Chapter 3

The Topology of the Real Line

A *topology* on a set X is a family of distinguished subsets of X called *open sets*. A space with a topology is called a *topological space*. A topology is in a certain sense a generalization of a metric, in that it allows us to carry over many of the ideas and methods metric spaces to topological spaces, without having a distance function.

In a metric space (X, ρ), we usually define a topology in terms of the metric ρ, called the *metric topology*. Everything that can be said or proven in terms of the metric topology can be proven directly in terms of ρ without reference to the topology, but the topological language can often make the discussion more transparent.

We focus primarily on $\mathbb{R}$ and its standard metric topology in an effort to keep the discussion as concrete as possible for students who may need more time to feel comfortable with a higher level of abstraction. On the other hand, much of what we discuss in this chapter is true in general metric spaces, and we state (and prove) such results in general when appropriate.

3.1. Closed sets and open sets in $\mathbb{R}$

To define the topology on $\mathbb{R}$, we need to describe which sets are open. Equivalently, we can define which sets are *closed*. A set F in a topological space X is closed if it is the complement of an open set $O \subset X$. Both concepts are defined in terms of open balls, which in $\mathbb{R}$ are open intervals.

3.1.1. Limit points and closed sets

A set F is closed if any point that "touches" F is contained in F. By "touches", we mean that the point is a limit point of F.

Definition. A point $x \in \mathbb{R}$ is a *limit point* of a set $E \subset \mathbb{R}$ if for all $\varepsilon > 0$, the set $(x-\varepsilon, x+\varepsilon) \cap E$ is infinite. A set $E \subset \mathbb{R}$ is *closed* if and only if E contains all of its limit points.

The set of limit points of a set $E \subset \mathbb{R}$ is denoted by E'.

More generally, if X is a metric space, then x is a limit point of the set $E \subset X$ if for every $\varepsilon > 0$, the set $B(x,\varepsilon) \cap E$ is infinite. A set $E \subset X$ is closed if and only if it contains all of its limit points.

We observe that the condition "for all $\varepsilon > 0$ the set $B(x,\varepsilon) \cap E$ is infinite" is equivalent to the seemingly weaker condition "for all $\varepsilon > 0$, the set $B(x,\varepsilon) \cap E$ contains at least one point other than x." See **ex3.1.2**.

Proposition. *The (closed) interval $[a,b]$ is closed for every $a, b \in \mathbb{R}$, with $a < b$.*

Proof. Suppose that $a < b$. We will show that if $x \notin [a,b]$, then x is not a limit point of $[a,b]$. Hence, $[a,b]$ must contain all of its limit points.

If $x \notin [a,b]$, then either $x < a$ or $b < x$. In the first case, setting $\varepsilon = (a-x)/2 > 0$, we have $(x-\varepsilon, x+\varepsilon) \cap [a,b] = \emptyset$, and in the second case, setting $\varepsilon = (x-b)/2 > 0$, we have $(x-\varepsilon, x+\varepsilon) \cap [a,b] = \emptyset$. Thus, if $x \notin [a,b]$, there is an $\varepsilon > 0$ such that $(x-\varepsilon, x+\varepsilon) \cap [a,b] = \emptyset$, so x cannot be a limit point of $[a,b]$. □

Comment. The limits of convergent subsequences of a sequence $\{x_n\}_{n=1}^{\infty} \subset \mathbb{R}$ are called the "limit points of the sequence". It is important to distinguish between the limit points of a sequence and the limit points of the *range* of the sequence.

A sequence $\{x_n\}_{n=1}^{\infty} \subset \mathbb{R}$ represents a function $f : \mathbb{N} \to \mathbb{R}$, where $x_n = f(n)$. The range E of a sequence is the same as the range of the function that it represents:

$$E = f(\mathbb{N}) = \{x \in \mathbb{R} : x = f(n) \text{ for some } n \in \mathbb{N}\}.$$

The point a is a limit point of the sequence $\{x_n\}$ (i.e., subsequential limit) if the set $I_\varepsilon = \{n \in \mathbb{N} : |x_n - a| < \varepsilon\}$ is infinite for every $\varepsilon > 0$. On the other hand, the point a is a limit point of the range E of $\{x_n\}$ if the set $E_\varepsilon = \{x \in E : |x-a| < \varepsilon\}$ is infinite for every $\varepsilon > 0$.

Theorem. *If $F \subset \mathbb{R}$ is closed, $x_n \in F$ for all $n \in \mathbb{N}$ and $\lim x_n = x^*$, then $x^* \in F$.*

Proof. If $x^* = x_n$ for some $n \in \mathbb{N}$, then $x^* \in F$ and we are done. Otherwise $x^* \neq x_n$ for all $n \in \mathbb{N}$, and we show that in this case, x^* is a limit point of F and hence contained in F.

Let $\varepsilon > 0$, then since $x_n \to x^*$, there is an $N_\varepsilon \in \mathbb{N}$ such that $x_n \in (x^*-\varepsilon, x^*+\varepsilon)$ for all $n \geq N_\varepsilon$. If $x_n = x_m$ for all $n, m \geq N_\varepsilon$, then $x^* = x_N$ contradicting our assumption. Hence there are $n > m \geq N_\varepsilon$ such that $x_n \neq x_m$, and therefore the interval $(x^*-\varepsilon, x^*+\varepsilon) \cap F$ contains at least two points, from which it follows that x^* is a limit point of F. □

3.1.2. Open sets

Open sets may be defined as the complements of closed sets, which they are, but they are naturally defined as follows.

Definition. A set $E \subset \mathbb{R}$ is *open* if for every $x \in E$ there is a positive ε such that the interval $(x-\varepsilon, x+\varepsilon)$ is contained in E.

More generally, if (X,ρ) is a metric space, then $A \subset X$ is open if for any $x \in A$, there is an $\varepsilon > 0$ such that $B(x,\varepsilon) \subset A$.

We observe that in any metric space X, the empty set satisfies the condition of being open, but "vacuously", because the condition "if $x \in E$, then…" is satisfied for $E = \varnothing$ since there is no $x \in \varnothing$.

Proposition. *If (X,ρ) is a metric space then for every $x \in X$ and $\varepsilon > 0$, the sets $B(x,\varepsilon)$ and $R(x,\varepsilon) = \{y \in X : \rho(x,y) > \varepsilon\}$ are both open.*

Proof. Let $x \in X$ and $\varepsilon > 0$. If $y \in B(x,\varepsilon)$, then $\rho(y,x) < \varepsilon$, and $\delta_y = \varepsilon - \rho(y,x) > 0$. If $z \in B(y,\delta_y)$, then

$$\rho(z,x) \le \rho(z,y) + \rho(y,x) < \delta + \rho(y,x) = \varepsilon,$$

so $z \in B(x,\varepsilon)$ and hence $B(y,\delta_y) \subset B(x,\varepsilon)$.

If $y \in R(x,\varepsilon)$, then $\eta_y = \rho(x,y) - \varepsilon > 0$. If $z \in B(y,\eta_y)$, then

$$\rho(x,y) \le \rho(x,z) + \rho(z,y) < \rho(x,z) + \eta_y = \rho(x,z) + \rho(x,y) - \varepsilon,$$

from which it follows that $\varepsilon < \rho(x,z)$. Hence if $y \in R(x,\varepsilon)$, then $B(y,\eta_y) \subset R(x,\varepsilon)$, so $R(x,\varepsilon)$ is open. □

Theorem. *The set E in the metric space (X,ρ) is open if and only if E^c is closed.*

Proof. Suppose that E^c is closed and $x \in E$. Then x is not a limit point of E^c and it follows that there exists a positive ε such that the set $A = B(x,\varepsilon) \cap E^c$ is *finite*. This means that $\{\rho(x,y) : y \in A\}$ is a finite set of (strictly) positive numbers, so $\min\{\rho(x,y) : y \in A\} = \varepsilon_A > 0$. Hence, $B(x,\varepsilon_A) \cap E^c = \varnothing$ so $B(x,\varepsilon_A) \subset E$ and E is open.

Conversely, if E is an open set and $x \in E$, then there is a positive ε such that $B(x,\varepsilon) \subset E$. It follows that $B(x,\varepsilon) \cap E^c$ is empty, so x is not a limit point of E^c. Hence, E^c contains all of its limit points and is therefore closed. □

Corollary. *If X is a metric space, then for any $x \in X$ and $\varepsilon > 0$, the set $\overline{B}(x,\varepsilon) = \{y \in X : \rho(x,y) \le \varepsilon\}$ is closed.*

Proof. We have $\overline{B}(x,\varepsilon) = (R(x,\varepsilon))^c$ and by Proposition 3.1.2, the set $R(x,\varepsilon)$ is open. □

Comment. We call $\overline{B}(x,\varepsilon)$ the *closed* ball of radius ε around x. Note that in $\mathbb{R}$, $[a,b] = \overline{B}(c,\varepsilon)$, where $c = (a+b)/2$ and $\varepsilon = (b-a)/2$, so that the corollary above includes Proposition 3.1.1 as a special case.

3.1.3. Unions and intersections of open sets and closed sets

Certain combinations of open sets are themselves open, and the corresponding combinations of closed sets are closed.

Theorem. *The union of any family of open sets is open and the intersection of any finite collection of open sets is open.*

Proof. Suppose that E_a is an open set for every $a \in A$. If $x \in \bigcup_{a\in A} E_a$, then there is an $a_1 \in A$ such that $x \in E_{a_1}$. Since E_{a_1} is open by assumption, there is a positive ε such that $B(x,\varepsilon) \subset E_{a_1}$, and hence $B(x,\varepsilon) \subset \bigcup_{a\in A} E_a$. Since x was arbitrary, it follows that $\bigcup_{a\in A} E_a$ is open.

Now suppose that E_j is open for $1 \le j \le n$ and let $x \in \bigcap_{j=1}^n E_j$, so that $x \in E_j$ for $1 \le j \le n$. Since E_j is open for each j, there is an $\varepsilon_j > 0$ such that $B(x,\varepsilon_j) \subset E_j$. If $\varepsilon = \min(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)$, then $\varepsilon > 0$ and $B(x,\varepsilon) \subset B(x,\varepsilon_j) \subset E_j$ for each $1 \le j \le n$. Hence $B(x,\varepsilon) \subset \bigcap_{j=1}^n E_j$, and since x was arbitrary, it follows that $\bigcap_{j=1}^n E_j$ is open. ☐

Corollary. *The intersection of any family of closed sets is closed and the union of any finite collection of closed sets is closed.*

Proof. Exercise ☐

We observe that there is no mention of $\mathbb{R}$ in the statement or proof of Theorem 3.1.3. This theorem and its corollary are true in any metric space with the metric topology, and are in fact characteristic of topological spaces in general.

3.1.4. Closures and interiors

Given a nonempty subset $E \subset \mathbb{R}$, we might think of "closing it" by adding all of its limits points. Likewise, by removing points from E that are limit points of its complement in $\mathbb{R}$, we might construct the 'open part' of E — the largest open set contained in E. These intuitions are justified after the following definitions.

Definition. Suppose that $E \subset \mathbb{R}$ is nonempty. Let $\mathcal{F}$ be the collection of all closed subsets of $\mathbb{R}$ that contain E and let $\mathcal{O}$ be the collection of all open subsets of $\mathbb{R}$ that are contained in E. The *closure* of E is the set

$$\overline{E} = \bigcap_{F\in\mathcal{F}} F$$

and the *interior* of E is the set

$$E^o = \bigcup_{O\in\mathcal{O}} O.$$

We observe that $E \subset \mathbb{R}$ may contain no open subsets, so the interior of E may be empty.[1] On the other hand, every subset of $\mathbb{R}$ is contained in the closed set $\mathbb{R}$, so the closure of a nonempty set is never empty.

Theorem. *For any $E \subset \mathbb{R}$,*

(i) *$\overline{E}$ is closed, $E \subset \overline{E}$ and if F is closed and contains E, then $\overline{E} \subset F$.*
(ii) *E^o is open, $E^o \subset E$, and if O is open and contained in E, then $O \subset E^o$.*

I.e., the closure of E is the smallest closed set containing E and the interior of E is the largest open set contained in E.

Proof. Let $E \subset \mathbb{R}$ be nonempty, and let $\mathcal{F}$ and $\mathcal{O}$ be the collections of sets described in the definitions of closure and interior, above.

[1] For example, the interior of $\mathbb{Q}$ is empty.

(i) $\overline{E}$ is closed by Corollary 3.1.3 as the intersection of a family of closed sets. If $x \in E$, then $x \in F$ for all $F \in \mathcal{F}$ and therefore $x \in \bigcap_{F\in\mathcal{F}} F = \overline{E}$, so $E \subset \overline{E}$. If F_0 is closed and $E \subset F_0$, then $F \in \mathcal{F}$, so $\overline{E} = \bigcap_{F\in\mathcal{F}} F \subset F_0$.
(ii) E^o is open by Theorem 3.1.3 as the union of a family of open sets. If $x \in E^o = \bigcup_{O\in\mathcal{O}} O$, then $x \in O \subset E$ for some $O \in \mathcal{O}$, hence $E^o \subset E$. If O is an open set contained in E, then $O \in \mathcal{O}$, so $O \subset \bigcup_{O\in\mathcal{O}} O = E^o$. □

Proposition. *Let $E \subset \mathbb{R}$. Then $\overline{E} = E \cup E'$ and $E^o = E \setminus (E^c)'$.*

Recall that E' denotes the set of limit points of E. Also, for any two sets A and B, $A \setminus B$ is the set $A \cap B^c$, of points in A that are not in B.

Proof. First, if x is a limit point of E, then x is a limit point of $\overline{E}$, so $x \in \overline{E}$, because $\overline{E}$ is closed. It follows that $F = E \cup E' \subset \overline{E}$. To show that $F = \overline{E}$, it is enough to show that F is closed, because $E \subset F$.

To that end, Let x be a limit point of F. Then for any $\varepsilon > 0$, the set $B(x, \varepsilon/2) \cap F$ contains infinitely many points, and so contains a point $y \neq x$. If $y \in E$, we set $x_\varepsilon = y$. If $y \notin E$, then y is a limit point of E, and the set $B(y, \varepsilon/2) \cap E$ contains infinitely many points, and therefore contains a point $z \neq x$. In this case, we set $x_\varepsilon = z$ and observe that

$$|x - x_\varepsilon| \leq |x - y| + |y - \varepsilon| < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

In either case, for any $\varepsilon > 0$, the set $B(x, \varepsilon) \cap E$ contains a point $x_\varepsilon \neq x$, and it follows that x is a limit point of E, by the observation following Definition 3.1.1.

Second, if $x \in E$ is not a limit point of E^c, then there is a $\varepsilon > 0$ such that $B(x, \varepsilon) \cap E^c$ is empty, so $B(x, \varepsilon) \subset E$. It follows that $E \setminus (E^c)'$ is open. Moreover, if $O \subset E$ is open and $x \in O$, then $x \notin (E^c)'$, so $x \in E \setminus (E^c)'$, and it follows that $E \setminus (E^c)'$ contains every open subset of E. □

3.1.5. Topology, more generally

At the beginning of the chapter we wrote that a topology on a set X is a collection of special subsets of X called open sets. However not every collection of subsets can be a topology. Certain conditions have to be met.

Definition. A topology on a set X is a family $\mathcal{O}$ of subsets of X, called open sets, with the following properties.

top1. $\varnothing \in \mathcal{O}$ and $X \in \mathcal{O}$.

top2. If $E_a \in \mathcal{O}$ for all $a \in A$, then $\bigcup_{a\in A} E_a \in \mathcal{O}$.

top3. If $E_j \in \mathcal{O}$ for $1 \leq j \leq n$, then $\bigcap_{j=1}^{n} E_j \in \mathcal{O}$.

Properties **top**2. and **top**3. are often summarized by saying "$\mathcal{O}$ is closed under arbitrary unions and finite intersections".

A set $F \subset X$ is *closed* if its complement in X is open, and we note that the set X itself is both open and closed, as is the empty set.

A general topological space X with topology $\mathcal{O} \subseteq \mathcal{P}(X)$ is denoted by $(X, \mathcal{O})$.

Theorem 3.1.3 shows that the collection of sets in $\mathbb{R}$ that we defined to be open in Definition 3.1.2 satisfies the definition of a topology. I.e., the metric topology is a topology on $\mathbb{R}$ and the same is true for the metric topology of any metric space.

There are always two topologies that can be defined on any nonempty set X. The first is the trivial topology that includes only the sets X and $\varnothing$ as open sets. At the other extreme is the *discrete topology* that includes all subsets of X as open sets. In the discrete topology for example, each singleton $\{x\} \subset X$ is open.

3.1.6. The relative topology

Nonempty subsets of topological spaces are naturally endowed with a topologies that they "inherit" from the larger space.

Theorem. *If $(X, \mathcal{O})$ is a topological space and Y is a nonempty subset of X, then the family of sets $\mathcal{O}_Y = \{U \cap Y : U \in \mathcal{O}\}$ is a topology on Y.*

The topology $\mathcal{O}_Y$ is called the *relative topology* on $Y \subset X$.

Proof. We need to show that $\mathcal{O}_Y$ satisfies properties **top**1.–**top**3., and we leave this as an exercise. ☐

If $Y \subset X$ is open and $U \in \mathcal{O}$, then $U \cap Y \in \mathcal{O}$. In this case $\mathcal{O}_Y = \{U \in \mathcal{O} : U \subset Y\}$ is a subset of $\mathcal{O}$.

If Y is not an open subset of X and U is open, then $Y \cap U$ need not be an open set in X (though it can be). For example, if $I = [0, 1]$, then the interval $[0, 1/2) = I \cap (-1/2, 1/2)$ is open in the relative topology on I even though it is not open in $\mathbb{R}$.

3.1.7. Connected sets

One of the nice properties of intervals in $\mathbb{R}$ is that they do not have any gaps. More formally, we say that intervals in $\mathbb{R}$ are *connected*.

Definition. A nonempty subset E of a metric space X is *disconnected* if there are disjoint, open sets $U, V \subset X$ such that $E \cap U \neq \varnothing$, $E \cap V \neq \varnothing$ and $E \subset U \cup V$. A set is *connected* if it is not disconnected.

We observe that in any topological space X, a singleton $\{x\}$ is always connected, however in general, connected sets can have many forms. In $\mathbb{R}$, the only connected sets are intervals (and singletons).

Theorem. *Let $E \subset \mathbb{R}$ contain at least two points. Then E is connected if and only if E is an interval.*

Proof. If E is not an interval, then it follows from Proposition 1.5.1 that there exist points $x, y \in E$ and $z \notin E$ such that $x < z < y$. It follows that $E \subset (-\infty, z) \cup (z, \infty)$, so E is disconnected.

To prove that all intervals are connected, we first prove that bounded closed intervals are connected. Suppose that $a < b$ and that, by way of contradiction, $[a, b] \subset U \cup V$, where U and V are disjoint open sets, such that $U \cap [a, b]$ and $V \cap [a, b]$ are both nonempty. Without loss of generality, we may assume that $a \in U$ and since

U is open, there is an $\varepsilon > 0$, such that $(a - \varepsilon, a + \varepsilon) \subset U$. It follows that the set $\{x \in \mathbb{R} : [a, x] \subset U\}$ is not empty, and we denote by c its supremum. If $c \geq b$, then $[a, b] \subset U$, so $[a, b] \cap V = \emptyset$, which contradicts one of our assumptions. If $c < b$, then $c \in U$ and $(c, b] \subset V$. Since $c \in U$ and U is open, there is an $\delta > 0$, such that $(c - \delta, c + \delta) \subset U$, but $(c - \delta, c + \delta) \cap (c, b] \neq \emptyset$ which implies that $U \cap V \neq \emptyset$. This contradicts our other assumption, so $[a, b]$ must be connected.

Finally, suppose that I is any other type of interval, open, half-open or unbounded, and let $I \subset U \cup V$, where U and V are disjoint open sets. If $[a, b] \subset U$ for all $[a, b] \subset I$ or $[a, b] \subset V$ for all $[a, b] \subset I$, then either $I \subset U$ or $I \subset V$, and U is not disconnected by U and V. Otherwise, there exist closed bounded, intervals $[a, b], [c, d] \subset I$ such that $[a, b] \subset U$ and $[c, d] \subset V$.

If $x \in [a, b] \cap [c, d]$, then $U \cap V \neq \emptyset$, so these intervals are not disjoint, and we may assume that $b < c$. Since $b, c \in I$ and I is an interval, then I must contain the interval $[b, c]$. But $b \in [b, c] \cap U$ and $c \in [b, c] \cap V$, which contradicts the fact that closed bounded intervals are connected. □

3.1.8. Separability and complete separability

We have seen that $\mathbb{Q}$ is dense in $\mathbb{R}$ in two different senses: (i) between every two real numbers, there is a rational number (order density) and (ii) for every real number there sequence of rational numbers converging to it (sequential density). To these two notions, we now add a third, topological sense.

Definition. A subset A in a topological space X is (topologically) *dense* if every nonempty open set $O \subset X$ contains an element from A. If the topological space X has a *countable* dense subset, then X is said to be *separable*. If there is a countable collection $\mathcal{O} = \{O_n\}_{n \in \mathbb{N}}$ of open sets in X such that every open set in X is a union of sets from $\mathcal{O}$, then X is said to be *completely separable*.

Lemma. *Every nonempty open subset of $\mathbb{R}$ contains a rational number.*

I.e., $\mathbb{Q}$ is (topologically) dense in $\mathbb{R}$, and since $\mathbb{Q}$ is countable, $\mathbb{R}$ is a separable metric space.

Proof. If $O \subset \mathbb{R}$ is nonempty and open, then there is an $x \in O$ and an $\varepsilon > 0$ such that $(x - \varepsilon, x + \varepsilon) \subset O$. By the order density of $\mathbb{Q}$ in $\mathbb{R}$, it follows that there is an $r \in \mathbb{Q}$ such that $x - \varepsilon < r < x + \varepsilon$, and hence, $r \in O$. □

The collection of open intervals with rational endpoints, $\mathcal{Q} = \{(a, b) : a, b \in \mathbb{Q}\}$, is countable because $\mathbb{Q}$ is countable, so it follows from the next theorem that $\mathbb{R}$ is completely separable.

Theorem. *Every nonempty open set $O \in \mathbb{R}$ is a union of open intervals with rational endpoints.*

Proof. If $O \subset \mathbb{R}$ is open and $x \in O$, then there is an $\varepsilon_x > 0$ such that $(x - \varepsilon_x, x + \varepsilon_x) \subset O$ and it follows that $O = \bigcup_{x \in O}(x - \varepsilon_x, x + \varepsilon_x)$. To prove the theorem, it is therefore enough to prove that every bounded open interval is a union of intervals with rational endpoints.

Suppose then that $(a, b) \subset \mathbb{R}$. From the density of $\mathbb{Q}$ in $\mathbb{R}$, it follows that for every $n \in \mathbb{N}$, there are $a_n, b_n \in \mathbb{Q}$ such that $a_n \in (a, a + \delta_n)$ and $b_n \in (b - \delta_n)$, where

$$\delta_n = \frac{b-a}{3n}.$$

We have $b_n - a_n > \frac{b-a}{3n} > 0$ and therefore $a < a_n < b_n < b$ for all $n \in \mathbb{N}$ and furthermore, $\delta_n \to 0$, so $\lim a_n = a$ and $\lim b_n = b$.

Let $x \in (a, b)$, and let $\varepsilon = \min(x - a, b - x)$, so that $\varepsilon > 0$. For $n \in \mathbb{N}$ that is sufficiently large, we have $\delta_n < \varepsilon$ and for such n, we have $a_n < x < b_n$. In other words, for every $x \in (a, b)$, there is an $n \in \mathbb{N}$ such that $x \in (a_n, b_n)$, and since $(a_n, b_n) \subset (a, b)$ for all $n \in \mathbb{N}$ this means that

$$(a, b) = \bigcup_{n=1}^{\infty} (a_n, b_n). \qquad \square$$

The intervals with rational endpoints guaranteed by the theorem are not necessarily pairwise disjoint. On the other hand, if we drop the condition that the endpoints are rational, we have the following fact.

Proposition. *If $G \subseteq \mathbb{R}$ is open, then there is a countable set of pairwise disjoint open intervals $\{I_n\}$ such that $G = \bigcup_n I_n$.*

Proof. Exercise. □

Exercises

ex3.1.1. Show that any finite subset F of $\mathbb{R}$ is closed.

ex3.1.2. Prove that x is a limit point of a set E if and only if for every $\varepsilon > 0$ the set $(x - \varepsilon, x + \varepsilon)$ contains a point $y \in E$, $y \neq x$.

ex3.1.3. Prove that x is a limit point of a set E if, and only if, there exists a sequence $\{x_n\}_{n \in \mathbb{N}}$ of distinct elements of E such that $x = \lim_{n\to\infty} x_n$.

ex3.1.4. Show that if a is a limit point of the range of the sequence $\{x_n\}$, then it is a limit point of the sequence.

Hint: You may rely on the previous exercise, but there is a little more work to be done beyond quoting it.

ex3.1.5. Suppose that X is a metric space and $\{x_n\}_{n=1}^{\infty} \subset X$ is a sequence with no convergent subsequences. Show that the set $\{x_n\}_{n \in \mathbb{N}}$ is closed in X.

ex3.1.6. Prove that the intersection of any collection of closed sets is closed, and that the union of finitely many closed sets is closed.

Hint: Theorem 1.1 may be useful here, and in the next exercise.

ex3.1.7. Complete the proof of Theorem 3.1.6.

ex3.1.8. Prove that for any set E, the set E' of limit points of E is closed.

Hint: Show that every limit point of E' is also a limit point of E. Adapt (part of) the proof of Proposition 3.1.4 to do this.

ex3.1.9. Prove Proposition 3.1.8.

Hint: If two open intervals overlap, then their union is a single open interval. If $\{I_a\}_{a\in A}$ is a collection of disjoint open intervals then each one of them contains a rational number not contained in any other.

3.2. Compactness

In introductory calculus classes, we learn that a continuous function f defined on a closed interval $[a, b]$ always attains both a minimum and maximum value in the interval. This is an important theorem, with many applications.

The proof of this theorem relies on the continuity of the function f, properly defined, and (at least implicitly) the fact that closed intervals in $\mathbb{R}$ are *compact*. We will return to continuous functions in Chapter 4, and in this section explore the concept of compactness in $\mathbb{R}$ and in metric spaces in general.

3.2.1. Open covers

Compact subsets of $\mathbb{R}$ (and compact sets in general) behave in certain respects like finite sets. For example, if E is a finite set and $x_n \in E$ for all $n \in \mathbb{N}$, then the sequence $\{x_n\}$ has a limit point in E. As we shall see below, the same is true for compact sets.

Compactness can be defined in several ways, and we begin with a definition that incorporates finiteness in a crucial way.

Definition. Let X be a topological space.[2]

(i) An *open cover* of a subset E of X is a family $\{O_a\}_{a\in A}$ of open sets in X such that $E \subset \bigcup_{a\in A} O_a$.

(ii) A *subcover* of an open cover $\{O_a\}_{a\in A}$ of E is a family of open sets $\{O_b\}_{b\in B}$, where $B \subset A$ and $E \subset \bigcup_{b\in B} O_b$. The subcover is a *finite subcover* if the index set B is a finite set. If $\{O_a : a \in A\}$ contains a finite subcover of E, we also say that E *admits a finite subcover from* $\mathcal{O}$.

(iii) A set E in X is *compact* if *every* open cover of E admits a finite subcover of E.

(iv) A metric space X is a *compact metric space* if X is a compact set in itself.

In line with the remarks before the definition, we observe that finite sets are always compact (in any topological space) and leave the proof as an exercise. Moreover, compact sets have a variety of useful characteristics, all of which can be seen as generalizations of corresponding properties of finite sets.

Theorem. *If K is a compact subset of the metric space (X, ρ), then K is closed and bounded in X.*

[2] You can think of X as a metric space with the metric topology or even more specifically as $\mathbb{R}$.

Proof. Choose a point $x_0 \in X$, then $K \subset X = \bigcup_{n\in\mathbb{N}} B(x_0, n)$, so $\{B(x_0, n)\}_{n\in\mathbb{N}}$ is an open cover of K. It follows from the compactness of K that there is a finite set of open balls $\{B(x_0, n_1), \dots, B(x_0, n_k)\}$, with $n_1 < \cdots < n_k$ in $\mathbb{N}$, such that

$$K \subset \bigcup_{j=1}^{k} B(x_0, n_j) = B(x_0, n_k),$$

so K is bounded.

We prove that K is closed by showing that K^c is open. Let $y \in K^c$ and note that if $x \in K$, then $\rho(x, y) > 0$, so there exists $n \in \mathbb{N}$ such that $\rho(x, y) > 1/n$. It follows that $\{R(y, 1/n)\}_{n\in\mathbb{N}}$ is an open cover of K, using the notation and conclusion of Proposition 3.1.2.[3] Since K is compact, there exist $n_1 < n_2 < \ldots < n_k$ in $\mathbb{N}$ such that

$$K \subset \bigcup_{j=1}^{k} R(y, 1/n_j) = R(y, 1/n_k),$$

and it follows that $\rho(x, y) > 1/n_k$ for all $x \in K$ so that $B(y, 1/n_k) \subset K^c$. Hence, K^c is open and K is closed. □

The proof that a compact subset of a metric space is bounded can be adjusted to show a bit more.

Corollary. *If X is a metric space and $K \subset X$ is compact, then for every $\varepsilon > 0$ there is a finite set $\{x_1, \dots, x_n\} \subset K$ such that*

$$K \subset \bigcup_{j=1}^{n} B(x_j, \varepsilon).$$

We say that K is *totally bounded* in X in this case.

Proof. Suppose that $K \subset X$ is compact and let $\varepsilon > 0$. The set $\{B(x, \varepsilon) : x \in K\}$ is an open cover of K, so it admits a finite subcover, $\{B(x_1, \varepsilon), \dots, B(x_n, \varepsilon)\}$. □

If E and F are subsets of a metric space X, then the distance between them is defined by

$$d(E, F) = \inf\{d(x, y) : x \in E,\ y \in F\}.$$

If E and F are disjoint, finite sets, then $d(E, F) > 0$. The same is true of disjoint compact sets.

Proposition. *If K_1 and K_2 are compact subsets of a metric space X, and $K_1 \cap K_2 = \emptyset$, then $d(K_1, K_2) > 0$.*

Proof. For every $x \in K_1$, let $\delta_x = \inf\{d(x, y) : y \in K_2\}$. If $\delta_x = 0$ for some $x \in K_1$, then there is a sequence of points $\{y_n\} \subset K_2$ such that $d(x, y_n) < 1/n$, whence $x \in K_2$ since K_2 is closed. This is impossible because $K_1 \cap K_2 = \emptyset$, so $\delta_x > 0$ for all $x \in K_1$.

[3] Where we defined $R(y, \varepsilon)$ to be the interior of $B(y, \varepsilon)^c$.

The set $\{B(x, \delta_x/2) : x \in K_1\}$ is an open cover of K_1, so there is a finite set $\{x_1, \ldots, x_n\} \subseteq K_1$ such that

$$K_1 \subseteq \bigcup_{j=1}^{n} B(x_j, \delta_{x_j}/2).$$

If $x \in K_1$ and $y \in K_2$, then $x \in B(x_j, \delta_{x_j}/2)$, for some $1 \le j \le n$, so $d(x,y) > \delta_{x_j}/2$, by the triangle inequality. It follows that $d(x,y) \ge \frac{1}{2}\min(\delta_{x_j} : 1 \le j \le n) > 0$, for all $x \in K_1$ and $y \in K_2$. ☐

3.2.2. The Heine-Borel Theorem

The converse of Theorem 3.2.1 is not true in all metric spaces—there are metric spaces containing closed and bounded sets that are not compact.

Example. Let $X = (0,1)$ with the standard metric inherited from $\mathbb{R}$. The set $F = \{1/n : n \ge 2\}$ is bounded because $F \subset B(1/2, 1)$. Furthermore, F doesn't have any limit points in X because $0 \notin X$, so F is closed in X.

For each $n \in \mathbb{N}$, we have $\frac{1}{n} - \frac{1}{n+1} = \frac{1}{n(n+1)} > \frac{1}{(n+1)^2}$, and therefore

$$\left(\frac{1}{n} - \frac{1}{(n+1)^2}, \frac{1}{n} + \frac{1}{(n+1)^2}\right) \cap F = \{1/n\}$$

for all $n \ge 2$. Hence $\left\{\left(\frac{1}{n} - \frac{1}{(n+1)^2}, \frac{1}{n} + \frac{1}{(n+1)^2}\right)\right\}_{n=2}^{\infty}$ is an open cover of F that does not contain a finite subcover, so F is not compact in X.

A metric space in which the converse of Theorem 3.2.1 is true is said to have the *Heine-Borel* property. The canonical example of such a space is $\mathbb{R}$.

Theorem (Heine-Borel). *A set $E \subset \mathbb{R}$ is compact if and only if it is closed and bounded.*

It follows from this theorem that every closed and bounded interval is compact in $\mathbb{R}$. These are the archetypical compact subsets of $\mathbb{R}$, and they play an important role in the proof of the theorem.

Proof. The first half of this statement, that compact sets are closed and bounded, has been proved in Theorem 3.2.1 for all metric spaces. We prove the converse, that closed and bounded sets in $\mathbb{R}$ are compact, by contradiction.

Suppose that E is closed and bounded in $\mathbb{R}$, and let $\mathcal{O} = \{O_\alpha\}_{\alpha \in A}$ be an open cover of E, that contains no finite subcover of E.

The boundedness of E implies that there is an interval $J_1 = [a,b]$, such that $E \subset J_1$. Let $J_1^1 = [a, (a+b)/2]$ be the left half of J_1 and let $J_1^2 = [(a+b)/2, b]$ be the right half. If both $E \cap J_1^1$ and $E \cap J_1^2$ admit finite subcovers from $\mathcal{O}$, then the union of these subcovers is a finite subcover of E (because $E = (E \cap J_1^1) \cup (E \cap J_1^2)$) so at least one of $E \cap J_1^1$ and $E \cap J_1^2$ fails to admit a finite subcover. We set $J_2 = J_1^1$ if $E \cap J_1^1$ does not admit a finite subcover from $\mathcal{O}$ and $J_2 = J_1^2$ otherwise.

Once again, we split J_2 into left and right closed half-intervals, J_2^1 and J_2^2. It follows as above that at least one of $E \cap J_2^1$ or $E \cap J_2^2$ must fail to admit a finite subcover from $\mathcal{O}$ and we set $J_3 = J_2^1$ if $E \cap J_2^1$ doesn't admit a finite subcover from $\mathcal{O}$ and $J_3 = I_2^2$ otherwise.

We continue in this way to recursively define a sequence of closed intervals $\{J_n\}$. Assume that $n \geq 1$ and that the intervals $J_1, J_2, \ldots, J_n$ have been chosen in such a way that (i) J_k is the closed left or right half of J_{k-1} for $2 \leq k \leq n$ and (ii) $E \cap J_k$ does not admit a finite subcover from $\mathcal{O}$ for $1 \leq k \leq n$. Letting J_n^1 and J_n^2 be the closed left and right half of J_n, we observe as before that either $E \cap J_n^1$ or $E \cap J_n^2$ must fail to admit a finite subcover from $\mathcal{O}$, and we set $J_{n+1} = J_n^1$ in the first case and $J_{n+1} = J_n^2$ otherwise.

This produces a sequence $\{J_n\}_{n\in\mathbb{N}}$, where each J_n is either the left or the right half of J_{n-1}, and for every n the set $E \cap J_n$ does not admit a finite subcover from $\mathcal{O}$. In particular, $E \cap J_n$ is not empty, and for each n, we choose a point $x_n \in E \cap J_n$.

For all n, we write $J_n = [a_n, b_n]$ (with $a_1 = a$ and $b_1 = b$), and note that by construction, $b_{n+1} - a_{n+1} = (b_n - a_n)/2$, and therefore $b_n - a_n = (b-a)/2^{n-1}$ for all $n \geq 1$, as follows by induction. Now, if $n, m \geq N$, then $x_n, x_m \in J_N$, because $J_n, J_m \subset J_N$ in this case, and therefore

$$|x_n - x_m| \leq b_N - a_N = \frac{b-a}{2^{N-1}}.$$

It follows that $\{x_n\}$ is a Cauchy sequence, and therefore converges to a limit $x^* \in \mathbb{R}$, because $\mathbb{R}$ is complete.

By Theorem 3.1.1, $x^* \in E$, because E is closed and $\{x_n\} \subset E$. Since $x^* \in E \subset \bigcup_{a\in A} O_a$, there exists $O_a \in \mathcal{O}$ such that $x^* \in O_a$, and so there is an $\varepsilon > 0$ such that $(x^* - \varepsilon, x^* + \varepsilon) \subset O_a$ (because O_a is open).

We now choose $N \in \mathbb{N}$ large enough so that both $|x_n - x^*| < \varepsilon/2$ for all $n \geq N$ and $(b-a)/2^{N-1} < \varepsilon/2$. With this choice, if $x \in J_N$, then $|x - x^*| \leq |x - x_N| + |x_N - x^*| < \varepsilon$, which implies that

$$E \cap J_N \subset J_N \subset (x^* - \varepsilon, x^* + \varepsilon) \subset O_a,$$

contradicting the assumption that $E \cap J_N$ does not admit a finite subcover from $\mathcal{O}$. This contradiction shows that E must be compact. □

From the Heine-Borel and Bolzano-Weierstrass theorems we have the following corollary.

Corollary. *A set $K \subset \mathbb{R}$ is compact if and only if every sequence $\{x_n\}$ in K has a subsequence that converges to a limit in K.*

I.e., compact subsets of $\mathbb{R}$ are *sequentially compact*.

Proof. Suppose that $K \subset \mathbb{R}$ is compact, and therefore closed and bounded. If $\{x_n\} \subset K$, then the sequence is bounded because K is bounded, and from Theorem 2.5.3 (Bolzano-Weierstrass) it follows that $\{x_n\}$ has a subsequence $\{x_{n_k}\}$ that converges to a limit $x^* \in \mathbb{R}$. Since K is also closed, it follows that $x^* \in K$ (by Theorem 3.1.1).

If K is not compact, then by the Heine-Borel theorem, K is not closed or K is not bounded. If K is not bounded, then for every $n \in \mathbb{N}$, there is an $x_n \in K$ such that $|x_n| > n$. The sequence $\{x_n\}$ does not have a convergent subsequence because for every $a \in \mathbb{R}$, the set $\{n \in \mathbb{N} : |x_n - a| < 1\}$ is finite.

If K is not closed, then there is a limit point x^* of K that does not belong to K. It follows that for every $n \in \mathbb{N}$, there is an $x_n \in K$ such that $|x_n - x^*| < 1/n$, and hence

$x_n \to x^*$, so $\lim_{k\to\infty} x_{n_k} = x^* \notin K$ for *every* subsequence $\{x_{n_k}\}$ of $\{x_n\}$. It follows that this sequence has no subsequence that converges to a point in K. □

⋆3.2.3. Sequential compactness and Lebesgue numbers

The property of compact subsets of $\mathbb{R}$ described in Corollary 3.2.2 has a name.

Definition. The subset E of the metric space (X, ρ) is *sequentially compact* if every sequence in E has a subsequence that converges to a point in E.

Using this language, Proposition 3.2.2 says that sets in $\mathbb{R}$ are compact if and only if they are sequentially compact. This statement generalizes to all metric spaces, but neither the Bolzano-Weierstrass theorem nor the Heine-Borel theorem are true in all metric spaces, so we need to use different tools to prove it. One of these tools is the *Lebesgue number lemma*.

Lemma. *If (X, ρ) is a metric space and $K \subset X$ is sequentially compact, then for every open cover $\mathcal{U} = \{U_a\}_{a\in A}$, there is a constant $\delta = \delta_{\mathcal{U},K} > 0$ such that for every $x \in K$, there is a $U \in \mathcal{U}$ that contains $B(x, \delta)$.*

The positive constant $\delta_{\mathcal{U},K}$ is called the *Lebesgue number* of the cover.

Proof. Let $K \subset X$ be sequentially compact and suppose to the contrary, that there is no $\delta > 0$ as claimed. Then for any $n \in \mathbb{N}$, there is an $x_n \in K$ such that $B(x_n, 1/n) \not\subset U_a$ for all $a \in A$.

By the sequential compactness of K, the sequence $\{x_n\}$ has a subsequence $\{x_{n_k}\}_{k=1}^{\infty}$ that converges to a limit $x^* \in K$. Since $\mathcal{U}$ is an open cover of K, there is a $U \in \mathcal{U}$ such that $x^* \in U$, and therefore there is an $\varepsilon > 0$ such that $B(x^*, \varepsilon) \subset U$.

Let $K \in \mathbb{N}$ be such that $\rho(x_{n_k}, x^*) < \varepsilon/2$ for all $k > K$. This and the triangle inequality together imply that $B(x_{n_k}, \varepsilon/2) \subset B(x^*, \varepsilon) \subset U$ for all $k > K$. Hence, if $k > \max(K, 2/\varepsilon)$, then $1/n_k \le 1/k < \varepsilon/2$, and therefore $B(x_{n_k}, 1/n_k) \subset B(x_{n_k}, \varepsilon/2) \subset U$. This contradicts the assumption we made about the elements of the sequence $\{x_n\}$. □

We can now prove the generalization of Corollary 3.2.2.

Theorem. *If (X, ρ) is a metric space then for any $K \subset X$, K is compact if and only if K is sequentially compact.*

Proof. First suppose that $K \subset X$ is compact, and that the sequence $\{x_n\} \subset K$ has no convergent subsequences.

Let $A_m = \{x_k : k > m\}$ be the subsequence of $\{x_n\}$ obtained by removing the first m elements, then the assumption that $\{x_n\}$ has no convergent subsequences has two implications for the subsequences A_m. First,

$$\bigcap_{m=1}^{\infty} A_m = \varnothing, \tag{3.2.1}$$

because if $x \in \bigcap_{m=1}^{\infty} A_m$, then for every $m \in \mathbb{N}$, there must be an $n_m > m$ such that $x_{n_m} = x$, and the *constant* subsequence $\{x_{n_m}\}$ would converge to x.

Second, for every $m \in \mathbb{N}$, A_m has no convergent subsequence because such a subsequence would also be a convergent subsequence of $\{x_n\}_{n=1}^{\infty}$. Hence, A_m has no limit points (see **ex3.2.5**) and is therefore closed, so $O_m = A_m^c$ is open and it follows from (3.2.1) that

$$\bigcup_{m=1}^{\infty} O_m = \left(\bigcap_{n=1}^{\infty} A_m\right)^c = X.$$

In other words, $\{O_m\}_{m\in\mathbb{N}}$ is an open cover of X and hence, an open cover of $K \subset X$.

From the compactness of K it follows that $\{O_n : n \in \mathbb{N}\}$ contains a finite subcover of K, and since $O_1 \subset O_2 \subset \ldots \subset O_n \subset \ldots$, it follows that $K \subset O_N = A_N^c$ for some $N \in \mathbb{N}$. This implies that $x_j \notin K$ for all $j > N$, contradicting the assumption that $\{x_n\} \subset K$. This contradiction shows that K must be sequentially compact.

In the other direction, suppose that $K \subset X$ is sequentially compact and let $\mathcal{U} = \{U_a\}_{a\in A}$ be an open cover of K. We want to show that $\mathcal{U}$ contains a finite subcover of K, but the difficulty is that we don't know anything about the nature of the open sets in $\mathcal{U}$. This is where the Lebesgue number lemma comes in.

Let $\delta = \delta_{\mathcal{U},K} > 0$ be the Lebesgue number for $\mathcal{U}$. The family $\mathcal{O} = \{B(x,\delta)\}_{x\in K}$ is also an open cover of K, and it follows from the Lebesgue number theorem that if $\mathcal{O}$ admits a finite subcover of K, $\{B(x_1,\delta),\ldots,B(x_n,\delta)\}$, then so does $\mathcal{U}$. This is because for each i, there is an open set $U_i \in \mathcal{U}$ such that $B(x_i,\delta) \subset U_i$ for $1 \le i \le n$, so

$$K \subset \bigcup_{i=1}^{n} B(x_i,\delta) \subset \bigcup_{i=1}^{n} U_i.$$

Suppose that $\mathcal{O}$ does not admit a finite subcover of K, then in particular, K is infinite. Let $x_1 \in K$, then from our assumption, there must be an $x_2 \in K$ such that $x_2 \notin B(x_1,\delta)$. Likewise, there must be an $x_3 \in K$ such that $x_3 \notin B(x_1,\delta) \cup B(x_2,\delta)$, and more to the point, $\rho(x_1,x_2) > \delta$, $\rho(x_1,x_3) > \delta$ and $\rho(x_2,x_3) > \delta$. We continue in this way to recursively construct a sequence in K: having found $x_1, \ldots, x_n \in K$ satisfying $\rho(x_i,x_j) > \delta$ for $1 \le i < j \le n$, the assumption that $\mathcal{O}$ does not admit a finite subcover of K implies that there is an $x_{n+1} \in K$ such that $x_{n+1} \notin \bigcup_{i=1}^{n} B(x_i,\delta)$.

The sequence $\{x_i\}_{i=1}^{\infty} \subset K$, thus constructed, has the property that $\rho(x_i,x_j) > \delta$ for all $i \neq j$. It follows that for any $x \in X$, the ball $B(x,\delta/2)$ contains at most one element x_j from this sequence and therefore the sequence has no convergent subsequence, contradicting the assumption that K is sequentially compact. Hence, $\mathcal{O}$ admits a finite subcover of K and therefore, as explained above, so does $\mathcal{U}$. □

*3.2.4. The Heine-Borel theorem in $\mathbb{R}^n$

In any metric space, a compact set is closed and bounded, as we proved in Theorem 3.2.1. In $\mathbb{R}$, the converse is also true — this is the Heine-Borel theorem. This statement generalizes to closed and bounded sets in $\mathbb{R}^n$.

This can be proved by mimicking the proof of Theorem 3.2.2, but using n-dimensional boxes instead of (1-dimensional) intervals. We leave this approach as an exercise, and instead rely on the fact that compactness and sequential compactness are equivalent in any metric space and on the Bolzano-Weierstrass theorem (in Euclidean space).

Theorem. *If $K \subset \mathbb{R}^n$ is closed and bounded, then K is sequentially compact, and hence compact.*

Proof. Suppose that $K \subset \mathbb{R}^n$ is closed and bounded and let $\{\mathbf{x}_j\}_{j=1}^{\infty}$ be a sequence in K. The sequence is bounded because K is bounded, and hence by Theorem* 2.5.6 (the Bolzano-Weierstrass theorem), it contains a subsequence $\{\mathbf{x}_{n_k}\}_{k=1}^{\infty}$ that converges to a limit $\mathbf{x}^*$. Since K is also closed, it follows that $\mathbf{x}^* \in K$. □

Exercises

ex3.2.1. Prove that finite subsets of $\mathbb{R}$ are compact.

ex3.2.2. Suppose that K_1 and K_2 are disjoint, compact sets in a metric space X. Show that there are disjoint, open sets $U, V \subseteq X$, such that $K_1 \subset U$ and $K_2 \subset V$.

ex3.2.3. Suppose that K_1 and K_2 are disjoint compact subsets of $\mathbb{R}$, with $d(K_1, K_2) = \delta > 0$. Show that if $(a,b) \cap (K_1 \cup K_2) \neq \varnothing$, then there is a closed subinterval $[c,d] \subset (a,b)$ such that $d - c > \delta/2$ and $[c,d] \cap (K_1 \cup K_2) = \varnothing$.

ex3.2.4. Let $\{K_n : n \in \mathbb{N}\}$ be a collection of compact subsets of $\mathbb{R}$ satisfying $\bigcap_{n=1}^{N} K_n \neq \varnothing$ for all $N \in \mathbb{N}$ (i.e., the intersection of any finite subcollection of these sets is nonempty).[4] Show that $\bigcap_{n=1}^{\infty} K_n \neq \varnothing$.

Hint: Prove this by contradiction: if $\bigcap_{n=1}^{\infty} K_n = \varnothing$, then $\mathbb{R} = \left(\bigcap_{n=1}^{\infty} K_n\right)^c$, so

$$K_1 \subset \mathbb{R} = \bigcup_{n=1}^{\infty} K_n^c,$$

which means that $\{K_n^c : n \in \mathbb{N}\}$ is an open cover of K_1 (why?). Now what?

ex3.2.5. Show that if K is a sequentially compact set in a metric space X, then K is closed and bounded. Do this directly from the definition of sequential compactness (not by quoting Theorems* 3.2.3 and 3.2.1.)

ex3.2.6. Suppose that X is a complete metric space and that X is totally bounded. Show that X is a compact metric space.

3.3. Small sets in $\mathbb{R}$

We conclude this chapter with a brief discussion of different notions of "small" for subsets of $\mathbb{R}$.

From a purely set theoretic point of view, cardinality is the standard measure of size. The integers and rational numbers are both countable sets, while all intervals in $\mathbb{R}$ have the same uncountable cardinality.[5] We think of countable subsets of $\mathbb{R}$ as small and uncountable subsets as big, and as we will see, countable subsets of $\mathbb{R}$ are small in every other sense that we describe below.

[4] This property of a collection of sets is called the *finite intersection property*.
[5] See Appendix 2.2.

3.3.1. Sets of measure zero

There is another, more natural way to describe the size of an interval, namely its length. The length of an interval takes the order and field structure on $\mathbb{R}$ into account. This notion of size (and its generalization to *measure*, that we discuss in Chapter 8) is important to the development of the integral (that we discuss in Chapters 7 and 8).

For an interval $I \subset \mathbb{R}$, we denote its length by $\ell(I)$. If I is bounded with endpoints $a < b$, then $\ell(I) = b - a$, whether I is open, closed or half open. If I is unbounded, e.g., $I = (a, \infty), (-\infty, b]$ or $(-\infty, \infty)$, then $\ell(I) = \infty$. We think of shorter intervals as smaller sets and longer intervals as larger sets in this context.

More generally, if a set can be covered by (i.e., is contained in) a union of intervals, the sum of whose lengths is small, then the set is small. This intuition leads to the following definition.

Definition. A set $E \subset \mathbb{R}$ has *measure zero* if for every $\varepsilon > 0$, there is a countable set of intervals $\{I_n\}$ such that covers E (i.e., $E \subset \bigcup I_n$) and $\sum \ell(I_n) \leq \varepsilon$.

If $E \subset \mathbb{R}$ contains an interval (a, b), then the sum of the lengths of any (countable) collection of intervals covering E must also cover (a, b). It follows in this case that if $E \subset \bigcup I_n$, then $\sum \ell(I_n) \geq (b - a)$. In other words, sets of measure zero cannot contain any intervals.

Proposition. *If $E \subset \mathbb{R}$ is countable, then E has measure zero.*

Since intervals are uncountable, it is certainly true that a countable set cannot contain an interval, but this doesn't prove that a countable set must have measure zero.

Proof. Suppose first that $E = \{x_1, \ldots, x_n\}$ is finite and let $\varepsilon > 0$. For $1 \leq j \leq n$, let $I_j = (x_j - \varepsilon/2n, x_j + \varepsilon/2n)$, then $x_j \in I_j$ for each $1 \leq j \leq n$, so $E \subset \bigcup_{j=1}^{n} I_n$. Moreover, $\sum_{j=1}^{n} \ell(I_n) = n \cdot 2\varepsilon/2n = \varepsilon$.

Now suppose that $E = \{x_n\}_{n\in\mathbb{N}}$ is countably infinite. Given $\varepsilon > 0$, let $I_n = (x_n - \varepsilon/3^n, x_n + \varepsilon/3^n)$ for $n \in \mathbb{N}$. Then $x_n \in I_n$ for each n, so $E \subset \bigcup_{n=1}^{\infty} I_n$, and

$$\sum_{n=1}^{\infty} \ell(I_n) = \sum_{n=1}^{\infty} \frac{2\varepsilon}{3^n} = 2\varepsilon \sum_{n=1}^{\infty} \frac{1}{3^n} = 2\varepsilon \cdot \frac{1}{2} = \varepsilon. \qquad \square$$

Thus for example, the set of rational numbers has measure zero, even though this set is dense in the set of real numbers.

A countable union of countable sets is countable.[6] The same is true for sets of measure zero.

Theorem. *If $E_n \subset \mathbb{R}$ has measure zero for every $n \in \mathbb{N}$, then $\bigcup_{n=1}^{\infty} E_n$ has measure zero.*

In other words, a countable union of sets of measure zero has measure zero.

[6] See Proposition A.2.1.

Proof. Let $\varepsilon > 0$ be given. For each $n \in \mathbb{N}$, E_n has measure zero, so there is a $J_n \subset \mathbb{N}$ and a collection of intervals $\{I_{n,j}\}_{j\in J_n}$ such that $E_n \subset \bigcup_{j\in J_n} I_{n,j}$ and

$$\sum_{j\in J_n} \ell(I_{n,j}) < \frac{\varepsilon}{2^n}.$$

The collection $\{I_{n,j} \,:\, n \in \mathbb{N} \text{ and } j \in J_n\}$ is a countable collection of intervals (as a countable union of countable sets) that covers $\bigcup_{n=1}^{\infty} E_n$ and

$$\sum_{n=1}^{\infty}\left(\sum_{j\in J_n} \ell(I_{n,j})\right) < \sum_{n=1}^{\infty} \frac{\varepsilon}{2^n} = \varepsilon.$$

It follows that $\bigcup_{n=1}^{\infty} E_n$ has measure zero, as claimed. □

*3.3.2. The Cantor set

There are uncountable sets of measure zero, the best known of which is the Cantor set. Starting with $C_0 = [0,1]$, let

$$C_1 = C_0 \setminus (1/3, 2/3) = [0,1/3] \cup [2/3,1],$$

i.e., C_1 is obtained by removing the *open middle third* of C_0. Next, C_2 is obtained by removing the open middle third of each of the two closed intervals comprising C_1:

$$C_2 = C_1 \setminus \big((1/9,2/9) \cup (7/9,8/9)\big) = [0,1/9] \cup [2/9,1/3] \cup [2/3,7/9] \cup [8/9,1].$$

Likewise, we obtain C_3 from C_2 by removing the open middle thirds of each of the four (disjoint) closed intervals comprising C_2, so that C_3 is the union of eight disjoint closed intervals, etc.

Definition. Let $C_0 = [0,1]$ and for $n \in \mathbb{N}$, let C_n be obtained from by removing the open middle third of each of the closed intervals whose disjoint union comprises C_{n-1}. The *Cantor set* is the set

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

Proposition. *For every $n \geq 0$, C_n is the union of 2^n disjoint, closed intervals, of length 3^{-n} each.*

Proof. The proof is by induction on n, beginning with the observation that statement is clearly true for $n = 0$. Now suppose that $n \geq 0$, the statement is true for n, and $I_1, \dots, I_{2^n}$ are the disjoint closed intervals comprising C_n. If $I_j = [a_j, b_j]$ for $1 \leq j \leq 2^n$, then removing the open middle third from I_j leaves us with the two closed intervals $[a_j, c_j]$ and $[d_j, b_j]$, where $c_j = a_j + 3^{-(n+1)}$ and $d_j = a_j + 2\cdot 3^{-(n+1)}$, since $b_j - a_j = 3^{-n}$. Thus, C_{n+1} is the union of $2 \cdot 2^n = 2^{n+1}$ closed intervals, each of length $3^{-(n+1)}$. Moreover, these intervals are disjoint because $I_1, \dots, I_{2^n}$ are disjoint. □

It follows from this proposition that C_n is closed, for all n, as a finite union of closed sets, and hence $\mathscr{C}$ is closed as an intersection of closed sets.

In what follows, we say *the endpoints of* C_n, to mean the endpoints of the 2^n intervals comprising C_n. We observe that it follows from the definition (as detailed in the proof of the lemma) that if c is a left (right) endpoint of C_n, then c is also a left (right) endpoint of C_m for all $m \geq n$. Moreover, every endpoint of C_n is either an endpoint or an interior point of C_{n-1}, for $n \geq 1$. Hence, for every n, the set of endpoints of C_n is contained in $\mathscr{C}$.

The elements of the Cantor set are characterized by their ternary expansions (see §2.9.2). We show this first for the endpoints in $\mathscr{C}$.

Lemma. *The number x is a left endpoint of C_n if and only if*

$$x = \sum_{k=1}^{\infty} \frac{t_k(x)}{3^k},$$

with $t_k(x) = 0$ or 2 for all k and $t_k = 0$ for all $k > n$. Similarly, x is a right endpoint of C_n, if and only if

$$x = \sum_{k=1}^{\infty} \frac{t_k(x)}{3^k},$$

with $t_k(x) = 0$ or 2 for all k and $t_k = 2$ for all $k > n$.

Proof. We prove this by induction on n, starting with $n = 0$. The endpoints of C_0 are 0, which satisfies the claim with $t_k = 0$ for all k, and 1 which satisfies the claim because

$$1 = \sum_{k=1}^{\infty} \frac{2}{3^k}.$$

Now suppose that $n \geq 1$, that the claim is true for $n - 1$ and that x is a right endpoint of C_n. The claim is true for right endpoints of C_{n-1}, so we may assume that x is not a right endpoint of C_{n-1}. In this case there is a left endpoint y of C_{n-1} such that

$$x = y + \frac{1}{3^n} = \sum_{k=1}^{n} \frac{t_k(y)}{3^k} + \frac{1}{3^n} = \sum_{k=1}^{n} \frac{t_k(y)}{3^k} + \sum_{k=n+1}^{\infty} \frac{2}{3^k}, \tag{3.3.1}$$

so x has the claimed ternary expansion. In the other direction, if x satisfies the identity (3.3.1), then $y = x - \frac{1}{3^n}$ must be a left endpoint of C_{n-1}, based on the induction hypothesis, so x is a right endpoint of C_n.

The inductive argument for left endpoints is analogous, and we leave it as an exercise. □

Corollary. *The number $x \in \mathscr{C}$ if and only if*

$$x = \sum_{n=1}^{\infty} \frac{t_n(x)}{3^n},$$

where $t_n(x) \in \{0, 2\}$ for all n.

Proof. Let $x \in [0, 1]$ have the ternary expansion

$$x = \sum_{n=1}^{\infty} \frac{t_n(x)}{3^n},$$

where $t_n(x) \in \{0,1,2\}$ for all n. In the case that x has two ternary expansions, one that is finite and terminates with a $1/3^m$, and one that terminates with $\sum_{n=m+1}^{\infty} 2/3^n$, we choose the latter.

If $x \in \mathscr{C}$, then it follows from the construction that for every $n \in \mathbb{N}$, $x \in I_j$ for one of the closed intervals comprising C_n, so there is an endpoint $y_n \in C_n$ such that $|x - y_n| \leq \frac{1}{2 \cdot 3^n}$. It follows that $t_n(x) = t_n(y_n)$ so $t_n(x) = 0$ or 2, by the lemma, and hence x has the claimed expansion.

If $t_n(x) \in \{0,2\}$ for all n, then

$$y_n = \sum_{k=1}^{n} \frac{t_k(x)}{3^k}$$

is a left endpoint in $\mathscr{C}$, and $y_n \to x$. Hence $x \in \mathscr{C}$ because $\mathscr{C}$ is closed. □

Theorem. *The Cantor set is an uncountable, closed set of measure zero.*

Proof. We have already shown that $\mathscr{C}$ is closed.

Next, for each $n \in \mathbb{N}$, $\mathscr{C} \subset C_n$ and the sum of the lengths of the intervals comprising C_n is $(2/3)^n$, so $\mathscr{C}$ has measure zero because $(2/3)^n \to 0$.

Finally, to show that $\mathscr{C}$ is uncountable, we construct a bijection from $\mathscr{P}(\mathbb{N})$ to $\mathscr{C}$.[7] Specifically, let $\Lambda = \{n_k : k \in \mathbb{N}\}$ be a (finite or infinite) subset of $\mathbb{N}$, with the elements listed in ascending order. If $\Lambda = \varnothing$, we set $\varphi(\Lambda) = 0 \in \mathscr{C}$. If $\Lambda = \{n_1, \ldots, n_k, \ldots\}$ is nonempty, we set

$$\varphi(\Lambda) = \sum_{n_k \in \Lambda} \frac{2}{3^{n_k}},$$

so $\varphi(\Lambda) \in \mathscr{C}$ in this case too, by the preceding corollary. In particular, $\varphi(\Lambda)$ is a left endpoint if Λ is finite and $\varphi(\Lambda)$ is a right endpoint if the complement of Λ is finite.

For any $x \in \mathscr{C}$, we have $x = \varphi(\Lambda_x)$, where $\Lambda_x = \{n \in \mathbb{N} : t_n(x) = 2\}$, so the map φ is surjective. If $\Lambda_1 \neq \Lambda_2$, let $n_0 \in \mathbb{N}$ be the smallest integer that is not in $\Lambda_1 \cap \Lambda_2$. Then,

$$|\varphi(\Lambda_1) - \varphi(\Lambda_2)| \geq \frac{2}{3^{n_0}} - \sum_{n=n_0+1}^{\infty} \frac{2}{3^n} = \frac{1}{3^{n_0}} > 0,$$

since $\sum_{n=n_0+1}^{\infty} \frac{2}{3^n} = \frac{1}{3^{n_0}}$, so φ is injective as well. □

*3.3.3. Nowhere dense sets

A set E is dense in $\mathbb{R}$ if $E \cap O \neq \varnothing$ for every nonempty open set $O \subset \mathbb{R}$. The set of rational numbers is an important example of a dense subset of $\mathbb{R}$. Sets which have the opposite property, in the sense described below, are considered to be small, topologically speaking.

Definition. A subset F of a metric space X is *nowhere dense* if for every ball $B(x,\varepsilon) \subset X$, there is a ball $B(y,\delta) \subset B(x,\varepsilon)$ such that $F \cap B(x,\varepsilon) = \varnothing$. In other words a nowhere dense set misses an open piece of every open set.

[7] $\mathscr{P}(\mathbb{N})$ is the set of all subsets of $\mathbb{N}$, which is uncountable. See the Appendix.

A single point is clearly nowhere dense in $\mathbb{R}$, and more generally any finite set is nowhere dense. Some countable sets are nowhere dense in $\mathbb{R}$, like $\mathbb{Z}$, while others are not, like $\mathbb{Q}$. There are also uncountable sets that are nowhere dense.

Proposition. *The Cantor set is nowhere dense in* $\mathbb{R}$.

Proof. Let I be an arbitrary, nonempty open interval in $[0,1]$. We need to show that there is a nonempty open interval $J \subset I$ such that $J \subset \mathscr{C}^c$. If $I \cap \mathscr{C} = \emptyset$, then we can set $J = I$ and there is nothing more to prove. Otherwise, $I \cap \mathscr{C} \neq \emptyset$ and hence, $I \cap C_n \neq \emptyset$ for all $n \in \mathbb{N}$ (referring to the sets described in the previous section).

In this case, let $n \in \mathbb{N}$ be large enough that $\ell(I) > 1/3^n$, and denote by $\{I_k : 1 \leq k \leq 2^n\}$ the collection of intervals comprising C_n. Since $I \cap C_n \neq \emptyset$, it follows that $I \cap I_k \neq \emptyset$ for some k. The interval I cannot be properly contained in I_k because $\ell(I_k) = 1/3^n < \ell(I)$ and writing $I_k = [r_k, s_k]$, it follows that either $r_k \in I$ or $s_k \in I$.

Suppose that $r_k \in I$, then since I is open there is an $\varepsilon > 0$ such that $(r_k - \varepsilon, r_k + \varepsilon) \subset I$. Since the open interval $(r_k - 1/3^n, r_k) \subset \mathscr{C}^c$, it follows that $J = (r_k - \delta, r_k) \subset I \cap \mathscr{C}^c$, where $\delta = \min(\varepsilon, 1/3^n)$.[8] The same argument shows that if $s_k \in I$, then there is an $\eta > 0$ such that $J = (s_k, s_k + \eta) \subset I \cap \mathscr{C}^c$. □

Nowhere dense sets have certain properties in common with closed sets.

Theorem. *The closure of a nowhere dense set is nowhere dense and the union of finitely many nowhere dense sets is nowhere dense.*

Proof. Suppose (X, ρ) is a metric space and that $E \subset X$ is nowhere dense. Given any open ball $B_1 = B(x, \varepsilon) \subset X$, there is a second ball $B_2 = B(y, \delta) \subset B_1$, such that $B_2 \subset E^c$. If $z \in B_2$ then there is some $\eta > 0$, such that $B(z, \eta) \subset B_2 \subset E^c$, since B_2 is open. It follows that z is not a limit point of E for every $z \in B_2$, so $B_2 \cap \overline{E}$ is empty and hence $\overline{E}$ is nowhere dense.

We prove that a finite union of nowhere dense sets is nowhere dense by induction on the number n of sets in the union.

If $n = 1$ then there is nothing to prove, so we consider first the case that $n = 2$. Let E_1 and E_2 be nowhere dense sets and let $B = B(x, \varepsilon) \subset X$ be a nonempty open ball. Then there is a ball $B_1 = B(y_1, \delta) \subset B$ such that $B_1 \cap E_1 = \emptyset$. Likewise, there is a ball $B_2 = B(y_2, \eta) \subset B_1$ such that $B_2 \cap E_2 = \emptyset$. It follows that $B_2 \cap (E_1 \cup E_2) = \emptyset$, and since $B_2 \subset B$, it follows that $E_1 \cup E_2$ is nowhere dense.

Assume now that $n \geq 2$ and that the claim is true for n. If $E_1, \ldots E_{n+1}$ are nowhere dense sets, then $\bigcup_{j=1}^{n} E_j$ is nowhere dense by the induction hypothesis, and therefore

$$\bigcup_{j=1}^{n=1} E_j = \left(\bigcup_{j=1}^{n} E_j \right) \cup E_{n+1}$$

is nowhere dense by the case $n = 2$. □

[8] See the paragraph preceding Theorem 3.3.2.

*3.3.4. Baire's category theorem

One of the indications that countable sets and sets of measure zero are small is that a countable union of such sets, is still small. I.e., A countable union of countable sets is still countable, and a countable union of sets of measure zero still has measure zero.

The union of finitely many nowhere dense sets is nowhere dense, but a countable union of nowhere dense sets need not be. For example $\mathbb{Q}$ is a countable union of singletons (each of which is nowhere dense), but $\mathbb{Q}$ is dense. Nonetheless, countable unions of nowhere dense sets are still topologically small, and they are important enough to have a name.

Definition. A subset E of a metric space X is of the *first category* if E is a countable union of nowhere dense sets. Otherwise, E is of the *second category*.

Since a countable union of countable sets is countable, it follows that a countable union of sets of first category is also a countable union of nowhere dense sets. This means that the collection of sets of first category is closed under countable unions like countable sets and sets of measure zero. Baire's category theorem shows that sets of first category are also small in the sense that no set of first category can cover $\mathbb{R}$. In fact this statement is true for all complete metric spaces.

Theorem (Baire). *If X is a complete metric space then X is of the second category.*

Proof. Let X be a complete metric space and let $\{F_n\}_{n\in\mathbb{N}}$ be a countable collection of nowhere dense sets. We want to show that $X \neq \bigcup F_n$. We will do this by constructing a Cauchy sequence $\{x_n\} \subset X$ whose limit x^* is not in $\bigcup F_n$. Throughout the proof, we rely on the fact that $\bigcup_{j=1}^{n} F_j$ is nowhere dense for each $n \in \mathbb{N}$

Let $x_0 \in X$ be an arbitrary point and let $B_0 = B(x_0, 1)$. Since F_1 is nowhere dense, there is an $x_1 \in B_0$ and an $\varepsilon_1 \in (0, 1/2)$, such that $B(x_1, \varepsilon_1) \subset B_0$ and $F_1 \cap B(x_1, \varepsilon_1) = \varnothing$.[9] We have $\rho(y, x_1) \geq \varepsilon_1$ for all $y \in F_1$, and setting $B_1 = B(x_1, \varepsilon_1/2)$, it follows from the triangle inequality that for all $y \in F_1$ and $x \in B_1$,

$$\rho(y, x) \geq \rho(y, x_1) - \rho(x, x_1) \geq \varepsilon_1 - \varepsilon_1/2 = \varepsilon_1/2.$$

Since $F_1 \cup F_2$ is nowhere dense, there is an $x_2 \in B_1$ and an $\varepsilon_2 \in (0, \varepsilon_1/2)$ such that $B(x_2, \varepsilon_2) \subset B_1$ and $(F_1 \cup F_2) \cap B(x_2, \varepsilon_2) = \varnothing$. Setting $B_2 = B(x_2, \varepsilon_2/2)$, the same reasoning we used above shows that $\rho(y, x) \geq \varepsilon_2/2$ for all $y \in F_1 \cup F_2$ and all $x \in B_2$.

Summarizing, after two steps we have open balls

$$B_2 = B(x_2, \varepsilon_2/2) \subset B(x_1, \varepsilon_1/2) = B_1,$$

with $\varepsilon_2/2 < (\varepsilon_1/2)/2 < 1/4$, such that $\rho(x, y) \geq \varepsilon_2/2$ for all $x \in B_2$ and $y \in F_1 \cup F_2$.

Suppose now that $n \geq 2$ and we have found points $x_1, x_2, \ldots, x_n \in X$ and balls $B_1 \supset B_2 \supset \ldots \supset B_n$, with $B_k = B(x_k, \varepsilon_k/2)$ for $1 \leq k \leq n$, and $\varepsilon_{k+1} < \varepsilon_k/2$ for $1 \leq k \leq n-1$, such that $\rho(x, y) \geq \varepsilon_n/2$ for all $x \in B_n$ and all $y \in \bigcup_{j=1}^{n} F_j$. Since $\bigcup_{j=1}^{n+1} F_j$ is nowhere dense, there is an $x_{n+1} \in B_n$ and an $\varepsilon_{n+1} \in (0, \varepsilon_n/2)$ such that $B(x_{n+1}, \varepsilon_{n+1}) \subset B_n$ and $\left(\bigcup_{j=1}^{n+1} F_j\right) \cap B(x_{n+1}, \varepsilon_{n+1}) = \varnothing$. Setting $B_{n+1} = B(x_{n+1}, \varepsilon_{n+1}/2)$, the same reasoning

[9] If $F_1 \cap B(x_1, \varepsilon) = \varnothing$ for some $\varepsilon \geq 1/2$, then it is certainly true that $F_1 \cap B(x_1, \varepsilon) = \varnothing$ for all $\varepsilon < 1/2$.

we have used twice before shows that $\rho(y,x) \geq \varepsilon_{n+1}/2$ for all $y \in \bigcup_{j=1}^{n+1} F_j$ and all $x \in B_{n+1}$.

In this way, we recursively construct a sequence of open balls, $\{B_n\}_{n=1}^{\infty} = \{B(x_n, \varepsilon_n/2)\}_{n=1}^{\infty}$, with $\varepsilon_n/2 > \varepsilon_{n+1}$ and $B_{n+1} \subset B_n$ for all n. For every $n \in \mathbb{N}$, these balls also have the property that for all $x \in B_n$ and for all $y \in \bigcup_{j=1}^{n} F_j$, we have $\rho(x,y) \geq \varepsilon_n/2$.

Since $\varepsilon_1 < 1/2$ and $\varepsilon_{n+1} < \varepsilon_n/2$ for all $n \in \mathbb{N}$, it follows by induction on n that $\varepsilon_n < 1/2^n$ for all n. Hence, if $n, m > N$, then $x_n, x_m \in B_N$ and therefore

$$\rho(x_n, x_m) \leq \rho(x_n, x_N) + \rho(x_N, x_m) < \varepsilon_N/2 + \varepsilon_N/2 = \varepsilon_N < 1/2^N.$$

It follows that $\{x_n\}$ is a Cauchy sequence in X and hence converges to a limit $x^* \in X$ because X is complete.

If $y \in \bigcup_{n=1}^{\infty} F_n$, then $y \in \bigcup_{n=1}^{n_0} F_n$ for some $n_0 \in \mathbb{N}$ and therefore $\rho(y,x) \geq \varepsilon_{n_0}/2$ for all $x \in B_{n_0}$. Let n be large enough that $n > n_0$ and $\rho(x_n, x^*) < \varepsilon_{n_0}/2$. With this choice of n, we have $x_n \in B_{n_0}$ and therefore

$$\rho(y, x^*) \geq \rho(y, x_n) - \rho(x_n, x^*) > \varepsilon_{n_0}/2 - \varepsilon_{n_0}/2 = 0,$$

so $y \neq x^*$. It follows that $x^* \notin \bigcup_{n=1}^{\infty} F_n$, so $\bigcup_{n=1}^{\infty} F_n \neq X$. □

⋆3.3.5. A partition of ℝ into small sets

No countable union of sets of measure zero can cover ℝ and no countable union of sets of first category can cover ℝ. One way of understanding this is that sets of measure zero are all small in the same way, so combining them in a countable union produces another set of the same type. The same can be said for sets of the first category.

On the other hand, sets of measure zero are small in a different way than sets of first category. In particular, a set of measure zero can be of second category, and a set of first category can have “full measure”.[10] We illustrate this by finding a subset $A \subset \mathbb{R}$ of measure zero whose complement $B = \mathbb{R} \setminus A$ is of first category.

We begin with a simple observation.

Lemma. *A subset F of a metric space X is closed and nowhere dense in X if and only if F^c is open and dense in X.*

Note that the complement of an arbitrary dense set need not be nowhere dense. For example, the complement of ℚ in ℝ is dense.

Proof. Suppose that F is closed and nowhere dense in X. Then $G = F^c$ is open and for any ball $B_0 = B(x, \varepsilon) \subset X$, there is a ball $B_1 = B(y, \delta) \subset B_0$ such that $F \cap B_1 = \varnothing$, so that $B_1 \subset G$, showing that G is dense.

In the other direction, if G is open and dense in X then $F = G^c$ is closed. If B_0 is an open ball in X, then there is a point $x_0 \in G \cap B_0$ since G is dense. Since $B_0 \cap G$ is open, there is a $\delta > 0$ such that $B_1 = B(x_0, \delta) \subset B_0 \cap G$, and hence $F \cap B_1 = \varnothing$. □

From this and Baire’s theorem we have the following corollary.

[10] See Chapter 8.

Corollary. *If X is a complete metric space, and G_n is a dense open set in X for every $n \in \mathbb{N}$, then $\bigcap_{n\in\mathbb{N}} G_n$ is not empty.*

Proof. If G_n is a dense open set, then G_n^c is a nowhere dense and therefore, by Baire's category theorem,

$$\left(\bigcap_{n=1}^{\infty} G_n\right)^c = \bigcup_{n=1}^{\infty} G_n^c \neq X.$$

Hence $\bigcap_{n=1}^{\infty} G_n$ is not empty. □

Theorem. *There exist sets $A, B \subset \mathbb{R}$ such that A has measure zero, B is of first category and $\mathbb{R} = A \cup B$.*

Proof. The rational numbers form a countable set, so we can list them in a sequence: $\mathbb{Q} = \{r_k\}_{k=1}^{\infty}$. For each $n \in \mathbb{N}$, we use this sequence to form the set

$$G_n = \bigcup_{k=1}^{\infty} \left(r_k - \frac{1}{n^k}, r_k + \frac{1}{n^k}\right) = \bigcup_{k=1}^{\infty} I_{n,k}.$$

The set G_n is open for every n, as a union of open intervals, and it is dense in $\mathbb{R}$ because $\mathbb{Q} \subset G_n$. Writing $A = \bigcap_{n\in\mathbb{N}} G_n$, it follows from Lemma 3.3.5 that

$$B = A^c = \left(\bigcap_{n\in\mathbb{N}} G_n\right)^c = \bigcup_{n\in\mathbb{N}} G_n^c$$

is of first category. It remains to show that A has measure zero.

Let $\varepsilon > 0$ be given, and let $n_\varepsilon \in \mathbb{N}$ be such that $2/(n_\varepsilon - 1) < \varepsilon$. By construction, we have $A \subset G_{n_\varepsilon} = \bigcup_{k\in\mathbb{N}} I_{n_\varepsilon,k}$, and

$$\sum_{k=1}^{\infty} \ell(I_{n_\varepsilon,k}) = \sum_{k=1}^{\infty} \frac{2}{n_\varepsilon^k} = \frac{2}{n_\varepsilon - 1} < \varepsilon.$$

Hence, A has measure zero. □

Exercises

ex3.3.1. Complete the proof of Corollary 3.3.2.

ex3.3.2. Prove that $\mathscr{C}$ is uncountable by showing that

$$\psi(x) = \sum_{n=1}^{\infty} \frac{t_n(x)/2}{2^n}$$

is a *surjection* from $\mathscr{C}$ to $[0,1]$. It is understood here that if $x \in \mathscr{C}$ has two ternary expansions, we choose the one without 1s.

ex3.3.3. A closed subset F of a metric space X is a *perfect set* if every point in F is a limit point of F. Show that the Cantor set is perfect.

ex3.3.4. Generalize Theorem* 3.3.3 to any metric space X.

ex3.3.5. Prove Corollary* 3.3.5.