functions $f(t) = 5t - 1$ and $g(t) = 9t^3 - 18t^2 + 11t - 1$. The region between the graphs is shaded.

Here is a different norm on a space of functions. In order to define this norm we need to restrict our attention to functions that can be integrated. For now we restrict to continuous functions on the domain $[0, 1]$, for which the *Riemann integral* is defined (later, in Chapter 4 we will introduce the more general *Lebesgue integral* of functions).

Example 1′.1.8. Let $C[0, 1]$ denote the space of all continuous functions f on the domain $[0, 1]$. This is a subspace of $\mathcal{F}_b[0, 1]$. Consequently if we restrict the uniform norm to just $C[0, 1]$ then we obtain a norm on $C[0, 1]$.

However, $\|\cdot\|_{\mathrm{u}}$ is not the only norm that we can define on $C[0, 1]$. Recall that every continuous function on $[0, 1]$ is Riemann integrable, i.e., the Riemann integral $\int_0^1 f(t)\,dt$ exists for every continuous function f on $[0, 1]$. The L^1-*norm* of f is defined to be the integral of the absolute value of f:

$$\|f\|_1 = \int_0^1 |f(t)|\,dt. \qquad (1'.5)$$

With this definition, the corresponding *distance* between points in $C[0, 1]$ is

$$\|f - g\|_1 = \int_0^1 |f(t) - g(t)|\,dt, \qquad f, g \in C[0, 1]. \qquad (1'.6)$$

The proof that $\|\cdot\|_1$ satisfies the requirements of a norm on $C[0, 1]$ is assigned as Problem 1′.1.11.

Let $f(t) = 5t - 1$ and $g(t) = 9t^3 - 18t^2 + 11t - 1$ be the two functions pictured in Figure 1′.2. We saw in Example 1′.1.7 that $\|f - g\|_{\mathrm{u}} = 3$. This is because the maximum value of $|f(t) - g(t)|$ over $t \in [0, 1]$ is 3 units. The L^1-norm measures the distance between these functions in a different way. Instead of basing the distance on just the *supremum* of $|f(t) - g(t)|$, the L^1-norm takes all values of $|f(t) - g(t)|$ into account by computing the *integral* of $|f(t) - g(t)|$ over all t. For these two functions f and g, the L^1-norm is the *area* of the shaded region depicted in Figure 1′.2, which is

$$\begin{aligned}\|f-g\|_1 &= \int_0^1 |-9t^3+18t^2-6t|\,dt \\ &= \int_0^{1-\frac{\sqrt{3}}{3}} (9t^3-18t^2+6t)\,dt \;+\; \int_{1-\frac{\sqrt{3}}{3}}^1 (-9t^3+18t^2-6t)\,dt \\ &= \frac{16\sqrt{3}-15}{12} \;\approx\; 1.059401\ldots.\end{aligned}$$

Thus f and g are fairly close as measured by the L^1-norm, being only a little over one unit apart. In contrast, f and g are a fairly distant 3 units apart as measured by the uniform norm. Neither of these values is any more "correct" than the other. However, depending on our application, one of these ways of measuring may potentially be more *useful* than the other. ♢

Problems

1′.1.9. Given a fixed finite dimension $d \ge 1$ prove that there exist finite, positive numbers A_d and B_d such that the following inequality holds simultaneously for every $x \in \mathbb{R}^d$:

$$A_d\,\|x\|_\infty \;\le\; \|x\|_1 \;\le\; B_d\,\|x\|_\infty, \qquad \text{for all } x \in \mathbb{R}^d. \tag{1′.7}$$

The numbers A_d and B_d can depend on the dimension, but they must be independent of the choice of vector x. What are the *optimal values* for A_d and B_d, i.e., what is the largest number A_d and the smallest number B_d such that equation (1′.7) holds for every $x \in \mathbb{R}^d$?

1′.1.10. Prove the following statements.

(a) The function $\|\cdot\|_\infty$ defined in equation (1′.2) is a norm on ℓ^∞.

(b) $\|x\|_\infty \le \|x\|_1$ for every $x \in \ell^1$.

(c) There does not exist a finite constant $B > 0$ such that the inequality $\|x\|_1 \le B\,\|x\|_\infty$ holds simultaneously for *every* $x \in \ell^1$.

1′.1.11. Prove the following statements.

(a) The function $\|\cdot\|_{\mathrm{u}}$ defined in equation (1′.3) is a norm on $\mathcal{F}_b[0,1]$, and is therefore also a norm on the subspace $C[0,1]$.

(b) The function $\|\cdot\|_1$ defined in equation (1′.6) is a norm on $C[0,1]$.

(c) $\|f\|_1 \le \|f\|_{\mathrm{u}}$ for every $f \in C[0,1]$.

(d) There is a function $f \in C[0,1]$ such that $\|f\|_{\mathrm{u}} = 1000$ yet $\|f\|_1 = 1$.

(e) There does not exist a finite constant $B > 0$ such that the inequality $\|f\|_{\mathrm{u}} \le B\,\|f\|_1$ holds simultaneously for *every* $f \in C[0,1]$.

1′.2 Convergence and Completeness

1′.2.1 Convergence

If $\|\cdot\|$ is a norm on a vector space X, then the number $\|x-y\|$ represents the distance from the point $x \in X$ to the point $y \in X$ with respect to this norm. We will say that points x_n are *converging* to a point x if the distance from x_n to x shrinks to zero as n increases. This is made precise in the following definition.

Definition 1′.2.1 (Convergent Sequence). Let X be a normed space. We say that a sequence of points $\{x_n\}_{n\in\mathbb{N}}$ in X *converges* to the point $x \in X$ if

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

That is, for every $\varepsilon > 0$ there must exist some integer $N > 0$ such that

$$n \geq N \implies \|x - x_n\| \,<\, \varepsilon.$$

In this case, we write $x_n \to x$ *as* $n \to \infty$, or simply $x_n \to x$ for short. ◊

Convergence implicitly depends on the choice of norm for X, so if we want to emphasize that we are using a particular norm, we may write $x_n \to x$ *with respect to the norm* $\|\cdot\|$.

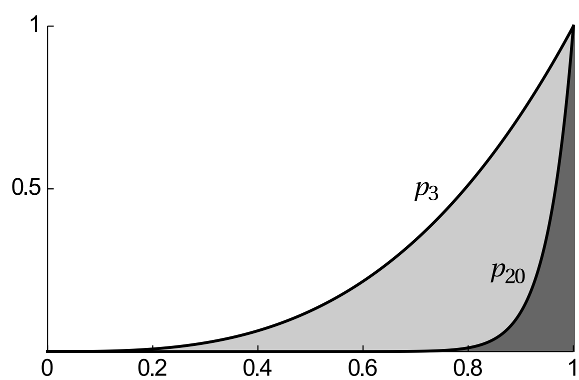

Fig. 1′.3 Graphs of the functions $p_3(t) = t^3$ and $p_{20}(t) = t^{20}$. The area of the region under the graph of p_3 is $1/4$, while the area of the region under the graph of p_{20} is $1/21$.

Example 1′.2.2. Let $X = C[0,1]$, the space of continuous functions on the domain $[0,1]$. One norm on $C[0,1]$ is the uniform norm defined in equation (1′.3), and another is the L^1-norm defined in equation (1′.5). For each $n \in N$, let $p_n(t) = t^n$. With respect to the L^1-norm, the distance from p_n to the zero function is the area between the graphs of these two functions, which is

$$\|p_n - 0\|_1 \;=\; \int_0^1 |t^n - 0|\,dt \;=\; \int_0^1 t^n\,dt \;=\; \frac{1}{n+1}.$$

This distance decreases to 0 as n increases, so p_n *converges to* 0 *with respect to the* L^1*-norm* (consider Figure 1′.3). However, if we change the norm then the meaning of distance and convergence changes. For example, if we instead measure distance using the uniform norm, then for every n we have

$$\|p_n - 0\|_{\mathrm{u}} \;=\; \sup_{0\le t\le 1} |t^n - 0| \;=\; \sup_{0\le t\le 1} t^n \;=\; 1.$$

Using this norm the two vectors p_n and 0 are 1 unit apart, no matter how large we choose n. The distance between p_n and 0 does not decrease with n, so p_n *does not* converge to 0 with respect to the uniform norm. ♢

1′.2.2 Cauchy Sequences

Closely related to *convergence* is the idea of a *Cauchy sequence*, which is a sequence of points $\{x_n\}_{n\in\mathbb{N}}$ where the distance $\|x_m - x_n\|$ between two points x_m and x_n decreases to zero as m and n increase.

Definition 1′.2.3 (Cauchy Sequence). Let X be a normed space. A sequence of points $\{x_n\}_{n\in\mathbb{N}}$ in X is a *Cauchy sequence* if for every $\varepsilon > 0$ there exists an integer $N > 0$ such that

$$m, n \ge N \implies \|x_m - x_n\| < \varepsilon. \qquad \diamondsuit$$

Thus, if $\{x_n\}_{n\in\mathbb{N}}$ is a *convergent* sequence, then there exists a point x such that x_n gets closer and closer to x as n increases, while if $\{x_n\}_{n\in\mathbb{N}}$ is a *Cauchy* sequence, then the elements x_m, x_n of the sequence get closer and closer to each other as m and n increase.

We prove now that *every convergent sequence is Cauchy.*

Lemma 1′.2.4 (Convergent Implies Cauchy). *If* $\{x_n\}_{n\in\mathbb{N}}$ *is a convergent sequence in a normed space* X, *then* $\{x_n\}_{n\in\mathbb{N}}$ *is a Cauchy sequence in* X.

Proof. Assume that $x_n \to x$. If we fix an $\varepsilon > 0$ then, by the definition of a convergent sequence, there exists some $N > 0$ such that $\|x - x_n\| < \varepsilon/2$ for all $n \ge N$. Consequently, if $m, n \ge N$ then the Triangle Inequality implies that

$$\|x_m - x_n\| \;\le\; \|x_m - x\| + \|x - x_n\| \;<\; \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \;=\; \varepsilon.$$

Therefore $\{x_n\}_{n\in\mathbb{N}}$ is Cauchy. □

Example 1′.2.5. The contrapositive formulation of Lemma 1′.2.4 is that *a sequence that is not Cauchy cannot converge.* To illustrate, consider the sequence of standard basis vectors $\{\delta_n\}_{n\in\mathbb{N}}$ in ℓ^1. If $m \neq n$, then $\delta_m - \delta_n$ has a 1 in the mth component, a -1 in the nth component, and zeros in all other components, and hence $\|\delta_m - \delta_n\|_1 = 2$. Consequently, if $\varepsilon < 2$ then we can *never* have $\|\delta_m - \delta_n\|_1 < \varepsilon$, no matter how large we take m and n. Hence $\{\delta_n\}_{n\in\mathbb{N}}$ is not a Cauchy sequence in ℓ^1, and therefore it does not converge to any element of ℓ^1. ♢

1′.2.3 Completeness

As we have seen, every convergent sequence is Cauchy, and a sequence that is not Cauchy cannot converge. This does not tell us whether Cauchy sequences converge. Here is an example of a Cauchy sequence in a particular normed space that *does not converge*.

Example 1′.2.6. Let c_{00} be the space of all sequences of scalars that have only finitely many nonzero components:

$$c_{00} \;=\; \Bigl\{x = (x_1, \dots, x_N, 0, 0, \dots) \;:\; N > 0,\; x_1, \dots, x_N \text{ are scalars}\Bigr\}.$$

A vector $x \in c_{00}$ is sometimes called a "finite sequence" (although this is an abuse of language since x does have infinitely many components, but only finitely many of these can be nonzero). Since c_{00} is a subspace of ℓ^1, it is a normed space with respect to the ℓ^1-norm.

For each $n \in \mathbb{N}$, let x_n be the sequence $x_n = (2^{-1}, \dots, 2^{-n}, 0, 0, 0, \dots)$, and consider the sequence of vectors $\{x_n\}_{n\in\mathbb{N}}$. This sequence is contained in both c_{00} and ℓ^1. If $m < n$, then

$$x_m - x_n \;=\; (0, \dots, 0, 2^{-m-1}, 2^{-m-2}, \dots, 2^{-n}, 0, 0, \dots),$$

so

$$\|x_n - x_m\|_1 \;=\; \sum_{k=m+1}^{n} 2^{-k} \;<\; \sum_{k=m+1}^{\infty} 2^{-k} \;=\; 2^{-m}.$$

The reader should verify that this implies that $\{x_n\}_{n\in\mathbb{N}}$ is a Cauchy sequence with respect to $\|\cdot\|_1$. If we take our normed space to be $X = \ell^1$, then then this sequence does converge. In fact x_n converges in ℓ^1-norm to the sequence $x = (2^{-1}, 2^{-2}, \dots) = (2^{-k})_{k\in\mathbb{N}}$ because

$$\|x - x_n\|_1 \;=\; \sum_{k=n+1}^{\infty} 2^{-k} \;=\; 2^{-n} \;\to\; 0 \quad \text{as } n \to \infty.$$

However, this vector x does not belong to c_{00}, and there is no other sequence $y \in c_{00}$ such that $x_n \to y$ (why not?). Therefore, if we take our normed space to be $X = c_{00}$, then $\{x_n\}_{n\in\mathbb{N}}$ is not a convergent sequence *in the space* c_{00}, even though it is a Cauchy sequence. □

To emphasize, for a sequence to be convergent in a normed space X it must converge *to an element of* X and not just to an element of some larger space. Example $1'.2.6$ shows that there are sequences in c_{00} that are Cauchy but do not converge to an element of c_{00}. The following theorem states that every Cauchy sequence in the real line $\mathbb{R}$ does converge to an element of $\mathbb{R}$ with respect to its standard norm (absolute value). For one proof of Theorem $1'.2.7$, see [Rud76, Thm. 3.11]. A similar result holds for the complex plane $\mathbb{C}$ with respect to absolute value.

Theorem $1'.2.7$. *If $\{x_n\}_{n\in\mathbb{N}}$ is a Cauchy sequence in $\mathbb{R}$, then there exists some $x \in \mathbb{R}$ such that $x_n \to x$.* ◊

In summary, *some* normed spaces have the property that every Cauchy sequence in the space converges to an element of the space. Since we can test for Cauchyness without having the limit vector x in hand, this is often very useful. We give such spaces the following name.

Definition $1'.2.8$ (Banach Space). Let X be a normed space. If every Cauchy sequence in X converges to an element of X, then we say that X is *complete*, and in this case we also say that X is a *Banach space*. ◊

Thus a Banach space is precisely a normed space that is complete. The terms "Banach space" and "complete normed space" are entirely interchangeable, and we will use whichever is more convenient in a given context.

Remark $1'.2.9$. The reader should be aware that the term "complete" is heavily overused and has a number of distinct mathematical meanings. In particular, the notion of a *complete space* as given in Definition $1'.2.8$ is quite different from the notion of a *complete sequence* that will be introduced in Definition 8.2.17. ◊

We will show that ℓ^1 is a Banach space. Recall that, unless otherwise specified, we take the norm on ℓ^1 to be $\|\cdot\|_1$, so $x_n \to x$ (convergence in ℓ^1) means that $\|x - x_n\|_1 \to 0$ as $n \to \infty$. To emphasize that the convergence is taking place with respect to this norm, we often say that $x_n \to x$ *in ℓ^1-norm* if $\|x - x_n\|_1 \to 0$.

Theorem $1'.2.10$. *If $\{x_n\}_{n\in\mathbb{N}}$ is a Cauchy sequence in ℓ^1, then there exists a vector $x \in \ell^1$ such that $x_n \to x$ in ℓ^1-norm. Consequently ℓ^1 is complete, and therefore it is a Banach space.*

Proof. Assume that $\{x_n\}_{n\in\mathbb{N}}$ is a Cauchy sequence in ℓ^1. Each x_n is a vector in ℓ^1, which means that x_n is an infinite sequence of scalars whose components are summable. For this proof we will write the components of x_n as

$$x_n \;=\; \bigl(x_n(1), x_n(2), \dots\bigr) \;=\; \bigl(x_n(k)\bigr)_{k\in\mathbb{N}}.$$

That is, $x_n(k)$ is the kth component of the vector x_n.

Choose any $\varepsilon > 0$. Then, by the definition of a Cauchy sequence, there is an integer $N \ge 0$ such that $\|x_m - x_n\|_1 < \varepsilon$ for all $m, n \ge N$. Therefore, if we fix a particular index $k \in \mathbb{N}$, then for all $m, n > N$ we have

$$|x_m(k) - x_n(k)| \;\le\; \sum_{j=1}^{\infty} |x_m(j) - x_n(j)| \;=\; \|x_m - x_n\|_1 \;<\; \varepsilon.$$

Thus, with k fixed, $\bigl(x_n(k)\bigr)_{n\in\mathbb{N}}$ is a Cauchy sequence of *scalars* and therefore, by Theorem 1′.2.7, it must converge to some scalar. Let $x(k)$ be the limit of this sequence, i.e., define

$$x(k) \;=\; \lim_{n\to\infty} x_n(k). \tag{1′.8}$$

Then let x be the sequence $x = \bigl(x(k)\bigr)_{k\in\mathbb{N}} = \bigl(x(1), x(2), \dots\bigr)$.

Now, for each fixed integer k, as $n \to \infty$ the scalar $x_n(k)$, which is the kth component of x_n, converges to the scalar $x(k)$, which is the kth component $x(k)$ of x. We therefore say that x_n *converges componentwise to* x (see the illustration in Figure 1′.4). However, this is not enough. We need to show that $x \in \ell^1$, and that x_n *converges to* x *in* ℓ^1*-norm.*

$$\begin{array}{rcllllll}
x_1 & = & (x_1(1), & x_1(2), & x_1(3), & x_1(4), & \dots) & \text{components of } x_1\\
x_2 & = & (x_2(1), & x_2(2), & x_2(3), & x_2(4), & \dots) & \text{components of } x_2\\
x_3 & = & (x_3(1), & x_3(2), & x_3(3), & x_3(4), & \dots) & \text{components of } x_3\\
 & & \quad\vdots & \quad\vdots & \quad\vdots & \quad\vdots & & \qquad\vdots\\
 & & \quad\downarrow & \quad\downarrow & \quad\downarrow & \quad\downarrow & & \qquad\downarrow\\
x & = & (x(1), & x(2), & x(3), & x(4), & \dots) & \text{components of } x
\end{array}$$

Fig. 1′.4 Illustration of componentwise convergence. For each k, the kth component of x_n converges to the kth component of x.

To prove that $x_n \to x$ in ℓ^1-norm, fix any $\varepsilon > 0$. Since $\{x_n\}_{n\in\mathbb{N}}$ is Cauchy, there is an $N > 0$ such that $\|x_m - x_n\|_1 < \varepsilon$ for all $m, n \ge N$. Choose any particular $n \ge N$, and fix an integer $M > 0$. Then, since M is finite, we compute that

$$\begin{aligned}
\sum_{k=1}^{M} |x(k) - x_n(k)| \;&=\; \sum_{k=1}^{M} \lim_{m\to\infty} |x_m(k) - x_n(k)| && \text{(by equation (1′.8))}\\
&=\; \lim_{m\to\infty} \sum_{k=1}^{M} |x_m(k) - x_n(k)| && \text{(since the sum is finite)}\\
&\le\; \lim_{m\to\infty} \sum_{k=1}^{\infty} |x_m(k) - x_n(k)| && \text{(all terms nonnegative)}\\
&=\; \lim_{m\to\infty} \|x_m - x_n\|_1 \;\le\; \varepsilon.
\end{aligned}$$

As this is true for every M, we conclude that

$$\|x - x_n\|_1 \;=\; \sum_{k=1}^{\infty} |x(k) - x_n(k)| \;=\; \lim_{M\to\infty} \sum_{k=1}^{M} |x(k) - x_n(k)| \;\le\; \varepsilon. \qquad (1'.9)$$

Even though we do not know yet that $x \in \ell^1$, this computation implies that the vector $y = x - x_n$ has finite ℓ^1-norm (because $\|y\|_1 = \|x - x_n\|_1 \le \varepsilon$, which is finite). Therefore y belongs to ℓ^1. Since ℓ^1 is closed under addition and since both y and x_n belong to ℓ^1, it follows that their sum $y + x_n$ belongs to ℓ^1. But $x = y + x_n$, so $x \in \ell^1$. Thus our "candidate sequence" x is in ℓ^1. Further, equation (1′.9) establishes that $\|x - x_n\|_1 \le \varepsilon$ for all $n \ge N$, so we have shown that x_n converges to x in ℓ^1-norm as $n \to \infty$. Therefore ℓ^1 is complete. □

We mentioned componentwise convergence in the proof of Theorem 1′.2.10. We make that notion precise in the following definition.

Definition 1′.2.11 (Componentwise Convergence). For each $n \in \mathbb{N}$ let $x_n = \big(x_n(k)\big)_{k\in\mathbb{N}}$ be a sequence of scalars, and let $x = \big(x(k)\big)_{k\in\mathbb{N}}$ be another sequence of scalars. We say that x_n *converges componentwise to* x if

$$\lim_{n\to\infty} x_n(k) \;=\; x(k), \qquad \text{for every } k \in \mathbb{N}. \qquad \diamondsuit$$

The proof of Theorem 1′.2.10 shows that if $x_n \to x$ in ℓ^1-norm (i.e., if $\|x - x_n\|_1 \to 0$), then x_n converges componentwise to x. However, the converse statement fails in general. For example, write the components of the nth standard basis vector as $\delta_n = (\delta_n(k))_{k\in\mathbb{N}}$, and let $0 = (0, 0, \dots)$ be the zero sequence. If we fix any particular index k, then $\delta_n(k) = 0$ for all $n > k$. Therefore

$$\lim_{n\to\infty} \delta_n(k) \;=\; 0.$$

Thus, for each fixed k the kth component of δ_n converges to the kth component of the zero vector as $n \to \infty$ (see Figure 1′.5). Consequently δ_n converges componentwise to the zero vector. However, we showed in Example 1′.2.5 that

$$\begin{array}{rcllllll}
\delta_1 & = & (1, & 0, & 0, & 0, & \dots) & \text{components of } \delta_1 \\
\delta_2 & = & (0, & 1, & 0, & 0, & \dots) & \text{components of } \delta_2 \\
\delta_3 & = & (0, & 0, & 1, & 0, & \dots) & \text{components of } \delta_3 \\
\delta_4 & = & (0, & 0, & 0, & 1, & \dots) & \text{components of } \delta_4 \\
 & & \vdots & \vdots & \vdots & \vdots & & \quad\quad\vdots \\
 & & \downarrow & \downarrow & \downarrow & \downarrow & & \quad\quad\downarrow \\
0 & = & (0, & 0, & 0, & 0, & \dots) & \text{components of the zero sequence}
\end{array}$$

Fig. 1′.5 For each k, the kth component of δ_n converges to the kth component of the zero sequence. However, $\|\delta_n - 0\|_1 = 1 \not\to 0$, so δ_n does not converge to the zero sequence in ℓ^1-norm.

δ_n does not converge in ℓ^1-norm. Thus componentwise convergence does not imply convergence in ℓ^1 in general.

Here are some properties of norms and convergence.

Lemma 1′.2.12. *If X is a normed space and x_n, x, y_n, y are vectors in X, then the following statements hold.*

(a) Uniqueness of Limits: *If $x_n \to x$ and $x_n \to y$, then $x = y$.*

(b) Reverse Triangle Inequality: $\big|\|x\| - \|y\|\big| \le \|x - y\|$.

(c) Convergent implies Cauchy: *If $x_n \to x$, then $\{x_n\}_{n\in\mathbb{N}}$ is Cauchy.*

(d) Boundedness of Cauchy sequences: *If $\{x_n\}_{n\in\mathbb{N}}$ is a Cauchy sequence, then* $\sup \|x_n\| < \infty$.

(e) Continuity of the norm: *If $x_n \to x$, then $\|x_n\| \to \|x\|$.*

(f) Continuity of vector addition: *If $x_n \to x$ and $y_n \to y$, then $x_n+y_n \to x+y$.*

(g) Continuity of scalar multiplication: *If $x_n \to x$ and $c_n \to c$ (where c_n and c are scalars), then $c_nx_n \to cx$.*

Proof. We will prove one part, and assign the proof of the remaining statements to the reader.

(d) Suppose that $\{x_n\}_{n\in\mathbb{N}}$ is Cauchy and consider $\varepsilon = 1$. Then there exists an $N > 0$ such that $\|x_m - x_n\| < 1$ for all $m, n \ge N$. Therefore, for $n \ge N$ we have

$$\|x_n\| \;=\; \|x_n - x_N + x_N\| \;\le\; \|x_n - x_N\| \,+\, \|x_N\| \;\le\; 1 \,+\, \|x_N\|.$$

Hence, if we let $R = \max\big\{\|x_1\|, \dots, \|x_{N-1}\|,\, \|x_N\| + 1\big\}$, then $\|x_n\| \le R$ for every n.

Exercise: Prove statements (a), (b), (c), (e), (f), and (g). □

Problems

1′.2.13. Given points x_n and x in a normed space X, prove that the following four statements are equivalent.

(a) $x_n \to x$, i.e., for each $\varepsilon > 0$ there exists an integer $N > 0$ such that

$$n \geq N \quad \Longrightarrow \quad \|x - x_n\| < \varepsilon.$$

(b) For each $\varepsilon > 0$ there exists an integer $N > 0$ such that

$$n \geq N \quad \Longrightarrow \quad \|x - x_n\| \leq \varepsilon.$$

(c) For each $\varepsilon > 0$ there exists an integer $N > 0$ such that

$$n > N \quad \Longrightarrow \quad \|x - x_n\| < \varepsilon.$$

(d) For each $\varepsilon > 0$ there exists an integer $N > 0$ such that

$$n > N \quad \Longrightarrow \quad \|x - x_n\| \leq \varepsilon.$$

Formulate and prove an analogous set of equivalent statements for Cauchy sequences.

1′.2.14. Let p_n be the function whose rule is $p_n(t) = t^n$, for $t \in [0,1]$.

(a) Prove directly that $\{p_n\}_{n\in\mathbb{N}}$ is a Cauchy sequence in $C[0,1]$ with respect to the L^1-norm.

(b) Prove directly that $\{p_n\}_{n\in\mathbb{N}}$ is a not Cauchy sequence in $C[0,1]$ with respect to the uniform norm.

1′.2.15. Let $x = (x_k)_{k\in\mathbb{N}}$ be any particular element of ℓ^1. Compute $\|x - \delta_n\|_1$, and show that $\|x - \delta_n\|_1 \not\to 0$ as $n \to \infty$. Conclude that the sequence $\{\delta_n\}_{n\in\mathbb{N}}$ does not converge to x, no matter which $x \in \ell^1$ that we choose.

1′.2.16. (a) Show that ℓ^∞ is complete with respect to the norm $\|\cdot\|_\infty$.

(b) Show that c_{00} is not complete with respect to the norm $\|\cdot\|_\infty$.

(c) Find a proper subspace of ℓ^∞ that is complete with respect to $\|\cdot\|_\infty$ (other than the trivial subspace $\{0\}$). Can you find an *infinite-dimensional* subspace that is complete?

1′.2.17. Suppose that $\{x_n\}_{n\in\mathbb{N}}$ is a Cauchy sequence in a normed space X, and suppose there exists a subsequence $\{x_{n_k}\}_{k\in\mathbb{N}}$ that converges to $x \in X$, i.e., $x_{n_k} \to x$ as $k \to \infty$. Prove that $x_n \to x$ as $n \to \infty$.

1′.2.18. Given a sequence $\{x_n\}_{n\in\mathbb{N}}$ in a normed space X, prove the following statements.

(a) If $\|x_n - x_{n+1}\| < 2^{-n}$ for every $n \in \mathbb{N}$, then $\{x_n\}_{n\in\mathbb{N}}$ is Cauchy.

(b) If $\{x_n\}_{n\in\mathbb{N}}$ is Cauchy, then there exists a subsequence $\{x_{n_k}\}_{k\in\mathbb{N}}$ such that $\|x_{n_k} - x_{n_{k+1}}\| < 2^{-k}$ for each $k \in \mathbb{N}$.

1′.2.19. (a) Let $\{x_n\}_{n\in\mathbb{N}}$ be a sequence in a normed space X, and fix a point $x \in X$. Suppose that every subsequence $\{y_n\}_{n\in\mathbb{N}}$ of $\{x_n\}_{n\in\mathbb{N}}$ has a subsequence $\{z_n\}_{n\in\mathbb{N}}$ of $\{y_n\}_{n\in\mathbb{N}}$ such that $z_n \to x$. Prove that $x_n \to x$.

(b) Give an example of a normed space X and a sequence $\{x_n\}_{n\in\mathbb{N}}$ such that every subsequence $\{y_n\}_{n\in\mathbb{N}}$ has a convergent subsequence $\{z_n\}_{n\in\mathbb{N}}$, yet $\{x_n\}_{n\in\mathbb{N}}$ does not converge. What hypothesis of part (a) does your sequence $\{x_n\}_{n\in\mathbb{N}}$ not satisfy?

1′.2.20. Let X be a normed space. Extend the definition of convergence to families indexed by a real parameter by declaring that if $x \in X$ and $x_t \in X$ for t in the interval $(0, c)$, where $c > 0$, then $x_t \to x$ as $t \to 0^+$ if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that $\|x - x_t\| < \varepsilon$ whenever $0 < t < \delta$. Show that $x_t \to x$ as $t \to 0^+$ if and only if $x_{t_k} \to x$ for every sequence of real numbers $\{t_k\}_{k\in\mathbb{N}}$ in $(0, c)$ that satisfy $t_k \to 0$.

1′.3 Open Sets

An *open ball* in a normed space is the set of all points that lie within a fixed distance from a central point x. These sets will appear frequently throughout the text, so we introduce a notation to represent them.

Definition 1′.3.1 (Open Balls). Let X be a normed space. If $x \in X$ and $r > 0$, then the *open ball in X with radius r centered at x* is

$$B_r(x) \;=\; \big\{y \in X : \|x - y\| < r\big\}. \qquad \diamondsuit \tag{1′.10}$$

We emphasize that "$r > 0$" means that r is a positive *real number*. In particular, every ball has a *finite* radius. We do not allow a ball to have an infinite radius.

We also emphasize that the definition of a ball in a given space *implicitly depends on the choice of norm*! In Figure 1′.6 we show the unit ball $B_1(0)$ in $\mathbb{R}^2$ with respect to each the norms $\|\cdot\|_1$, $\|\cdot\|_2$, and $\|\cdot\|_\infty$. While the colloquial meaning of the word "ball" suggests a sphere or disk, we can see in the figure that only the unit ball that is defined with respect to the Euclidean norm is "spherical" in the ordinary sense. Still, all of the sets depicted in Figure 1′.6 are open balls in the sense of Definition 1′.3.1, each corresponding to a different choice of norm on $\mathbb{R}^2$. Although the actual shape of a ball depends on the norm that we choose, for purposes of illustration we often depict them as if they looked like disks in $\mathbb{R}^2$ (for example, this is the case in Figure 1′.7).

We use open balls to define the meaning of *boundedness* in a normed space.

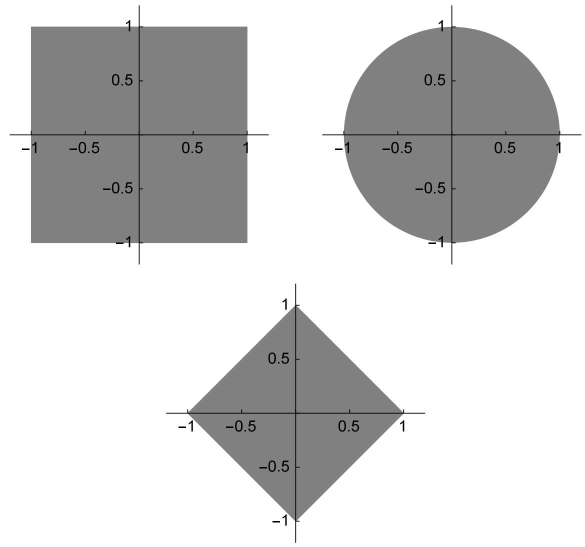

Fig. 1′.6 Unit open balls $B_1(0)$ with respect to different norms on $\mathbb{R}^2$. Top left: $\|\cdot\|_\infty$. Top right: $\|\cdot\|_2$ (the Euclidean norm). Bottom: $\|\cdot\|_1$.

Definition 1′.3.2 (Bounded Set). Let X be a normed space. We say that a set $E \subseteq X$ is *bounded* if it is contained in some open ball, i.e., if there exists some $x \in X$ and $r > 0$ such that $E \subseteq B_r(x)$. ♢

Next we use the open balls to define *open sets* in a normed space. According to the following definition, a set U is open if each point x in U has an open ball centered at x that is entirely contained within U.

Definition 1′.3.3 (Open Sets). Let U be a subset of a normed space X. We say that U is an *open subset* of X if for each point $x \in U$ there exists some $r > 0$ such that $B_r(x) \subseteq U$. ♢

We prove next that open balls are indeed open sets.

Lemma 1′.3.4. *If X is a normed space, then every open ball $B_r(x)$ is an open subset of X.*

Proof. Fix $x \in X$ and $r > 0$, and let y be any element of $B_r(x)$. We must show that there is some open ball $B_s(y)$ centered at y that is entirely contained in $B_r(x)$. We will show that the ball $B_s(y)$ with radius $s = r - \|x-y\|$ has this property (see Figure 1′.7). To prove this, choose any point $z \in B_s(y)$. Then $\|y-z\| < s$, so by applying the Triangle Inequality we see that

$$\|x-z\| \;\le\; \|x-y\| + \|y-z\| \;<\; \|x-y\| + s \;=\; r.$$

Hence $z \in B_r(x)$. Thus $B_s(y) \subseteq B_r(x)$, so $B_r(x)$ is an open set. □

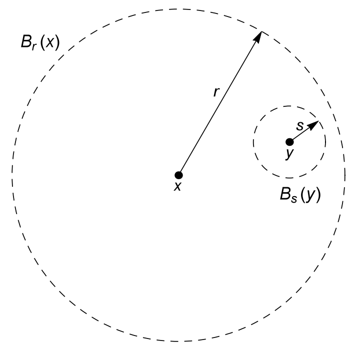

Fig. 1′.7 In order for $B_s(y)$ to fit inside $B_r(x)$, we need $\|x-y\|+s<r$.

We claim that every open set U is the union of some collection of open balls. To see this, recall that for each point $x \in U$ there must exist some radius $r_x > 0$ such that $B_{r_x}(x) \subseteq U$. This radius r_x will depend on the point x, but for each x we will have $x \in B_{r_x}(x) \subseteq U$. Consequently

$$U \;=\; \cup_{x\in U}\, \{x\} \;\subseteq\; \cup_{x\in U} B_{r_x}(x) \;\subseteq\; U,$$

and therefore $U = \cup_{x\in U} B_{r_x}(x)$. Thus every open set is a union of open balls.

The next lemma states three fundamental properties of open sets.

Lemma 1′.3.5. *If X is a normed space, then the following statements hold.*

(a) *X and $\varnothing$ are open subsets of X.*

(b) *Any union of open subsets of X is open. That is, if I is an index set and $\{U_i\}_{i\in I}$ is a collection of open subsets of X, then $\cup_{i\in I} U_i$ is an open set.*

(c) *The intersection of* finitely many *open subsets of X is open. That is, if $U_1,\dots,U_n$ are open subsets of X, then $U_1\cap\cdots\cap U_n$ is an open set.*

Proof. (a) If we choose any $x \in X$ then $B_r(x) \subseteq X$ for every $r > 0$, so X is open. On the other hand, the empty set is open because it contains no points, so it is vacuously true that "for each $x \in \varnothing$ there is some $r > 0$ such that $B_r(x) \subseteq \varnothing$."

(b) Suppose that $x \in \cup U_i$, where each U_i is open. Then $x \in U_j$ for some particular index $j \in I$. Since U_j is open there exists some $r > 0$ such that $B_r(x) \subseteq U_j$. Hence $B_r(x) \subseteq \cup U_i$, so $\cup U_i$ is open.

(c) Suppose that $x \in U \cap V$, where U and V are open. Then $x \in U$, so there is some $r > 0$ such that $B_r(x) \subseteq U$. Likewise, since $x \in V$ there is some $s > 0$ such that $B_s(x) \subseteq V$. Let t be the smaller of r and s. Then $B_t(x) \subseteq U$ and $B_t(x) \subseteq V$, so $B_t(x) \subseteq U\cap V$. Therefore $U\cap V$ is open. Using induction, this extends to the intersection of any finite number of open sets. ☐

Next we prove the *Hausdorff property* of normed spaces, which states that any two distinct points $x \neq y$ can be "separated" by open sets.

Lemma 1′.3.6 (Normed Spaces are Hausdorff). *If X is a normed space and $x \neq y$ are two distinct elements of X, then there exist disjoint open sets U and V such that $x \in U$ and $y \in V$.*

Proof. Suppose that $x \neq y$, and let $r = \|x-y\|/2$. If $z \in B_r(x) \cap B_r(y)$ then, by the Triangle Inequality,

$$\|x-y\| \;\le\; \|x-z\| + \|z-y\| \;<\; 2r \;=\; \|x-y\|,$$

which is a contradiction. Therefore $B_r(x) \cap B_r(y) = \varnothing$. Since open balls are open sets, the proof is finished by taking $U = B_r(x)$ and $V = B_r(y)$. □

We also define the interior of a set as follows.

Definition 1′.3.7 (Interior). Let X be a normed space. The *interior* of a set $E \subseteq X$, denoted E°, is the union of all of the open sets that are contained in E, i.e.,

$$E^\circ \;=\; \bigcup \{U \subseteq X \,:\, U \text{ is open and } U \subseteq E\}. \qquad \diamondsuit$$

According to Problem 1′.3.10, we have that E° is an open set, $E^\circ \subseteq E$, and if U is any open subset of E, then $U \subseteq E^\circ$. In this sense, E° is the "largest open set that is contained in E."

Since a normed space is a vector space, we can define lines, planes, and other related notions. In particular, if x and y are vectors in X, then the *line segment* joining x to y is the set of all points of the form $tx + (1-t)y$ where $0 \le t \le 1$.

Definition 1′.3.8 (Convex Set). We say that a subset K of a vector space X is *convex* if given any two points $x, y \in K$, the line segment joining x to y is entirely contained within K. That is, K is convex if

$$x, y \in K,\; 0 \le t \le 1 \qquad \Longrightarrow \qquad tx + (1-t)y \,\in\, K. \qquad \diamondsuit$$

All subspaces are convex by definition, and we prove now that every open ball $B_r(x)$ in a normed space X is convex.

Lemma 1′.3.9. *If X is a normed space, $x \in X$, and $r > 0$, then the open ball $B_r(x) = \{y \in X : \|x-y\| < r\}$ is convex.*

Proof. Choose any two points $y, z \in B_r(x)$ and fix $0 \le t \le 1$. Then

$$\begin{aligned} \big\|x - \big((1-t)y + tz\big)\big\| \;&=\; \big\|(1-t)(x-y) \,+\, t(x-z)\big\| \\ &\le\; (1-t)\,\|(x-y)\| \,+\, t\,\|x-z\| \\ &<\; (1-t)\,r \,+\, tr \;=\; r, \end{aligned}$$

so $(1-t)y + tz \in B_r(x)$. □

Problems

1′.3.10. Let A and B be subsets of a normed space X.

(a) Prove that A° is open, $A^\circ \subseteq A$, and if $U \subseteq A$ is open then $U \subseteq A^\circ$.

(b) Prove that A is open if and only if $A = A^\circ$.

(c) Prove that $(A \cap B)^\circ = A^\circ \cap B^\circ$.

(d) Show by example that $(A \cup B)^\circ$ need not equal $A^\circ \cup B^\circ$.

1′.3.11. For this problem we take scalars to be real.

(a) Let Q be the "open first quadrant" in $\mathbb{R}^d$, i.e.,

$$Q \;=\; \bigl\{x = (x_1, \dots, x_d) \in \mathbb{R}^d \,:\, x_1, \dots, x_d > 0\bigr\}.$$

Prove that Q is an open subset of $\mathbb{R}^d$.

(b) Let R be the "open first quadrant" in ℓ^1, i.e.,

$$R \;=\; \bigl\{x = (x_k)_{k\in\mathbb{N}} \in \ell^1 \,:\, x_k > 0 \text{ for every } k\bigr\}.$$

Prove that R is not an open subset of ℓ^1.

(c) Let S be the "open first quadrant" in $C[0,1]$, i.e.,

$$S \;=\; \bigl\{f \in C[0,1] \,:\, f(x) > 0 \text{ for all } x \in [0,1]\bigr\}.$$

Determine, with proof, whether S is an open subset of $C[0,1]$ with respect to the uniform norm. Hint: A continuous real-valued function on a closed finite interval achieves a maximum and a minimum on that interval.

(d) Same as part (c), except this time use the L^1-norm.

1′.3.12. Let $\{x_n\}_{n\in\mathbb{N}}$ be a sequence in a normed space X. Show that if $x_n \to x$, then either:

(a) there exists an open set U that contains x and there is some $N > 0$ such that $x_n = x$ for all $n > N$, or

(b) every open set U that contains x must also contain infinitely many *distinct* x_n (that is, the set $\{x_n : n \in \mathbb{N} \text{ and } x_n \in U\}$ contains infinitely many elements).

1′.3.13.* Prove that every open subset of $\mathbb{R}^d$ can be written as the union of *countably many* open balls.

1′.4 Closed Sets

The complements of the open sets are very important; we call these the *closed subsets* of X.

Definition $1'.4.1$. Let X be a normed space. We say that a set $F \subseteq X$ is *closed* if its complement $F^{\mathrm{C}} = X \backslash F$ is open. $\diamondsuit$

By taking complements in Lemma $1'.3.5$ we see that if X is a normed space then:

- the empty set and the entire space are closed,
- an *arbitrary intersection* of closed subsets of X is closed, and
- a *union of finitely many* closed subsets of X is closed.

Definition $1'.4.1$ is "indirect" in the sense that it is worded in terms of the complement of E rather than E itself. The next theorem gives a "direct" characterization of closed sets in terms of limits of elements of E.

Theorem $1'.4.2$. *If E is a subset of a normed space X, then the following two statements are equivalent.*

(a) *E is closed.*

(b) *If $\{x_n\}_{n\in\mathbb{N}}$ is a sequence of points in E and $x_n \to x \in X$, then $x \in E$.*

Proof. (a) $\Rightarrow$ (b). Assume that E is closed. Suppose that $x \in X$ is a limit of elements of E, i.e., there exist points $x_n \in E$ such that $x_n \to x$. Suppose that x did not belong to E. Then x belongs to E^{C}, which is an open set, so there must exist some $r > 0$ such that $B_r(x) \subseteq E^{\mathrm{C}}$. Since $x_n \in E$, none of the x_n belong to $B_r(x)$. This implies that $\|x - x_n\| \ge r$ for every n, which contradicts the assumption that $x_n \to x$. Therefore x must belong to E.

(b) $\Rightarrow$ (a). Suppose that statement (b) holds, but E is not closed. Then E^{C} is not open, so there must exist some point $x \in E^{\mathrm{C}}$ such that no open ball $B_r(x)$ centered at x is entirely contained in E^{C}. Considering $r = \frac{1}{n}$ in particular, this tells us that there must exist a point $x_n \in B_{1/n}(x)$ that is not in E^{C}. But then $x_n \in E$, and we have $\|x - x_n\| < \frac{1}{n}$, so $x_n \to x$. Statement (b) therefore implies that $x \in E$, which is a contradiction since x belongs to E^{C}. Consequently E must be closed. $\square$

In other words, Theorem $1'.4.2$ says that

> *a set E is closed if and only if the limit of every convergent sequence of points of E belongs to E.*

In practice this is often (though not always) the best way to prove that a particular set E is closed.

If X is a Banach space and Y is a subspace of X, then by restricting the norm on X to vectors in Y we obtain a norm on Y. When is Y complete with respect to this inherited norm? The following exercise addresses this.

Exercise $1'.4.3$. Let Y be a subspace of a Banach space X, and let the norm on Y be the restriction of the norm on X to the set Y. Prove that Y is a Banach space with respect to this norm if and only if Y is a closed subset of X. That is,

$$Y \text{ is complete} \quad \Longleftrightarrow \quad Y \text{ is closed.} \qquad \diamondsuit$$

To give an application, let c_0 be the set of all sequences of scalars that "vanish at infinity," i.e.,

$$c_0 \;=\; \Bigl\{x = (x_k)_{k\in\mathbb{N}} : \lim_{k\to\infty} x_k = 0\Bigr\}. \tag{1'.11}$$

The reader should check that we have the inclusions

$$c_{00} \;\subsetneq\; \ell^1 \;\subsetneq\; \ell^p \;\subsetneq\; c_0 \;\subsetneq\; \ell^\infty, \qquad \text{for every } 1 < p < \infty.$$

We will use Exercise $1'.4.3$ to prove that c_0 is a Banach space with respect to the norm of ℓ^∞.

Lemma $1'.4.4$. *c_0 is a closed subspace of ℓ^∞ with respect to $\|\cdot\|_\infty$, and hence is a Banach space with respect to that norm.*

Proof. We will show that c_0 is closed by proving that the ℓ^∞-norm limit of any sequence of elements of c_0 belongs to c_0. To do this, suppose that $\{x_n\}_{n\in\mathbb{N}}$ is a sequence of vectors from c_0 and there exists some vector $x \in \ell^\infty$ such that $\|x-x_n\|_\infty \to 0$. For convenience, denote the components of x_n and x by $x_n = \bigl(x_n(k)\bigr)_{k\in\mathbb{N}}$ and $x = \bigl(x(k)\bigr)_{k\in\mathbb{N}}$, respectively. We must show that $x \in c_0$, which means that we must prove that $x(k) \to 0$ as $k \to \infty$.

Fix any $\varepsilon > 0$. Since $x_n \to x$ in ℓ^∞-norm, there exists some n such that $\|x - x_n\|_\infty < \frac{\varepsilon}{2}$ (in fact, this will be true for all large enough n, but we need only one n for this proof). Since $x_n \in c_0$, we know that $x_n(k) \to 0$ as $k \to \infty$. Therefore there exists some $K > 0$ such that $|x_n(k)| < \frac{\varepsilon}{2}$ for all $k \ge K$. Consequently, for any $k \ge K$ we have

$$|x(k)| \;\le\; |x(k) - x_n(k)| \,+\, |x_n(k)| \;<\; \|x - x_n\|_\infty + \frac{\varepsilon}{2} \;<\; \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \;=\; \varepsilon.$$

This shows that $x(k) \to 0$ as $k \to \infty$, so x belongs to c_0. Therefore Theorem $1'.4.2$ implies that c_0 is a closed subset of ℓ^∞, and hence Exercise $1'.4.3$ implies that c_0 is a Banach space with respect to the sup-norm. □

Problems

$1'.4.5$. Let X be a normed space. Define the *distance* from a point $x \in X$ to a subset $A \subseteq X$ to be $\operatorname{dist}(x,A) = \inf\bigl\{\|x-y\| : y \in A\bigr\}$. Prove the following statements.

(a) If A is closed, then $x \in A$ if and only if $\operatorname{dist}(x,A) = 0$.

(b) $\operatorname{dist}(x,A) \le \|x-y\| + \operatorname{dist}(y,A)$ for all $x, y \in X$.

(c) $\bigl|\operatorname{dist}(x,A) - \operatorname{dist}(y,A)\bigr| \le \|x-y\|$ for all $x, y \in X$.

1′.4.6. Let X be a *complete* normed space (i.e., a Banach space), and let E be a subset of X. Prove that the following two statements are equivalent.

(a) E is closed.

(b) E is complete, i.e., every Cauchy sequence of points in E converges to a point of E.

1′.4.7. Let X be a normed space. We define the *diameter* of $E \subseteq X$ to be $\operatorname{diam}(E) = \sup\{\|x-y\| : x, y \in E\}$.

(a) Suppose that X is complete and $F_1 \supseteq F_2 \supseteq \cdots$ is a nested decreasing sequence of closed nonempty subsets of X such that $\operatorname{diam}(F_n) \to 0$. Prove that there exists some $x \in X$ such that $\cap F_n = \{x\}$.

(b) Show by example that the conclusion of part (a) can fail if X is not complete or if the F_n are not closed.

1′.4.8. For this problem we take scalars to be real.

(a) Let Q be the "closed first quadrant" in $\mathbb{R}^d$, i.e.,

$$Q = \{x = (x_1, \dots, x_d) \in \mathbb{R}^d : x_1, \dots, x_d \geq 0\}.$$

Prove that Q is a closed subset of $\mathbb{R}^d$.

(b) Let R be the "closed first quadrant" in ℓ^1, i.e.,

$$R = \{x = (x_k)_{k \in \mathbb{N}} \in \ell^1 : x_k \geq 0 \text{ for every } k\}.$$

Determine, with proof, whether R is a closed subset of ℓ^1.

(c) Let S be the "closed first quadrant" in $C[0,1]$, i.e.,

$$S = \{f \in C[0,1] : f(x) \geq 0 \text{ for all } x \in [0,1]\}.$$

Determine, with proof, whether S is a closed subset of $C[0,1]$ with respect to the uniform norm.

1′.4.9. Let E be the subset of ℓ^1 consisting of sequences whose even components are all zero: $E = \{x = (x_k)_{k \in \mathbb{N}} : x_{2j} = 0 \text{ for all } j \in \mathbb{N}\}$.

(a) Prove that E is a proper, closed *subspace* of ℓ^1 (not just a subset). Also prove that E is not an open set.

(b) Is c_{00} a closed subspace of ℓ^1? Is it an open subspace?

(c) Let $N > 0$ be a *fixed* positive integer, and let

$$S_N = \{x = (x_1, \dots, x_N, 0, 0, \dots) : x_1, \dots, x_N \text{ are scalars}\}.$$

Prove that S_N is a closed subspace of ℓ^1. Is it open? What is the union of the S_N over $N \in \mathbb{N}$?

1′.4.10. Show that if S is an open *subspace* of a normed space X then $S = X$.

$1'.5$ Examples: The ℓ^p Spaces

We introduced the space ℓ^1 in Section $1'.1.1$. Now we will define an entire family of related spaces.

Given a finite number $1 \le p < \infty$, we say that a sequence of scalars $x = (x_k)_{k\in\mathbb{N}}$ is *p-summable* if $\sum_{k=1}^{\infty} |x_k|^p < \infty$. In this case, we set

$$\|x\|_p \;=\; \|(x_k)_{k\in\mathbb{N}}\|_p \;=\; \biggl(\sum_{k=1}^{\infty} |x_k|^p\biggr)^{1/p}. \tag{1'.12}$$

If the sequence x is not p-summable then we set $\|x\|_p = \infty$. If $p = 1$ then we usually just write "summable" instead of "1-summable," and for $p = 2$ we usually write "square summable" instead of "2-summable."

We also allow $p = \infty$, although in this case the definition is different. As we declared in Example $1'.1.6$, the *ℓ^∞-norm* (or the *sup-norm*) of a sequence $x = (x_k)_{k\in\mathbb{N}}$ is

$$\|x\|_\infty \;=\; \sup_{k\in\mathbb{N}} |x_k|.$$

The sup-norm of x is finite if and only if x is a bounded sequence.

If $1 \le p \le \infty$ then we call $\|x\|_p$ the *ℓ^p-norm* of the sequence x (though we should note that have not yet established that it is a norm on any space). We collect the sequences that have *finite* ℓ^p-norm to form a space that we call ℓ^p:

$$\ell^p \;=\; \bigl\{x = (x_k)_{k\in\mathbb{N}} : \|x\|_p < \infty\bigr\}. \tag{1'.13}$$

Thus, if $1 \le p < \infty$ then ℓ^p is the set of all p-summable sequences, while ℓ^∞ is the set of all bounded sequences.

Each element of ℓ^p is a sequence of scalars, but because we will often need to consider *sequences of elements* of ℓ^p we will often refer to $x \in \ell^p$ as a *point* or *vector* in ℓ^p. For example, the standard basis $\{\delta_n\}_{n\in\mathbb{N}}$ is a sequence of vectors in ℓ^p, and each particular vector δ_n is a sequence of scalars.

The ℓ^p spaces do not all contain the same vectors. For example, the sequence $x = \bigl(\frac{1}{k}\bigr)_{k\in\mathbb{N}} = \bigl(1, \frac{1}{2}, \frac{1}{3}, \dots\bigr)$ belongs to ℓ^p for every $1 < p \le \infty$, but it does not belong to ℓ^1. Problem $1'.5.8$ shows that the ℓ^p spaces are nested in the sense that

$$1 \;\le\; p \;<\; q \;\le\; \infty \qquad \Longrightarrow \qquad \ell^p \subsetneq \ell^q. \tag{1'.14}$$

It is clear that $\|\cdot\|_p$ satisfies the nonnegativity, homogeneity, and uniqueness properties of a norm, but it is not obvious whether the Triangle Inequality is satisfied. We will prove that $\|\cdot\|_p$ is a norm on ℓ^p, but first we need to establish a fundamental result known as *Hölder's Inequality*. This inequality gives a relation between ℓ^p and $\ell^{p'}$, where p' is the *dual index* to p. If $1 < p < \infty$, then p' is the unique number that satisfies

$$\frac{1}{p} + \frac{1}{p'} \;=\; 1. \tag{1'.15}$$

Explicitly, $p' = p/(p-1)$ if $1 < p < \infty$. If we adopt the standard real analysis convention that

$$\frac{1}{\infty} = 0, \tag{1'.16}$$

then equation (1′.15) still has a unique solution if $p = 1$ ($p' = \infty$ in that case) or if $p = \infty$ (in which case $p' = 1$). We take these to be the definition of p' for these endpoint cases. Hence some particular examples of dual indices are

$$1' = \infty, \quad \left(\frac{4}{3}\right)' = 4, \quad \left(\frac{3}{2}\right)' = 3, \quad 2' = 2, \quad 3' = \frac{3}{2}, \quad 4' = \frac{4}{3}, \quad \infty' = 1.$$

We have $(p')' = p$ for each $1 \le p \le \infty$. If $1 \le p \le 2$ then $2 \le p' \le \infty$, while if $2 \le p \le \infty$ then $1 \le p' \le 2$ (hence analysts often consider $p = 2$ to be the "midpoint" of the extended interval $[1, \infty]$).

To motivate the next lemma, recall the arithmetic-geometric mean inequality, $\sqrt{ab} \le (a+b)/2$ for $a, b \ge 0$. Replacing a with a^2 and b with b^2 gives

$$ab \le \frac{a^2}{2} + \frac{b^2}{2}, \qquad \text{for all } a, b \ge 0.$$

We generalize this inequality to values of p between 1 and ∞ as follows (one proof of Lemma 1′.5.1 is outlined in Problem 1′.5.10).

Lemma 1′.5.1. *If $1 < p < \infty$, then*

$$ab \le \frac{a^p}{p} + \frac{b^{p'}}{p'}, \qquad \text{for all } a, b \ge 0. \qquad \diamondsuit \tag{1'.17}$$

Now we prove Hölder's Inequality. In this result, given two sequences $x = (x_k)_{k\in\mathbb{N}}$ and $y = (y_k)_{k\in\mathbb{N}}$ we let xy be the sequence obtained by multiplying corresponding components of x and y, i.e.,

$$xy = (x_k y_k)_{k\in\mathbb{N}} = (x_1y_1, x_2y_2, x_3y_3, \dots).$$

Theorem 1′.5.2 (Hölder's Inequality). *Fix $1 \le p \le \infty$ and let p' be the dual index to p. If $x = (x_k)_{k\in\mathbb{N}} \in \ell^p$ and $y = (y_k)_{k\in\mathbb{N}} \in \ell^{p'}$, then the sequence $xy = (x_k y_k)_{k\in\mathbb{N}}$ belongs to ℓ^1, and*

$$\|xy\|_1 \le \|x\|_p \, \|y\|_{p'}. \tag{1'.18}$$

If $1 < p < \infty$, then equation (1′.18) is

$$\sum_{k=1}^{\infty} |x_k y_k| \le \left(\sum_{k=1}^{\infty} |x_k|^p\right)^{1/p} \left(\sum_{k=1}^{\infty} |y_k|^{p'}\right)^{1/p'}. \tag{1'.19}$$

If $p = 1$, then equation (1′.18) is

$$\sum_{k=1}^{\infty} |x_k y_k| \;\le\; \biggl(\sum_{k=1}^{\infty} |x_k|\biggr) \biggl(\sup_{k\in\mathbb{N}} |y_k|\biggr). \tag{1'.20}$$

If $p = \infty$, *then equation* (1′.18) *is*

$$\sum_{k=1}^{\infty} |x_k y_k| \;\le\; \biggl(\sup_{k\in\mathbb{N}} |x_k|\biggr) \biggl(\sum_{k=1}^{\infty} |y_k|\biggr). \tag{1'.21}$$

Proof. Case $p = 1$. In this case we have $p' = \infty$, and we assign the proof of equation (1′.20) as an exercise. The case $p = \infty$ is entirely symmetrical, because $p' = 1$ when $p = \infty$.

Case $1 < p < \infty$. If either x or y is the zero sequence, then equation (1′.19) holds trivially, so we may assume that $x \ne 0$ and $y \ne 0$ (i.e., neither x nor y is the sequence of all zeros, and hence $\|x\|_p > 0$ and $\|y\|_{p'} > 0$).

Suppose first that $x \in \ell^p$ and $y \in \ell^{p'}$ are unit vectors in their respective spaces, i.e., $\|x\|_p = 1$ and $\|y\|_{p'} = 1$. Then by applying equation (1′.17) we see that

$$\begin{aligned} \|xy\|_1 \;=\; \sum_{k=1}^{\infty} |x_k y_k| \;&\le\; \sum_{k=1}^{\infty} \biggl(\frac{|x_k|^p}{p} + \frac{|y_k|^{p'}}{p'}\biggr) \\ &=\; \frac{\|x\|_p^p}{p} + \frac{\|y\|_{p'}^{p'}}{p'} \;=\; \frac{1}{p} + \frac{1}{p'} \;=\; 1. \end{aligned} \tag{1'.22}$$

Now let x be any nonzero sequence in ℓ^p, and let y be any nonzero sequence in $\ell^{p'}$. Let

$$u \;=\; \frac{x}{\|x\|_p} \qquad \text{and} \qquad v \;=\; \frac{y}{\|y\|_{p'}}.$$

Then u is a unit vector in ℓ^p, and v is a unit vector in $\ell^{p'}$, so equation (1′.22) implies that $\|uv\|_1 \le 1$. However,

$$uv \;=\; \frac{xy}{\|x\|_p \, \|y\|_{p'}},$$

so by homogeneity we obtain

$$\frac{\|xy\|_1}{\|x\|_p \, \|y\|_{p'}} \;=\; \biggl\| \frac{x}{\|x\|_p} \, \frac{y}{\|y\|_{p'}} \biggr\|_1 \;=\; \|uv\|_1 \;\le\; 1.$$

Rearranging yields $\|xy\|_1 \le \|x\|_p \, \|y\|_{p'}$. □

We will use Hölder's Inequality to prove that $\|\cdot\|_p$ satisfies the Triangle Inequality (which for $\|\cdot\|_p$ is often called *Minkowski's Inequality*).

Theorem 1′.5.3 (Minkowski's Inequality). *If* $1 \le p \le \infty$, *then for all* $x, y \in \ell^p$ *we have*

$$\|x+y\|_p \;\le\; \|x\|_p + \|y\|_p. \tag{1′.23}$$

If $1 \le p < \infty$, *then equation* (1′.23) *is*

$$\left(\sum_{k=1}^{\infty} |x_k + y_k|^p\right)^{1/p} \;\le\; \left(\sum_{k=1}^{\infty} |x_k|^p\right)^{1/p} \;+\; \left(\sum_{k=1}^{\infty} |y_k|^p\right)^{1/p},$$

while if $p = \infty$, *then equation* (1′.23) *is*

$$\sup_{k\in\mathbb{N}} |x_k + y_k| \;\le\; \Big(\sup_{k\in\mathbb{N}} |x_k|\Big) \;+\; \Big(\sup_{k\in\mathbb{N}} |y_k|\Big).$$

Proof. Assume first that $1 < p < \infty$, and let $x = (x_k)_{k\in\mathbb{N}}$ and $y = (y_k)_{k\in\mathbb{N}}$ be given. Since $p > 1$ we have $p - 1 > 0$, so we can write

$$\begin{aligned}
\|x+y\|_p^p \;&=\; \sum_{k=1}^{\infty} |x_k + y_k|^p \\
&=\; \sum_{k=1}^{\infty} |x_k + y_k|\, |x_k + y_k|^{p-1} \\
&\le\; \sum_{k=1}^{\infty} |x_k|\, |x_k + y_k|^{p-1} \;+\; \sum_{k=1}^{\infty} |y_k|\, |x_k + y_k|^{p-1} \;=\; S_1 + S_2.
\end{aligned}$$

To simplify the series S_1, let $z_k = |x_k + y_k|^{p-1}$, so that

$$S_1 \;=\; \sum_{k=1}^{\infty} |x_k|\, |x_k + y_k|^{p-1} \;=\; \sum_{k=1}^{\infty} |x_k|\, |z_k|.$$

We apply Hölder's Inequality, and then substitute $p' = p/(p-1)$, to compute as follows:

$$\begin{aligned}
S_1 \;=\; \sum_{k=1}^{\infty} |x_k|\, |z_k| \;&\le\; \left(\sum_{k=1}^{\infty} |x_k|^p\right)^{1/p} \left(\sum_{k=1}^{\infty} |z_k|^{p'}\right)^{1/p'} && \text{(Hölder)} \\
&=\; \left(\sum_{k=1}^{\infty} |x_k|^p\right)^{1/p} \left(\sum_{k=1}^{\infty} |x_k + y_k|^p\right)^{(p-1)/p} && \text{(substitute)} \\
&=\; \|x\|_p\, \|x+y\|_p^{p-1}.
\end{aligned}$$

A similar calculation shows that $S_2 \le \|y\|_p\, \|x+y\|_p^{p-1}$. Combining these inequalities, we see that

$$\|x+y\|_p^p \;\le\; S_1 + S_2 \;\le\; \|x+y\|_p^{p-1} \big(\|x\|_p + \|y\|_p\big).$$

If $x+y$ is not the zero vector then we can divide both sides by $\|x+y\|_p^{p-1}$ to obtain $\|x+y\|_p \le \|x\|_p + \|y\|_p$. On the other hand, this inequality holds trivially if $x+y=0$, so we are done.

Exercise: Complete the proof for the cases $p=1$ and $p=\infty$. □

Since $\|\cdot\|_p$ also satisfies the nonnegativity, homogeneity, and uniqueness requirements of a norm, we obtain the following theorem.

Theorem 1′.5.4. *If $1 \le p \le \infty$, then $\|\cdot\|_p$ is a norm on ℓ^p.* ♢

We saw in Theorem 1′.2.10 that the space ℓ^1 is complete and hence is a Banach space. A similar argument, which we assign as the following exercise, shows that ℓ^p is a Banach space for all indices in the range $1 \le p \le \infty$.

Exercise 1′.5.5. Prove that ℓ^p is a Banach space for each $1 \le p \le \infty$. ♢

By making minor changes to the arguments above, we obtain a corresponding family of norms on $\mathbb{R}^d$. An entirely similar result holds for $\mathbb{C}^d$ using complex scalars.

Exercise 1′.5.6. Given $1 \le p < \infty$ define

$$\|x\|_p \;=\; \big(|x_1|^p + \cdots + |x_d|^p\big)^{1/p}, \qquad x \in \mathbb{R}^d,$$

while for $p=\infty$ set

$$\|x\|_\infty \;=\; \max\big\{|x_1|, \dots, |x_d|\big\}, \qquad x \in \mathbb{R}^d.$$

Prove that $\|\cdot\|_p$ is a norm on $\mathbb{R}^d$ for each $1 \le p \le \infty$, and $\mathbb{R}^d$ is a Banach space with respect to each of these norms. ♢

Problems

1′.5.7. Given $1 \le p \le \infty$, prove the following statements.

(a) If $x_n \to x$ in ℓ^p, then x_n converges componentwise to x.

(b) There exist vectors $x_n, x \in \ell^p$ such that x_n converges componentwise to x but x_n does not converge to x in ℓ^p-norm.

1′.5.8. Given fixed indices $1 \le p < q \le \infty$, prove the following statements.

(a) $\ell^p \subseteq \ell^q$.

(b) There exists a sequence x that belongs to ℓ^q but not ℓ^p.

(c) $\|x\|_q \le \|x\|_p$ for each $x \in \ell^p$.

1′.5.9. Prove that if $x \in \ell^q$ for some finite q, then $\lim_{p\to\infty} \|x\|_p = \|x\|_\infty$. Exhibit a sequence $x \in \ell^\infty$ for which this equality fails.

1′.5.10. (a) Show that if $0 < \theta < 1$, then $t^\theta \le \theta t + (1-\theta)$ for all $t \ge 0$, and equality holds if and only if $t = 1$.

(b) Suppose that $1 < p < \infty$ and $a, b \ge 0$. Apply part (a) with $t = a^p b^{-p'}$ and $\theta = 1/p$ to show that

$$ab \;\le\; \frac{a^p}{p} + \frac{b^{p'}}{p'}.$$

(c) Prove that equality holds in part (b) if and only if $b = a^{p-1}$.

1′.5.11. Given $1 < p < \infty$, show that equality holds in Hölder's Inequality if and only if there exist scalars α, β, not both zero, such that $\alpha\, |x_k|^p = \beta\, |y_k|^{p'}$ for each $k \in \mathbb{N}$. What about the cases $p = 1$ or $p = \infty$?

1′.5.12. For each $n \in \mathbb{N}$, let $y_n = (1, \frac{1}{2}, \dots, \frac{1}{n}, 0, 0, \dots)$. Note that $y_n \in \ell^1$ for every n.

(a) Assume that the norm on ℓ^1 is its usual norm $\|\cdot\|_1$. Prove that $\{y_n\}_{n\in\mathbb{N}}$ is not a Cauchy sequence in ℓ^1 with respect to this norm.

(b) Now assume that the norm on ℓ^1 is $\|\cdot\|_2$. Prove that $\{y_n\}_{n\in\mathbb{N}}$ is a Cauchy sequence in ℓ^1 with respect to $\|\cdot\|_2$. Even so, prove that there is no vector $y \in \ell^1$ such that $\|y - y_n\|_2 \to 0$. Conclude that ℓ^1 is not complete with respect to the norm $\|\cdot\|_2$.

1′.5.13. Fix $1 \le p < \infty$.

(a) Let $x = (x_k)_{k\in\mathbb{N}}$ be a sequence of scalars that decays on the order of $k^{-\alpha}$ where $\alpha > 1/p$. That is, assume that $\alpha > 1/p$ and there exists a constant $C > 0$ such that

$$|x_k| \;\le\; C\,k^{-\alpha} \quad \text{for all } k \in \mathbb{N}. \tag{1′.24}$$

Show that $x \in \ell^p$.

(b) Set $\alpha = 1/p$. Exhibit a sequence $x \notin \ell^p$ that satisfies equation (1′.24) for some $C > 0$, and another sequence $x \in \ell^p$ that satisfies equation (1′.24) for some $C > 0$.

(c) Given $\alpha > 0$, show that there exists a sequence $x = (x_k)_{k\in\mathbb{N}} \in \ell^p$ such that there is *no* constant $C > 0$ that satisfies equation (1′.24). Conclude that no matter how small we choose α, there exist sequences in ℓ^p whose decay rate is slower than $k^{-\alpha}$.

(d) Suppose that the components of $x = (x_k)_{k\in\mathbb{N}} \in \ell^p$ are nonnegative and *monotonically decreasing*. Show that there exists some $\alpha \ge 1/p$ and some $C > 0$ such that equation (1′.24) holds.

Hint: Consider $\sum_{k=n+1}^{2n} |x_k|^p$.

1′.5.14. Prove the following generalization of Hölder's Inequality. Assume that $1 \le p,\, q,\, r \le \infty$ satisfy $\frac{1}{p} + \frac{1}{q} = \frac{1}{r}$. Given $x = (x_k)_{k\in\mathbb{N}} \in \ell^p$ and $y = (y_k)_{k\in\mathbb{N}} \in \ell^q$, prove that $xy = (x_k y_k)_{k\in\mathbb{N}}$ belongs to ℓ^r, and

$$\|xy\|_r \;\le\; \|x\|_p\, \|y\|_q.$$

Hint: Consider $w = (|x_k|^r)_{k\in\mathbb{N}}$ and $z = (|y_k|^r)_{k\in\mathbb{N}}$. Be careful with the endpoint cases.

1′.5.15. We say that a nonempty set X (not necessarily a vector space) is a *metric space* if for each $x,\, y \in X$ there exists a real number $\mathrm{d}(x,y)$ such that for all $x,\, y,\, z \in X$ we have:

i. *Nonnegativity*: $\mathrm{d}(x,y) \ge 0$,

ii. *Symmetry*: $\mathrm{d}(x,y) = \mathrm{d}(y,x)$,

iii. *The Triangle Inequality*: $\mathrm{d}(x,y) \le \mathrm{d}(x,z) + \mathrm{d}(z,y)$, and

iv. *Uniqueness*: $\mathrm{d}(x,y) = 0$ if and only if $x = y$.

In this case we call d a *metric* on X, and we refer to $\mathrm{d}(x,y)$ as the *distance* between x and y.

(a) Show that if X is a normed vector space, then $\mathrm{d}(x,y) = \|x-y\|$ is a metric on X.

(b) Make a definition of convergent and Cauchy sequences in a metric space, and show that every convergent sequence is Cauchy.

(c) Define a metric space to be *complete* if every Cauchy sequence in X converges to an element of X. Let $X = \mathbb{Q}$ and set $\mathrm{d}(x,y) = |x-y|$, the absolute value of the difference of x and y. Show that d is a metric on $\mathbb{Q}$, but $\mathbb{Q}$ is incomplete with respect to this metric.

1′.5.16. Fix $0 < p < 1$. Given a sequence $x = (x_k)_{k\in\mathbb{N}}$, define $\|x\|_p$ just as we did before, i.e., $\|x\|_p = \|(x_k)\|_{\ell^p} = \Bigl(\sum_k |x_k|^p\Bigr)^{1/p}$, and let ℓ^p be the set of all sequences such that $\|x\|_p < \infty$.

(a) Show $\|\cdot\|_p$ fails the Triangle Inequality, hence is not a norm on ℓ^p.

(b) Prove that $(1+t)^p \le 1 + t^p$ for all $t > 0$.

(c) Prove that if $a,\, b > 0$, then $(a+b)^p \le a^p + b^p$.

(d) Show that $\|x+y\|_p^p \le \|x\|_p^p + \|y\|_p^p$, and use this to prove that ℓ^p is a vector space and $\mathrm{d}(x,y) = \|x-y\|_p^p$ is a metric on ℓ^p.

(e) Let $B = \{x \in \ell^p : \mathrm{d}(x,0) < 1\}$ be the "open unit ball" with respect to the metric d, and show that B is not convex. Use this to show that there is no norm $\|\cdot\|$ on ℓ^p such that $\mathrm{d}(x,y) = \|x-y\|$.

1′.6 Compact Sets in Normed Spaces

A *compact set* is a special type of closed set. The abstract definition of a compact set is phrased in terms of coverings by open sets, but we will also derive several equivalent reformulations of the definition for subsets of normed spaces. One of these will show if K is a compact set, then every infinite sequence $\{x_n\}_{n\in\mathbb{N}}$ of points of K has a subsequence that converges to an element of K. We will also prove that in the Euclidean space $\mathbb{R}^d$, a subset is compact if and only if it is both closed and bounded. However, in other normed spaces the question of whether a set is compact or not can be a more subtle issue. For example, we will see that the closed unit disk in ℓ^1 is a *closed and bounded set that is not compact*!

By a *cover* of a set S, we mean a collection of sets $\{E_i\}_{i\in I}$ whose union contains S. If each set E_i is open, then we call $\{E_i\}_{i\in I}$ an *open cover* of S. The index set I may be finite or infinite (even uncountable). If I is finite then we call $\{E_i\}_{i\in I}$ a *finite cover* of S. Thus a *finite open cover* of S is a collection of finitely many open sets whose union contains S. Also, if $\{E_i\}_{i\in I}$ is a cover of S, then a *finite subcover* is a collection of finitely many of the E_i, say $E_{i_1}, \dots, E_{i_N}$, that still cover S.

Definition $1'.6.1$ (Compact Set). A subset K of a normed space X is *compact* if every open cover of K contains a finite subcover. Stated precisely, K is compact if it is the case that whenever

$$K \subseteq \bigcup_{i\in I} U_i,$$

where $\{U_i\}_{i\in I}$ is *any* collection of open subsets of X, then there exist *finitely* many indices $i_1, \dots, i_N \in I$ such that

$$K \subseteq \bigcup_{k=1}^{N} U_{i_k}. \qquad \diamondsuit$$

In order for a set K to be called *compact*, it must be the case that *every* open cover of K has a finite subcover. If there is even one particular open covering of K that does not have a finite subcover, then K is not compact.

Example $1'.6.2$. (a) Consider the interval $(0,1]$ in the real line. For each integer $k\in\mathbb{N}$, let U_k be the open interval $U_k = \left(\frac{1}{k}, 2\right)$. Then $\{U_k\}_{k\in\mathbb{N}}$ is an open cover of $(0,1]$, but we cannot cover the interval $(0,1]$ using only finitely many of the sets U_k (why not?). Consequently the open cover $\{U_k\}_{k\in\mathbb{N}}$ contains no finite subcover, so $(0,1]$ is not compact.

Even though $(0,1]$ is not compact, there do exist *some* open coverings that have finite subcoverings. For example, if we set $U_1 = (0,1)$ and $U_2 = (0,2)$, then $\{U_1, U_2\}$ is an open cover of $(0,1]$, and $\{U_2\}$ is a finite subcover. However, as we showed above, it is not true that *every* open cover of $(0,1]$ contains a finite subcover. This is why $(0,1]$ is not compact.

(b) Now consider the interval $[1,\infty)$. Even though this interval is a closed subset of $\mathbb{R}$, we will prove that it is not compact. To see this, set $U_k = (k-1, k+1)$ for each $k\in\mathbb{N}$. Then $\{U_k\}_{k\in\mathbb{N}}$ is an open cover of $[1,\infty)$, but we cannot cover $[0, \infty)$ using only finitely many of the intervals U_k. Hence there is at least one open cover of $[0, \infty)$ that has no finite subcover, so this set is not compact. ◇

We prove next that all compact subsets of a normed space are both closed and bounded.

Lemma 1′.6.3. *If K is a compact subset of a normed space X, then K is closed and bounded.*

Proof. Suppose that K is compact. Then the union of the open balls $B_n(0)$ over $n \in \mathbb{N}$ is all of X, so $\{B_n(0)\}_{n\in\mathbb{N}}$ an open cover of K. This cover must have a finite subcover, say

$$\{B_{n_1}(0),\, B_{n_2}(0),\, \dots,\, B_{n_M}(0)\}.$$

By reordering these sets if necessary we can assume that $n_1 < \cdots < n_M$. With this ordering the balls are *nested*:

$$B_{n_1}(0) \subseteq B_{n_2}(0) \subseteq \cdots \subseteq B_{n_M}(0).$$

Since K is contained in the union of these balls it is therefore contained in the ball with largest radius, i.e., $K \subseteq B_{n_M}(0)$. According to Definition 1′.3.2, this says that K is bounded.

It remains to show that K is closed. If $K = X$ then we are done, so assume that $K \neq X$. Fix any point y in $K^{\mathrm{C}} = X \backslash K$. If x is a point in K then $x \neq y$, so by the *Hausdorff property* stated in Lemma 1′.3.6 there exist disjoint open sets U_x and V_x such that $x \in U_x$ and $y \in V_x$. The collection $\{U_x\}_{x\in K}$ is an open cover of K, so it must contain some finite subcover. That is, there must exist finitely many points $x_1, \dots, x_N \in K$ such that

$$K \subseteq U_{x_1} \cup \cdots \cup U_{x_N}. \tag{1′.25}$$

Each V_{x_j} is disjoint from U_{x_j}, so it follows from equation (1′.25) that the set

$$V = V_{x_1} \cap \cdots \cap V_{x_N}$$

is entirely contained in the complement of K. Thus, V is an open set that satisfies $y \in V \subseteq K^{\mathrm{C}}$. This shows that K^{C} is open, and so K is closed. □

While every compact subset of a normed space is closed and bounded, there can exist sets that are closed and bounded but not compact. For example, we will see in Example 1′.6.6 that the "closed unit disk" in ℓ^1 is closed and bounded but *not compact*. On the other hand, we prove next that, *in the Euclidean space $\mathbb{R}^d$*, a set is compact *if and only if* it closed and bounded. An entirely similar result holds for $\mathbb{C}^d$ using complex scalars.

Theorem 1′.6.4 (Heine–Borel Theorem). *If $K \subseteq \mathbb{R}^d$, then K is compact if and only if K is closed and bounded.*

Proof. $\Rightarrow$. According to Lemma $1'.6.3$, compact sets are closed and bounded.

$\Leftarrow$. First we will show that the closed cube $Q = [-R, R]^d$ is a compact subset of $\mathbb{R}^d$. If Q were not compact, then there would exist an open cover $\{U_i\}_{i \in I}$ of Q that has no finite subcover. By bisecting each side of Q, we can divide Q into 2^d closed subcubes whose union is Q. At least one of these subcubes cannot be covered by finitely many of the sets U_i (why not?). Call that subcube Q_1. Then subdivide Q_1 into 2^d closed subcubes. At least one of these, which we call Q_2, cannot be covered by finitely many U_i. Continuing in this way we obtain nested decreasing closed cubes $Q \supseteq Q_1 \supseteq Q_2 \supseteq \cdots$. Appealing to Problem $1'.6.14$, $\bigcap Q_k$ is nonempty. Hence there is a point $x \in Q$ that belongs to every Q_k. This point must belong to some set U_i (since these sets cover Q), and since U_i is open there is some $r > 0$ such that $B_r(x) \subseteq U_i$. However, the sidelengths of the cubes Q_k decrease to zero as k increases, so if we choose k large enough then we will have $x \in Q_k \subseteq B_r(x) \subseteq U_i$ (check this!). Hence Q_k is covered by a single set U_i, which is a contradiction. Therefore Q must be compact.

Now let K be an arbitrary closed and bounded subset of $\mathbb{R}^d$. Then K is contained in $Q = [-R, R]^d$ for some $R > 0$. Hence K is a closed subset of the compact set Q. Therefore K is compact by Problem $1'.6.12$. □

We will give several equivalent reformulations of compactness in terms of the following concepts.

Definition $1'.6.5$. Let E be a subset of a normed space X.

(a) E is *sequentially compact* if every sequence $\{x_n\}_{n \in \mathbb{N}}$ of points from E contains a convergent subsequence $\{x_{n_k}\}_{k \in \mathbb{N}}$ whose limit belongs to E.

(b) E is *totally bounded* if given any $r > 0$ we can cover E by finitely many open balls of radius r. That is, for each $r > 0$ there must exist finitely many points $x_1, \dots, x_N \in X$ such that

$$E \subseteq \bigcup_{k=1}^{N} B_r(x_k). \qquad \diamondsuit$$

Here are some examples illustrating sequential compactness.

Example $1'.6.6$. (a) The interval $(0, 1]$ is not sequentially compact, because the sequence $\left\{\frac{1}{n}\right\}_{n \in \mathbb{N}}$ is contained in this interval, but it (and every subsequence) converges to 0, which does not belong to the interval.

(b) Suppose that E is a subset of normed space that is not closed. Then E does not contain every limit of elements of E. Hence there must be some point $x \in X \setminus E$ for which there exist points $x_n \in E$ such that $x_n \to x$. Hence $\{x_n\}_{n \in \mathbb{N}}$ is an infinite sequence in E that converges to a point outside of E. Consequently (why?), no subsequence can converge to an element of E, so E is not sequentially compact. Therefore we have proved, by a contrapositive argument, that all sequentially compact sets are closed.

(c) The interval $[1,\infty)$ is closed, but it is not sequentially compact because the sequence $\mathbb{N} = \{1,2,3,\dots\}$ has no convergent subsequences.

(d) Let $D = \{x \in \ell^1 : \|x\|_1 \le 1\}$ be the closed unit disk in ℓ^1. This set is both closed and bounded, but it is not sequentially compact because the standard basis $\{\delta_n\}_{n\in\mathbb{N}}$ is contained in D but contains no convergent subsequences (prove this). ♢

Here are some examples illustrating total boundedness.

Example 1′.6.7. (a) We will show that the interval $[0,1)$ is totally bounded. Given $r>0$, let n be large enough that $\frac{1}{n} < r$. Open balls in one dimension are just open intervals, and the finitely many balls with radius $\frac{1}{n}$ centered at the points $0, \frac{1}{n}, \frac{2}{n}, \dots, 1$ cover $[0,1)$. Consequently the balls with radius r centered at these points cover $[0,1)$. Therefore $[0,1)$ is totally bounded (even though it is not closed).

(b) The interval $[1,\infty)$ is not totally bounded, because we cannot cover it with finitely many balls with a fixed finite radius.

(c) Based on our intuition from finite dimensions, it may seem that every bounded set is totally bounded. However, according to Problem 1′.6.15, the closed unit disk in ℓ^1 is not totally bounded, even though it is both closed and bounded. ♢

In order to prove some equivalent characterizations of compactness, we will need the following lemma.

Lemma 1′.6.8. *Let E be a sequentially compact subset of a normed space X. If $\{U_i\}_{i\in I}$ is an open cover of E, then there exists a number $\delta>0$ such that if B is an open ball of radius δ that intersects E, then there is an $i \in I$ such that $B \subseteq U_i$.*

Proof. Let $\{U_i\}_{i\in I}$ be an open cover of E. We want to prove that there is a $\delta>0$ such that

$$B_\delta(x) \cap E \ne \varnothing \quad \Longrightarrow \quad B_\delta(x) \subseteq U_i \text{ for some } i \in I.$$

Suppose that no $\delta>0$ has this property. Then for each positive integer n, there must exist some open ball with radius $\frac{1}{n}$ that intersects E but is not contained in any U_i. Call this open ball G_n. For each n, choose a point $x_n \in G_n \cap E$. Then since $\{x_n\}_{n\in\mathbb{N}}$ is a sequence in E and since E is sequentially compact, there must be a subsequence $\{x_{n_k}\}_{k\in\mathbb{N}}$ that converges to a point $x \in E$. Since $\{U_i\}_{i\in I}$ is a cover of E, we must have $x \in U_i$ for some $i \in I$, and since U_i is open, there must exist some $r>0$ such that $B_r(x) \subseteq U_i$. Now choose k large enough that we have both

$$\frac{1}{n_k} < \frac{r}{3} \qquad \text{and} \qquad \|x - x_{n_k}\| < \frac{r}{3}.$$

Keeping in mind that G_{n_k} contains x_{n_k}, G_{n_k} is an open ball with radius $1/n_k$, the distance from x to x_{n_k} is less than $r/3$, and $B_r(x)$ has radius r, it follows (why?) that $G_{n_k} \subseteq B_r(x) \subseteq U_i$. But this is a contradiction. □

Now we prove some reformulations of compactness that hold for subsets of normed spaces.

Theorem 1′.6.9. *If K is a subset of a normed space X, then the following three statements are equivalent.*

(a) *K is compact.*

(b) *K is sequentially compact.*

(c) *K is totally bounded and complete (where* complete *means that every Cauchy sequence of points from K converges to a point in K).*

Proof. (a) $\Rightarrow$ (b). We will prove the contrapositive statement. Suppose that K is not sequentially compact. Then there exists a sequence $\{x_n\}_{n\in\mathbb{N}}$ in K that has no subsequence that converges to an element of K. Choose any point $x \in K$. If every open ball centered at x contains infinitely many of the points x_n, then by considering radii $r = \frac{1}{k}$ we can construct a subsequence $\{x_{n_k}\}_{k\in\mathbb{N}}$ that converges to x. This is a contradiction, so there must exist some open ball centered at x that contains only finitely many x_n. If we call this ball B_x, then $\{B_x\}_{x\in K}$ is an open cover of K that contains no finite subcover (because any finite union $B_{x_1} \cup \cdots \cup B_{x_M}$ can contain only finitely many of the x_n). Consequently, K is not compact.

(b) $\Rightarrow$ (c). Suppose that K is sequentially compact, and suppose that $\{x_n\}_{n\in\mathbb{N}}$ is a Cauchy sequence in K. Then since K is sequentially compact, there is a subsequence $\{x_{n_k}\}_{k\in\mathbb{N}}$ that converges to some point $x \in K$. Hence $\{x_n\}_{n\in\mathbb{N}}$ is Cauchy and has a convergent subsequence. Appealing to Problem 1′.2.17, this implies that $x_n \to x$. Therefore K is complete.

Suppose that K were not totally bounded. Then there would be a radius $r > 0$ such that K cannot be covered by finitely many open balls of radius r centered at points of X. Choose any point $x_1 \in K$. Since K cannot be covered by a single r-ball, K cannot be a subset of $B_r(x_1)$. Hence there exists a point $x_2 \in K \setminus B_r(x_1)$. In particular, $\|x_2 - x_1\| \geq r$. But K cannot be covered by two r-balls, so there must exist a point x_3 that belongs to $K \setminus \big(B_r(x_1) \cup B_r(x_2)\big)$. In particular, we have both $\|x_3 - x_1\| \geq r$ and $\|x_3 - x_2\| \geq r$. Continuing in this way, we obtain a sequence of points $\{x_n\}_{n\in\mathbb{N}}$ in K that has no convergent subsequence, which is a contradiction.

(c) $\Rightarrow$ (b). Assume that K is complete and totally bounded, and let $\{x_n\}_{n\in\mathbb{N}}$ be any sequence of points in K. Since K is totally bounded, it can be covered by finitely many open balls of radius $\frac{1}{2}$. Each x_n belongs to K and hence must be contained in one or more of these balls. But there are only finitely many of the balls, so at least one ball must contain x_n for infinitely many different indices n. That is, there is some infinite subsequence

$\{x_n^{(1)}\}_{n\in\mathbb{N}}$ of $\{x_n\}_{n\in\mathbb{N}}$ that is contained in an open ball of radius $\frac{1}{2}$. The Triangle Inequality therefore implies that

$$\forall\, m,n \in \mathbb{N}, \quad \|x_m^{(1)} - x_n^{(1)}\| \;<\; 1.$$

Similarly, since K can be covered by finitely many open balls of radius $\frac{1}{4}$, there is some subsequence $\{x_n^{(2)}\}_{n\in\mathbb{N}}$ of $\{x_n^{(1)}\}_{n\in\mathbb{N}}$ such that

$$\forall\, m,n \in \mathbb{N}, \quad \|x_m^{(2)} - x_n^{(2)}\| \;<\; \frac{1}{2}.$$

Continuing by induction, for each $k > 1$ we find a subsequence $\{x_n^{(k)}\}_{n\in\mathbb{N}}$ of $\{x_n^{(k-1)}\}_{n\in\mathbb{N}}$ such that $\|x_m^{(k)} - x_n^{(k)}\| < \frac{1}{k}$ for all $m, n \in \mathbb{N}$.

Now consider the diagonal subsequence $\{x_k^{(k)}\}_{k\in\mathbb{N}}$. Given $\varepsilon > 0$, let N be large enough that $\frac{1}{N} < \varepsilon$. If $j \ge k > N$, then $x_j^{(j)}$ is one element of the sequence $\{x_n^{(k)}\}_{n\in\mathbb{N}}$ (why?), say $x_j^{(j)} = x_n^{(k)}$. Hence

$$\|x_j^{(j)} - x_k^{(k)}\| \;=\; \|x_n^{(k)} - x_k^{(k)}\| \;<\; \frac{1}{k} \;<\; \frac{1}{N} \;<\; \varepsilon.$$

Thus $\{x_k^{(k)}\}_{k\in\mathbb{N}}$ is a Cauchy subsequence of the original sequence $\{x_n\}_{n\in\mathbb{N}}$. Since K is complete, this subsequence must converge to some element of E. Hence K is sequentially compact.

(b) $\Rightarrow$ (a). Assume that K is sequentially compact. Since we have already proved that statement (b) implies statement (c), we know that K is complete and totally bounded.

Suppose that $\{U_i\}_{i\in I}$ is any open cover of K. By Lemma 1′.6.8, there exists a $\delta > 0$ such that if B is an open ball of radius δ that intersects E, then there is an $i \in I$ such that $B \subseteq U_i$. However, K is totally bounded, so we can cover K by finitely many open balls of radius δ. Each of these balls is contained in some U_i, so K is covered by finitely many U_i. □

The type of reasoning used to prove the implication (c) $\Rightarrow$ (b) in Theorem 1′.6.9 is often called a *Cantor diagonalization argument*.

Remark 1′.6.10. (a) By Problem 1′.4.6, if X is a Banach space then a subset K is closed if and only if it is complete. Therefore, *when X is a Banach space* we can equivalently word statement (c) of Theorem 1′.6.9 as "K is closed and totally bounded."

(b) We saw in Example 1′.6.6 that the closed unit disk D in ℓ^1 is not sequentially compact. Applying Theorem 1′.6.9, it follows that D is neither compact nor totally bounded, even though it is both closed and bounded. ♢

Problems

1′.6.11. (a) Prove directly from the definition that every finite subset of a normed space, including the empty set, is compact.

(b) In $X = \mathbb{R}$, exhibit an open cover of $E = \{\frac{1}{n}\}_{n\in\mathbb{N}}$ that has no finite subcover.

1′.6.12. Let K be a compact subset of a normed space X, and let E be a subset of K. Prove that E is closed if and only if E is compact.

1′.6.13. (a) Suppose that X is a normed space, and F and K are nonempty disjoint subsets of X. Prove that if F is closed and K is compact, then $\operatorname{dist}(F, K) > 0$, where $\operatorname{dist}(F, K) = \inf\{\|x - y\| : x \in F,\ y \in K\}$.

Hint: Use the definition of compactness directly, or use the equivalence with sequential compactness.

(b) Show by example that the distance between two nonempty disjoint *closed* sets E and F can be zero.

1′.6.14. (a) Prove the *Cantor Intersection Theorem* for normed spaces: If $K_1 \supseteq K_2 \supseteq \cdots$ is a nested decreasing sequence of nonempty compact subsets of a normed space X, then $\cap K_n \neq \varnothing$.

(b) Show by example that it is possible for the intersection of a nested decreasing sequence of nonempty *closed* sets to be empty.

1′.6.15. The closed unit disk in ℓ^1 is $D = \{x \in \ell^1 : \|x\|_1 \le 1\}$. Give direct proofs of the following statements.

(a) D is a closed and bounded subset of ℓ^1.

(b) D is not sequentially compact.

(c) D is not totally bounded.

1′.7 Continuity for Functions on Normed Spaces

Our focus up to now has mostly been on particular normed spaces and on particular points in those spaces. Now we will look at *functions* that transform points in one normed space into points in another space. We will be especially interested in *continuous functions*, which are defined as follows.

Definition 1′.7.1 (Continuous Function). Assume that X and Y are normed spaces. We say that a function $f\colon X \to Y$ is *continuous* if given any open set $V \subseteq Y$, its inverse image $f^{-1}(V)$ is an open subset of X. ◊

That is, f is continuous if the *inverse image* of every open set is open. However, the *direct image* of an open set under a continuous function need not be open. For example, $f(x) = \sin x$ is continuous on $X = \mathbb{R}$, but the direct image of the open interval $(0, 2\pi)$ is $f((0, 2\pi)) = [-1, 1]$. Thus, a continuous function need not send open sets *to* open sets.

A continuous function also need not map closed sets to closed sets. For example, $f(x) = \arctan x$ is continuous on $\mathbb{R}$ and $\mathbb{R}$ is a closed set, but $f(\mathbb{R}) = (-\frac{\pi}{2}, \frac{\pi}{2})$, which is not closed. Consequently, the following lemma, which states that *a continuous function maps compact sets to compact sets*, may seem a bit surprising at first glance.

Lemma 1′.7.2. *Let X and Y be normed spaces. If $f\colon X \to Y$ is continuous and K is a compact subset of X, then $f(K)$ is a compact subset of Y.*

Proof. Let $\{V_i\}_{i\in I}$ be any open cover of $f(K)$. Then each set $U_i = f^{-1}(V_i)$ is open, and $\{U_i\}_{i\in I}$ is an open cover of K (why?). Since K is compact, this cover must have a finite subcover, say $\{U_{i_1}, \dots, U_{i_N}\}$. But then $\{V_{i_1}, \dots, V_{i_N}\}$ is a finite subcover of $f(K)$, so we conclude that $f(K)$ is compact. □

As a corollary, we prove that a continuous function on a compact set must be bounded.

Corollary 1′.7.3. *Let X and Y be a normed spaces, and assume that K is a compact subset of X. If $f\colon K \to Y$ is continuous, then f is bounded on K, i.e.,* $\mathrm{range}(f)$ *is a bounded subset of Y.*

Proof. Lemma 1′.7.2 implies that $\mathrm{range}(f) = f(K)$ is a compact subset of Y. By Lemma 1′.6.3, all compact sets are bounded. □

The next lemma gives a useful reformulation of continuity in terms of preservation of limits. In particular, part (c) of this lemma says that a function on a normed space is continuous if and only if f *maps convergent sequences to convergent sequences.*

Lemma 1′.7.4. *Let X be a normed space with norm $\|\cdot\|_X$, and let Y be a normed space with norm $\|\cdot\|_Y$. If $f\colon X \to Y$, then the following three statements are equivalent.*

(a) *f is continuous.*

(b) *If x is any point in X, then for every $\varepsilon > 0$ there exists a $\delta > 0$ such that for all $y \in X$ we have*

$$\|x - y\|_X < \delta \quad \Longrightarrow \quad \|f(x) - f(y)\|_Y < \varepsilon.$$

(c) *Given any point $x \in X$ and any sequence $\{x_n\}_{n\in\mathbb{N}}$ in X,*

$$x_n \to x \text{ in } X \quad \Longrightarrow \quad f(x_n) \to f(x) \text{ in } Y.$$

Proof. For this proof we let $B_r^X(x)$ and $B_s^Y(y)$ denote open balls in X and Y, respectively.

(a) $\Rightarrow$ (b). Suppose that f is continuous, and choose any point $x \in X$ and any $\varepsilon > 0$. Then the ball $V = B_\varepsilon^Y\big(f(x)\big)$ is an open subset of Y, so $U = f^{-1}(V)$ must be an open subset of X. As $x \in U$, there exists some $\delta > 0$ such that $B_\delta^X(x) \subseteq U$. If $y \in X$ is any point that satisfies $\|x-y\|_X < \delta$, then $y \in B_\delta^X(x) \subseteq U$, and therefore $f(y) \in f(U) \subseteq V = B_\varepsilon^Y\big(f(x)\big)$. Consequently $\|f(x) - f(y)\|_Y < \varepsilon$.

(b) $\Rightarrow$ (c). Assume that statement (b) holds, choose $x \in X$, and let $x_n \in X$ be any points such that $x_n \to x$. Fix $\varepsilon > 0$, and let $\delta > 0$ be the number whose existence is given by statement (b). Since $x_n \to x$, there must exist some $N > 0$ such that $\|x - x_n\|_X < \delta$ for all $n > N$. Statement (b) therefore implies that $\|f(x) - f(x_n)\|_Y < \varepsilon$ for all $n > N$, so we conclude that $f(x_n) \to f(x)$.

(c) $\Rightarrow$ (a). Suppose that statement (c) holds, and let V be any open subset of Y. Suppose that $f^{-1}(V)$ were not open in X. Then there is some point $x \in f^{-1}(V)$ such that there is no radius $r > 0$ for which the open ball $B_r(x)$ is a subset of $f^{-1}(V)$. In particular, we have for each $n \in \mathbb{N}$ that the ball $B_{1/n}(x)$ is not contained in $f^{-1}(V)$, and therefore some point $x_n \in B_{1/n}(x)$ such that $x_n \notin f^{-1}(V)$. As a consequence, $\|x - x_n\| < 1/n$ for every n, but $f(x_n) \notin V$ for any n.

Now, $x \in f^{-1}(V)$, so $f(x)$ does belong to V. Since V is open, there is some open ball centered at $f(x)$ that is entirely contained in V. That is, there is some radius $\varepsilon > 0$ such that $B_\varepsilon(f(x)) \subseteq V$.

On the other hand, we have $x_n \to x$, so by applying statement (c) we must have $f(x_n) \to f(x)$. Consequently, there is some $N > 0$ such that $\|f(x) - f(x_n)\| < \varepsilon$ for all $n \geq N$. But then

$$f(x_N) \in B_\varepsilon(f(x)) \subseteq V,$$

which contradicts the fact that $f(x_N) \notin V$. This is a contradiction, so $f^{-1}(V)$ must be open, and therefore f is continuous. $\square$

The number δ that appears in statement (b) of Lemma $1'.7.4$ depends on both x and ε. We say that a function f is *uniformly continuous* if δ can be chosen independently of x. Here is the precise definition.

Definition $1'.7.5$ (Uniform Continuity). Let X be a normed space with norm $\|\cdot\|_X$, and let Y be a normed space with norm $\|\cdot\|_Y$. If $E \subseteq X$, then we say that a function $f\colon E \to Y$ is *uniformly continuous* on E if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that for all $x, y \in E$ we have

$$\|x - y\|_X < \delta \quad \Longrightarrow \quad \|f(x) - f(y)\|_Y < \varepsilon. \qquad \diamondsuit$$

The reader should be aware that *uniform continuity* as introduced in Definition $1'.7.5$ is distinct from *uniform convergence*, which is convergence with respect to the uniform norm (see Definition $1'.8.5$).

Our next result states that any function that is continuous on a compact domain is uniformly continuous on that set. We will omit the proof, which can be found in [Heil18, Lemma 2.9.6].

Lemma 1′.7.6. *Let X and Y be normed spaces. If $K \subseteq X$ is compact and $f\colon K \to Y$ is continuous, then f is uniformly continuous on K.* $\diamondsuit$

Problems

1′.7.7. Let $X = [0,\infty)$, and prove the following statements.

(a) $f(x) = x^2$ is continuous but not uniformly continuous on $[0,\infty)$.

(b) $h(x) = \sin x^2$ is continuous but not uniformly continuous on $[0,\infty)$, even though it is bounded.

(c) If f is continuous on $[0,\infty)$ and there is some $R > 0$ such that $f(x) = 0$ for all $x > R$, then f is uniformly continuous on $[0,\infty)$.

(d) If f is continuous on $[0,\infty)$ and f "vanishes at infinity," i.e., if $\lim_{x\to\infty} f(x) = 0$, then f is uniformly continuous on $[0,\infty)$.

1′.7.8. Let X, Y, and Z be normed spaces, and suppose that $f\colon X \to Y$ and $g\colon Y \to Z$ are continuous. Prove that $g \circ f\colon X \to Z$ is continuous.

1′.7.9. Let X be a normed space, and let f and g be continuous scalar-valued functions on X. Prove the following statements.

(a) If a and b are scalars, then $af + bg$ is continuous.

(b) fg is continuous.

(c) If $g(x) \neq 0$ for every x, then $1/g$ and f/g are continuous.

(d) $h(x) = |f(x)|$ is continuous.

(e) The functions $m(x) = \min\{f(x), g(x)\}$ and $M(x) = \max\{f(x), g(x)\}$ are continuous.

Hint: $m(x) = \frac{1}{2}\,(f + g - |f-g|)$ and $M(x) = \frac{1}{2}\,(f+g+|f-g|)$.

(f) $Z_f = \{x \in X : f(x) = 0\}$ is a closed subset of X.

1′.7.10. Let K be a compact subset of a normed space X, and assume that $f\colon K \to \mathbb{R}$ is continuous and real-valued. Prove that f achieves its maximum and its minimum on K, i.e., there exist points $x, y \in K$ such that

$$f(x) \;=\; \inf\bigl\{f(t) : t \in K\bigr\} \qquad \text{and} \qquad f(y) \;=\; \sup\bigl\{f(t) : t \in K\bigr\}.$$

1′.7.11. (a) Let X and Y be normed spaces, and suppose that $f\colon X \to Y$ is uniformly continuous. Prove that f maps Cauchy sequences to Cauchy sequences, i.e., if $\{x_n\}_{n\in\mathbb{N}}$ is Cauchy in X then $\{f(x_n)\}_{n\in\mathbb{N}}$ is Cauchy in Y.

(b) Show by example that part (a) can fail if we only assume that f is continuous instead of uniformly continuous. (Contrast this with Lemma 1′.7.4, which shows that every continuous function must map convergent sequences to convergent sequences.)

1′.7.12. We say that a function $f\colon \mathbb{R}^d \to \mathbb{R}$ is *upper semicontinuous* (abbreviated usc) at a point $x \in \mathbb{R}^d$ if for every $\varepsilon > 0$ there is a $\delta > 0$ such that

$$|x-y| < \delta \quad\Longrightarrow\quad f(y) \le f(x)+\varepsilon.$$

An analogous definition is made for *lower semicontinuity* (lsc). Prove the following statements.

(a) If $g\colon \mathbb{R}^d \to \mathbb{R}$ and $r > 0$, then $h(x) = \inf\bigl\{g(y) : y \in B_r(x)\bigr\}$ is usc at every point where $h(x) \ne -\infty$.

(b) If $f\colon \mathbb{R}^d \to \mathbb{R}$ is given, then f is continuous at x if and only if f is both usc and lsc at x.

(c) If $\{f_i\}_{i\in J}$ is a family of functions on $\mathbb{R}^d$ that are each usc at a point x, then $g(t) = \inf_{i\in J} f_i(t)$ is usc at x.

(d) $f\colon \mathbb{R}^d \to \mathbb{R}$ is usc at every point $x \in \mathbb{R}^d$ if and only if

$$f^{-1}[a,\infty) \;=\; \{x \in \mathbb{R}^d : f(x) \ge a\}$$

is closed for each $a \in \mathbb{R}$. Likewise, f is lsc at every point x if and only if $f^{-1}(a,\infty) = \{x \in \mathbb{R}^d : f(x) > a\}$ is open for each $a \in \mathbb{R}$.

(e) If K is a compact subset of $\mathbb{R}^d$ and $f\colon \mathbb{R}^d \to \mathbb{R}$ is usc at every point of K, then f is bounded above on K.

1′.8 Uniform Convergence of Functions

Suppose that X is a set, $f_1, f_2, \dots$ are scalar-valued functions on X, and f is another scalar-valued function on X. There are many ways in which we can quantify what it means for the functions f_n to converge to f. Perhaps the simplest is the following notion of *pointwise convergence*.

Definition 1′.8.1 (Pointwise Convergence). Let X be a set, and let scalar-valued functions f_n and f on X be given. We say that f_n *converges pointwise to f on X* if

$$\lim_{n\to\infty} f_n(x) \;=\; f(x), \quad \text{for every } x \in X.$$

In this case we write $f_n \to f$ *pointwise.* $\diamondsuit$

Example 1′.8.2 (Shrinking Triangles). Set $X = [0,1]$, and for each $n \in \mathbb{N}$ let $f_n \colon [0,1] \to \mathbb{R}$ be the continuous function given by

$$f_n(x) \;=\; \begin{cases} 0, & x = 0, \\ \text{linear}, & 0 < x < \frac{1}{2n}, \\ 1, & x = \frac{1}{2n}, \\ \text{linear}, & \frac{1}{2n} < x < \frac{1}{n}, \\ 0, & \frac{1}{n} \le x \le 1. \end{cases}$$

Then $f_n(x) \to 0$ for each $x \in [0,1]$ (see the illustration in Figure 1′.8). Therefore f_n converges pointwise to the zero function on $[0,1]$. $\diamondsuit$

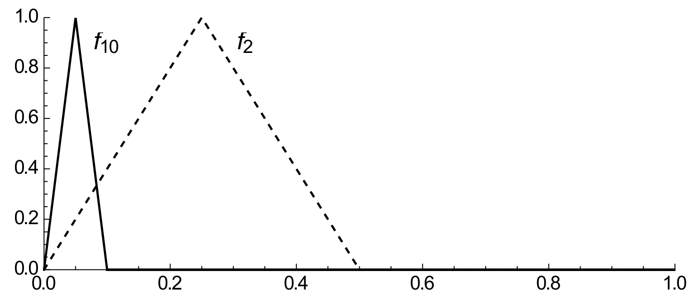

Fig. 1′.8 Graphs of the functions f_2 (dashed) and f_{10} (solid) from Example 1′.8.2.

However, we often need other, more stringent, notions of convergence of functions. We defined the *uniform norm* of functions on the domain $[0,1]$ in equation (1′.3). Now we extend that definition to functions on other sets, and then in Definition 1′.8.5 we will use this notion of the uniform norm to define *uniform convergence* of functions.

Definition 1′.8.3 (Uniform Norm). Let X be a set. The *uniform norm* of a scalar-valued function f on X is

$$\|f\|_{\mathrm{u}} \;=\; \sup_{x \in X} |f(x)|. \qquad \diamondsuit$$

The uniform norm of f can be infinite; indeed, $\|f\|_{\mathrm{u}} < \infty$ if and only if f is bounded on X. We collect the bounded functions to form the space

$$\mathcal{F}_b(X) \;=\; \bigl\{f : f \text{ is a bounded function on } X\bigr\}.$$

By definition, the uniform norm of each element of $\mathcal{F}_b(X)$ is finite.

Lemma 1′.8.4. *The uniform norm* $\|\cdot\|_{\mathrm{u}}$ *is a norm on* $\mathcal{F}_b(X)$.

Proof. The nonnegativity, homogeneity, and uniqueness properties of a norm are immediate. The Triangle Inequality follows from basic properties of suprema, for if f and g are bounded functions on X, then

$$\begin{aligned} \|f+g\|_{\mathrm{u}} \;&=\; \sup_{x\in X} |f(x)+g(x)| \;\le\; \sup_{x\in X} \big(|f(x)|+|g(x)|\big) \\ &\le\; \Big(\sup_{x\in X} |f(x)|\Big) \;+\; \Big(\sup_{x\in X} |g(x)|\Big) \\ &=\; \|f\|_{\mathrm{u}} \;+\; \|g\|_{\mathrm{u}}. \qquad \square \end{aligned}$$

Convergence with respect to the norm $\|\cdot\|_{\mathrm{u}}$ is called *uniform convergence.*

Definition 1′.8.5 (Uniform Convergence). Let X be a set, and let f_n and f be scalar-valued functions on X. Then we say that f_n *converges uniformly to* f on X, and write $f_n \to f$ *uniformly*, if

$$\lim_{n\to\infty} \|f-f_n\|_{\mathrm{u}} \;=\; \lim_{n\to\infty} \Big(\sup_{x\in X} |f(x)-f_n(x)|\Big) \;=\; 0.$$

Similarly, a sequence $\{f_n\}_{n\in\mathbb{N}}$ that is Cauchy with respect to the uniform norm is said to be a *uniformly Cauchy sequence.* ♢

Suppose that f_n converges uniformly to f. Then, by the definition of a supremum, for each individual point x we have

$$|f(x)-f_n(x)| \;\le\; \|f-f_n\|_{\mathrm{u}} \;\to\; 0 \quad \text{as } n\to\infty. \tag{1′.26}$$

Thus $f_n(x) \to f(x)$ for each individual x. This shows that *uniform convergence implies pointwise convergence.* However, the next example shows that pointwise convergence does not imply uniform convergence in general.

Example 1′.8.6 (Shrinking Triangles Revisited). Let $\{f_n\}_{n\in\mathbb{N}}$ be the sequence of "Shrinking Triangles" from Example 1′.8.2. Then $\lim_{n\to\infty} f_n(x) = 0$ for each $x \in [0,1]$, so f_n converges pointwise to the zero function on $[0,1]$. However, for each $n \in \mathbb{N}$ we have

$$\|0-f_n\|_{\mathrm{u}} \;=\; \sup_{x\in[0,1]} |0-f_n(x)| \;=\; 1 \;\not\to\; 0 \quad \text{as } n\to\infty.$$

Therefore f_n does not converge uniformly to the zero function. In fact, $\{f_n\}_{n\in\mathbb{N}}$ is not uniformly Cauchy (why?), so there is *no function* f that f_n converges to uniformly. ♢

Next we will show that the space of bounded functions on X is *complete* with respect to the uniform norm.

Theorem 1′.8.7. *Let X be a set. If $\{f_n\}_{n\in\mathbb{N}}$ is a sequence in $\mathcal{F}_b(X)$ that is Cauchy with respect to $\|\cdot\|_u$, then there exists a function $f \in \mathcal{F}_b(X)$ such that f_n converges uniformly to f. Consequently $\mathcal{F}_b(X)$ is a Banach space with respect to the uniform norm.*

Proof. Suppose that $\{f_n\}_{n\in\mathbb{N}}$ is a uniformly Cauchy sequence in $\mathcal{F}_b(X)$. If we fix any particular point $x \in X$, then for all m and n we have

$$|f_m(x) - f_n(x)| \le \|f_m - f_n\|_u. \qquad (1'.27)$$

Hence, for this fixed x the sequence $\{f_n(x)\}_{n\in\mathbb{N}}$ is a *Cauchy sequence of scalars*. Since the set of scalars (i.e., the real line $\mathbb{R}$ or the complex plane $\mathbb{C}$) is complete by Theorem 1′.2.7, this sequence of scalars must converge. Define

$$f(x) = \lim_{n\to\infty} f_n(x), \qquad \text{for } x \in X.$$

We will show that $f \in \mathcal{F}_b(X)$ and $f_n \to f$ uniformly.

Choose any $\varepsilon > 0$. Since $\{f_n\}_{n\in\mathbb{N}}$ is uniformly Cauchy, there exists an N such that $\|f_m - f_n\|_u < \varepsilon$ for all $m, n \ge N$. If $n \ge N$, then for every $x \in X$ we have

$$\begin{aligned}
|f(x) - f_n(x)| &= \lim_{m\to\infty} |f_m(x) - f_n(x)| && \left(\text{since } f(x) = \lim_{m\to\infty} f_m(x)\right)\\
&\le \lim_{m\to\infty} \|f_m - f_n\|_u && (\text{by equation } (1'.27))\\
&\le \varepsilon && (\text{since } m, n \ge N).
\end{aligned}$$

Hence for each $n > N$ we have

$$\|f - f_n\|_u = \sup_{x\in X} |f(x) - f_n(x)| \le \varepsilon.$$

This shows that $f_n \to f$ uniformly as $n \to \infty$. Also, since $\|f - f_n\|_u \le \varepsilon$, the function $g_n = f - f_n$ is bounded. Since f_n and g_n are bounded, their sum, which is f, is bounded. Thus $f \in \mathcal{F}_b(X)$, so $\mathcal{F}_b(X)$ is complete. □

1′.8.1 Spaces of Continuous Functions

Now we will consider the case where the domain X is a normed space. In this setting we can distinguish between continuous and discontinuous functions. We introduce two spaces of continuous functions on a normed space.

Definition 1′.8.8. If X is a normed space then we let $C(X)$ be the space of all continuous scalar-valued functions f on X, and we let $C_b(X)$ be the subspace of bounded continuous functions on X. That is,

$$C(X) \;=\; \bigl\{f : f \text{ is continuous and scalar-valued on } X\bigr\},$$

and

$$C_b(X) \;=\; C(X) \cap \mathcal{F}_b(X) \;=\; \bigl\{f \in C(X) : f \text{ is bounded}\bigr\}. \qquad \diamondsuit$$

Problem 1′.7.9 shows that $C_b(X)$ is a subspace of $\mathcal{F}_b(X)$, and therefore $C_b(X)$ is a normed space with respect to the uniform norm. Unless we explicitly specify otherwise, we always assume that the norm on $C_b(X)$ is the uniform norm. Note that if X is compact, then every function on X is bounded by Corollary 1′.7.3, so $C_b(X) = C(X)$ when X is compact.

According to the following lemma, the uniform limit of a sequence of continuous functions is continuous.

Theorem 1′.8.9. *Let X be a normed space. If $f_n \in C_b(X)$ for $n \in \mathbb{N}$ and $f\colon X \to \mathbb{R}$ is a function on X such that f_n converges uniformly to f, then $f \in C_b(X)$.*

Proof. Fix any $\varepsilon > 0$. Then, by the definition of uniform convergence, there exists some integer $n > 0$ such that $\|f - f_n\|_{\mathrm{u}} < \varepsilon/3$. In fact, this will be true for all large enough n, but we need only one particular n for this proof. Choose any point $x \in X$. Since f_n is continuous, there is a $\delta > 0$ such that for all $y \in X$ we have

$$\|x - y\| < \delta \quad\Longrightarrow\quad |f_n(x) - f_n(y)| < \frac{\varepsilon}{3}.$$

Consequently, if $y \in X$ satisfies $\|x - y\| < \delta$, then

$$\begin{aligned}
|f(x) - f(y)| &\le |f(x) - f_n(x)| + |f_n(x) - f_n(y)| + |f_n(y) - f(y)| \\
&\le \|f - f_n\|_{\mathrm{u}} + \frac{\varepsilon}{3} + \|f_n - f\|_{\mathrm{u}} \\
&< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.
\end{aligned}$$

Hence f is continuous by Lemma 1′.7.4. Since f_n and $f - f_n$ are both bounded, their sum, which is f, is also bounded. Therefore $f \in C_b(X)$. □

Now we prove that $C_b(X)$ is complete.

Theorem 1′.8.10. *$C_b(X)$ is a closed subspace of $\mathcal{F}_b(X)$ with respect to the uniform norm, and therefore it is a Banach space with respect to $\|\cdot\|_{\mathrm{u}}$.*

Proof. Suppose functions $f_n \in C_b(X)$ and $f \in \mathcal{F}_b(X)$ are given such that $f_n \to f$ uniformly. If we can show that f belongs to $C_b(X)$, then Theorem

$1'.4.2$ will imply that $C_b(X)$ is closed. But the work for this is already done—we just apply Theorem $1'.8.9$ and conclude that $f \in C_b(X)$. Consequently $C_b(X)$ is a closed subspace of $\mathcal{F}_b(X)$ with respect to the uniform norm, and therefore it is complete since $\mathcal{F}_b(X)$ is complete. $\square$

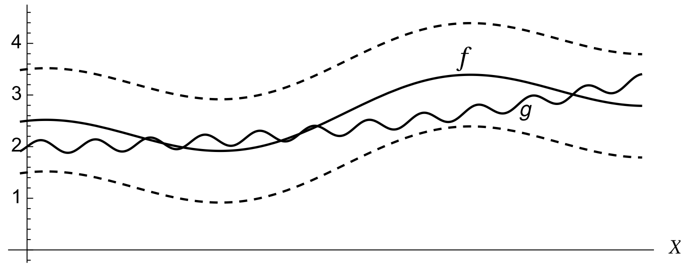

Fig. 1′.9 The function g belongs to the open ball $B_1(f)$ centered at f with radius 1, because $\|f-g\|_{\mathrm{u}} < 1$. The dashed curves are the graphs of $f(x)+1$ and $f(x)-1$.

It is not as easy to visualize the appearance of an open ball in $C_b(X)$ as it is in $\mathbb{R}^2$ or $\mathbb{R}^3$, but we can draw a picture that illustrates the idea. Suppose that I is an interval in $\mathbb{R}$, and f is some particular function in $C_b(I)$. By definition, the open ball $B_r(f)$ of radius r centered at f is

$$B_r(f) \;=\; \bigl\{g \in C_b(I) : \|f-g\|_{\mathrm{u}} < r\bigr\}.$$

That is, $B_r(f)$ consists of all functions g whose uniform distance from f is strictly less than r. Since $\|f-g\|_{\mathrm{u}} = \sup_{x \in I} |f(x)-g(x)|$, a function g belongs to $B_r(f)$ if and only if there exists a number $0 \le s < r$ such that

$$|f(x)-g(x)| \;\le\; s \quad \text{for every } x \in I.$$

Thus if $g \in B_r(f)$, then $g(x)$ can never stray from $f(x)$ by more than $s < r$ units. Figure $1'.9$ depicts a function g that belongs to the open ball $B_1(f)$ for one particular f. The ball $B_1(f)$ consists of all such continuous functions g whose graphs lie between the dashed lines in Figure $1'.9$.

We state a useful result on the approximation of continuous functions by polynomials on a finite interval. There are many different proofs of this theorem; one can be found in [Heil18, Thm. 4.6.2].

Theorem 1′.8.11 (Weierstrass Approximation Theorem). *Let $[a,b]$ be a finite closed interval. If $f \in C[a,b]$ and $\varepsilon > 0$, then there exists a polynomial $p(x) = \sum_{k=0}^{n} c_k x^k$ that satisfies*

$$\|f-p\|_{\mathrm{u}} \;=\; \sup_{x \in [a,b]} |f(x)-p(x)| \;<\; \varepsilon. \qquad \diamondsuit$$

1′.8.2 A First Look at L^p-Norms

We created a family of spaces ℓ^p in Section 1′.5. By substituting integrals for sums we can create a related family of norms on the space of continuous functions $C[a,b]$. We call the norm $\|\cdot\|_p$ defined in the following exercise the L^p-*norm* on $C[a,b]$.

Exercise 1′.8.12. Fix $1 \le p < \infty$, and for each $f \in C[a,b]$ define

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

Prove that $\|\cdot\|_p$ is a norm on $C[a,b]$. ♢

We saw in Exercise 1′.5.5 that ℓ^p is a Banach space with respect to the norm $\|\cdot\|_p$. We prove next that $C[a,b]$ is *not* a Banach space with respect to $\|\cdot\|_p$ for any finite p. The problem is that the Riemann integral is not a "robust enough" notion. We will see in Chapter 7 how to create a larger space of functions that is complete with respect to an L^p-norm that is defined in terms of *Lebesgue integrals* instead of Riemann integrals. The Lebesgue integral extends the Riemann integral, but is much more robust. We can define the Lebesgue integral for functions whose domain is any measurable set (not just intervals as for Riemann integrals), and we can prove more powerful results about Lebesgue integrals than we can for Riemann integrals. We will see the details of the construction of Lebesgue measure and Lebesgue integrals in Chapter 2 on onward in the main text.

Lemma 1′.8.13. *If* $1 \le p < \infty$, *then* $C[a,b]$ *is not complete with respect to the* L^p*-norm* $\|\cdot\|_p$.

Proof. For simplicity of presentation we will take $a=-1$ and $b=1$. For each $n \in \mathbb{N}$, define a continuous function $f_n\colon[-1,1]\to\mathbb{R}$ by

$$f_n(x) \;=\; \begin{cases} -1, & \text{if } -1 \le x \le -\frac{1}{n}, \\ \text{linear}, & \text{if } -\frac{1}{n} < x < \frac{1}{n}, \\ 1, & \text{if } \frac{1}{n} \le x \le 1. \end{cases}$$

Figure 1′.10 shows the graph of f_5. Choose any positive integers $m \le n$. By inspection, we can see that $|f_m(x) - f_n(x)| \le 1$ for every x. Since we also have $f_m(x) = f_n(x) = 0$ for $x \notin [-\frac{1}{m}, \frac{1}{m}]$, we compute that

$$\begin{aligned} \|f_m - f_n\|_p^p \;&=\; \int_{-1}^{1} |f_m(x) - f_n(x)|^p\,dx \\ &=\; \int_{-1/m}^{1/m} |f_m(x) - f_n(x)|^p\,dx \;\le\; \int_{-1/m}^{1/m} 1\,dx \;=\; \frac{1}{m}, \end{aligned}$$

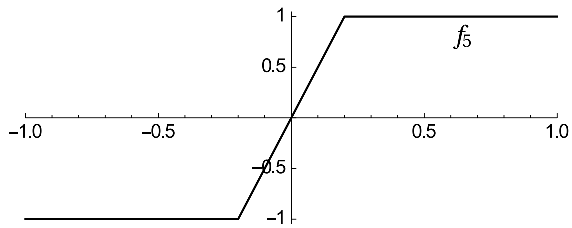

Fig. 1′.10 Graph of the function f_5.

and therefore $\|f_m - f_n\|_p \le 1/m^{1/p}$ for all $n \ge m$. Consequently, if we fix $\varepsilon > 0$, then for all large enough m and n we will have $\|f_m - f_n\|_p < \varepsilon$. This shows that $\{f_n\}_{n\in\mathbb{N}}$ is Cauchy in $C[-1,1]$ with respect to $\|\cdot\|_p$. However, we will prove that there is *no function* $g \in C[-1,1]$ such that $\|g - f_n\|_p \to 0$ as $n \to \infty$.

Suppose that there were some $g \in C[-1,1]$ such that $\|g - f_n\|_p \to 0$. Fix $0 < c < 1$, and suppose that g is not identically 1 on $[c,1]$. Then $|g-1|^p$ is continuous but not identically zero on $[c,1]$, so its integral on this interval must be nonzero, say

$$C \;=\; \int_c^1 |g(x) - 1|^p\,dx \;>\; 0.$$

On the other hand, f_n is identically 1 on $[c,1]$ for all $n > 1/c$, so for all large enough n we have

$$\begin{aligned}
\|g - f_n\|_p^p \;&=\; \int_{-1}^1 |g(x) - f_n(x)|^p\,dx \\
&\ge\; \int_c^1 |g(x) - f_n(x)|^p\,dx \\
&=\; \int_c^1 |g(x) - 1|^p\,dx \;=\; C.
\end{aligned}$$

Since C is a fixed positive constant, it follows that $\|g - f_n\|_p \ge C^{1/p} \not\to 0$, which is a contradiction. Therefore g must be identically 1 on $[c,1]$. This is true for every $c > 1$, so $g(x) = 1$ for all $0 < x \le 1$. A similar argument shows that $g(x) = -1$ for all $-1 \le x < 0$. But there is no continuous function that takes these values, so we have obtained a contradiction.

Thus, although $\{f_n\}_{n\in\mathbb{N}}$ is Cauchy in $C[-1,1]$, there is no function in $C[-1,1]$ that this sequence can converge to with respect to $\|\cdot\|_p$. Therefore $C[-1,1]$ is not complete with respect to this norm. □

Problems

1′.8.14. Let $p_n(x) = x^n$ for $x \in [0,1]$. Prove that the sequence $\{p_n\}_{n\in\mathbb{N}}$ converges pointwise to a *discontinuous* function, but p_n does not converge *uniformly* to that function.

1′.8.15. For the following choices of functions g_n, determine whether the sequence $\{g_n\}_{n\in\mathbb{N}}$ is: pointwise convergent, uniformly convergent, uniformly Cauchy, or bounded with respect to the uniform norm.

(a) $g_n(x) = x^n/n$ for $x \in [0,1]$.

(b) $g_n(x) = \sin 2\pi n x$ for $x \in [0,1]$.

(c) $g_n(x) = n f_n(x)$ for $x \in [0,1]$, where f_n is the Shrinking Triangle from Example 1′.8.2.

(d) $g_n(x) = e^{-n|x|}$ for $x \in \mathbb{R}$.

1′.8.16. Let X be a set. Suppose that $f_n,\ f,\ g_n,\ g \in \mathcal{F}_b(X)$ are such that $f_n \to f$ and $g_n \to g$ uniformly. Prove the following statements.

(a) $f_n + g_n \to f + g$ uniformly.

(b) $f_n g_n \to fg$ uniformly.

(c) If there is a $\delta > 0$ such that $|g_n(x)| \ge \delta$ for every $x \in X$ and $n \in \mathbb{N}$, then $f_n/g_n \to f/g$ uniformly.

1′.8.17. Prove that if $f \in C(\mathbb{R})$ is uniformly continuous and $f_n(x) = f(x - \frac{1}{n})$, then $f_n \to f$ uniformly. Show by example that this can fail if $f \in C(\mathbb{R})$ is not uniformly continuous.

1′.8.18. The space of continuous functions that “vanish at infinity” is

$$C_0(\mathbb{R}) \;=\; \Big\{ f \in C(\mathbb{R}) \;:\; \lim_{x\to\pm\infty} f(x) = 0 \Big\}.$$

Prove the following statements.

(a) $C_0(\mathbb{R})$ is a closed subspace of $C_b(\mathbb{R})$ with respect to the uniform norm, and it therefore is a Banach space with respect to $\|\cdot\|_{\mathrm{u}}$.

(b) Every function in $C_0(\mathbb{R})$ is uniformly continuous on $\mathbb{R}$, but there exist functions in $C_b(\mathbb{R})$ that are not uniformly continuous.

(c) A continuous function $f \in C(\mathbb{R})$ has *compact support* if it is identically zero outside of some finite interval. Let

$$C_c(\mathbb{R}) \;=\; \big\{ f \in C(\mathbb{R}) \;:\; f \text{ has compact support} \big\}.$$

Given $g \in C_0(\mathbb{R})$, prove that there exist functions $g_n \in C_c(\mathbb{R})$ such that $g_n \to f$ uniformly (consider Figure 1′.11).

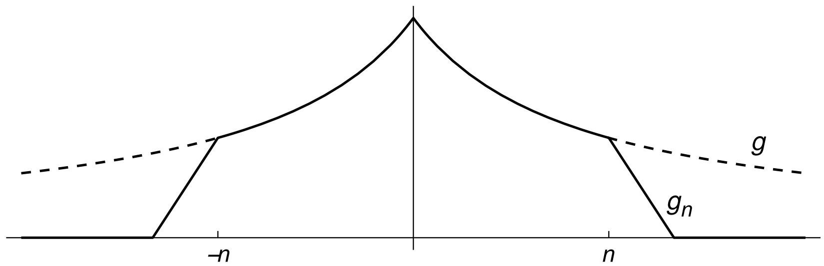

Fig. 1′.11 A function g and a compactly supported approximation g_n.

1′.8.19. Let I be an open interval in the real line (either finite or infinite). Let $C_b^1(I)$ be the set of all bounded differentiable functions on I that have a bounded continuous derivative:

$$C_b^1(I) \;=\; \bigl\{f \in C_b(I) \,:\, f \text{ is differentiable and } f' \in C_b(I)\bigr\}.$$

Prove the following statements.

(a) $\|\cdot\|_{\mathrm{u}}$ is a norm on $C_b^1(I)$, but $C_b^1(I)$ is not complete with respect to this norm.

(b) $\|f\| = \|f'\|_{\mathrm{u}}$ defines a seminorm on $C_b^1(I)$, but it is not a norm.

(c) $\|f\|_{C_b^1} = \|f\|_{\mathrm{u}} + \|f'\|_{\mathrm{u}}$ defines a norm on $C_b^1(I)$.

(d) $C_b^1(I)$ is complete with respect to the norm $\|\cdot\|_{C_b^1}$.

Remark: If I is a closed or half-open interval then the same results hold if *differentiable on I* means that f is differentiable at any point in the interior of I and differentiable from the right or left (as appropriate) at any endpoint of I.

1′.8.20. The *unit disk D* in $C_b(\mathbb{R})$ is the set of all functions in $C_b(\mathbb{R})$ whose uniform norm is at most 1, i.e., $D = \{f \in C_b(\mathbb{R}) : \|f\|_{\mathrm{u}} \le 1\}$.

(a) Prove that D is a closed and bounded subset of $C_b(\mathbb{R})$.

(b) The *hat function* or *tent function* on the interval $[-1,1]$ is

$$W(x) \;=\; \max\bigl\{1-|x|, 0\bigr\} \;=\; \begin{cases} 1-x, & 0 \le x \le 1, \\ 1+x, & -1 \le x \le 0, \\ 0, & |x| \ge 1. \end{cases}$$

Let $f_k(x) = T_k W(x) = W(x-k)$. Observe that $\|f_k\|_{\mathrm{u}} = 1$, so the sequence $\{f_k\}_{k\in\mathbb{N}}$ is contained in the unit disk D. Prove that $\{f_k\}_{k\in\mathbb{N}}$ is not a Cauchy sequence and contains no Cauchy subsequences.

(c) Prove that D is not compact.

1′.9 Infinite Series in Normed Spaces

Suppose that $x_1, x_2, \dots$ are infinitely many vectors in a vector space X. Does it make sense to sum all of these vectors, i.e., to form the *infinite series*

$$\sum_{n=1}^{\infty} x_n \;=\; x_1 + x_2 + \cdots \,?$$

Since a vector space has an operation of *vector addition*, we do have a way to add two vectors together, and hence by induction we can compute the sum of *finitely many* vectors. Thus, we can define the *partial sums*

$$s_N \;=\; \sum_{n=1}^{N} x_n \;=\; x_1 + x_2 + \cdots + x_N$$

for any finite integer $N = 1, 2, \dots$. However, this does not by itself tell us how to compute a sum of *infinitely many* vectors. To make sense of this we have to know what happens to the partial sums s_N as N increases. And to do this we need to be able to compute the *limit* of the partial sums, which requires that we have a notion of *convergence* in our space. If all we know is that X is a vector space, then we do not have any way to measure distance or to determine convergence in X. But if X is a *normed space* then we do know what convergence and limits mean, and so we can determine whether the partial sums $s_N = x_1 + \cdots + x_N$ converge to some limit as $N \to \infty$. If they do, then this is what it means to form an infinite sum. We make all this precise in the following definition.

Definition 1′.9.1 (Convergent Series). Let $\{x_n\}_{n \in \mathbb{N}}$ be a sequence of vectors in a normed space X. We say that the infinite series $\sum_{n=1}^{\infty} x_n$ *converges* if there is a vector $x \in X$ such that the *partial sums* $s_N = \sum_{n=1}^{N} x_n$ converge to x, i.e., if

$$\lim_{N\to\infty} \|x - s_N\| \;=\; \lim_{N\to\infty} \left\| x - \sum_{n=1}^{N} x_n \right\| \;=\; 0. \tag{1′.28}$$

In this case, we write

$$x \;=\; \sum_{n=1}^{\infty} x_n,$$

and we also use the shorthands

$$x \;=\; \sum x_n \qquad \text{or} \qquad x \;=\; \sum_n x_n. \qquad \diamondsuit$$

In order for an infinite series to converge in X, the *norm of the difference* between x and the partial sum s_N must converge to zero. If we wish to emphasize which norm we are referring to, we may write that $x = \sum x_n$ *converges with respect to* $\|\cdot\|$, or we may say that $x = \sum x_n$ *converges in* X.

Example 1′.9.2. Let $X = \ell^\infty$, and let $\{\delta_n\}_{n\in\mathbb{N}}$ be the sequence of standard basis vectors. Does the series $\sum_{n=1}^\infty \delta_n$ converge in ℓ^∞? The Nth partial sum of this series is

$$s_N \;=\; \sum_{n=1}^{N} \delta_n \;=\; (1,\dots,1,0,0,\dots),$$

where the 1 is repeated N times. It may appear that s_N converges to the constant vector $x = (1,1,\dots)$. However, while s_N does converge *componentwise* to x, *it does not converge in* ℓ^∞*-norm*, because

$$\|x - s_N\|_\infty \;=\; \big\|(0,\dots,0,1,1,\dots)\big\|_\infty \;=\; 1 \;\not\to\; 0.$$

In fact, the sequence of partial sums $\{s_N\}_{N\in\mathbb{N}}$ is not even Cauchy in ℓ^∞, so the partial sums do not converge to any vector. Therefore $\sum_{n=1}^\infty \delta_n$ is not a convergent series in ℓ^∞. Does this series converge in ℓ^p if p is finite? ◇

Example 1′.9.3. (a) Consider the series $\sum_{n=1}^\infty \frac{(-1)^n}{n}\delta_n$ in the space $X = \ell^1$. The partial sums of this series are

$$s_N \;=\; \sum_{n=1}^{N} \frac{(-1)}{n}\delta_n \;=\; \Big(-1, \tfrac{1}{2}, -\tfrac{1}{3}, \dots, \tfrac{(-1)^N}{N}, 0, 0, \dots\Big).$$

These partial sums converge componentwise to

$$x \;=\; \Big(\tfrac{(-1)^n}{n}\Big)_{n\in\mathbb{N}} \;=\; \big(-1, \tfrac{1}{2}, -\tfrac{1}{3}, \dots\big), \qquad (1'.29)$$

but this vector does not belong to ℓ^1. The partial sums s_N do not converge to x in ℓ^1-norm, or to any vector in ℓ^1, because convergent sequences must be bounded but the s_N are not:

$$\sup_{N\in\mathbb{N}} \|s_N\|_1 \;=\; \sup_{N\in\mathbb{N}} \sum_{n=1}^{N} \left|\frac{(-1)^n}{n}\right| \;=\; \sup_{N\in\mathbb{N}} \sum_{n=1}^{N} \frac{1}{n} \;=\; \infty.$$

Consequently $\sum_{n=1}^\infty \frac{(-1)^n}{n}\delta_n$ is not a convergent series in ℓ^1.

(b) Again consider $\sum_{n=1}^\infty \frac{(-1)^n}{n}\delta_n$, but this time in ℓ^p where $1 < p < \infty$. In this case the vector x defined in equation (1′.29) does belong to ℓ^p, because $\sum \frac{1}{n^p} < \infty$. Furthermore,

$$x - s_N \;=\; \Big(0,\dots,0, \tfrac{(-1)^{N+1}}{N+1}, \tfrac{(-1)^{N+2}}{N+2}, \dots\Big). \qquad (1'.30)$$

Since $\sum \frac{1}{n^p} < \infty$ we know that $\lim_{N\to\infty} \sum_{n=N+1}^\infty \frac{1}{n^p} = 0$ (see Problem 1′.9.5), and therefore

$$\begin{aligned}\lim_{N\to\infty} \|x - s_N\|_1 &= \lim_{N\to\infty} \sum_{n=N+1}^{\infty} \left|\frac{(-1)^n}{n}\right|^p && \text{(by equation (1}'\text{.30))}\\ &= \lim_{N\to\infty} \sum_{n=N+1}^{\infty} \frac{1}{n^p}\\ &= 0 && \text{(by Problem 1}'\text{.9.5).}\end{aligned}$$

Thus $\sum_{n=1}^{\infty} \frac{(-1)^n}{n}\delta_n$ is a convergent series in ℓ^p when $1 < p < \infty$. ♢

Do the $(-1)^n$ factors play any role in Example 1′.9.3, i.e., does anything change in that example if we instead consider the series $\sum_{n=1}^{\infty} \frac{1}{n}\delta_n$?

Problems

1′.9.4. (a) Suppose that $\sum x_n$ and $\sum y_n$ are convergent series in a normed space X. Show that $\sum (x_n + y_n)$ is convergent and equals $\sum x_n + \sum y_n$.

(b) Show by example that $\sum (x_n + y_n)$ can converge even if $\sum x_n$ and $\sum y_n$ do not converge.

(c) If $\sum x_n$ converges but $\sum y_n$ does not, is it possible that $\sum (x_n + y_n)$ could converge?

1′.9.5. Suppose that $c_n \ge 0$ for every $n \in \mathbb{N}$. Prove that if $\sum_{n=1}^{\infty} c_n < \infty$, i.e., if the series converges to a finite scalar, then

$$\lim_{N\to\infty} \left(\sum_{n=N}^{\infty} c_n\right) = 0.$$

1′.9.6. Assume that $\sum_{n=1}^{\infty} x_n$ is a convergent series in a normed space X. Prove that

$$\left\|\sum_{n=1}^{\infty} x_n\right\| \le \sum_{n=1}^{\infty} \|x_n\|.$$

Note that the right-hand side of this inequality could be ∞.

1′.9.7. Given a sequence of scalars $c = (c_n)_{n\in\mathbb{N}}$, prove the following statements.

(a) If $1 \le p < \infty$, then $\sum_{n=1}^{\infty} c_n\delta_n$ converges in ℓ^p if and only if $c \in \ell^p$.

(b) The series $\sum_{n=1}^{\infty} c_n\delta_n$ converges in ℓ^∞ if and only if $c \in c_0$.

1′.9.8. This problem will characterize the functions $f \in C[a,b]$ for which the infinite series $\sum_{n=1}^{\infty} f(x)^n$ converges uniformly. Let $s_N(x) = \sum_{n=1}^{N} f(x)^n$ denote the Nth partial sum of the series, and prove the following statements.

(a) If $\|f\|_u < 1$, then the partial sums $\{s_N\}_{N\in\mathbb{N}}$ are uniformly Cauchy, and therefore $\sum_{n=1}^{\infty} f(x)^n$ converges in the space $C[a,b]$.

(b) If $\|f\|_u \ge 1$, then there is some $x \in [a,b]$ such that $\sum_{n=1}^{\infty} f(x)^n$ does not converge. Explain why this implies that the infinite series $\sum_{n=1}^{\infty} f(x)^n$ does not converge uniformly.

1′.10 Closure, Density, and Separability

The set of rationals $\mathbb{Q}$ is a subset of the real line, but it is not a *closed* subset because a sequence of rational points can converge to an irrational point. Can we find a closed set F that "surrounds" $\mathbb{Q}$, i.e., a closed set that satisfies $\mathbb{Q} \subseteq F \subseteq \mathbb{R}$? Of course, we could just take $F = \mathbb{R}$, but is there a smaller closed set that contains $\mathbb{Q}$? By Theorem 1′.4.2, any such set F would have to contain every limit of elements of $\mathbb{Q}$, but since *every* $x \in \mathbb{R}$ is a limit of rational points we conclude that $F = \mathbb{R}$ is the only closed subset of $\mathbb{R}$ that contains $\mathbb{Q}$.

This type of issue arises quite often. Specifically, given a set E it can be important to find a closed set that not only contains E but is the "smallest closed set" that contains E. The next definition gives a name to this "smallest closed set."

Definition 1′.10.1. Let E be a subset of a normed space X. The *closure of* E, denoted $\overline{E}$, is the intersection of all the closed sets that contain E:

$$\overline{E} = \bigcap \{F \subseteq X : F \text{ is closed and } F \supseteq E\}. \qquad \diamondsuit \tag{1′.31}$$

$\overline{E}$ is closed because the intersection of closed sets is closed, and it contains E by definition. Furthermore, if F is any particular closed set that contains E, then F is one of the sets appearing in the intersection on the right-hand side of equation (1′.31), so F must contain $\overline{E}$. Thus:

- $\overline{E}$ is a closed set that contains E, and
- if F is any closed set with $E \subseteq F$, then $\overline{E} \subseteq F$.

Thus $\overline{E}$ is closed and contains E, and it is contained within any other closed set that contains E. Hence *$\overline{E}$ is the smallest closed set that contains E.*

The following lemma shows that the closure of E is the set of all possible limits of elements of E.

Theorem 1′.10.2. *If E is a subset of a normed space X, then*

$$\overline{E} = \bigl\{y \in X : \text{there exist } x_n \in E \text{ such that } x_n \to y\bigr\}. \tag{1′.32}$$

Proof. Let F be the set that appears on the right-hand side of equation (1′.32), i.e., F is the set of all limits of elements of E. We must show that $F = \overline{E}$.

Choose any point $y \in F$. Then, by definition, there exist points $x_n \in E$ such that $x_n \to y$. Since $E \subseteq \overline{E}$, the points x_n all belong to $\overline{E}$. Hence y is a limit of points of $\overline{E}$. But $\overline{E}$ is a closed set, so it must contain all of these limits. Therefore y belongs to $\overline{E}$, so we have shown that $F \subseteq \overline{E}$.

In order to prove that $\overline{E}$ is a subset of F, we will first prove that F^{C} is an open set. To do this, choose a point $y \in F^{\mathrm{C}}$. We must show that there is a ball centered at y that is entirely contained in F^{C}. That is, we must show that there is some $r > 0$ such that $B_r(y)$ contains no limits of elements of E.

Suppose that for each $k \in \mathbb{N}$, the ball $B_{1/k}(y)$ contained some point from E, say $x_k \in B_{1/k}(y) \cap E$. Then these x_k are points of E that converge to y (why?). Hence y is a limit of points of E, which contradicts the fact that $y \notin F$. Hence there must be at least one k such that $B_{1/k}(y)$ contains no points of E. We will show that $r = 1/k$ is the radius that we seek. That is, we will show that the ball $B_r(y)$, where $r = 1/k$, not only contains no elements of E but furthermore contains no *limits* of elements of E.

Suppose that $B_r(y)$ did contain some point z that was a limit of points of E, i.e., suppose that there did exist some $x_n \in E$ such that $x_n \to z \in B_r(y)$. Then, since $\|y - z\| < r$ and $\|z - x_n\|$ becomes arbitrarily small, by choosing n large enough we will have $\|y - x_n\| \le \|y - z\| + \|z - x_n\| < r$. But then this point x_n belongs to $B_r(y)$, which contradicts the fact that $B_r(y)$ contains no points of E.

Thus, $B_r(y)$ contains no limits of elements of E. Since F is the set of all limits of elements of E, this means that $B_r(y)$ contains no points of F. That is, $B_r(y) \subseteq F^{\mathrm{C}}$.

In summary, we have shown that each point $y \in F^{\mathrm{C}}$ is the center of some ball $B_r(y)$ that is entirely contained in F^{C}. Therefore F^{C} is an open set. Hence, by definition, F is a closed set. We also know that $E \subseteq F$ (why?), so F is one of the closed sets that contains E. But $\overline{E}$ is the smallest closed set that contains E, so we conclude that $\overline{E} \subseteq F$. □

We saw earlier that the closure of the set of rationals $\mathbb{Q}$ is the entire real line $\mathbb{R}$. We introduce some terminology for this type of situation.

Definition $1'.10.3$ (Dense Subset). Let E be a subset of a normed space X. If $\overline{E} = X$ (i.e., the closure of E is all of X), then we say that E is *dense* in X. ♢

Theorem $1'.10.2$ tells us that the closure of a set E equals the set of all limits of elements of E. If E is dense, then the closure of E is the entire space X, so *every* point in X must be a limit of points of E. The converse is also true, and so we obtain the following useful equivalent reformulation of the meaning of density (additional reformulations appear in Problem $1'.10.11$).

Corollary $1'.10.4$. *Let E be a subset of a normed space X. Then E is dense in X if and only if for each $x \in X$ there exist points $x_n \in E$ such that $x_n \to x$.*

Proof. $\Rightarrow$. Suppose that E is dense, i.e., $\overline{E} = X$. If we choose any point $x \in X$, then x belongs to the closure of E. Theorem 1′.10.2 therefore implies that x is a limit of elements of E.

$\Leftarrow$. Suppose that every point $x \in X$ is a limit of elements of E. Theorem 1′.10.2 tells us that the closure of E is the set of all limits of elements of E, so we conclude that $\overline{E} = X$. Therefore E is dense in X. □

Here is a short summary of the facts that we proved in Theorem 1′.4.2, Theorem 1′.10.2, and Corollary 1′.10.4:

- *E is closed if and only if it contains the limit of every convergent sequence of elements of E,*
- *the closure of E is the set of all possible limits of elements chosen from E, and*
- *E is dense in X if and only if every point in X is a limit of points of E.*

For example, the set of rationals $\mathbb{Q}$ is not a closed subset of $\mathbb{R}$ because a limit of rational points need not be rational. The closure of $\mathbb{Q}$ is $\mathbb{R}$ because every point in $\mathbb{R}$ can be written as a limit of rational points. Similarly, $\mathbb{Q}$ is a dense subset of $\mathbb{R}$ because every real number can be written as a limit of rationals.

To illustrate, we will show that c_{00}, the space of "finite sequences" introduced in Example 1′.2.6, is a dense subspace of ℓ^1.

Theorem 1′.10.5. *c_{00} is a dense subspace of ℓ^1.*

Proof. Observe that c_{00} is contained in ℓ^1, and it is a subspace since it is closed under vector addition and scalar multiplication. So, we just have to prove that it is dense. We will do this by showing that every vector in ℓ^1 is a limit of elements of c_{00}.

Choose any $x = (x_k)_{k\in\mathbb{N}} \in \ell^1$. For each $n \in \mathbb{N}$, let y_n be the sequence whose first n components are the same as those of x, but whose remaining components are zero. That is, $y_n = (x_1, \dots, x_n, 0, 0, \dots)$. Note that y_n belongs to c_{00}. By definition, y_n *converges componentwise* to x. However, this is not enough. We must show that y_n converges *in ℓ^1-norm* to x. Note that

$$x - y_n \;=\; (0, \dots, 0, x_{n+1}, x_{n+2}, \dots). \tag{1′.33}$$

Applying the definition of the ℓ^1-norm and the fact that $\sum_{k=1}^{\infty} |x_k| < \infty$, it follows that

$$\begin{aligned} \lim_{n\to\infty} \|x - y_n\|_1 \;&=\; \lim_{n\to\infty} \biggl(\sum_{k=n+1}^{\infty} |x_k|\biggr) && \text{(by equation (1′.33))} \\ &=\; 0 && \text{(by Problem 1′.9.5).} \end{aligned}$$

This shows that $y_n \to x$ in ℓ^1-norm. Hence every element of ℓ^1 is a limit of elements of c_{00}, and therefore c_{00} is dense by Corollary 1′.10.4. □

We introduce another notion related to density. For motivation, recall that the set of rationals $\mathbb{Q}$ is a dense subset of the real line. Since $\mathbb{Q}$ is countable, it follows that $\mathbb{R}$ contains a subset that is "small" in terms of cardinality, but is "large" in the sense that it is dense. In higher dimensions, the set $\mathbb{Q}^d$ consisting of vectors with rational components is a countable yet dense subset of $\mathbb{R}^d$. It may seem unlikely that an infinite-dimensional space could contain such a subset, but we show next that this is true of the infinite-dimensional space ℓ^1.

Example 1′.10.6. Recall that c_{00}, the set of "finite sequences," is dense in ℓ^1. However, c_{00} is not a countable set (why not)? To construct a countable dense subset of ℓ^1, first consider the sets

$$A_N \;=\; \bigl\{(r_1,\dots,r_N,0,0,\dots) : r_1,\dots,r_N \text{ rational}\bigr\}, \qquad \text{for } N \in \mathbb{N}.$$

Each A_N is countable (why?). Since the union of countably many countable sets is countable, it follows that

$$A \;=\; \bigl\{(r_1,\dots,r_N,0,0,\dots) : N>0,\; r_1,\dots,r_N \text{ rational}\bigr\}. \tag{1′.34}$$

is also countable. Further, A is dense in ℓ^1 (see Problem 1′.10.18). Therefore ℓ^1 contains a countable dense subset, so it is separable. ◇

Thus, some normed spaces (even including some infinite-dimensional vector spaces) contain countable dense subsets, but it is also true that there exist normed spaces that do not contain any countable dense subsets. This gives us one way to distinguish between "small" and "large" spaces. We introduce a name for the "small" spaces that contain countable dense subsets.

Definition 1′.10.7 (Separable Space). A normed space that contains a countable dense subset is said to be *separable*. ◇

Here is a nonseparable space.

Theorem 1′.10.8. *ℓ^∞ does not contain a countable dense subset and therefore it is not a separable space.*

Proof. Recall that the distance between two elements of ℓ^∞ is

$$\|x-y\|_\infty \;=\; \sup_{k\in\mathbb{N}} |x_k - y_k|.$$

Let S be the set of all sequences whose components are either 0 or 1, i.e.,

$$S \;=\; \bigl\{x = (x_k)_{k\in\mathbb{N}} \;:\; x_k \in \{0,1\} \text{ for every } k\bigr\}.$$

This is an uncountable set (why?), and if $x \neq y$ are any two distinct elements of S, then $\|x-y\|_\infty = 1$ (why?).

Suppose that ℓ^∞ were separable. Then there would exist countably many vectors $y_1, y_2, \ldots \in \ell^\infty$ such that the set $A = \{y_n\}_{n\in\mathbb{N}}$ was dense in ℓ^∞. Choose any point $x \in S$. Since A is dense in ℓ^∞, Corollary 1′.10.4 implies that there must exist points of A that lie as close to x as we like. In particular, there must exist some integer n_x such that $\|x - y_{n_x}\|_\infty < 1/2$. If we define $f(x) = n_x$, then this gives us a mapping $f\colon S \to \mathbb{N}$. Since S is uncountable and $\mathbb{N}$ is countable, f cannot be injective. Hence there must exist two distinct points $x, z \in S$ such that $n_x = n_z$. Since x and z are two different points from S we have $\|x-z\|_\infty = 1$. Applying the Triangle Inequality and the fact that $y_{n_x} = y_{n_z}$, this implies that

$$1 \;=\; \|x - z\|_\infty \;\le\; \|x - y_{n_x}\|_\infty \;+\; \|y_{n_z} - z\|_\infty \;<\; \frac{1}{2} + \frac{1}{2} \;<\; 1.$$

This is a contradiction, so no countable subset of ℓ^∞ can be dense. □

Problems

1′.10.9. Take $X = \mathbb{R}$, and find E° and $\overline{E}$ for each of the following sets.

(a) $E = \bigl\{1, \frac{1}{2}, \frac{1}{3}, \dots\bigr\}$.

(b) $E = [0, 1)$.

(c) $E = \mathbb{Z}$, the set of integers.

(d) $E = \mathbb{Q}$, the set of rationals.

(e) $E = \mathbb{R}\backslash\mathbb{Q}$, the set of irrationals.

1′.10.10. Let E be a subset of a normed space X. Show that if E is not dense in X, then $X\backslash E$ contains an open ball.

1′.10.11. Let E be a subset of a normed space X. Prove that the following four statements are equivalent.

(a) E is dense in X.

(b) If $x \in X$ and $\varepsilon > 0$, then there exists some $y \in E$ such that $\|x-y\| < \varepsilon$.

(c) $E \cap U \ne \varnothing$ for every nonempty open set $U \subseteq X$.

1′.10.12. Let E be a subset of a normed space X. Prove that E has empty interior ($E^\circ = \varnothing$) if and only if $X\backslash E$ is dense in X.

1′.10.13. Given a subset E of a normed space X, prove that E is closed if and only if $E = \overline{E}$.

1′.10.14. Let X be a Banach space.

(a) Prove that if M is a subspace of X, then its closure $\overline{M}$ is also a subspace.

(b) We define the *closed span* of a set $S \subseteq X$ to be the closure of the span of S, and we denote this closed span by $\overline{\text{span}}(S)$. Prove that $\overline{\text{span}}(S)$ is the *smallest closed subspace* of X that contains S. That is, $\overline{\text{span}}(S)$ is a closed subspace of X, and if M is any other closed subspace such that $S \subseteq M$, then $\overline{\text{span}}(S) \subseteq M$.

1′.10.15. Let X be a normed space.

(a) Given sets $A, B \subseteq X$, prove that $\overline{A \cup B} = \overline{A} \cup \overline{B}$. Use induction to extend this to the union of finitely many subsets of X.

(b) Let I be an arbitrary index set. Given sets $E_i \subseteq X$ for $i \in I$, prove that

$$\overline{\bigcap_{i \in I} E_i} \subseteq \bigcap_{i \in I} \overline{E_i}.$$

Show by example that equality can fail (even if the index set I is finite).

1′.10.16. In this problem we take $X = \ell^\infty$, with respect to the norm $\|\cdot\|_\infty$.

(a) Is c_{00} a dense subspace of ℓ^∞? Is ℓ^1 a dense subspace of ℓ^∞?

(b) Let c_0 be the space of all sequences whose components converge to zero:

$$c_0 = \Bigl\{x = (x_k)_{k \in \mathbb{N}} : \lim_{k \to \infty} x_k = 0\Bigr\} \qquad (1'.35)$$

Prove that $c_{00} \subsetneq c_0 \subsetneq \ell^\infty$, and show that the closure of c_{00} with respect to the ℓ^∞-norm is c_0, i.e., $\overline{c_{00}} = c_0$.

(c) Why does part (b) not contradict the statement made after Theorem 1′.10.5 that $\overline{c_{00}} = \ell^1$?

1′.10.17. Let S be the subset of $\mathbb{R}^d$ consisting of all sequences with rational components: $S = \bigl\{(r_1, \dots, r_d) : r_1, \dots, r_d \text{ are rational}\bigr\}$. Prove that S is a countable, dense subset of $\mathbb{R}^d$. Conclude that $\mathbb{R}^d$ is separable.

1′.10.18. Let S be the set of all "finite sequences" with rational components defined in equation (1′.34).

(a) Explain why S is contained in ℓ^1, but is not a subspace of ℓ^1.

(b) Prove that S is a countable, dense subset of ℓ^1, and conclude that ℓ^1 is separable.

1′.10.19. Given $1 \le p \le \infty$, prove the following statements.

(a) ℓ^p is complete.

(b) If p is finite then ℓ^p is separable (compare Problem 1′.10.18).

1′.10.20. Let E be a nonempty subset of a normed space X. The distance from a point $x \in X$ to E is $\text{dist}(x, E) = \inf\bigl\{\text{d}(x, y) : y \in E\bigr\}$. For each $r > 0$, let $G_r(E)$ be the set of all points that are within a distance of r from E:

$$G_r(E) \;=\; \bigl\{x \in X : \operatorname{dist}(x,E) < r\bigr\}.$$

Prove that $G_r(E)$ is an open set, and $\overline{E} = \bigcap \{G_r(E) : r > 0\}$.