II. Metric Spaces, 83-135

DOI: 10.3792/euclid/9781429799997-2

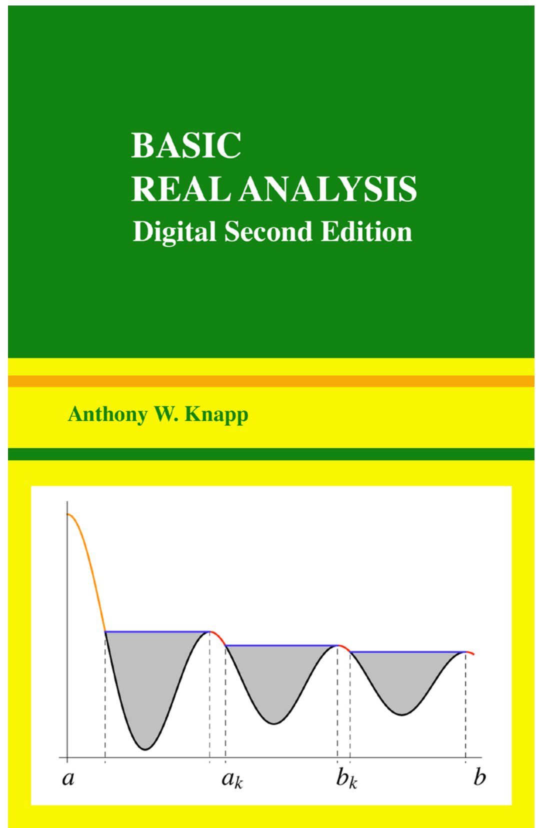
from

Basic Real Analysis

Digital Second Edition

Anthony W. Knapp

Full Book DOI: 10.3792/euclid/9781429799997
ISBN: 978-1-4297-9999-7



Anthony W. Knapp
81 Upper Sheep Pasture Road
East Setauket, N.Y. 11733–1729, U.S.A.
Email to: aknapp@math.stonybrook.edu
Homepage: www.math.stonybrook.edu/~aknapp

Title: Basic Real Analysis, with an appendix "Elementary Complex Analysis"
Cover: An instance of the Rising Sun Lemma in Section VII.1.

Mathematics Subject Classification (2010): 28–01, 26–01, 42–01, 54–01, 34–01, 30–01, 32–01.

First Edition, ISBN-13 978-0-8176-3250-2

Published by Birkhäuser Boston

Digital Second Edition, not to be sold, no ISBN

Published by the author

CHAPTER II

Metric Spaces

Abstract. This chapter is about metric spaces, an abstract generalization of the real line that allows discussion of open and closed sets, limits, convergence, continuity, and similar properties. The usual distance function for the real line becomes an example of a metric. The other notions are defined in terms of the metric. The advantage of the generalization is that proofs of certain properties of the real line immediately go over to all other examples.

Section 1 gives the definition of metric space and open set, and it lists a number of important examples, including Euclidean spaces and certain spaces of functions.

Sections 2 through 4 develop properties of open and closed sets, continuity, and convergence of sequences that are simple generalizations of known facts about $\mathbb{R}$.

Section 5 shows how a subset of a metric space can be made into a metric space so that the restriction of a continuous function from the whole space to the subset remains continuous. It also shows that three natural metrics for the product of two metric spaces lead to the same open sets, continuous functions, and convergent sequences.

Section 6 shows that any metric space is "Hausdorff," "regular," and "normal," and it goes on to exhibit three different countability hypotheses about a metric space as equivalent. A metric space with these properties is called "separable."

Section 7 concerns compactness and completeness. A metric space is defined to be "compact" if every open cover has a finite subcover. This property is equivalent to the condition that every sequence has a convergent subsequence. The Heine–Borel Theorem says that the compact sets of $\mathbb{R}^n$ are exactly the closed bounded sets. A number of the results early in Chapter I that were proved by the Bolzano–Weierstrass Theorem in the context of the real line are seen to extend to any compact metric space. A metric space is "complete" if every Cauchy sequence is convergent. A metric space is compact if and only if it is complete and "totally bounded."

Section 8 concerns connectedness, which is an abstraction of the property of an interval of the line that accounts for the Intermediate Value Theorem.

Section 9 proves a fundamental result known as the Baire Category Theorem. A sample consequence of the theorem is that the pointwise limit of a sequence of continuous complex-valued functions on a complete metric space must have points where it is continuous.

Section 10 studies the spaces of real-valued and complex-valued continuous functions on a compact metric space. A generalization of Ascoli's Theorem from the setting of Chapter I provides a characterization of compact sets in either of these spaces of continuous functions. A generalization of the Weierstrass Approximation Theorem, known as the Stone–Weierstrass Theorem, gives sufficient conditions for a subalgebra of either of these spaces of continuous functions to be dense. One consequence is that these spaces of continuous functions are separable.

Section 11 constructs the "completion" of a metric space out of Cauchy sequences in the given space. The result is a complete metric space and a distance-preserving map of the given metric space into the completion such that the image is dense.

1. Definition and Examples

Let X be a nonempty set. A function d from $X \times X$, the set of ordered pairs of members of X, to the real numbers is a **metric**, or distance function, if

(i) $d(x, y) \geq 0$ always, with equality if and only if $x = y$,
(ii) $d(x, y) = d(y, x)$ for all x and y in X,
(iii) $d(x, y) \leq d(x, z) + d(z, y)$ for all x, y, and z, the **triangle inequality**.

In this case the pair (X, d) is called a **metric space**.

The real line $\mathbb{R}^1$ with metric $d(x, y) = |x - y|$ is the motivating example. Properties (i) and (ii) are apparent, and property (iii) is readily verified one case at a time according as z is less than both x and y, z is between x and y, or z is greater than both x and y.

We come to further examples in a moment. Particularly in the case that X is a space of functions, a space may turn out to be almost a metric space but not to satisfy the condition that $d(x, y) = 0$ implies $x = y$. Accordingly we introduce a weakened version of (i) as

(i′) $d(x, y) \geq 0$ and $d(x, x) = 0$ always,

and we say that a function d from $X \times X$ to the real numbers is a **pseudometric** if (i′), (ii), and (iii) hold. In this case, (X, d) is called a **pseudometric space**.

Let (X, d) be a pseudometric space. If $r > 0$, the **open ball** of radius r and center x, denoted by $B(r; x)$, is the set of points at distance less than r from x, namely

$$B(r; x) = \{y \in X \mid d(x, y) < r\}.$$

The name "ball" will be appropriate in Euclidean space in dimension three, which is part of the Example 1 below, and "ball" is adopted for the corresponding notion in a general pseudometric space.

A subset U of X is **open** if for each x in U and some sufficiently small $r > 0$, the open ball $B(r; x)$ is contained in U. For the line the open balls in the above sense are just the bounded open intervals, and the open sets in the above sense are the usual open sets in the sense of Chapter I.

Lemma 2.1. In any pseudometric space (X, d), every open ball is an open set. The open sets are exactly all possible unions of open balls.

PROOF. Let an open ball $B(r; x)$ be given. If y is in $B(r; x)$, then the open ball $B(r - d(x, y),\ y)$ has center y and positive radius; we show that it is contained in $B(r; x)$. In fact, if z is in $B(r - d(x, y),\ y)$, then the triangle inequality gives

$$d(x, z) \leq d(x, y) + d(y, z) < d(x, y) + (r - d(x, y)) = r,$$

and the containment follows.

For the second assertion it follows from the definition of open set that every open set is the union of open balls. In the reverse direction, let U be a union of open balls. If y is in U, then y lies in one of these balls, say in $B(r; x)$. We have just shown that some open ball $B(s; y)$ is contained in $B(r; x)$, and $B(r; x)$ is contained in U. Thus $B(s; y)$ is contained in U, and U is open. $\square$

EXAMPLES.

(1) **Euclidean space** $\mathbb{R}^n$. Fix an integer $n > 0$. Let $\mathbb{R}^n$ be the space of all n-tuples of real numbers $x = (x_1, \dots, x_n)$. We define addition of n-tuples componentwise, and we define scalar multiplication by $cx = (cx_1, \dots, cx_n)$ for real c. Following the normal convention in linear algebra, we identify this space with the real vector space, also denoted by $\mathbb{R}^n$, of all n-component column vectors of real numbers $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$. Generalizing the notion of absolute value when $n = 1$, we let $|x| = \big(\sum_{j=1}^n x_j^2\big)^{1/2}$ for $x = (x_1, \dots, x_n)$ in $\mathbb{R}^n$. The quantity $|x|$ is the **Euclidean norm** of x. The Euclidean norm satisfies the properties

(a) $|x| \ge 0$ always, with equality if and only if x equals the zero tuple $0 = (0, \dots, 0)$,
(b) $|cx| = |c||x|$ for all x and for all real c,
(c) $|x + y| \le |x| + |y|$ for all x and y.

Properties (a) and (b) are apparent, but (c) requires proof. The proof makes use of the familiar **dot product**, given by $x \cdot y = \sum_{j=1}^n x_j y_j$ if $x = (x_1, \dots, x_n)$ and $y = (y_1, \dots, y_n)$. In terms of dot product, the Euclidean norm is nothing more than $|x| = (x \cdot x)^{1/2}$. The dot product satisfies the important inequality $|x \cdot y| \le |x||y|$, known as the **Schwarz inequality** and proved for this context in Section A5 of Appendix A at the end of the book. A more general version of the Schwarz inequality will be stated and proved in Lemma 2.2 below. The Schwarz inequality implies (c) above because we then have

$$\begin{aligned} |x + y|^2 &= (x + y) \cdot (x + y) = x \cdot x + 2(x \cdot y) + y \cdot y \\ &= |x|^2 + 2(x \cdot y) + |y|^2 \le |x|^2 + 2|x||y| + |y|^2 = (|x| + |y|)^2. \end{aligned}$$

We make $X = \mathbb{R}^n$ into a metric space (X, d) by defining

$$d(x, y) = |x - y|.$$

Properties (i) and (ii) of a metric are immediate from (a) and (b), respectively; property (iii) follows from (c) in the form $|a + b| \le |a| + |b|$ if we substitute $a = x - z$ and $b = z - y$. For $n = 1$, this example reduces to the line as discussed above. For $n = 2$, open balls are geometric open disks, while for $n = 3$, open balls are geometric open balls. For any n, the open sets in the metric space coincide with the open sets as defined in calculus of several variables.

(2) **Complex Euclidean space** $\mathbb{C}^n$. The space $\mathbb{C}$ of complex numbers, with distance function $d(z, w) = |z - w|$ as in Section I.5, can be seen in two ways to be a metric space. One way was carried out in Section I.5 and directly uses the properties of the absolute value function $|z|$ in Section A4 of Appendix A. The other way is to identify $z = x + iy$ with the member (x, y) of $\mathbb{R}^2$, and then the absolute value $|z|$ equals the Euclidean norm $|(x, y)|$ in the sense of Example 1; hence the construction of Example 1 makes the set of complex numbers into a metric space. More generally the complex vector space $\mathbb{C}^n$ of n-tuples

$$z = (z_1, \dots, z_n) = (x_1, \dots, x_n) + i(y_1, \dots, y_n) = x + iy$$

becomes a metric space in two equivalent ways. One way is to define the norm $|z| = \big(\sum_{j=1}^n |z_j|^2\big)^{1/2}$ as a generalization of the Euclidean norm for $\mathbb{R}^n$; then we put $d(z, w) = |z - w|$. The argument that d satisfies the triangle inequality is a variant of the one for $\mathbb{R}^n$: The object for $\mathbb{C}^n$ that generalizes the dot product for $\mathbb{R}^n$ is the **Hermitian inner product**

$$(z, w) = \big((z_1, \dots, z_n), (w_1, \dots, w_n)\big) = \sum_{j=1}^n z_j \overline{w_j}.$$

The **Euclidean norm** is given in terms of this expression by $|z| = (z, z)^{1/2}$, and the version of the Schwarz inequality in Section A5 of Appendix A is general enough to show that $|(z, w)| \le |z||w|$. The same argument as for Example 1 shows that the norm satisfies the triangle inequality, and then it follows that d satisfies the triangle inequality. The other way to view $\mathbb{C}^n$ as a metric space is to identify $\mathbb{C}^n$ with $\mathbb{R}^{2n}$ by $(z_1, \dots, z_n) \mapsto (x_1, \dots, x_n, y_1, \dots, y_n)$ and then to use the metric on $\mathbb{R}^{2n}$ from Example 1. This is the same metric, since $\sum_{j=1}^n |z_j|^2 = \sum_{j=1}^n x_j^2 + \sum_{j=1}^n y_j^2$. We still get the same metric if we instead use the identification $(z_1, \dots, z_n) \mapsto (x_1, y_1, \dots, x_n, y_n)$. With either identification the Hermitian inner product (z, w) for $\mathbb{C}^n$ corresponds to the ordinary dot product for $\mathbb{R}^{2n}$.

(3) System $\mathbb{R}^*$ of extended real numbers. The function $f(x) = x/(1 + x)$ carries $[0, +\infty)$ into $[0, +1)$ and has $g(y) = y/(1 - y)$ as a two-sided inverse. Therefore f is one-one and onto. We can extend f so that it carries $(-\infty, +\infty)$ one-one onto $(-1, +1)$ by putting $f(x) = x/(1 + |x|)$. We can extend f further by putting $f(-\infty) = -1$ and $f(+\infty) = +1$, and then f carries $[-\infty, +\infty]$, i.e., all of $\mathbb{R}^*$, one-one onto $[-1, +1]$. The function f is nondecreasing on $[-\infty, +\infty]$. For x and x' in $\mathbb{R}^*$, let

$$d(x, x') = |f(x) - f(x')|.$$

We shall show that d is a metric. By inspection, d satisfies properties (i) and (ii) of a metric, and we are to prove the triangle inequality (iii), namely that

$$d(x, x') \le d(x, x'') + d(x'', x').$$

The critical fact is that f is nondecreasing. Since d satisfies (ii), we may assume that $x \le x'$, and then

$$d(x, x') = f(x') - f(x).$$

We divide the proof into three cases, depending on the location of x'' relative to x and x'. The first case is that $x'' \le x$, and then

$$d(x, x'') + d(x'', x') = f(x) - f(x'') + f(x') - f(x'').$$

Thus the question is whether

$$f(x') - f(x) \overset{?}{\le} f(x) - f(x'') + f(x') - f(x''),$$

hence whether

$$2f(x'') \overset{?}{\le} 2f(x).$$

This inequality holds, since f is nondecreasing. The second case is that $x \le x'' \le x'$, and then

$$d(x, x'') + d(x'', x') = f(x'') - f(x) + f(x') - f(x'') = f(x') - f(x).$$

Hence equality holds in the triangle inequality. The third case is that $x' \le x''$, and then

$$d(x, x'') + d(x'', x') = f(x'') - f(x) + f(x'') - f(x').$$

The triangle inequality comes down to the question whether

$$2f(x') \overset{?}{\le} 2f(x'').$$

This inequality holds, since f is nondecreasing. We conclude that $(\mathbb{R}^*, d)$ is a metric space. It is not hard to see that the open balls in $\mathbb{R}^*$ are all intervals (a, b), $[-\infty, b)$, $(a, +\infty]$, and $[-\infty, +\infty]$ with $-\infty \le a < b \le +\infty$. Each of these open balls in $\mathbb{R}^*$ intersects $\mathbb{R}$ in an ordinary open interval, bounded or unbounded. The open sets in $\mathbb{R}$ therefore coincide with the intersections of $\mathbb{R}$ with the open sets of $\mathbb{R}^*$.

(4) Bounded functions in the **uniform metric**. Let S be a nonempty set, and let $X = B(S)$ be the set of all "scalar"-valued functions f on S that are **bounded** in the sense that $|f(s)| \le M$ for all $s \in S$ and for a constant M depending on f. The **scalars** are allowed to be the members of $\mathbb{R}$ or the members of $\mathbb{C}$, and it will ordinarily make no difference which one is understood. If it does make a difference, we shall write $B(S, \mathbb{R})$ or $B(S, \mathbb{C})$ to be explicit about the range. For f and g in $B(S)$, let

$$d(f, g) = \sup_{s \in S} |f(s) - g(s)|.$$

It is easy to verify that (X, d) is a metric space. Let us not lose sight of the fact that the members of X are functions. When we discuss convergence of sequences in a metric space, we shall see that a sequence of functions in this X converges if and only if the sequence of functions converges uniformly on S.

(5) Generalization of Example 4. We can replace the range $\mathbb{R}$ or $\mathbb{C}$ of the functions in Example 4 by any metric space (R, ρ). Fix a point r_0 in the range R. A function $f : S \to R$ is **bounded** if $\rho(f(s), r_0) \le M$ for all s and for some M depending on f. This definition is independent of the choice of r_0 because ρ is assumed to satisfy the triangle inequality. If we let X be the space of all such bounded functions from S to R, we can make X into a metric space by defining $d(f, g) = \sup_{s \in S} \rho(f(s), g(s))$.

(6) Sequence space ℓ^2. This is the space of all sequences $\{c_n\}_{n=-\infty}^{\infty}$ of scalars with $\sum |c_n|^2 < \infty$. A metric is given by

$$d(\{c_n\}, \{d_n\}) = \Big(\sum_{n=-\infty}^{\infty} |c_n - d_n|^2 \Big)^{1/2}.$$

In the case of complex scalars, this example arises as a natural space containing all systems of Fourier coefficients of Riemann integrable functions on $[-\pi, \pi]$, in the sense of Chapter I. Proving the triangle inequality involves arguing as in Examples 1 and 2 above and then letting the number of terms tend to infinity. The role of the dot product is played by $(\{c_n\}, \{d_n\}) = \sum_{n=-\infty}^{\infty} c_n \overline{d_n}$.

(7) **Indiscrete space.** If X is any nonempty set and if $d(x, y) = 0$ for all x and y, then d is a pseudometric and the only open sets are X and the empty set $\varnothing$. If X contains more than one element, then d is not a metric.

(8) **Discrete metric.** If X is any nonempty set and if

$$d(x, y) = \begin{cases} 1 & \text{if } x \neq y, \\ 0 & \text{if } x = y, \end{cases}$$

then d is a metric, and every subset of X is open.

(9) Let S be a nonempty set, fix an integer $n > 0$, and let X be the set of n-tuples of members of S. For n-tuples $x = (x_1, \dots, x_n)$ and $y = (y_1, \dots, y_n)$, define

$$d(x, y) = \#\{j \mid x_j \neq y_j\},$$

the number of components in which x and y differ. Then (X, d) is a metric space. The proof of the triangle inequality requires a little argument, but we leave that for Problem 1 at the end of the chapter. Every subset of X is open, just as with the discrete metric in Example 8.

(10) **Hedgehog space.** Let X be $\mathbb{R}^2$, and single out the origin for special attention. Let d be the metric of Euclidean space, and define

$$\rho(x, y) = \begin{cases} d(x, y) & \text{if } x \text{ and } y \text{ are on the same ray from } 0, \\ d(x, 0) + d(0, y) & \text{otherwise.} \end{cases}$$

Then ρ is a metric. Every open set in (X, d) is open in (X, ρ), but a set like the one in Figure 2.1 is open in (X, ρ) but not in (X, d).

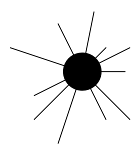


FIGURE 2.1. An open set centered at the origin in the hedgehog space.

(11) **Hilbert cube.** Let X be the set of all sequences $\{x_m\}_{m \geq 1}$ of real numbers satisfying $0 \leq x_m \leq 1$ for all m, and put

$$d(\{x_m\}, \{y_m\}) = \sum_{m=1}^{\infty} 2^{-m} |x_m - y_m|.$$

Then (X, d) is a metric space. To verify the triangle inequality, we can argue as follows: Let $\{x_m\}$, $\{y_m\}$, and $\{z_m\}$ be in X. For each m, we have

$$2^{-m}|x_m - y_m| \leq 2^{-m}|x_m - z_m| + 2^{-m}|z_m - y_m|.$$

Thus

$$\begin{aligned} \sum_{m=1}^{N} 2^{-m}|x_m - y_m| &\leq \sum_{m=1}^{N} 2^{-m}|x_m - z_m| + \sum_{m=1}^{N} 2^{-m}|z_m - y_m| \\ &\leq \sum_{m=1}^{\infty} 2^{-m}|x_m - z_m| + \sum_{m=1}^{\infty} 2^{-m}|z_m - y_m| \end{aligned}$$

for each N. Letting N tend to infinity yields the desired inequality.

(12) L^1 metric on Riemann integrable functions. Fix a nontrivial bounded interval $[a, b]$ of the line, let X be the set of all Riemann integrable complex-valued functions on $[a, b]$ in the sense of Chapter I, and define

$$d_1(f, g) = \int_a^b |f(x) - g(x)|\, dx$$

for f and g in X. Then (X, d_1) is a pseudometric space. It can happen that $\int_a^b |f(x) - g(x)|\, dx = 0$ without $f = g$; for example, f could differ from g at a single point. Therefore d_1 is not a metric.

(13) L^2 metric on complex-valued $\mathcal{R}[-\pi, \pi]$. This example arose in the discussion of Fourier series in Section I.10, and it was convenient to include a factor $\frac{1}{2\pi}$ in front of integrals. Let $X = \mathcal{R}[-\pi, \pi]$, and define

$$d_2(f, g) = \Big(\frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x) - g(x)|^2\, dx\Big)^{1/2}.$$

Then (X, d_2) is a pseudometric metric space. The triangle inequality was proved in Lemma 1.64 using the version of the Schwarz inequality in Lemma 1.63; that version of the Schwarz inequality needed a special argument given in Lemma 1.62 in order to handle functions f whose norm satisfies $\|f\|_2 = 0$.

The constructions of metric spaces in Examples 1, 2, 6, and 13 are sufficiently similar to warrant abstracting what was involved. We start with a real or complex vector space V, possibly infinite-dimensional, and with a generalization $(\,\cdot\,,\,\cdot\,)$ of dot product. This generalization is a function from $V \times V$ to $\mathbb{R}$ in the case that V is real, and it is a function from $V \times V$ to $\mathbb{C}$ in the case that V is complex. We shall write the scalars as if they are complex, but only real scalars are to be used if the vector space is real. The function is written $(\,\cdot\,,\,\cdot\,)$ and is assumed to satisfy the following properties:

(i) it is linear in the first variable, i.e., $(x_1 + x_2, y) = (x_1, y) + (x_2, y)$ and $(cx, y) = c(x, y)$,

(ii) it is conjugate linear in the second variable, i.e., $(x, y_1 + y_2) = (x, y_1) + (x, y_2)$ and $(x, cy) = \bar{c}(x, y)$,

(iii) it is **symmetric** in the real case and **Hermitian symmetric** in the complex case, i.e., $(y, x) = \overline{(x, y)}$,

(iv) it is **definite**, i.e., $(x, x) > 0$ if $x \neq 0$.

The form $(\,\cdot\,,\,\cdot\,)$ is called an **inner product** if V is real or complex and is often called also a **Hermitian inner product** if V is complex; in either case, V with the form is called an **inner-product space**. Two vectors x and y with $(x, y) = 0$ are said to be **orthogonal**; the notion of orthogonality generalizes perpendicularity in the case of the dot product.

For either kind of scalars, we define $\|x\| = (x, x)^{1/2}$, and the function $\|\cdot\|$ is called the associated **norm**. We shall see shortly that a version of the Schwarz inequality is valid in this generality, the proof being no more complicated than the one in Section A5 of Appendix A.

In many cases in practice, item (iv) is replaced by the weaker condition that

(iv′) $(\cdot, \cdot)$ is **semidefinite**, i.e., $(x, x) \geq 0$ if $x \neq 0$.

This was what happened in Example 13 above. In order to have a name for this kind of space, let us call V with the semidefinite form $(\cdot, \cdot)$ a **pseudo inner-product space**. It is still meaningful to speak of orthogonality. It is still meaningful also to define $\|x\| = (x, x)^{1/2}$, and this is called the **pseudonorm** for the space. The Schwarz inequality is still valid, but its proof is more complicated than for an inner-product space. The extra complication was handled by Lemma 1.62 in the case of Example 13 in order to obtain a little extra information; the general argument proceeds along different lines.

Lemma 2.2 (Schwarz inequality). Let V be a pseudo inner-product space with form $(\cdot, \cdot)$. If x and y are in V, then $|(x, y)| \leq \|x\|\|y\|$.

PROOF. First suppose that $\|y\| \neq 0$. Then

$$\begin{aligned}0 &\leq \big\|x - \|y\|^{-2}(x, y)y\big\|^2 = \big((x - \|y\|^{-2}(x, y)y),\ (x - \|y\|^{-2}(x, y)y)\big)\\ &= \|x\|^2 - 2\|y\|^{-2}|(x, y)|^2 + \|y\|^{-4}|(x, y)|^2\|y\|^2 = \|x\|^2 - \|y\|^{-2}|(x, y)|^2,\end{aligned}$$

and the inequality follows in this case.

Next suppose that $\|y\| = 0$. It is enough to prove that $(x, y) = 0$ for all x. If c is a real scalar, we have

$$\|x+cy\|^2 = (x+cy, x+cy) = \|x\|^2 + 2\operatorname{Re}(x, cy) + |c|^2\|y\|^2 = \|x\|^2 + 2c\operatorname{Re}(x, y).$$

The left side is ≥ 0 as c varies, but the right side can be < 0 unless $\operatorname{Re}(x, y) = 0$. Thus we must have $\operatorname{Re}(x, y) = 0$ for all x. Replacing x by ix gives us $\operatorname{Im}(x, y) = -\operatorname{Re} i(x, y) = -\operatorname{Re}(ix, y)$, and this we have just shown is 0 for all x. Thus $\operatorname{Re}(x, y) = \operatorname{Im}(x, y) = 0$, and $(x, y) = 0$. ☐

Proposition 2.3 (triangle inequality). If V is a pseudo inner-product space with form $(\cdot, \cdot)$ and pseudonorm $\|\cdot\|$, then the pseudonorm satisfies

(a) $\|x\| \geq 0$ for all $x \in V$,
(b) $\|cx\| = |c|\|x\|$ for all scalars c and all $x \in V$,
(c) $\|x + y\| \leq \|x\| + \|y\|$ for all x and y in V.

Moreover, the definition $d(x, y) = \|x - y\|$ makes V into a pseudometric space. The space V is a metric space if the pseudo inner-product space is an inner-product space.

PROOF. Properties (a) and (b) of the pseudonorm are immediate, and (c) follows because

$$\begin{aligned}\|x+y\|^2 &= (x+y,\ x+y) = (x,x) + 2\operatorname{Re}(x,y) + (y,y)\\ &= \|x\|^2 + 2\operatorname{Re}(x,y) + \|y\|^2 \le \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = (\|x\|+\|y\|)^2.\end{aligned}$$

Putting $x = a - c$ and $y = c - b$ gives $d(a,b) \le d(a,c) + d(c,b)$, and thus d satisfies the triangle inequality for a pseudometric. The other properties of a pseudometric are immediate from (a) and (b). If the form is definite and $d(f,g) = 0$, then $(f-g,\ f-g) = 0$ and hence the definiteness yields $f - g = 0$. $\square$

EXAMPLES, CONTINUED.

14) Let us take double integrals of continuous functions of nice subsets of $\mathbb{R}^2$ as known. (The detailed study of general Riemann integrals in several variables occurs in Chapter III.) Let V be the complex vector space of all power series $F(z) = \sum_{n=0}^{\infty} c_n z^n$ with infinite radius of convergence. Since any such $F(z)$ is bounded on the open unit disk $D = \{z \in \mathbb{C} \mid |z| < 1\}$, the form $(F,G) = \int_D F(z)\overline{G(z)}\,dx\,dy$ is meaningful and makes V into an inner-product space. The proposition shows that V becomes a metric space with metric given by $d(F,G) = \big(\int_D |F(z) - G(z)|^2\,dx\,dy\big)^{1/2}$.

2. Open Sets and Closed Sets

In this section we generalize the Euclidean notions of open set, closed set, neighborhood, interior, limit point, and closure so that they make sense for all pseudometric spaces, and we prove elementary properties relating these metric-space notions. In working with metric spaces and pseudometric spaces, it is often helpful to draw pictures as if the space in question were $\mathbb{R}^2$, even computing distances that are right for $\mathbb{R}^2$. We shall do that in the case of the first lemma but not afterward in this section. Let (X,d) be a pseudometric space.

Lemma 2.4. If z is in the intersection of open balls $B(r;x)$ and $B(s;y)$, then there exists some $t > 0$ such that the open ball $B(t;z)$ is contained in that intersection. Consequently the intersection of two open balls is open.

REMARK. Figure 2.2 shows what $B(t;z)$ looks like in the metric space $\mathbb{R}^2$.

PROOF. Take $t = \min\{r - d(x,z),\ s - d(y,z)\}$. If w is in $B(t;z)$, then the triangle inequality gives

$$d(x,w) \le d(x,z) + d(z,w) < d(x,z) + t \le d(x,z) + (r - d(x,z)) = r,$$

and hence w is in $B(r;x)$. Similarly w is in $B(s;y)$. $\square$

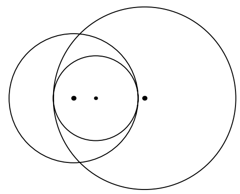


FIGURE 2.2. Open ball contained in an intersection of two open balls.

Proposition 2.5. The open sets of X have the properties that

(a) X and the empty set $\varnothing$ are open,
(b) an arbitrary union of open sets is open,
(c) any finite intersection of open sets is open.

PROOF. We know from Lemma 2.1 that a set is open if and only if it is the union of open balls. Then (b) is immediate, and (a) follows, since X is the union of all open balls and $\varnothing$ is an empty union. For (c), it is enough to prove that $U \cap V$ is open if U and V are open. Write $U = \bigcup_\alpha B_\alpha$ and $V = \bigcup_\beta B_\beta$ as unions of open balls. Then $U \cap V = \bigcup_{\alpha,\beta} (B_\alpha \cap B_\beta)$, and Lemma 2.4 shows that $U \cap V$ is exhibited as the union of open balls. Thus $U \cap V$ is open. $\square$

A **neighborhood** of a point in X is any set that contains an open set containing the point. An **open neighborhood** is a neighborhood that is an open set.[1] A **neighborhood** of a subset E of X is a set that is a neighborhood of each point of E. If A is a subset of X, then the set A^o of all points x in A for which A is a neighborhood of x is called the **interior** of A. For example, the interior of the half-open interval $[a, b)$ of the real line is the open interval (a, b).

Proposition 2.6. The interior of a subset A of X is the union of all open sets contained in A; that is, it is the largest open set contained in A.

PROOF. Suppose that $U \subseteq A$ is open. If x is in U, then U is an open neighborhood of x, and hence A is a neighborhood of x. Thus x is in A^o, and A^o contains the union of all open sets contained in A. For the reverse inclusion, let x be in A^o. Then A is a neighborhood of x, and there exists an open subset U of A containing x. So x is contained in the union of all open sets contained in A. $\square$

Corollary 2.7. A subset A of X is open if and only if $A = A^o$.

A subset F of X is **closed** if its complement is open. Every closed interval of the real line is closed. A half-open interval $[a, b)$ on the real line is neither open nor closed if a and b are both finite.

[1] Some authors use the term "neighborhood" to mean what is here called "open neighborhood."

Proposition 2.8. The closed sets of X have the properties that

(a) X and the empty set $\varnothing$ are closed,
(b) an arbitrary intersection of closed sets is closed,
(c) any finite union of closed sets is closed.

PROOF. This result follows from Proposition 2.5 by taking complements. In (a), the complements of X and $\varnothing$ are $\varnothing$ and X, respectively. For (b) and (c), we use the formulas $\big(\bigcap_\alpha F_\alpha\big)^c = \bigcup_\alpha F_\alpha^c$ and $\big(\bigcup_\alpha F_\alpha\big)^c = \bigcap_\alpha F_\alpha^c$ for the complements of intersections and unions. $\square$

If A is a subset of X, then x in X is a **limit point** of A if each neighborhood of x contains a point of A distinct from x. The **closure**[2] A^{cl} of A is the union of A with the set of all limit points of A. For example, the limit points of the set $[a, b) \cup \{b+1\}$ on the real line are the points of the closed interval $[a, b]$, and the closure of the set is $[a, b] \cup \{b+1\}$.

Proposition 2.9. A subset A of X is closed if and only if it contains all its limit points.

PROOF. Suppose A is closed, so that A^c is open. If x is in A^c, then A^c is an open neighborhood of x disjoint from A, so that x cannot be a limit point of A. Thus all limit points of A lie in A. In the reverse direction suppose that A contains all its limit points. If x is in A^c, then x is not a limit point of A, and hence there exists an open neighborhood of x lying completely in A^c. Since x is arbitrary, A^c is open, and thus A is closed. $\square$

Proposition 2.10. The closure A^{cl} of a subset A of X is closed. The closure of A is the intersection of all closed sets containing A; that is, it is the smallest closed set containing A.

PROOF. We shall apply Proposition 2.9. If x is given as a limit point of A^{cl}, we are to see that x is in A^{cl}. Assume the contrary. Then x is not in A, and x is not a limit point of A. Because of the latter condition, there exists an open neighborhood U of x that does not meet A except possibly in x. Because of the former condition, U does not meet A at all. Since x is a limit point of A^{cl}, U contains a point y of A^{cl}. Since U does not meet A, y has to be a limit point of A. Since U is an open neighborhood of y, U has to contain a point of A, and we have a contradiction. We conclude that x is in A^{cl}, and Proposition 2.9 shows that A^{cl} is closed.

Any closed set F containing A contains all its limit points, by Proposition 2.9, and hence contains all the limit points of A. Thus $F \supseteq A^{\mathrm{cl}}$. Since A^{cl}

[2]Some authors write $\overline{A}$ instead of A^{cl} for the closure of A.

itself is a closed set containing A, it follows that A^{cl} is the smallest closed set containing A. $\square$

Corollary 2.11. A subset A of X is closed if and only if $A = A^{\mathrm{cl}}$. Consequently $(A^{\mathrm{cl}})^{\mathrm{cl}} = A^{\mathrm{cl}}$ for any subset A of X.

Two remarks are in order. The first remark is that the proofs of all the results from Proposition 2.6 through Corollary 2.11 use only that the family of open subsets of X satisfies properties (a), (b), and (c) in Proposition 2.5 and do not actually depend on the precise definition of "open set." This observation will be of importance to us in Chapter X, when properties (a), (b), and (c) will be taken as an axiomatic definition of a "topology" of open sets for X, and then all the results from Proposition 2.6 through Corollary 2.11 will still be valid.

The second remark is that the mathematics of pseudometric spaces can always be reduced to the mathematics of metric spaces, and we shall normally therefore work only with metric spaces. The device for this reduction is given in the next proposition, which uses the notion of an equivalence relation. Equivalence relations are taken as known but are reviewed in Section A6 of Appendix A.

Proposition 2.12. Let (X, d) be a pseudometric space. If members x and y of X are called equivalent whenever $d(x, y) = 0$, then the result is an equivalence relation. Denote by $[x]$ the equivalence class of x and by X_0 the set of all equivalence classes. The definition $d_0([x], [y]) = d(x, y)$ consistently defines a function $d_0 : X_0 \times X_0 \to \mathbb{R}$, and (X_0, d_0) is a metric space. A subset A is open in X if and only if two conditions are satisfied: A is a union of equivalence classes, and the set A_0 of such classes is an open subset of X_0.

PROOF. The reflexive, symmetric, and transitive properties of the relation "equivalent" are immediate from the defining properties of a metric. Let x and x' be equivalent, and let y and y' be equivalent. Then

$$d(x, y) \le d(x, x') + d(x', y') + d(y', y) = 0 + d(x', y') + 0 = d(x', y'),$$

and similarly

$$d(x', y') \le d(x, y).$$

Thus $d(x, y) = d(x', y')$, and d_0 is well defined. The properties showing that d_0 is a metric are immediate from the corresponding properties for d.

Next let x be in an open set A, and let x' be equivalent to x. Since A is open, some open ball $B(r; x)$ is contained in A. Since x' has $d(x, x') = 0$, x' lies in $B(r; x)$. Thus x' lies in A, and A is the union of equivalence classes.

Finally let A be any union of equivalence classes, and let A_0 be the set of those classes. If x is in A, then the set of points in some equivalence class lying in $B(r; [x])$ is just $B(r; x)$, and it follows that A is open in X if and only if A_0 is open in X_0. $\square$

3. Continuous Functions

Before we discuss continuous functions between metric spaces, let us take note of some properties of inverse images for abstract functions as listed in Section A1 of Appendix A. If $f : X \to Y$ is a function between two sets X and Y and E is a subset of Y, we denote by $f^{-1}(E)$ the inverse image of E under f, i.e., $\{x \in X \mid f(x) \in E\}$. The properties are that inverse images of functions respect unions, intersections, and complements.

Let (X, d) and (Y, ρ) be metric spaces. A function $f : X \to Y$ is **continuous at a point** $x \in X$ if for each $\epsilon > 0$, there is a $\delta > 0$ such that $\rho(f(x), f(y)) < \epsilon$ whenever $d(x, y) < \delta$. This definition is consistent with the definition when (X, d) and (Y, ρ) are both equal to $\mathbb{R}$ with the usual metric.

Proposition 2.13. If (X, d) and (Y, ρ) are metric spaces, then a function $f : X \to Y$ is continuous at the point $x \in X$ if and only if for any open neighborhood V of $f(x)$ in Y, there is a neighborhood U of x such that $f(U) \subseteq V$.

PROOF. Let f be continuous at x and let V be given. Choose $\epsilon > 0$ such that $B(\epsilon; f(x))$ is contained in V, and choose $\delta > 0$ such that $\rho(f(x), f(y)) < \epsilon$ whenever $d(x, y) < \delta$. Then $y \in B(\delta; x)$ implies $f(y) \in B(\epsilon; f(x)) \subseteq V$. Thus $U = B(\delta; x)$ has $f(U) \subseteq V$.

Conversely suppose that f satisfies the condition in the statement of the proposition. Let $\epsilon > 0$ be given, and choose a neighborhood U of x such that $f(U) \subseteq B(\epsilon; f(x))$. Since U is a neighborhood of x, we can find an open ball $B(\delta; x)$ lying in U. Then $f(B(\delta; x)) \subseteq B(\epsilon; f(x))$, and hence $\rho(f(x), f(y)) < \epsilon$ whenever $d(x, y) < \delta$. □

Corollary 2.14. Let $f : X \to Y$ and $g : Y \to Z$ be functions between metric spaces. If f is continuous at x and g is continuous at $f(x)$, then the composition $g \circ f$, given by $(g \circ f)(y) = g(f(y))$, is continuous at x.

PROOF. Let W be an open neighborhood of $g(f(x))$. By continuity of g at $f(x)$, we can choose a neighborhood V of $f(x)$ such that $g(V) \subseteq W$. Possibly by passing to a subset of V, we may assume that V is an open neighborhood of $f(x)$. By continuity of f at x, we can choose a neighborhood U of x such that $f(U) \subseteq V$. Then $g(f(U)) \subseteq W$. Taking Proposition 2.13 into account, we see that $g \circ f$ is continuous at x. □

Proposition 2.15. If (X, d) and (Y, ρ) are metric spaces and f is a function from X into Y, then the following are equivalent:

(a) the function f is continuous at every point of X,
(b) the inverse image under f of every open set in Y is open in X,
(c) the inverse image under f of every closed set in Y is closed in X.

PROOF. Suppose (a) holds. If V is open in Y and x is in $f^{-1}(V)$, then $f(x)$ is in V. Since f is continuous at x by (a), Proposition 2.13 gives us a neighborhood U of x, which we may take to be open, such that $f(U) \subseteq V$. Then we have $x \in U \subseteq f^{-1}(V)$. Since x is arbitrary in $f^{-1}(V)$, $f^{-1}(V)$ is open. Thus (b) holds. In the reverse direction, suppose (b) holds. Let x in X be given, and let V be an open neighborhood of $f(x)$. By (b), $U = f^{-1}(V)$ is open, and U is then an open neighborhood of x mapping into V. This proves (a), and thus (a) and (b) are equivalent. Conditions (b) and (c) are equivalent, since $f^{-1}(V)^c = f^{-1}(V^c)$. □

A function $f : X \to Y$ that is continuous at every point of X, as in Proposition 2.15, will simply be said to be **continuous**. A function $f : X \to Y$ is a **homeomorphism** if f is continuous, if f is one-one and onto, and if $f^{-1} : Y \to X$ is continuous. The relation "is homeomorphic to" is an equivalence relation. Namely, the identity function shows that the relation is reflexive, the symmetry of the relation is built into the definition, and the transitivity follows from Corollary 2.14.

If (X, d) is a metric space and if A is a nonempty subset of X, then the **distance** from x to A, denoted by $D(x, A)$, is defined by

$$D(x, A) = \inf_{y \in A} d(x, y).$$

Proposition 2.16. Let A be a fixed nonempty subset of a metric space (X, d). Then the real-valued function f defined on X by $f(x) = D(x, A)$ is continuous.

PROOF. If x and y are in X and z is in A, then the triangle inequality gives

$$D(x, A) \le d(x, z) \le d(x, y) + d(y, z).$$

Taking the infimum over z gives $D(x, A) \le d(x, y) + D(y, A)$. Reversing the roles of x and y, we obtain $D(y, A) \le d(x, y) + D(x, A)$, since $d(y, x) = d(x, y)$. Therefore

$$|f(x) - f(y)| = |D(x, A) - D(y, A)| \le d(x, y).$$

Fix x, let $\epsilon > 0$ be given, and take $\delta = \epsilon$. If $d(x, y) < \delta = \epsilon$, then our inequality gives us $|f(x) - f(y)| < \epsilon$. Hence f is continuous at x. Since x is arbitrary, f is continuous. □

Corollary 2.17. If (X, d) is a metric space, then the real-valued function $d(x, y)$ for fixed y is continuous in x.

PROOF. This is the special case of the proposition in which A is the set $\{y\}$. □

Corollary 2.18. Let (X, d) be a metric space, and let x be in X. Then the closed ball $\{y \in X \mid d(x, y) \le r\}$ is a closed set.

REMARK. Nevertheless, the closed ball is not necessarily the closure of the open ball $B(r; x) = \{y \in X \mid d(x, y) < r\}$. A counterexample is provided by any open ball of radius 1 in a space with the discrete metric.

PROOF. If $f(y) = d(x, y)$, the set in question is $f^{-1}([0, r])$. Corollary 2.17 says that f is continuous, and the equivalence of (a) and (c) in Proposition 2.15 shows that the set in question is closed. $\square$

Proposition 2.19. If A is a nonempty subset of a metric space (X, d), then $A^{\mathrm{cl}} = \{x \mid D(x, A) = 0\}$.

PROOF. The set $\{x \mid D(x, A) = 0\}$ is closed by Propositions 2.16 and 2.15, and it contains A. By Proposition 2.10 it contains A^{cl}. For the reverse inclusion, suppose x is not in A^{cl}, hence that x is not in A and x is not a limit point of A. These conditions imply that there is some $\epsilon > 0$ such that $B(\epsilon; x)$ is disjoint from A, hence that $d(x, y) \ge \epsilon$ for all y in A. Taking the infimum over y gives $D(x, A) \ge \epsilon > 0$. Hence $D(x, A) \ne 0$. $\square$

4. Sequences and Convergence

For a set S, we have already defined in Section I.1 the notion of a **sequence** in S as a function from a certain kind of subset of integers into S. In this section we work with sequences in metric spaces.

A sequence $\{x_n\}$ in a metric space (X, d) is **eventually in** a subset A of X if there is an integer N such that x_n is in A whenever $n \ge N$. The sequence $\{x_n\}$ **converges** to a point x in X if the sequence is eventually in each neighborhood of x. It is apparent that if $\{x_n\}$ converges to x, then so does every subsequence $\{x_{n_k}\}$.

Proposition 2.20. If (X, d) is a metric space, then no sequence in X can converge to more than one point.

PROOF. Suppose on the contrary that $\{x_n\}$ converges to distinct points x and y. The number $m = d(x, y)$ is then > 0. By the assumed convergence, x_n lies in both open balls $B(\frac{m}{2}; x)$ and $B(\frac{m}{2}; y)$ if n is large enough. Thus x_n lies in the intersection of these balls. But this intersection is empty, since the presence of a point z in both balls would mean that $d(x, y) \le d(x, z) + d(z, y) < \frac{m}{2} + \frac{m}{2} = m$, contradiction. $\square$

If a sequence $\{x_n\}$ in a metric space (X, d) converges to x, we shall call x the **limit** of the sequence and write $\lim_{n\to\infty} x_n = x$ or $\lim_n x_n = x$ or $\lim x_n = x$ or $x_n \to x$. A sequence has at most one limit, by Proposition 2.20. If the definition of convergence is extended to pseudometric spaces, then sequences need not have unique limits.

Let us identify convergent sequences in some of the examples of metric spaces in Section 1.

EXAMPLES OF CONVERGENCE IN METRIC SPACES.

(0) The real line. On $\mathbb{R}$ with the usual metric, the convergent sequences are the sequences convergent in the usual sense of Section I.1.

(1) Euclidean space $\mathbb{R}^n$. Here the metric is given by

$$d(x, y) = \Big(\sum_{k=1}^{n} (x_k - y_k)^2\Big)^{1/2}$$

if $x = (x_1, \dots, x_n)$ and $y = (y_1, \dots, y_n)$. Another metric $d'(x, y)$ is given by

$$d'(x, y) = \max_{1\le k\le n} |x_k - y_k|,$$

and we readily check that

$$d'(x, y) \le d(x, y) \le \sqrt{n}\, d'(x, y).$$

From this inequality it follows that the convergent sequences in $(\mathbb{R}^n, d)$ are the same as the convergent sequences in $(\mathbb{R}^n, d')$. On the other hand, the definition of d' as a maximum means that we have convergence in $(\mathbb{R}^n, d')$ if and only if we have ordinary convergence in each entry. Thus convergence of a sequence of vectors in $(\mathbb{R}^n, d)$ means convergence in the k^{th} entry for all k with $1 \le k \le n$.

(2) Complex Euclidean space $\mathbb{C}^n$. As a metric space, $\mathbb{C}^n$ gets identified with $\mathbb{R}^{2n}$. Thus a sequence of vectors in $\mathbb{C}^n$ converges if and only if it converges entry by entry.

(3) Extended real line $\mathbb{R}^*$. Here the metric is given by $d(x, y) = |f(x) - f(y)|$ with $f(x) = x/(1 + |x|)$ if x is in $\mathbb{R}$, $f(-\infty) = -1$, and $f(+\infty) = +1$. We saw in Section 1 that the intersections with $\mathbb{R}$ of the open balls of $\mathbb{R}^*$ are the open intervals in $\mathbb{R}$. Thus convergence of a sequence in $\mathbb{R}^*$ to a point x in $\mathbb{R}$ means that the sequence is eventually in $(-\infty, +\infty)$ and thereafter is an ordinary convergent sequence in $\mathbb{R}$. Convergence to $+\infty$ of a sequence $\{x_n\}$ means that for each real number M, there is an integer N such that $x_n \ge M$ whenever $n \ge N$. Convergence to $-\infty$ is analogous.

(4) Bounded scalar-valued functions on S in the uniform metric. A sequence $\{f_n\}$ in $B(S)$ converges in the uniform metric on $B(S)$ if and only if $\{f_n\}$ converges uniformly, in the sense below, to some member f of $B(S)$. The definition of **uniform convergence** here is the natural generalization of the one in Section I.3: $\{f_n\}$ converges to f uniformly if for each $\epsilon > 0$, there is an integer N such that $n \geq N$ implies $|f_n(s) - f(s)| < \epsilon$ for all s simultaneously. An important fact in this case is that the sequence $\{f_n\}$ is **uniformly bounded**, i.e., that there exists a real number M such that $|f_n(s)| \leq M$ for all n and s. In fact, choose some integer N for $\epsilon = 1$. Then the triangle inequality gives

$$|f_n(s)| \leq |f_n(s) - f(s)| + |f(s) - f_N(s)| + |f_N(s)| \leq 2 + |f_N(s)|$$

for all s if $n \geq N$, so that M can be taken to be $\max_{1 \leq n \leq N} \big\{ \sup_{s \in S} |f_n(s)| \big\} + 2$.

(5) Bounded functions from S into a metric space (R, ρ). Convergence here is the expected generalization of **uniform convergence**: $\{f_n\}$ converges to f uniformly if for each $\epsilon > 0$, there is an integer N such that $n \geq N$ implies $\rho(f_n(s), f(s)) < \epsilon$ for all s simultaneously. As in Example 4, a uniformly convergent sequence of bounded functions is **uniformly bounded** in the sense that $\rho(f_n(s), r_0) \leq M$ for all n and s, M being some real number. Here r_0 is any fixed member of R.

(7) Indiscrete space X. The function $d(x, y)$ in this case is a pseudometric, not a metric, unless X has only one point. Every sequence in X converges to every point in X.

(8) Discrete metric. Convergence of a sequence $\{x_n\}$ in a space X with the discrete metric means that $\{x_n\}$ is eventually constant.

(11) Hilbert cube. For each n, let $(\{x_m\}_{m=1}^{\infty})_n$ be a member of the Hilbert cube, and write x_{mn} for the m^{th} term of the n^{th} sequence. As n varies, the sequence of sequences converges if and only if $\lim_n x_{mn}$ exists for each m.

(12) L^1 metric on Riemann integrable functions. The function $d(f, g)$ defined in this case is a pseudometric, not a metric. Convergence in the corresponding metric space as in Proposition 2.12 therefore really means a certain kind of convergence of equivalence classes: If $\{f_n\}$ and f are given, the sequence of classes $\{[f_n]\}$ converges to the class $[f]$ if and only if $\lim_n \int_a^b |f_n(x) - f(x)|\, dx = 0$. The use of classes in the notation is rather cumbersome and not very helpful, and consequently it is common practice to treat the L^1 space as a metric space and to work with its members as if they were functions rather than equivalence classes. We return to this point in Chapter V.

Let us elaborate a little on Examples 4 and 5, concerning the space $B(S)$ of bounded scalar-valued functions on a set S or, more generally, the space of bounded functions from S into a metric space (R, ρ). Suppose that S has the additional structure of a metric space (S, d). We let $C(S)$ be the subset of $B(S)$ consisting of bounded continuous functions on S, and we write $C(S, \mathbb{R})$ or $C(S, \mathbb{C})$ if we want to be explicit about the range. More generally we consider the space of bounded continuous functions from S into the metric space R. All of these are metric spaces in their own right.

Proposition 2.21. Let (S, d) and (R, ρ) be metric spaces, let x_0 be in S, and let $f_n : S \to R$ be a sequence of bounded functions from S into R that converge uniformly to $f : S \to R$ and are continuous at x_0. Then f is continuous at x_0. In particular, the uniform limit of continuous functions is continuous.

PROOF. For x in S, we write

$$\rho(f(x), f(x_0)) \le \rho(f(x), f_n(x)) + \rho(f_n(x), f_n(x_0)) + \rho(f_n(x_0), f(x_0)).$$

Given $\epsilon > 0$, we choose an integer N by the uniform convergence such that the first and third terms on the right side are $< \epsilon$ for $n \ge N$. With N fixed, we choose $\delta > 0$ by the continuity of f_N at x_0 such that $\rho(f_N(x), f_N(x_0)) < \epsilon$ whenever $d(x, x_0) < \delta$. Then the displayed inequality shows that $d(x, x_0) < \delta$ implies $\rho(f(x), f(x_0)) < 3\epsilon$, and the proposition follows. □

We conclude this section with some elementary results involving convergence of sequences in metric spaces.

Proposition 2.22. If (X, d) is a metric space, then

(a) for any subset A of X and limit point x of A, there exists a sequence in $A - \{x\}$ converging to x,

(b) any convergent sequence in X with limit $x \in X$ either has infinite image, with x as a limit point of the image, or else is eventually constantly equal to x.

REMARK. This result and the first corollary below are used frequently—and often without specific reference.

PROOF OF (a). For each $n \ge 1$, the open ball $B(1/n; x)$ is an open neighborhood of x and must contain a point x_n of A distinct from the limit point x. Then $d(x_n, x) < 1/n$, and thus $\lim x_n = x$. Hence $\{x_n\}$ is the required sequence. □

PROOF OF (b). Suppose that $\{x_n\}$ converges to x and has infinite image. By discarding the terms equal to x, we obtain a subsequence $\{x_{n_k}\}$ with limit x. If U is an open neighborhood of x, then $\{x_{n_k}\}$ is eventually in U, by the assumed convergence. Since no term of the subsequence equals x, U contains a member of the image of $\{x_n\}$ different from x. Thus x is a limit point of the image of $\{x_n\}$.

Now suppose that $\{x_n\}$ converges to x and has finite image $\{p_1, \dots, p_r\}$. If x_n is equal to some particular p_{j_0} for infinitely many n, then $\{x_n\}$ has an infinite subsequence converging to p_{j_0}. Since $\{x_n\}$ converges to x, every convergent subsequence converges to x. Therefore $p_{j_0} = x$. For $j \neq j_0$, only finitely many x_n can then equal p_j, and it follows that $\{x_n\}$ is eventually constantly equal to $p_{j_0} = x$. □

Corollary 2.23. If (X, d) is a metric space, then a subset F of X is closed if and only if every convergent sequence in F has its limit in F.

PROOF. Suppose that F is closed and $\{x_n\}$ is a convergent sequence in F with limit x. By Proposition 2.22b, either x is in the image of the sequence or x is a limit point of the sequence. In either case, x is in F; thus the limit of any convergent sequence in F is in F.

Conversely suppose every convergent sequence in F has its limit in F. If x is a limit point of F, then Proposition 2.22a produces a sequence in $F - \{x\}$ converging to x. By assumption, the limit x is in F. Therefore F contains all its limit points and is closed. □

Corollary 2.24. If (S, d) is a metric space, then the set $C(S)$ of bounded continuous scalar-valued functions on S is a closed subset of the metric space $B(S)$ of all bounded scalar-valued functions on S.

PROOF. Proposition 2.21 shows for any sequence in $C(S)$ convergent in $B(S)$ that the limit is actually in $C(S)$. By Corollary 2.23, $C(S)$ is closed in $B(S)$. □

Proposition 2.25. Let $f : X \to Y$ be a function between metric spaces. Then f is continuous at a point x in X if and only if whenever $\{x_n\}$ is a convergent sequence in X with limit x, then $\{f(x_n)\}$ is convergent in Y with limit $f(x)$.

REMARK. In the special case of domain and range $\mathbb{R}$, this result was mentioned in Section I.1 after the definition of continuity. We deferred the proof of the special case until now to avoid repetition.

PROOF. Suppose that f is continuous at x and that $\{x_n\}$ is a convergent sequence in X with limit x. Let V be any open neighborhood of $f(x)$. By continuity, there exists an open neighborhood U of x such that $f(U) \subseteq V$. Since $x_n \to x$, there exists N such that x_n is in U whenever $n \geq N$. Then $f(x_n)$ is in $f(U) \subseteq V$ whenever $n \geq N$. Hence $\{f(x_n)\}$ converges to $f(x)$.

Conversely suppose that $x_n \to x$ always implies $f(x_n) \to f(x)$. We are to show that f is continuous. Let V be an open neighborhood of $f(x)$. We are to show that some open neighborhood of x maps into V under f. Assuming the contrary, we can find, for each $n \geq 1$, some x_n in $B(1/n; x)$ such that $f(x_n)$ is not in V. Then $x_n \to x$, but the distance of $f(x_n)$ from $f(x)$ is bounded away from 0. Thus $f(x_n)$ cannot converge to $f(x)$. This is a contradiction, and we conclude that some $B(1/n; x)$ maps into V under f; since V is arbitrary, f is continuous. $\square$

5. Subspaces and Products

When working with functions on the real line, one frequently has to address situations in which the domain of the function is just an open interval or a closed interval, rather than the whole line. When one uses the ϵ-δ definition of continuity, the subject does not become much more cumbersome, but it can become more cumbersome if one uses some other definition, such as one involving limits. The theory of metric spaces has a device for addressing smaller domains than the whole space—the notion of a subspace—and then the theory of functions on a subspace stands on an equal footing with the theory of functions on the whole space.

Let (X, d) be a metric space, and let A be a nonempty subset of X. There is a natural way of making A into a metric space, namely by taking the restriction $d\big|_{A\times A}$ as a metric for A. When we do so, we speak of A as a **subspace** of X. When there is a need to be more specific, we may say that A is a **metric subspace** of X. If A is an open subset of X, we may say that A is an **open subspace**; if A is a closed subset of X, we may say that A is a **closed subspace**.

Proposition 2.26. If A is a subspace of a metric space (X, d), then the open sets of A are exactly all sets $U \cap A$, where U is open in X, and the closed sets of A are all sets $F \cap A$, where F is closed in X.

PROOF. The open balls in A are the intersections with A of the open balls of X, and the statement about open sets follows by taking unions. The closed sets of A are the complements within A of all the open sets of A, thus all sets of the form $A - (U \cap A)$ with U open in X. Since $A - (U \cap A) = A \cap U^c$, the statement about closed sets follows. $\square$

Corollary 2.27. If A is a subspace of (X, d) and if $f : X \to Y$ is continuous at a point a of A, then the restriction $f\big|_A$, mapping A into Y, is continuous at a. Also, f is continuous at a if and only if the function $f_0 : X \to f(X)$ obtained by redefining the range to be the image is continuous at a.

PROOF. Let V be an open neighborhood of $f(a)$ in Y. By continuity of f at a as a function on X, choose an open neighborhood U of a in X with $f(U) \subseteq V$. Then $U \cap A$ is an open neighborhood of a in A, and $f(U \cap A) \subseteq V$. Hence $f\big|_A$ is continuous at a.

The most general open neighborhood of $f(a)$ in $f(X)$ is of the form $V \cap f(X)$ with V an open neighborhood of $f(a)$ in Y. Since $f^{-1}(V) = f_0^{-1}(V \cap f(X))$, the condition for continuity of f_0 at a is the same as the condition for continuity of f at a. $\square$

We now turn our attention to product spaces. Product spaces are a convenient device for considering functions of several variables.

If (X, d) and (Y, d') are metric spaces, there are several natural ways of making the product set $X \times Y$, the set of ordered pairs with the first member from X and the second from Y, into a metric space, but all such ways lead to the same class of open sets and therefore also the same class of convergent sequences. We discussed an instance of this phenomenon in Example 1 of Section 4. For general X and Y, three such metrics on $X \times Y$ are

$$\begin{aligned}
\rho_1\big((x_1, y_1), (x_2, y_2)\big) &= d(x_1, x_2) + d'(y_1, y_2), \\
\rho_2\big((x_1, y_1), (x_2, y_2)\big) &= \big(d(x_1, x_2)^2 + d'(y_1, y_2)^2\big)^{1/2}, \\
\rho_\infty\big((x_1, y_1), (x_2, y_2)\big) &= \max\{d(x_1, x_2),\ d'(y_1, y_2)\}.
\end{aligned}$$

Each satisfies the defining properties of a metric. Simple algebra gives

$$\max\{a, b\} \le (a^2 + b^2)^{1/2} \le a + b \le 2\max\{a, b\}$$

whenever a and b are nonnegative reals, and therefore

$$\rho_\infty \le \rho_2 \le \rho_1 \le 2\rho_\infty.$$

Let us check that this chain of inequalities implies that the neighborhoods of a point (x_0, y_0) are the same in all three metrics, hence that the open sets are the same in all three metrics. For any $r > 0$, the open balls about (x_0, y_0) in the three metrics satisfy

$$B_1(r; (x_0, y_0)) \subseteq B_2(r; (x_0, y_0)) \subseteq B_\infty(r; (x_0, y_0)) \subseteq B_1(2r; (x_0, y_0)).$$

The first and second inclusions show that open balls about (x_0, y_0) in the metrics ρ_2 and ρ_∞ are neighborhoods of (x_0, y_0) in the metric ρ_1. Similarly the second and third inclusions show that open balls in the metrics ρ_∞ and ρ_1 are neighborhoods in the metric ρ_2, and the third and first inclusions show that open balls in the metrics ρ_1 and ρ_2 are neighborhoods in the metric ρ_∞.

We shall refer to the metric ρ_∞ as the **product metric** for $X \times Y$. If $X \times Y$ is being regarded as a metric space and no metric has been mentioned, ρ_∞ is to be understood. But it is worth keeping in mind that ρ_1 and ρ_2 yield the same open sets. In the case of Euclidean space, it is the metric ρ_2 on $\mathbb{R}^m \times \mathbb{R}^n$ that gives the

Euclidean metric on $\mathbb{R}^{m+n}$; thus the product metric and the Euclidean metric are distinct but yield the same open sets.

A sequence $\{(x_n, y_n)\}$ in the product metric converges to (x_0, y_0) in $X \times Y$ if and only if $\{x_n\}$ converges to x_0 and $\{y_n\}$ converges to y_0. Since the three metrics on $X \times Y$ yield the same convergent sequences, this statement is valid in the metrics ρ_1 and ρ_2 as well.

It is an elementary property of the arithmetic operations in $\mathbb{R}$ that if $\{x_n\}$ converges to x_0 and $\{y_n\}$ converges to y_0, then $\{x_n + y_n\}$ converges to $x_0 + y_0$. Similar statements apply to subtraction, multiplication, maximum, and minimum, and then to absolute value and to division except where division by 0 is involved. Further similar statements apply to those operations on vectors that make sense. Applying Proposition 2.25, we obtain (a) through (e) in the following proposition. Conclusions (a′) through (e′) are proved similarly.

Proposition 2.28. The following operations are continuous:

(a) addition and subtraction from $\mathbb{R}^n \times \mathbb{R}^n$ into $\mathbb{R}^n$,
(b) scalar multiplication from $\mathbb{R} \times \mathbb{R}^n$ into $\mathbb{R}^n$,
(c) the map $x \mapsto x^{-1}$ from $\mathbb{R} - \{0\}$ to $\mathbb{R} - \{0\}$,
(d) the map $x \mapsto |x|$ from $\mathbb{R}^n$ to $\mathbb{R}$,
(e) the operations from $\mathbb{R}^2$ to $\mathbb{R}$ of taking the maximum of two real numbers and taking the minimum of two real numbers,
(a′) addition and subtraction from $\mathbb{C}^n \times \mathbb{C}^n$ into $\mathbb{C}^n$,
(b′) scalar multiplication from $\mathbb{C} \times \mathbb{C}^n$ into $\mathbb{C}^n$,
(c′) the map $x \mapsto x^{-1}$ from $\mathbb{C} - \{0\}$ to $\mathbb{C} - \{0\}$,
(d′) the map $x \mapsto |x|$ from $\mathbb{C}^n$ to $\mathbb{R}$,
(e′) the map $x \mapsto \bar{x}$ from $\mathbb{C}$ to $\mathbb{C}$.

Corollary 2.29. Let (X, d) be a metric space, and let f and g be continuous functions from X into $\mathbb{R}^n$ or $\mathbb{C}^n$. If c is a scalar, then $f+g$, cf, $f-g$, and $|f|$ are continuous. If $n = 1$, then the product fg is continuous, and the function $1/f$ is continuous on the set where f is not zero. If $n = 1$ and the functions take values in $\mathbb{R}$, then $\max\{f, g\}$ and $\min\{f, g\}$ are continuous. If $n = 1$ and the functions take values in $\mathbb{C}$, then the complex conjugate $\bar{f}$ is continuous.

REMARKS. If (S, d) is a metric space, then it follows that the metric space $C(S)$ of bounded continuous scalar-valued functions on S is a vector space. As such, it is a vector subspace of the metric space $B(S)$ of bounded scalar-valued functions on S, and it is a metric subspace as well.[3]

[3]The word "subspace" can now be used in two senses, that of a metric subspace of a metric space and that of a vector subspace of a vector space. The latter kind of subspace we shall always refer to as a "vector subspace," retaining the word "vector" for clarity. A "closed vector subspace" of $B(S)$ then has to mean a closed metric subspace that is also a vector subspace.

PROOF. The argument for $f+g$ and for functions with values in $\mathbb{R}^n$ will illustrate matters sufficiently. We set up $x \mapsto f(x)+g(x)$ as a suitable composition, expressing the composition in a diagram:

$$X \xrightarrow{x \mapsto (x,x)} X \times X \xrightarrow{(x,y) \mapsto (f(x),g(y))} \mathbb{R}^n \times \mathbb{R}^n \xrightarrow{(u,v) \mapsto u+v} \mathbb{R}^n.$$

Each function in the diagram is continuous, the last of them by Proposition 2.28a, and then the composition is continuous by Corollary 2.14. $\square$

We conclude this section with one further remark. When (X,d) is a metric space, we saw in Corollary 2.17 that $x \mapsto d(x,y)$ and $y \mapsto d(x,y)$ are continuous functions from X to $\mathbb{R}$. Actually, $(x,y) \mapsto d(x,y)$ is a continuous function from $X \times X$ into $\mathbb{R}$ if we use the product metric. In fact, if ρ_∞ denotes the product metric with $\rho_\infty\big((x,y),(x_0,y_0)\big) = \max\{d(x,x_0), d(y,y_0)\}$, then we have $d(x,y) \leq d(x,x_0) + d(x_0,y_0) + d(y_0,y)$ and therefore

$$d(x,y) - d(x_0,y_0) \leq d(x,x_0) + d(y,y_0).$$

Reversing the roles of (x,y) and (x_0,y_0), we see that

$$\begin{aligned} |d(x,y) - d(x_0,y_0)| &\leq d(x,x_0) + d(y,y_0) \\ &\leq 2\max\{d(x,x_0), d(y,y_0)\} \\ &= 2\rho_\infty\big((x,y),(x_0,y_0)\big). \end{aligned}$$

From this chain of inequalities, it follows that d is continuous with $\delta = \epsilon/2$.

6. Properties of Metric Spaces

This section contains two results about metric spaces. One lists a number of "separation properties" of sets within any metric space. The other concerns the completely different property of "separability," which is satisfied by some metric spaces and not by others, and it says that separability may be defined in any of three equivalent ways.

Proposition 2.30 (separation properties). Let (X,d) be a metric space. Then

(a) every one-point subset of X is a closed set, i.e., X is $\mathbf{T}_1$,

(b) for any two distinct points x and y of X, there are disjoint open sets U and V with $x \in U$ and $y \in V$, i.e., X is **Hausdorff**,

(c) for any point $x \in X$ and any closed set $F \subseteq X$ with $x \notin F$, there are disjoint open sets U and V with $x \in U$ and $F \subseteq V$, i.e., X is **regular**,

(d) for any two disjoint closed subsets E and F of X, there are disjoint open sets U and V such that $E \subseteq U$ and $F \subseteq V$, i.e., X is **normal**,

(e) for any two disjoint closed subsets E and F of X, there is a continuous function $f : X \to [0,1]$ such that f is 0 exactly on E and f is 1 exactly on F.

PROOF. For (a), the set $\{x\}$ is the intersection of all closed balls $B(r; x)$ for $r > 0$ and hence is closed by Corollary 2.18 and Proposition 2.8b. For (e), the function $f(x) = D(x; E)/(D(x; E) + D(x; F))$ is continuous by Proposition 2.16 and Corollary 2.29 and takes on the values 0 and 1 exactly on E and F, respectively, by Proposition 2.19.

For (d), we need only apply (e) and Proposition 2.15b with $U = f^{-1}\big((-\infty, \tfrac{1}{2})\big)$ and $V = f^{-1}\big((\tfrac{1}{2}, +\infty)\big)$. Conclusions (a) and (d) imply (c), and conclusions (a) and (c) imply (b). This completes the proof. □

A **base** $\mathcal{B}$ for a metric space (X, d) is a family of open sets such that every open set is a union of members of $\mathcal{B}$. The family of all open balls is an example of a base.

Proposition 2.31. If (X, d) is a metric space, then a family $\mathcal{B}$ of subsets of X is a base for (X, d) if and only if

(a) every member of $\mathcal{B}$ is open and
(b) for each $x \in X$ and open neighborhood U of x, there is some member B of $\mathcal{B}$ such that x is in B and B is contained in U.

PROOF. If $\mathcal{B}$ is a base, then (a) holds by definition of base. If U is open in X, then $U = \bigcup_\alpha B_\alpha$ for some members B_α of $\mathcal{B}$, and any such B_α containing x can be taken as the set B in (b).

Conversely suppose that $\mathcal{B}$ satisfies (a) and (b). By (a), each member of $\mathcal{B}$ is open in X. If U is open in X, we are to show that U is a union of members of $\mathcal{B}$. For each $x \in U$, choose some set $B = B_x$ as in (b). Then $U = \bigcup_{x \in U} B_x$, and hence each open set in X is a union of members of $\mathcal{B}$. Thus $\mathcal{B}$ is a base. □

This book uses the word **countable** to mean finite or countably infinite. It is then meaningful to ask whether a particular metric space (X, d) has a countable base. On the real line $\mathbb{R}$, the open intervals with rational endpoints form a countable base.

A subset D of X is **dense in** a subset A of X if $D^{\text{cl}} \supseteq A$; D is **dense**, or **everywhere dense**, if D is dense in X. A set D is dense if and only if there is some point of D in each nonempty open set of X.

A family $\mathcal{U}$ of open sets is an **open cover** of X if the union of the sets in $\mathcal{U}$ is X. An **open subcover** of $\mathcal{U}$ is a subfamily of $\mathcal{U}$ that is itself an open cover.

Proposition 2.32. The following three conditions are equivalent for a metric space (X, d):

(a) X has a countable base,
(b) every open cover of X has a countable open subcover,
(c) X has a countable dense subset.

PROOF. If (a) holds, let $\mathcal{B} = \{B_n\}_{n\geq 1}$ be a countable base, and let $\mathcal{U}$ be an open cover of X. Any $U \in \mathcal{U}$ is the union of the $B_n \in \mathcal{B}$ with $B_n \subseteq U$. If $\mathcal{B}_0 = \{B_n \in \mathcal{B} \mid B_n \subseteq U \text{ for some } U \in \mathcal{U}\}$, then it follows that $\bigcup_{B_n \in \mathcal{B}_0} = \bigcup_{U \in \mathcal{U}} = X$. For each B_n in $\mathcal{B}_0$, select some U_n in $\mathcal{U}$ with $B_n \subseteq U_n$. Then $\bigcup_n U_n \supseteq \bigcup_{B_n \in \mathcal{B}_0} = X$, and $\{U_n\}$ is a countable open subcover of $\mathcal{U}$. Thus (b) holds.

If (b) holds, form, for each fixed $n \geq 1$, the open cover of X consisting of all open balls $B(1/n; x)$. For that n, let $\{B(1/n; x_{mn})\}_{m\geq 1}$ be a countable open subcover. We shall prove that the set D of all x_{mn}, with m and n arbitrary, is dense in X. It is enough to prove that each nonempty open set in X contains a member of D, hence to prove, for each n, that each open ball of radius $1/n$ contains a member of D. Thus consider $B(1/n; x)$. Since the open balls $B(1/n; x_{mn})$ with $m \geq 1$ cover X, x is in some $B(1/n; x_{mn})$. Then that x_{mn} has $d(x_{mn}, x) < 1/n$, and hence x_{mn} is in $B(1/n; x)$. Thus D is dense, and (c) holds.

If (c) holds, let $\{x_n\}_{n\geq 1}$ be a countable dense set. Form the collection of all open balls centered at some x_n and having rational radius. Let us use Proposition 2.31 to see that this collection of open sets, which is certainly countable, is a base. Let U be an open neighborhood of x. We are to see that there is some member B of our collection such that x is in B and B is contained in U. Since U is a neighborhood of x, we can find an open ball $B(r; x)$ such that $B(r; x) \subseteq U$; we may assume that r is rational. The given set $\{x_n\}_{n\geq 1}$ being dense, some x_n lies in $B(r/2; x)$. If y is in $B(r/2; x_n)$, then $d(x, y) \leq d(x, x_n) + d(x_n, y) < \frac{r}{2} + \frac{r}{2} = r$. Hence x lies in $B(r/2; x_n)$ and $B(r/2; x_n) \subseteq B(r; x) \subseteq U$. Since $r/2$ is rational, the open ball $B(r/2; x_n)$ is in our countable collection, and our countable collection is a base. This proves (a). $\square$

A metric space satisfying the equivalent conditions of Proposition 2.32 is said to be **separable**. Among the examples of metric spaces in Section 1, the ones in Examples 1, 2, 3, 6, 8 if X is countable, 9, 11, 12, 13, and 14 are separable. A countable dense set in Examples 1, 2, and 3 is given by all points with all coordinates rational. In Example 6, one countable dense set consists of all sequences with only finitely many nonzero entries, those being rational, and in Examples 8 and 9, X itself is a countable dense set. In Example 11, the sequences that are 0 in all but finitely many entries, those being rational, form a countable dense set. In Example 13, the set of finite linear combinations of exponentials e^{inx} using scalars in $\mathbb{Q} + i\mathbb{Q}$ is dense as a consequence of Parseval's equality. In Example 12, when $[a, b] = [-\pi, \pi]$, the same countable set as for Example 13 is dense by Proposition 2.25 because the sets of functions in Examples 12 and 13 coincide and the inclusion of $\mathcal{R}[-\pi, \pi]$ relative to L^2 into $\mathcal{R}[-\pi, \pi]$ relative to L^1 is continuous. In Example 14, the set of polynomials with coefficients in $\mathbb{Q}+i\mathbb{Q}$ is countable and can be shown to be dense.

Example 10 is not separable, and Example 8 is not separable if X is uncountable.

7. Compactness and Completeness

In Section 6 we introduced the notions of open cover and subcover for a metric space. We call a metric space **compact** if every open cover of the space has a *finite* subcover. A subset E of a metric space (X, d) is **compact** if it is compact as a subspace of the whole space, i.e., if every collection of open sets in X whose union contains E has a finite subcollection whose union contains E.

Historically this notion was embodied in the Heine–Borel Theorem, which says that any closed bounded subset of Euclidean space has the property that has just been defined to be compactness. As we shall see in Theorem 2.36 and Corollary 2.37 below, the Heine–Borel Theorem can be proved from the Bolzano–Weierstrass Theorem (Theorem 1.8) and leads to faster, more transparent proofs of some of the consequences of the Bolzano–Weierstrass Theorem. Even more important is that it generalizes beyond metric spaces and produces useful conclusions about certain spaces of functions when statements about pointwise convergence of a sequence of functions are inadequate.

Easily established examples of compact sets are hard to come by. For one example, consider in a metric space (X, d) a convergent sequence $\{x_n\}$ along with its limit x. The subset $E = \{x\} \cup \bigcup_n \{x_n\}$ of X is compact. In fact, if $\mathcal{U}$ is an open cover of E, some member U of $\mathcal{U}$ has x as an element, and then all but finitely many elements of the sequence must be in U as well. Say that U contains x and all x_n with $n \geq N$. For $1 \leq n < N$, let U_n be a member of $\mathcal{U}$ containing x_n. Then $\{U, U_1, \ldots, U_{N-1}\}$ is a finite subcover of $\mathcal{U}$.

It is easier to exhibit noncompact sets. The open interval $(0, 1)$ is not compact, as is seen from the open cover $\{(\frac{1}{n}, 1)\}$. Nor is an infinite discrete space, since one-point sets form an open cover. A subtle dramatic example is the closed unit ball C of the hedgehog space X, Example 10 in Section 1; this set is not compact. In fact, the open ball of radius $1/2$ about the origin is an open set in X, and so is each open ray from the origin out to infinity. Let $\mathcal{U}$ be this collection of open sets. Then $\mathcal{U}$ is an open cover of C. However, no member of $\mathcal{U}$ is superfluous, since for each U in $\mathcal{U}$, there is some point x in C such that x is in C but x is in no other member of $\mathcal{U}$. Thus $\mathcal{U}$ does not contain even a countable subcover.

Let us now work directly toward a proof of the equivalence of compactness and the Bolzano–Weierstrass property in a metric space.

Proposition 2.33. A compact metric space is separable.

PROOF. This is immediate from equivalent condition (b) for the definition of separability in Proposition 2.32. □

Proposition 2.34. In any metric space (X, d),

(a) every compact subset is closed and bounded and
(b) any closed subset of a compact set is compact.

PROOF. For (a), let E be a compact subset of X, fix x_0 in X, and let U_n for $n \geq 1$ be the open ball $\{x \in X \mid d(x_0, x) < n\}$. Then $\{U_n\}$ is an open cover of E. Since the U_n's are nested, the compactness of E implies that E is contained in a single U_N for some N. Then every member of E is at distance at most N from x_0, and E is bounded.

To see that E is closed, we argue by contradiction. Let x_0' be a limit point of E that is not in E. By the Hausdorff property (Proposition 2.30b), we can find, for each $x \in E$, open sets U_x and V_x with $x \in U_x$, $x_0' \in V_x$, and $U_x \cap V_x = \varnothing$. The sets U_x form an open cover of E. By compactness let $\{U_{x_1}, \dots, U_{x_n}\}$ be a finite subcover. Then $E \subseteq U_{x_1} \cup \cdots \cup U_{x_n}$, which is disjoint from the neighborhood $V_{x_1} \cap \cdots \cap V_{x_n}$ of x_0'. Thus x_0' cannot be a limit point of E, and we have arrived at a contradiction. This proves (a).

For (b), let E be compact, and let F be a closed subset of E. Because of (a), F is a closed subset of X. Let $\mathcal{U}$ be an open cover of F. Then $\mathcal{U} \cup \{F^c\}$ is an open cover of E. Passing to a finite subcover and discarding F^c, we obtain a finite subcover of F. Thus F is compact. $\square$

A collection of subsets of a nonempty set is said to have the **finite-intersection property** if each intersection of finitely many of the subsets is nonempty.

Proposition 2.35. A metric space (X, d) is compact if and only if each collection of closed subsets of X with the finite-intersection property has nonempty intersection.

PROOF. Closed sets with the finite-intersection property have complements that are open sets, no finite subcollection of which is an open cover. $\square$

Theorem 2.36. A metric space (X, d) is compact if and only if every sequence has a convergent subsequence.

PROOF. Suppose that X is compact. Arguing by contradiction, suppose that $\{x_n\}_{n\geq 1}$ is a sequence in X with no convergent subsequence. Put $F = \bigcup_{n=1}^{\infty}\{x_n\}$. The subset F of X is closed by Corollary 2.23, hence compact by Proposition 2.34b. Since no x_n is a limit point of F, there exists an open set U_n in X containing x_n but no other member of F. Then $\{U_n\}_{n\geq 1}$ is an open cover of F with no finite subcover, and we have arrived at a contradiction.

Conversely suppose that every sequence has a convergent subsequence. We first show that X is separable. Fix an integer n. There cannot be infinitely many disjoint open balls of radius $1/n$, since otherwise we could find a sequence from among their centers with no convergent subsequence. Thus we can choose a finite disjoint collection of these open balls that is not contained in a larger such finite collection. Let their centers be $x_1, \dots, x_N$. The claim is that every point of X is at distance $< 2/n$ from one of these finitely many centers. In fact, if $x \in X$ is given, form $B(\frac{1}{n}; x)$. This must meet some $B(\frac{1}{n}; x_i)$ at a point y, and then

$$d(x, x_i) \leq d(x, y) + d(y, x_i) < \tfrac{1}{n} + \tfrac{1}{n} = \tfrac{2}{n}.$$

Thus x is at distance $< 2/n$ from one of the finitely many centers, as asserted. Now let n vary, and let D be the set of all these centers for all n. Then every point of X has members of D arbitrarily close to it, and hence D is a countable dense set in X. Thus X is separable.

Let $\mathcal{U}$ be an open cover of X having no finite subcover. By the separability and condition (b) in Proposition 2.32, we may assume that $\mathcal{U}$ is countable, say $\mathcal{U} = \{U_1, U_2, \dots\}$. Since $U_1 \cup U_2 \cup \cdots \cup U_n$ is not a cover, there exists a point x_n not in the union of the first n sets. By hypothesis the sequence $\{x_n\}$ has a convergent subsequence $\{x_{n_k}\}$, say with limit x. Since $\mathcal{U}$ is a cover, some member U_N of $\mathcal{U}$ contains x. Then $\{x_{n_k}\}$ is eventually in U_N, and some n_k with $n_k > N$ has x_{n_k} in U_N. But x_{n_k} is not in $U_1 \cup \cdots \cup U_{n_k}$ by construction, and this union contains U_N, since $n_k > N$. We have arrived at a contradiction, and we conclude that $\mathcal{U}$ must have had a finite subcover. $\square$

Corollary 2.37 (Heine–Borel Theorem) In Euclidean space $\mathbb{R}^n$, every closed bounded set is compact.

REMARK. Conversely we saw in Proposition 2.34a that every compact subset of any metric space is closed and bounded.

PROOF. Let C be a closed rectangular solid in $\mathbb{R}^n$, and let $x^{(k)} = (x_1^{(k)}, \dots, x_n^{(k)})$ be the members of a sequence in C. By the Bolzano–Weierstrass Theorem (Theorem 1.8) for $\mathbb{R}^1$, we can find a subsequence convergent in the first coordinate, a subsequence of that convergent in the second coordinate, and so on. Thus $\{x^{(k)}\}$ has a subsequence convergent in $\mathbb{R}^n$. By Corollary 2.23 the limit is in C. By Theorem 2.36, C is compact. Applying Corollary 2.34b, we see that every closed bounded subset of $\mathbb{R}^n$ is compact. $\square$

The next few results will show how the use of compactness both simplifies and generalizes some of the theorems proved in Section I.1.

Proposition 2.38. Let (X, d) and (Y, ρ) be metric spaces with X compact. If $f : X \to Y$ is continuous, then $f(X)$ is a compact subset of Y.

PROOF. If $\{U_\alpha\}$ is an open cover of $f(X)$, then $\{f^{-1}(U_\alpha)\}$ is an open cover of X. Let $\{f^{-1}(U_j)\}_{j=1}^n$ be a finite subcover. Then $\{U_j\}_{j=1}^n$ is a finite subcover of $f(X)$. $\square$

Corollary 2.39. Let (X, d) be a compact metric space, and let $f : X \to \mathbb{R}$ be a continuous function. Then f attains its maximum and minimum values.

REMARK. Theorem 1.11 was the special case of this result with $X = [a, b]$. This particular space X is compact by the Heine–Borel Theorem (Corollary 2.37), and the corollary applies to yield exactly the conclusion of Theorem 1.11.

PROOF. By Proposition 2.38, $f(X)$ is a compact subset of $\mathbb{R}$. By Proposition 2.34a, $f(X)$ is closed and bounded. The supremum and infimum of the members of $f(X)$ in $\mathbb{R}^*$ lie in $\mathbb{R}$, since $f(X)$ is bounded, and they are limits of sequences in $f(X)$. Since $f(X)$ is closed, Proposition 2.23 shows that they must lie in $f(X)$. $\square$

Corollary 2.40. Let (X, d) and (Y, ρ) be metric spaces with X compact. If $f : X \to Y$ is continuous, one-one, and onto, then f is a homeomorphism.

REMARK. In the hypotheses of the change of variables formula for integrals in $\mathbb{R}^1$ (Theorem 1.34), a function $\varphi : [A, B] \to [a, b]$ was given as strictly increasing, continuous, and onto. Another hypothesis of the theorem was that φ^{-1} was continuous. Corollary 2.40 shows that this last hypothesis was redundant.

PROOF. Let E be a closed subset of X, and consider $(f^{-1})^{-1}(E) = f(E)$. The set E is compact by Proposition 2.34b, $f(E)$ is compact by Proposition 2.38, and $f(E)$ is closed by Proposition 2.34a. Proposition 2.15b thus shows that f^{-1} is continuous. $\square$

If (X, d) and (Y, ρ) are metric spaces, a function $f : X \to Y$ is **uniformly continuous** if for each $\epsilon > 0$, there is some $\delta > 0$ such that $d(x_1, x_2) < \delta$ implies $\rho(f(x_1), f(x_2)) < \epsilon$. This is the natural generalization of the definition in Section I.1 for the special case of a real-valued function of a real variable.

Proposition 2.41. Let (X, d) and (Y, ρ) be metric spaces with X compact. If $f : X \to Y$ is continuous, then f is uniformly continuous.

REMARK. This result generalizes Theorem 1.10, which is the special case $X = [a, b]$ and $Y = \mathbb{R}$.

PROOF. Let $\epsilon > 0$ be given. For each $x \in X$, choose $\delta_x > 0$ such that $d(x', x) < \delta_x$ implies $\rho(f(x'), f(x)) < \epsilon/2$. The open balls $B(\frac{1}{2}\delta_x; x)$ cover X; let the balls with centers $x_1, \dots, x_n$ be a finite subcover. Put $\delta = \frac{1}{2}\min\{\delta_{x_1}, \dots, \delta_{x_n}\}$. Now suppose that $d(x', x) < \delta$. The point x is in some ball in the finite subcover; suppose x is in $B(\frac{1}{2}\delta_{x_j}; x_j)$. Then $d(x, x_j) < \frac{1}{2}\delta_{x_j}$, so that

$$d(x', x_j) \le d(x', x) + d(x, x_j) < \delta + \tfrac{1}{2}\delta_{x_j} \le \delta_{x_j}.$$

By definition of δ_{x_j}, $\rho(f(x'), f(x_j)) < \epsilon/2$ and $\rho(f(x_j), f(x)) < \epsilon/2$. Therefore

$$\rho(f(x'), f(x)) \le \rho(f(x'), f(x_j)) + \rho(f(x_j), f(x)) < \tfrac{\epsilon}{2} + \tfrac{\epsilon}{2} = \epsilon,$$

and the proof is complete. $\square$

One final application of compactness is the Fundamental Theorem of Algebra, which is discussed in Section A8 of Appendix A in the context of properties of polynomials.

Theorem 2.42 (Fundamental Theorem of Algebra). Every polynomial with complex coefficients and degree ≥ 1 has a complex root.

PROOF. Let $P : \mathbb{C} \to \mathbb{C}$ be the function $P(z) = \sum_{j=0}^{n} a_j z^j$, where $a_0, \dots, a_n$ are in $\mathbb{C}$ with $a_n \ne 0$ and with $n \ge 1$. We may assume that $a_n = 1$. Let $m = \inf_{z\in\mathbb{C}} |P(z)|$. Since $P(z) = z^n\big(1 + a_{n-1}z^{-1} + \cdots + a_1 z^{-(n-1)} + a_0 z^{-n}\big)$, we have $\lim_{z\to\infty} P(z)/z^n = 1$. Thus there exists an R such that $|P(z)| \ge \frac{1}{2}|z|^n$ whenever $|z| \ge R$. Choosing $R = R_0$ such that $\frac{1}{2}R_0^n \ge 2m$, we see that $|P(z)| \ge 2m$ for $|z| \ge R_0$. Consequently $m = \inf_{|z|\le R_0} |P(z)|$. The set $S = \big\{z \in \mathbb{C} \bigm| |z| \le R_0\big\}$ is compact by the Heine–Borel Theorem (Corollary 2.37), and Corollary 2.39 shows that $|P(z)|$ attains its minimum on S at some point z_0 in S. Then $|P(z)|$ attains its minimum on $\mathbb{C}$ at z_0. We shall show that this minimum value m is 0.

Assuming the contrary, define $Q(z) = P(z + z_0)/P(z_0)$, so that $Q(z)$ is a polynomial of degree $n \ge 1$ with $Q(0) = 1$ and $|Q(z)| \ge 1$ for all z. Write

$$Q(z) = 1 + b_k z^k + b_{k+1} z^{k+1} + \cdots + b_n z^n \qquad \text{with } b_k \ne 0.$$

Corollary 1.45 produces a real number θ such that $e^{ik\theta} b_k = -|b_k|$. For any $r > 0$ with $r^k |b_k| < 1$, we then have

$$\big|1 + b_k r^k e^{ik\theta}\big| = 1 - r^k |b_k|.$$

For such r and that θ, this equality implies that

$$\begin{aligned} |Q(re^{i\theta})| &\le \big|1 + b_k r^k e^{ik\theta}\big| + r^{k+1}|b_{k+1}| + \cdots + r^n |b_n| \\ &\le 1 - r^k\big(|b_k| - r|b_{k+1}| - \cdots - r^{n-k}|b_n|\big). \end{aligned}$$

For sufficiently small $r > 0$, the expression in parentheses on the right side is positive, and then $|Q(re^{i\theta})| < 1$, in contradiction to hypothesis. Thus we must have had $m = 0$, and we obtain $P(z_0) = 0$. $\square$

Another theme discussed in Section I.1 is that Cauchy sequences in $\mathbb{R}^1$ are convergent. This convergence was proved in Theorem 1.9 as a consequence of the Bolzano–Weierstrass Theorem. Actually, many sequences in metric spaces of importance in analysis are shown to converge without one's knowing the limit in advance and without using any compactness, and we therefore isolate the forced convergence of Cauchy sequences as a definition. In a metric space (X, d), a sequence $\{x_n\}$ is a **Cauchy sequence** if for any $\epsilon > 0$, there is some integer N

such that $d(x_m, x_n) < \epsilon$ whenever m and n are $\geq N$. A familiar 2ϵ argument shows that convergent sequences are Cauchy. Other familiar arguments show that any Cauchy sequence with a convergent subsequence is convergent and that any Cauchy sequence is bounded.

We say that the metric space (X, d) is **complete** if every Cauchy sequence in X converges to a point in X. We know that the line $\mathbb{R}^1$ is complete. It follows that $\mathbb{R}^n$ is complete because a Cauchy sequence in $\mathbb{R}^n$ is Cauchy in each coordinate. A nonempty subset E of X is **complete** if E as a subspace is a complete metric space. The next two propositions and corollary give three examples of complete metric spaces.

Proposition 2.43. A subset E of a complete metric space X is complete if and only if it is closed.

REMARK. In particular every closed subset of $\mathbb{R}^n$ is a complete metric space.

PROOF. Suppose E is closed. Let $\{x_n\}$ be a Cauchy sequence in E. Then $\{x_n\}$ is Cauchy in X, and the completeness of X implies that $\{x_n\}$ converges, say to some $x \in X$. By Corollary 2.23, x is in E. Thus $\{x_n\}$ is convergent in E. The converse is immediate from Corollary 2.23. □

Proposition 2.44. If S is a nonempty set, then the vector space $B(S)$ of bounded scalar-valued functions on S, with the uniform metric, is a complete metric space.

PROOF. Let $\{f_n\}$ be a Cauchy sequence in $B(S)$. Then $\{f_n(x)\}$ is a Cauchy sequence in $\mathbb{C}$ for each x in S. Define $f(x) = \lim_n f_n(x)$. For any $\epsilon > 0$, we know that there is an integer N such that $|f_n(x) - f_m(x)| < \epsilon$ whenever n and m are $\geq N$. Taking into account the continuity of the distance function on $\mathbb{C}$, i.e., the continuity of absolute value, we let m tend to infinity and obtain $|f_n(x) - f(x)| \leq \epsilon$ for $n \geq N$. Thus $\{f_n\}$ converges to f in $B(S)$. □

Corollary 2.45. Let (S, d) be a metric space. Then the vector space $C(S)$ of bounded continuous scalar-valued functions on S, with the uniform metric, is a complete metric space.

REMARK. $C(S)$ was observed to be a vector subspace in the remarks with Corollary 2.29.

PROOF. The space $B(S)$ is complete by Proposition 2.44, and $C(S)$ is a closed metric subspace by Corollary 2.24. Then $C(S)$ is complete by Proposition 2.43. □

Now we shall relate compactness and completeness. A metric space (X, d) is said to be **totally bounded** if for any $\epsilon > 0$, finitely many open balls of radius ϵ cover X.

Theorem 2.46. A metric space (X, d) is compact if and only if it is totally bounded and complete.

PROOF. Let (X, d) be compact. If $\epsilon > 0$ is given, the open balls $B(\epsilon; x)$ cover X. By compactness some finite number of the balls cover X. Therefore X is totally bounded. Next let a Cauchy sequence $\{x_n\}$ be given. By Theorem 2.36, $\{x_n\}$ has a convergent subsequence. A Cauchy sequence with a convergent subsequence is necessarily convergent, and it follows that X is complete.

In the reverse direction, let X be totally bounded and complete. Theorem 2.36 shows that it is enough to prove that any sequence $\{x_n\}$ in X has a convergent subsequence. By total boundedness, find finitely many open balls of radius 1 covering X. Then infinitely many of the x_n's have to lie in one of these balls, and hence there is a subsequence $\{x_{n_k}\}$ that lies in a single one of these balls of radius 1. Next finitely many open balls of radius 1/2 cover X. In the same way there is a subsequence $\{x_{n_{k_l}}\}$ of $\{x_{n_k}\}$ that lies in a single one of these balls of radius 1/2. Continuing in this way, we can find successive subsequences, the m^{th} of which lies in a single ball of radius $1/m$. The Cantor diagonal process, used in the proof of Theorem 1.22, allows us to form a single subsequence $\{x_{j_i}\}$ of $\{x_n\}$ such that for each m, $\{x_{j_i}\}$ is eventually in a ball of radius $1/m$. If $\epsilon > 0$ is given, find m such that $1/m < \epsilon$, and let c_m be the center of the ball of radius $1/m$. Choose an integer N such that x_{j_i} lies in $B(1/m; c_m)$ whenever $j_i \geq N$. If $j_i \geq N$ and $j_{i'} \geq N$, then $d(c_m, x_{j_i}) < \epsilon$ and $d(c_m, x_{j_{i'}}) < \epsilon$, whence $d(x_{j_i}, x_{j_{i'}}) < 2\epsilon$. Therefore the subsequence $\{x_{j_i}\}$ is Cauchy. By completeness it converges. Hence $\{x_n\}$ has a convergent subsequence, and the theorem is proved. □

Let (X, d) and (Y, ρ) be metric spaces, and let $f : X \to Y$ be uniformly continuous. Then f carries Cauchy sequences to Cauchy sequences. In fact, if $\{x_n\}$ is Cauchy in X and if $\epsilon > 0$ is given, choose some δ of uniform continuity for f and ϵ, and find an integer N such that $d(x_n, x_{n'}) < \delta$ whenever n and n' are $\geq N$. Then $\rho(f(x_n), f(x_{n'})) < \epsilon$ for the same n's and n''s, and hence $\{f(x_n)\}$ is Cauchy.

Proposition 2.47. Let (X, d) and (Y, ρ) be metric spaces with Y complete, let D be a dense subset of X, and let $f : D \to Y$ be uniformly continuous. Then f extends uniquely to a continuous function $F : X \to Y$, and F is uniformly continuous.

PROOF OF UNIQUENESS. If x is in X, apply Proposition 2.22a to choose a sequence $\{x_n\}$ in D with $x_n \to x$. Continuity of F forces $F(x_n) \to F(x)$. But $F(x_n) = f(x_n)$ for all n. Thus $F(x) = \lim_n f(x_n)$ is forced. □

PROOF OF EXISTENCE. If x is in X, choose $x_n \in D$ with $x_n \to x$. Since $\{x_n\}$ is convergent, it is Cauchy. Since f is uniformly continuous, $\{f(x_n)\}$ is

Cauchy. The completeness of Y then allows us to define $F(x) = \lim f(x_n)$, but we must see that F is well defined. For this purpose, suppose also that $\{y_n\}$ is a sequence in D that converges to x. Let $\{z_n\}$ be the sequence $x_1, y_1, x_2, y_2, \ldots$. This sequence is Cauchy, and $\{x_n\}$ and $\{y_n\}$ are subsequences of it. Therefore $\lim f(y_n) = \lim f(z_n) = \lim f(x_n)$, and $F(x)$ is well defined.

For the uniform continuity of F, let $\epsilon > 0$ be given, and choose some δ of uniform continuity for f and $\epsilon/3$. Suppose that x and x' are in X with $d(x, x') < \delta/3$. Choose x_n in D with $d(x_n, x) < \delta/3$ and $\rho(f(x_n), F(x)) < \epsilon/3$, and choose x_n' in D with $d(x_n', x') < \delta/3$ and $\rho(f(x_n'), F(x')) < \epsilon/3$. Then $d(x_n, x_n') < \delta$ by the triangle inequality, and hence $\rho(f(x_n), f(x_n')) < \epsilon/3$. Thus $\rho(F(x), F(x')) < \epsilon$ by the triangle inequality. □

8. Connectedness

Although the Intermediate Value Theorem (Theorem 1.12) in Section I.1 was derived from the Bolzano–Weierstrass Theorem, the Intermediate Value Theorem is not to be regarded as a consequence of compactness. Instead, the relevant property is "connectedness," which we discuss in this section.

A metric space (X, d) is **connected** if X cannot be written as $X = U \cup V$ with U and V open, disjoint, and nonempty. A subset E of X is **connected** if E is connected as a subspace of X, i.e., if E cannot be written as a disjoint union $(E \cap U) \cup (E \cap V)$ with U and V open in X and with $E \cap U$ and $E \cap V$ both nonempty. The disjointness in this definition is of $E \cap U$ and $E \cap V$; the open sets U and V may have nonempty intersection.

Proposition 2.48. The connected subsets of $\mathbb{R}$ are the intervals—open, closed, and half open.

PROOF. Let E be a connected subset of $\mathbb{R}$, and suppose that there are real numbers a, b, c such that $a < c < b$, a and b are in E, and c is not in E. Forming the open sets $U = (-\infty, c)$ and $V = (c, +\infty)$ in $\mathbb{R}$, we see that E is the disjoint union of $E \cap U$ and $E \cap V$ and that these two sets are nonempty. Thus E is not connected.

Conversely suppose that I is an open, closed, or half-open interval of $\mathbb{R}$ from a to b, with $a \neq b$ but with a or b or both allowed to be infinite. Arguing by contradiction, suppose that I is not connected. Choose open sets U and V in $\mathbb{R}$ such that I is the disjoint union of $I \cap U$ and $I \cap V$ and these two sets are nonempty. Without loss of generality, there exist members c and c' of $I \cap U$ and $I \cap V$, respectively, with $c < c'$. Since U is open and c has to be $< b$, all real numbers $c + \epsilon$ with $\epsilon > 0$ sufficiently small are in $I \cap U$. Let $d = \sup \{x \mid [c, x) \subseteq I \cap U\}$, so that $d > c$.

If $d < b$, then the fact that U is open implies that d is not in $I \cap U$. Thus d is in $I \cap V$. Since V is open and $d > a$, $d - \epsilon$ is in $I \cap V$ if $\epsilon > 0$ is sufficiently small. But then $d - \epsilon$ is in both $I \cap U$ and $I \cap V$ for ϵ sufficiently small. This is a contradiction, and we conclude that $d = b$.

If $d = b$ is in $I \cap V$, then the same argument shows that $b - \epsilon$ is in both $I \cap U$ and $I \cap V$ for ϵ positive and sufficiently small, and we again have a contradiction. Consequently all points from c to the right end of I are in $I \cap U$. This is again a contradiction, since c' is known to be in $I \cap V$. □

Proposition 2.49. The continuous image of a connected metric space is connected.

PROOF. Let (X, d) and (Y, ρ) be metric spaces with X connected, and let $f : X \to Y$ be continuous. We are to prove that $f(X)$ is connected. Corollary 2.27 shows that there is no loss of generality in assuming that $f(X) = Y$, i.e., f is onto. Arguing by contradiction, suppose that Y is the union $Y = U \cup V$ of disjoint nonempty open sets. Then $X = f^{-1}(U) \cup f^{-1}(V)$ exhibits X as the disjoint union of nonempty sets, and these sets are open as a consequence of Proposition 2.15a. Thus X is not connected. □

Corollary 2.50 (Intermediate Value Theorem). For real-valued functions of a real variable, the continuous image of any interval is an interval.

PROOF. This is immediate from Propositions 2.48 and 2.49. □

Further connected sets beyond those in $\mathbb{R}$ are typically built from other connected sets. One tool is a **path** in X, which is a continuous function from a closed bounded interval $[a, b]$ into X. The image of a path is connected by Propositions 2.48 and 2.49. A metric space (X, d) is **pathwise connected** if for any two points x_1 and x_2 in X, there is some path p from x_1 to x_2, i.e., if there is some continuous $p : [a, b] \to X$ with $p(a) = x_1$ and $p(b) = x_2$.

A pathwise-connected metric space (X, d) is necessarily connected. In fact, otherwise we could write X as a disjoint union of two nonempty open sets U and V. Let x_1 be in U and x_2 be in V, and let $p : [a, b] \to X$ be a path from x_1 to x_2. Then $p([a, b]) = (p([a, b]) \cap U) \cup (p([a, b]) \cap V)$ exhibits $p([a, b])$ as a disjoint union of relatively open sets, and these sets are nonempty, since x_1 is in the first set and x_2 is in the second set. Consequently $p([a, b])$ is not connected, in contradiction to the fact that the image of any path is connected.

We can view a pathwise-connected metric space as the union of images of paths from a single point to all other points, and such a union is then connected. The following proposition generalizes this construction.

Proposition 2.51. If (X, d) is a metric space and $\{E_\alpha\}$ is a system of connected subsets of X with a point x_0 in common, then $\bigcup_\alpha E_\alpha$ is connected.

PROOF. Assuming the contrary, find open sets U and V in X such that $\bigcup_\alpha E_\alpha$ is the disjoint union of its intersections with U and V and these two intersections are both nonempty. Say that x_0 is in U. Since E_α is connected and x_0 is in $E_\alpha \cap U$, the decomposition $E_\alpha = (E_\alpha \cap U) \cup (E_\alpha \cap V)$ forces $E_\alpha \cap V$ to be empty. Then $\big(\bigcup_\alpha E_\alpha\big) \cap V = \bigcup_\alpha (E_\alpha \cap V)$ is empty, and we have arrived at a contradiction. $\square$

It follows from Proposition 2.51 that the union of all connected subsets of X that contain x_0 is connected. This set is called the **connected component** of x_0 in X. The metric space X is the disjoint union of its connected components. The next result implies that these connected components are closed sets.

Proposition 2.52. If (X, d) is a metric space and E is a connected subset of X, then the closure E^{cl} is connected.

PROOF. Suppose that U and V are open sets in X such that E^{cl} is contained in $U \cup V$ and $E^{\text{cl}} \cap U \cap V$ is empty. We are to prove that $E^{\text{cl}} \cap U$ and $E^{\text{cl}} \cap V$ cannot both be nonempty. Arguing by contradiction, let x be in $E^{\text{cl}} \cap U$ and let y be in $E^{\text{cl}} \cap V$. Since E is connected, $E \cap U$ and $E \cap V$ cannot both be nonempty, and thus x and y cannot both be in E. Thus at least one of them, say x, is a limit point of E. Since U is a neighborhood of x, U contains a point e of E different from x. Thus e is in $E \cap U$. Since y cannot then be in $E \cap V$, y is a limit point of E. Since V is a neighborhood of y, V contains a point f of E different from y. Thus f is in $E \cap V$, and we have arrived at a contradiction. $\square$

EXAMPLE. The graph in $\mathbb{R}^2$ of $\sin(1/x)$ for $0 < x \le 1$ is pathwise connected, and we have seen that pathwise-connected sets are connected. The closure of this graph consists of the graph together with all points $(0, t)$ for $-1 \le t \le 1$, and this closure is connected by Proposition 2.52. One can show, however, that this closure is not pathwise connected. Thus we obtain an example of a connected set in $\mathbb{R}^2$ that is not pathwise connected.

9. Baire Category Theorem

A number of deep results in analysis depend critically on the fact that some metric space is complete. Already we have seen that the metric space $C(S)$ of bounded continuous scalar-valued functions on a metric space is complete, and we shall see as not too hard a consequence in Chapter XII that there exists a continuous periodic function whose Fourier series diverges at a point. One of the features of the Lebesgue integral in Chapter V will be that the metric spaces of integrable functions and of square-integrable functions, with their natural metrics, are further examples of complete metric spaces. Thus these spaces too are available for applications that make use of completeness.

The main device through which completeness is transformed into a powerful hypothesis is the Baire Category Theorem below. A closed set in a metric space is **nowhere dense** if its interior is empty. Its complement is an open dense set, and conversely the complement of any open dense set is closed nowhere dense.

EXAMPLE. A nontrivial example of a closed nowhere dense set is a **Cantor set**[4] in $\mathbb{R}$. This is a set constructed from a closed bounded interval of $\mathbb{R}$ by removing an open interval in the middle of length a fraction r_1 of the total length with $0 < r_1 < 1$, removing from each of the 2 remaining closed subintervals an open interval in the middle of length a fraction r_2 of the total length of the subinterval, removing from each of the 4 remaining closed subintervals an open interval in the middle of length a fraction r_3 of the total length of the interval, and so on indefinitely. The Cantor set is obtained as the intersection of the approximating sets. It is closed, being the intersection of closed sets, and it is nowhere dense because it contains no interval of more than one point. For the **standard Cantor set**, the starting interval is $[0, 1]$, and the fractions are given by $r_1 = r_2 = \cdots = \frac{1}{3}$ at every stage. In general, the "length" of the resulting set[5] is the product of the length of the starting interval and $\prod_{n=1}^{\infty} (1 - r_n)$.

Theorem 2.53 (Baire Category Theorem). If (X, d) is a complete metric space, then

(a) the intersection of countably many open dense sets is nonempty,
(b) X is not the union of countably many closed nowhere dense sets.

PROOF. Conclusions (a) and (b) are equivalent by taking complements. Let us prove (a). Suppose that U_n is open and dense for $n \geq 1$. Since U_1 is nonempty and open, let E_1 be an open ball $B(r_1; x_1)$ whose closure is in U_1 and whose radius is $r_1 \leq 1$. We construct inductively open balls $E_n = B(r_n; x_n)$ with $r_n \leq \frac{1}{n}$ such that $E_n \subseteq U_1 \cap \cdots \cap U_n$ and $E_n^{\text{cl}} \subseteq E_{n-1}$. Suppose E_n with $n \geq 1$ has been constructed. Since U_{n+1} is dense and E_n is nonempty and open, $U_{n+1} \cap E_n$ is not empty. Let x_{n+1} be a point in $U_{n+1} \cap E_n$. Since $U_{n+1} \cap E_n$ is open, we can find an open ball $E_{n+1} = B(r_{n+1}; x_{n+1})$ with radius $r_{n+1} \leq \frac{1}{n+1}$ and center a point x_{n+1} in U_{n+1} such that $E_{n+1}^{\text{cl}} \subseteq U_{n+1} \cap E_n$. Then E_{n+1} has the required properties, and the inductive construction is complete. The sequence $\{x_n\}$ is

[4] Often a mathematician who refers to "the" Cantor set is referring to what is called the "standard Cantor set" later in the present paragraph.

[5] To be precise, the length is the "Lebesgue measure" of the set in the sense to be defined in Chapter V.

Cauchy because whenever $n \geq m$, the points x_n and x_m are both in E_m and thus have $d(x_n, x_m) < \frac{1}{m}$. Since X is by assumption complete, let $x_n \to x$. For any integer N, the inequality $n > N$ implies that x_n is in E_{N+1}. Thus the limit x is in $E_{N+1}^{\text{cl}} \subseteq E_N \subseteq U_1 \cap \cdots \cap U_N$. Since N is arbitrary, x is in $\bigcap_{n=1}^{\infty} U_n$. $\square$

REMARK. In (a), the intersection in question is dense, not merely nonempty. To see this, we observe in the first part of the proof that since U_1 is dense, E_1 can be chosen to be arbitrarily close to any member of X and to have arbitrarily small radius. Following through the construction, we see that x is in E_1 and hence can be arranged to be as close as we want to any member of X. The corresponding conclusion in (b) is that a nonempty open subset of X is never contained in the countable union of closed nowhere dense sets.

EXAMPLES.

(1) The subset $\mathbb{Q}$ of rationals in $\mathbb{R}$ is not the countable intersection of open sets. In fact, assume the contrary, and write $\mathbb{Q} = \bigcap_{n=1}^{\infty} U_n$ with U_n open. Each set U_n contains $\mathbb{Q}$ and hence is dense in $\mathbb{R}$. Also, for $q \in \mathbb{Q}$, the set $\mathbb{R} - \{q\}$ is open and dense. Thus the equality $\mathbb{Q} = \bigcap_{n=1}^{\infty} U_n$ implies that

$$\Big(\bigcap_{n=1}^{\infty} U_n\Big) \cap \Big(\bigcap_{q \in \mathbb{Q}} (\mathbb{R} - \{q\})\Big)$$

is empty, in contradiction to Theorem 2.53.

(2) Let us start with a Cantor set as at the beginning of this section. The total interval is to be $[0, 1]$, and the set is to be built with middle segments of fractions $r_1, r_2, \ldots$. Within the closure of each removed open interval, we insert a Cantor set for that interval, possibly with different fractions $r_1, r_2, \ldots$ for each inserted Cantor set. This insertion involves further removed open intervals, and we insert a Cantor set into each of these. We continue this process indefinitely. The union of the constructed sets is dense. Can it be the entire interval $[0, 1]$? The answer is "no" because each of the Cantor sets is closed nowhere dense and because by Theorem 2.53, the interval $[0, 1]$ is not the countable union of closed nowhere dense sets.

A subset E of a metric space is said to be of the **first category** if it is contained in the countable union of closed nowhere dense sets. Theorem 2.53 and the remark after it together imply that no nonempty open set in a complete metric space is of the first category.

Theorem 2.54. Let (X, d) be a complete metric space, and let U be an open subset of X. Suppose for $n \geq 1$ that $f_n : U \to \mathbb{C}$ is a continuous function and that f_n converges pointwise to a function $f : U \to \mathbb{C}$. Then the set of discontinuities of f is of the first category.

The proof will make use of the notion of the **oscillation** of a complex-valued function on a metric space U. For any function $g : U \to \mathbb{C}$, define

$$\operatorname{osc}_g(x_0) = \lim_{\delta \downarrow 0} \sup_{x \in B(\delta; x_0)} |g(x) - g(x_0)|,$$

so that g is continuous at x_0 if and only if $\operatorname{osc}_g(x_0) = 0$. At first glance it might seem that the sets $\{x \mid \operatorname{osc}_g(x) \geq r\}$ are always closed, no matter what discontinuities g has. Actually, these sets need not be closed. Take, for example, the function $g : \mathbb{R} \to \mathbb{R}$ that is 1 at every nonzero rational, 0 at every irrational, and 1/2 at 0. Then $\operatorname{osc}_g(x)$ is 1 at every x in $\mathbb{R}$ except for $x = 0$, where it is 1/2. Thus, in this example, the set $\{x \mid \operatorname{osc}_g(x) \geq 1\}$ is $\mathbb{R} - \{0\}$ and is not closed.

Lemma 2.55. Let (U, d) be a metric space. If $g : U \to \mathbb{C}$ is a function and $\epsilon > 0$ is a positive number, then

$$\{x \in U \mid \operatorname{osc}_g(x) \geq 2\epsilon\}^{\mathrm{cl}} \subseteq \{x \in U \mid \operatorname{osc}_g(x) \geq \tfrac{\epsilon}{2}\}.$$

PROOF. We need to see that the limit points of the set on the left are in the set on the right. Thus suppose that $\operatorname{osc}_g(x_n) \geq 2\epsilon$ for all n and that $x_n \to x_0$. For each n, choose $x_{n,m}$ such that $\lim_m x_{n,m} = x_n$ and $|g(x_{n,m}) - g(x_n)| \geq \epsilon$ for all m. Because of the convergence of $x_{n,m}$ to x_n, we may choose, for each n, an integer $m = m_n$ such that $d(x_{n,m_n}, x_n) < d(x_0, x_n)$, and then $\lim_n x_{n,m_n} = x_0$ by the triangle inequality. From $|g(x_{n,m_n}) - g(x_n)| \geq \epsilon$, the triangle inequality forces

$$|g(x_{n,m_n}) - g(x_0)| \geq \tfrac{\epsilon}{2} \qquad \text{or} \qquad |g(x_n) - g(x_0)| \geq \tfrac{\epsilon}{2}. \tag{$*$}$$

Defining y_n to be x_{n,m_n} or x_n according as the first or second inequality is the case in $(*)$, we have $y_n \to x_0$ and $|g(y_n) - g(x_0)| \geq \frac{\epsilon}{2}$. This proves the lemma. □

PROOF OF THEOREM 2.54. In view of Lemma 2.55 and the fact that U is open, it is enough to prove for each $\epsilon > 0$ that $\{x \mid \operatorname{osc}_f(x) \geq \epsilon\}$ does not contain a nonempty open subset of X. Assuming the contrary, suppose that it contains the nonempty open set V. Define

$$A_{mn} = \{x \in V \mid |f_m(x) - f_n(x)| \leq \tfrac{\epsilon}{4}\}.$$

This is a relatively closed subset of V. Then $A_m = \bigcap_{n \geq m} A_{mn}$ is closed in V. If x is in V, the fact that $\{f_n(x)\}$ is a Cauchy sequence implies that there is some m such that x is in A_{mn} for all $n \geq m$. Hence $\bigcup_{m=1}^{\infty} A_m = V$. Since V is open in a complete metric space, Theorem 2.53 and the remark after it show that some A_m has nonempty interior. Fix that m, and let W be its nonempty interior. Since

$$A_m \subseteq \{x \in V \mid |f_m(x) - f(x)| \leq \tfrac{\epsilon}{4}\},$$

every point of W has $|f_m(x) - f(x)| \leq \frac{\epsilon}{4}$ and $\operatorname{osc}_f(x) \geq \epsilon$. Let x_0 be in W and choose x_n tending to x_0 with $|f(x_n) - f(x_0)| \geq \frac{3\epsilon}{4}$. From $|f_m(x_n) - f(x_n)| \leq \frac{\epsilon}{4}$ and $|f_m(x_0) - f(x_0)| \leq \frac{\epsilon}{4}$, we obtain $|f_m(x_n) - f_m(x_0)| \geq \frac{\epsilon}{4}$. Since x_n converges to x_0, this inequality contradicts the continuity of f_m at x_0. □

10. Properties of $C(S)$ for Compact Metric S

If (S, d) is a metric space, then we saw in Proposition 2.44 that the vector space $B(S)$ of bounded scalar-valued functions on S, in the uniform metric, is a complete metric space. We saw also in Corollary 2.45 that the vector subspace $C(S)$ of bounded continuous functions is a complete subspace. In this section we shall study the space $C(S)$ further under the assumption that S is compact. In this case Propositions 2.38 and 2.34 tell us that every continuous scalar-valued function on S is automatically bounded and hence is in $C(S)$.

The first result about $C(S)$ for S compact is a generalization of Ascoli's Theorem from its setting in Theorem 1.22 for real-valued functions on a bounded interval $[a, b]$. The generalized theorem provides an insight that is not so obvious from the special case that S is a closed bounded interval of $\mathbb{R}$. The insight is a characterization of the compact subsets of $C(S)$ when S is compact, and it is stated precisely in Corollary 2.57 below. The relevant definitions for Ascoli's Theorem are generalized in the expected way. Let $\mathcal{F} = \{f_\alpha \mid \alpha \in A\}$ be a set of scalar-valued functions on the compact metric space S. We say that $\mathcal{F}$ is **equicontinuous** at $x \in S$ if for each $\epsilon > 0$, there is some $\delta > 0$ such that $d(t, x) < \delta$ implies $|f(t) - f(x)| < \epsilon$ for all $f \in \mathcal{F}$. The set $\mathcal{F}$ of functions is **pointwise bounded** if for each $t \in S$, there exists a number M_t such that $|f(t)| \le M_t$ for all $f \in \mathcal{F}$. The set is **uniformly equicontinuous** on S if it is equicontinuous at each point $x \in S$ and if the δ can be taken independent of x. The set is **uniformly bounded** on S if it is pointwise bounded at each $t \in S$ and the bound M_t can be taken independent of t; this last definition is consistent with the definition of a uniformly bounded sequence of functions given in Section 4.

Theorem 2.56 (Ascoli's Theorem). Let (S, d) be a compact metric space. If $\{f_n\}$ is a sequence of scalar-valued functions on S that is equicontinuous at each point of S and pointwise bounded on S, then

(a) $\{f_n\}$ is uniformly equicontinuous and uniformly bounded on S,
(b) $\{f_n\}$ has a uniformly convergent subsequence.

REMARKS. The proof involves only notational changes from the special case Theorem 1.22; there are enough such changes, however, so that it is worth writing out the details. Inspection of this proof shows also that the range $\mathbb{R}$ or $\mathbb{C}$ may be replaced by any compact metric space. We shall see a further generalization of this theorem in Chapter X, and the proof at that time will look quite different.

PROOF. Since each f_n is continuous at each point, we know from Propositions 2.38, 2.34a, and 2.41 that each f_n is uniformly continuous and bounded. The proof of (a) amounts to an argument that the estimates in those theorems can be arranged to apply simultaneously for all n.

First consider the question of uniform boundedness. Choose, by Corollary 2.39, some x_n in S with $|f_n(x_n)|$ equal to $K_n = \sup_{x\in S} |f_n(x)|$. Then choose a subsequence on which the numbers K_n tend to $\sup_n K_n$ in $\mathbb{R}^*$. There will be no loss of generality in assuming that this subsequence is our whole sequence. By compactness of S, apply the Bolzano–Weierstrass property given in Theorem 2.36 to find a convergent subsequence $\{x_{n_k}\}$ of $\{x_n\}$, and let x_0 be the limit of this subsequence. By pointwise boundedness, find M_{x_0} with $|f_n(x_0)| \le M_{x_0}$ for all n. Then choose some δ of equicontinuity at x_0 for $\epsilon = 1$. As soon as k is large enough so that $d(x_{n_k}, x_0) < \delta$, we have

$$K_{n_k} = |f_{n_k}(x_{n_k})| \le |f_{n_k}(x_{n_k}) - f_{n_k}(x_0)| + |f_{n_k}(x_0)| < 1 + M_{x_0}.$$

Thus $1 + M_{x_0}$ is a uniform bound for the functions f_n.

For the uniform equicontinuity, fix $\epsilon > 0$. The uniform continuity of f_n for each n, as given in Proposition 2.41, means that it makes sense to define

$$\delta_n(\epsilon) = \min\left\{1,\ \sup\left\{\delta' > 0 \;\middle|\; \begin{array}{l} |f_n(x) - f_n(y)| < \epsilon \text{ whenever} \\ d(x,y) < \delta' \text{ and } x \text{ and } y \text{ are in } S \end{array}\right\}\right\}.$$

If $d(x, y) < \delta_n(\epsilon)$, then $|f_n(x) - f_n(y)| < \epsilon$. Put $\delta(\epsilon) = \inf_n \delta_n(\epsilon)$. Let us see that it is enough to prove that $\delta(\epsilon) > 0$: If x and y are in S with $d(x, y) < \delta(\epsilon)$, then $d(x, y) < \delta(\epsilon) \le \delta_n(\epsilon)$. Hence $|f_n(x) - f_n(y)| < \epsilon$ as required.

Thus we are to prove that $\delta(\epsilon) > 0$. If $\delta(\epsilon) = 0$, then we first choose a strictly increasing sequence $\{n_k\}$ of positive integers such that $\delta_{n_k}(\epsilon) < \frac{1}{k}$, and we next choose x_k and y_k in S with $d(x_k, y_k) < 1/k$ and $|f_{n_k}(x_k) - f_{n_k}(y_k)| \ge \epsilon$. Using the Bolzano–Weierstrass property again, we obtain a subsequence $\{x_{k_l}\}$ of $\{x_k\}$ such that $\{x_{k_l}\}$ converges, say to a limit x_0. Then

$$\limsup_l d(y_{k_l}, x_0) \le \limsup_l d(y_{k_l}, x_{k_l}) + \limsup_l d(x_{k_l}, x_0) = 0 + 0 = 0,$$

so that $\{y_{k_l}\}$ converges to x_0. Now choose, by equicontinuity at x_0, a number $\delta' > 0$ such that $|f_n(x) - f_n(x_0)| < \frac{\epsilon}{2}$ for all n whenever $d(x, x_0) < \delta'$. The convergence of $\{x_{k_l}\}$ and $\{y_{k_l}\}$ to x_0 implies that for large enough l, we have $d(x_{k_l}, x_0) < \delta'$ and $d(y_{k_l}, x_0) < \delta'$. Therefore $|f_{n_{k_l}}(x_{k_l}) - f_{n_{k_l}}(x_0)| < \frac{\epsilon}{2}$ and $|f_{n_{k_l}}(y_{k_l}) - f_{n_{k_l}}(x_0)| < \frac{\epsilon}{2}$, from which we conclude that $|f_{n_{k_l}}(x_{k_l}) - f_{n_{k_l}}(y_{k_l})| < \epsilon$. But we saw that $|f_{n_k}(x_k) - f_{n_k}(y_k)| \ge \epsilon$ for all k, and thus we have arrived at a contradiction. This proves the uniform equicontinuity and completes the proof of (a).

To prove (b), let R be a compact set containing all sets $\operatorname{image}(f_n)$. Choose a countable dense set D in S by Proposition 2.33. Using the Cantor diagonal process and the Bolzano–Weierstrass property of R, we construct a subsequence

$\{f_{n_k}\}$ of $\{f_n\}$ that is convergent at every point in D. Let us prove that $\{f_{n_k}\}$ is uniformly Cauchy. Redefining our indices, we may assume that $n_k = k$ for all k. Let $\epsilon > 0$ be given, and let δ be some corresponding number exhibiting equicontinuity. The balls $B(\delta; r)$ centered at the members r of D cover S, and the compactness of S gives us finitely many of their centers $r_1, \dots, r_l$ such that any member of S is within δ of at least one of $r_1, \dots, r_l$. Then choose N with $|f_n(r_j) - f_m(r_j)| < \epsilon$ for $1 \le j \le l$ whenever n and m are $\ge N$. If x is in S, let $r(x)$ be an r_j with $d(x, r(x)) < \delta$. Whenever n and m are $\ge N$, we then have

$$\begin{aligned}
&|f_n(x) - f_m(x)| \\
&\quad \le |f_n(x) - f_n(r(x))| + |f_n(r(x)) - f_m(r(x))| + |f_m(r(x)) - f_m(x)| \\
&\quad < \epsilon + \epsilon + \epsilon = 3\epsilon.
\end{aligned}$$

Hence $\{f_{n_k}\}$ is uniformly Cauchy, and (b) follows since the metric space $C(S)$ is complete. □

Corollary 2.57. If (S, d) is a compact metric space, then a subset E of $C(S)$ in the uniform metric has compact closure if and only if E is uniformly bounded and uniformly equicontinuous.

PROOF. First let us see that if E is uniformly bounded and uniformly equicontinuous, then so is E^{cl}. In fact, if $|f(x)| \le M$ for $f \in E$, then the same thing is true of any uniform limit of such functions. Hence E^{cl} is uniformly bounded. For the uniform equicontinuity of E^{cl}, let ϵ be given, and find some δ of equicontinuity for ϵ and the members of E. If f is a limit point of E, we can find a sequence $\{f_n\}$ in E converging uniformly to f. If $d(x, y) < \delta$, then the inequality

$$|f(x) - f(y)| \le |f(x) - f_n(x)| + |f_n(x) - f_n(y)| + |f_n(y) - f(y)|$$

and the uniform convergence show that we obtain $|f(x) - f(y)| < 3\epsilon$ by fixing any sufficiently large n. Thus E^{cl} is uniformly equicontinuous.

Now suppose that E is a closed subset of $C(S)$ that is uniformly bounded and equicontinuous. Then Theorem 2.56 shows that any sequence in E has a subsequence that is convergent in $C(S)$. Since E is closed, the sequence is convergent in E. Theorem 2.36 then shows that E is compact.

Conversely suppose that E is compact in $C(S)$. Distance from 0 in $C(S)$ is a continuous real-valued function by Corollary 2.17, and this continuous function has to be bounded on the compact set E. Thus E is uniformly bounded. For the uniform equicontinuity, let $\epsilon > 0$ be given. Theorem 2.46 shows that E is totally bounded. Hence we can find a finite set $f_1, \dots, f_l$ in E such that each member f of E has $\sup_{x \in S} |f(x) - f_j(x)| < \epsilon$ for some j. By uniform continuity of each f_i, choose some number $\delta > 0$ such that $d(x, y) < \delta$ implies $|f_i(x) - f_i(y)| < \epsilon$

for $1 \leq i \leq l$. If f_j is the member of the finite set associated with f, then $d(x, y) < \delta$ implies

$$|f(x) - f(y)| \leq |f(x) - f_j(x)| + |f_j(x) - f_j(y)| + |f_j(y) - f(y)| < 3\epsilon.$$

Hence E is uniformly equicontinuous. $\square$

The second result about $C(S)$ when S is compact generalizes the Weierstrass Approximation Theorem (Theorem 1.52) of Section I.9. We shall make use of a special case of the Weierstrass theorem in the proof—that $|x|$ is the uniform limit on $[-1, 1]$ of polynomials $P_n(x)$ with $P_n(0) = 0$. This special case was proved also by a direct argument in Section I.8.

Let us distinguish the case of real-valued functions from that of complex-valued functions, writing $C(S, \mathbb{R})$ and $C(S, \mathbb{C})$ in the two cases. The theorem in question gives a sufficient condition for a "subalgebra" of $C(S, \mathbb{R})$ or $C(S, \mathbb{C})$ to be dense in the whole space in the uniform metric. Pointwise addition and scalar multiplication make $C(S, \mathbb{R})$ into a real vector space and $C(S, \mathbb{C})$ into a complex vector space, and each space has also the operation of pointwise multiplication; all of these operations on functions preserve continuity as a consequence of Corollary 2.29. By a **subalgebra** of $C(S, \mathbb{R})$ or $C(S, \mathbb{C})$, we mean any nonempty subset that is closed under *all* these operations. The space $C(S, \mathbb{C})$ has also the operation of complex conjugation; this again preserves continuity by Corollary 2.29.

We shall work with a subalgebra of $C(S, \mathbb{R})$ or of $C(S, \mathbb{C})$, and we shall assume that the subalgebra is closed under complex conjugation in the case of complex scalars. The closure of such a subalgebra in the uniform metric is again a subalgebra. To see that this closure is a subalgebra requires checking each operation separately, and we confine our attention to pointwise multiplication. If sequences $\{f_n\}$ and $\{g_n\}$ converge uniformly to f and g, then $\{f_n g_n\}$ converges uniformly to fg because

$$\begin{aligned}
&\sup_{x \in S} |f_n(x) g_n(x) - f(x) g(x)| \\
&\qquad \leq \sup_{x \in S} |f_n(x)(g_n(x) - g(x))| + \sup_{x \in S} |(f_n(x) - f(x)) g(x)| \\
&\qquad \leq \Big(\sup_{x \in S} |f_n(x)|\Big)\Big(\sup_{x \in S} |g_n(x) - g(x)|\Big) + \Big(\sup_{x \in S} |g(x)|\Big)\Big(\sup_{x \in S} |f_n(x) - f(x)|\Big)
\end{aligned}$$

with $\sup_{x \in S} |g(x)|$ finite and $\sup_{x \in S} |f_n(x)|$ convergent to $\sup_{x \in S} |f(x)|$.

We say that a subalgebra of $C(S, \mathbb{R})$ or $C(S, \mathbb{C})$ **separates points** if for each pair of distinct points x_1 and x_2 in S, there is some f in the subalgebra with $f(x_1) \neq f(x_2)$.

Theorem 2.58 (Stone–Weierstrass Theorem). Let (S, d) be a compact metric space.

(a) If $\mathcal{A}$ is a subalgebra of $C(S, \mathbb{R})$ that separates points and contains the constant functions, then $\mathcal{A}$ is dense in $C(S, \mathbb{R})$ in the uniform metric.

(b) If $\mathcal{A}$ is a subalgebra of $C(S, \mathbb{C})$ that separates points, contains the constant functions, and is closed under complex conjugation, then $\mathcal{A}$ is dense in $C(S, \mathbb{C})$ in the uniform metric.

PROOF OF (a). Let $\mathcal{A}^{\mathrm{cl}}$ be the closure of $\mathcal{A}$ in the uniform metric. We recalled above from Chapter I that $|t|$ is the limit of polynomials $t \mapsto P_n(t)$ uniformly on $[-1, 1]$. It follows that $|t|$ is the limit of polynomials $t \mapsto Q_n(t) = MP_n(M^{-1}t)$ uniformly on $[-M, M]$. Taking $M = \sup_{x \in S} |f(x)|$, we see that $|f|$ is in $\mathcal{A}^{\mathrm{cl}}$ whenever f is in $\mathcal{A}$.

Since $\mathcal{A}^{\mathrm{cl}}$ is a subalgebra closed under addition and scalar multiplication as well, the formulas

$$\max\{f, g\} = \tfrac{1}{2}(f + g) + \tfrac{1}{2}|f - g|,$$
$$\min\{f, g\} = \tfrac{1}{2}(f + g) - \tfrac{1}{2}|f - g|,$$

show that $\mathcal{A}^{\mathrm{cl}}$ is closed under pointwise maximum and pointwise minimum for two functions. Iterating, we see that $\mathcal{A}^{\mathrm{cl}}$ is closed under pointwise maximum and pointwise minimum for n functions for any integer $n \geq 2$.

The heart of the proof is an argument that if $f \in C(S, \mathbb{R})$, $x \in S$, and $\epsilon > 0$ are given, then there exists g_x in $\mathcal{A}^{\mathrm{cl}}$ such that $g_x(x) = f(x)$ and

$$g_x(s) > f(s) - \epsilon$$

for all $s \in S$. The argument is as follows: For each $y \in S$ other than x, there exists a function in $\mathcal{A}$ taking distinct values at x and y. Some linear combination of this function and the constant function 1 is a function h_y in $\mathcal{A}$ with $h_y(x) = f(x)$ and $h_y(y) = f(y)$. To complete the definition of h_y for all $y \in S$, we set h_x equal to the constant function $f(x)1$. The continuity of h_y and the equality $h_y(y) = f(y)$ imply that there exists an open neighborhood U_y of y such that $h_y(s) > f(s) - \epsilon$ for all $s \in U_y$. As y varies, these open neighborhoods cover S, and by compactness of S, finitely many suffice, say $U_{y_1}, \dots U_{y_k}$. Then the function $g_x = \max\{h_{y_1}, \dots, h_{y_k}\}$ has $g_x(s) > f(s) - \epsilon$ for all $s \in S$. Also, it has $g_x(x) = f(x)$, and it is in $\mathcal{A}^{\mathrm{cl}}$, since $\mathcal{A}^{\mathrm{cl}}$ is closed under pointwise maxima.

To complete the proof of (a), we continue with $f \in C(S, \mathbb{R})$ and $\epsilon > 0$ as above. We shall produce a member h of $\mathcal{A}^{\mathrm{cl}}$ such that $|h(s) - f(s)| < \epsilon$ for all $s \in S$. For each x, the continuity of g_x and the equality $g_x(x) = f(x)$ imply that there is an open neighborhood V_x of x such that $g_x(s) < f(s) + \epsilon$ for all

$s \in V_x$. As x varies, these open neighborhoods cover S, and by compactness of S, finitely many suffice, say $V_{x_1}, \dots V_{x_l}$. The function $h = \min\{g_{x_1}, \dots, g_{x_l}\}$ has $h(s) < f(s) + \epsilon$ for all $s \in S$, and it is in $(\mathcal{A}^{\mathrm{cl}})^{\mathrm{cl}} = \mathcal{A}^{\mathrm{cl}}$, since each g_{x_j} is in $\mathcal{A}^{\mathrm{cl}}$. Since each g_{x_j} has $g_{x_j}(s) > f(s) - \epsilon$ for all $s \in S$, we have $h(s) > f(s) - \epsilon$ as well. Thus $|h(s) - f(s)| < \epsilon$ for all $s \in S$.

Since ϵ is arbitrary, we conclude that f is a limit point of $\mathcal{A}^{\mathrm{cl}}$. But $\mathcal{A}^{\mathrm{cl}}$ is closed, and hence f is in $\mathcal{A}^{\mathrm{cl}}$. Therefore $\mathcal{A}^{\mathrm{cl}} = C(S, \mathbb{R})$. □

PROOF OF (b). Let $\mathcal{A}_{\mathbb{R}}$ be the subset of members of $\mathcal{A}$ that take values in $\mathbb{R}$. Then $\mathcal{A}_{\mathbb{R}}$ is certainly closed under addition, multiplication by real scalars, and pointwise multiplication, and the real-valued constant functions are in $\mathcal{A}_{\mathbb{R}}$. If $f = u + iv$ is in $\mathcal{A}$ and has real and imaginary parts u and v, then $\bar{f}$ is in $\mathcal{A}$ by assumption, and hence so are $u = \frac{1}{2}(f + \bar{f})$ and $v = \frac{1}{2i}(f - \bar{f})$. We are given that $\mathcal{A}$ separates points of S. If x_1 and x_2 are distinct points of S with $f(x_1) \neq f(x_2)$, then either $u(x_1) \neq u(x_2)$ or $v(x_1) \neq v(x_2)$, and it follows that $\mathcal{A}_{\mathbb{R}}$ separates points. By (a), $\mathcal{A}_{\mathbb{R}}$ is dense in $C(S, \mathbb{R})$. Finally let $f = u + iv$ be in $C(S, \mathbb{C})$, and let $\{u_n\}$ and $\{v_n\}$ be sequences in $\mathcal{A}_{\mathbb{R}}$ converging uniformly to u and v, respectively. Then $\{u_n + iv_n\}$ is a sequence in $\mathcal{A}$ converging uniformly to f. Hence $\mathcal{A}$ is dense in $C(S, \mathbb{C})$. □

EXAMPLES.

(1) On a closed bounded interval $[a, b]$ of the line, the scalar-valued polynomials form an algebra that separates points, contains the constants, and is closed under conjugation. The Stone–Weierstrass Theorem in this case reduces to the Weierstrass Theorem (Theorem 1.52), saying that the polynomials are dense in $C([a, b])$.

(2) Consider the algebra of continuous complex-valued periodic functions on $[-\pi, \pi]$ and the subalgebra of complex-valued trigonometric polynomials $\sum_{n=-N}^{N} c_n e^{inx}$; here N depends on the trigonometric polynomial. Neither the algebra nor the subalgebra separates points, since all functions in question have $f(-\pi) = f(\pi)$. To make the theorem applicable, we consider the domain of these functions to be the unit circle of $\mathbb{C}$, parametrized by e^{ix}; this parametrization is permissible by Corollary 1.45, and continuity is preserved. The Stone–Weierstrass Theorem then applies and gives a new proof that the trigonometric polynomials are dense in the space of complex-valued continuous periodic functions; our earlier proof was constructive, deducing the result as part of Fejér's Theorem (Theorem 1.59).

(3) Let S^{n-1} be the unit sphere $\{x \in \mathbb{R}^n \mid |x| = 1\}$ in $\mathbb{R}^n$. The restrictions to S^{n-1} of all scalar-valued polynomials $P(x_1, \dots, x_n)$ in n variables form a subalgebra of $C(S^{n-1})$ that separates points, contains the constants, and is closed

under conjugation. The Stone–Weierstrass Theorem says that this subalgebra is dense in $C(S^{n-1})$.

(4) Let S be the closed unit disk $\{z \mid |z| \le 1\}$ in $\mathbb{C}$. The set $\mathcal{A}$ of restrictions to S of sums of power series having infinite radius of convergence is a subalgebra of $C(S, \mathbb{C})$ that separates points and contains the constants. However, the continuous function $\bar{z}$ is not in the closure, because it has integral 0 over S with every member of $\mathcal{A}$ and also with uniform limits on S of members of $\mathcal{A}$. This example shows the need for some hypothesis like "closed under complex conjugation" in Theorem 2.58b.

Corollary 2.59. If (S, d) is a compact metric space, then $C(S)$ is separable as a metric space.

PROOF. It is enough to consider $C(S, \mathbb{C})$, since $C(S, \mathbb{R})$ is a metric subspace of $C(S, \mathbb{C})$. Being compact metric, S is separable by Proposition 2.33. Let $\mathcal{B}$ be a countable base of S. The number of pairs (U, V) of members of $\mathcal{B}$ such that $U^{\text{cl}} \subseteq V$ is countable. By Proposition 2.30e, there exists a continuous function $f_{UV} : S \to \mathbb{R}$ such that f_{UV} is 1 on U^{cl} and f_{UV} is 0 on V^c. Let us show that the system of functions f_{UV} separates points of S.

If x_1 and x_2 are given, the $\mathbf{T}_1$ property of S (Proposition 2.30a), when combined with Proposition 2.31, gives us a member V of $\mathcal{B}$ such that x_1 is in V and $V \subseteq \{x_2\}^c$. Since the set V^c is closed and does not contain x_1, the property that S is regular (Proposition 2.30c) gives us disjoint open sets U_1 and V_1 with $x_1 \in U_1$ and $V^c \subseteq V_1$. The latter condition means that $V \supseteq V_1^c$. By Proposition 2.31 let U be a basic open set with $x_1 \in U$ and $U \subseteq U_1$. Then we have $x_1 \in U \subseteq U_1 \subseteq U_1^{\text{cl}} \subseteq V_1^c \subseteq V$ and hence also $x_1 \in U \subseteq U^{\text{cl}} \subseteq V$. The function f_{UV} is therefore 1 on x_1 and 0 on x_2, and the system of functions f_{UV} separates points.

The set of all finite products of functions f_{UV} and the constant function 1 is countable, and so is the set D of linear combinations of all these functions with coefficients of the form $q_1 + iq_2$ with q_1 and q_2 rational. The claim is that this countable set D is dense in $C(S, \mathbb{C})$. The closure of D certainly contains the algebra $\mathcal{A}$ of all complex linear combinations of the function 1 and arbitrary finite products of functions f_{UV}, and $\mathcal{A}$ is closed under complex conjugation. By the Stone–Weierstrass Theorem (Theorem 2.58), $\mathcal{A}^{\text{cl}} = C(S, \mathbb{C})$. Since D^{cl} contains $\mathcal{A}$, we have $C(S, \mathbb{C}) = \mathcal{A}^{\text{cl}} \subseteq (D^{\text{cl}})^{\text{cl}} = D^{\text{cl}}$. In other words, D is dense. □

11. Completion

If (X, d) and (Y, ρ) are two metric spaces, an **isometry** of X into Y is a function $\varphi : X \to Y$ that preserves distances: $\rho(\varphi(x_1), \varphi(x_2)) = d(x_1, x_2)$ for all x_1 and x_2 in X. For example, a rotation $(x, y) \mapsto (x\cos\theta - y\sin\theta,\ x\sin\theta + y\cos\theta)$ is an isometry of $\mathbb{R}^2$ with itself. An isometry is necessarily continuous (with $\delta = \epsilon$). However, an isometry need not have the whole range as image. For example, the map $x \mapsto (x, 0)$ of $\mathbb{R}^1$ into $\mathbb{R}^2$ is an isometry that is not onto $\mathbb{R}^2$. In the case that there exists an isometry of X onto Y, we say that X and Y are **isometric**.

Theorem 2.60. If (X, d) is a metric space, then there exist a complete metric space (X^*, Δ) and an isometry $\varphi : X \to X^*$ such that the image of X in X^* is dense.

REMARK. It is observed in Problems 25–26 at the end of the chapter that (X^*, Δ) and $\varphi : X \to X^*$ are essentially unique. The metric space (X^*, Δ) is called a **completion** of (X, d), or sometimes "the" completion because of the essential uniqueness. There is more than one construction of X^*, and the proof below will use a construction by Cauchy sequences that is immediately suggested if X is the set of rationals and X^* is the set of reals.

PROOF. Let Cauchy(X) be the set of all Cauchy sequences in X. Define a relation $\sim$ on Cauchy(X) as follows: if $\{p_n\}$ and $\{q_n\}$ are in Cauchy(X), then $\{p_n\} \sim \{q_n\}$ means $\lim d(p_n, q_n) = 0$.

Let us prove that $\sim$ is an equivalence relation. It is reflexive, i.e., has $\{p_n\} \sim \{p_n\}$, because $d(p_n, p_n) = 0$ for all n. It is symmetric, i.e., has the property that $\{p_n\} \sim \{q_n\}$ implies $\{q_n\} \sim \{p_n\}$, because $d(p_n, q_n) = d(q_n, p_n)$. It is transitive, i.e., has the property that $\{p_n\} \sim \{q_n\}$ and $\{q_n\} \sim \{r_n\}$ together imply $\{p_n\} \sim \{r_n\}$, because

$$0 \le d(p_n, r_n) \le d(p_n, q_n) + d(q_n, r_n)$$

and each term on the right side is tending to 0. Thus $\sim$ is an equivalence relation.

Let X^* be the set of equivalence classes. If P and Q are two equivalence classes, we set

$$\Delta(P, Q) = \lim d(p_n, q_n), \tag{$*$}$$

where $\{p_n\}$ is a member of the class P and $\{q_n\}$ is a member of the class Q. We have to prove that the limit in $(*)$ exists in $\mathbb{R}$ and then that the limit is independent of the choice of representatives of P and Q.

For the existence of the limit $(*)$, it is enough to prove that the sequence $\{d(p_n, q_n)\}$ is Cauchy. The triangle inequality gives

$$d(p_n, q_n) \le d(p_n, p_m) + d(p_m, q_m) + d(q_m, q_n)$$

and hence $d(p_n, q_n) - d(p_m, q_m) \le d(p_n, p_m) + d(q_m, q_n)$. Reversing the roles of m and n, we obtain

$$|d(p_n, q_n) - d(p_m, q_m)| \le d(p_n, p_m) + d(q_m, q_n).$$

The two terms on the right side tend to 0, since $\{p_k\}$ and $\{q_k\}$ are Cauchy, and hence $\{d(p_n, q_n)\}$ is Cauchy. Thus the limit $(*)$ exists.

We have also to show that the limit $(*)$ is independent of the choice of representatives. Let $\{p_n\}$ and $\{p'_n\}$ be in P, and let $\{q_n\}$ and $\{q'_n\}$ be in Q. Then

$$d(p_n, q_n) \leq d(p_n, p'_n) + d(p'_n, q'_n) + d(q'_n, q_n).$$

Since the first and third terms on the right side tend to 0 and the other terms in the inequality have limits, we obtain $\lim_n d(p_n, q_n) \leq \lim_n d(p'_n, q'_n)$. Reversing the roles of the primed and unprimed symbols, we obtain $\lim d(p'_n, q'_n) \leq \lim d(p_n, q_n)$. Therefore $\lim d(p_n, q_n) = \lim d(p'_n, q'_n)$, and $\Delta(P, Q)$ is well defined.

Let us see that (X^*, Δ) is a metric space. Certainly $\Delta(P, P) = 0$ and $\Delta(P, Q) = \Delta(Q, P)$. To prove the triangle inequality

$$\Delta(P, Q) \leq \Delta(P, R) + \Delta(R, Q), \tag{$**$}$$

let $\{p_n\}$ be in P, $\{q_n\}$ be in Q, and $\{r_n\}$ be in R. Since

$$d(p_n, q_n) \leq d(p_n, r_n) + d(r_n, q_n),$$

we obtain $(**)$ by passing to the limit. Finally if two unequal classes P and Q are given, and if $\{p_n\}$ and $\{q_n\}$ are representatives, then $\lim d(p_n, q_n) \neq 0$ by definition of $\sim$. Therefore $\Delta(P, Q) > 0$. Thus (X^*, Δ) is a metric space.

Now we can define the isometry $\varphi : X \to X^*$. If x is in X, then $\varphi(x)$ is the equivalence class of the constant sequence $\{p_n\}$ in which $p_n = x$ for all n. To see that φ is an isometry, let x and y be in X, let $p_n = x$ for all n, and let $q_n = y$ for all n. Then $\Delta(\varphi(x), \varphi(y)) = \lim d(p_n, q_n) = \lim d(x, y) = d(x, y)$, and φ is an isometry.

Let us prove that $\varphi(X)$ is dense in X^*. In fact, if P is in X^* and $\{p_n\}$ is a representative, we show that $\varphi(p_n) \to P$. If $\varphi(p_n) = P$ for all sufficiently large n, then P is in $\varphi(X)$; otherwise this limit relation will exhibit P as a limit point of $\varphi(X)$, and we can conclude that P is in $\varphi(X)^{\mathrm{cl}}$ in any case. In other words, $\varphi(p_n) \to P$ implies that $\varphi(X)$ is dense. To prove that we actually do have $\varphi(p_n) \to P$, let $\epsilon > 0$ be given. Choose N such that $k \geq m \geq N$ implies $d(p_m, p_k) < \epsilon$. Then $\Delta(\varphi(p_m), P) = \lim_k d(p_m, p_k) \leq \epsilon$ for $m \geq N$. Hence $\lim_m \Delta(\varphi(p_m), P) = 0$ as required.

Finally let us prove that X^* is complete by showing directly that any Cauchy sequence $\{P_n\}$ converges. Since $\varphi(X)$ is dense in X^*, we can choose $x_n \in X$ with $\Delta(\varphi(x_n), P_n) < 1/n$. First let us prove that $\{x_n\}$ is Cauchy in X. Let $\epsilon > 0$ be given, and choose N large enough so that $\Delta(P_n, P_{n'}) < \epsilon/3$ when n and n' are

$\geq N$. Possibly by taking N still larger, we may assume that $1/N < \epsilon/3$. Then whenever n and n' are $\geq N$, we have

$$\begin{aligned} d(x_n, x_{n'}) &= \Delta(\varphi(x_n), \varphi(x_{n'})) \\ &\leq \Delta(\varphi(x_n), P_n) + \Delta(P_n, P_{n'}) + \Delta(P_{n'}, \varphi(x_{n'})) \\ &\leq \tfrac{1}{n} + \tfrac{\epsilon}{3} + \tfrac{1}{n'} \leq \tfrac{1}{N} + \tfrac{\epsilon}{3} + \tfrac{1}{N} < \tfrac{\epsilon}{3} + \tfrac{\epsilon}{3} + \tfrac{\epsilon}{3} = \epsilon. \end{aligned}$$

Thus $\{x_n\}$ is Cauchy in X. Let $P \in X^*$ be the equivalence class to which $\{x_r\}$ belongs. We prove completeness by showing that $P_n \to P$. Let $\epsilon > 0$ be given, and choose N large enough so that $r \geq n \geq N$ implies $d(x_n, x_r) < \epsilon/2$. Possibly by taking N still larger, we may assume that $\frac{1}{N} < \frac{\epsilon}{2}$. Then $r \geq n \geq N$ implies

$$\Delta(P_n, P) \leq \Delta(P_n, \varphi(x_n)) + \Delta(\varphi(x_n), P) < \tfrac{1}{n} + \lim_r d(x_n, x_r) < \tfrac{\epsilon}{2} + \tfrac{\epsilon}{2} = \epsilon.$$

Thus $P_n \to P$. Hence every Cauchy sequence in X^* converges, and X^* is complete. $\square$

An important application of Theorem 2.60 for algebraic number theory is to the construction of the p-adic numbers, p being prime. The metric space that is completed is the set of rationals with a certain nonstandard metric. This application appears in Problems 27–31 at the end of this chapter.

12. Problems

1. As in Example 9 of Section 1, let S be a nonempty set, fix an integer $n > 0$, and let X be the set of n-tuples of members of S. For n-tuples $x = (x_1, \dots, x_n)$ and $y = (y_1, \dots, y_n)$, define $d(x, y) = \#\{j \mid x_j \neq y_j\}$, the number of components in which x and y differ. Prove that d satisfies the triangle inequality, so that (X, d) is a metric space.
2. Prove that a separable metric space is the disjoint union of a countable open set and a closed set in which every point is a limit point.
3. Give an example of a function $f : [0, 1] \to \mathbb{R}$ for which the graph of f, given by $\big\{(x, f(x)) \bigm| 0 \leq x \leq 1\big\}$, is a closed subset of $\mathbb{R}^2$ and yet f is not continuous.
4. If A is a dense subset of a metric space (X, d) and U is open in X, prove that $U \subseteq (A \cap U)^{\mathrm{cl}}$.
5. Let (X, d) be a metric space, let U be an open set, and let $E_1 \supseteq E_2 \supseteq \cdots$ be a decreasing sequence of closed bounded sets with $\bigcap_{n=1}^{\infty} E_n \subseteq U$.
   (a) For X equal to $\mathbb{R}^n$, show that $E_N \subseteq U$ for some N.
   (b) For X equal to the subspace $\mathbb{Q}$ of rationals in $\mathbb{R}^1$, give an example to show that $E_N \subseteq U$ can fail for every N.

6. Let $F : X \times Y \to Z$ be a function from the product of two metric spaces into a metric space.
   (a) Suppose that $(x, y) \mapsto F(x, y)$ is continuous and that Y is compact. Prove that $F(x, \cdot)$ tends to $F(x_0, \cdot)$ uniformly on Y as x tends to x_0.
   (b) Conversely suppose $\mapsto F(x, y)$ is continuous except possibly at points $(x, y) = (x_0, y)$, and suppose that $F(x, \cdot) \to F(x_0, \cdot)$ uniformly. Prove that F is continuous everywhere.
7. Give an example of a continuous function between two metric spaces that fails to carry some Cauchy sequence to a Cauchy sequence.
8. **(Contraction mapping principle)** Let (X, d) be a complete metric space, let r be a number with $0 \leq r < 1$, and let $f : X \to X$ be a **contraction mapping**, i.e., a function such that $d(f(x), f(y)) \leq r d(x, y)$ for all x and y in X. Prove that there exists a unique x_0 in X such that $f(x_0) = x_0$.
9. Prove that a countable complete metric space has an isolated point.
10. A metric space (X, d) is called **locally connected** if each point has arbitrarily small open neighborhoods that are connected. Let C be a Cantor set in $[0, 1]$, as described in Section 9, and let $X \subset \mathbb{R}^2$ be the union of the three sets $C \times [0, 1]$, $[0, 1] \times \{0\}$, and $[0, 1] \times \{1\}$. Prove that X is compact and connected but is not locally connected.

Problems 11–13 concern the relationship between connected and pathwise connected. It was observed in Section 8 that pathwise connected implies connected. A metric space is called **locally pathwise connected** if each point has arbitrarily small open neighborhoods that are pathwise connected.

11. Prove that a metric space (X, d) that is connected and locally pathwise connected is pathwise connected.
12. Deduce from the previous problem that for an open subset of $\mathbb{R}^n$, connected implies pathwise connected.
13. Prove that any open subset of $\mathbb{R}^1$ is uniquely the disjoint union of open intervals.

Problems 14–17 concern almost periodic functions. Let $f : \mathbb{R}^1 \to \mathbb{C}$ be a bounded uniformly continuous function. If $\epsilon > 0$, an ϵ **almost period** for f is a number t such that $|f(x + t) - f(x)| \leq \epsilon$ for all real x. A subset E of $\mathbb{R}^1$ is called **relatively dense** if there is some $L > 0$ such that any interval of length $\geq L$ contains a member of E. The function f is **Bohr almost periodic** if for every $\epsilon > 0$, its set of ϵ almost periods is relatively dense. The function f is **Bochner almost periodic** if every sequence of translates $\{f_{t_n}\}$, where $f_t(x) = f(x + t)$, has a uniformly convergent subsequence. Any function $x \mapsto e^{icx}$ with c real is an example.

14. As usual, let $B(\mathbb{R}^1, \mathbb{C})$ be the metric space of bounded complex-valued functions on $\mathbb{R}^1$ in the uniform metric. Show that the subspace of bounded uniformly continuous functions is closed, hence complete.

15. Show that a bounded uniformly continuous function $f : \mathbb{R}^1 \to \mathbb{C}$ is Bohr almost periodic if and only if the set $\{f_t \mid t \in \mathbb{R}^1\}$ is totally bounded in $B(\mathbb{R}^1, \mathbb{C})$.

16. Prove that a bounded uniformly continuous function $f : \mathbb{R}^1 \to \mathbb{C}$ is Bohr almost periodic if and only if it is Bochner almost periodic. Thus the names Bohr and Bochner can be dropped.

17. Prove that the set of almost periodic functions on $\mathbb{R}^1$ is an algebra closed under complex conjugation and containing the constants. Prove also that it is closed under uniform limits.

Problems 18–20 concern the special case whose proof precedes that of the Stone–Weierstrass Theorem (Theorem 2.58). In the text in Section 10, this preliminary special case was the function $|x|$ on $[-1, 1]$, and it was handled in two ways—in Section I.8 by the binomial expansion and Abel's Theorem and in Section I.9 as a special case of the Weierstrass Approximation Theorem. The problems in the present group handle an alternative preliminary special case, the function $\sqrt{x}$ on $[0, 1]$. This is just as good because $|x| = \sqrt{x^2}$.

18. **(Dini's Theorem)** Let X be a compact metric space. Suppose that $f_n : X \to \mathbb{R}$ is continuous, that $f_1 \le f_2 \le f_3 \le \cdots$, and that $f(x) = \lim f_n(x)$ is continuous and is nowhere $+\infty$. Use the defining property of compactness to prove that f_n converges to f uniformly on X.

19. Define a sequence of polynomial functions $P_n : [0, 1] \to \mathbb{R}$ by $P_0(x) = 0$ and $P_{n+1}(x) = P_n(x) + \frac{1}{2}(x - P_n(x)^2)$. Prove that $0 = P_0 \le P_1 \le P_2 \le \cdots \le \sqrt{x} \le 1$ and that $\lim_n P_n(x) = \sqrt{x}$ for all x in $[0, 1]$.

20. Combine the previous two problems to prove that $\sqrt{x}$ is the uniform limit of polynomial functions on $[0, 1]$.

Problems 21–24 concern the effect of removing from the Stone–Weierstrass Theorem (Theorem 2.58) the hypothesis that the given algebra contains the constants. Let (S, d) be a compact metric space, and let $\mathcal{A}$ be a subalgebra of $C(S, \mathbb{R})$ that separates points. There can be no pair of points $\{x, y\}$ such that all members of $\mathcal{A}$ vanish at x and y.

21. If for each $s \in S$, there is some member of $\mathcal{A}$ that is nonzero at s, prove in the following way that $\mathcal{A}$ is still dense in $C(S, \mathbb{R})$: Observe that the only place in the proof of Theorem 2.58a that the presence of constant functions is used is in the construction of the function h_y in the third paragraph. Show that a function h_y still exists in $\mathcal{A}$ with $h_y(x) = f(x)$ and $h_y(y) = f(y)$ under the weaker hypothesis that for each $s \in S$, there is some member of $\mathcal{A}$ that is nonzero at s.

22. Suppose that the members of $\mathcal{A}$ all vanish at some s_0 in S. Let $\mathcal{B} = \mathcal{A} + \mathbb{R}1$, so that Theorem 2.58a applies to $\mathcal{B}$. Use the linear function $L : C(S, \mathbb{R}) \to \mathbb{R}$ given by $L(f) = f(s_0)$, together with the fact that $\mathcal{B}^{\mathrm{cl}} = C(S, \mathbb{R})$, to prove that $\mathcal{A}$ is uniformly dense in the subalgebra of all members of $C(S, \mathbb{R})$ that vanish at s_0.

23. Adapt the above arguments to prove corresponding results about the algebra $C(S, \mathbb{C})$ of complex-valued continuous functions.

24. Let $C_0([0, +\infty), \mathbb{R})$ be the algebra of continuous functions from $[0, +\infty)$ into $\mathbb{R}$ that have limit 0 at $+\infty$.
    (a) Prove that the set of all finite linear combinations of functions e^{-nx} for positive integers n is dense in $C_0([0, +\infty), \mathbb{R})$.
    (b) Suppose that f is in $C_0([0, +\infty), \mathbb{R})$, that $f(x) = 0$ for $x \geq b$, and that $\int_0^b f(x)e^{-nx}\,dx = 0$ for all integers $n > 0$. Prove that f is the 0 function.

Problems 25–26 concern completions of a metric space. They use the notation of Theorem 2.60. The first problem says that the completion is essentially unique, and the second problem addresses the question of what happens if the original space is already complete; in particular it shows that the completion of the completion is the completion.

25. Suppose that (X, d) is a metric space, that (X_1^*, Δ_1) and (X_2^*, Δ_2) are complete metric spaces, and that $\varphi_1 : X \to X_1^*$ and $\varphi_2 : X \to X_2^*$ are isometries such that $\varphi_1(X)$ is dense in X_1^* and $\varphi_2(X)$ is dense in X_2^*. Prove that there exists a unique isometry ψ of X_1^* onto X_2^* such that $\varphi_2 = \psi \circ \varphi_1$.

26. Prove that a metric space X is complete if and only if $X^* = X$, i.e., if and only if the standard isometry φ of X into its completion X^* is onto.

Problems 27–31 concern the field $\mathbb{Q}_p$ of p-**adic numbers**. The problems assume knowledge of unique factorization for the integers; the last problem in addition assumes knowledge of rings, ideals, and quotient rings. Let $\mathbb{Q}$ be the set of rational numbers with their usual arithmetic, and fix a prime number p. Each nonzero rational number r can be written, via unique factorization of integers, as $r = mp^k/n$ with p not dividing m or n and with k a well-defined integer (positive, negative, or zero). Define $|r|_p = p^{-k}$. For $r = 0$, define $|0|_p = 0$. The function $|\cdot|_p$ plays a role in the relationship between $\mathbb{Q}$ and $\mathbb{Q}_p$ similar to the role played by absolute value in the relationship between $\mathbb{Q}$ and $\mathbb{R}$.

27. Prove that $|\cdot|_p$ on $\mathbb{Q}$ satisfies (i) $|r|_p \geq 0$ with equality if and only if $r = 0$, (ii) $|-r|_p = |r|_p$, (iii) $|rs|_p = |r|_p|s|_p$, and (iv) $|r+s|_p \leq \max\{|r|_p, |s|_p\}$. Property (iv) is called the **ultrametric inequality**.

28. Show that $(\mathbb{Q}, d)$ is a metric space under the definition $d(r, s) = |r - s|_p$.

29. Let $(\mathbb{Q}_p, d)$ be the completion of the metric space $(\mathbb{Q}, d)$. Since $|r|_p$ can be recovered from the metric by $|r|_p = d(r, 0)$, the function $|\cdot|_p$ extends to a continuous function $|\cdot|_p : \mathbb{Q}_p \to \mathbb{R}$.
    (a) Using Proposition 2.47, show that addition, as a function from $\mathbb{Q} \times \mathbb{Q}$ to $\mathbb{Q}_p$, extends to a continuous function from $\mathbb{Q}_p \times \mathbb{Q}_p$ to $\mathbb{Q}_p$. Argue similarly that the operation of passing to the negative, as a function from $\mathbb{Q}$ to $\mathbb{Q}_p$, extends to a continuous function from $\mathbb{Q}_p$ to $\mathbb{Q}_p$. Then prove that $\mathbb{Q}_p$ is an abelian group under addition.

(b) Show that multiplication, as a function from $\mathbb{Q} \times \mathbb{Q}$ to $\mathbb{Q}_p$, extends to a continuous function from $\mathbb{Q}_p \times \mathbb{Q}_p$ to $\mathbb{Q}_p$. (This part is subtler than (a) because multiplication is not uniformly continuous as a function of two variables.)

(c) Let $\mathbb{Q}^\times = \mathbb{Q} - \{0\}$ and $\mathbb{Q}_p^\times = \mathbb{Q}_p - \{0\}$. Show that the operation of taking the reciprocal, as a function from $\mathbb{Q}^\times$ to $\mathbb{Q}_p^\times$, extends to a continuous function from $\mathbb{Q}_p^\times$ to itself. Then prove that $\mathbb{Q}_p^\times$ is an abelian group under multiplication.

(d) Complete the proof that $\mathbb{Q}_p$ is a field by establishing the distributive law $t(r+s) = tr + ts$ within $\mathbb{Q}_p$.

30. (a) Prove that the subset $\{t \in \mathbb{Q}_p \mid |t|_p \le 1\}$ of $\mathbb{Q}_p$ is totally bounded.
    (b) Prove that a subset of $\mathbb{Q}_p$ is compact if and only if it is closed and bounded.

31. Prove that the subset $\mathbb{Z}_p$ of $\mathbb{Q}_p$ with $|x|_p \le 1$ is a commutative ring with identity, that the subset P with $|x|_p \le p^{-1}$ is an ideal in $\mathbb{Z}_p$, and that the quotient $\mathbb{Z}_p/P$ is a field of p elements.