THE UNIVERSITY OF MELBOURNE

Department of Mathematics and Statistics

620-311

Metric Spaces

Lecture Notes

Semester 1, 2007

Notes by: Kris Wysocki

Revised by: Hyam Rubinstein

620-311.

Metric and Topological Spaces

Contents

1 Introduction

The ideas of limit and continuity which we encounter in Euclidean spaces occur in various other contexts, e.g in function spaces. Point set topology is the study of limits and continuity in a general setting. The notion of limit is based on the idea of nearness. These concepts are easier to understand when the notion of nearness is defined in terms of some *distance* function. The corresponding spaces are called metric spaces. These are introduced in Chapter 2 and applications to function spaces are discussed early. The desirability of finding limits leads to the notion of completeness and compactness. As we go on, we find that many of the arguments do not really need the notion of distance. This leads to the concept of topological spaces which are discussed from Chapter 6 onward. The idea of compactness is discussed in a general setting in Chapter 7 and the notion of connectedness (which is related to the intermediate value theorem) is discussed in Chapter 8. Under mild assumptions we can study abstract toplogical spaces by constructing continuous functions to the real line; the results known as Urysohn and Tietze's theorem are discussed in Chapter 10. The concepts of completeness and compactness come again in the guise of the important Ascoli-Arzela theorem discussed in Chapter 9. The necessary preliminary material is collected in Chapter 11. This set of notes is only a brief introduction to the subject and we refer to the books by Munkres [M] or Patty [P] for more comprehensive treatment.

2 Metric Spaces

Basic Concepts

Consider a non-empty set X, whose elements will be refered to as **points**. A distance function or a **metric** on a set X is a function $d : X \times X \to \mathbb{R}$ which assigns to each pair of points x and y in X a real number $d(x, y)$ having the following properties:

M1. $d(x, y) \geq 0$ and $d(x, y) = 0$ if and only if $x = y$;

M2. $d(x, y) = d(y, x)$ for all $x, y \in X$;

M3. $d(x, z) \leqslant d(x, y) + d(y, z)$ for all x, y and $z \in X$.

Definition 2.1. *A* **metric space** *is a pair* (X, d) *where* d *is a metric defined on the set* X*.*

The axiom M2 says that a metric is symmetric, and the axiom M3 is called the **triangle inequality** since it reflects the geometrical fact that the length of one side of a triangle is less than or equal to the sum of the lengths of the other two sides.

Examples.
(1) Let X be any nonempty set. The **discrete metric** on X is defined by

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

(2) If (X, d) is a metric space and Y is a non-empty subset of X, then

$$d_Y(x, y) = d(x, y) \qquad \text{for all } x, y \in Y$$

is a metric on Y. The pair (Y, d_Y) is called a **metric subspace** of (X, d). (We will usually refer to Y as a subspace of X, rather than (Y, d_Y) as a subspace of (X, d).)

(3) The **standard metric** in $\mathbb{C}$ is defined by

$$d(z, w) = |z - w|, \quad z, w \in \mathbb{C}.$$

The set of real numbers $\mathbb{R}$ inherits the metric from $\mathbb{C}$, namely $d(x, y) = |x - y|$, $x, y \in \mathbb{R}$. Identifying $\mathbb{C} = \{x_1 + ix_2 | \; x_1, x_2 \in \mathbb{R}\}$ with $\mathbb{R}^2 = \{(x_1, x_2) | \; x_1, x_2 \in \mathbb{R}\}$, the standard metric takes the form

$$d((x_1, x_2), (y_1, y_2)) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}.$$

(4) Cartesian product of a finite number of metric spaces. Consider a finite collection of metric spaces (X_i, d_i) , $1 \leqslant i \leqslant n$, and let $X = \prod_{i=1}^n X_i = X_1 \times \ldots \times X_n$. For $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n) \in X$, set

$$d(x, y) = \sum_{i=1}^{n} d_i(x_i, y_i).$$

Then d is a metric on X. The pair (X, d) defined above is called a **metric product** (or just a product) of (X_i, d_i), $1 \leqslant i \leqslant n$, and the metric d is called a **product metric**. (Other metrics are also used on $\prod_{i=1}^n X_i$).

Norms and normed vector spaces

We next define the class of metric spaces which are the most interesting in analysis. Let X be a vector space over $\mathbb{R}$ (or $\mathbb{C}$).

Definition 2.2. *A* **norm** *is a function* $\|\cdot\| : X \to \mathbb{R}$ *having the following properties:*

N1. $\|x\| \geq 0$ *and* $\|x\| = 0$ *if and only if* $x = 0$.

N2. $\|\alpha x\| = |\alpha| \cdot \|x\|$ *for all* $x \in X$ *and* $\alpha \in \mathbb{R}$ *(or* $\mathbb{C}$*).*

N3. $\|x + y\| \leqslant \|x\| + \|y\|$ *for all* $x, y \in X$.

The pair $(X, \|\cdot\|)$ *is called a* **normed vector space**.

Proposition 2.3. *Let* X *be a normed space. Then*

$$d(x, y) = \|x - y\|, \quad x, y \in X,$$

defines a metric on X.

Proof. The axioms M1 and M2 are clear. If x, y and $z \in X$, then, in view of N3,

$$\begin{aligned} d(x, z) &= \|x - z\| = \|(x - y) + (y - z)\| \\ &\leqslant \|x - y\| + \|y - z\| = d(x, y) + d(y, z), \end{aligned}$$

and so the triangle inequality follows. ∎

Examples of normed spaces

Example. (1) Euclidean Space. For $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, define

$$\|x\| = \left[\sum_{i=1}^{n} x_i^2\right]^{1/2}.$$

Clearly, N1 and N2 are satisfied. To see that N3 holds we need the following lemma.

Lemma 2.4. Cauchy inequality. *If $x, y \in \mathbb{R}^n$, then*

$$\left|\sum_{i=1}^{n} x_i y_i\right| \leqslant \left[\sum_{i=1}^{n} |x_i|^2\right]^{1/2} \cdot \left[\sum_{i=1}^{n} |y_i|^2\right]^{1/2}.$$

Proof.

$$\begin{aligned}
0 \leqslant \sum_{i,j=1}^{n} (x_i y_j - x_j y_i)^2 &= \sum_{i,j=1}^{n} (x_i^2 y_j^2 - 2 x_i x_j y_i y_j + x_j^2 y_i^2) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} x_i^2 y_j^2 + \sum_{i=1}^{n}\sum_{j=1}^{n} x_j^2 y_i^2 - 2\sum_{i=1}^{n}\sum_{j=1}^{n} x_i x_j y_i y_j \\
&= \sum_{i=1}^{n} \|y\|^2 x_i^2 + \sum_{j=1}^{n} \|x\|^2 y_j^2 - 2\left[\sum_{i=1}^{n} x_i y_i\right]^2 \\
&= 2\|x\|^2 \cdot \|y\|^2 - 2\left[\sum_{i=1}^{n} x_i y_i\right]^2.
\end{aligned}$$

■

As a corollary we have

Corollary 2.5.

$$\|x + y\| \leqslant \|x\| + \|y\| \quad \textit{for all } x, y \in \mathbb{R}^n.$$

Proof. In view of the Cauchy inequality we have

$$\begin{aligned}
\|x + y\|^2 = \sum_{i=1}^{n} |x_i + y_i|^2 &= \sum_{i=1}^{n} x_i^2 + 2\sum_{i=1}^{n} x_i y_i + \sum_{i=1}^{n} y_i^2 \\
&= \|x\|^2 + 2\sum_{i=1}^{n} x_i y_i + \|y\|^2 \leqslant \|x\|^2 + 2\|x\| \cdot \|y\| + \|y\|^2 \\
&= (\|x\| + \|y\|)^2.
\end{aligned}$$

By taking square roots of both sides the desired inequality follows. ■

Consequently,

$$d(x,y) = \|x-y\| = \left[\sum_{i=1}^{n}(x_i-y_i)^2\right]^{1/2}$$

defines a metric on $\mathbb{R}^n$. We shall call this metric the **Euclidean metric** or the **standard metric** on $\mathbb{R}^n$.

Example. (2) Space of bounded functions. Let X be a non-empty set. A function $f : X \to \mathbb{R}$ is **bounded** if $|f(x)| \leqslant M$ for some M and all $x \in X$. Introduce the set $B(X) = B(X,\mathbb{R})$ of all bounded functions from X to $\mathbb{R}$, and define

$$\|f\| = \sup\{|f(x)| : x \in X\}.$$

Then $\|\cdot\|$ is a norm on $B(X)$, and, in view of Proposition 2.3, this norm defines a metric on $B(X)$ by

$$d(f,g) = \|f-g\| = \sup\{|f(x)-g(x)| : \ x \in X\},$$

for $f, g \in B(X)$.

Example. (3) Let X be the set of all continuous functions $f : [a,b] \to \mathbb{R}$. For any $f \in X$, we set

$$\|f\| = \int_a^b |f(x)|dx.$$

Then $\|\cdot\|$ defines a norm on X which induces a metric on X by

$$d(f,g) = \int_a^b |f(x)-g(x)|dx, \quad f,g \in X.$$

Balls and diameter

If x is an element of X and r is a positive real number, then we write $B(x,r)$ for the **open ball** in X with centre at x and radius $r > 0$, defined by

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

Similarly, the **closed ball** with centre at x and radius $r > 0$ is denoted by $\overline{B}(x,r)$ and defined by

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

The **diameter** of a non-empty subset A of X is defined as

$$\operatorname{diam} A = \sup\{d(x, y) \mid x, y \in A\}.$$

Clearly, if $A \subset B$, then $\operatorname{diam} A \leqslant \operatorname{diam} B$. A subset $A \subset X$ is **bounded** if its diameter is finite, that is, $\operatorname{diam} A < \infty$. This is equivalent to saying that A is contained in some ball.
We can also define the concept of distance between subsets of a metric space. If A and B are non-empty subsets of X, then the **distance between A and B** is defined as

$$d(A, B) = \inf\{d(x, y) \mid x \in A,\ y \in B\}.$$

In particular, if $x \in X$, then the distance between sets $\{x\}$ and A

$$d(x, A) := d(\{x\}, A) = \inf\{d(x, y) \mid y \in A\}$$

is called the **distance between x and A**.

Sequences and Convergence

Convergence of a sequence in a metric space is defined as in calculus.

Definition 2.6. *The sequence $\{x_n\}$ is said to* **converge** *to a point x in X, if for every $\varepsilon > 0$ there exists a positive integer k such that*

$$d(x_n, x) < \varepsilon \qquad \text{for all } n \geq k.$$

In this case we write

$$\lim_{n\to\infty} x_n = x \quad \text{or} \quad x_n \to x.$$

The point x is called the **limit** *of $\{x_n\}$.*

The definition can be expressed in terms of the convergence of sequences of real numbers. Namely, a sequence $\{x_n\}$ converges to $x \in X$ if and only if $d(x_n, x) \to 0$ as $n \to \infty$. We are justified in referring to *the* limit because of the following proposition.

Proposition 2.7. *Let $\{x_n\}$ be a sequence in a metric space (X, d).*

- *Then there is at most one point $x \in X$ such that $\{x_n\}$ converges to x.*
- *If $\{x_n\}$ is convergent then it is bounded.*

- *if $\{x_n\}$ converges to x and $\{y_n\}$ converges to y then $d(x_n, y_n)$ converges to $d(x, y)$.*

Proof. Arguing by contradiction we assume that $x_n \to x$ and $x_n \to y$ with $x \neq y$. Then $d(x, y) > 0$. Take $\varepsilon = d(x, y)/2$. Then we find a positive integer k such that

$$d(x_n, x) < \varepsilon \quad \text{and} \quad d(x_n, y) < \varepsilon, \quad \text{for } n \geq k.$$

By the triangle inequality

$$d(x, y) \leqslant d(x, x_n) + d(x_n, y) < \varepsilon + \varepsilon = d(x, y)$$

which gives a contradiction. Hence we conclude that it is impossible for a sequence $\{x_n\}$ to converge to two different points.

Next, if $\{x_n\}$ converges to x, then we know that there is a positive integer k so that

$$d(x_n, x) < \varepsilon \quad \text{for } n \geq k.$$

Then picking $M = max\{|x_1|, \cdots, |x_{k-1}|, |x_k| + \varepsilon\}$, it is easy to see that M is a bound for all numbers in $\{|x_n|\}$. So the sequence is bounded.

Finally, by the 'quadrilateral' inequality,

$$|d(x_n, y_n) - d(x, y)| \leq d(x_n, x) + d(y_n, y) \to 0$$

(To check this, note that either sign for the absolute value gives an inequality which follows by two applications of the triangle inequality). ■

Given a sequence $\{x_n\}$ of points in X, consider a sequence of indices $\{n_k\}$ such that $n_1 < n_2 < n_3 < \cdots$. Then $\{x_{n_k}\}$ is called a **subsequence** of $\{x_n\}$.

Proposition 2.8. *If $X = \prod_{i=1}^n X_i$ is the product of metric spaces (X_i, d_i), $1 \leqslant i \leqslant n$, and $x^m = (x_1^m, x_2^m, \ldots, x_n^m) \in X$, then $x^m \to x = (x_1, \ldots, x_n) \in X$ if and only if $x_i^m \to x_i$ in X_i for $i = 1, \ldots, n$.*

Proof. Recall that we consider X with the metric

$$d(x, y) = \sum_{i=1}^{n} d_i(x_i, y_i)$$

for $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, x_n) \in X$. Observe that

$$d_i(x_i, y_i) \leqslant d(x, y) \leqslant n \cdot \max\{d_i(x_i, y_i) \mid 1 \leqslant i \leqslant n\}, \quad x, y \in X. \tag{1}$$

Let $x^m \to x$, where $x = (x_1, \ldots, x_n)$. Then given $\varepsilon > 0$ there exists $k \in \mathbb{N}$ such that

$$d(x^m, x) < \varepsilon \quad \text{for } m \geq k.$$

By (1) we have

$$d_j(x_j^m, x_j) < \varepsilon \quad \text{for } m \geq k \text{ and } j = 1, \ldots, n.$$

So $x_j^m \to x_j$ as required. Conversely, assume that $x_j^m \to x_j$ for $j = 1, \ldots, n$. Hence for a given $\varepsilon > 0$, there exists $k(j) \in \mathbb{N}$ such that

$$d_j(x_j^m, x_j) < \varepsilon/n \quad \text{for } m \geq k(j).$$

In view of the right hand side inequality in (1) we get

$$d(x^m, x) \leqslant n \max\{d_j(x_j^m, x_j) \mid j = 1, \ldots, n\} < \varepsilon$$

for all $m > k := \max\{k(j) \mid j = 1, \ldots, n\}$. Hence $x_n \to x$ as required. ∎

Definition 2.9. *Two metrics d and d' in X are called* **equivalent** *if*

$$d(x_n, x_0) \to 0 \quad \textit{if and only if} \quad d'(x_n, x_0) \to 0.$$

Examples. (1) Let d be any metric on X. Define

$$d'(x, y) = \frac{d(x, y)}{1 + d(x, y)}, \quad x, y \in X. \tag{2}$$

Then d' is a metric on X (show this!) which is equivalent to d. Indeed, if $d(x_n, x_0) \to 0$, then $d'(x_n, x_0) = \dfrac{d(x_n, x_0)}{1 + d(x_n, x_0)} \to 0$. Conversely, $d(x, y) = \dfrac{d'(x, y)}{1 - d'(x, y)}$. So if $d'(x_n, x_0) \to 0$, then $d(x_n, x_0) \to 0$. Note that with respect to this equivalent metric, the space X is *bounded* since $d'(x, y) < 1$ for all $x, y \in X$.

(2) Consider the product (X, d) of metric spaces (X_i, d_i). Recall that

$$d(x, y) = \sum_{i=1}^{n} d_i(x_i, y_i), \quad x = (x_1, \ldots, x_n),\ y = (y_1, \ldots, y_n) \in X.$$

Set

$$\sigma(x, y) = \max\{d_i(x_i, y_i) |\ 1 \leqslant i \leqslant n\}$$

$$\rho(x, y) = \left[\sum_{i=1}^{n} d_i(x_i, y_i)^2\right]^{1/2}.$$

Then d is equivalent to σ and ρ.

Open and closed sets

Definition 2.10. *Let $A \subset X$. A point $x \in A$ is called an* **interior point** *of A if $B(x,r) \subset A$ for some $r > 0$. The collection of all interior points of a set A is called the* **interior** *of A, and is denoted by A° or int$A$. A set A is called* **open** *if $A = A^\circ$.*

Note that if x is an interior point of $A \subset X$, then it is also an interior point of any set B such that $A \subset B$.

Examples. (1) The empty set $\emptyset$ and the whole space X are open in any metric space X. If X is equipped with the discrete metric d, then any subset of X is open.

(2) The set $\mathbb{Q}$ is not open in $\mathbb{R}$ with the usual metric but it is open in $(\mathbb{R}, d)$, where d is the discrete metric in $\mathbb{R}$. Indeed, if $x \in \mathbb{Q}$ and $r > 0$, then for large $n \in \mathbb{N}$, we have

$$x < x + \frac{\sqrt{2}}{n} < x + r$$

so that $x + \sqrt{2}/n \in B(x,r)$ but $x + \sqrt{2}/n \notin \mathbb{Q}$. In the case of the discrete metric, for every $x \in \mathbb{Q}$, $B(x, 1/2) = \{x\} \subset \mathbb{Q}$, so $\mathbb{Q}$ is open in $(\mathbb{R}, d)$.

Proposition 2.11. *Let X be a metric space. Then*

(a) *Every open ball $B(x,r)$ is an open set.*

(b) *The intersection of a finite collection $A_1, \ldots, A_k$ of open sets is open.*

(c) *The union of any collection of open sets is open.*

Proof. (a) Take any point $y \in B(x,r)$ and set $R := r - d(x,y) > 0$. If $z \in B(y,R)$, then the triangle inequality

$$d(x,z) \leqslant d(x,y) + d(y,z) < d(x,y) + R = d(x,y) + [r - d(x,y)] = r$$

shows that $z \in B(x,r)$, that is, $B(y,R) \subset B(x,r)$ and y is an interior point of $B(x,r)$ as claimed.

(b) If $x \in \bigcap_{1 \leqslant i \leqslant k} A_i$, then there are numbers $r_i > 0$ such that $B(x, r_i) \subset A_i$ for $1 \leqslant i \leqslant k$. Take r to be the smallest of the numbers r_j. Then $B(x,r) \subset \bigcap_{1 \leqslant i \leqslant k} A_i$.

(c) If x belongs to the union of open sets, then it belongs to one of these sets, say to A. Since x is an interior point of A, it is also an interior point of the union. ■

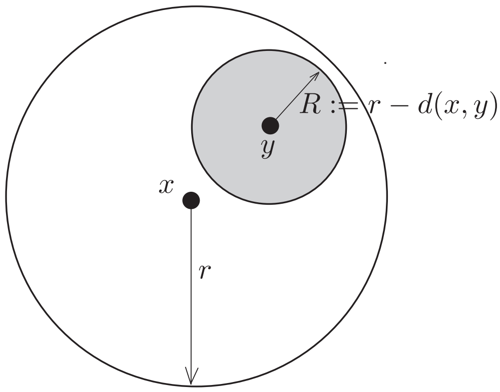


Figure 1: An open ball is an open set

Proposition 2.12. *The interior A° of a set A is an open set and it is the largest open set contained in A.*

Proof. Let $x \in A^\circ$. Since x is an interior point of A, $B(x,r) \subset A$ for some $r > 0$. In view of Proposition 2.11 (a), every point in $B(x,r)$ is an interior point of $B(x,r)$ and therefore of A. Hence every point of $B(x,r)$ is an interior point of A implying that $B(x,r) \subset A^\circ$.
If $B \subset A$ is open, then each point of B is an interior point of A and so belongs to A°. ∎

If $x \in X$, then a set $A \subset X$ is called a **neighbourhood** of x if $x \in A^\circ$.

Definition 2.13. *A point $x \in X$ is called an* **adherent point** *of a subset A if for every $\varepsilon > 0$ there exists a point $y \in A$ such that $y \in B(x,\varepsilon)$. The* **closure** *$\overline{A}$ of A is the set consisting of all the adherent points of A. If $A = \overline{A}$, then A is called* **closed***.*

If x is an adherent point of A, then it is also an adherent point of every larger set B, $A \subset B \subset X$. An adherent point x is either a limit point or an isolated point of a set A. An isolated point satisfies $x \in A$ and there is an open set U with $A \cap U = \{x\}$. (See Definition 2.22) An equivalent definition to adherent point is contained in the following proposition.

Proposition 2.14. *A point x is an adherent point of A if and only if there exists a sequence of points $x_n \in A$ converging to x.*

Proof. Suppose that x is an adherent point of A. Then for every positive integer n, there is $x_n \in A$ such that $d(x_n, x) < 1/n$. But this implies that $x_n \to x$. Conversely, suppose that $\{x_n\} \subset A$ and $x_n \to x$. Let $\varepsilon > 0$. Then $d(x_n, x) < \varepsilon$ for n greater than some k. Hence x is an adherent point of A. ∎

Examples. (1) A closed ball $\overline{B}(x, r)$ is closed set in X. Indeed, let y be an adherent point of $\overline{B}(x, r)$. Then there exists a sequence $\{y_n\}$ in $\overline{B}(x, r)$ such that $y_n \to y$. By the triangle inequality

$$d(y, x) \leqslant d(y, y_n) + d(y_n, x) \leqslant d(y, y_n) + r \to r.$$

Hence $y \in \overline{B}(x, r)$ as claimed. Note that the closure of an open ball $B(x, r)$ does not have to coincide with a closed ball $\overline{B}(x, r)$. Indeed, let X be a set containing at least two points and let d be the discrete metric on X. Then $\overline{B(x, 1)} = \{x\} \neq X = \overline{B}(x, 1)$.

(2) A subset of a metric space may be neither open nor closed. For instance, $[0, 1)$ is neither open nor closed in $\mathbb{R}$. The same is true for $\mathbb{Q}$. On the other hand, a subset may be open and at the same time closed. In a metric space equipped with the discrete metric any subset is both open and closed.

Proposition 2.15. *A subset A of X is open if and only if its complement A^c is closed in X.*

Proof. The proof is based on the following observation. A point x in A is either an adherent point of A^c or an interior point of A, but not both. Hence, if A is open, then its points are not adherent points of A^c which implies that adherent points of A^c belong to A^c, That is, A^c is closed. Conversely, if A^c is closed, then the points of A can't be adherent points of A^c. Hence they are interior points of A and A is open. ∎

Proposition 2.16. *Let X be a metric space. Then*

(a) *The union of a finite collection of closed sets $A_1, \dots, A_k$ is closed.*

(b) *The intersection of arbitrary collection of closed sets is closed.*

Proof. (a) Let $A = \bigcup_{1 \leqslant j \leqslant k} A_j$. Then $A^c = \bigcap_{1 \leqslant j \leqslant k} A_j^c$ is open in view of Proposition 2.11 (b) and Proposition 2.15. Using again Proposition 2.15 the set A is closed.
(b) Let $A = \bigcap_{i \in I} A_i$ be the intersection of closed sets. Then $A^c = \bigcup_{i \in I} A_i^c$ is open in view of Proposition 2.11 (b) and Proposition 2.15. By Proposition 2.15 again the set A is closed. ∎

Proposition 2.17. *The closure $\overline{A}$ of a set A is closed and it is the smallest closed set containing A.*

Proof. If $x \in (\overline{A})^c$, then there is $r > 0$ such that $B(x,r) \subset A^c$. None of the points of $B(x,r)$ belongs to $\overline{A}$ since $B(x,r)$ is open. Thus, $B(x,r) \subset (\overline{A})^c$ and $(\overline{A})^c$ is open, so $\overline{A}$ is closed.
If $A \subset B$ and B is closed, then every adherent point of A is also an adherent point of B, hence belongs to B. Thus, $\overline{A} \subset B$. ∎

Theorem 2.18. *Let Y be a subspace of X.*

(a) $B \subset Y$ is open in Y if and only if $B = Y \cap A$ for some open set A in X.

(b) $B \subset Y$ is closed in Y if and only if $B = Y \cap F$, where F is closed in X.

Proof.
(a) First assume that $B = Y \cap A$ for some open set A in X. Take $x \in B$. Then there exists an open ball $B(x,r)$ in X such that $B(x,r) \subset A$. But then $Y \cap B(x,r) \subset Y \cap A = B$. Since the open ball in the subspace Y with centre $x \in X$ and radius $r > 0$ is the intersection $Y \cap B(x,r)$, the set B is open in Y. Conversely, suppose that B is an open subset of the subspace Y. Then for every $x \in B$ there exists r_x such that the open ball $B(x,r_x) \cap Y$ in Y is contained in B. Then the open subset $A = \bigcup_{x \in B} B(x,r_x)$ of X satisfies $Y \cap A \subset B$. Since any $x \in B$ also belongs to A, $Y \cap A = B$ as required.
(b) A set B is closed in Y if and only if $Y \setminus B$ is open in Y, hence if and only if $Y \setminus B = Y \cap A$ for some open subset A of X. Let $F = X \setminus A$. Then F is closed in X and $B = Y \setminus [Y \cap A] = Y \setminus A = Y \cap [X \setminus A] = Y \cap F$, as required. ∎

Theorem 2.19. *Let X be the product of metric spaces (X_i, d_i), $1 \leqslant i \leqslant m$.*

(a) If A_i is open in X_i, $1 \leqslant i \leqslant m$, then the product $A = \prod_{i=1}^n A_i$ is an open subset of $X = \prod_{i=1}^n X_i$.

(b) If F_i is closed in X_i, $1 \leqslant i \leqslant m$, then $F = \prod_{i=1}^m F_i$ is closed in the product $X = \prod_{i=1}^m X_i$.

Proof.
(a) We prove the result for the product of two metric spaces X_1 and X_2. Let $a = (a_1, a_2) \in A \subset X$. Since A_i is open in X_i, there exists r_i such that an open ball $B(a_i, r_i)$ in X_i is contained in A_i. Let $r = \min\{r_1, r_2\}$. We

claim that $B(a,r) \subset A$. Indeed, if $x = (x_1, x_2) \in B(a,r)$, then $d(a,x) < r$ where $x = (x_1, x_2)$, and since $d_i(a_i, x_i) < d(a,x) < r \leqslant r_i$ we conclude that $x_i \in B(a_i, r_i)$. Hence $x_i \in A_i$, $i = 1, 2$, so that $x \in A$.
(b) The proof follows from Proposition 2.8. ■

Definition 2.20. *The* **boundary** *of A in X, denoted by ∂A, is the set $\overline{A} \cap \overline{X \setminus A}$.*

Hence $x \in \partial A$ if for every $r > 0$ the open ball $B(x,r)$ intersects A and $X \setminus A$ as well. Clearly, the boundary is a closed set as it is an intersection of closed sets.

Example 2.21. Consider $\mathbb{R}$ with the usual metric. Then

$$\partial([0,1]) = \partial((0,1)) = \{0,1\}$$
$$\partial(\mathbb{Q}) = \partial(\mathbb{R} \setminus \mathbb{Q}) = \mathbb{R}.$$

We shall show the last equality. Fix $x \in \mathbb{R}$. If $x \in \mathbb{Q}$, then

$$x \neq x + \frac{1}{n} \in \mathbb{Q} \quad \text{and} \quad x \neq x + \frac{\sqrt{2}}{n} \in \mathbb{Q}^c, \quad \text{for} \quad n \in \mathbb{N}.$$

Since

$$x = \lim_n (x + 1/n) = \lim_n (x + \sqrt{2}/n),$$

it follows that $x \in \overline{\mathbb{Q}} \cap \overline{\mathbb{Q}^c}$. So $\mathbb{Q} \subset \overline{\mathbb{Q}} \cap \overline{\mathbb{Q}^c} = \partial\mathbb{Q}$. If $x \in \mathbb{Q}^c$, then $x + 1/n \in \mathbb{Q}^c$ and there exists a sequence of rational numbers x_n such that

$$x = \lim_n (x + 1/n) = \lim_n x_n.$$

Hence $x \in \overline{\mathbb{Q}} \cap \overline{\mathbb{Q}^c}$ and $\partial(\mathbb{Q}) = \partial(\mathbb{Q}^c) = \overline{\mathbb{Q}} \cap \overline{\mathbb{Q}^c} = \mathbb{R}$.

Definition 2.22. *A point $x \in X$ is called* **isolated** *if $\{x\}$ is open. A space X is called* **discrete** *if all of its points are isolated.*

If x is an isolated point, then for some $\varepsilon > 0$, an open ball $B(x, \varepsilon) \subseteq \{x\}$, that is, $B(x, \varepsilon) = \{x\}$ and if $y \neq x$, then $d(x,y) \geq \varepsilon$. Conversely, if $\inf\{d(x,y) \mid y \neq x\} > 0$, then $\{x\}$ is open. Note also that $\{x\}$ is always closed. For example, consider $\mathbb{N}$ as subspace of $\mathbb{R}$. Then it is discrete. Also the space $J = \{1/n \mid n \in \mathbb{N}\}$ is discrete. In a discrete space any set is open, since it is a union of one-point sets which are open. Also any set is closed being a complement of an open set. Finally, a space is discrete if and only if the only convergent sequences are those which are eventually constant (Prove this!).

Definition 2.23. *A subset A of a metric space is* **dense** *if $\overline{A} = X$.*

Example 2.24. The sets $\mathbb{Q}$ and $\mathbb{Q}^c$ are dense in $\mathbb{R}$ with the usual metric.

Proposition 2.25. *Let X be a metric space and $A \subset X$. Then A is dense if and only if for every non-empty open set U of X, the intersection $U \cap A \neq \emptyset$.*

Definition 2.26. *A subset A of X is called* **nowhere dense** *if $(\overline{A})^\circ = \emptyset$.*

Example 2.27. The sets of all natural numbers $\mathbb{N}$ or all integers $\mathbb{Z}$ are nowhere dense in $\mathbb{R}$ with the usual metric. The set of real numbers $\mathbb{R}$ is nowhere dense in $\mathbb{R}^2$ with the standard metric.

Example 2.28. [**Cantor set**] The Cantor set is a subset of $[0, 1]$ constructed as follows:
Consider the interval $C_0 = [0, 1]$.
Step 1. We divide C_0 into three equal intervals $[0, 1/3]$, $[1/3, 2/3]$ and $[2/3, 1]$ and remove the middle open interval $(1/3, 2/3)$. Denote the remaining intervals by $C_1 = [0, 1/3] \cup [2/3, 1]$. The length of intervals which constitute C_1 is equal to $2/3$.
Step 2. We perform the same operations as in the first step on each of the intervals of C_1. We remove intervals $(1/9, 2/9)$ and $(7/9, 8/9)$. Denote the four remaining intervals by C_2.
Having finished the step $(n - 1)$, we perform the nth step and obtain the set C_n consisting of 2^n intervals.
Each of the sets C_n is closed and bounded, and $C_{n+1} \subset C_n$. The Cantor set is defined as

$$C = \bigcap_{n=1}^{\infty} C_n$$

It is non-empty and since for every n, C_n is closed, C is closed. The set C does not contain any open interval (show this!), and so, C has empty interior. Hence C is nowhere dense.

3 Continuity

The definition of continuity is the $\varepsilon - \delta$ definition of calculus.

Definition 3.1. *Let (X, d) and (Y, ρ) be metric spaces and let $f : X \to Y$ be a function. The function f is said to be* **continuous at the point** $x_0 \in X$ *if the following holds: for every $\varepsilon > 0$, there exists $\delta > 0$ such that for all $x \in X$ if $d(x, x_0) < \delta$, then $\rho(f(x), f(x_0)) < \varepsilon$. The function f is said to be* **continuous** *if it is continuous at each point of X.*

The following proposition rephrases the definition in terms of open balls.

Proposition 3.2. *Let $f : X \to Y$ be a function from a metric space X to another metric space Y and let $x_0 \in X$. Then f is continuous at x_0 if and only if for every $\varepsilon > 0$ there exists $\delta > 0$ such that*

$$f(B(x_0, \delta)) \subset B(f(x_0), \varepsilon).$$

Theorem 3.3. *Let $f : X \to Y$ be a function from a metric space (X, d) to another metric space (Y, ρ) and let $x_0 \in X$. Then f is continuous at x_0 if and only if for every sequence $\{x_n\}$ such that $x_n \to x_0$, $f(x_n) \to f(x_0)$. Also, f is continuous if and only if for every convergent sequence $\{x_n\}$ in X,*

$$\lim_n f(x_n) = f(\lim_n x_n).$$

Proof. Suppose that f is continuous at x_0 and let $x_n \to x_0$. We will prove that $f(x_n) \to f(x_0)$. Let $\varepsilon > 0$ be given. By the definition of continuity at x_0, there exists $\delta > 0$ such that for all $x \in X$,

$$\text{if } d(x, x_0) < \delta, \text{ then } \rho(f(x), f(x_0)) < \varepsilon \ . \tag{3}$$

Since $x_n \to x_0$, there exists an integer k such that for all $n \geq k$,

$$d(x_n, x_0) < \delta. \tag{4}$$

Combining (3) and (4), we get

$$\rho(f(x_n), f(x_0)) < \varepsilon \qquad \text{for all } n \geq k. \tag{5}$$

Hence $f(x_n) \to f(x_0)$ as required. Conversely, arguing by contradiction assume that f is not continuous at x_0. To obtain a contradiction we will construct a sequence $\{x_n\}$ such that $x_n \to x_0$ but the sequence $\{f(x_n)\}$ does not converge to $f(x_0)$. Since f is not continuous at x_0, there is positive $\varepsilon > 0$ such that for all $\delta > 0$ there exists x satisfying $d(x, x_0) < \delta$ but $\rho(f(x), f(x_0)) \geq \varepsilon$. For each n, take $\delta = 1/n$ and then choose x_n so that $d(x_n, x_0) < 1/n$ but $\rho(f(x_n), f(x_0)) \geq \varepsilon$. Hence $x_n \to x_0$ but the sequence $\{f(x_n)\}$ does not converge to $f(x_0)$. The second part of the theorem is an immediate consequence of the first. ∎

Global continuity has a simple formulation in terms of open and closed sets.

Theorem 3.4. *Let f be a function from a metric space (X, d) to (Y, ρ). Then*

(a) *f is continuous if and only if for every open set $U \subset Y$, the preimage $f^{-1}(U)$ of U is open in X. (Recall that the preimage $f^{-1}(U)$ is defined as $f^{-1}(U) = \{x \in X \mid f(x) \in U\}$).*

(b) *f is continuous if and only if for every closed set $F \subset Y$, $f^{-1}(F)$ is closed in X.*

Proof. (a) Suppose first that f is continuous and U is open in Y. If $x \in f^{-1}(U)$, then $f(x) \in U$. Since U is open in Y and $f(x) \in U$, there exists a positive number ε such that $B(f(x), \varepsilon) \subset U$. In view of Proposition 3.2, there exists $\delta > 0$ such that $f(B(x, \delta)) \subset B(f(x), \varepsilon)$. Hence $B(x, \delta) \subset f^{-1}(f(B(x, \delta))) \subset f^{-1}(U)$, so $f^{-1}(U)$ is open in X. Conversely, suppose that $f^{-1}(U)$ is open in X for every open set U in Y. Let $x \in X$ and let $\varepsilon > 0$ be given. Since $B(f(x), \varepsilon)$ is open in Y, the set $f^{-1}(B(f(x), \varepsilon))$ is open in X. Since $x \in f^{-1}(B(f(x), \varepsilon))$, there exists $\delta > 0$ such that $B(x, \delta) \subset f^{-1}(B(f(x), \varepsilon))$. This implies that $f(B(x, \delta)) \subset B(f(x), \varepsilon)$, and in view of Proposition 3.2, f is continuous.
(b) The proof is left as an exercise ■

Theorem 3.5. *Let X, Y and Z be three metric spaces.*

(a) If $f : X \to Y$ and $g : Y \to Z$ are continuous, then the composition $g \circ f$ is continuous.

(b) If $f : X \to Y$ is continuous, and A is a subspace of X, then the restriction of f to A, $f|_A : A \to Y$, is continuous.

Proof. (a) Let $x_n \to x_0$. Since f is continuous at x_0, $f(x_n) \to f(x_0)$. Since g is continuous at $f(x_0)$, $g(f(x_n)) \to g(f(x_0))$. Hence $g \circ f(x_n) \to g \circ f(x_0)$. The second statement follows from the first. Here is another proof of (a). Let U be an open subset of Z. Since g is continuous, $g^{-1}(U)$ is open in Y, and since f is continuous, $f^{-1}(g^{-1}(U))$ is open in X. But $f^{-1}(g^{-1}(U)) = (g \circ f)^{-1}(U)$ and so, $(g \circ f)^{-1}(U)$ is open in X. Hence $g \circ f$ is continuous.
(b) Note that $f|_A = f \circ j$, where $j : A \to X$ is the inclusion, i.e., defined by $j(x) = x$ for $x \in A$. Since for any open set U in X, $j^{-1}(U) = U \cap A$ which is open in A, it follows that j is continuous. So (b) follows from (a). ■

Theorem 3.6. *Let (X, d), (Y_1, ρ_1) and (Y_2, ρ_2) be metric spaces. Let f be a function from X to Y_1 and g a function from X to Y_2.*

- *Define the function h from X to the product $Y_1 \times Y_2$ by*

$$h(x) = (f(x), g(x)), \qquad \text{for } x \in X.$$

  Then h is continuous at x_0 if and only if f and g are continuous at x_0. Thus h is continuous if and only if both functions f and g are continuous.

- *If f and g are continuous functions from (X, d) to $\mathbb{R}$, so are $f + g, f \cdot g, f - g$. Similarly f/g is also continuous so long as $g(x) \neq 0$ for all x.*

- *$d : X \times X \to \mathbb{R}$ is continuous.*

The similar statement about functions from the direct product does not hold in general. Suppose that f is a function from $X \times Y$ to Z. It may happen that f is discontinuous, though the maps $x \mapsto f(x, y)$ for every $y \in Y$ and $y \mapsto f(x, y)$ for every $x \in X$ are all continuous. For example, consider a function from $\mathbb{R} \times \mathbb{R}$ to $\mathbb{R}$ defined by

$$f(x, y) = \begin{cases} \dfrac{xy}{x^2 + y^2} & \text{for } (x, y) \neq (0, 0); \\ 0 & \text{for } (x, y) = (0, 0). \end{cases}$$

The function f is discontinuous at $(0, 0)$ but all the functions $x \mapsto f(x, y)$ and $y \mapsto f(x, y)$ are continuous.

Theorem 3.7 (The pasting lemma). *Let $X = A \cup B$, where A and B are closed subspaces of X. Let $f : A \to Y$ and $g : B \to Y$ be continuous. If $f(x) = g(x)$ for all $x \in A \cap B$, then the function $h : X \to Y$ defined by*

$$h(x) = \begin{cases} f(x) & \text{if } x \in A; \\ g(x) & \text{if } x \in B \end{cases}$$

is continuous.

Proof. Let C be a closed subset of Y. Then $h^{-1}(C) = f^{-1}(C) \cup g^{-1}(C)$. Since f is continuous, $f^{-1}(C)$ is closed in A. But since A is closed $f^{-1}(C)$ is closed in X. Similarly, $g^{-1}(C)$ is closed in X. So $h^{-1}(C)$ is closed in X and the proof is finished. ■

Uniform Continuity and Uniform Convergence

Definition 3.8. *A mapping f from a metric space (X, d) to a metric space (Y, ρ) is said to be* **uniformly continuous** *if for every $\varepsilon > 0$, there exists $\delta > 0$ such that $\rho(f(x), f(y)) < \varepsilon$ for all $x, y \in X$ satisfying $d(x, y) < \delta$.*

Obviously, a uniformly continuous function is continuous.

Examples. (1) The function $f(x) = x/(1 + x^2)$ from $\mathbb{R}$ to $\mathbb{R}$ is uniformly continuous. To see this observe that for any $x < y$, in view of the mean value theorem of calculus, there exists $t \in (0, 1)$ such that

$$|f(x) - f(y)| = |f'(t)| \cdot |x - y| = \left| \frac{1 - t^2}{(1 + t^2)^2} \right| \cdot |x - y| \leqslant |x - y|.$$

since $|f'(t)| \leqslant 1$. Hence for given ε, choose $\delta = \varepsilon$. Then for any x, y such that $d(x, y) = |x - y| < \delta$, we have

$$d(f(x), f(y)) = |f(x) - f(y)| \leqslant |x - y| = d(x, y) < \delta = \varepsilon.$$

So f is uniformly continuous.

(2) The function $f(x) = x^2$ for $x \in \mathbb{R}$ is not uniformly continuous. Indeed, choose $\delta > 0$ and set

$$x = 1/\delta + \delta/2 \quad \text{and} \quad y = 1/\delta.$$

Then $|x - y| = \delta/2 < \delta$ but $|x^2 - y^2| > 1$. However, if we consider the same function on some bounded interval, say $[-a, a]$, then the function is uniformly continuous since if $\delta < \varepsilon/2a$ and $x, y \in [-a, a]$ with $|x - y| < \delta$, then $|x^2 - y^2| = |x - y| \cdot |x + y| < 2a|x - y| < \varepsilon$.

Let (X, d) and (Y, ρ) be metric space. Consider a sequence $\{f_n\}$ of functions $f_n : X \to Y$ and let $f : X \to Y$.

Definition 3.9. *The sequence $\{f_n\}$ is said to* **converge pointwise** *to f if for every $x \in X$ and for every $\varepsilon > 0$, there exists an index $N = N(x, \varepsilon)$ such that*

$$\rho(f_n(x), f(x)) < \varepsilon \qquad \textit{for all } n \geq N.$$

The sequence $\{f_n\}$ is said to **converge uniformly** *to f if for every $\varepsilon > 0$, there exists an index $N = N(\varepsilon)$ such that*

$$\rho(f_n(x), f(x)) < \varepsilon \qquad \textit{for all } n \geq N \textit{ and all } x \in X.$$

Equivalently, $\{f_n\}$ converges uniformly to f on X if

$$\sup\{\rho(f_n(x), f(x)) \mid x \in X\} \to 0.$$

The notion of uniform convergence of a sequence of functions is, in general, more useful than that of pointwise convergence.

Theorem 3.10. *Let $\{f_n\}$ be a sequence of continuous functions from a metric space (X, d) to a metric space (Y, ρ). Suppose that $\{f_n\}$ converges uniformly to f from X to Y. Then f is continuous.*

In words, the uniform limit of a sequence of continuous functions is continuous.

Proof. Let $x_0 \in X$ and let $\varepsilon > 0$ be given. Since $\{f_n\}$ converges uniformly to f, there exists an index N such that for all $n \geq N$ and all $x \in X$,

$$\rho(f_n(x), f(x)) < \varepsilon/3. \tag{6}$$

Since f_N is continuous at x_0, we can choose $\delta > 0$ so that

$$\rho(f_N(x), f_N(x_0)) < \varepsilon/3 \tag{7}$$

for all $d(x, x_0) < \delta$. Now if $d(y, x_0) < \delta$, then

$$\rho(f(y), f(x_0)) \leqslant \rho(f(y), f_N(y)) + \rho(f_N(y), f_N(x_0)) + \rho(f_N(x_0), f(x_0)).$$

Each term of the right-hand side is less than $\varepsilon/3$, the first and the third in view of (6) and the second in view of (7). Thus

$$\rho(f(y), f(x_0)) < \varepsilon$$

for all $d(y, x_0) < \delta$. This proves that f is continuous. ■

4 Complete Spaces

Definition 4.1. *Let (X, d) be a given metric space and let $\{x_n\}$ be a sequence of points of X. We say that $\{x_n\}$ is* **Cauchy** *(or satisfies the* **Cauchy condition***) if for every $\varepsilon > 0$ there exists $k \in \mathbb{N}$ such that*

$$d(x_n, x_m) < \varepsilon \qquad \text{for all } n, m \geq k.$$

Properties of Cauchy sequences are summarized in the following propositions.

Proposition 4.2. *If $\{x_n\}$ is a Cauchy sequence, then $\{x_n\}$ is bounded.*

Proof. Take $\varepsilon = 1$. Since $\{x_n\}$ is Cauchy, there exists an index k such that $d(x_n, x_k) < 1$ for all $n \geq k$. Let $R > 1$ be such that than $d(x_i, x_k) < R$ for $1 \leqslant i \leqslant k-1$. Then $x_n \in B(x_k, R)$ for all n, so $\{x_n\}$ is bounded. ∎

Proposition 4.3. *If $\{x_n\}$ is convergent, then $\{x_n\}$ is a Cauchy sequence.*

Proof. Assume that $x_n \to x$. Then for a given $\varepsilon > 0$ there exists $k \in \mathbb{N}$ such that $d(x_n, x) < \varepsilon/2$ for all $n \geq k$. Hence taking any $n, m \geq k$,

$$d(x_n, x_m) \leqslant d(x_n, x) + d(x, x_m) < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

So $\{x_n\}$ is Cauchy. ∎

Proposition 4.4. *If $\{x_n\}$ is Cauchy and it contains a convergent subsequence, then $\{x_n\}$ converges.*

Proof. Assume that $\{x_n\}$ is Cauchy and $x_{k_n} \to x$. We will show that $x_n \to x$. Let $\varepsilon > 0$. Since $\{x_n\}$ is Cauchy, there exists k' such that $d(x_n, x_{k_n}) < \varepsilon/2$ for all $n \geq k'$. Also since $x_{k_n} \to x$, there exists k'' such that $d(x_{k_n}, x) < \varepsilon/2$ for all $n \geq k''$. Set $k = \max\{k', k''\}$. Then for $n \geq k$,

$$d(x_n, x) \leqslant d(x_n, x_{k_n}) + d(x_{k_n}, x) < \varepsilon/2 + \varepsilon/2 = \varepsilon$$

showing that $x_n \to x$. ∎

A Cauchy sequence need not converge. For example, consider $\{1/n\}$ in the metric space $((0,1), |\cdot|)$. Clearly, the sequence is Cauchy in $(0,1)$ but does not converge to any point of the interval.

Definition 4.5. *A metric space (X, d) is called* **complete** *if every Cauchy sequence $\{x_n\}$ in X converges to some point of X. A subset A of X is called* **complete** *if A as a metric subspace of (X, d) is complete, that is, if every Cauchy sequence $\{x_n\}$ in A converges to a point in A.*

By the above example, not every metric space is complete; $(0,1)$ with the usual metric is not complete. Also $\mathbb{Q}$ is not complete.

Theorem 4.6. *The space $\mathbb{R}$ with the usual metric is complete.*

Proof. Let $\{x_n\}$ be a Cauchy sequence in $\mathbb{R}$. Then it is bounded, say $|x_n| \leqslant M$. Set $S_n = \{x_k \mid k \geq n\}$ and $y_n = \inf S_n$. Then $S_{n+1} \subset S_n$ and so $\{y_n\}$ is increasing and $y_n \leqslant M$ for all n. Hence $\{y_n\}$ converges, say to x (see

Proposition 11.11 in Appendix). We claim that also $x_n \to x$. To see this choose N so that $|x_n - x_m| < \varepsilon/2$ for $n, m \geq N$. In particular,

$$x_N - \varepsilon/2 < x_k < x_N + \varepsilon/2 \quad \text{for all } k \geq N.$$

Hence

$$x_N - \varepsilon/2 \leqslant y_n \leqslant x_N + \varepsilon/2 \quad \text{for all } n \geq N.$$

Let $n \to \infty$. Then

$$x_N - \varepsilon/2 \leqslant x \leqslant x_N + \varepsilon/2,$$

or equivalently, $|x_N - x| \leqslant \varepsilon/2$. Hence for $n \geq N$,

$$|x_n - x| \leqslant |x_n - x_N| + |x_N - x| < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

Thus $\{x_n\}$ converges to x. ■

A subspace of a complete metric space may not be complete. However, the following holds.

Theorem 4.7. *If (X, d) is a complete metric space and Y is a closed subspace of X, then (Y, d) is complete.*

Proof. Let $\{x_n\}$ be a Cauchy sequence of points in Y. Then $\{x_n\}$ also satisfies the Cauchy condition in X, and since (X, d) is complete, there exists $x \in X$ such that $x_n \to x$. But Y is also closed, so $x \in Y$ showing that Y is complete. ■

Theorem 4.8. *If (X, d) is a metric space, $Y \subset X$ and (Y, d) is complete, then Y is closed.*

Proof. Let $\{x_n\}$ be a sequence of points in Y such that $x_n \to x$. We have to show that $x \in Y$. Since $\{x_n\}$ converges in X, it satisfies the Cauchy condition in X and so, it also satisfies the Cauchy condition in Y. Since (Y, d) is complete, it converges to some point in Y, say to $y \in Y$. Since any sequence can have at most one limit, $x = y$. So $x \in Y$ and Y is closed. ■

Theorem 4.9. *If (X_i, d_i) are complete metric spaces for $i = 1, \ldots, m$, then the product (X, d) is a complete metric space.*

Proof. Let $x_n = (x_n^1, \ldots, x_n^m)$ and $\{x_n\}$ be a Cauchy sequence in (X, d). Then for a given $\varepsilon > 0$ there exists k such that $d(x_n, x_m) < \varepsilon$ for all $n, m \geq k$. Since

$$d_j(x_n^j, x_m^j) \leqslant d(x_n, x_m) < \varepsilon,$$

it follows that $\{x_n^j\}$ is Cauchy in (X_j, d_j) for $j = 1, \dots m$. Since (X_j, d_j) is complete, for $j = 1, \dots, m$ there exists $x^j \in X_j$ such that $x_n^j \to x^j$. Then, in view of Proposition 2.8, $x_n \to x$, where $x = (x^1, \dots, x^m)$. ∎

We write $C(X, Y)$ for the space of continuous functions from X to Y. A function $f : X \to Y$ is said to be **bounded** if the image of X, $f(X)$, is contained in a bounded subset of Y, and we write $C_b(X, Y)$ for the space of bounded continuous function $f : X \to Y$. If $Y = \mathbb{R}$, we simply write $C(X)$ instead of $C(X, \mathbb{R})$ and $C_b(X)$ instead of $C_b(X, \mathbb{R})$. For $f, g \in C_b(X, Y)$, we set

$$\rho(f, g) := \sup\{d'(f(x), g(x)) \mid x \in X\},$$

where d' denotes the metric on Y.

Theorem 4.10. *The space $(C_b(X, Y), \rho)$ is a complete metric space if (Y, d') is complete.*

Proof. The verification that ρ is a metric is left as an exercise. Suppose that Y is complete, and suppose that $\{f_n\}$ is a Cauchy sequence in $C_b(X, Y)$. Then for every $x \in X$,

$$d'(f_n(x), f_m(x)) \leqslant \rho(f_n, f_m)$$

so that $\{f_n(x)\}$ is a Cauchy sequence on Y. Hence there exists a point, denoted by $f(x) \in Y$, such that $d'(f_n(x), f(x)) \to 0$. In this way we obtain a function $f : X \to Y$ which associates with a point $x \in X$ a point which is the limit of $\{f_n(x)\}$. We must check that f is continuous and bounded, and that $\rho(f_n, f) \to 0$. Let $x \in X$, and $\varepsilon > 0$. Then there exists N such that $d'(f(x), f_N(x)) < \varepsilon/3$, and an open ball $B(x, \delta)$ such that $d'(f_N(x), f_N(y)) < \varepsilon/3$ for every $y \in B(x, \delta)$. It follows that for every $y \in B(x, \delta)$,

$$d'(f(x), f(y)) \leqslant d'(f(x), f_N(x)) + d'(f_N(x), f_N(y)) + d'(f_N(y), f(y)) < \varepsilon.$$

Hence f is continuous. Now given $\varepsilon > 0$, chose n_0 such that $\rho(f_n, f_m) < \varepsilon$ for all $n, m \geq n_0$. Then for every $x \in X$,

$$d'(f_n(x), f(x)) = \lim_{m \to \infty} d'(f_n(x), f_m(x)) \leqslant \varepsilon$$

for $n \geq n_0$. This says that $\rho(f_n, f) \leqslant \varepsilon$ for $n \geq n_0$. It remains to show that f is bounded. Take $x, y \in X$ and let $N \in \mathbb{N}$ be such that

$$d'(f(x), f_N(x)) < 1/2 \quad \text{and} \quad d'(f(y), f_N(y)) < 1/2$$

Note that we can find such an N since $\rho(f_n, f) \to 0$. Then

$$\begin{aligned} d'(f(x), f(y)) &\leqslant d'(f(x), f_N(x)) + d'(f_N(x), f_N(y)) + d'(f_N(y), f(y)) \\ &< 1 + d'(f_N(x), f_N(y)) \leqslant 1 + \text{diam} f_N(X). \end{aligned}$$

Since $x, y \in X$ were arbitrary, $\text{diam} f(X) \leqslant 1 + \text{diam} f_N(X)$. Hence f is bounded. The proof is completed. ■

Corollary 4.11. *The space $(C_b(X), \rho)$ is complete.*

Structure of complete metric spaces: Baire's theorem

Let (X, d) be a metric space. If U and V are open and dense, then $U \cap V$ is also open and dense. To see that $U \cap V$ is dense, we have to show that $O \cap U \cap V$ is non-empty for any non-empty open set O. Since U is dense, there is $u \in O \cap U$, and since $O \cap U$ is open, $B(u, r) \subset O \cap U$ for some $r > 0$. Since V is dense, $B(u, r) \cap V \neq \emptyset$ so that, $\emptyset \neq B(u, r) \cap V \subset O \cap U \cap V$. If U and V are assumed to be dense but not necessarily open, then the intersection $U \cap V$ does not have to be dense. For example, let U be the set of rational numbers and V the set of irrational numbers $\mathbb{Q}^c$. Then both sets are dense in $\mathbb{R}$ with the usual metric, however, $U \cap V = \emptyset$. Consider, now a sequence of dense and open sets U_n. In general, the intersection $\bigcap_{n \geq 1} U_n$ may be empty. For example, consider $(\mathbb{Q}, d)$ with the usual metric d. Let $\{q_n | n \in \mathbb{N}\}$ be an enumeration of rational numbers, and let $U_n = \mathbb{Q} \setminus \{q_n\}$. Then each U_n is open since it is a complement of a closed set $\{q_n\}$, and is dense . However, $\bigcap_{n \geq 1} U_n = \bigcap_{n \geq 1} \big[\mathbb{Q} \setminus \{q_n\}\big] = \mathbb{Q} \setminus \bigcup_{n \geq 1} \{q_n\} = \emptyset$. The Baire theorem says that if (X, d) is complete, then $\bigcap_{n \geq 1} U_n$ is dense.

Theorem 4.12 (Baire). *Let (X, d) be a complete metric space, and let $\{U_n\}$ be a sequence of open and dense subsets of X. Then $\bigcap_{n \geq 1} U_n$ is dense.*

Proof. It suffices to show that every open ball $B(x, r)$ contains a point belonging to $\bigcap_{n \geq 1} U_n$. Since U_1 is open and dense, $B(x, r) \cap U_1$ is non-empty and open. So, there exists an open ball $B(x_1, R)$ with $R < 1$ such that $B(x_1, R) \subseteq B(x, r)$ and $B(x_1, R) \subseteq U_1$. Taking $r_1 < R$, we get that $\overline{B}(x_1, r_1) \subseteq B(x, r)$ and $\overline{B}(x_1, r_1) \subseteq U_1$. Similarly, since U_2 is open and dense, there exists x_2 and $r_2 < 1/2$ such that $\overline{B}(x_2, r_2) \subset \overline{B}(x_1, r_1) \cap U_2$. Continuing in this way we find a sequence of balls $\overline{B}(x_n, r_n)$ with $r_n < 1/n$ and $\overline{B}(x_{n+1}, r_{n+1}) \subseteq \overline{B}(x_n, r_n) \cap U_{n+1}$. We claim that $\{x_n\}$ is Cauchy. By construction, $\overline{B_n}(x_n, r_n) \subset \overline{B}(x_k, r_k)$ for all $n \geq k$. Given $\varepsilon > 0$ choose $k \in \mathbb{N}$ so that $1/k < \varepsilon/2$. Then, if $n, m \geq k$,

$$d(x_n, x_m) \leqslant d(x_n, x_k) + d(x_k, x_m) < 1/k + 1/k < \varepsilon.$$

Because (X, d) is complete, $\{x_n\}$ converges, say to y. The point y lies in all balls $\overline{B}(x_k, r_k)$ since $x_n \in \overline{B}(x_k, r_k)$ for all $n \geq k$ and $\overline{B}(x_k, r_k)$ is closed for all k, so that after taking a limit as $n \to \infty$, $y \in \overline{B}(x_k, r_k)$ for all k. In particular, $y \in \overline{B}(x_1, r_1) \subseteq B(x, r)$ and $y \in \overline{B}(x_{n+1}, r_{n+1}) \subset U_{n+1}$ for all n. Consequently, $y \in B(x, r) \cap \bigcap_{n \geq 1} U_n$, and the proof is finished. ■

As a consequence we obtain the following theorem.

Theorem 4.13. *If (X, d) is a complete metric space and $\{F_n\}$ is a sequence of nowhere dense subsets of X, (i.e int $\overline{F_n} = \emptyset$) then $\bigcup F_n$ has empty interior.*

Proof. Arguing by contradiction assume that $\bigcup F_n$ has non-empty interior. So $B(x, r) \subseteq \bigcup F_n$ for some x and $r > 0$. Define $U_n = X \setminus \overline{F_n}$. Clearly, U_n is open and we claim that it is dense. Indeed, if for some non-empty open set V we have $V \cap U_n = \emptyset$, then $V \subseteq X \setminus U_n = \overline{F_n}$ contradicting that $\overline{F_n}$ has empty interior. Consequently, in view of the above theorem, $\bigcap_{n \geq 1} U_n$ is dense. So $B(x, r) \cap \bigcap_{n \geq 1} U_n \neq \emptyset$. On the other hand, $B(x, r) \subseteq \bigcup F_n \subseteq \bigcup \overline{F_n}$ so that $\emptyset = B(x, r) \cap \big[X \setminus \bigcup_{n \geq 1} \overline{F_n}\big] = B(x, r) \cap \bigcap_{n \geq 1} \big[X \setminus \overline{F_n}\big] = B(x, r) \cap \bigcap_{n \geq 1} U_n$, a contradiction. ■

Example 4.14. The metric space $\mathbb{R}$ with the standard metric cannot be written as a countable union of nowhere dense sets since it is complete. By contrast, $\mathbb{Q}$ with the standard metric can be written as the union of one point sets $\{q_n\}$, where $\{q_n \mid n \in \mathbb{N}\}$ is an enumeration of $\mathbb{Q}$. Every one point set $\{q_n\}$ is closed in $\mathbb{Q}$ and its interior is empty, so nowhere dense. This does not contradict Baire's theorem since $\mathbb{Q}$ with the standard metric is not complete.

Applications

Theorem 4.15. *Let (X, d) be a complete metric space, and let $\{f_n\}$ be a sequence of continuous functions $f_n : X \to \mathbb{R}$. Assume that the sequence $\{f_n(x)\}$ is bounded for every $x \in X$. Then there exists a non-empty open set $U \subset X$ on which the sequence $\{f_n\}$ is bounded, that is, there is a constant M such that $|f_n(x)| \leqslant M$ for all $x \in U$ and all $n \in \mathbb{N}$.*

Proof. Since the function f_n is continuous, the set $f_n^{-1}([-m,m]) = \{x \in X \mid |f_n(x)| \leqslant m\}$ is closed for any pair of positive integers n and m. Thus,

$$E_m = \{x \in X \mid |f_n(x)| \leqslant m \text{ for all } n \in \mathbb{N}\} = \bigcap_n f_n^{-1}([-m,m])$$

is closed for every $m \in \mathbb{N}$. If x is any point in X, then $|f_n(x)| \leqslant k$ for some $k \in \mathbb{N}$ and all n because $\{f_n(x)\}$ is bounded. Hence $X = \bigcup_m E_m$. In view of the Baire theorem, one of the sets, say E_m, has non-empty interior. Setting $U = E_m^\circ$ the conclusion follows. ∎

Theorem 4.16. *There exists a continuous function $f : [0,1] \to \mathbb{R}$ which is not differentiable at any point $x \in [0,1)$.*

Proof. Recall that f has a right-hand derivative at x if

$$\lim_{h\to 0^+} [(f(x+h) - f(x))/h] \text{ exists.}$$

We denote this limit by $f'_+(x)$. In particular, if f is differentiable at $x \in [0,1)$ then $f'_+(x)$ exists and is equal to $f'(x)$. Consider the complete metric space $C([0,1],\mathbb{R}) = C_b([0,1],\mathbb{R})$ with a metric d given by

$$d(f,g) = \sup\{|f(x) - g(x)||x \in [0,1]\}.$$

Let

$$M = \{f \in C([0,1],\mathbb{R})| \text{ there exists } x \in [0,1) \text{ such that } f'_+(x) \text{ exists}\}$$

and let M_m, for $m \geq 2$, be the set of all $f \in C([0,1],\mathbb{R})$ for which there exists some $x \in [0, 1-1/m]$ such that

$$|f(x+h) - f(x)| \leqslant m \cdot h \text{ for all } h \in [0, 1/m].$$

Claim 1: $M \subset \bigcup_{m\geq 2} M_m$. Let $f \in M$. Then there exists $x \in [0,1)$ such that $f'_+(x)$ exists. We will show that $|f(x+h) - f(x)| \leqslant m \cdot h$ for some $m \in \mathbb{N}$ and all $0 \leqslant h \leqslant 1/m$. Since

$$\lim_{h\to 0^+} \frac{f(x+h) - f(x)}{h} = f'_+(x),$$

we have

$$\lim_{h\to 0^+} \left|\frac{f(x+h) - f(x)}{h}\right| = |f'_+(x)|. \tag{1}$$

Take an integer $k \geq 2$ such that $|f'_+(x)| < k$ and $x \in [0, 1-1/k]$. In view of (1), there exists $0 < \delta < 1/k$ such that

$$|f(x+h) - f(x)| \leqslant k \cdot h \text{ for all } 0 \leqslant h \leqslant \delta.$$

Since f is continuous on a closed and bounded interval, there is $C > 0$ such that $|f(x)| \leqslant C$ for all $x \in [0,1]$ (this will be proved later on in the section on compactness). Let k' be any integer so that $2C/\delta < k'$. Then, for $\delta \leqslant h \leqslant 1$ such that $x + h \leqslant 1$,

$$|f(x+h) - f(x)| \leqslant |f(x+h)| + |f(x)| \leqslant 2C = \frac{2C}{\delta} \cdot \delta \leqslant \frac{2C}{\delta} \cdot h \leqslant k' \cdot h.$$

Taking $m = \max\{k, k'\}$, we have $x \in [0, 1-1/m]$ and $|f(x+h) - f(x)| \leqslant m \cdot h$ for all $h \in [0, 1/m]$, so that $f \in M_m$.
Claim 2: M_m is closed for all $m \geq 2$. To see this, take $f \in \overline{M_m}$. We will show that $f \in M_m$, that is, $|f(x+h) - f(x)| \leqslant m \cdot h$ for some $x \in [0, 1-1/m]$ and all $h \in [0, 1/m]$. There exists a sequence $(f_k) \subset M_m$ such that $d(f_k, f) = \sup\{|f_k(x) - f(x)| | \, x \in [0,1]\} \to 0$ as $k \to \infty$. Since $f_k \in M_m$, there exists $x_k \in [0, 1-1/m]$ such that

$$|f_k(x_k + h) - f_k(x_k)| \leqslant m \cdot h \tag{2}$$

for all $h \in [0, 1/m]$. Since $\{x_k\} \subseteq [0, 1-1/m]$ we may assume that there exists a subsequence, again denoted by $\{x_k\}$, such that $x_k \to x \in [0, 1-1/m]$. Hence, by the triangle inequality and by (2),

$$\begin{aligned}
|f(x+h) - f(x)| &\leqslant |f(x+h) - f(x_k+h)| + |f(x_k+h) - f_k(x_k+h)| \\
&\quad + |f_k(x_k+h) - f_k(x_k)| + |f_k(x_k) - f_k(x)| + |f_k(x) - f(x)| \\
&\leqslant |f(x+h) - f(x_k+h)| + d(f_k, f) + m \cdot h \\
&\quad + |f_k(x_k) - f_k(x)| + d(f_k, f)
\end{aligned}$$

for all $0 \leqslant h \leqslant 1/m$. Since $d(f_k, f) \to 0$, $|f(x+h) - f(x_k+h)| \to 0$, and $|f(x) - f(x_k)| \to 0$, as $k \to \infty$, we conclude that

$$|f(x+h) - f(x)| \leqslant m \cdot h$$

for all $0 \leqslant h \leqslant 1/m$. Consequently, $f \in M_m$ and M_m is closed.
Claim 3: $M_m^\circ = \emptyset$. Let $f \in M_m$, and let $\varepsilon > 0$. Then there exists a piecewise linear function $g : [0,1] \to \mathbb{R}$ such that $d(f,g) = \sup\{|f(x) - g(x)| \mid 0 \leqslant x \leqslant 1\} < \varepsilon$ and $|g'_+(x)| > m$ for all $x \in [0,1]$. (See Figure 2.) That is, $g \in B(f, \varepsilon)$ and $g \notin M_m$. (Here $B(f, \varepsilon)$ is a ball in $C([0,1], \mathbb{R})$ with centre

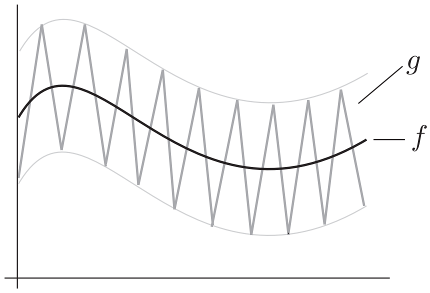


Figure 2: The black curve is the graph of f and the grey curve is the graph of g.

at f and radius ε). So $M_m^\circ = \emptyset$.

In view of the Baire's theorem, $C([0,1],\mathbb{R}) \neq \bigcup_{m\geq 2} M_m$ since otherwise $\bigcup_{m\geq 2} M_m$ has non-empty interior. Hence there exists $f \in C([0,1],\mathbb{R})$ so that $f \notin \bigcup_{m\geq 2} M_m$. Since $M \subseteq \bigcup_{m\geq 2} M_m$, $f \notin M$. Since M contains all functions which are differentiable at least at one point in $[0,1)$, f is not differentiable at any $x \in [0,1)$. ∎

Contraction mapping principle: Banach fixed point theorem

Let (X,d) be a metric space and let $f : X \to X$. A point $x \in X$ is a **fixed point** of f if $f(x) = x$. The solution of many classes of equations can be regarded as fixed points of appropriate functions. In this section we give conditions that guarantee the existence of fixed points of certain functions. A function $f : X \to X$ is called a **contraction** if there exists $\alpha \in (0,1)$ such that

$$d(f(x), f(y)) \leqslant \alpha d(x,y) \tag{8}$$

for all $x, y \in X$. Note that a contraction is uniformly continuous.

Theorem 4.17 (**Banach Fixed Point Theorem).** *Let $f : X \to X$ be a contraction of a complete metric space. Then f has a unique fixed point p. For any $x \in X$, define $x_0 = x$ and $x_{n+1} = f(x_n)$ for $n \geq 0$. Then $x_n \to p$, and*

$$d(x,p) \leqslant \frac{d(x, f(x))}{1-\alpha}. \tag{9}$$

Proof. We start with the uniqueness of the fixed point of f. Assume that $p \neq q$ and that $f(p) = p$ and $f(q) = q$. Then

$$d(p,q) = d(f(p), f(q)) \leqslant \alpha d(p,q)$$

so that $d(p,q) = 0$ since $\alpha \in (0,1)$. So $p = q$, contradicting our assumption. Hence f has at most one fixed point. Fix any point $x \in X$, and let $x_0 = x$ and $x_{n+1} = f(x_n)$ for $n \geq 0$. Then for any n,

$$d(x_{n+1}, x_n) = d(f(x_n), f(x_{n-1})) \leqslant \alpha d(x_n, x_{n-1})$$

and,

$$d(x_{n+1}, x_n) \leqslant \alpha d(x_n, x_{n-1}) \leqslant \alpha^2 d(x_{n-1}, x_{n-2}) \leqslant \cdots \leqslant \alpha^n d(x_1, x_0).$$

For $m > n$,

$$\begin{aligned}
d(x_m, x_n) &\leqslant d(x_n, x_{n+1}) + d(x_{n+1}, x_{n+2}) + \cdots + d(x_{m-1}, x_m) \\
&\leqslant \big(\alpha^n + \alpha^{n+1} + \cdots + \alpha^{m-1}\big) d(x_1, x_0) \leqslant \big(\sum_{i=n}^{\infty} \alpha^i\big) d(x_1, x_0) \\
&= \alpha^n \big(\sum_{i=0}^{\infty} \alpha^i\big) d(x_1, x_0) \ = \frac{\alpha^n d(x_1, x_0)}{1-\alpha}.
\end{aligned}$$

Since $\alpha^n \to 0$ as $n \to \infty$ (recall $\alpha \in (0,1)$), the sequence $\{x_n\}$ is Cauchy in X. Since (X,d) is complete, there exists $p \in X$ such that $x_n \to p$. Taking a limit $m \to \infty$ in the last inequality we find that

$$d(p, x_n) \leqslant \frac{\alpha^n d(x_1, x_0)}{1-\alpha}. \tag{10}$$

Thus,

$$\begin{aligned}
d(f(p), p) &\leqslant d(f(p), x_{n+1}) + d(x_{n+1}, p) = d(f(p), f(x_n)) + d(x_{n+1}, p) \\
&\leqslant d(p, x_n) + d(x_{n+1}, p) \leqslant \frac{\alpha^n d(x_1, x_0)}{1-\alpha} + \frac{\alpha^{n+1} d(x_1, x_0)}{1-\alpha} \\
&= \alpha^n \cdot \frac{(1+\alpha) d(x_1, x_0)}{1-\alpha} \to 0 \text{ as } n \to \infty,
\end{aligned}$$

and therefore $p = f(p)$. The inequality (9) follows from (10) by taking $n = 0$. ■

Here is an application of the Banach fixed point theorem to the local existence of solutions of ordinary differential equations.

Theorem 4.18 (Picard's Theorem). *Let U be an open subset of $\mathbb{R}^2$ and let $f : U \to \mathbb{R}$ be a continuous function which satisfies the Lipschitz condition with respect to the second variable, that is,*

$$|f(x, y_1) - f(x, y_2)| \leqslant \alpha |y_1 - y_2|$$

for all $(x, y_1), (x, y_2) \in U$, and some $\alpha > 0$. Then for a given $(x_0, y_0) \in U$ there is $\delta > 0$ so that the differential equation

$$y'(x) = f(x, y(x))$$

has a unique solution $y : [x_0 - \delta, x_0 + \delta] \to \mathbb{R}$ such that $y(x_0) = y_0$.

Proof. Note that it is enough to show that there are $\delta > 0$ and a unique function $y : [x_0 - \delta, x_0 + \delta] \to \mathbb{R}$ such that

$$y(x) = y_0 + \int_{x_0}^{x} f(t, y(t))dt.$$

Fix $(x_0, y_0) \in U$. Then we find $\delta > 0$ and $b > 0$ such that if $I = [x_0 - \delta, x_0 + \delta]$ and $J = [y_0 - b, y_0 + b]$, then $I \times J \subset U$. Since f is continuous and $I \times J$ is closed and bounded, f is bounded on $I \times J$. That is, $|f(x, y)| \leqslant M$ for some M and all $(x, y) \in I \times J$. Replacing δ by a smaller number we may assume that $\alpha\delta < 1$ and $M\delta < b$. Denote by X the set of all continuous functions $g : I \to J$. The set X with the metric $d(g, h) = \sup\{|g(x) - h(x)|, x \in I\}$ is a complete metric space. For $g \in X$, let

$$(Tg)(x) = y_0 + \int_{x_0}^{x} f(t, g(t))dt.$$

Then $Tg : I \to \mathbb{R}$ is continuous since if $x_1, x_2 \in I$ and $x_2 > x_1$, then

$$|(Tg)(x_2) - (Tg)(x_1)| = \left| \int_{x_1}^{x_2} f(t, g(t))dt \right| \leqslant \int_{x_1}^{x_2} |f(t, g(t))|dt \leqslant M|x_2 - x_1|.$$

For $x_0 \leqslant x \leqslant x_0 + \delta$,

$$|(Tg)(x) - y_0| = \left| \int_{x_0}^{x} f(t, g(t))dt \right| \leqslant \int_{x_0}^{x} |f(t, g(t))|dt \leqslant M|x - x_0| \leq M\delta < b$$

The same inequality holds for $x_0 - \delta \leqslant x \leqslant x_0$, and so $Tg \in X$ for any $g \in X$. Since f is Lipschitz with respect to the second variable, we obtain for $g, h \in X$ and $x \in [x_0, x_0 + \delta]$,

$$\begin{aligned}|(Tg)(x) - (Th)(x)| &= \left| \int_{x_0}^{x} [f(t, g(t)) - f(t, h(t))] \, dt \right| \\ &\leqslant \int_{x_0}^{x} |f(t, g(t)) - f(t, h(t))| dt \\ &\leqslant \alpha |x - x_0| d(g, h) < \alpha \delta d(g, h).\end{aligned}$$

Similarly, $|(Tg)(x) - (Th)(x)| \leqslant \alpha|x - x_0|d(g, h) < \alpha\delta d(g, h)$ for $x \in [x_0 - \delta, x_0]$. Since $\alpha\delta < 1$, T is a contraction and in view of Banach's fixed point theorem there exists a unique continuous function $y : I \to J$ such that

$$y(x) = (Ty)(x) = y_0 + \int_{x_0}^{x} f(t, y(t)) dt.$$

■

Completions

The space $(0, 1)$ with the usual metric is not complete but is a subspace of the complete metric space $[0, 1]$ with the usual metric. This example illustrates the general situation: every metric space X may be regarded as a subspace of a complete metric space $\widetilde{X}$ in such a way that $\overline{X} = \widetilde{X}$.
We will need the following concept.

Definition 4.19. *A map f from (X, d) to (Y, ρ) is called an* **isometry** *if*

$$\rho(f(x), f(y)) = d(x, y) \qquad \textit{for all } x, y \in X.$$

Note that an isometry is always injective. An isometry is distance preserving. If $f : X \to Y$ is a **surjective** isometry, then the inverse $f^{-1} : Y \to X$ is also an isometry, and the spaces (X, d) and (Y, ρ) are called **isometric**. Two isometric spaces can be regarded as indistinguishable for all practical purposes that involve only distance.

Definition 4.20. *A* **completion** *of a metric space (X, d) is a pair consisting of a complete metric space $(\widetilde{X}, \widetilde{d})$ and an isometry $\varphi : X \to \widetilde{X}$ such that $\varphi(X)$ is dense in $\widetilde{X}$.*

Theorem 4.21. *Let (X, d) be a metric space. Then (X, d) has a completion. The completion is unique in the following sense: If $((X_1, d_1), \varphi_1)$ and $((X_2, d_2), \varphi_2)$ are completions of (X, d), then (X_1, d_1) and (X_2, d_2) are isometric. That is, there exists a surjective isometry $f : X_1 \to X_2$ such that $f \circ \varphi_1 = \varphi_2$.*

Proof.
Existence: Let $B(X)$ be the space of bounded real valued functions defined on X equipped with the uniform norm $\sigma(f,g) = \sup_{y\in X} |f(y) - g(y)|$. Fix a point $a \in X$. With every $x \in X$ we associate a function $f_x : X \to \mathbb{R}$ defined by

$$f_x(y) = d(y,x) - d(y,a), \qquad y \in X.$$

We have

$$|f_x(y)| = |d(y,x) - d(y,a)| \leqslant d(x,a)$$

so that f_x is bounded. Since

$$|f_{x_1}(y) - f_{x_2}(y)| \leqslant d(x_1,x_2) \qquad \text{for all } y \in X,$$

$\sigma(f_{x_1}, f_{x_2}) = \sup_{y\in X}\{|f_{x_1}(y) - f_{x_2}(y)|\} \leqslant d(x_1,x_2)$. On the other hand,

$$\sigma(f_{x_1}, f_{x_2}) \geq |f_{x_1}(x_2) - f_{x_2}(x_2)| = d(x_1,x_2).$$

Hence

$$\sigma(f_{x_1}, f_{x_2}) = d(x_1,x_2),$$

and the map $f : X \to B(X)$ defined by $f(x) = f_x$ is an isometry onto $f(X)$,

$$\sigma(f(x_1), f(x_2)) = d(x_1,x_2).$$

Denote by X' the closure of $f(X)$ in $B(X)$ and let d' be the metric on X' induced by σ. Since $(B(X),\sigma)$ is complete and X' is closed in $B(X)$, the space (X',d') is complete.
Uniqueness:
The isometry $\varphi_1 : X \to \varphi_1(X)$ has an inverse $\varphi_1^{-1} : \varphi_1(X) \to X$. Then $\varphi_2 \circ \varphi_1^{-1}$ is an isometry from $\varphi_1(X)$ onto $\varphi_2(X_2)$. Since $\varphi_1(X)$ is dense in (X_1,d_1), $\varphi_2 \circ \varphi_1^{-1}$ extends to a map $\varphi : X_1 \to X_2$ satisfying

$$d_2(\varphi(x), \varphi(y)) = d_1(x,y), \quad \text{for all} \quad x,y \in X_1.$$

Since X_1 is complete, in view of the above equation, $\varphi(X_1)$ is closed in X_2. Since $\varphi \circ \varphi_1 = \varphi_2$, $\varphi_2(X) \subset \varphi(X_1)$. This implies that $X_2 = \overline{\varphi_2(X)} \subset \overline{\varphi(X_1)} = \varphi(X_1)$ since $\varphi(X_1)$ is closed in X_2. Consequently, $\varphi(X_1) = X_2$, i.e., φ is surjective and the proof is completed. ■

5 Compact Metric Spaces

We start with the classical theorem of Bolzano-Weierstrass.

Theorem 5.1 (**Bolzano-Weierstrass).** *Let I be a closed and bounded interval in $\mathbb{R}$, and let $\{x_n\}$ be a sequence in I. Then there exists a subsequence $\{x_{n_k}\}$ which converges to a point in I.*

Proof. Without loss of generality we may assume that $I = [0, 1]$. Bisect the interval $[0, 1]$ and consider the two intervals $[0, 1/2]$ and $[1/2, 0]$. One of these subintervals must contain x_n for infinitely many n. Call this subinterval I_1. Now bisect I_1. Again, one of the two subintervals contains x_n for infinitely many n. Denote this subinterval by I_2. Proceeding in this way we find a sequence of closed intervals I_n, each one contained in the preceding one, each one half of the length of the preceding one, and each containing x_n for infinitely many n. Choose an integer n_1 so that $x_{n_1} \in I_1$. Then choose $n_2 > n_1$ such that $x_{n_2} \in I_2$. Then choose $n_3 > n_2$ such that $x_{n_3} \in I_3$, and so on. Continuing this way we choose we find a sequence $\{x_{n_k}\}$ such that $x_{n_k} \in I_k$. If $i, j \geq k$, then $x_{n_i}, x_{n_j} \in I_k$ and so

$$|x_{n_i} - x_{n_j}| \leqslant 1/2^k.$$

Hence $\{x_{n_k}\}$ is Cauchy and since $[0, 1]$ is complete, $\{x_{n_k}\}$ converges to a point in $[0, 1]$. ■

Definition 5.2. *A metric space (X, d) is called* **compact** *if every sequence in X has a convergent subsequence. A subspace Y of X is compact if every sequence in Y has a subsequence converging to a point in Y.*

Proposition 5.3. *Let (X, d) be compact and Y a closed subset of X. Then Y is compact.*

Proof. Let $\{x_n\}$ be a sequence in Y. Since X is compact, the sequence $\{x_n\}$ has a converging subsequence, say $x_{n_k} \to x$. Since Y is closed, $x \in Y$. ■

Proposition 5.4. *Let X be a metric space and Y a compact subset of X. Then Y is closed and bounded.*

Proof. Take any $x \in \overline{Y}$. There exists a sequence $\{x_n\}$ in Y converging to x. Since Y is compact, the sequence $\{x_n\}$ has a converging subsequence, say $x_{n_k} \to y$ with $y \in Y$. In view of the uniqueness of the limit, $y = x$. Hence Y is closed. To see that Y is bounded, we argue by contradiction and construct a sequence $\{x_n\}$ which does not have a converging subsequence.

Fix any point $y \in X$. For every $n \in \mathbb{N}$, there exists a point $x_n \in Y$ so that $d(x_n, y) \geq n$ since otherwise $Y \subseteq \overline{B}(y, n)$ for some n. The sequence $\{x_n\}$ contains a converging subsequence since Y is compact. Say $x_{n_k} \to x \in Y$. Let $\varepsilon = d(x, y)$. Then $d(x_{n_k}, x) \leqslant 1$ for all $k \geq N$. Hence by the triangle inequality,

$$d(x, y) \geq d(y, x_{n_k}) - d(x, x_{n_k}) \geq n_k - 1 \geq k - 1$$

for all $k \geq N$; contradiction. Consequently, Y is bounded. ∎

Combining Proposition 5.4 with Theorem 5.1 we get

Theorem 5.5. *A subset Y of $\mathbb{R}$ is compact if and only if Y is bounded and closed.*

The result is also valid in $\mathbb{R}^n$ with the standard metric: A subset of $\mathbb{R}^n$ is compact if and only if it is bounded and closed. This follows from the fact that if A_i is a compact subset of (X_i, d_i) for $1 \leqslant i \leqslant n$, then $A_1 \times A_2 \times \cdots \times A_n$ is compact in the product space $X_1 \times X_2 \times \cdots \times X_n$. In particular, using Theorem 5.1, $[-a, a]^n$ is compact in $\mathbb{R}^n$. So if A is bounded and closed in $\mathbb{R}^n$, then A is a subset of a compact set $[-a, a]^n$, and then Proposition 5.3 implies that A is compact.

Theorem 5.5 does not hold true for general metric spaces.

Example 5.6. Consider the metric space $(C([0,1], \mathbb{R}), d)$ consisting of all continuous real valued functions on the interval $[0, 1]$ with the supremum metric $d(f, g) = \sup\{|f(x) - g(x)| \mid x \in [0, 1]\}$. Let $A = \{f \mid d(f, 0) \leq 1\}$ be the closed unit ball in $C([0, 1], \mathbb{R})$. Then A is closed since the distance function is continuous, and bounded since $A \subseteq B(0, 2)$. Let $f_i(x) = x^i$ for $x \in [0, 1]$. Then the sequence $\{f_1, f_2, f_3, \cdots\}$ lies in A but has no convergent subsequence. To see this, note that any subsequence converges pointwise to

$$f(x) = \begin{cases} 0 & \text{if } 0 \leq x < 1, \\ 1 & \text{if } x = 1, \end{cases}$$

so this is the only possible limit. But convergence in the supremum metric is the same as uniform convergence, and uniform limits of continuous functions are continuous (Theorem 3.10). So no subsequence has a limit.

Theorem 5.7. *Let (X, d) and (Y, d') be metric spaces and let $f : X \to Y$ be continuous. If a subset $K \subseteq X$ is compact, then $f(K)$ is compact in (Y, d'). In particular, if (X, d) is compact, then $f(X)$ is compact in Y.*

Proof. Let $\{y_n\}$ be any sequence in $f(K)$, and let $\{x_n\}$ be a sequence in K of points such that $f(x_n) = y_n$. Since K is compact, $\{x_n\}$ has a converging subsequence to a point in K; say $x_{n_k} \to x$ with $x \in K$. Since f is continuous, $f(x_{n_k}) \to f(x)$. That is, $y_{n_k} \to f(x)$ and since $f(x) \in f(K)$, $f(K)$ is compact. ■

As a corollary we get

Corollary 5.8. *Let $f : X \to \mathbb{R}$ be a continuous function on a compact metric space. Then f attains a maximum and a minimum value, that is, there exist a and $b \in X$ such that $f(a) = \inf\{f(x) \mid x \in X\}$ and $f(b) = \sup\{f(x) \mid x \in X\}$.*

Proof. By Theorem 5.7, $f(X)$ is compact, so it is bounded and $\sup\{f(x) \mid x \in X\}$ is finite. Set $C = \sup\{f(x) \mid x \in X\}$. By definition of supremum, for every $n \in \mathbb{N}$, there exists x_n such that $C - 1/n \leqslant f(x_n) \leqslant C$. The sequence $\{x_n\}$ has a converging subsequence, $x_{n_k} \to b$ because X is compact. In view of the continuity of f, $f(x_{n_k}) \to f(b)$, and since $C - 1/n \leqslant f(x_n) \leqslant C$ for all n, $f(b) = C$. Similarly, there exists $a \in X$ such that $f(a) = \inf\{f(x) \mid x \in X\}$. ■

Theorem 5.9. *Suppose $f : (X, d) \to (Y, d')$ is a continuous mapping defined on a compact metric space X. Then f is uniformly continuous.*

Proof. Suppose not. Then there is some $\varepsilon > 0$ such that for all $\delta > 0$ there exist points x, y with $d(x, y) < \delta$ but $d'(f(x), f(y)) \geq \varepsilon > 0$. Take $\delta = 1/n$ and let x_n, y_n be points such that $d(x_n, y_n) < 1/n$ but $d'(f(x_n), f(y_n)) \geq \varepsilon$. Compactness of X implies that there is a subsequence $\{x_{n_k}\}$ converging to some point $x \in X$. Since $d(x_{n_k}, y_{n_k}) < 1/n_k \to 0$ as $k \to \infty$, the sequence $\{y_{n_k}\}$ converges to the same point x. Continuity of f implies that the sequences $\{f(x_{n_k})\}$, $\{f(y_{n_k})\}$ converge to $f(x)$. Then $d'(f(x_{n_k}), f(x)) < \varepsilon/2$ and $d'(f(y_{n_k}), f(x)) < \varepsilon/2$ for k large, and so,

$$d'(f(x_{n_k}, f(y_{n_k})) \leqslant d'(f(x_{n_k}), f(x)) + d'(f(x), f(y_{n_k})) < \varepsilon$$

for k large, a contradiction to the fact that $d'(f(x_n), f(y_n)) \geq \varepsilon$ for all n. ■

Characterization of Compactness for Metric Spaces

Definition 5.10. *Let (X, d) be a metric space and let $A \subseteq X$. If $\{U_i\}_{i \in I}$ is a family of subsets of X such that $A \subseteq \bigcup_{i \in I} U_i$, then it is called a* **cover** *of A, and A is said to be* **covered** *by the U_i's. If each U_i is open, then $\{U_i\}_{i \in I}$ is an* **open cover**. *If $J \subset I$ and still $A \subseteq \bigcup_{i \in J} U_i$, then $\{U_i\}_{i \in J}$ is a* **subcover**.

Definition 5.11. *Let (X, d) be a metric space and let $A \subseteq X$. Then A has the* **Heine-Borel property** *if for every open cover $\{U_i\}_{i \in I}$ of A, there is a finite set $F \subseteq I$ such that $A \subseteq \bigcup_{i \in S} U_i$.*

Example 5.12. Consider a set X with a discrete metric. Then every one-point set is open and the collection of all one-point sets is an open cover of X. Clearly, this cover does not have any proper subcover. Hence, a discrete metric space X has the Heine-Borel property if and only if X consists of a finite number of points.

Definition 5.13. *Let (X, d) be a metric space and $A \subseteq X$. Let $\varepsilon > 0$. A subset S is called an* ε**-net** *for A if $A \subseteq \bigcup_{x \in S} B(x, \varepsilon)$. A set A is called* **totally bounded** *if, for every $\varepsilon > 0$, there is a finite ε-net for A. That is, for every $\varepsilon > 0$, there is a finite set S such that $A \subseteq \bigcup_{x \in S} B(x, \varepsilon)$.*

Every totally bounded set is bounded, for if $x, y \in \bigcup_{i=1}^{n} B(x_i, \varepsilon)$, say $x \in B(x_1, \varepsilon), y \in B(x_2, \varepsilon)$, then

$$d(x, y) \leqslant d(x, x_1) + d(x_1, x_2) + d(x_2, y) \leqslant 2\varepsilon + \max\{d(x_i, x_j) |\ 1 \leqslant i, j \leqslant n\}.$$

The converse is in general false.

Example 5.14. Consider $(\mathbb{R}, d)$ with $d(x, y) = \min\{|x - y|, 1\}$. Then (R, d) is bounded since $d(x, y) \leqslant 1$ for all $x, y \in \mathbb{R}$. But $(\mathbb{R}, d)$ is not totally bounded since it cannot be covered by a finite number of balls of radius $1/2$. Indeed, let S be any finite subset of $\mathbb{R}$, and let x be the largest number in S. If $y \in S$, then $d(x + 1, y) = \min\{|x + 1 - y|, 1\} = 1$ and so there is no $1/2$-net for $\mathbb{R}$.

Theorem 5.15. *Let A be a subset of a metric space (X, d). Then the following conditions are equivalent:*

(a) A is compact.

(b) A is complete and totally bounded.

(c) A has the Heine-Borel property.

Proof. We will show that (a) implies (b), (b) implies (c), (c) implies (a).
(a) implies (b). Let $\{x_n\}$ be a Cauchy sequence in A. We have to show that it converges to a point in A. By compactness of A, some subsequence,

$\{x_{n_k}\}$, converges to $x \in A$. Then $x_n \to x$ by Proposition 4.4. Hence we have proved that A is complete.

Next assume that A is not totally bounded. Then there exists $r > 0$ so that A cannot be covered by finitely many balls of radius r. We construct a sequence $\{x_n\}$ in A which does not have a converging subsequence. Take any $x_1 \in A$. Since $B(x_1, r)$ does not cover A, there is at least one point in $A \setminus B(x_1, r)$. Choose one such point and call it x_2. Having chosen points $x_1, \ldots, x_n$, we choose x_{n+1} so that it belongs to $X \setminus \bigcup_{i=1}^{n} B(x_i, r)$. This is possible since A is not covered by $B(x_1, r), \ldots, B(x_n, r)$. Continuing in this way we get a sequence $\{x_n\}$ such that $d(x_n, x_m) \geq r$ for all n and m. Such a sequence cannot have a convergent subsequence since if $\{x_{n_k}\}$ converges, then it is Cauchy and $d(x_{n_k}, x_{n_m}) < r$ for large k and m. Hence A is not compact; contradiction.

(b) implies (c). Let $\mathcal{U} = \{U_i\}_{i \in I}$ be a collection of open sets covering A. Arguing by contradiction we assume that $\mathcal{U}$ does not contain a finite subcover. Total boundedness of A implies that there is a finite set of closed balls $B_1, \ldots, B_n$ of radius 1 which cover A. If each of the sets $A \cap B_i$ can be covered by a finite number of sets from $\mathcal{U}$, then A can also be covered by a finite subcollection of sets from $\mathcal{U}$. Therefore some $A \cap B_i$, , denoted by B^1, cannot be covered by a finite number of sets from $\mathcal{U}$. Since B^1 is a subset of A and A is totally bounded, B^1 is totally bounded. So let $B_1^1, \ldots, B_m^1$ be a finite set of closed balls of radius $1/2$ which cover B^1. If each $B_i^i \cap B^1$ can be finitely covered by sets from $\mathcal{U}$, the same is true for B^1. Therefore, some $B_j^1 \cap B^1$, denoted by B^2, cannot be covered by a finite number of sets from $\mathcal{U}$. Continuing in this way we obtain a sequence of closed sets B^n such that $\cdots \subset B^n \subset B^{n-1} \subset \cdots \subset B^1$, none of which can be finitely covered and diam $B^n \leqslant 1/n$. From each B^n choose a point x_n. The sequence $\{x_n\}$ is Cauchy since for $n, m \geq k$, $x_n, x_n \in B_k$ and

$$d(x_n, x_m) \leqslant \text{diam } B^k \leqslant 1/k.$$

By completness of A, the sequence $\{x_n\}$ converges, say $x_n \to x$. In fact, $x \in B^k$ for all k since $x_n \in B^k$ for all $n \geq k$ and since B^k is closed. In particular, $x \in A$. Since $\mathcal{U}$ covers A, the point x belongs to some U_i, and therefore, $B(x, \varepsilon) \subset U_i$ for some ε. If $y \in B^n$, then

$$d(x, y) \leqslant d(x, x_n) + d(x_n, y) \leqslant d(x, x_n) + \text{diam } \; B^n \leqslant d(x, x_n) + 1/n.$$

For large n, the right side is less than ε. So for large n, $B^n \subset B(x, \varepsilon)$. Hence $B^n \subset U_i$ which shows that B^n can be finitely covered by sets from $\mathcal{U}$. This contradiction shows that A has the Heine-Borel property.

(c) implies (a). Suppose that A is not compact. Then there exists a sequence $\{x_n\}$ in A with no convergent subsequence in A. Then for every $x \in A$, there exists a ball $B(x, \varepsilon_x)$ which contains x_n for at most finitely many n. Otherwise, there exists x such that for every $r > 0$, $B(x, r)$ contains x_n for infinitely many n. Then, in particular, for every k, $B(x, 1/k)$ contains x_n for infinitely many n. Choose n_1 so that $x_{n_1} \in B(x, 1)$. Since $B(x, 1/2)$ contains x_n for infinitely many n, there is $n_2 > n_1$ such that $x_{n_2} \in B(x, 1/2)$. In this way we construct a subsequence $\{x_{n_k}\}$ such that $x_{n_k} \in B(x, 1/k)$. This implies $x_{n_k} \to x$ contradicting our assumption on $\{x_n\}$. Now the family $\{B(x, \varepsilon_x)\}_{x \in A}$ is an open cover of A from which it is impossible to choose a finite number of balls which will cover A since any finite cover by these balls contains x_n for finitely many n and since A contains x_n for all positive integers. Consequently, A is compact. ■

6 Topological Spaces

Our next aim is to push the process of abstraction a little further and define spaces without distances in which continuous functions still make sense. The motivation behind the definition is the criterion of continuity in terms of open sets. This criterion tells us that a function between metric spaces is continuous provided that the preimage of an open set is open. We make the following definition.

Definition 6.1. *Let X be a set. A* **topology** *on X is a collection $\mathcal{T}$ of subsets of X satisfying the following properties:*

O1 *$\emptyset$ and $X \in \mathcal{T}$.*

O2 *If $\{U_i\}_{i \in I} \subset \mathcal{T}$, then $\bigcup_{i \in I} U_i \in \mathcal{T}$.*

O3 *If $U_1, U_2, \ldots, U_n \in \mathcal{T}$, then $\bigcap_{i=1}^n \in \mathcal{T}$.*

The pair $(X, \mathcal{T})$ is called a **topological space***. If X is a topological space with topology $\mathcal{T}$, we say that a subset U of X is an* **open set** *in X if $U \in \mathcal{T}$.*

Here are some examples of topological spaces.

Example 6.2. Let (X, d) be a metric space. Then the family of open subsets of X with respect to the metric d is a topology on X.

Example 6.3. Let X be any set. The collection of *all* subsets of X, $\mathcal{P}(X)$, is a topology on X. This topology is called the **discrete topology**. Every subset U of X is an open set. On the other extreme, consider X and the collection $\{\emptyset, X\}$. It is also a topology on X, and is called the **indiscrete topology** or the **trivial topology**.

Example 6.4. Let $X = \mathbb{R}$ and let $\mathcal{T}_u$ be the collection of subsets of X consisting of $\emptyset$, $\mathbb{R}$, and the unbounded open intervals $(-\infty, a)$ for all $a \in \mathbb{R}$. Then $\mathcal{T}_u$ is a topology on $\mathbb{R}$. Similarly, we can define a topology $\mathcal{T}_l$ consisting of $\emptyset$, $\mathbb{R}$ and all unbounded intervals (a, ∞), $a \in \mathbb{R}$.

Example 6.5. Let $(X, \mathcal{T})$ be a topological space and $Y \subseteq X$. Then $\mathcal{T}_Y = \{U \cap Y \mid U \in \mathcal{T}\}$ is a topology on Y. It is called the **subspace topology** or **relative topology** induced by $\mathcal{T}$.

Definition 6.6. *Suppose that $\mathcal{T}$ and $\mathcal{T}'$ are two topologies on X. If $\mathcal{T} \subset \mathcal{T}'$ we say that $\mathcal{T}'$ is* **finer** *or* **larger** *than $\mathcal{T}$. In this case we also say $\mathcal{T}$ is* **coarser** *or* **smaller** *than $\mathcal{T}'$. Topologies $\mathcal{T}$ and $\mathcal{T}'$ are* **comparable** *if $\mathcal{T}' \subset \mathcal{T}$ or $\mathcal{T} \subset \mathcal{T}'$.*

Along with a concept of open sets there is the companion concept of closed set. If X is a topological space, then a set $F \subset X$ is **closed** if $F^c = X \setminus F$ is open. By de Morgan's laws, the family of closed sets is closed under arbitrary intersection of closed sets and finite unions. More precisely, the class of closed sets has the following properties:

C1 X and $\emptyset$ are closed.

C2 If F_i is a closed set for every $i \in I$, then $\bigcap_{i \in I} F_i$ is closed.

C3 If $F_1, \dots F_n$ are closed, then $\bigcup_{i=1}^n F_i$ is closed.

Given a subset A of a topological space X, its **closure** is the intersection of all closed subsets of X containing A. The closure of A is denoted by $\overline{A}$. The **interior** of A, denoted by A°, is the union of all open subsets of A.
If $x \in X$, then a set $A \subset X$ is called a **neighbourhood** of x if $x \in A^\circ$.

Basis

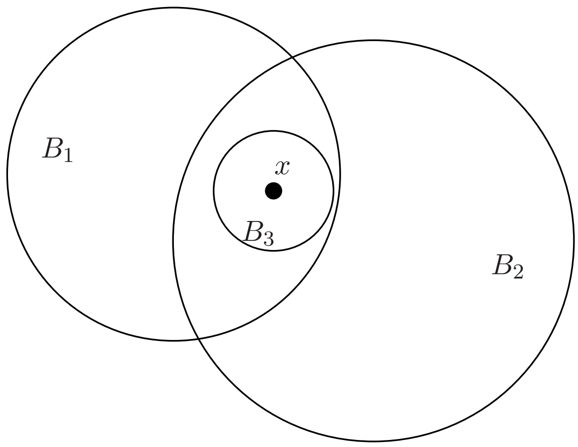
If X is a topological space with topology $\mathcal{T}$, then a **basis** for $\mathcal{T}$ is a collection $\mathcal{B} \subset \mathcal{T}$ such that every member of $\mathcal{T}$, i.e., every open set, is a union of elements of $\mathcal{B}$. (We allow the empty union giving the empty set.)

Example 6.7. The collection of all open balls forms a basis for the topology of any metric space.

Theorem 6.8. *Let X be a set. Then a collection $\mathcal{B}$ of subsets of X is a basis for a topology of X if and only if $\mathcal{B}$ has the following two properties:*

(1) For every $x \in X$, there exists $B \in \mathcal{B}$ such that $x \in B$.

(2) If $B_1, B_2 \in \mathcal{B}$ and $x \in B_1 \cap B_2$, then there exists $B_3 \in \mathcal{B}$ such that $x \in B_3 \subset B_1 \cap B_2$.



Proof. Any basis satisfies (1) since the whole space X is open, and (2) since the intersection of two open sets $B_1 \cap B_2$ is open. Conversely, assume that $\mathcal{B}$ is a collection of subsets of X with properties (1) and (2). Define $\mathcal{T}$ to be the collection of all subsets of X that are unions of sets in $\mathcal{B}$. We shall show that $\mathcal{T}$ is a topology. The condition (1) guarantees that $X \in \mathcal{T}$. Clearly, an arbitrary union of sets in $\mathcal{T}$ belongs to $\mathcal{T}$ in view of the definition of $\mathcal{T}$. Assume that $U, V \in \mathcal{T}$. We have to show that $U \cap V$ is the union of sets in $\mathcal{B}$. Take any $x \in U \cap V$. Since U and V are unions of sets in $\mathcal{B}$, there exist $B_1, B_2 \in \mathcal{B}$ such that $x \in B_1 \subset U$ and $x \in B_2 \subset V$. So $x \in B_1 \cap B_2$, and, in view of (2), there exists $B_x \in \mathcal{B}$ such that $x \in B_x \subset B_1 \cap B_2$. Hence $B_x \subset U \cap V$, and consequently,

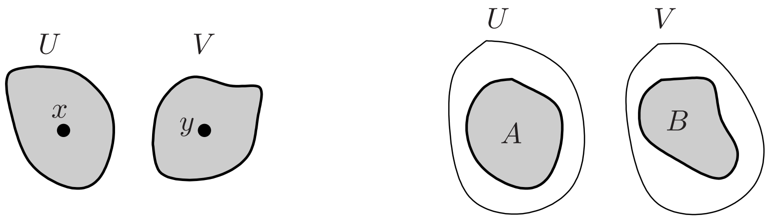
$$U \cap V = \bigcup_{x \in U \cap V} B_x.$$

This shows that $U \cap V \in \mathcal{T}$. ■

Example 6.9. Given two sets X, Y with topologies $\mathcal{S}, \mathcal{T}$ respectively, the *product topology* is defined on the Cartesian product set $X \times Y$ by taking as a basis $\mathcal{B} = \{U \times V : U \in \mathcal{T}, V \in \mathcal{S}\}$.

Hausdorff and normal spaces

Definition 6.10. *A topological space X is called a* **Hausdorff space** *if for every two points $x, y \in X$ such that $x \neq y$, there exist disjoint open sets U and V satisfying $x \in U$ and $y \in V$. A space X is* **normal** *if for each pair A, B of disjoint closed subsets of X, there exist disjoint open sets U and V such that $A \subset U$ and $V \subset V$.*



Continuity

Continuous functions in metric spaces were characterized in terms of open and closed sets. This suggests the definition of continuity in topological spaces.

Definition 6.11. *Let X and Y be topological spaces and let $f : X \to Y$. The map f is* **continuous at a point** *x_0 if for every neighbourhood U of $f(x_0)$ in Y there exists a neighbourhood V of x_0 in X such that $f(V) \subset U$. Global continuity of f is defined in terms of open sets: f is* **continuous** *if $f^{-1}(U)$ is open in X for every open set U in Y.*

Equivalently, f is continuous if $f^{-1}(C)$ is closed in X for every closed set C in Y. If $f : X \to Y$ is bijective and f and f^{-1} are both continuous, f is called a **homeomorphism** and X and Y are said to be **homeomorphic**. We call a property **topological** if it is invariant under homeomorphism.

Elementary properties of continuous functions

(1) If $f : X \to Y$ and $g : Y \to Z$ are continuous maps between topological spaces, then the composition $g \circ f : X \to Z$ is continuous.

(2) If $f : X \to \mathbb{R}$ and $g : X \to \mathbb{R}$ are continuous, then $h : X \to \mathbb{R}^2$ given by $h(x) = (f(x), g(x))$ is continuous.

(3) If A is a subspace of X, then the inclusion map $i : A \to X$ is continuous. This follows from the definition of the topology on the subspace A. If $f : X \to Y$ is continuous, where Y is another topological space, then the restriction map $h = f|_A : A \to Y$ defined by $h(x) = f(x)$ for $x \in A$, is continuous. This follows from (1) using the fact that $h = f \circ i$.

7 Compact Topological Spaces

Theorem 5.15 gives three equivalent characterizations of compactness for metric spaces: the Bolzano-Weierstrass property (every sequence has a convergent subsequence with limit in the space - often called *sequential; compactness*), completeness together with total boundedness, and the Heine-Borel property. In the case of general topological spaces the most useful is the Heine-Borel property. A subset Y of a topological space $(X, \mathcal{T})$ is called **compact** if for every collection $\mathcal{U} = \{U_i\}_{i\in I}$ of open sets of X such that $Y \subseteq \bigcup_{i\in I} U_i$, there is a finite $J \subseteq I$ for which $Y \subseteq \bigcup_{i\in J} U_i$. De Morgan's laws lead to the following characterization of compactness in terms of closed sets.

Definition 7.1. *A family $\{F_i\}_{i\in I}$ of closed subsets of X is said to have the* **finite intersection property** *if $\bigcap_{i\in J} F_i \neq \emptyset$ for all finite $J \subseteq I$.*

Theorem 7.2. *A topological space X has is compact if and only if for every family $\{F_i\}_{i\in I}$ of closed subsets of X having the finite intersection property, $\bigcap_{i\in I} F_i \neq \emptyset$.*

Proof. Assume that X is compact. Let $\{F_i\}_{i\in I}$ be a collection of closed sets having the finite intersection property. Arguing by contradiction assume that $\bigcap_{i \in F_i} = \emptyset$. Writing $U_i = X \setminus F_i$ we have $\bigcup_{i \in I} U_i = \bigcup_{i \in I}[X \setminus F_i] = X \setminus \bigcap_{i \in I} F_i = X$. So $\{U_i\}_{i \in I}$ is an open cover of X. Hence there are $U_{i_1}, \dots, U_{i_k}$ such that $X = U_{i_1} \cup \dots \cup U_{i_k}$. But then $\emptyset = X \setminus X = X \setminus \bigcup_{l=1}^{k} U_{i_l} = \bigcap_{l=1}^{n} F_{i_l}$, contradicting the assumption that $\{F_i\}$ has the finite intersection property. Conversely, suppose that for every collection $\{F_i\}_{i \in I}$ having the finite intersection property we have $\bigcap_{i \in I} F_i \neq \emptyset$. Take any open cover $\{U_i\}_{i \in I}$ of X, and define $F_i = X \setminus U_i$. Then the F_i's are closed and $\bigcap_{i \in I} F_i = \bigcap_{i \in I}[X \setminus U_i] = X \setminus \bigcup_{i \in I} U_i = \emptyset$. So $\{F_i\}$ does not have the finite intersection property (otherwise $\bigcap_{i \in I} F_i \neq \emptyset$). So there is a finite set $J \subseteq I$ such that $\bigcap_{i \in J} F_i = \emptyset$. But then $X = \bigcup_{i \in J}[X \setminus F_i] = \bigcup_{i \in J} U_i$ showing that X is compact. ■

Theorem 7.3. *A closed subspace of a compact topological space is compact.*

Proof. Let K be a closed subset of a topological space X, and let $\{U\}_{i \in J}$ be an open cover of K. Then the collection $\{U\}_{i \in J} \cup \{K^c\}$ is a family of open subsets of X that covers X. Since X is compact, there is a finite subfamily of this family that covers X. The corresponding subfamily of $\{U\}_{i \in J}$ covers K.

■

Theorem 7.4. *If X is a Hausdorff space, then every compact subset of X is closed.*

Proof. Let K be a compact subset of X. Since X is Hausdorff, for every $x \in K^c$ and every $y \in K$, there are disjoint open sets U_{xy} and V_{xy} such that $x \in U_{xy}$ and $y \in V_{xy}$. Then for every $x \in K^c$, $\{V_{xy}\}_{y \in K}$ is an open cover of K. Since K is compact, there exist $y_1, \dots, y_n \in K$ such $K \subseteq \bigcup_{i=1}^{n} V_{xy_i}$. Set $U = \bigcap_{i=1}^{n} U_{xy_i}$. Then U is open, $U \cap K = \emptyset$, and $x \in U$. Thus $x \in U \subseteq K^c$ showing that K^c is open, and consequently, that K is closed. ■

Theorem 7.5. *A compact Hausdorff space is normal.*

Proof. Let A and B be disjoint closed subsets of a compact Hausdorff space. In view of Theorem 7.3, the sets A and B are compact. Proceeding like we did in the proof of the previous theorem, we find for every $x \in B$ disjoint open sets V_x and U_x such that $x \in V_x$ and $A \subseteq U_x$. Then the open sets $\{V_x\}_{x \in B}$ cover B. Consequently, there exist $x_1, \dots, x_n \in B$ such that $B \subseteq V_{x_1} \cup \dots \cup V_{x_n} := V$. Then $U := U_{x_1} \cap \dots \cap U_{x_n}$ is open, $U \cap V = \emptyset$, and $A \subseteq U, B \subseteq V$. ■

Theorem 7.6. *Suppose that $f : X \to Y$ is a continuous map between topological spaces X and Y. If $K \subseteq X$ is a compact set, then $f(K)$ is a compact subset of Y. In particular, if X is compact, then $f(X)$ is compact.*

Proof. Let $\mathcal{U}$ be an open cover of $f(K)$. That is, $\mathcal{U}$ consists of open subsets of Y such that their union contains $f(K)$. The continuity of f implies that for any set $U \in \mathcal{U}$, $f^{-1}(U)$ is an open subset of X. Moreover, the family $\{f^{-1}(U) \mid U \in \mathcal{U}\}$ is an open cover of K. Indeed, if $x \in K$, then $f(x) \in f(K)$, and so $f(x) \in U$ for some $U \in \mathcal{U}$. This implies that $x \in f^{-1}(U)$. Since K is compact, $K \subseteq \bigcup_{i=1}^{n} f^{-1}(U_i)$ for some n. It follows that $f(K) \subseteq \bigcup_{i=1}^{n} U_i$ which proves that $f(K)$ is a compact subset of Y. This completes the proof of the theorem. ∎

Theorem 7.7. *Let f be a continuous bijective function from a compact topological space X to a Hausdorff topological space Y. Then the inverse function $f^{-1} : Y \to X$ is continuous.*

Proof. Denote the inverse function by g, i.e., $g = f^{-1} : Y \to X$. We have to show that $g^{-1}(K)$ is closed in Y for any closed set K in X. Since f is a bijection with inverse g, $g^{-1}(A) = f(A)$ for any subset of Y. So $g^{-1}(K) = f(K)$. Since K is closed and X compact, K is also compact. By the previous result, $f(K)$ is compact in Y and since Y is Hausdorff, $f(K)$ is closed. So $g^{-1}(K)$ is closed in Y, as required. ∎

Example 7.8. Let S^1 be the unit circle in $\mathbb{R}^2$ of radius 1 and centre $(0,0)$. We consider S^1 as a subspace of $\mathbb{R}^2$. Let $f : [0, 2\pi) \to S^1$ be given by $f(x) = (\cos x, \sin x)$ for $x \in [0, 2\pi)$. Show that f is a continuous bijection but the inverse map $f^{-1} : S^1 \to [0, 2\pi)$ is not continuous. Why doesn't this contradict Theorem 7.7?

8 Connected Spaces

A pair of non-empty and open sets U, V of a topological space X is called a **separation** of X if $U \cap V = \emptyset$ and $X = U \cup V$. A topological space X is called **disconnected** if there is a separation of X, and otherwise is called **connected**. A subset Y of X is said to be connected if it is connected as a subspace of X, that is, Y is not the union of two non-empty sets $U, V \in \mathcal{T}_Y$ such that $U \cap V = \emptyset$.

Example 8.1. The set X containing at least two points and considered with the discrete topology is disconnected. However, X with the indiscrete topology is connected.

Example 8.2. The subspace $\mathbb{R} \setminus \{0\}$ of $\mathbb{R}$ is disconnected since $\mathbb{R} \setminus \{0\} = A \cup B$, where $A = \{r \in \mathbb{R} \mid r < 0\}$ and $B = \{r \in \mathbb{R} \mid r > 0\}$. If $X = \mathbb{Q}$ is considered as subspace of $\mathbb{R}$, then X is disconnected since $X = A \cup B$ with $A = \mathbb{Q} \cap (-\infty, r)$ and $B = \mathbb{Q} \cap (r, \infty)$, where r is irrational.

A "2-valued" function is a function from X to $\{0, 1\}$, where $\{0, 1\}$ is considered with discrete topology.

Theorem 8.3. *A space X is connected if and only if every 2-valued continuous function on X is constant. Equivalently, X is disconnected if and only if there exists a 2-valued continuous function from X onto $\{0, 1\}$*

Proof. Suppose that X is connected and $f : X \to \{0, 1\}$ is continuous. Let $A = f^{-1}(\{0\})$ and $B = f^{-1}(\{1\})$. The sets A, B are open, disjoint and $X = A \cup B$. So one of A, B has to be empty. Conversely, assume that every continuous 2-valued function is constant. Assume that $X = A \cup B$, A and B are open, and $A \cap B = \emptyset$. Define

$$f(x) = \begin{cases} 0 & \text{if } x \in A, \\ 1 & \text{if } x \in B. \end{cases}$$

Clearly, the function f is continuous. So f is constant, say $f(x) = 0$ for all $x \in X$. But then $A = X$ and $B = \emptyset$. Hence X is connected as claimed. ∎

Theorem 8.4. *Let $f : X \to Y$ be a continuous function between spaces X and Y. If X is connected, then the image $f(X)$ is connected.*

Proof. Let $g : f(X) \to \{0, 1\}$ be continuous. Then the composition $g \circ f : X \to \{0, 1\}$ is continuous, hence constant since X is connected. Hence g is constant on $f(X)$ and the result follows in view of Theorem 8.3. ∎

Theorem 8.5. *If A is a connected subset of a space X, then $\overline{A}$ is also connected.*

Proof. Let $f : \overline{A} \to \{0, 1\}$ be continuous. Then $f_{|A}$ is continuous, and so, f is constant on A. Say $f = 0$ on A. We claim that $f = 0$ on $\overline{A}$. Suppose $f(x) = 1$ for some $x \in \overline{A}$. The set $\{1\}$ is open in $\{0, 1\}$ and since f is continuous $f^{-1}(\{1\})$ is an open subset of $\overline{A}$. Thus say $f^{-1}(\{1\}) = U \cap \overline{A}$ for some open set U in X. This mean that $f = 1$ on $U \cap \overline{A}$. Since $x \in \overline{A}$, $U \cap A \neq \emptyset$, say $y \in U \cap A$. Then $f(y) = 1$ since $y \in U \cap A \subseteq U \cap \overline{A}$, but one the other hand $f(y) = 0$ since $f = 0$ on A. Therefore, $\overline{A}$ is connected as claimed. ■

Example 8.6. The union of connected subspaces does not have to be connected. Consider $\mathbb{R}$ with the usual topology. Then the sets $(-\infty, 0)$ and $(0, \infty)$ are connected subspaces of $\mathbb{R}$, but the union $(-\infty, 0) \cup (0, \infty) = \mathbb{R} \backslash \{0\}$ is disconnected.

Theorem 8.7. *If $\{A_i\}_{i \in I}$ is a family of connected subsets of X such that $\bigcap_{i \in I} A_i \neq \emptyset$, then $A = \bigcup_{i \in I} A_i$ is connected.*

Proof. Let $f : A \to \{0, 1\}$ be continuous. Then $f_{|A_i}$ is continuous for every i, so it is constant. Since $\bigcap_{i \in I} A_i \neq \emptyset$, we must have the same constant on every A_i. Hence f is constant and A is connected. ■

As an application of this theorem we have the following

Theorem 8.8. *Suppose that for any two points in a space X there exists a connected subspace of X containing these two points. Then X is connected.*

Proof. Fix a point $a \in X$. For $b \in X$, denote by $C(b)$ a connected subspace of X containing a and b. Then $X = \bigcup_{b \in X} C(b)$. Since $a \in \bigcap_{b \in X} C(b)$, the result follows from the previous theorem. ■

Let $x \in X$ and let C_x be the union of all the connected subsets of X containing x. Each C_x is called a **component** (or **connected component**) of X.

Proposition 8.9. *Let C_x be the connected component of X containing x. Then*

(a) for each $x \in X$, C_x is connected and closed; and

(b) for any two $x, y \in X$, either $C_x = C_y$ or $C_x \cap C_y = \emptyset$.

Proof. The set C_x is connected in view of Theorem 8.7, and by Theorem 8.5, $\overline{C_x}$ is connected. Hence by the definition of C_x, $\overline{C_x} \subset C_x$, so $C_x = \overline{C_x}$ and C_x is closed. If $C_x \cap C_y \neq \emptyset$, then $C_x \cup C_y$ is connected by Theorem 8.7. So again by the definition of C_x, $C_x \cup C_y \subset C_x$. Hence $C_y \subset C_x$. Similarly, $C_x \subset C_y$, so $C_x = C_y$ as required. ■

Example 8.10. If X is equipped with the discrete topology, then every subset of X is open and closed. Hence the connected components of X are sets consisting of one point.

Example 8.11. X, Y are connected spaces if and only if $X \times Y$, with the product topology, is connected. In fact, if X is disconnected by disjoint open sets U, V , then $X \times Y$ is disconnected by $U \times Y . V \times Y$. Conversely, if $X \times Y$ is disconnected by disjoint open sets W, Z, then X (or Y) is disconnected, by considering $X \times \{y\}$ intersecting W, Z. For more details, see Theorem 9.6.

Next we shall determine the connected subsets of $\mathbb{R}$. By an **interval** $I \subset \mathbb{R}$ we mean a subset of $\mathbb{R}$ having the following property: if $x, y \in I$ and $x \leqslant z \leqslant y$, then $z \in I$.

Theorem 8.12. *A subset of* $\mathbb{R}$ *is connected if and only if it is an interval.*

Proof. Suppose that $J \subset \mathbb{R}$ is not an interval. Then there are $x, y \in J$ and $z \notin J$ with $x < z < y$. Then define $A = (-\infty, z) \cap J$ and $B = (z, \infty) \cap J$. Clearly, A, B are disjoint, non-empty, relatively open, and $A \cup B = J$. So J is not connected. Conversely, suppose that J is an interval. We will show that J is connected. Let $f : J \to \{0, 1\}$ be continuous, and suppose that f is not constant. Then there are x_1 and $y_1 \in J$ such that $f(x_1) = 0$ and $f(y_1) = 1$. For simplicity assume that $x_1 < y_1$. Let a be the midpoint of $[x_1, y_1]$. If $f(a) = 0$, then set $x_2 = 0$ and $y_2 = y_1$, and otherwise, $x_2 = x_1$ and $y_2 = a$. So $x_1 \leqslant x_2 \leqslant y_2 < y_1$, $|x_2 - y_2| \leqslant 2^{-1}|x_1 - y_1|$, and $f(x_i) \neq f(y_i)$. Iterating this procedure we find sequences $\{x_n\}$ and $\{y_n\}$ with the following properties: $x_1 \leqslant x_2 \leqslant \cdots \leqslant x_n < y_n \leqslant \cdots \leqslant y_1$, $|x_n - y_n| \leqslant 2^{-1}|x_{n-1} - y_{n-1}| \leqslant 2^{n-1}|x_1 - y_1|$, and $f(x_n) = 0$, $f(y_n) = 1$. Since $\mathbb{R}$ is complete, $\{x_n\}$ converges to some z, and since $|x_n - y_n| \to 0$, $y_n \to z$. Clearly, $z \in J$. Hence $0 = \lim_n f(x_n) = f(z) = \lim_n f(y_n) = 1$, a contradiction. So f is constant, and this implies that J is connected. ■

We can apply the last theorem to analyze the structure of open subsets of $\mathbb{R}$. We claim that any open set $U \subset \mathbb{R}$ is a countable union of pairwise disjoint open intervals. Indeed, let $x \in U$ and let I_x be the connected component of U containing x. Thus, I_x is an interval. If $y \in I_x$, then there is $\delta > 0$ such that $(y - \delta, y + \delta) \subset U$ since U is open. Hence $I_x \cup (y - \delta, y + \delta)$ is connected and since I_x is a connected component, $(y - \delta, y + \delta) \subset I_x$. So I_x is an open interval, and U is a union of open intervals (its components). Since each

must contain a different rational number, U is at most a countable union of disjoint open intervals.
Here is an important application of Theorem 8.12.

Theorem 8.13 (**Intermediate Value Theorem**). *Let $f : X \to \mathbb{R}$ be a continuous function defined on a connected space X. Then for any $x, y \in X$ and any $r \in \mathbb{R}$ such that $f(x) \leqslant r \leqslant f(y)$ there exists $c \in X$ such that $f(c) = r$.*

Proof. The set $f(X)$ is a connected subset of $\mathbb{R}$. Hence $f(X)$ is an interval, and since $f(x), f(y) \in f(X)$, it has to contain r. ■

Definition 8.14. *A space X is called* **path connected** *if for any two points p and $q \in X$, there exists a continuous function $f : [0,1] \to X$ such that $f(0) = p$ and $f(1) = q$. The function f is called a* **path** *from $f(0)$ to $f(1)$.*

If X is path connected, then X is connected but the converse is false in general as the following example shows.

Example 8.15. Write $X = \{(t, \sin(\pi/t)) \mid t \in (0,2]\} \subset \mathbb{R}^2$. Let $\varphi : \mathbb{R}^2 \to \mathbb{R}$ be the projection onto the first coordinate, that is, $\varphi(x,y) = x$. Then $\varphi : X \to (0,2]$ is a homeomorphism and since $(0,2]$ is connected so is X. Therefore, $\overline{X} = (\{0\} \times [-1,1]) \cup X = J \cup X$ is connected, where we abbreviated $J = \{0\} \times [-1,1]$. We shall show that $\overline{X}$ is not path connected. Arguing by contradiction assume that $f : [0,1] \to \overline{X}$ is a continuous path in $\overline{X}$ such that $f(0) \in J$ and $f(1) \in X$. Consider $f^{-1}(J)$. It is closed in $[0,1]$ and contains 0. Let $a = \sup\{t \in [0,1], t \in f^{-1}(J)\}$. Since $f(1) \in X$, $a < 1$. Since f is continuous, there exists $\delta > 0$ such that $f(a+\delta) \in X$. Write $f(t) = (x(t), y(t))$. Then $x(a) = 0$, $x(t) > 0$, and $y(t) = \sin(\pi/x(t))$ for $t \in (0, a+\delta]$. For every large n we can find r_n such that $0 < r_n < x(a+1/n)$ and $\sin(\pi/r_n) = (-1)^n$. Since the function x is continuous by the Intermediate Value Theorem there is $t_n \in (a, a+\delta]$ such that $x(t_n) = r_n$ and $y(t_n) = (-1)^n$. So $t_n \to a$ but $y(t_n)$ does not converge contradicting the fact that f is continuous. Hence $\overline{X}$ is not path connected.

Theorem 8.16. *An open subset of a normed space is connected if and only if it is path connected.*

Proof. We give a sketch of this important result. As above, we know that path connected spaces are always connected. So let's assume that a normed space X has a connected open set U and show it is path connected. The idea is that an open ball in a normed space is always path connected, by just connecting each point back to the centre of the ball. But then we can show that the largest open path connected subset of U is in fact all of U.

To do this, note that if two open path connected sets A, B have a non empty intersection, then their union is path connected. Given $a \in A, b \in B$ construct a path from a to b by going through an intermediate point $c \in A \cap B$. We can apply this to any proper open path connected subset A of U together with an open ball B inside U centred at a point of the closure of V. We leave this as an exercise to check the details.

■

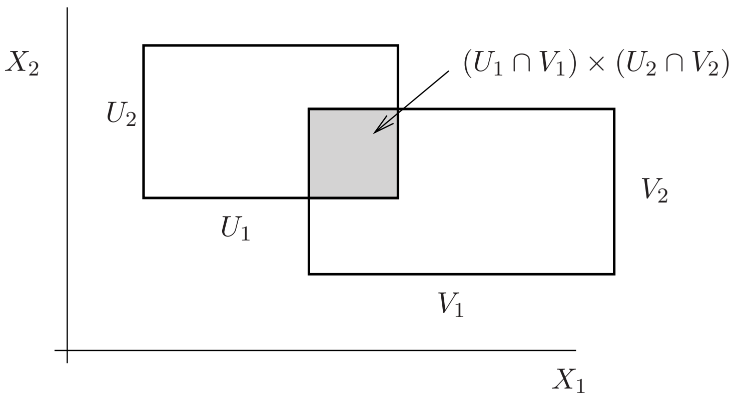
9 Product Spaces

We define a topology on a finite product of topological spaces. Consider a finite collection $X_1, \ldots, X_n$ of topological spaces. The **product topology** on the product $X = X_1 \times \cdots \times X_n$ is the topology for which a basis of open sets is given by "rectangles"

$$\{U_1 \times \cdots \times U_n \mid U_j \text{ is open in } X_j \text{ for } 1 \leqslant j \leqslant n\}. \tag{11}$$

Observe that the intersection of two such sets is again a set of this form. Indeed,

$$(U_1 \times \cdots \times U_n) \cap (V_1 \times \cdots \times V_n) = (U_1 \cap V_1) \times \cdots \times (U_n \cap V_n).$$



Consequently, the family (11) forms a basis. Let $\pi_j : X \to X_j$ be the projection of X onto the jth factor, defined by

$$\pi_j(x_1, \ldots, x_n) = x_j, \quad (x_1, \ldots, x_n) \in X.$$

For an open set $U_j \subset X_j$, we have

$$\pi_j^{-1}(U_j) = X_1 \times \cdots \times X_{j-1} \times U_j \times X_{j+1} \times \cdots \times X_n$$

which is a basic open set. Hence each projection π_j is continuous.

Theorem 9.1. *Let X be the product of the topological spaces $X_1, \ldots, X_n$, and let π_j be the projection of X onto X_j. The product topology for X is the smallest topology for which each of the projections π_j is continuous.*

Proof. Let $\mathcal{T}$ be another topology on X such that the projections π_j are $\mathcal{T}$-continuous. Take open sets $U_j \subset X_j$, $1 \leqslant j \leqslant n$. Then each $\pi_j^{-1}(U_j)$ belongs to $\mathcal{T}$ since π_j is $\mathcal{T}$-continuous. Since

$$\pi_1^{-1}(U_1) \cap \cdots \cap \pi_n^{-1}(U_n) = U_1 \times \cdots \times U_n$$

the basic set $U_1 \times \cdots \times U_n$ belongs to $\mathcal{T}$ and $\mathcal{T}$ includes the product topology. ■

Call a function f from one topological space to another **open** if it maps open sets onto open sets.

Theorem 9.2. *Let X be the product of the topological spaces $X_1, \ldots, X_n$. Then each projection π_j of X onto X_j is open.*

Proof. Let $U = U_1 \times \cdots \times U_n$ be a basic open set in X. Then $\pi_j(U) = U_j$, and since the maps preserve unions, the image of any open set is open. ■

Theorem 9.3. *Let Y be a topological space and let f be a continuous map from Y to the product $X = X_1 \times \cdots \times X_n$. Then f is continuous if and only if $\pi_j \circ f$ is continuous for all $1 \leqslant j \leqslant n$.*

Proof. If f is continuous, the $\pi_j \circ f$ is continuous as a composition of continuous maps. Conversely, suppose that $\pi_j \circ f$ is continuous for all $1 \leqslant j \leqslant n$. Take a basic open set $U = U_1 \times \cdots \times U_n$ in X. Then

$$f^{-1}(U) = (\pi_1 \circ f)^{-1}(U_1) \cap \cdots \cap (\pi_n \circ f)^{-1}(U_n)$$

is a finite intersection of open sets and hence is open. Since the inverses of functions preserve unions, the inverse image of any open set is open, and consequently, f is continuous. ■

We next study which properties of topological spaces are valid for the product $X = X_1 \times \cdots \times X_m$ whenever they hold for $X_1, \ldots, X_m$.

Theorem 9.4. *Let X be the product of Hausdorff spaces $X_1, \dots, X_n$. Then X is Hausdorff.*

Proof. Take two different points $x = (x_1, \dots, x_n)$, $y = (y_1, \dots, y_n)$, and choose an index i so that $x_i \neq y_i$. Since X_i is Hausdorff, there exist open sets U_i and V_i in X_i such that $U_i \cap V_i = \emptyset$. Then $\pi_i^{-1}(U_i)$ and $\pi_i^{-1}(V_i)$ are open and disjoint sets containing x and y, respectively. Consequently, X is Hausdorff as required.

■

Theorem 9.5. *Let X be the product of path-connected spaces $X_1, \dots, X_n$. Then X is path-connected.*

Proof. Take two points $x = (x_1, \dots, x_n)$, $y = (y_1, \dots, y_n)$ in X. Since each X_j is path-connected, for each $1 \leqslant j \leqslant n$ there exists a path $\gamma_j : [0,1] \to X_j$ from x_j to y_j. Define $\gamma : [0,1] \to X$ by setting

$$\gamma(t) = (\gamma_i(t), \dots, \gamma_n(t)), \quad t \in [0,1].$$

Then γ is a path connecting x with y. So X is path-connected. ■

To study connectedness of the product of connected spaces we will need the following fact. Fix points $x_2 \in X_2, \dots, x_n \in X_n$ and define a map $h : X_1 \to X$ by setting $h(x_1) = (x_1, \dots, x_n)$. Then h is a homeomorphsim of X_1 onto the "slice" $X_1 \times \{x_2\} \times \cdots \times \{x_n\}$ of X. Indeed, if $U = U_1 \times \cdots \times U_n$ is a basic open set in X, then $h^{-1}(U) = U_1$ is open so that h is continuous. Since the inverse of h is equal to $\pi_1|_{X_1 \times \{x_2\} \times \cdots \times \{x_n\}}$, h^{-1} is continuous and h is a homeomorphism. Similarly, for each fixed j and fixed points $x_i \in X_i$, $i \neq j$, the map $X_j \to \{x_1\} \times \cdots \{x_{j-1}\} \times X_j \times \{x_{j+1}\} \times \cdots \times \{x_n\}$ is a homeomorphism.

Theorem 9.6. *Let X be the product of connected spaces $X_1, \dots, X_n$. Then X is connected.*

Proof. We prove the theorem for the product of two connected spaces X_1 and X_2. We apply Theorem 8.8. Take any two points $a = (a_1, a_2), b = (b_1, b_2) \in X$ and consider sets $C_1 = \{(x, b_2) \in X \mid x \in X_1\}$ and $C_2 = \{(a_1, y) \in X \mid y \in X_2\}$. By the above remark, the sets C_1, C_2 are connected. Then, in view of Theorem 8.7, $C = C_1 \cup C_2$ is connected since $C_1 \cap C_2 = \{(a_1, b_2)\}$. Applying Theorem 8.8, the space X is connected since $a, b \in C$. ■

To study compactness of the product of compact spaces we need the following lemma.

Lemma 9.7. *Let Y be a topological space and let $\mathcal{B}$ be a basis for the topology of Y. If every open cover of Y by sets in $\mathcal{B}$ has a finite subcover, then Y is compact.*

Proof. Let $\{U_i\}_{i\in I}$ be an open cover of Y. For each $y \in Y$, choose $V_y \in \mathcal{B}$ and an index j so that $y \in V_y \subset U_j$. The family $\{V_y\}_{y\in Y}$ forms an open cover of Y by sets belonging to $\mathcal{B}$. In view of the assumption, there exists a finite number of the V_y's that cover Y. Since each of these V_y's is contained in at least one of the U_j's, we obtain a finite number of U_j's that cover Y. Hence Y is compact.

■

Theorem 9.8 (Tychonoff's Theorem for the finite product). *Let X be the product of compact spaces $X_1, \ldots, X_n$. Then X is compact.*

Proof. We consider only the product of two compact spaces X_1 and X_2. let $\mathcal{R}$ be a cover of $X_1 \times X_2$ by basic open sets of the form $U \times V$, U open in X_1 and V open in X_2. In view of Lemma 9.7, it is enough to show that $\mathcal{R}$ has a finite subcover. Fix $z \in X_2$. The slice $X_1 \times \{z\}$ is compact. Hence there are finitely many sets $U_1 \times V_1, \ldots, U_n \times V_n$ in $\mathcal{R}$ covering the slice $X \times \{z\}$. We may assume that $z \in V_j$ for all $1 \leqslant j \leqslant n$, by throwing out products where the second factor does not contain z. The set $V(z) = V_1 \cap \cdots \cap V_n$ is an open set containing z, and the set $\pi_2^{-1}(V(z))$ is covered by sets $U_j \times V_j$, $1 \leqslant j \leqslant n$. The collection $\{V(z)\}_{z\in X_2}$ is an open cover of X_2, and since X_2 is compact, $X_2 = V(z_1) \cup \cdots \cup V(z_l)$ for some finite number of points $z_j \in X_2$. Then $X = \pi_2^{-1}(V(z_1)) \cup \cdots \cup \pi_2^{-1}(V(z_l))$. Each $\pi_2^{-1}(V(z_j))$ is covered by finitely many sets in $\mathcal{R}$. Consequently, X can be covered by finitely many sets in $\mathcal{R}$, and, in view of Lemma 9.7, X is compact.

■

Compactness in function spaces: Ascoli-Arzela theorem

Next we study compact subsets of the space of continuous functions. Let X be a compact topological space and (M, σ) a complete metric space. By $C(X, M)$ we denote the set of all continuous functions from X to M. We consider $C(X, M)$ with the metric

$$d(f, g) = \sup\{\sigma(f(x), g(x)) \mid x \in X\}$$

Definition 9.9. *Let X be a topological space and (M, σ) a metric space, and let $\mathcal{F}$ be a family of functions from X to M. The family $\mathcal{F}$ is called* **equicontinuous at** $x \in X$ *if for every $\varepsilon > 0$ there exists a neighbourhood U_ε of x such that*

$$\sigma(f(y), f(x)) < \varepsilon \text{ for all } y \in U_\varepsilon \text{ and all } f \in \mathcal{F}.$$

The family $\mathcal{F}$ is called **equicontinuous** *if it is equicontinuous at each $x \in X$.*

Example 9.10. Consider two metric spaces (X, ρ) and (M, σ). Given $M > 0$ let $\mathcal{F}$ be a set of all functions $f : X \to Y$ such that

$$\sigma(f(x), f(y)) \leqslant M\rho(x, y)$$

for all $x, y \in X$. Then $\mathcal{F}$ is an equicontinuous family of functions. Indeed, take $\varepsilon > 0$ and let $U_\varepsilon = B(x, \varepsilon/M)$. Then if $y \in U_\varepsilon$ and $f \in \mathcal{F}$, we have

$$\sigma(f(x), f(y)) \leqslant M\rho(x, y) < M \cdot \varepsilon/M = \varepsilon.$$

Theorem 9.11 (**Ascoli-Arzela Theorem**)**.** *Let X be a compact topological space and let (M, σ) be a complete metric space. Let $\mathcal{F} \subset C(X, M)$. Then the closure $\overline{\mathcal{F}}$ is compact in $C(X, M)$ if and only if the two following conditions hold:*

(1) $\mathcal{F}$ is equicontinuous.

(2) For every $x \in X$, the set $\mathcal{F}(x) = \{f(x) \mid f \in \mathcal{F}\}$ has a compact closure in M.

Proof. Since $C(X, M)$ is a complete metric space, $\overline{\mathcal{F}}$ is compact if and only if $\mathcal{F}$ is totally bounded. Assume first that the conditions (1) and (2) are satisfied. In view of the above remark we have to show that $\mathcal{F}$ is totally bounded. Given $\varepsilon > 0$, for each $x \in X$ there exists an open neighbourhood $V(x)$ such that if $y \in V(x)$, then $\sigma(f(x), f(y)) < \varepsilon$ for all $f \in \mathcal{F}$. Since $\{V(x)\}_{x \in X}$ is an open cover of X and X is compact by assumption, there exist a finite number of points $x_1, \ldots, x_n$ such that $V(x_1), \ldots, V(x_n)$ cover X. The sets $\mathcal{F}(x_j)$ are totally bounded in M, hence so is the union $\mathcal{S} = \mathcal{F}(x_1) \cup \cdots \cup \mathcal{F}(x_n)$. Let $\{a_1, \ldots, a_m\}$ be an ε-net for $\mathcal{S}$. For every map $\varphi : \{1, \ldots, n\} \to \{1, \ldots, m\}$ denote by

$$B_\varphi = \{f \in \mathcal{F} \mid \sigma(f(x_j), a_{\varphi(j)}) < \varepsilon \text{ for all } j = 1, \ldots, n\}.$$

Observe that there is only a finite number of sets B_φ and every $f \in \mathcal{F}$ belongs to one of such sets. Moreover, if $f, g \in \mathcal{F}$, then

$$\begin{aligned}\sigma(f(y), g(y)) &\leqslant \sigma(f(y), f(x_k)) + \sigma(f(x_k), a_{\varphi(k)}) \\ &\quad + \sigma(a_{\varphi(k)}, g(x_k)) + \sigma(g(x_k), g(y)) \\ &< \varepsilon + \varepsilon + \varepsilon + \varepsilon = 4\varepsilon\end{aligned}$$

for all $y \in V(x_k)$. So if $f, g \in B_\varphi$, then $d(f, g) < 4\varepsilon$. Consequently, the diameter of B_φ is less than 4ε, and since there are finitely many such B_φ and they cover $\mathcal{F}$, the set $\mathcal{F}$ is totally bounded.
Conversely, assume that $\mathcal{F}$ is totally bounded. Note that the mapping $\Psi : \mathcal{F} \to M$ given by $\Psi(f) = f(x)$ is distance decreasing, i.e.,

$$\sigma(\Psi(f), \Psi(g)) = \sigma(f(x), g(x)) \leqslant d(f, g).$$

It follows that for every $x \in X$, the set $\mathcal{F}(x) \subset M$ is totally bounded and (2) holds. To see that (1) holds, let $\varepsilon > 0$ and let $f_1, \ldots, f_n$ be an ε-net of $\mathcal{F}$. Given $x \in X$ we find open neighbourhood $V(x)$ of x such that $\sigma(f_j(x), f_j(y)) < \varepsilon$ for all $y \in V(x)$ and all $j = 1, \ldots, n$. Then if $f \in \mathcal{F}$ choose an index j so that $d(f, f_k) < \varepsilon$. It follows that if $y \in V(x)$, then

$$\begin{aligned}\sigma(f(x), f(y)) &\leqslant \sigma(f(x), f_j(x)) + \sigma(f_j(x), f_j(y)) + \sigma(f_j(y), f(y)) \\ &< \varepsilon + \varepsilon + \varepsilon = 3\varepsilon.\end{aligned}$$

Therefore, the family $\mathcal{F}$ is equicontinuous at x, and since x was an arbitrary point of X, $\mathcal{F}$ is equicontinuous as required. ■

Corollary 9.12. *Let X be a compact topological space and Y a compact metric space. Let $\mathcal{F} \subseteq C(X, Y)$ be an equicontinuous family. Then every sequence in $\mathcal{F}$ has a uniformly convergent subsequence.*

Definition 9.13. *A family $\mathcal{F}$ of maps $f : X \to Y$, where Y is a metric space is called* **pointwise bounded** *if $\{f(x) \mid f \in \mathcal{F}\}$ is bounded in Y for every $x \in X$.*

Lemma 9.14. *Assume that X is a compact metric space and let $\mathcal{F}$ be an equicontinuous and pointwise bounded family in $C(X)$. Then there is a constant M such that $f(X) \subset [-M, M]$ for all $f \in \mathcal{F}$.*

Proof. For every $x \in X$, there exists M_x such that $|f(x)| \leqslant M_x$ for all $f \in \mathcal{F}$. Since $\mathcal{F}$ is equicontinuous, for each x there is an open set U_x such that $|f(x) - f(y)| \leqslant 1$ for all $f \in \mathcal{F}$ and $y \in U_x$. Then

$$|f(y)| \leqslant |f(y) - f(x)| + |f(x)| \leqslant 1 + M_x = K_x$$

for all $y \in U_x$. The sets U_x form an open covering of X and since X is compact, there exists a finite subcovering $U_{x_1}, \ldots, U_{x_n}$. Set now $M = \max\{K_{x_1}, \ldots, K_{x_n}\}$. Then $|f(x)| \leqslant M$ for all $x \in M$. ■

Corollary 9.15 (**Ascoli-Arzela Theorem, classical version**). *Let X be a compact topological space. Assume that $\mathcal{F}$ is a pointwise bounded and equicontinuous subset of $C(X)$. Then every sequence in $\mathcal{F}$ has a uniformly convergent subsequence.*

Proof. By Lemma 9.14, the set $\mathcal{F}$ is uniformly bounded, that is, $|f(x)| \leqslant M$ for all $f \in \mathcal{F}$ and $x \in X$. Set $Y = [-M, M]$. Then Y is compact in $\mathbb{R}$, and $\mathcal{F}$ is a subset of $C(X, Y)$. So the corollary follows from Corollary 9.12. ■

10 Urysohn's and Tietze's Theorems

We show the existence of continuous functions on normal topological spaces. We start with the following characterisation of normal spaces.

Lemma 10.1. *A topological space X is normal if and only if for every closed subset $A \subset X$ and every open subset $B \subset X$ containing A , there exists an open set U such that $A \subset U \subset \overline{U} \subset B$.*

Proof. Assume first that X is normal and A and B are as above. Then the sets A and $X \setminus B$ are closed and disjoint. So, in view of the normality of X, there exist open disjoint sets U and V such that $A \subset U$ and $X \setminus B \subset V$. Then $\overline{U} \subset X \setminus V \subset B$, so that U has the required properties.
Conversely, let A and B be closed disjoint subsets of X. Then $V = X \setminus B$ is open and $A \subset V$. By assumption there exists an open set U such that $A \subset U \subset \overline{U} \subset V$. Then U and $X \setminus \overline{U}$ are disjoint open sets satisfying $A \subset U$ and $B \subset X \setminus \overline{U}$. So X is normal and the proof is completed. ■

Theorem 10.2 (Urysohn's Lemma)**.** *Let A and B be disjoint closed subspaces of a normal space X. Then we can find a continuous function $f : X \to [0, 1]$ such that $f(a) = 0$ for all $a \in A$ and $f(b) = 1$ for all $b \in B$.*

Proof. For the proof recall that a dyadic rational number is a number which can be written in the form $p = \dfrac{m}{2^n}$ with n, m being integers. Set $V = X \setminus B$, an open set which contains A. By Lemma 10.1, there exists an open set $U_{1/2}$ such that

$$A \subset U_{1/2} \subset \overline{U}_{1/2} \subset V.$$

Applying Lemma 10.1 again to the open set $U_{1/2}$ containing A and to the open set V containing $\overline{U}_{1/2}$, we obtain open sets $U_{1/4}$ and $U_{3/4}$ such that

$$A \subset U_{1/4} \subset \overline{U}_{1/4} \subset U_{1/2} \subset \overline{U}_{1/2} \subset U_{3/4} \subset \overline{U}_{3/4} \subset V.$$

Continuing in this way, we associate to every such dyadic rational number $p \in (0,1)$ an open subset $U_p \subset X$ having the following properties

$$\overline{U_p} \subset U_q, \qquad 0 < p < q < 1, \tag{12}$$

$$A \subset U_p, \qquad 0 < p < 1, \tag{13}$$

$$U_p \subset V, \qquad 0 < p < 1. \tag{14}$$

Next we shall construct a function f which is continuous and such that the sets ∂U_p are level sets of f on which f assumes the value p. Define $f(x) = 0$ if $x \in U_p$ for all $p > 0$ and $f(x) = \sup\{p|\ x \notin U_p\}$ otherwise. Clearly, $0 \leqslant f \leqslant 1$, $f(x) = 0$ for all $x \in A$ and $f(x) = 1$ for all $x \in B$. It remains to show that f is continuous. Take $x \in X$. We only consider the case that $0 < f(x) < 1$. (The remaining cases $f(x) = 0$ and $f(x) = 1$ are left as an exercise.) Let $\varepsilon > 0$ and choose dyadic rationals p and q such that $0 < p, q < 1$ and

$$f(x) - \varepsilon < p < f(x) < q < f(x) + \varepsilon.$$

Then $x \notin U_r$ for dyadic rationals r between p and $f(x)$ so that, in view of (12), $x \notin \overline{U}_p$. On the other hand, $x \in U_q$. So $W = U_q \setminus \overline{U}_p$ is an open neighbourhood of x. Then $p \leqslant f(y) \leqslant q$ for any $y \in W$ which shows that $|f(x) - f(y)| < \varepsilon$ for all $y \in W$. Hence f is continuous and the proof is completed. ■

Theorem 10.3 (Tietze's extension theorem)**.** *Let A be a closed subset of a normal space X and let f be a bounded continuous real valued function on A. Then there exists a bounded continuous function $h : X \to \mathbb{R}$ such that $f = h$ on A.*

Proof. Set $a_0 = \sup\{|f(a)||\ a \in A\}$. Since f is bounded, the number a_0 is finite. Define sets

$$B_0 = \{a \in A|\ f(a) \leqslant -a_0/3\} \qquad C_0 = \{a \in A|\ f(a) \geq a_0/3\}.$$

Since f is continuous and A is closed, the sets B_0 and C_0 are closed and disjoint subsets of X. Taking a linear combination of the function from Urysohn's lemma and a suitable constant function we find a continuous function $g_0 : X \to \mathbb{R}$ satisfying $-a_0/3 \leqslant g_0 \leqslant a_0/3$ on X, $g_0 = -a_0/3$ on B_0 and $g_0 = a_0/3$ on C_0. Thus,

$$|g_0| \leqslant a_0/3 \quad \text{on } X$$
$$|f - g_0| \leqslant 2a_0/3 \quad \text{on } A.$$

Iterating this process we construct the sequence of functions $\{g_n\}$ satisfying

$$|g_n| \leqslant 2^n a_0/3^{n+1} \quad \text{on } X \tag{15}$$
$$|f_n| = |f - g_0 - g_1 - \cdots - g_n| \leqslant 2^n a_0/3^n \quad \text{on } A. \tag{16}$$

Indeed, suppose that the functions $g_0, \ldots, g_{n-1}$ have been constructed. To construct g_n, set

$$a_{n-1} = \sup\{|f(a) - g_0(a) - g_1(a) - \cdots - g_{n-1}(a)| \mid a \in A\},$$

and repeat the above argument with a_{n-1} replacing a_0 and $f - g_0 - g_1 - \cdots - g_{n-1}$ replacing f. This gives the function g_n such that

$$|g_n| \leqslant a_{n-1}/3$$
$$|f - g_0 - g_1 - \cdots - g_n| \leqslant 2a_{n-1}/3 \quad \text{on } A.$$

Since $a_{n-1} \leqslant 2^n a_0/3^n$, the function g_n satisfies (15)-(16). Set

$$h_n = g_0 + \cdots + g_n, \quad n \geq 1.$$

If $n > m$, then

$$|h_n - h_m| = |g_{m+1} + \cdots + g_n| \leqslant \left(\left(\frac{2}{3}\right)^{m+1} + \cdots + \left(\frac{2}{3}\right)^{n}\right) \cdot \frac{a_0}{3}$$
$$\leqslant \left(\frac{2}{3}\right)^{m+1} \cdot a_0.$$

Consequently, $\{h_n\}$ is Cauchy in $C(X, \mathbb{R})$. Hence there exists a continuous function $h : X \to \mathbb{R}$ such that $h_n \to h$. In addition,

$$|h| = |\lim h_n| = \lim |h_n| \leqslant \lim \sum_{k=0}^{n} |g_k| \leqslant \frac{a_0}{3} \sum_{k=1}^{\infty} \left(\frac{2}{3}\right)^n = a_0,$$

so that h is bounded. Finally, in view of (16), $|f - h| = \lim |f - h_n| \leqslant \lim 2^n a_0/3^{n+1} \to 0$ on A, so that $f = h$ on A. The proof is completed.

■

11 Appendix

Sets

A **set** is considered to be a collection of objects. The objects of a set A are called **elements** (or **members**) of A. If x is an element of a set A we write $x \in A$, and if x is not an element of A we write $x \notin A$. Two sets A and B are called **equal**, $A = B$, if A and B have the same elements. A set A is a **subset** of a set B, written $A \subset B$, if every element of A is also an element of B. The **empty set** $\emptyset$ has no elements; it has the property that it is a subset of any set, that is, $\emptyset \subset A$ for any set A. Given two sets A and B we define:

(a) the **union** $A \cup B$ of A and B as the set

$$A \cup B = \{x \mid x \in A \,\text{or}\, x \in B\};$$

(b) the **intersection** $A \cap B$ of A and B as the set

$$A \cap B = \{x \mid x \in A \,\text{and}\, x \in B\};$$

(c) the set **difference** $A \setminus B$ (or $A - B$) of A and B as the set

$$A \setminus B = \{x \mid x \in A \,\text{and}\, x \notin B\}.$$

Sets A and B are called **disjoint** if $A \cap B = \emptyset$. The concept of union and intersection of two sets extends to unions and intersections of arbitrary families of sets. By a **family of sets** we mean a nonempty set $\mathcal{F}$ whose elements are sets themselves. If $\mathcal{F}$ is a family of sets, then

$$\bigcup_{A \in \mathcal{F}} A = \{x \mid x \in A \text{ for some } A \in \mathcal{F}\}$$
$$\bigcap_{A \in \mathcal{F}} A = \{x \mid x \in A \text{ for all } A \in \mathcal{F}\}.$$

If $\mathcal{F} = \{A_i \mid i \in I\}$ is a family of sets A_i indexed by elements of a set I, then we also write

$$\bigcup_{i \in I} A_i = \{x \mid x \in A_i \text{ for some } i \in I\}$$
$$\bigcap_{i \in I} A_i = \{x \mid x \in A_i \text{ for all } i \in I\}.$$

When it is understood that all sets under consideration are subsets of a fixed set X, then the **complement** A^c of a set $A \subset X$ is defined by

$$A^c = X \setminus A = \{x \in X \mid x \notin A\}.$$

In this situation we have **de Morgan's laws:**

$$\left(\bigcup_{i \in I} A_i\right)^c = \bigcap_{i \in I} A_i^c, \qquad \left(\bigcap_{i \in I} A_i\right)^c = \bigcup_{i \in I} A_i^c.$$

The set of all subsets of a given set X is called the **power set** and is denoted by $\mathcal{P}(X)$.
If X and Y are sets, their **cartesian product** $X \times Y$ is the set consisting of all ordered pairs (x, y) with $x \in X$ and $y \in Y$. Similarly, given n sets $X_1, \ldots, X_n$ we can define their cartesian product $\prod_{i=1}^n X_i = X_1 \times \ldots \times X_n$.
Given two sets X and Y, a **relation** from X to Y is a subset R of $X \times Y$. We say that R is a relation on X if R is a subset of $X \times X$, that is, $R \subset X \times X$. Quite often we write xRy instead of $(x, y) \in R$.
The most important example of a relation is a **function**. A relation f from X to Y is called a function if for each $x \in X$ there exists exactly one $y \in Y$ such that xfy. If xfy, we write $y = f(x)$; y is called the **value** of f at x. We also will write $f : X \to Y$ to mean that f is a function from X to Y. Here X is called the **domain** of f, Y is called the **codomain** of f, and the set $\{f(x) \mid x \in X\}$ is called the **range** of f. If $f : X \to Y$ is a function, $A \subset X$ and $B \subset Y$, then the **image** of A and the **preimage** (or inverse image) of B under f are sets defined by

$$f(A) = \{f(x) \mid x \in A\}, \qquad f^{-1}(B) = \{x \mid f(x) \in B\}.$$

Unions and intersections behave nicely under inverse image:

$$f^{-1}\left(\bigcup_{i \in I} A_i\right) = \bigcup_{i \in I} f^{-1}(A_i).$$

$$f^{-1}\left(\bigcap_{i \in I} A_i\right) = \bigcap_{i \in I} f^{-1}(A_i).$$

$$f^{-1}(A^c) = \left(f^{-1}(A)\right)^c.$$

Given two functions $f : X \to Y$ and $g : Y \to Z$, we define the **composition** $g \circ f$ of f and g as the function $g \circ f : X \to Z$ defined by the equation $g \circ f(x) = g(f(x))$.

We say that f is **injective**, or **one-one**, if $f(x) = f(y)$ only when $x = y$, and we say that f is **surjective**, or **onto**, if $f(X) = Y$, that is, if the image of f is the whole of Y. A function which is both injective and surjective is called **bijective**. Sometimes we will use words a "map" or a "mapping" instead of a function.
If $f : X \to Y$ is bijective, then f has an **inverse** $f^{-1} : Y \to X$. This is given by the formula $f^{-1}(y) = x$ if and only if $f(x) = y$.

Countable and Uncountable Sets

A set A is called **finite** if for some $n \in \mathbb{N}$, there is a bijection f from $\{1, \ldots, n\}$ to A. The number n is uniquely determined and is called the **cardinality of** A. We denote this fact by $\sharp A = n$ or $\mathrm{card}(A) = n$. If A is not finite, then it is called **infinite**. If A is infinite, then there is an injective function f from the set of natural numbers $\mathbb{N}$ into A. If there exists a bijection between $\mathbb{N}$ and A, then we say that X is **countably infinite**.

So A is countably infinite if and only if its elements can be listed without repetitions in an infinite sequence $X = \{x_1, x_2, \ldots\}$. A set A is called **countable** if it is finite or countably infinite; otherwise A is called **uncountable**.

Example 11.1. The set $\mathbb{Z}$ of all integers is countably infinite. To see this consider the function $f : \mathbb{N} \to \mathbb{Z}$ defined by

$$f(n) = \begin{cases} n/2 & \text{if } n \text{ is even;} \\ -(n-1)/2 & \text{if } n \text{ is odd.} \end{cases}$$

Check that the function f is a bijection from $\mathbb{N}$ to $\mathbb{Z}$ so that $\mathbb{Z}$ is countably infinite.

Example 11.2. Consider the interval $I = [0, 1]$. We shall show that I is uncountable. Seeking a contradiction, suppose that I is countable. Hence all elements of I can be listed as an infinite sequence $\{x_1, x_2, \ldots\}$ with decimal expansions:

$$\begin{aligned} x_1 &= 0.a_1^1 a_2^1 a_3^1 \cdots \\ x_2 &= 0.a_1^2 a_2^2 a_3^2 \cdots \\ x_3 &= 0.a_1^3 a_2^3 a_3^3 \cdots \\ \vdots\ & \quad \vdots \end{aligned}$$

Define

$$b_n = \begin{cases} 1 & \text{if } a_n^n \neq 1, \\ 2 & \text{if } a_n^n = 1 \end{cases}$$

and $x = 0.b_1 b_2 b_3 \cdots$. Then $x \in [0, 1]$ but it is not a member of $\{x_n\}$; contradiction.

Proposition 11.3. *Let A be a non-empty set. Then the following are equivalent:*

(a) A is countable.

(b) There exists a surjection $f : \mathbb{N} \to A$.

(c) There exists an injection $g : A \to \mathbb{N}$.

Proof. Assume that A is countable. If A is countably infinite, then there exists a bijection $f : \mathbb{N} \to A$. If A is finite, then there is a bijection $h : \{1, \ldots, n\} \to A$ for some n. Define $f : \mathbb{N} \to A$ by

$$f(i) = \begin{cases} h(i) & \text{if } 1 \leqslant i \leqslant n, \\ h(n) & \text{if } i > n. \end{cases}$$

Check that f is a surjection. So the implication $(a) \Longrightarrow (b)$ is proved. Next we prove the implication $(b) \Longrightarrow (c)$. Let $f : \mathbb{N} \to A$ be a surjection. Define $g : A \to \mathbb{N}$ by the equation $g(a) =$ smallest number in $f^{-1}(a)$. Since f is a surjection, $f^{-1}(a)$ is non-empty for any $a \in A$, so that g is well-defined. Next check that if $a \neq a'$, then $f^{-1}(a)$ and $f^{-1}(a')$ are disjoint, so they have different smallest elements. The injectivity of g follows. Now the implication $(c) \Rightarrow (a)$. Assume that $g : A \to \mathbb{N}$ is injective. We want to show that A is countable. Note that g from A to $g(A)$ is a bijection. So it suffices to show that any subset B of $\mathbb{N}$ is countable. This is obvious when B is finite. Hence assume that B is an infinite subset of $\mathbb{N}$. We define a bijection $h : \mathbb{N} \to B$. Let $h(1)$ be the smallest element of B. Since B is infinite, it is non-empty and so $h(1)$ is well-defined. Having already defined $h(n-1)$, let $h(n)$ be the smallest element of the set $\{k \in B \mid k > h(n-1)\}$. Again this set is non-empty, so $h(n)$ is well-defined. Now check that the function h is a bijection from $\mathbb{N}$ to B. ■

Corollary 11.4. *The set* $\mathbb{N} \times \mathbb{N}$ *is countable.*

Proof. In view of the previous proposition, it is enough to construct an injective function $f : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$. For example, let $f(n, m) = 2^n 3^m$. Suppose that $f(n, m) = f(k, l)$, i.e. $2^n 3^m = 2^k 3^l$. If $n < k$, then $3^m = 2^{k-n} 3^l$. The left side of this equality is an odd number whereas the right is an even number. So $n = k$, and $3^m = 3^l$. But then also $m = l$. Hence f is injective as required. ■

Proposition 11.5. *If* A *and* B *are countable, then* $A \times B$ *is countable.*

Proof. Since A and B are countable, there exist surjective functions $f : \mathbb{N} \to A$ and $g : \mathbb{N} \to B$. Define $h : \mathbb{N} \times \mathbb{N} \to A \times B$ by $h(n, m) = (f(n), g(m))$. The function h is surjective and $\mathbb{N} \times \mathbb{N}$ is countable, so $A \times B$ is countable. ■

Corollary 11.6. *The set* $\mathbb{Q}$ *of all rational numbers is countable.*

Proposition 11.7. *If* I *is a countable set and* A_i *is countable for every* $i \in I$*, then* $\bigcup_{i \in I} A_i$ *is countable.*

Proof. For each $i \in I$, there exists a surjection $f_i : \mathbb{N} \to A_i$. Moreover, since I is countable, there exists a surjection $g : \mathbb{N} \to I$. Now define $h : \mathbb{N} \times \mathbb{N} \to \bigcup_{i \in I} A_i$ by $h(n, m) = f_{g(n)}(m)$. Check that h is surjective so that $\bigcup_{i \in i} A_i$ is countable. ■

Real numbers, Sequences

The set of all real numbers, $\mathbb{R}$, has the following properties:

(a) the arithmetic properties;

(b) the ordering properties; and

(c) the completeness property.

The arithmetic properties start with the fact that any two real numbers a, b can be added to produce a real number $a + b$, the sum of a and b. The rules for addition are $a + b = b + a$, $(a + b) + c = a + (b + c)$. There is a real number 0, called zero, such that $a + 0 = 0 + a = a$ for all real numbers a. Each real number a has a negative $-a$ such that $a + (-a) = 0$. Besides addition, we have multiplication; two real numbers a, b can be multiplied to produce the product of a and b, $a \cdot b$. The rules for multiplication are $ab = ba$ and $(ab)c = a(bc)$. There is a real number 1, called one, such that $a1 = 1a = a$, and for each $a \neq 0$, there is a reciprocal $1/a$ such that $a(1/a) = 1$.

The ordering properties start with the fact that there is a subset $\mathbb{R}^+$ of $\mathbb{R}$, the set of positive real numbers. The set $\mathbb{R}^+$ is characterized by the property: if $a, b \in \mathbb{R}^+$, then $a + b$ and $ab \in \mathbb{R}^+$. The fact that $a \in \mathbb{R}^+$ is denoted by $0 < a$ or $a > 0$. The set of negative real numbers $\mathbb{R}^- = -\mathbb{R}^+$ is the set of negatives of elements in $\mathbb{R}^+$. For every $a \in \mathbb{R}$, we have $a \in \mathbb{R}^+$ or $a = 0$ or $a \in \mathbb{R}^-$. The notation $a < b$ (or $b > a$) means that $b - a \in \mathbb{R}^+$. We also write $a \leqslant b$ to mean $a < b$ or $a = b$. The order properties of real numbers are as follows:

(a) $a < b$ and $b < c$, then $a < c$.

(b) $a < b$ and $c > 0$, then $ac < bc$.

(c) $a < b$ and $c \in \mathbb{R}$, then $a + c < b + c$.

(d) $a < b$ and $a, b > 0$, then $1/b < 1/a$.

If $A \subset \mathbb{R}$, a number M is called an **upper bound** for A if $a \leqslant M$ for all $a \in A$. Similarly, m is a **lower bound** for A if $m \leqslant a$ for all $a \in A$. A subset A of $\mathbb{R}$ is called **bounded above** if it has an upper bound, and is called **bounded below** if it has a lower bound. If A has an upper and lower bound, then it is called **bounded**. A given subset of $\mathbb{R}$ may have several upper bounds. If A has an upper bound M such that $M \leqslant b$ for any upper bound b of A, then we call M a **least upper bound** of A or **supremum** of A, and denote it by $M = \sup A$. Similarly, a real number m is called a **greatest lower bound** of A or an **infimum** of A if m is a lower bound of A and $b \leqslant m$ for all lower bounds b of A. If m is the greatest lower bound of A, we write $m = \inf A$.

The **completeness property** of $\mathbb{R}$ asserts that every non-empty subset $A \subset \mathbb{R}$ that is bounded above has a least upper bound, and that every non-empty subset $S \subset \mathbb{R}$ which is bounded below has a greatest lower bound. Useful characterisations of a least upper bound and a greatest lower bound are contained in the following propositions:

Proposition 11.8. *Let $A \subset \mathbb{R}$ be bounded above. Then $a = \sup A$ if and only if $x \leqslant a$ for any $x \in A$, and for any $\varepsilon > 0$ there exists $x \in A$ such that $a < x + \varepsilon$.*

Proof. Assume first that $a = \sup A$. Clearly, $x \leqslant a$ for any $x \in A$. Take $\varepsilon > 0$. If for all $x \in A$, $x + \varepsilon \leqslant a$, then $x \leqslant a - \varepsilon$ for all x. Hence $a - \varepsilon$ is an upper bound of A contradicting the definition of a as the least upper bound of A. Conversely, from $x \leqslant a$ for any $x \in A$ follows that a is an upper bound of A. Assume that there is an upper bound b such that $b < a$. Then we get a contradiction with the fact that for any $\varepsilon > 0$ there exists $x \in A$ such that $a < x + \varepsilon$. Let $\varepsilon := (a - b)/2$ and choose such an $x \in A$. Then $x + \varepsilon \leqslant b + \varepsilon = (a + b)/2 < a$. ■

There is also a similar characterisation of $\inf A$ provided that A is bounded from below.

Proposition 11.9. *Let $A \subset \mathbb{R}$ be bounded from below. Then $a = \inf A$ if and only if $a \leqslant x$ for any $x \in A$, and for any $\varepsilon > 0$ there exists $x \in A$ such that $x - \varepsilon < a$.*

The proof of the proposition follows from the previous one by observing two facts: if A is bounded from below then the set $-A = \{x \mid -x \in A\}$ is bounded from above and that $\sup(-A) = -\inf A$.

It is useful to introduce the **extended real number system**, $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$ by adjoining symbols ∞ and $-\infty$ subject to the ordering rule $-\infty < a < \infty$ for all $a \in \mathbb{R}$. If A is not bounded above, then we write $\sup A = \infty$, and if A is not bounded below we write $\inf A = -\infty$. For example, we have $\inf \mathbb{R} = -\infty$ and $\sup \mathbb{R} = \infty$. We also have $\sup \emptyset = -\infty$ and $\inf \emptyset = \infty$, and for all non-empty sets A, $\inf A \leqslant \sup A$. With this terminology, the completeness property asserts that every subset of $\mathbb{R}$ has a least upper bound and a greatest lower bound.
The arithmetic operations on $\mathbb{R}$ can be partially extended to $\overline{\mathbb{R}}$. In particular we have:

$$\pm\infty + r = r + \pm\infty = \pm\infty \quad \text{for } r \in \mathbb{R}$$
$$(\infty) + (\infty) = \infty, \quad \text{and} \quad (-\infty) + (-\infty) = -\infty.$$

Subtraction is defined in a similar way with the exception that

$$(\infty) + (-\infty) \quad \text{and} \quad (-\infty) + (\infty)$$

are not defined. We also define multiplication by

$$r(\pm\infty) = (\pm\infty)r = \begin{cases} \pm\infty, & \text{if } r > 0, \\ \mp\infty, & \text{if } r < 0, \end{cases}$$

and

$$(\pm\infty)(\pm\infty) = \infty, \quad (\pm\infty)(\mp\infty) = -\infty.$$

The multiplication $0 \cdot (\pm\infty)$ is not defined.

If a is an upper bound of A and $a \in A$, then a is a **maximum** of A, and we write $a = \max A$. Similarly, if $a \in A$ is a lower bound of A, then a is a **minimum** of A and this fact is denoted by $a = \min A$. If A and $B \subset \mathbb{R}$, then $A + B = \{a + b \mid a \in A,\ b \in B\}$, $A + a = \{x + a \mid x \in A\}$, and $aA = \{ax \mid x \in A\}$. Here are some properties of supremum and infimum:

(a) **monotonicity property:** $A \subset B$, then $\sup A \leqslant \sup B$ and $\inf B \leqslant \inf A$.

(b) **reflection property:** $\sup(-A) = -\inf A$ and $\inf(-A) = -\sup A$.

(c) **translation property:** $\sup(A + a) = \sup A + a$ and $\inf(A + a) = \inf A + a$.

(d) **dilation property:** $\sup(aA) = a \sup A$ and $\inf(aA) = a \inf A$, provided that $a > 0$

(e) **addition property:** $\sup(A + B) = \sup A + \sup B$ and $\inf(A + B) = \inf A + \inf B$.

A **sequence** of real numbers is a function $f : \mathbb{N} \to \mathbb{R}$. We often write the sequence as $\{f(n)\}$ or $\{f_n\}$. A sequence $\{a_n\}$ of real numbers is said to **converge** to a real number a if for every $\varepsilon > 0$ there is an integer n_0 such that if $n \geq n_0$, then $|a_n - a| < \varepsilon$. In this situation we call a the **limit** of $\{a_n\}$; a convergent sequence has a unique limit. We also write $a_n \to a$ or $\lim_{n\to\infty} a_n = a$. A sequence $\{a_n\}$ which does not converge to any limit in $\mathbb{R}$ is said to **diverge**. We say that $a_n \to \infty$ if for every $M > 0$, there is n_0 such that $a_n > M$ for all $n \geq n_0$. Similarly, $a_n \to -\infty$, if for every $M < 0$ there exists n_0 such that $a_n < M$ for all $n \geq n_0$. A sequence $\{a_n\}$ is **bounded** if $|a_n| < M$ for some number M and all $n \in \mathbb{N}$. A convergent sequence is always bounded. Here are some elementary properties of limits of sequences:

Proposition 11.10. *Let $\{a_n\}$ and $\{b_n\}$ be sequences converging to a and b, respectively. Let $c \in \mathbb{R}$. Then*

(a) $\{ca_n\}$ converges to ca;

(b) the sequence $\{a_n + b_n\}$ converges to $a + b$;

(c) the sequence $\{a_n \cdot b_n\}$ converges to $a \cdot b$;

(d) if $b_n \neq 0$ for all n and $b \neq 0$, then the sequence $\{a_n/b_n\}$ converges to a/b.

A sequence $\{a_n\}$ is called **monotone increasing** if $a_n \leqslant a_{n+1}$ for all $n \in \mathbb{N}$. It is **monotone decreasing** if $a_n \geq a_{n+1}$ for all $n \in \mathbb{N}$.

Proposition 11.11. *If $\{a_n\}$ is a monotone increasing sequence that is bounded above, $a_n \leqslant M$ for all n, then $\{a_n\}$ is convergent. If $\{a_n\}$ is monotone increasing and it is not bounded from above, then $a_n \to \infty$. If $\{a_n\}$ is monotone decreasing and it is bounded below, $M \leqslant a_n$ for all n, then $\{a_n\}$ is convergent, and if $\{a_n\}$ is not bounded from below, then $a_n \to -\infty$.*

Proof. If $\{a_n\}$ is unbounded from above, then for every M there is k such that $a_k > M$. Since the sequence is increasing, $a_n \geq a_k \geq M$ for all $n \geq k$. Thus $a_n \to \infty$. Next assume that $\{a_n\}$ is bounded above. Then $a := \sup\{a_n \mid n \in \mathbb{N}\} < \infty$. Let $\varepsilon > 0$. By the definition of supremum, $a_n \leqslant a$ for all n and there is an integer n_0 such that $a < a_{n_0} + \varepsilon$. Since $\{a_n\}$ is monotone increasing, $a_n \leqslant a < a_n + \varepsilon$ for all $n \geq n_0$, that is, $|a_n - a| < \varepsilon$ for all $n \geq n_0$. Thus the sequence converges to a. The proof for monotonically decreasing sequences is similar and is left as an exercise. ■

Let $\{a_n\}$ be a sequence. If $0 < n_1 < n_2 < \ldots$ are positive integers, then $\{a_{n_k}\}$ is called a **subsequence** of $\{a_n\}$.

Proposition 11.12. *If $\{a_n\}$ is a convergent sequence with the limit a, then every subsequence of $\{a_n\}$ converges to a. Conversely, if a sequence $\{a_n\}$ has the property that each of its subsequences is convergent, then $\{a_n\}$ itself converges.*

Proof. Let $\{a_{n_k}\}$ be a subsequence of $\{a_n\}$. For a given $\varepsilon > 0$, choose n_0 such that $|a_n - a| < \varepsilon$ for all $n > n_0$. Note that if $k > n_0$, then $n_k > n_0$ and so $|a_{n_k} - a| < \varepsilon$ for all $k > n_0$. Therefore, $\{a_{n_k}\}$ converges to a. The converse follows from the fact that the sequence $\{a_n\}$ is a subsequence of itself. ■

Let $\{a_n\}$ be a bounded sequence. For each $n \in \mathbb{N}$, let $b_n = \sup_{m \geq n} a_m = \sup\{a_n, a_{n+1}, \ldots\}$. Then $\{b_n\}$ is monotone decreasing, and it is bounded since $\{a_n\}$ is bounded. In view of Proposition 11.11, $\{b_n\}$ converges. The limit is called the **upper limit** of $\{a_n\}$. Similarly, let $c_n = \inf_{m \geq n} a_m = \inf\{a_n, a_{n+1}, \ldots\}$. Then $\{c_n\}$ is monotone increasing, and it is bounded since $\{a_n\}$ is bounded. The limit of $\{c_n\}$ is called the **lower limit** of $\{a_n\}$. If $\{a_n\}$ is not bounded above, then its upper limit is equal to ∞, and if $\{a_n\}$

is not bounded below, then its lower limit is equal to $-\infty$. Summarizing

$$\limsup a_n = \overline{\lim}\, a_n = \inf_{n\geq m} \sup_{k\geq n} a_k = \lim_{n\to\infty} \sup_{k\geq n} a_k$$
$$\liminf a_n = \underline{\lim}\, a_n = \sup_{n\geq m} \inf_{k\geq n} a_k = \lim_{n\to\infty} \inf_{k\geq n} a_k,$$

for any fixed integer $m \in \mathbb{N}$. A useful characterisation of the upper limit is the following proposition.

Proposition 11.13. *Let $\{a_n\}$ be a sequence in $\mathbb{R}$. Then the following are equivalent:*

(a) $\overline{\lim}\, a_n = a$.

(b) For every $b > a$, $a_n < b$ for all but finitely many n and for every $c < a$, $a_n > c$ for infinitely many n.

Proof. Assume $\overline{\lim}\, a_n = a$. Then for any $b > a$, there exists m such that $\sup_{n\geq m} a_n < b$. In particular, $a_n < b$ for all $n \geq m$. Since the sequence $\{\sup_{n\geq m} a_n\}$ is decreasing and convergent to a, it follows that $a \leqslant \sup_{n\geq m} a_n$ for all m. Hence if $c < a$, then for every m there exists $n \geq m$ such that $c < a_n$. This shows the implication $(a) \Longrightarrow (b)$. Conversely, assume that (b) holds. Then for every $b > a$, there exists m such that $a_n < b$ for all $n \geq m$. Hence $\sup_{n\geq m} a_n \leqslant b$. This implies that $\limsup a_n \leqslant b$ for every $b > a$ so that $\limsup a_n \leqslant a$. If for every $c < a$ and for every m there exists $n \geq m$ such that $a_n > c$, then for every m, $\sup_{n\geq m} a_n \geq c$. This gives $\limsup a_n > c$ and since this holds for every $c < a$, we have $\limsup a_n \geq a$. Thus $\limsup a_n = a$ and the implication $(b) \Longrightarrow (a)$ is proved. ∎

As an exercise formulate and prove the corresponding statement for the lower limit. The basic properties of the upper and the lower limits are listed in the following proposition:

Proposition 11.14. *If $\{a_n\}$ and $\{b_n\}$ are sequences of real numbers, then:*

(a) $\limsup(-a_n) = -\liminf a_n$ *and* $\liminf(-a_n) = -\limsup a_n$*;*

(b) $\limsup(ca_n) = c \limsup a_n$ *and* $\liminf(ca_n) = c \liminf a_n$ *for any* $c > 0$*;*

(c) $\limsup(a_n + b_n) \leqslant \limsup a_n + \limsup b_n$ *and* $\liminf a_n + \liminf b_n \leqslant \liminf(a_n + b_n)$*;*

(d) $\liminf a_n \leqslant \limsup a_n$, *with equality if and only if* $\{a_n\}$ *converges (in this case* $\limsup a_n = \lim a_n$*);*

(f) If $\{a_{n_k}\}$ *is a subsequence of* $\{a_n\}$*, then* $\liminf a_n \leqslant \liminf a_{n_k} \leqslant \limsup a_{n_k} \leqslant \limsup a_n$.

The proof is left as an exercise.

Theorem 11.15 (Bolzano-Weierstrass Theorem).
Let $\{a_n\}$ *be a bounded sequence in* $\mathbb{R}$*. Then it has a subsequence that converges.*

Proof. Set $a = \limsup a_n$. We will construct inductively a subsequence $\{a_{n_k}\}$ of $\{a_n\}$ which converges to a. In view of Proposition 11.13, there exists n_1 such that $a_{n_1} > a - 1$. Having obtained $n_1 < n_2 < \cdots < n_k$ such that $a_{n_j} > a - 1/j$ for $1 \leqslant j \leqslant k$, we find, again by applying Proposition 11.13, $n_{k+1} > n_k$ such that $a_{n_{k+1}} > a - 1/(k+1)$. Hence $a \leqslant \liminf a_{n_k} \leqslant \limsup a_{n_k} \leqslant \limsup a_n = a$. So $\lim a_{n_k} = a$ and the proof is finished. ■

References

[Me] B. Mendelson, *Introduction to Topology.*

[M] J. R. Munkres, *Topology: A First Course.*

[P] C. W. Patty, *Foundations of topology.*

[S] W. A. Sutherland. *Introduction to metric and topological spaces.*