Introduction to Analysis

William Fisher

Summer 2024

Contents

1 Introduction

The goals of these notes is to rigorously develop the theory of calculus and explore the definitions and concepts encountered along the way. They are heavily inspired by Walter Rudin's *Principles of Mathematical Analysis*.

2 The real numbers

The goal of this section is to rigorously define and study the set of real numbers and its structure. At first glance, this may appear like a superfluous endeavor. If I asked you to define the real numbers, you would probably be at a loss, but we all *know* what we mean, so why dwell?

Consider the following basic issue: Early in one's math career one often encounters the fact that

$$0.99999... = 1.$$

But suppose we weren't convinced of this. How would we add in a sensical manner? If decimal expansions are meant to exist then surely

$$1/3 = 0.33333...,$$

but if 1/3 is to mean anything then surely we also have

$$1 = 3 \cdot (1/3) = 0.99999....$$

We thus immediately see that non-obvious facts such as $0.9999... = 1$ are already integral to having a proper definition of $\mathbb{R}$ where both decimal expansion and addition makes sense.

On another level, we also need a proper definition of $\mathbb{R}$ to make sense of some of our favorite operations. Consider for example $\sqrt{\cdot}$. Morally $\sqrt{\cdot}$ should take in a real number and give back a number which squares to its input. Let us focus on $\sqrt{2}$.

Proposition 2.0.1. *$\sqrt{2}$ is not a rational number, i.e. there is no rational number whose square is* 2.

Proof. Suppose we could write

$$\sqrt{2} = \frac{a}{b}$$

where a/b is in reduced form, i.e. a and b share no common factors. Then we would have

$$2b^2 = a^2.$$

From this we see that a^2 is even, which is only possible if a is itself even. Thus write $a = 2n$ for some integer n. This then gives

$$2b^2 = (2n)^2 = 4n^2$$

so

$$b^2 = 2n^2.$$

However, this shows that b^2 is even and hence so is b. But then a/b was not in reduced form, since a and b are both even, a contradiction. $\square$

This proposition shows that $\sqrt{2}$ is an example of a real number which is not rational. This also tells us that there is something special about $\mathbb{R}$—why can we be certain that square roots exist in $\mathbb{R}$ and therefore define $\sqrt{\cdot}$? All of these concerns require a rigorous definition of $\mathbb{R}$ to address.

2.1 Fields and ordered sets

Our construction of $\mathbb{R}$ will be motivated as follows: First we will use our working understanding of $\mathbb{R}$ to decide what a rigorous definition *should* look like. We will then formulate the definition of an object worthy of being called the real numbers and prove that there is a unique such object satisfying this definition, which we will from there on out refer to as $\mathbb{R}$.

In this section we will abstract two essential structures of $\mathbb{R}$: the ability to add, multiply and divide elements as well as the ability to compare "sizes" of elements. The first of these structures we capture with the following definition.

Definition 2.1.1. A *field* is a set F with three operations—addition $(+)$, multiplication $(\cdot)$, and division satisfying the following axioms.

(A) (i) There exists an element of F denoted by 0 such that for all $a \in F$,

$$a + 0 = 0 + a = a.$$

0 is referred to both as *zero* and as the *additive identity* of F.

(ii) Addition is *associative*, that is for any three elements $a, b, c \in F$ we have

$$a + (b + c) = (a + b) + c.$$

(iii) Addition is *commutative*, that is for any two elements $a, b \in F$ we have

$$a + b = b + a.$$

(iv) Addition has *inverses*, i.e. for all $a \in F$ there exists some $b \in F$ such that

$$a + b = 0.$$

b is often denoted by $-a$ and we will write $c - d$ to mean $c + (-d)$.

(M) (i) There exists an element of F denoted by 1 such that for all $a \in F$,

$$a \cdot 1 = 1 \cdot a = a.$$

1 is referred to both as *one* and as the *multiplicative identity*

(ii) Multiplication is associative, i.e. for all $a, b, c \in F$

$$a \cdot (b \cdot c) = (a \cdot b) \cdot c.$$

(iii) Multiplication is commutative, i.e. for all $a, b \in F$

$$a \cdot b = b \cdot a.$$

(iv) Multiplication is *distributive* over addition, i.e. for all $a, b, c \in F$ we have

$$a \cdot (b + c) = (a \cdot b) + (a \cdot c).$$

(D) For all non-zero $a \in F$, i.e. all $a \neq 0$, there exists a $b \in F$ such that $a \cdot b = 1$. b is often denoted by $1/a$ and we will write c/d for $c \cdot (1/d)$ provided $d \neq 0$.

Remark 2.1.2. Often we will abuse notation by writing ab for $a \cdot b$ and writing expressions such as $ab + c$ instead of $(a \cdot b) + c$ with the understanding that multiplication is done before addition.

Example 2.1.3. The integers

$$\mathbb{Z} = \{0, 1, -1, 2, -2, 3, -3, \dots\}$$

have a multiplication and addition which satisfy axioms (A) and (M) of Definition 2.1.1 but does not satisfy (D), the ability to divide. Such an object is called a *ring*.

If we take the integers and allow division we get the rational numbers $\mathbb{Q}$ which should be your canonical example of a field.

Example 2.1.4. Let $\mathbb{Z}/k$ be the set

$$\mathbb{Z}/k = \{0, 1, \dots, k-1\}.$$

Define addition and multiplication for $x, y \in \mathbb{Z}/k$ by doing the usual operation as integers and then taking the remainder given by dividing by k. For example, in $\mathbb{Z}/11$ we have

$$7 \cdot 10 = 4$$

since $7 \times 10 = 70$ and 70 has a remainder of 4 after division by 11. Then we have that $\mathbb{Z}/3$, $\mathbb{Z}/5$, etc... is a field.

In fact, $\mathbb{Z}/k$ is a field if and only if k is prime.

Proposition 2.1.5. *If F is a field, then the additive and multiplicative identities are unique. In particular, in makes sense to refer to them as* 0 *and* 1.

Similarly, additive and multiplicative inverses are unique. In particular, it makes sense to denote them by $-a$ and $1/a$ and refer to them as the *additive and multiplicative inverse of a.*

Proof. Suppose we have two $e_1, e_2 \in F$ such that for all $a \in F$,

$$a + e_1 = e_1 + a = a \tag{2.1.1}$$

$$a + e_2 = e_2 + a = a. \tag{2.1.2}$$

Then

$$\begin{aligned} e_1 &= e_1 + e_2 \quad &\text{(by (2.1.2))} \\ &= e_2 &\text{(by (2.1.1))} \end{aligned}$$

so $e_1 = e_2$ as required. The proof is similar for uniqueness of the multiplicative identity.

For additive inverses, let $a \in F$. Suppose $b_1, b_2 \in F$ are such that

$$b_1 + a = a + b_1 = 0 \tag{2.1.3}$$

$$b_2 + a = a + b_2 = 0. \tag{2.1.4}$$

We want to show $b_1 = b_2$. Well,

$$\begin{aligned} b_1 &= b_1 + 0 && \text{(by definition of 0)} \\ &= b_1 + (a + b_2) && \text{(by (2.1.4))} \\ &= (b_1 + a) + b_2 && \text{(associativity)} \\ &= 0 + b_2 && \text{(by (2.1.3))} \\ &= b_2 \end{aligned}$$

as required. The proof is similar for multiplicative inverses. □

Proposition 2.1.6. *Let F be a field. Then the following statements hold for all $x, y, z \in F$:*

(i) If $x + y = x + z$ then $y = z$

(ii) If $x + y = x$ then $y = 0$

(iii) $-(-x) = x$

(iv) If $x \neq 0$ and $xy = xz$ then $y = z$

(v) If $x \neq 0$ and $xy = x$ then $y = 1$

(vi) If $x \neq 0$ then $1/(1/x) = x$.

Proof. We prove only some of these. For (i), we have that

$$x + y = x + z. \tag{2.1.5}$$

Adding $-x$ to both sides of (2.1.5) we get

$$(-x) + (x + y) = (-x) + (x + z). \tag{2.1.6}$$

However, by associativity, we have that

$$\begin{aligned} (-x) + (x + y) &= ((-x) + x) + y \\ &= 0 + y \\ &= y. \end{aligned}$$

Similarly, $(-x) + (x + z) = z$, so (2.1.6) gives $y = z$ as required.

Statement (ii) follows from (i) by taking $z = 0$.

We have that $x + (-x) = 0$ by definition. However, this also tells us that x is the inverse of $-x$, so $-(-x) = x$.

Statements (iv)-(vi) are proven similarly to the above. □

Proposition 2.1.7. *If F is a field then for all $a, b \in F$ the following statements hold:*

(i) $0 \cdot a = 0$

(ii) $(-a)b = -(ab)$.

Proof. We first prove (i). Notice that we have

$$\begin{aligned} 0 \cdot a + 0 \cdot a &= (0 + 0) \cdot a \\ &= 0 \cdot a. \end{aligned}$$

Then by Proposition 2.1.6(ii) we get that $0 \cdot a = 0$.

For (ii), it suffices to check that $(-a)b + ab = 0$ as we are then done by the defining property of $-(ab)$. By commutativity and distributivity of multiplication we have

$$\begin{aligned} (-a)b + ab &= (-a + a)b \\ &= 0 \cdot b \\ &= 0 \end{aligned}$$

where on the last line we used part (i). □

Looking at the above definition and comparing our intuition for what $\mathbb{R}$ should be, we see that any rigorous definition of $\mathbb{R}$ should have the structure of a field.

Next, we abstract the essential properties that allows us to compare real numbers. We first do this by defining a general notion of an ordering on a set.

Definition 2.1.8. An *ordered set* is a set S with a relation $<$ such that for all $a, b, c \in S$ the following holds:

(i) Exactly one of $a < b$, $b < a$ or $a = b$ holds

(ii) If $a < b$ and $b < c$, then $a < c$

We will write $a \leq b$ to mean that either $a < b$ or $a = b$ and write $a > b$ to mean $b < a$.

If we also have a field structure on our set S, then we would like this ordering to be compatible with the field operations.

Definition 2.1.9. An *ordered field* is a field F which is also an ordered set such that

(i) for all $a, b, c \in F$, if $a < b$ then $a + c < b + c$

(ii) for all $a, b, c \in F$ with $a \leq b$ and $0 \leq c$ we have $ac \leq bc$.

Given $x \in F$, if $x > 0$ we say that x is *positive* and if $x < 0$ we say that x is *negative*.

Example 2.1.10. We have already seen in Example 2.1.3 that $\mathbb{Q}$ is a field. The standard ordering on $\mathbb{Q}$ also gives $\mathbb{Q}$ the structure of an ordered field.

Example 2.1.11. The field structure on the complex numbers $\mathbb{C}$ and fields $\mathbb{Z}/p$ (see Example 2.1.4) cannot be given the structure of an ordered field. (Exercise: Prove this.)

Proposition 2.1.12. *If F is an ordered field then the following statements hold for $a, b, c \in F$:*

(i) if $a \geq 0$ then $-a \leq 0$

(ii) if $a \leq b$ and $c \leq 0$ then $ac \geq bc$

Proof. We first prove (i). By property (i) in Definition 2.1.9 we have that

$$a \geq 0 \Longrightarrow a + (-a) \geq -a$$

but $a + (-a) = 0$ so the second relation says $0 \geq -a$ as desired.

For (ii), by part (i) we know that since $c \leq 0$ we have $-c \geq 0$. Thus by property (ii) of Definition 2.1.9 we have

$$a \leq b \Longrightarrow (-c)a \leq (-c)b.$$

Now by Proposition 2.1.7 this second relation is just $-ca \leq -cb$. Now adding $ca + cb$ to both sides and using property (i) of Definition 2.1.9 we get the result. □

Definition 2.1.13. Let F be an ordered field (or simply an ordered set) and suppose that $S \subseteq F$. We say that $a \in F$ is an *upper bound* for S if for every $s \in S$ we have that $s \leq a$. We say that a is a *least upper bound* (or *supremum*) for S if a is an upper bound of S and for any other upper bound b of S, $a \leq b$. We denote the least upper bound of S (if it exists) by $\sup S$.

Similarly we define lower bounds and greatest lower bounds (called *infimums*) for S. We denote the greatest lower bound of S (if it exists) by $\inf S$.

Example 2.1.14. The infimum or supremum may or may not belong to the set S itself if it exists. Consider $\mathbb{Q}$ with its usual ordering. Then the infimum of

$$S = \{1/n : n \in \mathbb{N}\}$$

is $0 \in \mathbb{Q}$ which does not belong to S.

However, the infimum of $S = \{x \in \mathbb{Q} : x \geq 0\}$ is 0 which does belong to S.

Proposition 2.1.15. *If* $\sup S$ *or* $\inf S$ *exists then it is unique.*

Proof. Exercise to the reader. □

Definition 2.1.16. Let F be an ordered field (or simply an ordered set). We say that F has the *least-upper-bound property* if for every non-empty $S \subseteq F$ with an upper bound, $\sup S$ exists.

Theorem 2.1.17. *If F has the least-upper-bound property, then for every non-empty set $S \subseteq F$ which is bounded below,* $\inf S$ *exists.*

In fact, if L denotes the set of all lower bounds of S, then $\inf S = \sup L$.

Proof. We have assumed that S is bounded below, so L is non-empty. Moreover, since S is non-empty, there exists some $s \in S$. Since every $x \in L$ is a lower bound for S, $x \leq s$. Hence s is an upper bound for L. Thus L is bounded above and non-empty, so $\sup L$ exists. Set $\alpha = \sup L$.

We wish to show that α is a lower bound for S. The above shows that every element of S is an upper bound for L. Thus for every $x \in B$, since α is the *least* upper bound for L, $\alpha \leq x$. Thus $\alpha \in L$.

Moreover, since $\alpha = \sup L$ it is an upper bound for L so it is greater than or equal to every lower bound of S. Hence $\alpha = \inf S$ by definition. □

Remark 2.1.18. Following this section, we will assume minor results about how arithmetic works in fields similar to Propositions 2.1.5 and 2.1.7 without proof, and encourage the reader to prove them themselves.

2.2 Construction of the real numbers

In the previous section, we saw that any good definition of $\mathbb{R}$ should be such that $\mathbb{R}$ is an ordered field. However, Example 2.1.10 shows that $\mathbb{Q}$ is also an ordered field which means that this alone is not enough to determine a proper definition of the real line. In this section we will motivate what addition property should determine $\mathbb{R}$ and then prove its existence and uniqueness.

Example 2.2.1. Consider the set

$$S = \{a \in \mathbb{Q} : a^2 < 2\} \subseteq \mathbb{Q}.$$

Clearly S is bounded above, as any $a \in S$ must have $a \leq 2$, for example. However, we prove that S has no least upper bound in $\mathbb{Q}$.

Indeed, suppose that a is an upper bound for S. We must then have $a > 0$ (since, e.g., $1/2 \in S$ and a is an upper bound for S). Then by Proposition 2.0.1 we know that $a^2 \neq 2$. If we were to have $a^2 < 2$ then we would be able to slightly increase the size of a to get some $b > a$ such that $b^2 < 2$ still holds (check this!). Therefore we'd have $b \in S$ but $a < b$ violating the assumption that a is an upper bound for S. On the other hand, if $a^2 > 2$ then we could slightly decrease the size of a to find some $0 < b < a$ with $b^2 > 2$ (check this!). But then for any $c \in S$ we would have

$$c^2 < 2 < b^2$$

which forces $c < b$ (check this!). Hence b is still an upper bound for S, but with $b < a$ meaning that a was not a least upper bound. Therefore S has no least upper bound.

From the above argument, we also see that if a least upper bound $s = \sup S$ were to exist, it should satisfy $s^2 = 2$ and the fact that $\sup S$ does not exist essentially boils down to Proposition 2.0.1 which says that $\sqrt{2}$ is irrational. This should be thought of informally as saying "$\mathbb{Q}$ has holes" where certain least upper bounds *should* exist but are missing.

At this point, motivated by the above example, the reader should consider other subsets of $\mathbb{Q}$ which have upper bounds but no least upper bound and simultaneously convince themselves that in their intuitive model for $\mathbb{R}$ these least upper bounds do exist. This will become the defining property of $\mathbb{R}$.

Theorem 2.2.2. *There exists a (unique) ordered field, denoted $\mathbb{R}$, which contains $\mathbb{Q}$ as a sub-ordered field and has the least-upper-bound property.*

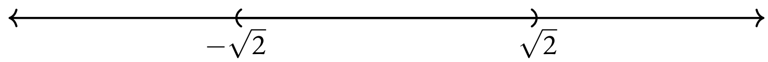
Corollary 2.2.3. *For every non-empty, lower bounded $S \subseteq \mathbb{R}$, $\inf S$ exists.*

Proof. This is Theorem 2.1.17. □

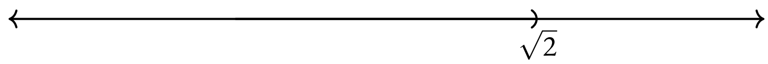
We now begin a sketch of the proof of Theorem 2.2.2. We saw in Example 2.2.1 that certain real numbers can be "encoded" by subsets of $\mathbb{Q}$ which ought to have a given real number as a supremum. This will be our approach to constructing the real number line—we will encode real numbers as subsets of $\mathbb{Q}$. However, we must take slight care in doing this. If we are to encode the real numbers by subsets of $\mathbb{Q}$ under the correspondence

$$\begin{aligned} \{\text{subsets of } \mathbb{Q}\} &\longrightarrow \mathbb{R} \\ S &\longmapsto \sup S \end{aligned}$$

we must constrain which subsets we consider. Indeed, in Example 2.2.1 the set S we considered looks like



on our intuitive real number line. But if we are encoding real numbers as supremums of subsets, it doesn't matter if we add smaller elements. Thus to get a unique subset encoding $\sqrt{2}$ an easy option would be to take the set which contain *all* things smaller than $\sqrt{2}$, i.e.



This motivates the following definition.

Definition 2.2.4. A *Dedekind cut* is a subset $S \subseteq \mathbb{Q}$ such that

(i) if $a \in S$ and $b < a$ then $b \in S$

(ii) S has no greatest element

This definition in hand we are ready to sketch a proof of Theorem 2.2.2.

(Sketch) Proof of Theorem 2.2.2. We also sketch the proof of existence. For this, we define $\mathbb{R}$ as a set by

$$\mathbb{R} = \{S \subseteq \mathbb{Q} : S \text{ is a Dedekind cut}\}.$$

We now need to give $\mathbb{R}$ a field structure and an ordering.

Morally, if a Dedekind cut S is supposed to represent the real number $\sup S$, then we may give the following natural ordering: We say $S_1 < S_2$ if S_2 contains an upper bound for S_1. We leave it to the reader to check that this gives an ordering on $\mathbb{R}$.

To define the field structure we have

$$S_1 + S_2 = \{s_1 + s_2 : s_1 \in S_1,\ s_2 \in S_2\}.$$

One needs to check this is still a Dedekind cut and gives an addition which satisfies all the necessary axioms. With respect to this addition we have

$$0 = \{q \in \mathbb{Q} : q < 0\}$$

and to define multiplication one needs to break into cases based on sign. This is left to the reader.

Finally, we confirm that $\mathbb{Q}$ is a subfield of $\mathbb{R}$. To see this, we encode the rational numbers as follows:

$$\begin{aligned} \mathbb{Q} &\longrightarrow \mathbb{R} \\ q &\longmapsto \{a \in \mathbb{Q} : a < q\}. \end{aligned}$$

□

We now take account of some properties of $\mathbb{R}$. One could prove these using the explicit construction of $\mathbb{R}$, however since Theorem 2.2.2 classifies $\mathbb{R}$ abstractly, we opt to prove these results using the least-upper-bound property.

Theorem 2.2.5. *The following statements are true:*

(i) (Archimedean Property) Given $x, y \in \mathbb{R}$ with $x > 0$, there exists some $n \in \mathbb{N}$ such that $nx > y$.

(ii) (Density of $\mathbb{Q}$) Given any two $x, y \in \mathbb{R}$ with $x < y$, there exists some $p \in \mathbb{Q}$ with $x < p < y$.

Proof. For (i), let

$$S = \{nx : n \in \mathbb{N}\}.$$

Suppose that for all $n \in \mathbb{N}$ we were to have $nx \leq y$. Then y is an upper bound for S. Since S is clearly non-empty, we can take $\alpha = \sup S$. Then, as $x > 0$ we have that $\alpha - x < \alpha$ and thus $\alpha - x$ cannot be an upper bound for S. Thus let $m \in \mathbb{N}$ be such that $\alpha - x < mx$. But then this gives $\alpha < (m+1)x$ contradicting the fact that α is an upper bound for S.

For (ii), since $x < y$ we have that $y - x > 0$, so (i) gives a positive integer n such that

$$n(y - x) > 1. \tag{2.2.1}$$

Now, apply (i) twice to find positive integers m_1, m_2 such that $m_1 > nx$ and $m_2 > -nx$. Then

$$-m_2 < nx < m_1.$$

But this then tells us that there exists an integer m (with $-m_2 \leq m \leq m_1$) such that

$$m - 1 \leq nx < m. \tag{2.2.2}$$

Combining (2.2.1) and (2.2.2) we get

$$nx < m \leq 1 + nx < ny$$

or in particular $nx < m < ny$. Dividing by n we get

$$x < \frac{m}{n} < y$$

as required. □

Theorem 2.2.6. *For every real $x > 0$ and every integer $n > 0$ there is exactly one $y > 0$ such that $y^n = x$.*

Proof. Uniqueness follows from the fact that if $0 < y_1 < y_2$, then $y_1^n < y_2^n$. Thus we are left with proving existence.

Let

$$S = \{t \in \mathbb{R} : t > 0,\ t^n < x\}.$$

If we let $t = x/(1+x)$ then $0 < t < 1$ so $t^n \le t < x$ and thus $t \in S$. Hence S is not empty. Additionally, if $t > 1+x$, then $t^n \ge t > x$ so $t \notin S$. In particular, if $t^n < x$ then $t \le 1+x$ so $1+x$ is an upper bound for S.

Thus by Theorem 2.2.2 we may take $y = \sup S$. To see that $y^n = x$, we will show that either one of the assumptions $y^n < x$ and $y^n > x$ leads to a contradiction.

First, note that we have an algebraic fact that

$$b^n - a^n = (b-a)(b^{n-1} + b^{n-2}a + \cdots + a^{n-1}).$$

If $0 < a < b$ then this yields

$$b^n - a^n < (b-a)nb^{n-1}. \tag{2.2.3}$$

Assume that $y^n < x$. Our goal is to show we can add a small $h > 0$ to y to get some $y+h$ with $(y+h)^n < x$ so that y is not an upper bound for S. For this, choose h such that $0 < h < 1$ and

$$h < \frac{x - y^n}{n(y+1)^{n-1}}. \tag{2.2.4}$$

Then applying (2.2.3) with $a = y$ and $b = y+h$ gives

$$\begin{aligned}(y+h)^n - y^n &< nh(y+h)^{n-1} \\ &< nh(y+1)^{n-1} \qquad (\text{as } h < 1) \\ &< x - y^n. \qquad (\text{by } (2.2.4))\end{aligned}$$

Thus $(y+h)^n < x$ and so $y+h \in S$ but $y < y+h$ contradicting the fact that y is an upper bound for S.

Now assume that $y^n > x$. We want to find some $h > 0$ such that $y - h$ is still an upper bound for

S, contradicting the fact that y is a least upper bound for S. Put

$$h = \frac{y^n - x}{ny^{n-1}}.$$

Then $0 < h < y$. If $t \geq y - h$ we get from (2.2.3) that

$$\begin{aligned} y^n - t^n &\leq y^n - (y-h)^n \\ &< hny^{n-1} \\ &= y^n - x. \end{aligned}$$

Thus $x < t^n$ so $t \notin S$. In particular, we get that $y - h \geq t$ for all $t \in S$ so $y - h$ is still an upper bound for S. □

Proposition 2.2.7. *If $x > 0$, and $p/q = m/n$ with p, q, m, n non-negative integers with $q \neq 0$, $n \neq 0$, then*

$$(x^m)^{1/n} = (x^p)^{1/q}.$$

Proof. Exercise to the reader. Use the uniqueness of roots. □

In particular, given any non-negative $p/q \in \mathbb{Q}$ we can define

$$x^{p/q} := (x^p)^{1/q}$$

and this is independent of the choice of how to represent p/q as a fraction.

2.3 Cardinality

Definition 2.3.1. Given two sets A, B, a *function* f from A to B, denoted by $f : A \to B$ is an assignment $f(a) \in B$ of an element of B to every element of A.

Example 2.3.2. For every set A we have the identity function $\mathrm{id}_A : A \to A$ which is given by $\mathrm{id}_A(a) = a$.

Example 2.3.3. Given two functions $f : A \to B$ and $g : B \to C$ we can form the composite $g \circ f$ given by the assignment $(g \circ f)(a) = g(f(a))$.

Definition 2.3.4. A function $f : A \to B$ is said to be

(i) *injective* if $f(a) = f(b)$ implies $a = b$

(ii) *surjective* if for every $b \in B$ there exists some $a \in A$ with $f(a) = b$

(iii) *bijective* if there exists a function $g : B \to A$ with $f \circ g = \mathrm{id}_B$ and $g \circ f = \mathrm{id}_A$. In this case g is said to be an *inverse* of f.

Proposition 2.3.5. *A function $f : A \to B$ is bijective if and only if it is injective and surjective.*

Proof. Exercise for the reader. □

With a formal notion of functions, we can now introduce the notion of cardinality, which is the mathematical formalism for the slogan that "some infinities are larger than others." In particular, we will be interested in sets whose elements we can count.

Definition 2.3.6. A set S is called *countable* if either S is finite, or there exists a bijection $f : \mathbb{N} \to S$ where $\mathbb{N} = \{1, 2, \dots\}$ is the natural numbers. Otherwise, S is said to be *uncountable*.

Note that being countable is the same as saying we can "list" the elements of S. Indeed, if $f : \mathbb{N} \to S$ is a bijection, then if we set $x_n = f(n)$ then $x_1, x_2, x_3, \dots$ lists or *enumerates* the set S.

Example 2.3.7. $\mathbb{N}$ is countable via the identity function $\mathrm{id}_{\mathbb{N}}$. Moreover, $\mathbb{Z}$ is also countable as we can enumerate it via

$$0, 1, -1, 2, -2, 3, -3, \dots.$$

Example 2.3.8. If A and B are countable, then $A \times B$ is countable. Indeed, enumerate A as $x_1, x_2, \dots$ and enumerate B as $y_1, y_2, \dots$. Then the elements of $A \times B$ fit into an array

$$\begin{matrix} (x_1, y_1) & (x_1, y_2) & (x_1, y_3) & \cdots \\ (x_2, y_1) & (x_2, y_2) & (x_2, y_3) & \cdots \\ (x_3, y_1) & (x_3, y_2) & (x_3, y_3) & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{matrix}$$

which we can enumerate by reading along the diagonals, i.e. as

$$(x_1, y_1), (x_2, y_1), (x_1, y_2), (x_3, y_1), (x_2, y_2), (x_1, y_3), \dots.$$

In particular, we have that $\mathbb{Z} \times \mathbb{Z}$ is countable.

Proposition 2.3.9. *If A is countable, then any $B \subseteq A$ is countable.*

Proof. If A is finite, we are done as any subset of a finite set is finite. Thus we may assume A is infinite and hence there exists a bijection $f : \mathbb{N} \to A$. Under this bijection, $B \subseteq A$ corresponds to some subset $C \subseteq \mathbb{N}$. Hence it suffices to show every subset of $\mathbb{N}$ is countable.

Thus suppose $C \subseteq \mathbb{N}$. If C is finite we are done. Otherwise we inductively construct a sequence $(x_n)_n$ which enumerates C. For this we define

$$x_1 = \min C$$

which exists as every non-empty subset of $\mathbb{N}$ has a minimal element. Now suppose $x_1, \ldots, x_n$ have been constructed. Then define

$$x_{n+1} = \min C \setminus \{x_1, \ldots, x_n\}.$$

This is well-defined as C is assumed to be infinite so $C \setminus \{x_1, \ldots, x_n\}$ is never empty. We now claim that every element of C is given by some x_n. Indeed, suppose that $m \in C$. Then m must belong to $\{x_1, \ldots, x_m\}$. Thus C is countable with an explicit bijection given by

$$\begin{aligned} \mathbb{N} &\longrightarrow C \\ n &\longmapsto x_n. \end{aligned}$$

Indeed, we just argued this map was surjective and if $m > n$ then as $x_m \in C \setminus \{x_1, \ldots, x_{m-1}\}$ we have $x_m \neq x_n$, so the map is injective. □

Corollary 2.3.10. $\mathbb{Q}$ *is countable.*

Proof. We have that $\mathbb{Q} \subseteq \mathbb{Z} \times \mathbb{Z}$ by sending a fraction p/q in reduced form to (p, q). Now, $\mathbb{Z} \times \mathbb{Z}$ is countable so we are done by Proposition 2.3.9. □

Our main goal of this section is to show that $\mathbb{R}$ is fundamentally larger than $\mathbb{Q}$ or $\mathbb{N}$. We will do this by showing that $\mathbb{R}$ is uncountable.

Theorem 2.3.11. $\mathbb{R}$ *is uncountable.*

Proof. We will prove this by an argument which is referred to as *Cantor's diagonalization argument*. Clearly $\mathbb{R}$ is infinite as it contains $\mathbb{Q}$. Thus suppose we were to have an enumeration $x_1, x_2, x_3, \ldots$ of $\mathbb{R}$. We will make use of decimal expansions to construct a real number not on this list.

Writing out the x_i as their decimal expansions we get a list

$$\begin{gathered} x_{10} \,.\, x_{11}x_{12}x_{13}x_{14}\ldots \\ x_{20} \,.\, x_{21}x_{22}x_{23}x_{24}\ldots \\ x_{30} \,.\, x_{31}x_{32}x_{33}x_{34}\ldots \\ \vdots \end{gathered}$$

where we are writing $x_i = x_{i0}.x_{i1}x_{i2}x_{i3}x_{i4}\ldots$ where x_{i0} is the integer part of x_i and $x_{ij}, j \geq 1$ is the j-th decimal digit of x_i. We now construct a real number x which does not belong to this list. To do this, let the integer part of x_i be 0, and let the j-th decimal digit be given as follows: If $x_{jj} = 1$ then the j-th decimal digit is 2, otherwise it is 1.

Then, by construction, $x \neq x_j$ for any j as x and x_j disagree at the j-th decimal digit. (Note that decimal expansions are not unique so there is something to check here—in theory we could have $x = 1.0000\ldots$ and $x_1 = 0.99999\ldots$ and thus $x = x_1$ even though as written the decimal expansions aren't equal, but we've ruled this possibility out by the rules for our construction of x). Thus x does not belong to $\{x_1, x_2, x_3, \ldots\}$ and hence $(x_n)_n$ was not an enumeration of $\mathbb{R}$. □

Question 2.3.12. Does there exist an uncountable set S with cardinality smaller than $\mathbb{R}$?

Remark 2.3.13. The above question has been proven to be unanswerable. The assumption that no such S exists is referred to as *the continuum hypothesis*, and it has been shown that this assumption is independent of the usual axioms of math. That is, it can be axiomatically assumed true or false without contradicting the other axioms in standard math.

Exercises

Exercise 2.1. Prove that additive and multiplicative identities and inverses are unique.

Exercise 2.2. Show that for a positive integer n, $\sqrt{n}$ is rational if and only if n is a square.

Exercise 2.3. Using the definition of $\mathbb{R}$ in terms of Dedekind cuts, give a definition for a function

$$f : \{a \in \mathbb{R} : a \geq 0\} \to \mathbb{R}$$

such that for every $a \in \mathbb{R}, a \geq 0$ we have $f(a)^2 = a$. That is, define rigorously a square root function and prove it is well-defined with the desired properties.

Exercise 2.4. Prove that for any two real numbers a, b with $a < b$ there exists a rational number q with $a < q < b$.

Exercise 2.5. Prove that the set of all subsets of $\mathbb{N}$, denoted by $\mathcal{P}(\mathbb{N})$ and referred to as the *power set* of $\mathbb{N}$, is uncountable.

Exercise 2.6. Show that there is never a surjection $f : X \to \mathcal{P}(X)$ where $\mathcal{P}(X)$ is the set of all subsets of X. This is a generalization of the previous exercise.

3 Metric spaces and their topology

In this section we attempt to abstract the properties of Euclidean space $\mathbb{R}^n$ which make in useful for doing math. $\mathbb{R}^n$ is a vector space, meaning we can add and multiply by scalars which in this case are real numbers. On top of this, it is given an *inner product*

$$\langle v, w\rangle = \sum_{i=1}^{n} v_i w_i,$$

i.e. a bracket which takes in two vectors and spits out a number and has certain geometric properties. Here v_i denotes the i-th coordinate of v in the standard basis of $\mathbb{R}^n$.

From this inner product, we get a *norm* which helps us measure the length of vectors.

Definition 3.0.1. The canonical norm on $\mathbb{R}^n$, denoted by $\|\cdot\|$ is given by

$$\|v\| = \sqrt{\langle v, v\rangle} = \left(\sum_i v_i^2\right)^{1/2}.$$

With the ability to measure the length of a vector, we are then able to measure the relative distance between two vectors.

Definition 3.0.2. Given two $v, w \in \mathbb{R}^n$, the *distance between v and w* is given by $\|v - w\|$.

Morally, one can think of $\|v-w\|$ as representing the length of the shortest path between v and w. In this framing, we expect certain things to hold: If we travel from v to w, that distance should always be less than the total distance had we first went from v to z then z to w, i.e. we expect

$$\|v - w\| \leq \|v - z\| + \|z - w\|.$$

Proposition 3.0.3 (The Triangle Inequality)**.** *For any $x, y, z \in \mathbb{R}^n$ we have*

$$\|x - y\| \leq \|x - z\| + \|z - x\|.$$

To prove this we first need a lemma.

Lemma 3.0.4 (Cauchy-Schwarz)**.** *For any two $v, w \in \mathbb{R}^n$ we have that $|\langle v, w\rangle| \leq \|v\| \cdot \|w\|$.*

Proof. Consider the polynomial in x given by

$$\begin{aligned}(v_1x+w_1)^2+\cdots+(v_nx+w_n)^2 &= \left(\sum_i v_i^2\right)x^2+2\left(\sum_i v_iw_i\right)x+\sum_i w_i^2\\ &= \|v\|^2x^2+2\langle v,w\rangle x+\|w\|^2.\end{aligned}$$

However, looking at the left hand side, we see that the polynomial is always non-negative, hence cannot have two distinct real roots. Thus the discriminant must be non-positive. Computing the discriminant with the right hand side we see that

$$4\langle v,w\rangle^2-4\|v\|^2\cdot\|w\|^2\leq 0$$

from which the result follows. □

Proof of Proposition 3.0.3. It suffices to prove that

$$\|u+v\|\leq\|u\|+\|v\|$$

as then taking $u=x-z$ and $v=z-y$ gives the result. For this, notice that

$$\begin{aligned}\|u+v\|^2 &= \langle u+v,u+v\rangle\\ &= \|u\|^2+2\langle u,v\rangle+\|v\|^2.\end{aligned}$$

By Cauchy-Schwarz we have that

$$\begin{aligned}\|u\|^2+2\langle u,v\rangle+\|v\|^2 &\leq \|u\|^2+2\|u\|\cdot\|v\|+\|v\|^2\\ &= (\|u\|+\|v\|)^2.\end{aligned}$$

Thus $\|u+v\|^2\leq(\|u\|+\|v\|)^2$ and the result follows. □

3.1 Definition of a metric space

Metric spaces

Definition 3.1.1. A *metric space* is a pair (X,d) where X is a set and $d:X\times X\to\mathbb{R}$ is a function satisfying the following properties for all $x,y,z\in X$:

(i) $d(x,y)=d(y,x)$

(ii) $d(x,y)\geq 0$ with equality if and only if $x=y$

(iii) (Triangle Inequality) $d(x, y) \leq d(x, z) + d(z, y)$

d is referred to as the *metric* on X.

Remark 3.1.2. When the metric d is clear from context we will often simply write X instead of (X, d).

Example 3.1.3. $\mathbb{R}^n$ with distance function $d(x, y) = \|x - y\|$ is a metric space. The triangle inequality is Proposition 3.0.3, while properties (i) and (ii) of Definition 3.1.1 follow from the analyzing the algebraic expression for $\|\cdot\|$.

Example 3.1.4 (Discrete Metric)**.** Every set X can be equipped with the *discrete metric* given by

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

Subspaces

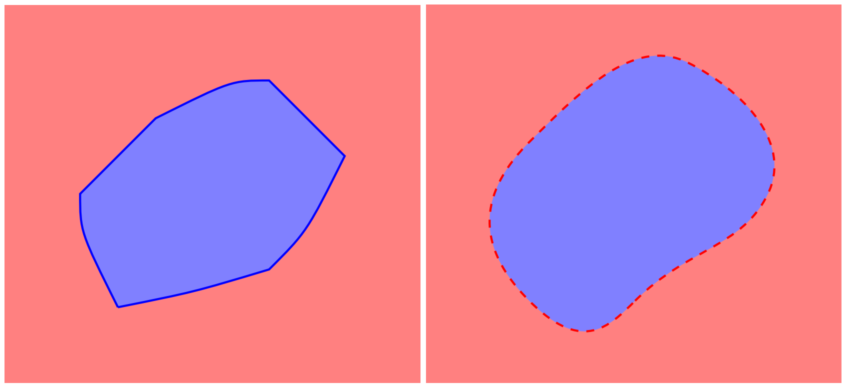
We often will want to restrict from one large ambient space to a smaller one. Given a metric space (X, d) and a subset $Y \subseteq X$, the restriction the metric $d : X \times X \to \mathbb{R}$ to Y gives a metric on Y. We give this a name.

Definition 3.1.5. Let (X, d) be a metric space and $Y \subseteq X$ a subset. The *subspace metric* or *subspace topology* on Y is $d|_{Y \times Y} : Y \times Y \to \mathbb{R}$.

Remark 3.1.6. Unless otherwise stated, all subsets of a metric space will be assumed to be equipped with the subspace metric.

3.2 Open and closed sets

We now wish to study two types of subsets that are important to the study of metric spaces—open and closed sets. Given the ability to measure distance on a set X and a point $x \in X$, it makes sense to speak of points that are "close to x." Often, we will want to talk about properties or features of a metric space that need only hold "close" to a certain point. Consider the following two pictures:



(A) (B)

Suppose that these two pictures represent subsets of a metric space X and the points in blue are where some property P holds and the points in red are where P does not hold. In picture (B), if one takes any point $x \in X$ where P holds and zooms in, we find that all nearby points are also blue, i.e. P holds. However, in picture (A) this is not the case—if one chooses a point on the "boundary" then no matter how much we zoom in there will be points that are both blue and red.

Informally, open sets will be sets "having no boundary points" and closed sets will be sets having "all their boundary points." This will be of great importance when discussing properties which should hold "locally," i.e. after zooming in.

Open sets

Definition 3.2.1. Let (X, d) be a metric space. For $x \in X$ and $r \geq 0$ the set

$$B_r(x) = \{y \in X : d(x, y) < r\}$$

is referred to as the *open ball of radius r centered at x*.

Definition 3.2.2. A subset $U \subseteq X$ is said to be *open* if for every $x \in U$ there exists some $r > 0$ such that $B_r(x) \subseteq U$.

Example 3.2.3. Being an open subset depends on the ambient space. For example, we have that $(0, 1)$ is open in $\mathbb{R}$ but not in $\mathbb{R}^2$.

Example 3.2.4. Every set in a discrete metric space is open. Indeed, let X have the discrete metric and let $S \subseteq X$. Then given $x \in S$ we have that

$$B_{1/2}(x) = \{x\} \subseteq S.$$

Proposition 3.2.5. *All open balls $B_r(x)$ are open.*

Proof. Take $y \in B_r(x)$, so by definition $d(x, y) < r$. Let

$$R = r - d(x, y) > 0.$$

We wish to show that $B_R(y) \subseteq B_r(x)$. Indeed, consider $z \in B_R(y)$. Then

$$\begin{aligned} d(x, z) &\leq d(x, y) + d(y, z) \\ &< d(x, y) + R \\ &= r \end{aligned}$$

so $z \in B_r(x)$ as required. □

Theorem 3.2.6. *The following statements are true:*

(i) A finite intersection of open sets is open.

(ii) An arbitrary union of open subsets is open.

Proof. For (i), let $U_1, \dots, U_n$ be open subsets and let

$$x \in U_1 \cap \cdots \cap U_n.$$

Then, since each U_i is open, we may find for every i some $r_i > 0$ such that $B_{r_i}(x) \subseteq U_i$. Then if we set $r = \min\{r_1, \dots, r_n\}$ we have that $r > 0$ and $B_r(x) \subseteq U_1 \cap \cdots \cap U_n$ as required.

For (ii), suppose that $\{U_\alpha : \alpha \in A\}$ is an arbitrary collection of open sets. Let

$$x \in \bigcup_{\alpha \in A} U_\alpha.$$

Then there exists some $\beta \in A$ such that $x \in U_\beta$. Since U_β is open, there exists some $r > 0$ such that

$$x \in B_r(x) \subseteq U_\beta \subseteq \bigcup_{\alpha \in A} U_\alpha$$

as required. □

Example 3.2.7. The finiteness assumption in Theorem 3.2.6(i) is necessary. Indeed, consider

$$\mathbb{R}_{\geq 0} = \{x \in \mathbb{R} : x \geq 0\} \subseteq \mathbb{R}$$

with the subspace topology. Then $[0, a)$ is open in $\mathbb{R}_{\geq 0}$ for all $a > 0$ but we have that

$$\bigcap_{n \in \mathbb{N}} [0, 1/n) = \{0\}$$

which is not open in $\mathbb{R}_{\geq 0}$.

Definition 3.2.8. Let (X, d) be a metric space and $x \in X$. A *neighborhood* of x refers to an open set U with $x \in U$.

By Proposition 3.2.5, saying that a set S contains a neighborhood of a point x is equivalent to saying that S contains a ball of positive radius centered at x.

Proposition 3.2.9. *Let $p, q \in X$ be two distinct points in X, i.e. $p \neq q$. Then there exist disjoint neighborhoods of p and q. That is, there exists a neighborhood V of p and W of q such that $V \cap W = \emptyset$.*

Proof. Set $r = d(p, q)$. Then $r > 0$ as $p \neq q$ and take $V = B_{r/2}(p)$ and $W = B_{r/2}(q)$. The fact that $V \cap W = \emptyset$ then follows from the triangle inequality. □

Before closing this section we remark that Example 3.2.3 shows that the property of being open is not intrinsic to the subspace—it depends on the parent or ambient space. This next theorem completely classifies this dependence.

Theorem 3.2.10. *Let $U \subseteq Y \subseteq X$. Then U is open as subset of Y if and only if it may be written is the form $U = Y \cap V$ for V an open subset of X.*

Proof. For the purpose of distinguishing between spaces, given $y \in Y$ we will write $B_r^Y(y)$ for the open ball of radius r centered at y taken in Y and $B_r^X(y)$ for the open ball of radius r centered at y taken in X. Note that because $d_Y = d_X|_{Y \times Y}$ we have that

$$B_r^Y(y) = Y \cap B_r^X(y). \tag{3.2.1}$$

Thus first suppose that $U = Y \cap V$ for some V open in X. Then take $y \in U$. Since V is open in X there exists some $r > 0$ such that $B_r^X(y) \subseteq V$. But then by (3.2.1) we have that

$$B_r^Y(y) = Y \cap B_r^X(y) \subseteq Y \cap V = U$$

so U is open as required.

Conversely, suppose that U is open in Y. Then for every $y \in U$ we may find some $r > 0$ such that $B_r^Y(y) \subseteq U$. Set $V_y = B_r^X(y)$. Then by (3.2.1) we have that

$$y \in V_y \cap Y = B_r^Y(y) \subseteq U. \tag{3.2.2}$$

Now set $V = \bigcup_{y \in U} V_y$. By Theorem 3.2.6 V is open in X. Moreover, by (3.2.2) we have that $V \cap Y \subseteq U$ and for every $y \in U$ we have $y \in V_y \cap Y \subseteq V \cap Y$. Thus $U \subseteq V \cap Y$ and $V \cap Y \subseteq U$ so $U = V \cap Y$ as required. □

Sequences

Equipped with a way of measuring distance, we can now speak of when points "approach" other points or when sequences converge.

Definition 3.2.11. Let (X, d) be a metric space. A *sequence* in X is a function $f : \mathbb{N} \to X$.

Remark 3.2.12. We will often denote a sequence by a subscript indexed sequence such as $(x_n)_n$ or $x_1, x_2, x_3, \ldots$. When doing this, we implicitly mean the sequence $f : \mathbb{N} \to X$ given by $f(n) = x_n$.

Definition 3.2.13. Let (X, d) be a metric space and $(x_n)_n$ a sequence in X. Then, given $x \in X$, we say that $(x_n)_n$ *converges to* x, denoted $x_n \to x$, if for all $\varepsilon > 0$ there exists some $N \in \mathbb{N}$ such that for all $n \geq N$ we have $d(x_n, x) < \varepsilon$.

Morally, convergence says that no matter how close we want to be to x (with distances measured by our metric d), if we go far enough down our sequence all of our terms are eventually at least that close to x.

Example 3.2.14. In $(\mathbb{R}, |\cdot|)$ we have that $1/n \to 0$. Indeed, fix $\varepsilon > 0$. Then we wish to find $N \in \mathbb{N}$ such that if $n \geq N$ we have

$$\frac{1}{n} = \left|\frac{1}{n} - 0\right| < \varepsilon.$$

This is true if and only if $n > 1/\varepsilon$ so we may take $N = \lceil 1/\varepsilon \rceil + 1$.

Example 3.2.15. In the discrete metric, the only convergent sequences are those which are eventually constant.

Closed sets

Definition 3.2.16. Let (X, d) be a metric space and $S \subseteq X$. Define $\lim S$ to be the set

$$\lim S = \{x \in X : \exists \text{ sequence } (x_n)_n \text{ in } S \text{ with } x_n \to x\}.$$

We say that S is *closed* if $S = \lim S$.

Example 3.2.17. For every metric space X, both $\emptyset$ and X are closed.

Proposition 3.2.18. $\lim S \supseteq S$.

Proof. For every $s \in S$ we have that the constant sequence $(s)_n$ converges to s. □

Example 3.2.19. $S = [0, 1] \subseteq \mathbb{R}$ is closed. By Proposition 3.2.18, we only need to show that $\lim S \subseteq S$. For this, suppose that $(x_n)_n$ is a sequence in $[0, 1]$ which converges to $x \in \mathbb{R}$. Suppose for contradiction that $x > 1$. Then let $\varepsilon = x - 1 > 0$. For n sufficiently large, we must have that $|x - x_n| < \varepsilon$. But then

$$x - x_n < \varepsilon = x - 1$$

which implies $x_n > 1$, contradicting that $x_n \in [0, 1]$. Similarly, if $x < 0$ we get a contradiction, so $x \in [0, 1]$. Thus $\lim S \subseteq S$ as required.

Generalizing this example we get the following proposition.

Proposition 3.2.20. *For $r \geq 0$ let*

$$D_r(x) = \{y \in X : d(x, y) \leq r\}$$

be the closed ball *of radius r centered at x. Then $D_r(x)$ is closed.*

Proof. Let $(y_n)_n$ be a sequence in $D_r(x)$ with $y_n \to y$. We want to show that $y \in D_r(x)$. Suppose not, so that $d(x, y) > r$. Then set

$$\varepsilon = d(x, y) - r > 0.$$

Since $y_n \to y$, for n sufficiently large we must have that $d(y, y_n) < \varepsilon$. But then

$$\begin{aligned} d(x, y_n) &\geq d(x, y) - d(y, y_n) \\ &> d(x, y) - \varepsilon \\ &= r \end{aligned}$$

contradicting that $y \in D_r(x)$. □

Proposition 3.2.21. $\lim S$ *is closed, i.e.* $\lim \lim S = \lim S$.

Proof. Suppose that $(x_n)_n$ is a sequence with $x_n \in \lim S$ which converges to some point x. We need to show that $x \in \lim S$.

By definition of $\lim S$, since each $x_n \in \lim S$ we may find for every n some $y_n \in S$ with $d(y_n, x_n) < 1/n$. We will show that $y_n \to x$ at which point we are done by definition of $\lim S$. Fix $\varepsilon > 0$. Since $x_n \to x$, we may find N such that for all $n \geq N$, $d(x_n, x) < \varepsilon/2$. Now, increasing N if necessary, we may assume that $1/N < \varepsilon/2$. Then for $n \geq N$ we have that

$$\begin{aligned} d(y_n, x) &\leq d(y_n, x_n) + d(x_n, x) \\ &< 1/n + \varepsilon/2 \\ &\leq 1/N + \varepsilon/2 \\ &< \varepsilon/2 + \varepsilon/2 = \varepsilon. \end{aligned}$$

This shows that $y_n \to x$ as required. □

Theorem 3.2.22. *A subset $E \subseteq X$ is closed if and only if $E^c = X \setminus E$ is open.*

Proof. First assume that E is closed. Suppose that E^c were not open, so we could find some $x \in E^c$ such that every ball centered at x does not lie in E^c, i.e. intersects E. Thus, for every n choose some $x_n \in E \cap B_{1/n}(x)$. Then by construction $x_n \to x$ so $x \in \lim E$. But E is closed so $E = \lim E$ and thus $x \in E$, a contradiction.

On the other hand, suppose that E^c is open. Then let $(x_n)_n$ be a sequence in E with $x_n \to x$. If we were to have $x \in E^c$ then as E^c is open there would exist some $r > 0$ such that $B_r(x) \subseteq E^c$ violating the fact that $x_n \to x$. Thus we must have $x \in E$ so $\lim E \subseteq E$. By Proposition 3.2.18 it follows that $E = \lim E$ so E is closed. □

Remark 3.2.23. It turns out that open sets are the more intrinsic collection of sets, and Theorem 3.2.22 is often taken as the *definition* of what it means to be closed.

Example 3.2.24. Theorem 3.2.22 gives another way of showing that $[0, 1] \subseteq \mathbb{R}$ is closed. Indeed,

$$[0, 1]^c = (-\infty, 0) \cup (1, \infty)$$

is the union of two opens, hence open. Thus $[0, 1]$ is closed.

Example 3.2.25. Every subset of a discrete metric space is closed. Indeed, by Example 3.2.4 every subset is open, and hence every subset has open complement.

Corollary 3.2.26. *Arbitrary intersections of closed subsets are closed, and finite unions of closed subsets are closed.*

Proof. Apply Theorem 3.2.6 and Theorem 3.2.22. □

Corollary 3.2.27. *Let $K \subseteq Y \subseteq X$. Then K is closed as a subset of Y if and only if we can write $K = S \cap Y$ for S a closed subset of X.*

Proof. Apply Theorem 3.2.10 and Theorem 3.2.22. □

Definition 3.2.28. Let (X, d) be a metric space and $S \subseteq X$. The *closure of* S, denoted by $\overline{S}$, is the smallest closed set containing S.

Note that by Corollary 3.2.26 closures exist. Indeed, X itself is a closed set containing S, and then

$$\overline{S} = \bigcap_{\substack{E \supseteq S \\ E \text{ closed}}} E$$

gives the closure of S.

Proposition 3.2.29. *One has that* $\overline{S} = \lim S$.

Proof. By Propositions 3.2.21 and 3.2.18, we have that $\lim S \supseteq \overline{S}$. However, since $\overline{S} \supseteq S$ we must have by definition of lim that $\lim \overline{S} \supseteq \lim S$. But $\overline{S}$ is closed so $\overline{S} = \lim \overline{S}$ and thus $\overline{S} \supseteq \lim S$ giving reverse containment. □

Corollary 3.2.30. *A set S is closed if and only if $S = \overline{S}$.*

Proof. Since $\overline{S} = \lim S$ by Proposition 3.2.29 this is just the definition of being closed. □

Proposition 3.2.31. *Let $S \subseteq \mathbb{R}$ be upper bounded. Then* $\sup S \in \lim S$. *In particular,* $\sup S \in S$ *if S is closed.*

Proof. We construct a sequence $(x_n)_n$ in S such that $x_n \to \sup S$. For this, let $x = \sup S$. Then for all $n \in \mathbb{N}$ we cannot have that $x - 1/n$ is an upper bound for S, so there must exist $x_n \in S$ with $x - 1/n < x_n$. But then since x is an upper bound for S we have that

$$x - 1/n < x_n \leq x$$

so $|x - x_n| \leq 1/n$. Thus we must have that $x_n \to x$ as required. □

Definition 3.2.32. Let (X, d) be a metric space. We say that $S \subseteq X$ is *clopen* if is both open and closed.

Example 3.2.33. The only clopen subsets of $\mathbb{R}$ are $\emptyset$ and $\mathbb{R}$. Indeed, suppose that S were a clopen set with $S \neq \emptyset$ and $S \neq \mathbb{R}$. Then both S and S^c are non-empty so let $a \in S$ and $b \in \mathbb{R} \setminus S$. Suppose that $a < b$ (the proof is similar if $b < a$). Consider

$$c = \sup(S \cap [a, b]).$$

Since S and $[a, b]$ are closed, we have that $S \cap [a, b]$ is closed and thus $c \in S \cap [a, b]$ by Proposition 3.2.31. But $c \neq b$ as $b \notin S$ so $a \leq c < b$. Moreover, by construction of S it follows that $(c, b] \subseteq \mathbb{R} \setminus S$ contradicting that S is open as no neighborhood of c can then be contained in S.

Example 3.2.34. Let $X = [0, 1] \cup [2, 3] \subseteq \mathbb{R}$ with the subspace metric. Then $[0, 1]$ and $[2, 3]$ are both clopen in X.

We will later see when studying connected sets that the only spaces which have non-trivial clopen sets are "disconnected" in some suitable sense. This is illustrated by Examples 3.2.33 and 3.2.34.

3.3 Compact sets

Compactness

Definition 3.3.1. Let (X, d) be a metric space and $S \subseteq X$. An *open cover* of S is a collection $\{U_\alpha\}_{\alpha \in A}$ of open subsets of X such that

$$\bigcup_{\alpha \in A} U_\alpha \supseteq S.$$

A *subcover* of $\{U_\alpha\}_{\alpha \in A}$ refers to an open cover of the form $\{U_\alpha\}_{\alpha \in B}$ for some $B \subseteq A$.

Definition 3.3.2. Let (X, d) be a metric space and $C \subseteq X$. C is said to be *compact* if every open cover of C has a finite subcover, i.e. a subcover with finitely many elements.

Example 3.3.3. Any finite subset of a metric space is compact.

Theorem 3.3.4. *Closed intervals* $[a, b] \subseteq \mathbb{R}$ *are compact.*

Proof. Suppose we have an open cover $\{U_\alpha\}_\alpha$ of $[a, b]$. Note that for all $x \in [a, b]$, $[a, x]$ is still covered by $\{U_\alpha\}$. Let S be the collection of all $x \in [a, b]$ such that $[a, x]$ has a finite subcover. Then S is non-empty as $a \in S$ and S is upper-bounded by definition. Thus let

$$x_0 = \sup S \in [a, b].$$

Our goal is to show that $x_0 = b$. Note that $x_0 > a$ since taking any open containing a produces an interval of points with a finite subcover, namely covered by a single open.

Suppose that $x_0 \neq b$ so $a < x_0 < b$ and let i_0 be such that $x_0 \in U_{i_0}$. Then we may find $\varepsilon > 0$ such that

$$a \leq x_0 - \varepsilon < x_0 < x_0 + \varepsilon \leq b$$

and such that

$$(x_0 - \varepsilon, x_0 + \varepsilon) \subseteq U_{i_0} \tag{3.3.1}$$

since U_{i_0} is open. Then since $x_0 = \sup S$ we may find some $x_1 \in S$ with $x_0 - \varepsilon < x_1 \leq x_0$. Thus we have by definition of S that $[a, x_1]$ has a finite subcover, and adding U_{i_0} if necessary to this subcover we see by (3.3.1) that $[a, x_0 + \varepsilon]$ has a finite subcover. Thus $x_0 + \varepsilon \in S$ contradicting that x_0 is an upper bound for S.

We conclude that $x_0 = b$. Moreover, the same argument shows that $x_0 \in S$ (or more strongly that S^c is open, so S is closed). Thus we conclude that $b \in S$ so $[a, b]$ has a finite subcover as required. □

Compactness should be thought of as a way of saying that a space is "not too big." We will prove various theorems throughout this section that justify this intuition. Moreover, this sense of size is intrinsic to the subspace itself, not its ambient space as the following theorem says.

Theorem 3.3.5. *Let $C \subseteq Y \subseteq X$. Then C is compact as a subset of Y if and only if it is compact as a subset of X.*

Proof. Suppose C is compact as a subset of Y. Then let $\{U_\alpha : \alpha \in A\}$ be an open cover of C as a subset of X. Then by Theorem 3.2.10 we have that $\{U_\alpha \cap Y : \alpha \in A\}$ is an open cover of C in Y. Since C is compact as a subset of Y we may find a finite subset $B \subseteq A$ such that $\{U_\alpha \cap Y : \alpha \in B\}$ is a cover of C. But then $\{U_\alpha : \alpha \in B\}$ is a cover of C in X. Thus C is compact as a subset of X.

Conversely, suppose that C is compact as a subset of X. Let $\{U_\alpha : \alpha \in A\}$ be an open cover of C in Y. Then by Theorem 3.2.10 for each $\alpha \in A$ we may write $U_\alpha = Y \cap V_\alpha$ for V_α an open subset of X. We then have that $\{V_\alpha : \alpha \in A\}$ is an open cover of C in X. Thus it has a finite subcover indexed by $B \subseteq A$ and we get that $\{U_\alpha : \alpha \in B\}$ is a finite subcover of C in Y as required. □

In light of this theorem, we could say that a metric space X is compact if it is compact as a subset $X \subseteq X$ of itself. Then, by Theorem 3.3.5, given $C \subseteq X$ saying C is compact as a subset of X is equivalent to saying C is compact as a metric space when given the subspace metric.

Theorem 3.3.6. *Let $C \subseteq X$ be compact. Then C is closed.*

Proof. We will show that the complement of C is open from which the result will follow from Theorem 3.2.22.

Fix some $p \in X \setminus C$. We wish to find some neighborhood of p lying outside C. To do this, note that for every $q \in C$ by Proposition 3.2.9 we may find disjoint neighborhoods V_q and W_q of q and p respectively. Now, $\{V_q : q \in C\}$ form an open cover of C, so by compactness we may find a finite subcover $V_{q_1}, \dots, V_{q_n}$ such that

$$C \subseteq V_{q_1} \cup \cdots \cup V_{q_n}. \tag{3.3.2}$$

Then take $W = W_{q_1} \cap \cdots \cap W_{q_n}$. Since W is a finite intersection of neighborhoods of p, W itself is a neighborhood of p. Moreover, by construction, $W \cap V_{q_i} = \emptyset$ for all i so by (3.3.2) we have $W \cap C = \emptyset$. Thus $p \in W \subseteq X \setminus C$ and so $X \setminus C$ is open as required. □

Theorem 3.3.7. *Let X be compact. Then every closed subset $E \subseteq X$ is compact.*

Proof. Given an open cover $\{U_\alpha : \alpha \in A\}$ of E in X, since E is closed in X we find that $\{X \setminus E\} \cup \{U_\alpha : \alpha \in A\}$ is an open cover of X. Since X is compact this has a finite subcover, which, after throwing out $X \setminus E$ if necessary, gives a finite subcover of E. □

Sequential compactness

Definition 3.3.8. A metric space X is said to be *sequentially compact* if every sequence $(x_n)_n$ in X has a convergent subsequence. That is, we may find some sequence $n_1 < n_2 < n_3 < \cdots$ of increasing integers such that $(x_{n_k})_k$ converges.

We have two goals in this section and the following: To prove that compactness and sequential compactness are the same (in the setting we are in), and to classify compact subsets of Euclidean space $\mathbb{R}^n$.

We begin with an easy direction.

Lemma 3.3.9. *Let $(x_n)_n$ be a sequence in a metric space X with no convergent subsequence and let $x \in X$. Then we may find an open neighborhood of x which does not contain any of the x_n (except possibly x itself).*

Proof. WLOG we may assume that x is not one of the x_n as if it is then we may remove all occurrences of x in $(x_n)_n$ and apply the result.

Next, assuming that x does not belong to $\{x_n : n \in \mathbb{N}\}$, notice that it suffices to find an open neighborhood of x containing finitely many x_n's. Indeed, suppose we have some neighborhood U of x containing only finitely many of the x_n's, i.e.

$$U \cap \{x_n : n \in \mathbb{N}\} = \{x_{n_1}, \ldots, x_{n_\ell}\}.$$

Then as $x \neq x_{n_i}$ for any i, we may find for each i a neighborhood U_i of x with $x_{n_i} \notin U_i$ by Proposition 3.2.9. Then setting

$$V = U \cap U_1 \cap \cdots \cap U_n$$

we have that V is a neighborhood of x with

$$V \cap \{x_n : n \in \mathbb{N}\} = \emptyset$$

as required. Thus for contradiction assume that every open neighborhood of x contains infinitely many of the x_n's. From this we will construct a subsequence of $(x_n)_n$ converging to x giving a contradiction.

To do this, let n_1 be any index such that $x_{n_1} \in B_1(x)$. Now, assume that $n_1 < \cdots < n_{k-1}$ have been constructed. Since $B_{1/k}(x)$ contains infinitely many of the x_n's, we may find $n_k > n_{k-1}$ such that $x_{n_k} \in B_{1/k}(x)$. Doing this, we inductively get a sequence $(n_k)_k$ such that $d(x, x_{n_k}) < 1/k$ and hence $(x_{n_k})_k$ is a subsequence converging to x, a contradiction. □

Proposition 3.3.10. *Let (X, d) be a metric space. If X is compact then X is sequentially compact.*

Proof. Let $(x_n)_n$ be a sequence in X. Suppose $(x_n)_n$ did not have any convergent subsequences. We will use this to construct an open cover with no finite subcover.

For this, fix $x \in X$. By Lemma 3.3.9, we may find an open neighborhood U_x of x such that

$$U_x \cap \{x_n : n \in \mathbb{N}\} \subseteq \{x\} \tag{3.3.3}$$

Then $\{U_x : x \in X\}$ is an open cover of X with no finite subcover.

Indeed, suppose that $\{U_x : x \in B\}$ were a subcover. Then since

$$X = \bigcup_{x \in B} U_x$$

we in particular have that

$$\begin{aligned} \{x_n : n \in \mathbb{N}\} &= \{x_n : n \in \mathbb{N}\} \cap \left(\bigcup_{x \in B} U_x\right) \\ &= \bigcup_{x \in B} (\{x_n : n \in \mathbb{N}\} \cap U_x) \\ &\subseteq B \end{aligned} \tag{3.3.4}$$

where in the last step we use (3.3.3). But $(x_n)_n$ has no convergent subsequence, so $\{x_n : n \in \mathbb{N}\}$ must be an infinite set, and hence B must be infinite by (3.3.4). □

The other direction requires some work. We begin with some definitions and preliminaries.

Definition 3.3.11. Let (X, d) be a metric space and let $S \subseteq X$. We say that S is *dense* if $\overline{S} = X$.

Proposition 3.3.12. *A set S is dense in X if and only if every non-empty open set in X intersects S.*

Proof. For one direction, suppose that U is a non-empty set not intersecting S. Then U^c is closed and we have that

$$\overline{S} \subseteq U^c \neq X$$

so S is not dense.

For the other direction, suppose that every non-empty open $U \subseteq X$ intersects S. Then if $\overline{S} \neq X$ we have that $\overline{S}^c$ is a non-empty open not intersecting S, a contradiction. $\square$

Definition 3.3.13. We say that a metric space X is *separable* if there exists a countable dense subset of X.

Example 3.3.14. $\mathbb{Q}$ is dense in $\mathbb{R}$. In particular, $\mathbb{R}$ is separable. Indeed, this follows from Theorem 2.2.5 which tells us that every open interval in $\mathbb{R}$ contains a rational number.

Proposition 3.3.15. *Let X be a sequentially compact metric space. Then X is separable.*

Proof. Fix $\varepsilon > 0$. X can be covered by finitely many open balls of radius ε. Indeed, suppose not. We will construct inductively a sequence $(x_n)_n$ with no convergent subsequence. For this, let x_1 be any $x_1 \in X$. Now suppose that $x_1, \ldots, x_n$ have been constructed. Then since X cannot be covered by finitely many balls of radius ε, we can find

$$x_{n+1} \in X \setminus (B_\varepsilon(x_1) \cup \cdots \cup B_\varepsilon(x_n)).$$

This sequence then has the property that $d(x_i, x_j) \geq \varepsilon$ for all $i \neq j$, so in particular $(x_n)_n$ has no convergent subsequence.

We now construct our countable dense subset. For every $n \in \mathbb{N}$, let

$$S_n = \{x_1, \ldots, x_m\}$$

be a finite set of points in X such that

$$X = B_{1/n}(x_1) \cup \cdots \cup B_{1/n}(x_m)$$

which exists by the above work. Then

$$S = \bigcup_{n \in \mathbb{N}} S_n$$

is countable and we show that $\overline{S} = X$.

Indeed, let U be a non-empty open subset of X. Then in particular, U contains some open ball $B_r(y)$ for $r > 0$. But taking n with $1/n < r$ we may find some $z \in S_n \subseteq S$ such that $d(z, y) < 1/n < r$. Thus in particular, $z \in S \cap U$ so S is dense by Proposition 3.3.12. $\square$

Proposition 3.3.16. *Let X be a separable metric space. Then every open cover $\{U_\alpha\}_\alpha$ of X has a countable subcover.*

Proof. Let S be a countable dense subset of X. Note that there are only countably many open balls of rational radius and center belonging to S (since $\mathbb{Q} \times S$ is countable). Thus it suffices to find a collection $\mathcal{C}$ of open balls of rational radius centered at points of S which cover X and have the property that for every $V \in \mathcal{C}$ there exists U_α in our original cover with $V \subseteq U_\alpha$.

To show this, notice that for every $x \in X$ we may find some U_β in our cover with $x \in U_\beta$. Then we may find some $q \in \mathbb{Q}$ with

$$B_q(x) \subseteq U_\beta.$$

Since S is dense, we may find some $y \in S$ with $d(x, y) < q/2$. Then

$$x \in B_{q/2}(y) \subseteq B_q(x) \subseteq U_\beta.$$

Let $V_x = B_{q/2}(y)$. Then $\mathcal{C} = \{V_x : x \in X\}$ is our desired collection. □

Lemma 3.3.17. *Let X be a sequentially compact space and let $F_1 \supseteq F_2 \supseteq F_3 \supseteq \cdots$ be a nested sequence of non-empty closed subsets. Then $\bigcap_n F_n \neq \emptyset$.*

Proof. For each n pick $x_n \in F_n$. Then as X is sequentially compact, $(x_n)_n$ has a convergent subsequence, say $x_{n_k} \to y$. We wish to show that $y \in \bigcap_n F_n$.

For this, notice that because the F_i's are nested, for any n the tail of the subsequence $(x_{n_k})_k$ lies in F_n. Thus $y \in \lim F_n$ so $y \in F_n$ as each F_n is closed. Thus $y \in F_n$ for all n as required. □

Theorem 3.3.18. *Let X be a sequentially compact metric space. Then X is compact.*

Proof. By Propositions 3.3.15 and 3.3.16 we have that every open cover of X has a countable subcover. Thus it remains to show that every countable open cover of X has a finite subcover.

Let $\{U_1, U_2, U_3, \dots\}$ be a countable open cover which we may WLOG assume is infinite. Suppose that no finite subset covers X. Then set

$$F_n = U_1^c \cap \cdots \cap U_n^c.$$

Then F_n is closed and non-empty as $X \neq U_1 \cup \cdots \cup U_n$. Moreover, we have that $F_1 \supseteq F_2 \supseteq F_3 \supseteq \cdots$. Thus by Lemma 8.3.1 we have that

$$\bigcap_n F_n = \bigcap_n U_n^c \neq \emptyset.$$

But then the U_i do not cover, a contradiction. □

Corollary 3.3.19. *A metric space X is compact if and only if it is sequentially compact.*

3.4 Heine-Borel and Bolzano-Weierstrass

Having proved that sequential compactness is the same as compactness, we are able to reap immediate benefits which we take account of in this section.

Definition 3.4.1. Let (X, d_X) and (Y, d_Y) be metric spaces. We define the *product metric* on $X \times Y$ to be the metric given by

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

Example 3.4.2. Usual Euclidean space $\mathbb{R}^n$ is equal to the product metric space

$$\underbrace{\mathbb{R} \times \cdots \mathbb{R}}_{n \text{ times}}.$$

There are many different ways we could have equipped the product set $X \times Y$ with a metric. Some of the natural ways are:

$$d_{\text{Eucl}}((x_1, y_1), (x_2, y_2)) = \sqrt{d_X(x_1, x_2)^2 + d_Y(y_1, y_2)^2}$$
$$d_1((x_1, y_1), (x_2, y_2)) = d_X(x_1, x_2) + d_Y(y_1, y_2)$$
$$d_\infty((x_1, y_1), (x_2, y_2)) = \max\{d_X(x_1, x_2), d_Y(y_1, y_2)\}.$$

It turns out that these metrics are *equivalent* in a suitable sense (see homework) and thus checking whether a set $A \subseteq X \times Y$ is open does not depend on which of the above metrics we choose.

Proposition 3.4.3. *Let X, Y be metric spaces. Then a sequence $((x_n, y_n))_n$ in $X \times Y$ converges to (x, y) if and only if $x_n \to x$ and $y_n \to y$.*

Proof. Exercise to the reader. □

Theorem 3.4.4. *Let $X_1, \ldots, X_n$ be compact metric spaces. Then $X_1 \times \cdots \times X_n$ is compact.*

Proof. By induction, we may reduce to the case $n = 2$. Since compactness is the same as sequential compactness, it suffices to show that every sequence in $X_1 \times X_2$ has a convergent subsequence.

For this, let $((x_n, y_n))_n$ be any sequence in $X_1 \times X_2$. Then $(x_n)_n$ has a convergent subsequence, say $(x_{n_k})_k$, since X_1 is compact. But then $(y_{n_k})_k$ also has a convergent subsequence by compactness of X_2, say $(y_{n_{k_\ell}})_\ell$. Then by Proposition 3.4.3 we have that $((x_{n_{k_\ell}}, y_{n_{k_\ell}}))_\ell$ is a convergent subsequence of $((x_n, y_n))_n$ as required. □

This gives us some more examples of compact subsets.

Definition 3.4.5. A *k-cell* in $\mathbb{R}^k$ is a set I of the form

$$I = [a_1, b_1] \times \cdots \times [a_k, b_k]$$

for $a_1, b_1, \ldots, a_k, b_k \in \mathbb{R}^n$.

Theorem 3.4.6. *k-cells are compact.*

Proof. Theorem 3.3.4 tells us that the closed intervals $[a_i, b_i] \subseteq \mathbb{R}$ are compact. Then the result follows from Theorem 3.4.4. $\square$

This is a strong result from which many famous corollaries follow.

Definition 3.4.7. A metric space (X, d) is said to be *bounded* if there exists some $M > 0$ such that $d(x, y) \leq M$ for all $x, y \in X$.

Theorem 3.4.8. *Every compact metric space is bounded.*

Proof. This is a consequence of our proof of Proposition 3.3.15, which we repeat here. Let X be a compact metric space. Notice that

$$\{B_1(x) : x \in X\}$$

is an open cover of X. Thus, by compactness, we may find a finite subcover and write

$$X = B_1(x_1) \cup \cdots \cup B_1(x_n). \tag{3.4.1}$$

Then let

$$M = \max\{d(x_j, x_i) : 1 \leq i, j \leq n\}.$$

For every $x, y \in X$, by (3.4.1), we may find x_i and x_j such that $d(x_i, x) < 1$ and $d(x_j, y) < 1$. Then

$$\begin{aligned} d(x, y) &\leq d(x, x_i) + d(x_i, x_j) + d(x_j, y) \\ &< M + 2 \end{aligned}$$

so X is bounded. $\square$

Combining Theorem 3.4.8 and Theorem 3.3.6 we see that every compact space is closed and bounded. The Heine-Borel theorem will give us the converse in the case of subsets of $\mathbb{R}^n$.

Theorem 3.4.9 (Heine-Borel)**.** *Let $A \subseteq \mathbb{R}^n$. Then A is compact if and only if A is closed and bounded.*

Proof. One direction follows from Theorems 3.3.6 and 3.4.8.

For the other direction, if A is bounded then we may find some n-cell I with $A \subseteq I$. But by Theorem 3.4.6 we have that I is compact, and since A is closed as a subset of $\mathbb{R}^n$, it is closed as a subset of I. Thus A is compact by Theorem 3.3.7. □

Example 3.4.10. This theorem only works in $\mathbb{R}^n$. Indeed, let X be any infinite discrete metric space. Then X is bounded and closed, but never compact.

Theorem 3.4.11 (Bolzano-Weierstrass)**.** *Every bounded sequence in $\mathbb{R}^n$ has a convergent subsequence.*

Proof. Every bounded sequence $(x_n)_n$ can be contained in some n-cell I. But I is compact, hence sequentially compact, and we are done. □

3.5 Connected sets

Definition 3.5.1. A metric space X is said to be *connected* if it cannot be written as a disjoint union of two non-empty open sets.

Proposition 3.5.2. *A space X is connected if and only if the only clopen subsets of X are $\emptyset$ and X.*

Proof. First suppose that X is connected and let $U \subseteq X$ be clopen. Then

$$X = U \cup U^c$$

and U and U^c are disjoint open sets. Since X is connected, one of U or U^c must be empty. Thus we get that $U = \emptyset$ or $U = X$ as required.

Next, suppose that the only clopen subsets of X are the trivial ones and suppose that

$$X = U \cup V$$

for U, V disjoint open subsets. Then $U^c = V$ is open, so U is clopen. Hence either $U = \emptyset$ or $U = X$ in which case $V = \emptyset$. Thus X cannot be written as a disjoint union of two open sets and is therefore connected. □

Theorem 3.5.3. *The connected subsets of $\mathbb{R}$ are precisely the intervals. That is, it is those $I \subseteq \mathbb{R}$ with the property that if $x, y \in I$ and $x < z < y$, then $z \in I$.*

Proof. Suppose that $A \subseteq \mathbb{R}$ is a subset without this property. That is, we may find $x, y \in A$ and $x < z < y$ with $z \notin A$. Then

$$A = (A \cap (-\infty, z)) \cup (A \cap (z, \infty))$$

witnesses A as a disjoint union of two non-empty open subsets.

On the other hand, suppose that $I \subseteq \mathbb{R}$ is an interval. We will replicate the proof in Example 3.2.33 which shows $\mathbb{R}$ has no non-trivial clopen subsets. Suppose that $S \subseteq I$ is a non-trivial clopen subset. Then choose $x \in S$ and $y \in I \setminus S$. Since $I \setminus S$ is also a non-trivial clopen subset we may WLOG assume that $x < y$. Set

$$z = \sup(S \cap [x, y]).$$

Then since S is closed we have that $z \in S$ and in particular $z \neq y$. Moreover, since S is open and $x \in S$, we must have $z > x$. Thus $x < z < y$. Since I is an interval, we know $[x, y] \subseteq I$ and it follows from the above that

$$(z, y] \subseteq I \setminus S$$

which contradicts that S is open, since no ball centered at z is contained in S, a contradiction. □

Theorem 3.5.4. *Let $\{X_\alpha : \alpha \in A\}$ be a family of connected subspaces of some space X with $\bigcap_{\alpha \in A} X_\alpha \neq \emptyset$. Then $\bigcup_{\alpha \in A} X_\alpha$ is connected.*

Proof. Let $x \in \bigcap_{\alpha \in A} X_\alpha$ and set

$$Y = \bigcup_{\alpha \in A} X_\alpha.$$

Suppose that we could write $Y = U \cup V$ for U, V disjoint open sets. Then WLOG $x \in U$. Now, for every $\alpha \in A$ we have that

$$X_\alpha = (X_\alpha \cap U) \cup (X_\alpha \cap V)$$

witnesses X_α as a disjoint union of two open sets. Since X_α is connected, one of these sets must be empty. But $x \in X_\alpha \cap U$ so it must be the case that $X_\alpha \cap U = X_\alpha$ and $X_\alpha \cap V = \emptyset$. In particular, we have that $U \supseteq X_\alpha$ for all α, so $U \supseteq Y$.

We conclude that $U = Y$ and $V = \emptyset$, so Y cannot be a written as a disjoint union of non-empty open sets, as required. □

Exercises

Exercise 3.1. (i) Show that the set $\{1/n : n \in \mathbb{N}\}$ is not closed a subset of $\mathbb{R}$, but is closed as a subset of

$$\mathbb{R}_{>0} = \{x \in \mathbb{R} : x > 0\} = (0, \infty).$$

(ii) Show from definition that $\{1/n : n \in \mathbb{N}\} \cup \{0\} \subseteq \mathbb{R}$ is compact.

Exercise 3.2. Let (X, d) be a metric space. Recall that we defined the open and closed balls of radius

r as follows:

$$B_r(x) = \{y \in X : d(x, y) < r\}$$
$$D_r(x) = \{y \in X : d(x, y) \leq r\}.$$

(i) Show that $\overline{B_r(x)} \subseteq D_r(x)$. (This should take one sentence if you quote results from lecture.)

(ii) Construct a metric space (X, d) and $x \in X$, $r > 0$ such that $\overline{B_r(x)} \neq D_r(x)$. That is to say, the closed ball of radius r centered at x is not necessarily the closure of the open ball of radius r centered at x.

Exercise 3.3. Let X be a set with two metrics d_1 and d_2. We say that d_1 is *equivalent* to d_2 if there exists constants $C_1, C_2 > 0$ such that for all $x, y \in X$ we have

$$C_1 d_1(x, y) \leq d_2(x, y) \leq C_2 d_1(x, y).$$

Show that if d_1 and d_2 are equivalent metrics on X, then a set $A \subseteq X$ is open with respect to d_1 if and only if A is open with respect to d_2.

Exercise 3.4. Let X be a metric space, and let $\mathcal{C} = \{U_\alpha : \alpha \in A\}$ be an open cover. We say that an open cover $\mathcal{S}$ is a *refinement* of $\mathcal{C}$ if for every $A \in \mathcal{S}$ there exists some $B \in \mathcal{C}$ with $A \subseteq B$.

(i) Show that a space X is compact if and only if every open cover has a finite refinement.

(ii) Show that every open cover has a refinement $\mathcal{S}$ consisting only of open balls.

(iii) Deduce using (i) and (ii) that a space X is compact if and only if every open cover by open balls has a finite subcover.

Exercise 3.5. Show that the product of two connected metric spaces is connected.

Exercise 3.6. Give an example of a nested sequence $C_1 \supseteq C_2 \supseteq C_3 \supseteq \cdots$ of closed, connected subsets of $\mathbb{R}^2$ such that $\bigcap_{n=1}^{\infty} C_n$ is not connected.

Exercise 3.7. Show that

$$\Delta = \{(x, x) : x \in X\} \subseteq X \times X$$

is closed for any metric space X.

Exercise 3.8 (∗)**.** Show that $\mathbb{R}$ cannot be written as a disjoint union of bounded, closed intervals of positive length.

4 Sequences and series

4.1 Convergence

Recall that a sequence $(x_n)_n$ is a metric space (X, d) is said to converge to x if for all $\varepsilon > 0$ there exists some $N \in \mathbb{N}$ such that for all $n \geq N$ we have $d(x_n, x) < \varepsilon$. In this scenario we will write

$$\lim_n x_n = x$$

and say that $(x_n)_n$ *converges*. If no such x exists we will say that $(x_n)_n$ *diverges*.

Note that convergence is dependent on the ambient space and is closely related to the notion of closed-ness that we have studied. Indeed, for example, $(1/n)_n$ converges in $\mathbb{R}$ but diverges in $(0, \infty)$.

We now collect some basic facts about sequences that we have implicitly encountered in the previous section.

Theorem 4.1.1. *Let $(x_n)_n$ be a sequence in a metric space X. Then the following statements hold:*

(i) $(x_n)_n$ converges to $x \in X$ if and only if every neighborhood of x contains x_n for all but finitely many n

(ii) If $x \in X$ and $x' \in X$ with $x_n \to x$ and $x_n \to x'$, then $x = x'$

(iii) If $(x_n)_n$ converges, then $(x_n)_n$ is bounded

Proof. For (i), suppose that $x_n \to x$ and let U be a neighborhood of x. Since U contains an open ball centered at x, we reduce to showing that every open ball about x contains all but finitely many x_n's. For this consider $B_r(x)$ for $r > 0$. Since we have $d(x_n, x) < r$ for n sufficiently large we are done.

On the other hand, suppose every neighborhood of x contains all but finitely many of the x_n's. Then given $\varepsilon > 0$, we have that $B_\varepsilon(x)$ contains every x_n for all but finitely many n. Then taking N to be the largest subscript not contained in $B_\varepsilon(x)$ we have that for all $n \geq N + 1$ that $d(x_n, x) < \varepsilon$ as required.

For (ii), suppose that $x \neq x'$. Then take

$$\varepsilon = d(x, x')/2 > 0.$$

Then for n sufficiently large we have that $d(x_n, x) < \varepsilon$ and $d(x_n, x') < \varepsilon$, but this is impossible since

$$B_\varepsilon(x) \cap B_\varepsilon(x') = \emptyset.$$

For (iii), let $x_n \to x$. Remark that, by (i), $B_1(x)$ contains all but finitely many terms say $x_{n_1}, \dots, x_{n_k}$. Then setting

$$M = \max\{d(p, q) : p, q \in \{x_{n_1}, \dots, x_{n_k}, x\}\} + 2$$

we have that $\{x_n\}$ is bounded by M. □

We now take account of how numerical sequences behave with respect to arithmetical operations. First, we equip $\mathbb{C}$ with the metric

$$d(z_1, z_2) = |z_1 - z_2|$$

where

$$|a + bi| = \sqrt{a^2 + b^2}.$$

Note that this is the same as the metric on $\mathbb{R}^2$ when we identify $\mathbb{C}$ with the plane via real and imaginary parts.

Theorem 4.1.2. *Let $(s_n)_n$ and $(t_n)_n$ be sequences of complex numbers with $s_n \to s$ and $t_n \to t$. Then*

(i) $s_n + t_n \to s + t$

(ii) $cs_n \to cs$ *and* $c + s_n \to c + s$ *for any number* $c \in \mathbb{C}$

(iii) $s_n t_n \to st$

(iv) if $s_n \neq 0$ *for all* n *and* $s \neq 0$*, then* $1/s_n \to 1/s$

Proof. For (i), let $\varepsilon > 0$. Then we may find N_1, N_2 such that for all $n \geq N_1$ we have

$$|s_n - s| < \varepsilon/2$$

and for all $n \geq N_2$ we have

$$|t_n - t| < \varepsilon/2.$$

Then for $N = \max\{N_1, N_2\}$ we have that

$$\begin{aligned} |(s_n + t_n) - (s + t)| &\leq |s_n - s| + |t_n - t| \\ &< \varepsilon/2 + \varepsilon/2 \\ &= \varepsilon \end{aligned}$$

as required.

The proof of (ii) is an exercise for the reader.

For (iii), we write

$$s_n t_n - st = (s_n - s)(t_n - t) + s(t_n - t) + t(s_n - s). \tag{4.1.1}$$

Now, fix $\varepsilon > 0$. Then for n sufficiently large we have that

$$|s_n - s| < \sqrt{\varepsilon}$$

and

$$|t_n - t| < \sqrt{\varepsilon}.$$

It follows that for n sufficiently large we have

$$|(s_n - s)(t_n - t)| < \sqrt{\varepsilon} \cdot \sqrt{\varepsilon} = \varepsilon.$$

Thus $\lim_n (s_n - s)(t_n - t) = 0$. Taking limits of both sides of (4.1.1) and applying (i) and (ii) we get that

$$\lim_n (s_n t_n - st) = 0.$$

Applying (ii) once more we get $\lim_n s_n t_n = st$ as required.

For (iv), since $s \neq 0$, for n sufficiently large (say $n \geq m$) we have that $|s_n - s| < |s|/2$. When this holds, we have that

$$|s_n| > |s|/2.$$

Fix $\varepsilon > 0$. For n sufficiently large (say $n \geq M$) we have that

$$|s_n - s| < \frac{1}{2}|s|^2 \varepsilon.$$

Then for $n \geq \max\{m, M\}$ we have that

$$\begin{aligned} \left| \frac{1}{s_n} - \frac{1}{s} \right| &= \left| \frac{s - s_n}{s s_n} \right| \\ &< \frac{|s|^2 \varepsilon / 2}{|s|^2 / 2} \\ &= \varepsilon \end{aligned}$$

as required. □

Limits also play well with arithmetic in $\mathbb{R}^n$.

Theorem 4.1.3. *(i) Let*

$$x_n = (x_{1,n}, \dots, x_{k,n})$$

be a sequence in $\mathbb{R}^k$*. Then* $x_n \to (y_1, \dots, y_k)$ *if and only if* $x_{i,n} \to y_i$ *for all* i*.*

(ii) Let $(x_n)_n$ *and* $(y_n)_n$ *be sequences in* $\mathbb{R}^k$ *and let* $\beta \in \mathbb{R}$*. If* $x_n \to x$ *and* $y_n \to y$ *then we have*

$$\lim_n x_n \cdot y_n = x \cdot y$$
$$\lim_n (x_n + y_n) = x + y$$
$$\lim_n \beta x_n = \beta x.$$

Proof. (i) follows from Proposition 3.4.3. (ii) then follows from (i) and Theorem 4.1.2. □

4.2 Cauchy sequences

Definition 4.2.1. A sequence $(x_n)_n$ in a metric space X is said to be *Cauchy* if for all $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that for all $n, m \geq N$ we have $d(x_n, x_m) < \varepsilon$.

Remark 4.2.2. Cauchy-ness is a property of the sequence itself. It does not required addition input to state, unlike convergence which requires reference to a point to which the sequence converges. As such, Cauchy-ness also does not depend on the ambient space unlike convergence. The following example illustrates this point.

Example 4.2.3. (i) $(1/n)_n$ is Cauchy in $(0, \infty)$ but not convergent.

(ii) $(n)_n$ in $\mathbb{R}$ is not Cauchy.

(iii) Only the eventually constant sequence in a discrete metric space are Cauchy.

(iv) If we equip $\mathbb{R}$ with the metric

$$d(x, y) = |e^{-x} - e^{-y}|$$

then $(n)_n$ is Cauchy but convergent.

Our goal for this section is to show that being Cauchy is the same as being convergent in certain settings.

Proposition 4.2.4. *Let* $(x_n)_n$ *be a Cauchy sequence with a convergent subsequence* $x_{n_k} \to x$*. Then* $(x_n)_n$ *is convergent with* $x_n \to x$*.*

Proof. Fix $\varepsilon > 0$. Then there exists some N such that for $n, m \geq N$ we have

$$d(x_n, x_m) < \varepsilon/2.$$

Now, since $x_{n_k} \to x$, we may find some n_ℓ such that $n_\ell \geq N$ and $d(x_{n_\ell}, x) < \varepsilon/2$. We then have that for $n \geq N$,

$$\begin{aligned} d(x_n, x) &\leq d(x_n, x_{n_\ell}) + d(x_{n_\ell}, x) \\ &< \varepsilon/2 + \varepsilon/2 \\ &= \varepsilon \end{aligned}$$

as required. □

Proposition 4.2.5. *Every Cauchy sequence is bounded.*

Proof. Let $(x_n)_n$ be a Cauchy sequence. Then there exists some N such that for $n, m \geq N$ we have that $d(x_n, x_m) < 1$. Then setting

$$M = \max\{d(p, q) : p, q \in \{x_1, \dots, x_N\}\} + 1$$

we have that $\{x_n : n \in \mathbb{N}\}$ is bounded by M. □

Theorem 4.2.6. *(i) In any metric space, every convergent sequence is Cauchy.*

(ii) Every Cauchy sequence in a compact metric space is convergent.

(iii) Every Cauchy sequence in $\mathbb{R}^n$ is convergent.

Proof. For (i), suppose that $x_n \to x$. Then fix $\varepsilon > 0$. Let N be such that for all $n \geq N$ we have $d(x_n, x) < \varepsilon/2$. Then for all $n, m \geq N$ we have

$$\begin{aligned} d(x_n, x_m) &\leq d(x_n, x) + d(x, x_m) \\ &< \varepsilon/2 + \varepsilon/2 \\ &= \varepsilon \end{aligned}$$

as required.

For (ii) and (iii), we remark that by Proposition 4.2.4 it suffices to show every Cauchy sequence has a convergent subsequence in these settings. For (ii), we know that every sequence in a compact metric space has a convergent subsequence. For (iii), Proposition 4.2.5 tells us every Cauchy sequence is bounded so we are done by Bolzano-Weierstrass. □

4.3 Completeness

We remark that Theorem 4.2.6(i) tells us that being convergent is stronger than being Cauchy and Example 4.2.3 shows that in general it is a strictly stronger notion. However, Theorem 4.2.6(ii) and (iii) gives us examples of spaces where the two notions are in fact equivalent. We give such spaces a name.

Definition 4.3.1. We say that a metric space X is *complete* if every Cauchy sequence in X is convergent.

Proposition 4.3.2. *Every closed subspace of a complete space is complete.*

Proof. Let X be complete and $Y \subseteq X$ a closed subset. Then given a Cauchy sequence $(y_n)_n$ in Y, since X is complete we know that $y_n \to x$ for some $x \in X$. But Y is closed so we must have that $x \in Y$, so $(y_n)_n$ is convergent in Y. □

Combining with Theorem 4.2.6, are examples of complete spaces now include compact spaces and closed subsets of Euclidean space.

4.4 $\limsup$ and $\liminf$

Similar to how every convergent sequence is Cauchy, but not necessarily the converse, we have by Theorem 4.1.1(iii) that every convergent sequence is bounded. The converse to this is also clearly not true: The sequence $((-1)^n)_n$ is $\mathbb{R}$ is bounded but not convergent.

There is however a special case in which being bounded does imply convergence.

Definition 4.4.1. A sequence of real numbers $(x_n)_n$ is said to be

(i) *monotonically increasing* if $x_n \leq x_{n+1}$ for all n

(ii) *monotonically decreasing* if $x_n \geq x_{n+1}$ for all n

If either of these hold, we say that the sequence $(x_n)_n$ is *monotone*.

Proposition 4.4.2. *A monotone sequence $(x_n)_n$ is $\mathbb{R}$ is convergent if and only if it is bounded. In this case, we have*

(i) $x_n \to \sup\{x_n\}$ if x_n is monotonically increasing

(ii) $x_n \to \inf\{x_n\}$ if x_n is monotonically decreasing

Proof. We have already seen that all convergent sequences are bounded, giving one direction.

For the other direction, we prove the proposition in the case of x_n monotonically increasing as the decreasing case is proved similarly. Since $\{x_n\}$ is bounded and non-empty we may take

$$\alpha = \sup\{x_n\}.$$

Now fix $\varepsilon > 0$. Since $\alpha - \varepsilon$ cannot be an upper bound for $\{x_n\}$ we must be able to find some N with $x_N > \alpha - \varepsilon$. But $(x_n)_n$ is increasing so for all $n \geq N$ we have that $x_n \geq x_N$. Combining this with the fact that α is an upper bound for $\{x_n\}$, we find that for all $n \geq N$

$$\alpha - \varepsilon < x_n \leq \alpha.$$

Hence $x_n \to \alpha$ as required. □

If we allow statements of the form $x_n \to \infty$ and $x_n \to -\infty$ then parts (i) and (ii) of Proposition 4.4.2 may be written without boundedness hypotheses. We thus make such a definition

Definition 4.4.3. Let $(x_n)_n$ be a sequence of real numbers. We write $x_n \to +\infty$ if for all $M \in \mathbb{R}$ there exists $N \in \mathbb{N}$ such that $x_n \geq M$ for all $n \geq N$.

Similarly, we write $x_n \to -\infty$ if for all $M \in \mathbb{R}$ there exists $N \in \mathbb{N}$ such that $x_n \leq M$ for all $n \geq N$.

We will also write $\lim_n x_n = +\infty$ and $\lim_n x_n = -\infty$ to mean the same thing as $x_n \to +\infty$ and $x_n \to -\infty$ respectively.

Remark 4.4.4. Even though we may now write $\lim_n x_n = \pm\infty$ sequences for which this hold are still said to be divergent (check that such sequences are indeed divergent per our definition!).

Using Proposition 4.4.2 it is possible to extract from every sequence of real numbers two sequences which are guaranteed to converge. Let $(s_n)_n$ be a sequence of real numbers. For now, suppose that $(s_n)_n$ is bounded below. Define the following modified sequence:

$$\underline{s}_n = \inf\{s_k : k \geq n\}.$$

If $(s_n)_n$ is bounded above, then we may also define

$$\bar{s}_n = \sup\{s_k : k \geq n\}.$$

Then one checks that $(\underline{s}_n)_n$ is monotonically increasing and $(\bar{s}_n)_n$ is monotonically decreasing. We use these to make the following definition

Definition 4.4.5. Let $(s_n)_n$ be a sequence of real numbers. We define the *lower limit* of $(s_n)_n$, denoted by $\liminf_n s_n$, as follows:

(i) $\liminf_n s_n = \lim_n \underline{s}_n$ if $(s_n)_n$ is bounded below

(ii) $\liminf_n s_n = -\infty$ otherwise.

Similarly, we define the *upper limit* of $(s_n)_n$, denoted by $\limsup_n s_n$, as follows:

(i) $\limsup_n s_n = \lim_n \bar{s}_n$ if $(s_n)_n$ is bounded above

(ii) $\limsup_n s_n = +\infty$ otherwise.

Theorem 4.4.6. *For every sequence of real numbers $(s_n)_n$ we have that*

$$\liminf_n s_n \leq \limsup_n s_n$$

with equality if and only if $s_n \to s$ with

$$s = \liminf_n s_n = \limsup_n s_n.$$

Proof. We remark that

$$\limsup_n s_n = \inf\{\sup\{s_k : k \geq n\} : n \in \mathbb{N}\}$$

and

$$\liminf_n s_n = \sup\{\inf\{s_k : k \geq n\} : n \in \mathbb{N}\}$$

by Proposition 4.4.2. Set $\alpha = \limsup_n s_n$ and $\beta = \liminf_n s_n$ and let $\varepsilon > 0$. Then we may find $N \in \mathbb{N}$ such that

$$\beta - \varepsilon < \inf\{s_k : k \geq N\}.$$

However, notice that we then in fact have for all $n \geq N$

$$\begin{aligned}\beta - \varepsilon &< \inf\{s_k : k \geq N\} \\ &\leq \inf\{s_k : k \geq n\} \\ &\leq \sup\{s_k : k \geq n\}.\end{aligned}$$

We find then that $\beta - \varepsilon$ is then a lower bound for $\{\sup\{s_k : k \geq n\} : n \in \mathbb{N}\}$ so $\beta - \varepsilon \leq \alpha$. But this holds for all $\varepsilon > 0$, so $\beta \leq \alpha$.

For the second part, suppose that $s_n \to s$ for $s \in \mathbb{R}$. Then fix $\varepsilon > 0$. We may find $N \in \mathbb{N}$ such that for all $n \geq N$ we have $|s_n - s| < \varepsilon$. In particular, we have that

$$s_n < s + \varepsilon$$

for all $n \geq N$ so

$$\limsup_n s_n \leq s + \varepsilon.$$

But also we have that $s - \varepsilon < s_n$ for $n \geq N$ so

$$\liminf_n s_n \geq s - \varepsilon.$$

Thus

$$\limsup_n s_n - \varepsilon \leq s \leq \liminf_n + \varepsilon$$

for all $\varepsilon > 0$ so in fact

$$\limsup_n \leq s \leq \liminf_n s_n.$$

Combined with first part of the proposition we find that

$$s = \limsup_n s_n = \liminf_n s_n$$

as required.

Next suppose that $s = \limsup_n s_n = \liminf_n s_n$. We wish to show that $s_n \to s$. We will check the case when $s \in \mathbb{R}$ and the case of $s = \pm\infty$ is an exercise. Let $\varepsilon > 0$. Then we may find $N \in \mathbb{N}$ such that

$$s - \varepsilon < \inf\{s_k : k \geq N\} \leq \sup\{s_k : k \geq N\} < s + \varepsilon$$

(check this!). But then for all $n \geq N$ we have that

$$s - \varepsilon < s_n < s + \varepsilon$$

as required. □

Example 4.4.7. (i) Let $s_n = (-1)^n/(1 + 1/n)$. Then

$$\limsup_n s_n = 1, \quad \liminf_n s_n = -1$$

and since $\limsup_n s_n \neq \liminf_n s_n$, $(s_n)_n$ does not converge.

(ii) Let $(s_n)_n$ be an enumeration of the rational numbers. Then $(s_n)_n$ is neither bounded above nor below so

$$\limsup_n s_n = +\infty, \quad \liminf_n s_n = -\infty.$$

4.5 Series

From here on out, it will be assumed that all sequences are complex valued.

Given a sequence $(a_n)_n$ we denote by

$$\sum_{n=p}^{q} a_n$$

the sum $a_p + a_{p+1} + \cdots + a_q$ of all sequential terms starting at the index p and ending at the index q.

Definition 4.5.1. Let $(a_n)_n$ be a sequence. The sequence of *partial sums* is the sequence

$$s_n = \sum_{k=1}^{n} a_k.$$

Further, we will write

$$\sum_{n=1}^{\infty} a_n \tag{4.5.1}$$

for the limit of the partial sums $\lim_n s_n$. We refer to the above as a *series* or *infinite series*. We will say that (4.5.1) converges if $(s_n)_n$ converges, otherwise we will say that (4.5.1) diverges.

Remark 4.5.2. Since $\sum_{n=1}^{\infty} a_n$ is simply notation for $\lim_n s_n$, it makes sense to write $\sum_{n=1}^{\infty} a_n = \infty$ using the notation of the previous sections. However, in this case we still say that $\sum_n a_n$ diverges.

Collecting the results from the previous sections about sequences and applying them to the sequence of partial sums we get the following results.

Theorem 4.5.3. *Let $(a_n)_n$ be a sequence. Then $\sum_n a_n$ converges if and only if for all $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that for all $n, m \geq N$ we have*

$$\left| \sum_{k=n}^{m} a_k \right| < \varepsilon.$$

Proof. This criterion is precisely the assertion that the sequence $(s_n)_n$ of partial sums is Cauchy, and we have seen that $\mathbb{R}^2$ (and hence $\mathbb{C}$) is complete. □

Theorem 4.5.4 (Divergence Test)**.** *If $\sum_n a_n$ converges then $a_n \to 0$.*

Proof. For any Cauchy sequence $(s_n)_n$ we have that $s_{n+1} - s_n \to 0$ (check this!). Applying this to the sequence of partial sums this precisely says that $a_n \to 0$. □

Example 4.5.5. The converse to Theorem 4.5.4 does not hold. Indeed, one has that

$$\sum_{n=1}^{\infty} \frac{1}{n} = \infty$$

(see homework or later in notes). However, it does give a criterion for showing that a series diverges. In particular, if $a_n \not\to 0$ then $\sum_n a_n$ necessarily diverges. This is sometimes called the *divergence test*.

Theorem 4.5.6. *Let $(a_n)_n$ be a sequence of non-negative terms. Then $\sum_n a_n$ converges if and only if the sequence of partial sums is bounded.*

Proof. If the a_n's are non-negative, then the sequence of partial sums is monotone increasing and the result follows from Proposition 4.4.2. □

The above theorem justifies the intuition that a sequence on non-negative terms either converges or "blows up." Since sequences of non-negative terms have simple behavior in regard to convergence, we define a stronger notion of convergence.

Definition 4.5.7. Let $(a_n)_n$ be a sequence. We say that $\sum_n a_n$ is *absolutely convergent* if $\sum_n |a_n|$ converges.

Proposition 4.5.8. *If $\sum_n a_n$ is absolutely convergent, then $\sum_n a_n$ is convergent.*

Proof. We have that

$$\left| \sum_{k=n}^{m} a_k \right| \leq \sum_{k=n}^{m} |a_k|$$

and so the result follows from Theorem 4.5.3. □

Remark 4.5.9. The above proof shows that if $\sum_n |a_n|$ converges then $\sum_n |a_n|$ is Cauchy. One might wonder whether there is a more direct proof that does not use that $\mathbb{C}$ is a complete metric space. However, it turns out that if one considers series in more general metric spaces where it makes sense to add things, then completeness is necessary for absolute convergence to convergence.

Example 4.5.10. Being absolutely convergent is a strictly stronger notion than being convergent. Indeed, $\sum_n (-1)^n/n$ is convergent (see Homework 3) whereas $\sum_n 1/n$ is divergent (see Theorem 4.6.4).

Theorem 4.5.11 (Comparison Test)**.** *Let $(a_n)_n$ be a sequence.*

(i) If $|a_n| \leq c_n$ *for all n sufficiently large and* $\sum_n c_n$ *converges then* $\sum_n a_n$ *is absolutely convergent.*

(ii) If $a_n \geq b_n \geq 0$ *and* $\sum_n b_n$ *diverges, then* $\sum_n a_n$ *diverges.*

Proof. For (i), to show that $\sum_n |a_n|$ converges, it suffices by Theorem 4.5.6 to show that the sequence of partial sums is bounded. However, we have that

$$\begin{aligned}\sum_{k=1}^{n} |a_k| &\leq \sum_{k=1}^{n} c_k \\ &\leq \sum_{n=1}^{\infty} c_n < \infty\end{aligned}$$

as required.

For (ii), again by Theorem 4.5.6, it suffices to show that $\sum_n a_n$ is unbounded. Since the terms b_n are non-negative and $\sum_n b_n$ diverges, we know that the sequence of partial sums of the b_n's are unbounded. But we have that

$$\sum_{k=1}^{n} a_k \geq \sum_{k=1}^{n} b_k$$

so the sequence of partial sums of the a_n's are unbounded as well. □

4.6 Special series

In the previous section, we saw some basic tests for establishing convergence and divergence of series. However, to make use of these test, especially the comparison test, we need families of series for which their convergence and divergence is known. In this section we take account of some common series and their convergence properties.

Theorem 4.6.1. *For any* $x \in \mathbb{C}$ *with* $|x| < 1$ *we have that*

$$\sum_{n=0}^{\infty} x^n = 1 + x + x^2 + \cdots = \frac{1}{1-x}.$$

If $|x| \geq 1$ *then the above series diverges.*

Proof. We have an explicit expression for the partial sums of these series. In fact, we have for $x \neq 1$ that

$$s_n = 1 + x + \cdots + x^{n-1} = \frac{1-x^n}{1-x}. \tag{4.6.1}$$

Letting $n \to \infty$, if $|x| < 1$ then the above expression approaches the claimed one. If $|x| \geq 1$, then the limit of (4.6.1) does not exist or is infinite (check this!). □

Remark 4.6.2. The series $\sum_n x^n$, or more generally those of the form $\sum_n ax^n$ for $a \in \mathbb{C}$, are referred to as *geometric series*.

We now want to address convergence of series of the form $\sum_n n^{-p}$. To do this we need the following peculiar looking lemma.

Lemma 4.6.3. *Let $(a_n)_n$ be a monotone decreasing sequence of non-negative terms, i.e. $a_1 \geq a_2 \geq a_3 \geq \cdots \geq 0$. Then $\sum_n a_n$ converges if and only if*

$$\sum_{n=1}^{\infty} 2^n a_{2^n} = a_1 + 2a_2 + 4a_4 + 8a_8 + \cdots$$

converges.

Proof. Since all terms in sight are non-negative, convergence is equivalent to bounded-ness of partial sums. For this, let

$$\begin{aligned} s_n &= a_1 + a_2 + \cdots + a_n \\ t_k &= a_1 + 2a_2 + \cdots + 2^k a_{2^k}. \end{aligned}$$

For $n < 2^k$ we have that

$$\begin{aligned} s_n &\leq a_1 + (a_2 + a_3) + \cdots + (a_{2^k} + \cdots + a_{2^{k+1}-1}) \\ &\leq a_1 + 2a_2 + \cdots + 2^k a_{2^k} \\ &= t_k. \end{aligned}$$

Thus if $(t_k)_k$ is bounded, then $(s_n)_n$ is bounded.

On the other hand, if $n > 2^k$, then

$$\begin{aligned} s_n &\geq a_1 + a_2 + (a_3 + a_4) + \cdots + (a_{2^{k-1}-1} + \cdots + a_{2^k}) \\ &\geq \frac{1}{2}a_1 + a_2 + 2a_4 + \cdots + 2^{k-1}a_{2^k} \\ &= \frac{1}{2}t_k \end{aligned}$$

so if $(s_n)_n$ is bounded then $(t_k)_k$ is bounded, as required. □

The above lemma lets us turn certain non-geometric series into geometric series, as the next theorem shows.

Theorem 4.6.4. *The series*
$$\sum_{n=1}^{\infty} \frac{1}{n^p} = 1 + \frac{1}{2^p} + \frac{1}{3^p} + \cdots$$
converges if and only if $p > 1$.

Proof. By Lemma 4.6.3 we have that $\sum_n n^{-p}$ converges if and only if

$$\sum_{k=1}^{\infty} 2^k \cdot \frac{1}{2^{kp}} = \sum_{k=1}^{\infty} \frac{1}{2^{k(p-1)}}$$

converges. By Theorem 4.6.1, this converges if and only if $|1/2^{p-1}| < 1$ which is if and only if $p > 1$. □

4.7 More convergence tests

In this section we take account of more useful convergence tests. The next two tests, the root and ratio test, may be thought of morally as combining the geometric series test and comparison test.

Theorem 4.7.1 (Root test)**.** *Let* $(a_n)_n$ *be a sequence and set* $\alpha = \limsup_n \sqrt[n]{|a_n|}$*. Then*

(i) if $\alpha < 1$*, then* $\sum_n a_n$ *converges absolutely*

(ii) if $\alpha > 1$*, then* $\sum_n a_n$ *diverges*

(iii) if $\alpha = 1$*, then the test is inconclusive.*

Proof. For (i), take any β with $\alpha < \beta < 1$. Since

$$\alpha = \inf_n \{\sup\{\sqrt[k]{|a_k|} : k \geq n\}\} < \beta$$

we may find $N \in \mathbb{N}$ such that
$$\sup\{\sqrt[k]{|a_k|} : k \geq N\} < \beta.$$

Thus for all $n \geq N$ we have that $\sqrt[n]{|a_n|} < \beta$. Thus we have that $|a_n| < \beta^n$ for all $n \geq N$ and $\sum_n \beta^n$ converges by Theorem 4.6.1. Hence by the Comparison Test (Theorem 4.5.11) we have that $\sum_n a_n$ converges absolutely.

For (ii), notice that if $\lim_n a_n = 0$ then for n sufficiently large we have that $|a_n| < 1$ and thus $\sqrt[n]{a_n} < 1$ for n sufficiently large. It follows that if $\lim_n a_n = 0$, then $\limsup_n \sqrt[n]{|a_n|} \leq 1$. Thus if $\alpha > 1$, then $\lim_n a_n \neq 0$ and we must diverge by the divergence test. □

Example 4.7.2. The inconclusivity of Theorem 4.7.1 when $\alpha = 1$ is necessary. Indeed, we have that for all $p \in \mathbb{R}$

$$\limsup_n \sqrt[n]{1/n^p} = \limsup_n \left(\frac{1}{n^{1/n}}\right)^p = 1.$$

Thus the root test is inconclusive for all sequences of the form $\sum_n n^{-p}$. However, we know by Theorem 4.6.4 that these series are convergent when $p > 1$ and divergent when $p \leq 1$.

Example 4.7.3. The series

$$\sum_{n=1}^{\infty} \frac{1}{n^n}$$

is convergent. Indeed,

$$\lim_n \left(\frac{1}{n^n}\right)^{1/n} = \lim_n \frac{1}{n} = 0 < 1.$$

Theorem 4.7.4 (Ratio test)**.** *Let $(a_n)_n$ be a sequence of non-zero terms. Then*

(i) if $\limsup_n \left|\frac{a_{n+1}}{a_n}\right| < 1$ *then* $\sum_n a_n$ *converges absolutely*

(ii) if $\left|\frac{a_{n+1}}{a_n}\right| \geq 1$ *for n sufficiently large, then* $\sum_n a_n$ *diverges.*

Proof. For (i), choose any β with $\limsup_n |a_{n+1}/a_n| < \beta < 1$. Then there exists $N \in \mathbb{N}$ such that for $n \geq N$ we have that

$$\left|\frac{a_{n+1}}{a_n}\right| < \beta.$$

Hence for $n > N$ we have that

$$|a_n| < \beta^{n-N}|a_N| = \frac{|a_N|}{\beta^N} \cdot \beta^n.$$

However, $\sum_n \beta^{n-N}|a_N|$ converges since $\beta < 1$ by Theorem 4.6.1. Thus $\sum_n a_n$ converges absolutely by the comparison test.

For (ii), suppose that $|a_{n+1}/a_n|$ for all $n \geq N$. Then past the first N terms, we have that the sequence becomes monotone increasing. In particular, for $n \geq N$, we have

$$|a_n| \geq |a_N| > 0$$

so $\lim_n a_n \neq 0$ and so $\sum_n a_n$ must diverge. □

Remark 4.7.5. Case (ii) of Theorem 4.7.4 holds in particular when

$$\lim_n \left| \frac{a_{n+1}}{a_n} \right| \geq 1.$$

Thus an alternative, but weaker statement, would be to say: Suppose

$$\alpha = \lim_n \left| \frac{a_{n+1}}{a_n} \right| \tag{4.7.1}$$

exists. Then

(i) if $\alpha < 1$, then $\sum_n a_n$ converges absolutely

(ii) if $\alpha > 1$, then $\sum_n a_n$ diverges

(iii) if $\alpha = 1$, then the test is inconclusive.

However, as stated, Case (i) of Theorem 4.7.4 does not require that the limit (4.7.1) exists.

Example 4.7.6. Consider the series

$$\frac{1}{2} + \frac{1}{3} + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{2^3} + \frac{1}{3^3} + \cdots .$$

Then we have that

$$\begin{aligned} \limsup_n \sqrt[n]{|a_n|} &= \lim_n \left(\frac{1}{2^n} \right)^{\frac{1}{2n}} = \frac{1}{\sqrt{2}} \\ \limsup_n \left| \frac{a_{n+1}}{a_n} \right| &= \lim_n \frac{1}{2} \left(\frac{3}{2} \right)^n = \infty \end{aligned}$$

so the Root Test implies convergence but the Ratio Test does not apply.

It turns out the above example is indicative of a broader phenomenon. In fact, we have the following result which says that the root test is a strictly stronger test for determining convergence than the ratio test.

Proposition 4.7.7. *For any sequence $(c_n)_n$ of positive numbers, we have that*

$$\limsup_n \sqrt[n]{c_n} \leq \limsup_n \left| \frac{c_{n+1}}{c_n} \right|.$$

Proof. See [Rud64, Theorem 3.37]. □

In practice, however, the ratio test tends to be easier to apply.

The root and ratio tests above only allow us to conclude absolute convergence of series. For series which converge but not absolutely, such as $\sum_n (-1)^n/n$, a different test is needed.

Theorem 4.7.8 (Alternating series test)**.** *Let $(a_n)_n$ be a decreasing sequence of non-negative real numbers. Then*

$$\sum_{n=1}^{\infty} (-1)^n a_n$$

converges if and only if $\lim_n a_n = 0$.

Proof. One direction is the divergence test, so we prove the other direction. Let $(s_n)_n$ be the sequence of partial sums. We show that the sequence of even partial sums converges, i.e. $\lim_n s_{2n}$ exists. Then since

$$\lim_n (s_{n+1} - s_n) = \lim_n a_n = 0,$$

this tells us that the whole sequence $(s_n)_n$ converges (check this!).

Notice that

$$\begin{aligned} s_{2n+2} &= s_{2n} - a_{2n+1} + a_{2n+2} \\ &\leq s_{2n} \end{aligned}$$

since $a_{2n+1} \geq a_{2n+2}$. Thus $(s_{2n})_n$ forms a decreasing sequence, so it suffices to show that it is lower bounded. However,

$$\begin{aligned} s_{2n} &= -a_1 + \underbrace{(a_2 - a_3)}_{\geq 0} + \cdots + \underbrace{(a_{2n-2} - a_{2n-1})}_{\geq 0} + \underbrace{a_{2n}}_{\geq 0} \\ &\geq -a_1 \end{aligned}$$

as required. □

Example 4.7.9. Theorem 4.7.8 can be used to immediately conclude that $\sum_n (-1)^n/n$ is convergent. In fact, it shows that

$$\sum_{n=1}^{\infty} \frac{(-1)^n}{n^p}$$

converges for all $p > 0$.

4.8 Power series

Definition 4.8.1. Let $(c_n)_n$ be a sequence of complex numbers. Series of the form

$$\sum_{n=0}^{\infty} c_n z^n$$

where z is a complex number are referred to as *power series*.

In a power series, z is often treated as a variable and thus we get a function

$$f(z) = \sum_{n=0}^{\infty} c_n z^n.$$

However, this definition only makes sense when the right hand side converges. Thus it is of interest to understand when power series converge.

Theorem 4.8.2. *Let $(c_n)_n$ be a sequence. Set*

$$\alpha = \limsup_n \sqrt[n]{|c_n|}, \quad R = \frac{1}{\alpha}$$

where we take $R = \infty$ when $\alpha = 0$. Then the power series $\sum_n c_n z^n$ converges absolutely when $|z| < R$ and diverges when $|z| > R$.

Proof. We have that

$$\limsup_n \sqrt[n]{|c_n z^n|} = |z|\alpha.$$

Then $|z|\alpha < 1$ if and only if $|z| < R$ so the conclusion follows from the root test (Theorem 4.7.1). ☐

Definition 4.8.3. Given a power series $\sum_n c_n z^n$, the value of R in Theorem 4.8.2 is referred to as the *radius of convergence* of $\sum_n c_n z^n$.

Example 4.8.4. We define the exponential function by

$$\exp(z) = \sum_{n=0}^{\infty} \frac{z^n}{n!}. \tag{4.8.1}$$

We have that

$$\limsup_n \left(\frac{1}{n!}\right)^{1/n} = 0 \tag{4.8.2}$$

so exp has radius of convergence ∞. Thus exp defines a function $\mathbb{C} \to \mathbb{C}$.

Alternatively, since (4.8.2) is a difficult limit, we could apply the ratio test. Indeed,

$$\begin{aligned}\lim_n \left| \frac{z^{n+1}/(n+1)!}{z^n/n!} \right| &= \lim_n \frac{|z|}{n+1} \\ &= 0.\end{aligned}$$

Thus (4.8.1) converges for all $z \in \mathbb{C}$ by the ratio test, so the radius of convergence of (4.8.1) is necessarily ∞.

4.9 Rearrangements

Normally, addition is commutative. That is, we have that $a + b = b + a$. It is thus tempting to think that series, i.e. infinite sums, should also be unaffected by rearranging its terms.

Definition 4.9.1. We say that a series is *conditionally convergent* if is is convergent but not absolutely convergent.

Definition 4.9.2. Let $(a_n)_n$ be a sequence. A *rearrangement* of $(a_n)_n$ is any sequence of the form $(a_{k_n})_n$ where

$$\begin{aligned}\mathbb{N} &\longrightarrow \mathbb{N} \\ n &\longmapsto k_n\end{aligned}$$

is a bijection.

Example 4.9.3. Consider the series

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots .$$

We know by the alternating series test that this sequence converges to some s. We also know that the odd partial sums form a decreasing sequence so

$$s < 1 - \frac{1}{2} + \frac{1}{3} = \frac{5}{6}.$$

However, consider the following rearrangement of the terms giving the series

$$1 + \frac{1}{3} - \frac{1}{2} + \frac{1}{5} + \frac{1}{7} - \frac{1}{4} + \cdots \tag{4.9.1}$$

where we add two positive terms then one negative each time. Every term in this series still appears, however since

$$\frac{1}{4k-3} + \frac{1}{4k-1} - \frac{1}{2k} > 0$$

if $(s'_n)_n$ denotes the sequence of partial sums then $(s'_{3n})_n$ is an increasing sequence. In particular,

$$\limsup_n s'_{3n} > s'_3 = \frac{5}{6}.$$

Thus if the series (4.9.1) converges, it certainly cannot converge to s.

The above example shows that rearranging terms may change the value of series, and potentially even affect convergence. In fact, in the case of conditionally convergent series, the situation is as bad as it could be.

Theorem 4.9.4. *Let $\sum_n a_n$ be conditionally convergent. Then for any*

$$-\infty \leq \alpha \leq \beta \leq \infty$$

there exists a rearrangement $(a'_n)_n$ of $(a_n)_n$, such that, if $(s'_n)_n$ are the corresponding partial sums,

$$\liminf_n s'_n = \alpha$$
$$\limsup_n s'_n = \beta.$$

Proof (Sketch). For a complete proof, see [Rud64, Theorem 3.54].

Since the series $\sum_n a_n$ is convergent, we know that $\lim_n a_n = 0$. However, since it is not absolutely convergent one can show that the sum of all the negative terms, when ordered in decreasing absolute value, must equal $-\infty$ and the sum of all the positive terms, when ordered in decreasing absolute value, must equal ∞.

Using this, we may pull from the negative and positive terms and required to oscillate the value of our partial sums. We can change its value as much as desired, since the sum of these terms is infinite, and the granularity at which we can adjust the value increases as we include more terms since $\lim_n a_n = 0$. Thus we can add positive terms until our partial sums approach β, then add negative terms until we approach down to α, and repeat, each time getting closer. □

However, when we are absolutely convergent, we can rearrange.

Theorem 4.9.5. *Let $\sum_n a_n$ be an absolutely convergent series with $\sum_n a_n = s$. Then for any rearrangement $(a'_n)_n$ of $(a_n)_n$ we have that $\sum_n a'_n = s$.*

Proof. Let $(a_{k_n})_n$ be our rearrangement of $(a_n)_n$ and denote by $(s'_n)_n$ the partial sums of $(a_{k_n})_n$. Fix

$\varepsilon > 0$. Then we may find $N \in \mathbb{N}$ such that for all $n, m \geq N$ we have that

$$\sum_{k=n}^{m} |a_k| < \varepsilon. \tag{4.9.2}$$

Now, choose p sufficiently large so that $\{1, \ldots, N-1\} \subseteq \{k_1, \ldots, k_p\}$. Then we have that

$$\begin{aligned} s_n - s_n' &= \sum_{\ell=1}^{n} a_\ell - \sum_{\ell=1}^{n} a_{k_\ell} \\ &= \sum_{\ell=N}^{n} a_\ell - \sum_{\substack{\ell=1 \\ k_\ell \geq N}}^{n} a_{k_\ell} \end{aligned}$$

since the terms a_i for $1 \leq i \leq N-1$ will cancel out. But then by (4.9.2) we have that

$$|s_n - s_n'| < 2\varepsilon.$$

Since ε was arbitrary and $s_n \to s$, we get that $s_n' \to s$ as required. □

Exercises

Exercise 4.1. (i) For every n, let S_n be the set of all rational numbers that can be represented as a fraction p/q with $|q| \leq n$. Show that S_n is complete.

(ii) Show that

$$\mathbb{Q} = \bigcup_{n \in \mathbb{N}} S_n.$$

(iii) Show that $\mathbb{Q}$ is not complete.

Exercise 4.2. Let $(s_n)_n$ be a sequence of real numbers. Show that if $(s_n)_n$ converges then $(|s_n|)_n$ converges. Given an example to show that the converse is not true.

Exercise 4.3. Give an example of two sequences $(s_n)_n$ and $(t_n)_n$ of real numbers such that

$$\liminf_n (s_n t_n) \neq \liminf_n s_n \cdot \liminf_n t_n$$

where all terms in the above are finite (i.e. not $\pm\infty$).

Exercise 4.4. Show that for two real sequences $(s_n)_n$ and $(t_n)_n$ that

$$\limsup_n (s_n + t_n) \leq \limsup_n s_n + \limsup_n t_n$$

(whenever the right hand side is not of the form $\infty - \infty$).

Exercise 4.5. Consider the sequence defined by

$$s_n = 1 + \frac{1}{2} + \cdots + \frac{1}{n}.$$

(i) Show that $(s_n)_n$ is monotone.

(ii) Show that

$$s_{2^{n+1}} - s_{2^n} \geq \frac{1}{2}.$$

Deduce that $(s_n)_n$ is unbounded.

(iii) Deduce from (i) and (ii) that $(s_n)_n$ does not converge.

Exercise 4.6. Suppose that we have a series

$$\sum_{n=1}^{\infty} a_n = a_1 + a_2 + a_3 + \cdots . \tag{†}$$

We will call a *regrouping* of $\sum_n a_n$ to be any series arising by grouping terms in the summation, e.g.

$$a_1 + (a_2 + a_3) + (a_4 + a_5 + a_6 + a_7) + a_8 + a_9 + (a_{10} + a_{11}) + \cdots$$

and

$$(a_1 + a_2 + a_3) + a_4 + a_5 + a_6 + \cdots$$

are examples of regroupings of (†).

(i) Show that if $\sum_n a_n$ converges, then every regrouping converges to the same value.

(ii) Give an example of a divergent series $\sum_n a_n$ which has a convergent regrouping.

Exercise 4.7. Suppose that $\sum_n (-1)^n 2^n a_n$ converges, does it follow that $\sum_n a_n$ converges? Prove or give a counterexample.

Exercise 4.8. Compute the radius of convergence of

$$\sum_{n=0}^{\infty} \frac{n!}{n^n} \cdot z^n.$$

Hint: You may use that

$$\lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n = e.$$

5 Continuity

5.1 Limits of functions and continuity

Our goal in this section is to consider the behavior of functions $f : X \to Y$ between two metric spaces X and Y. We would like to produce machinery for rigorously making statements such as "as the input x approaches a point p, $f(x)$ approaches the point q." To begin we first need to make rigorous which points in a metric space can be "approached."

Definition 5.1.1. Let X be a metric space and $E \subseteq X$. A point $p \in X$ is a *limit point* of E if every neighborhood (taken in X) of p contains a point in $E \setminus \{p\}$.

If $E = X$ and $p \in X$ is not a limit point, we say that it is an *isolated point*.

Remark 5.1.2. $p \in X$ being isolated is equivalent to $\{p\}$ being an open set.

Limit points are our rigorous notion for points that can be "approached."

Definition 5.1.3. Let X, Y be metric spaces, $E \subseteq X$ and $f : E \to Y$. If $p \in X$ is a limit point of E, then given some $q \in Y$ we say that $f(x)$ *approaches q as x approaches p*, written $f(x) \to q$ as $x \to p$ or

$$\lim_{x \to p} f(x) = q,$$

if for all $\varepsilon > 0$ there exists $\delta > 0$ such that

$$d_Y(f(x), q) < \varepsilon$$

for all points $x \in E$ with $0 < d_X(x, p) < \delta$.

Morally, this definition is saying that $f(x)$ can be made arbitrarily closed to q provided we make x sufficiently close (but not equal) to p.

Example 5.1.4. It's important to note that the statement $\lim_{x \to p} f(x) = q$ depends only on the behaviour of f at points near p which *are not equal* to p itself. As an example of this, consider the function $f : \mathbb{R} \to \mathbb{R}$ given by

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

Then $\lim_{x \to 0} f(x) = 1$ even though $f(0) = 0$.

Definition 5.1.5. Let $f : X \to Y$ be a function between two metric spaces. Given $p \in X$, we say that f is *continuous at p* if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$d_Y(f(x), f(p)) < \varepsilon$$

whenever $d_X(x, p) < \delta$. Otherwise we say that f is *discontinuous at p*.

We say that f is *continuous* if it is continuous at every $p \in X$.

Proposition 5.1.6. *A function $f : X \to Y$ is continuous at $p \in X$ if and only if either p is isolated or* $\lim_{x\to p} f(x) = f(p)$.

Proof. It is clear that if p is a limit point of X, then the definition of f being continuous at p implies that $\lim_{x\to p} f(x) = f(p)$. This gives one direction.

Now, suppose that p is isolated. Then there exists some $\delta > 0$ such that $B_\delta(p) = \{p\}$. Thus f is continuous at p since for any $\varepsilon > 0$ we have that

$$\begin{aligned} d_X(x, p) < \delta &\implies x = p \\ &\implies d_Y(f(x), f(p)) = 0 < \varepsilon. \end{aligned}$$

Lastly, suppose that p is a limit point and $\lim_{x\to p} f(x) = f(p)$. Then one verifies from the definitions that f is continuous at p. $\square$

Example 5.1.7. (i) For any metric space X, the identity function $\mathrm{id}_X : X \to X$ is continuous.

(ii) Constant functions are continuous.

(iii) The function $f : \mathbb{R} \to \mathbb{R}$ given by

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

is continuous at every point $x \neq 0$ and discontinuous at $x = 0$.

(iv) The function $f : \mathbb{R} \to \mathbb{R}$ given by

$$f(x) = \begin{cases} x & \text{if } x \in \mathbb{Q} \\ 0 & \text{if } x \notin \mathbb{Q} \end{cases}$$

is continuous at $x = 0$ and discontinuous at every $x \neq 0$.

The intuition for continuous functions is that there are no abrupt "jumps"—as we perturb the input by small amounts the output also only ever changes by small amounts.

Theorem 5.1.8. *Let $f : X \to Y$ and $g : Y \to Z$. Suppose that f is continuous at $p \in X$ and g is continuous at $f(p) \in Y$. Then $g \circ f$ is continuous at p.*

Proof. Fix $\varepsilon > 0$. Then as g is continuous at $f(p)$, there exists $\eta > 0$ such that

$$d_Y(y, f(p)) < \eta \implies d_Z(g(y), g(f(p))) < \varepsilon. \tag{5.1.1}$$

Then, since f is continuous at p, there exists $\delta > 0$ such that

$$d_X(x, p) < \delta \implies d_Y(f(x), f(p)) < \eta. \tag{5.1.2}$$

Combining (5.1.1) and (5.1.2) we see that

$$d_X(x, p) < \delta \implies d_Z(g(f(x)), g(f(p))) < \varepsilon.$$

Thus $g \circ f$ is continuous at p. □

Corollary 5.1.9. *The composition of two continuous functions is continuous.*

It turns out that being continuous is a *topological* property. That is, the continuity of a function $f : X \to Y$ can be checked with only the knowledge of what the open sets of X and Y are—the precise metric is not important.

Definition 5.1.10. Given a function $f : A \to B$ of sets and $V \subseteq B$, we write

$$f^{-1}(V) = \{a \in A : f(a) \in V\}$$

for the collection of all elements of A which are mapped into V by f.

Theorem 5.1.11. *A function $f : X \to Y$ is continuous if and only if $f^{-1}(V)$ is open in X for all open subsets $V \subseteq Y$.*

Proof. First suppose that f is continuous and let $V \subseteq Y$ be open. Then we may write V as a union of open balls, i.e.

$$V = \bigcup_{B \in \mathcal{C}} B$$

where $\mathcal{C}$ is some collection of open balls in Y. Then

$$f^{-1}(V) = \bigcup_{B \in \mathcal{C}} f^{-1}(B),$$

so it suffices to show that for every open ball $B_r(y)$ in Y, $f^{-1}(B_r(y))$ is open. For this, let $p \in f^{-1}(B_r(y))$ so

$$d_Y(f(p), y) < r.$$

Set $\varepsilon = r - d_Y(f(p), y) > 0$. By continuity of f, there exists a $\delta > 0$ such that

$$\begin{aligned} d_X(x, p) < \delta &\implies d_Y(f(x), f(p)) < \varepsilon \\ &\implies d_Y(f(x), y) < r. \end{aligned}$$

Thus $B_\delta(p) \subseteq f^{-1}(B_r(y))$ so $f^{-1}(B_r(y))$ is open as required.

Conversely, suppose that $f^{-1}(V)$ is open for all $V \subseteq Y$ and let $p \in X$. Fix $\varepsilon > 0$. Then $f^{-1}(B_\varepsilon(f(p)))$ is open and contains p. Hence we may find $\delta > 0$ such that

$$B_\delta(p) \subseteq f^{-1}(B_\varepsilon(f(p))).$$

But this exactly says that

$$d_X(x, p) < \delta \implies d_Y(f(x), f(p)) < \varepsilon$$

so f is continuous at p, as required. □

Remark 5.1.12. The equivalent formulation of continuity in Theorem 5.1.11 is often given as the definition of continuity and in some ways is the "correct" definition since it generalizes to spaces where we have a notion of open sets but not necessarily any notion of distance.

Since $f^{-1}(Y \setminus V) = X \setminus f^{-1}(V)$ we also have the following corollary:

Corollary 5.1.13. *A function $f : X \to Y$ is continuous if and only if $f^{-1}(S)$ is closed in X for every closed subset $S \subseteq Y$.*

Definition 5.1.14. Let $f : A \to B$ be a function of sets. Given $S \subseteq A$ we define the *restriction of f to S*, denoted $f|_S$, to be the function $f|_S : S \to B$ given by the formula $f|_S(x) = f(x)$.

Corollary 5.1.15. *Let $f : X \to Y$ be a continuous function and $Z \subseteq X$. The restriction $f|_Z : Z \to Y$ is continuous.*

Proof. Notice that for $V \subseteq Y$ we have that

$$(f|_Z)^{-1}(V) = f^{-1}(V) \cap Z.$$

Thus the result follows from Theorem 5.1.11 and Theorem 3.2.10. $\square$

5.2 Arithmetic and continuity

We now want to take account of operations which preserve continuity so that we may build large classes of continuous functions. First we prove a sequential characterization of continuity.

Proposition 5.2.1. *Let X, Y be metric spaces, $E \subseteq X$, p a limit point of E and $f : E \to Y$. Then $\lim_{x\to p} f(x) = q$ if and only if for every sequence $(p_n)_n$ in E, $p_n \neq p$, with $p_n \to p$ we have that $f(p_n) \to q$.*

Proof. For one direction, let $(p_n)_n$ be any such sequence and fix $\varepsilon > 0$. Then as $\lim_{x\to p} f(x) = q$ we have that there exists some $\delta > 0$ such that for all $x \in E$,

$$0 < d_X(x, p) < \delta \implies d_Y(f(x), q) < \varepsilon.$$

Since $p_n \to p$, for n sufficiently large we have that $d_X(p_n, p) < \delta$. Moreover, $p_n \neq p$ so for n sufficiently large $0 < d_X(p_n, p) < \delta$. Thus for n sufficiently large we have $d_Y(f(p_n), q) < \varepsilon$ and so $f(p_n) \to q$ as required.

Conversely, suppose the sequence formulation holds but $\lim_{x\to p} f(x) \neq q$. Let $\varepsilon > 0$ be such that no such $\delta > 0$ as in Definition 5.1.3 exists. Then for every natural n we may find $p_n \in E$ with $0 < d_X(p_n, p) < 1/n$ and $d_Y(f(p_n), q) \geq \varepsilon$. Then $(p_n)_n$ is a sequence satisfying the necessary hypotheses but having $f(p_n) \not\to q$, a contradiction. $\square$

Corollary 5.2.2. *A function $f : X \to Y$ is continuous at $p \in X$ if and only if for every sequence $p_n \to p$ in X we have that $f(p_n) \to f(p)$.*

Using this sequential characterization of continuity the following theorems are corollaries of previous work with sequences.

Definition 5.2.3. Given functions $f : X \to \mathbb{F}$ and $g : X \to \mathbb{F}$ where $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, we define new functions $f + g$, $f \cdot g$ and f/g (provided g is never zero) pointwise, i.e. via

$$(f + g)(x) = f(x) + g(x)$$
$$(f \cdot g)(x) = f(x)g(x)$$
$$(f/g)(x) = f(x)/g(x).$$

Similarly, given vector valued functions $\mathbf{f} : X \to \mathbb{R}^k$ and $\mathbf{g} : X \to \mathbb{R}^k$ we may define $\mathbf{f} + \mathbf{g}$ and $\mathbf{f} \cdot \mathbf{g}$ by pointwise addition and dot product respectively. From here on out, operations on functions will assumed to be computed pointwise unless otherwise stated.

Theorem 5.2.4. *Let $f : X \to \mathbb{F}$ and $g : X \to \mathbb{F}$ be continuous functions where $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. Then*

(i) $f + g$ is continuous

(ii) fg is continuous

(iii) f/g is continous provided g is never zero

Proof. This follows immediately from Corollary 5.2.2 and Theorem 4.1.2. □

Theorem 5.2.5. *Let $f : X \to Y$ and $g : X \to Z$ be continuous. Then $(f, g) : X \to Y \times Z$ is continuous.*

Proof. This follows immediately from Corollary 5.2.2 and Proposition 3.4.3. □

Remark 5.2.6. In fact, more is true. A function $h = (f, g) : X \to Y \times Z$ into a product is continuous *if and only if* both f, g are continuous. To show this, it suffices to show that each projection $\pi_Y : Y \times Z \to Y$ and $\pi_Z : Y \times Z \to Z$ is continuous. Then if h is continuous the compositions $f = \pi_Y \circ h$ and $g = \pi_Z \circ h$ are also necessarily continuous.

Corollary 5.2.7. *All polynomials from $\mathbb{R} \to \mathbb{R}$ are continuous.*

Proof. All polynomials can be built using arithmetic out of the identity function $f(x) = x$ and constant functions which are continuous (Example 5.1.7), thus are continuous by Theorem 5.2.4. □

Example 5.2.8. The map $f : \mathbb{R}^k \to \mathbb{R}$ given by $f(x) = \|x\|$ is continuous. Indeed, by the triangle inequality we have that

$$\|f(x) - f(y)\| = \| \|x\| - \|y\| \| \leq \|x - y\|. \tag{5.2.1}$$

Thus fix any $\varepsilon > 0$ and $x \in \mathbb{R}^k$. Then by (5.2.1) we have that

$$\|x - y\| < \varepsilon \Longrightarrow \|f(x) - f(y)\| < \varepsilon.$$

Since ε was arbitrary, we see that f is continuous at x, so f is continuous.

More generally, we say that a function $f : X \to Y$ is *Lipschitz* if there is a constant $C > 0$ such that

$$d_Y(f(x), f(y)) \leq C d_X(x, y).$$

In this case we say that f is *C-Lipschitz*. We then have that any Lipschitz function is continuous as we may take $\delta = \varepsilon/C$ is the definition of continuity.

Example 5.2.9. Given any metric space X, the metric $d : X \times X \to \mathbb{R}$ is continuous. Indeed, since continuity depends only on the opens sets (Theorem 5.1.11) it does not matter which of the equivalent metrics we put on $X \times X$. We then have that

$$\begin{aligned} |d(x_1, y_1) - d(x_2, y_2)| &\leq |d(x_1, y_1) - d(x_1, y_2)| + |d(x_1, y_2) - d(x_2, y_2)| \\ &\leq d(y_1, y_2) + d(x_1, x_2) \\ &= d_1((x_1, y_1), (x_2, y_2)) \end{aligned}$$

via the triangle inequality. Hence d is 1-Lipschitz with respect to the d_1 metric on the product. Thus d is continuous by the discussion in Example 5.2.8.

Since constant map and identity maps are continuous, we have by Theorem 5.2.5 that for any $x_0 \in X$ the map

$$\begin{aligned} X &\longrightarrow X \times X \\ x &\longmapsto (x_0, x) \end{aligned}$$

is continuous. Composing with d, we then see that for any $x_0 \in X$, the map

$$\begin{aligned} X &\longrightarrow \mathbb{R} \\ x &\longmapsto d(x_0, x) \end{aligned}$$

is also continuous.

5.3 Interaction with compact and connected sets

Continuous functions behave well with compact and connected sets. Intuitively, continuous functions send points that are close together to points that are close together. Thus, intuitively, continuous functions should not be able to separate connected spaces in two nor create too large of spaces from sufficiently "finite" spaces.

Definition 5.3.1. Let $f : A \to B$ be a function of sets. For $S \subseteq A$ we define $f(S)$ to be the image of all points in S under f, i.e.

$$f(S) = \{f(s) : s \in S\}.$$

Theorem 5.3.2. *Let $f : X \to Y$ be a continuous function and $K \subseteq X$ compact. Then $f(K)$ is compact.*

Proof. Let $\mathcal{C}$ be an open cover of $f(K)$. Then we have that

$$\mathcal{S} = \{f^{-1}(V) : V \in \mathcal{C}\}$$

is an open cover of K. Indeed, each $f^{-1}(V)$ is open by Theorem 5.1.11 and it covers K. This is because if $x \in K$, then $f(x) \in f(K)$ so $f(x) \in V$ for some $V \in \mathcal{C}$, so $x \in f^{-1}(V)$.

It follows by compactness of K that $\mathcal{S}$ has a finite subcover, say

$$K \subseteq f^{-1}(V_1) \cup \cdots \cup f^{-1}(V_n).$$

But then $f(K) \subseteq V_1 \cup \cdots \cup V_n$ so we are done. $\square$

This has some useful corollaries.

Theorem 5.3.3. *Let $f : K \to \mathbb{R}$ be a continuous function where K is compact. Then f achieves a maximum and minimum, i.e. there exists $p, q \in K$ such that*

$$f(p) = \sup_{x \in K} f(x) = \sup f(K)$$
$$f(q) = \inf_{x \in K} f(x) = \inf f(K).$$

Proof. By Theorem 5.3.2, we know that $f(K) \subseteq \mathbb{R}$ is compact, hence closed. It follows that $\sup f(K) \in f(K)$ and $\inf f(K) \in f(K)$. In particular, we may find points $p, q \in K$ such that $f(p) = \sup f(K)$ and $f(q) = \inf f(K)$. $\square$

Definition 5.3.4. We say that a function $f : X \to Y$ is *closed* if $f(E) \subseteq Y$ is closed for every closed subset $E \subseteq X$.

Theorem 5.3.5. *Let $f : K \to Y$ be a continuous map where K is compact. Then f is closed.*

Proof. Let $E \subseteq K$ be closed. Then E is compact as it is a closed subset of a compact space. Hence $f(E) \subseteq Y$ is compact by Theorem 5.3.2, hence closed, as required. $\square$

Example 5.3.6. Let X be any metric space and $K_1, K_2 \subseteq X$ disjoint compact subsets. Then there exists $\varepsilon > 0$ such that

$$d(x, y) \geq \varepsilon$$

for all $x \in K_1$ and $y \in K_2$.

Indeed, we know that the metric $d : X \times X \to \mathbb{R}$ is continuous by Example 5.2.9. Moreover, by Theorem 3.4.4 we know that $K_1 \times K_2$ is compact. Hence $d|_{K_1 \times K_2} : K_1 \times K_2 \to \mathbb{R}$ is continuous and thus achieves a minimum, say at (x_0, y_0), so set

$$\varepsilon = \inf_{(x,y) \in K_1 \times K_2} d(x, y) = d(x_0, y_0).$$

Then $d(x, y) \geq \varepsilon$ for all $x \in K_1$ and $y \in K_2$ by definition, and since $K_1 \cap K_2 = \emptyset$ we know that $\varepsilon = d(x_0, y_0) > 0$.

We now return to connected sets.

Theorem 5.3.7. *Let $f : X \to Y$ be a continuous function and let $C \subseteq X$ be connected. Then $f(C)$ is connected.*

Proof. Suppose that $f(C) \subseteq U \cup V$ for U, V open. Our goal is to show that either $U \supseteq f(C)$ or $V \supseteq f(V)$. However, this then gives that

$$C \subseteq f^{-1}(U) \cup f^{-1}(V).$$

Since C is connected and both $f^{-1}(U)$ and $f^{-1}(V)$ are open, we may WLOG assume that $f^{-1}(U) \supseteq C$. But then $f(C) \subseteq U$ as required. □

Corollary 5.3.8 (Intermediate Value Theorem)**.** *Let $f : \mathbb{R} \to \mathbb{R}$ be continuous. Given $a \leq b$ and $f(a) \leq y \leq f(b)$ (or $f(b) \leq y \leq f(a)$), there exists $a \leq c \leq b$ with $f(c) = y$.*

Proof. We have seen (Theorem 3.5.3) that the connected subsets of $\mathbb{R}$ are precisely the intervals. It follows that $f([a, b])$ is connected, thus an interval. It follows that any y as in the hypothesis satisfies $y \in f([a, b])$ and thus we may find $a \leq c \leq b$ with $f(c) = y$. □

Example 5.3.9. This theorem does not have a converse—there are functions $\mathbb{R} \to \mathbb{R}$ which send intervals to intervals which are not continuous. Consider for example

$$f(x) = \begin{cases} \sin\left(\frac{1}{x}\right) & \text{if } x \neq 0 \\ 0 & \text{if } x = 0. \end{cases}$$

Then f is continuous at every $x \neq 0$ (assuming for now that we know $\sin(x)$ to be continuous) as it is a composition of two continuous functions. However, f is not continuous at 0. Indeed, we have that

$$\frac{1}{\pi/2 + 2\pi n} \to 0$$

but

$$f\left(\frac{1}{\pi/2 + 2\pi n}\right) = 1 \to 1 \neq f(0).$$

Thus f is not continuous.

f does, however, send intervals to intervals. To see this, let $I \subseteq \mathbb{R}$ be an interval. If $0 \notin I$, then $f(I)$ is the image of a connected set under a continuous function $\mathbb{R} \setminus \{0\} \to \mathbb{R}$, hence connected.

If $0 \in I$, we have two cases. Either $I = \{0\}$ in which case $f(I) = \{0\}$ is an interval. Or I is not the singleton zero set and $f(I) = [-1, 1]$ which is again an interval.

5.4 Homeomorphisms and topological properties of metric spaces

Now that we have the correct notion of "maps" between metric spaces, i.e. continuous maps, we may speak of when two spaces are "the same."

Definition 5.4.1. We say that a continuous function $f : X \to Y$ is a *homeomorphism* if there exists a continuous function $g : Y \to X$ such that $f \circ g = \mathrm{id}_Y$ and $g \circ f = \mathrm{id}_X$.

We say that two spaces X, Y are *homeomorphic*, denoted $X \cong Y$, if there exists a homeomorphism $X \to Y$.

Definition 5.4.2. We say that a continuous function $f : X \to Y$ is *open* if $f(V) \subseteq Y$ is open for all open subsets $V \subseteq X$.

Proposition 5.4.3. *Let $f : X \to Y$ be a homeomorphism. Then $f(V)$ is open if and only if V is open and $f(E)$ in closed if and only if E is closed. In particular, homeomorphisms are open and closed.*

Proof. If $f(V)$ is open, then since f is bijective we have

$$V = f^{-1}(f(V))$$

is open by continuity of f. Conversely, let g be the inverse of f. If V is open then

$$f(V) = g^{-1}(V)$$

is open by continuity of g.

The proof for closed subsets is similar. □

Remark 5.4.4. Having a continuous inverse is *strictly* stronger than saying that f is continuous and bijective. Indeed, let $(\mathbb{R}, d_{\mathrm{disc}})$ be the real line with the discrete metric and let $\mathbb{R}$ be the real line with the standard metric. Then the identity function

$$\begin{aligned} f : (\mathbb{R}, d_{\mathrm{disc}}) &\longrightarrow \mathbb{R} \\ x &\longmapsto x \end{aligned}$$

is continuous (since the pullback of any open is open as all subsets of discrete metric spaces are open) and it is clearly bijective as a map of sets. However, it is not an open map since $f(\{0\}) = \{0\}$ is not open. Hence, f cannot be a homeomorphism by Proposition 5.4.3.

Proposition 5.4.5. *Let $f : X \to Y$ be a continuous bijection. Then the following are equivalent:*

(i) f is a homeomorphism

(ii) f is open

(iii) f is closed.

Proof. Let $g : Y \to X$ be the set-theoretic inverse to f. Then (i) is equivalent to g being continuous. But

$$g^{-1}(V) = f(V)$$

so the pullback of opens along g are open if and only if f is open. Similarly, the pullbacks of closed along g are closed if and only if f is closed. □

In light of the above proposition, we do have a special case where being continuous and bijective *is* sufficient for being a homeomorphism.

Theorem 5.4.6 (Topological inverse function theorem)**.** *Let $f : K \to Y$ be a continuous bijection where K is compact. Then f is a homeomorphism.*

Proof. This follows from Proposition 5.4.5 and Theorem 5.3.5. □

Example 5.4.7. In the above theorem, it is import that the *domain* is the one that is compact. We may have continuous bijections $f : Y \to K$ into compact spaces which are not homeomorphisms. Indeed, take any connected, compact space K with more than one point and let (K, d_{disc}) be the metric space with the same underlying set as K and the discrete metric. Then, just as in Remark 5.4.4 we have that the set-theoretic identity function $(K, d_{\text{disc}}) \to K$ is a continuous bijection, but not open, hence not a homeomorphism.

Example 5.4.8. Consider the subspace $\mathrm{SU}(2) \subseteq \mathbb{C}^4$ given by all 2×2 matrices

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

with complex entries satisfying

$$\det(M) = ad - bc = 1 \tag{5.4.1}$$

and

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = MM^{\dagger} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} \overline{a} & \overline{c} \\ \overline{b} & \overline{d} \end{pmatrix}. \tag{5.4.2}$$

This subspace is called the *special unitary group*. We will show that

$$\mathrm{SU}(2) \cong S^3 = \{v \in \mathbb{C}^2 : \|v\| = 1\}.$$

Now, (5.4.2) forces that

$$1 = a\overline{a} + b\overline{b} = |a|^2 + |b|^2 \tag{5.4.3}$$
$$1 = |c|^2 + |d|^2 \tag{5.4.4}$$
$$0 = a\overline{c} + b\overline{d}. \tag{5.4.5}$$

Multiplying both sides of (5.4.5) by d and using (5.4.1), we get that

$$\begin{aligned} 0 &= ad\overline{c} + bd\overline{d} \\ &= (1 + bc)\overline{c} + b|d|^2 \\ &= \overline{c} + b(|c|^2 + |d|^2) \\ &= \overline{c} + b. \end{aligned}$$

Thus $b = -\overline{c}$. Similarly, one finds that $d = \overline{a}$.

Thus consider the map

$$\begin{aligned} f : S^3 = \{(\alpha, \beta) \in \mathbb{C}^2 : |\alpha|^2 + |\beta|^2 = 1\} &\longrightarrow \mathrm{SU}(2) \\ (\alpha, \beta) &\longmapsto \begin{pmatrix} \alpha & -\overline{\beta} \\ \beta & \overline{\alpha} \end{pmatrix}. \end{aligned}$$

It is continuous as each projection is and by the above it is bijective. Since S^3 is compact by Heine-Borel (as it is closed and bounded), we get by Theorem 5.4.6 that f is a homeomorphism.

By Proposition 5.4.3, given a homeomorphism $f : X \to Y$ we may use f to identify points of X with points of Y in such a way that preserves open sets. As such, definitions which care only about open sets and not the metric itself will be invariant under homeomorphisms. We give such properties a name.

Definition 5.4.9. Let $\mathcal{P}$ be a property of metric spaces. We say that $\mathcal{P}$ is a *topological property* if given any two homeomorphic spaces X, Y, $\mathcal{P}$ holds for X if and only if $\mathcal{P}$ holds for Y.

Theorem 5.4.10. *Compactness and connectedness are topological properties.*

Proof. This follows from Theorems 5.3.2 and 5.3.7. □

Example 5.4.11. Combining the above theorem and Example 5.4.8, we see that SU(2) is compact.

5.5 Uniform continuity

Definition 5.5.1. Let $f : X \to Y$ be a function between two metric spaces. We say that f is *uniformly continuous* if for all $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$d_X(x, y) < \delta \implies d_Y(f(x), f(y)) < \varepsilon$$

for all $x, y \in X$.

Uniformly continuity should be thought of as regular continuity, except the δ we pick must work for *every* point. As such, it also only makes sense to speak of a function being uniformly continuous on its whole domain—speaking of uniform continuity at a single point is meaningless.

Example 5.5.2. Lipschitz functions (see Example 5.2.8) are uniformly continuous.

Example 5.5.3. $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^2$ is not uniformly continuous. Indeed, we have that, for $\delta > 0$,

$$\begin{aligned} |f(x) - f(x + \delta)| &= |x^2 - (x + \delta)^2| \\ &= \delta|2x + \delta|. \end{aligned}$$

Thus by making x sufficiently large we may find two points distance δ apart such that $|f(x) - f(x + \delta)|$ is arbitrarily large. It follows that f cannot be uniformly continuous.

Theorem 5.5.4. *Let $f : K \to Y$ be a continuous function where K is compact. Then f is uniformly continuous.*

Proof. Fix $\varepsilon > 0$. Since f is continuous, for each point $p \in K$ we may find some $\varphi(p) > 0$ such that

$$d_K(x, p) < \varphi(p) \implies d_Y(f(x), f(p)) < \varepsilon/2. \tag{5.5.1}$$

We then have that

$$\{B_{\varphi(p)/2}(p) : p \in K\}$$

forms an open cover of K, so by compactness we may write

$$K = B_{\varphi(p_1)/2}(p_1) \cup \cdots \cup B_{\varphi(p_n)/2}(p_n)$$

for some finite collection of points $p_1, \ldots, p_n \in K$.

Now set

$$\delta = \frac{1}{2}\min\{\varphi(p_1), \dots, \varphi(p_n)\}.$$

We claim that this δ works. For this, let $p, q \in K$ with $d_K(p, q) < \delta$. First, we may find p_i such that $d_K(p, p_i) < \varphi(p_i)/2$. We then have that

$$\begin{aligned} d_K(q, p_i) &\leq d_K(q, p) + d_K(p, p_i) \\ &< \delta + \varphi(p_i)/2 \\ &\leq \varphi(p_i) \end{aligned}$$

so $p, q \in B_{\varphi(p_i)}(p_i)$. Hence we have that

$$\begin{aligned} d_Y(f(p), f(q)) &\leq d_Y(f(p), f(p_i)) + d_Y(f(q), f(p_i)) \\ &< \varepsilon/2 + \varepsilon/2 \\ &= \varepsilon \end{aligned}$$

using (5.5.1) as required. □

Example 5.5.5. We saw a failure of Theorem 5.5.4 in the case of non-compact domain in Example 5.5.3. We can give another family of examples when the domain is bounded, but not compact.

Let $E \subseteq \mathbb{R}$ be a bounded, non-compact subset. By Heine-Borel, this means that E must not be closed. Thus let $x_0 \in \lim E \setminus E$ and consider

$$f(x) = \frac{1}{x - x_0}.$$

Then given any $\delta > 0$, we may find $x \in E$ with $|x - x_0| < \delta/2$. Since f is unbounded as $x \to x_0$, we may then make $|f(x) - f(y)|$ as large as desired by taking y sufficiently close to x. In particular, considering only $y \in E$ with $|y - x_0| < \delta/2$, we may make $|f(x) - f(y)|$ as large as desired while keeping $|x - y| < \delta$.

It follows that $f : E \to \mathbb{R}$ cannot be uniformly continuous.

5.6 Discontinuities of functions on the real line

We now focus in on discontinuities of functions $E \to \mathbb{R}$ where $E \subseteq \mathbb{R}$.

Definition 5.6.1. Let $f : I \to \mathbb{R}$, $I \subseteq \mathbb{R}$ an interval, be a function and let $c \in I$. We say that $f(x) \to r$ *as* $x \to c^-$, denoted

$$\lim_{x \to c^-} f(x) = r,$$

if for every sequence $t_n \to c$ in $(-\infty, c) \cap I$ we have $f(t_n) \to r$. We also refer to this as the *left limit of f as $x \to c$*.

Similarly, we write

$$\lim_{x \to c^+} f(x) = r$$

to mean that $f(t_n) \to r$ for every sequence $t_n \to c$ is $(c, \infty) \cap I$.

Remark 5.6.2. For notational simplicity, in this section we will write $f(c+)$ for $\lim_{x \to c^+} f(x)$ and $f(c-)$ for $\lim_{x \to c^-} f(x)$.

Definition 5.6.3. We say that $f : I \to \mathbb{R}$ has a *simple discontinuity*, or *discontinuity of the first type*, at $x \in (a, b)$ if f is discontinuous at x and $f(x+)$ and $f(x-)$ both exist.

If f has a discontinuity at x which is not simple, we say f has a *discontinuity of the second type* at x.

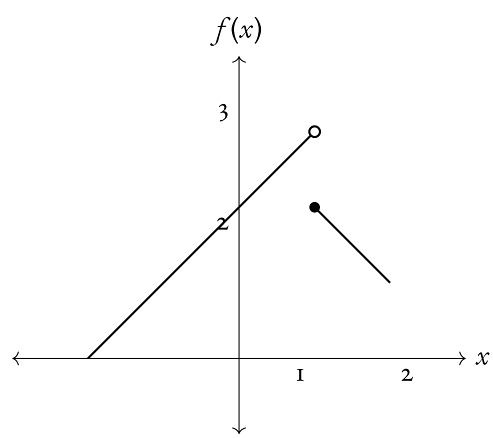
Comparing the definitions of $f(x-)$ and $f(x+)$ with the definition of continuity at x, we see that f has a simple discontinuity at x if and only if either $f(x+) \neq f(x-)$ or $f(x+) = f(x-) \neq f(x)$.

Example 5.6.4. The function

$$f(x) = \begin{cases} \sin\left(\frac{1}{x}\right) & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$$

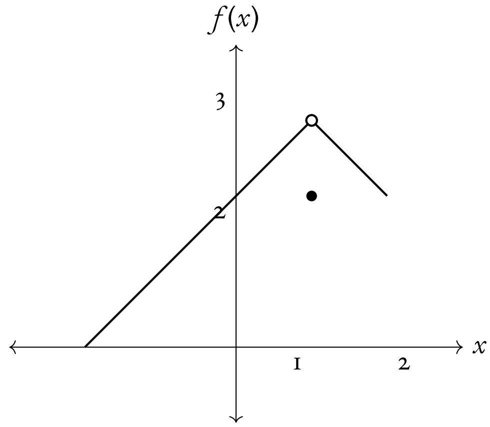
has a discontinuity of the second type at $x = 0$ as neither $f(0+)$ nor $f(0-)$ exists.

Example 5.6.5. The function given by the following graph



has $f(1-) = 3$ and $f(1-) = 1$, so f has a simple discontinuity at $x = 1$ since $f(1-) \neq f(1-)$.

The function given by the following graph



has $f(1-) = f(1+) = 3$ but $f(1) = 2$, so f again has a simple discontinuity at $x = 1$.

Definition 5.6.6. We say that a function $f : I \to \mathbb{R}$ is *monotone increasing* if we have that $f(x) \le f(y)$ whenever $x \le y$.

Similarly, we define *monotone decreasing* functions. We say that a function $f : I \to \mathbb{R}$ is *monotone* if it is either monotone increasing or monotone decreasing.

Proposition 5.6.7. *Let $f : I \to \mathbb{R}$ be a monotone increasing function. Then we have that*

$$f(x-) = \sup_{t<x} f(t)$$
$$f(x+) = \inf_{x<t} f(t).$$

Moreover, $f(x-) \le f(x) \le f(x+)$ and if $x < y$ then $f(x+) \le f(y-)$.

Proof. We have that $f(x)$ is an upper bound for $\{f(t) : t < x\}$, so $A = \sup\{f(t) : t < x\}$ exists and $A \le f(x)$. Thus we need to show that $f(x-) = A$. To do this, fix $\varepsilon > 0$. Then we may find $\delta > 0$ such that

$$A - \varepsilon < f(x - \delta) \le A.$$

Then for $x - \delta < t < x$ we have, as f is monotone increasing, that

$$A - \varepsilon < f(x - \delta) \le f(t) \le A.$$

Hence $\lim_{t\to x^-} f(t) = A$ as required.

The case for $f(x-)$ is similar.

To show the last statement, notice that

$$\begin{aligned} f(x+) &= \inf_{x<t} f(t) \\ &= \inf_{x<t<y} f(t) \\ &\leq \sup_{t<y} f(t) \\ &= f(y-) \end{aligned}$$

using that f is monotone increasing. □

Corollary 5.6.8. *Monotone functions only have simple discontinuities.*

Theorem 5.6.9. *A monotone function $f : I \to \mathbb{R}$ has at most countably many discontinuities.*

Proof. WLOG assume that f is monotone increasing and let x be a discontinuity. Then we have by Proposition 5.6.7 that

$$f(x-) < f(x) < f(x+).$$

Thus we may find a rational $q(x) \in (f(x-), f(x+))$.

Given two discontinuities x and y, say with $x < y$, then we have that $f(x+) \leq f(y-)$ by Proposition 5.6.7. Hence

$$(f(x-), f(x+)) \cap (f(y-), f(y+)) = \emptyset$$

so $q(x) \neq q(y)$. Thus

$$\begin{aligned} \{\text{discontinuities of } f\} &\longrightarrow \mathbb{Q} \\ x &\longmapsto q(x) \end{aligned}$$

is injective so f has countably many discontinuities. □

Exercises

Exercise 5.1. Let $f : X \to \mathbb{R}$ be a continuous function. Define

$$Z(f) = \{x \in X : f(x) = 0\}$$

to be the *zero set* of f. Show that $Z(f)$ is closed.

Exercise 5.2. Let X, Y be metric spaces. Show that the two projection maps

$$\begin{aligned} \pi_X : X \times Y &\longrightarrow X \\ (x, y) &\longmapsto x \end{aligned}$$

and

$$\begin{aligned} \pi_Y : X \times Y &\longrightarrow Y \\ (x, y) &\longmapsto y \end{aligned}$$

are continuous.

Exercise 5.3. Is the function $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = 2^x$ uniformly continuous?

Exercise 5.4. Let $f : [a, b] \to \mathbb{R}$ be a continuous function and $(x_n)_n$ a Cauchy sequence in $[a, b]$. Prove that $(f(x_k))_k$ is a Cauchy sequence.

Exercise 5.5. Let $f : S^1 \to \mathbb{R}$ be a continuous function where

$$S^1 = \{v \in \mathbb{R}^2 : \|v\| = 1\}$$

is the unit circle. Show that there exists some $x \in S^1$ such that $f(x) = f(-x)$.
[Hint: Consider the function $g(x) = f(x) - f(-x)$.]

Exercise 5.6. Let $(r_n)_n$ be an enumeration of the rational numbers and consider the function

$$f(x) = \sum_{n : r_n < x} \frac{1}{2^n}.$$

Show that f is discontinuous on $\mathbb{Q}$ and continuous on $\mathbb{R} \setminus \mathbb{Q}$.

6 Differentiation

6.1 The derivative

Definition 6.1.1. Let $f : I \to \mathbb{R}$ for $I \subseteq \mathbb{R}$ an interval. For $x \in I$, form the quotient

$$\varphi(t) = \frac{f(t) - f(x)}{t - x} : I \setminus \{x\} \to \mathbb{R}.$$

We set

$$f'(x) = \lim_{t \to x} \varphi(t) \tag{6.1.1}$$

provided the limit exists, and call $f'(x)$ the *derivative of f at x*.

When the limit in (6.1.1) exists, we say that f is *differentiable at x*. If f is differentiable at all points in I, we say that f is *differentiable*. The function

$$f' : \{x \in I : f \text{ is differentiable at } x\} \longrightarrow \mathbb{R}$$

is referred to as the *derivative* of f.

Theorem 6.1.2. *If $f : I \to \mathbb{R}$ is differentiable at $x \in I$, then f is continuous at x.*

Proof. We have that

$$\begin{aligned}\lim_{t\to x} f(t) &= \lim_{t\to x} \big(f(x) + \varphi(t)\cdot(t-x)\big)\\ &= f(x) + f'(x)\cdot 0\\ &= f(x)\end{aligned}$$

as required. □

Example 6.1.3. (i) Constant functions are differentiable with derivative zero. Indeed, let $f(x) = c$. Then

$$\varphi(t) = \frac{f(t)-f(x)}{t-x} = 0$$

so $\lim_{t\to x}\varphi(t) = 0$.

(ii) If $f(x) = x$, then $\varphi(t) = 1$ so $f'(x) = 1$.

Theorem 6.1.4. *Let $f, g : I \to \mathbb{R}$ be differentiable at $x \in I$. Then $f+g$, fg and f/g (provided $g(x) \neq 0$) are differentiable at x with derivatives given by*

(i) $(f+g)'(x) = f'(x) + g'(x)$

(ii) $(fg)'(x) = f'(x)g(x) + f(x)g'(x)$

(iii) $\left(\dfrac{f}{g}\right)'(x) = \dfrac{f'(x)g(x) - f(x)g'(x)}{g(x)^2}$.

Proof. Given a function h we will write φ_h for the φ occurring in the definition of the derivative of h (Definition 6.1.1).

For (i), we have that

$$\varphi_{f+g}(t) = \varphi_f(t) + \varphi_g(t)$$

from which the result follows.

For (ii), let $h = fg$. Then

$$h(t) - h(x) = f(t)[g(t) - g(x)] + g(x)[f(t) - f(x)]$$

so

$$\varphi_h(t) = f(t)\varphi_g(t) + g(x)\varphi_f(t).$$

Taking the limit as $t \to x$ and using Theorem 6.1.2 the result follows.

For (iii), since $g(x) \neq 0$ and g is continuous at x by Theorem 6.1.2, we know that $g(t) \neq 0$ for t near x. We then have that

$$\varphi_{f/g}(t) = \frac{1}{g(t)g(x)}\left(g(x)\varphi_f(t) - f(x)\varphi_g(t)\right).$$

Letting $t \to x$ and using Theorem 6.1.2 the result follows. □

Example 6.1.5. Combining Example 6.1.3 with Theorem 6.1.4 we see that every polynomial is differentiable. Moreover, using that the derivative of x is 1 and inductively implying Theorem 6.1.4(ii) we see that the derivative of x^n is nx^{n-1}.

Theorem 6.1.6 (Chain rule). *Let $f : I \to \mathbb{R}$ and $g : J \to \mathbb{R}$ where $I, J \subseteq \mathbb{R}$ are intervals and $f(I) \subseteq J$ so that $g \circ f$ exists. Given $x \in I$ such that $f'(x)$ exists and $g'(f(x))$ exists, then $(g \circ f)'(x)$ exists with*

$$(g \circ f)'(x) = g'(f(x))f'(x).$$

Proof. We may write

$$\begin{aligned} f(t) &= f(x) + (t - x) \cdot [f'(x) + u(t)] \\ g(t) &= g(f(x)) + (t - f(x)) \cdot [g'(f(x)) + v(t)] \end{aligned}$$

where u, v are functions such that $u \to 0$ as $t \to x$ and $v \to 0$ as $t \to f(x)$. Indeed, for f we take $u(t) = \varphi_f(t) - f'(x)$ and similarly for g.

It follows that

$$\begin{aligned} g(f(t)) &= g(f(x)) + (f(t) - f(x)) \cdot [g'(f(x)) + v(f(t))] \\ &= g(f(x)) + (t - x) \cdot [f'(x) + u(t)] \cdot [g'(f(x)) + v(f(t))]. \end{aligned}$$

Thus

$$\varphi_{g \circ f}(t) = (f'(x) + u(t)) \cdot (g'(f(x)) + v(f(t)).$$

Letting $t \to x$ and using that $f(t) \to f(x)$ as $t \to x$ so that $u(t) \to 0$ and $v(f(t)) \to 0$ as $t \to x$, we get the result. □

Example 6.1.7. (i) Consider the function

$$f(x) = \begin{cases} x \sin\left(\frac{1}{x}\right) & x \neq 0 \\ 0 & x = 0. \end{cases}$$

Assuming for now that $\sin'(x) = \cos(x)$, we may apply Theorems 6.1.4 and 6.1.6 to find that, when $x \neq 0$,

$$f'(x) = \sin\left(\frac{1}{x}\right) - \frac{1}{x}\cos\left(\frac{1}{x}\right).$$

However,

$$\frac{f(t) - f(0)}{t} = \sin\left(\frac{1}{t}\right)$$

which has no limit as $t \to 0$ so $f'(0)$ does not exist.

(ii) Consider the function

$$f(x) = \begin{cases} x^2 \sin\left(\frac{1}{x}\right) & x \neq 0 \\ 0 & x = 0. \end{cases}$$

Then for $x \neq 0$ we have

$$f'(x) = 2x\sin\left(\frac{1}{x}\right) - \cos\left(\frac{1}{x}\right)$$

and at $x = 0$ we have that

$$\begin{aligned} f'(0) &= \lim_{t\to 0} \frac{f(t) - f(0)}{t} \\ &= \lim_{t\to 0} t\sin\left(\frac{1}{t}\right) \\ &= 0. \end{aligned}$$

Thus f is differentiable at every point, but f' is not continuous. Indeed, $\lim_{t\to 0} f'(t)$ does not exist due to the $\cos(1/x)$ term.

6.2 Local extrema

Definition 6.2.1. Let $f : X \to \mathbb{R}$ be a function where X is any metric space. We say that $p \in X$ is a *local maximum* if there exists an open neighborhood U of p such that $f(p) \geq f(q)$ for all $q \in U$.

Similarly, we define what it means for $p \in X$ to be *local minimum* of f. We say that $p \in X$ is a *local extremum* of f if it is either a local minimum or a local maximum of f.

Proposition 6.2.2. *Let $f : [a, b] \to \mathbb{R}$ and $x \in (a, b)$ be a local extremum of f. If f is differentiable at x, then $f'(x) = 0$.*

Proof. We will only prove the case where x is a local maximum, as the proof for local minima is similar. For this, let

$$\varphi(t) = \frac{f(t) - f(x)}{t - x}.$$

Since $f(t) \le f(x)$ whenever t is near x, we have that $\varphi(t) \le 0$ when $t < x$ and $\varphi(t) \ge 0$ when $t > x$. Thus $\varphi(x-) \le 0$ and $\varphi(x+) \ge 0$. Hence

$$f'(x) = \lim_{t \to x} \varphi(t) = \varphi(x-) = \varphi(x+)$$

must be zero. □

Remark 6.2.3. This gives a strategy for maximizing and minimizing a differentiable function $f : [a, b] \to \mathbb{R}$. By Theorems 5.3.3 and 6.1.2 we know a maximum and minimum are achieved. By Proposition 6.2.2, these extrema must occur either a the endpoints a, b or where $f'(x) = 0$. In most cases this gives a finite set of points to check.

6.3 The mean value theorem

Theorem 6.3.1. *Let $f, g : [a, b] \to \mathbb{R}$ be continuous functions which are differentiable on (a, b). Then there exists a point $x \in (a, b)$ such that*

$$[f(b) - f(a)] \cdot g'(x) = [g(b) - g(a)] \cdot f'(x).$$

Proof. Set

$$h(t) = [f(b) - f(a)] \cdot g(t) - [g(b) - g(a)] \cdot f(t).$$

Then h is continuous on $[a, b]$ and differentiable on (a, b). Moreover, we have that $h(a) = h(b)$. To prove the theorem it suffices to find $x \in (a, b)$ with $h'(x) = 0$.

Now, by Theorem 5.3.3, we know that h achieves a minimum and maximum. If there are no extrema in (a, b), then since $h(a) = h(b)$ we must have that h is constant on $[a, b]$. In this case, $h'(x) = 0$ for all $x \in [a, b]$ and we are done. Otherwise, h has an extrema at $x \in (a, b)$ and so by Proposition 6.2.2 we have that $h'(x) = 0$. □

Theorem 6.3.2 (Mean Value Theorem). *Let $f : [a, b] \to \mathbb{R}$ be continuous on $[a, b]$ and differentiable on (a, b). Then there exists $x \in (a, b)$ such that*

$$f'(x) = \frac{f(b) - f(a)}{b - a}.$$

Proof. Take $g(x) = x$ in Theorem 6.3.1. □

6.4 Continuity of the derivative

We saw in Example 6.1.7 that there exists differentiable functions $f : I \to \mathbb{R}$ such that the derivative f' is not continuous. However, the derivative does satisfy some important properties enjoyed by continuous functions.

Theorem 6.4.1. *Let $f : [a, b] \to \mathbb{R}$ be differentiable and suppose that $f'(a) < \lambda < f'(b)$ (or $f'(b) < \lambda < f'(a)$). Then there exists $x \in (a, b)$ such that $f'(x) = \lambda$.*

Proof. Assume that $f'(a) < \lambda < f'(b)$ as the proof when $f'(a) > f'(b)$ is similar.

Take $g(t) = f(t) - \lambda t$. Then we have that $g'(a) < 0$ so by the definition of the derivative, there exists $t_1 \in (a, b)$ with $g(t_1) < g(a)$. Similarly, $g'(b) > 0$ so there exists $t_2 \in (a, b)$ with $g(t_2) < g(b)$. In particular, a, b are not minima of g. But g must attain a minimum on $[a, b]$ by Theorem 5.3.3, so this minimum must be at some $x \in (a, b)$. Then by Proposition 6.2.2 we have $g'(x) = 0$ so $f'(x) = \lambda$ as required. □

The same conclusion holds for continuous functions by Corollary 5.3.8, so this is saying that derivatives satisfy the same intermediate value property.

Corollary 6.4.2. *Let $f : [a, b] \to \mathbb{R}$ be a differentiable function. Then f' cannot have any discontinuities of the first kind.*

Proof. Suppose that x were a discontinuity of the first kind of f'. Then we would be able to find an interval containing x whose image under f' is not an interval, contradicting Theorem 6.4.1. □

Example 6.4.3. The function

$$f(x) = \begin{cases} x & x < 0 \\ 1 + x & x \geq 0 \end{cases}$$

cannot be the derivative of any function.

6.5 L'Hôpital's Theorem

To state L'Hôpital's theorem in proper generality, we need a notion of "limits at infinity."

Definition 6.5.1. Let $f : X \to \mathbb{R}$ be a function for X a metric space. Given $a \in X$ we write $\lim_{x\to a} f(x) = \infty$ if for all $M \in \mathbb{R}$ there exists $\delta > 0$ such that

$$0 < |x - a| < \delta \implies f(x) \geq M.$$

Similarly, we write $\lim_{x\to a} f(x) = -\infty$ if for all $M \in \mathbb{R}$ there exists $\delta > 0$ such that

$$0 < |x-a| < \delta \Longrightarrow f(x) \leq M.$$

Definition 6.5.2. Let $f : I \to X$ be a function where $I \subseteq \mathbb{R}$ and $(a, \infty) \subseteq I$ for some $a \in \mathbb{R}$. Then we write $\lim_{x\to\infty} f(x) = a$ if for all $\varepsilon > 0$ there exists $M \in \mathbb{R}$ such that

$$x > M \Longrightarrow d_X(f(x), a) < \varepsilon.$$

Similarly, we define $\lim_{x\to-\infty} f(x) = a$.

We also may write expressions such as $\lim_{x\to\infty} f(x) = \infty$ and $\lim_{x\to\infty} f(x) = -\infty$ when $f : \mathbb{R} \to \mathbb{R}$ which we leave to the reader to define rigorously.

Theorem 6.5.3 (L'Hôpital's Theorem)**.** *Let $f, g : (a, b) \to \mathbb{R}$ be differentiable functions with $g'(x) \neq 0$ for all $x \in (a, b)$ where $-\infty \leq a < b \leq \infty$. Then if*

$$\lim_{x\to a} \frac{f'(x)}{g'(x)} = A$$

and we are in one of the following situations:

(i) $\lim_{x\to a} f(x) = \lim_{x\to a} g(x) = 0$

(ii) $\lim_{x\to a} g(x) = \pm\infty$

then

$$\lim_{x\to a} \frac{f(x)}{g(x)} = A.$$

Remark 6.5.4. The same result holds if we replace all limits as $x \to a$ with limits as $x \to b$.

Proof. In case (ii) we will assume that $\lim_{x\to a} g(x) = \infty$. The proof for the case $\lim_{x\to a} g(x) = -\infty$ is similar.

First assume that $-\infty \leq A < \infty$ and let $A < M' < M$. Then as $f'(x)/g'(x) \to A$ we may find some $c \in (a, b)$ such that

$$a < x < c \Longrightarrow \frac{f'(x)}{g'(x)} < M'.$$

Then given any $a < x < y < c$ by the generalized mean value theorem (Theorem 6.3.1) we may find $t \in (x, y)$ such that

$$\frac{f(x) - f(y)}{g(x) - g(y)} = \frac{f'(t)}{g'(t)} < M' < M. \tag{6.5.1}$$

If case (i) holds, then letting $x \to a$ and keeping y fixed we see that

$$a < y < c \Longrightarrow \frac{f(y)}{g(y)} < M. \tag{6.5.2}$$

If case (ii) holds, keep y fixed we may find $c_1 \in (a, y)$ such that for $a < x < c_1$ we have that $g(x) > g(y)$ and $g(x) > 0$. Then multiplying (6.5.1) by $[g(x) - g(y)]/g(x)$ we get that

$$\frac{f(x)}{g(x)} < M' - M'\frac{g(y)}{g(x)} + \frac{f(y)}{g(x)}.$$

Since the right hand side goes to M' as $x \to a$, we may find $c_2 \in (a, c_1)$ such that

$$a < x < c_2 \Longrightarrow \frac{f(x)}{g(x)} < M. \tag{6.5.3}$$

Combining (6.5.2) and (6.5.3) we see that for any $A < M$ we may find $c \in (a, b)$ such that

$$a < x < c \Longrightarrow \frac{f(x)}{g(x)} < M.$$

Similarly, when $-\infty < A \leq \infty$, given $M < A$ we may find $c \in (a, b)$ such that

$$a < x < c \Longrightarrow M < \frac{f(x)}{g(x)}.$$

Combining these two results the claim follows. □

A more down to earth proof may be given for the slightly less general statement: Suppose we have $f, g : (a, b) \to \mathbb{R}$ differentiable functions and $c \in (a, b)$ such that $f(c) = g(c) = 0$ and $g'(c) \neq 0$. Then $f(x)/g(x) \to f'(c)/g'(c)$ as $x \to c$.

To do this, by the definition of the derivative, we may write

$$\begin{aligned} f(x) &= f(c) + (x - c) \cdot (f'(c) + u(x)) = (x - c) \cdot (f'(c) + u(x)) \\ g(x) &= g(c) + (x - c) \cdot (g'(c) + v(x)) = (x - c) \cdot (g'(c) + v(x)) \end{aligned}$$

where $u, v \to 0$ as $x \to c$. It follows that

$$\begin{aligned} \lim_{x \to c} \frac{f(x)}{g(x)} &= \lim_{x \to c} \frac{f'(c) + u(x)}{g'(c) + v(x)} \\ &= \frac{f'(c)}{g'(c)} \end{aligned}$$

as required.

The hard part of Theorem 6.5.3 is extending the above argument to work when c is instead one of the limit points a or b.

Example 6.5.5. Taking for granted at the moment that the derivative of $\ln(x)$ is $1/x$, we find that

$$\lim_{x\to\infty} \frac{\ln(x)}{x} = \lim_{x\to\infty} \frac{1/x}{1} = 0.$$

6.6 Taylor's theorem

Let $f : I \to \mathbb{R}$. We saw in previous sections that we may consider the derivative f' of f where it is defined. Having taken f', we may then consider its own derivative $(f')'$ (where it exists) which we denote by f'' or $f^{(2)}$ and call the *second derivative of* f. Continuing this process of taking derivatives where possible, we get a sequence

$$f, f', f'', f''', \ldots$$

which we refer to as the *higher derivative of* f. We denote the function obtained from f by taking the derivative n times by $f^{(n)}$.

Note that in order for $f^{(n)}$ to be defined at a point x, $f^{(n-1)}$ must be differentiable at x which requires $f^{(n-1)}$ to be defined on a (possibly one-sided) neighborhood of x. This in turn requires $f^{(n-2)}$ to be differentiable on a neighborhood of x.

Definition 6.6.1. Let $f : I \to \mathbb{R}$. We say that f is C^k if $f^{(k)}$ exists on I and is continuous.

We will write $C^k(I)$ to denote the set of all C^k functions $I \to \mathbb{R}$.

Example 6.6.2. C^0 functions are continuous functions, C^1 functions are differentiable functions with continuous derivatives, etc... Since differentiable functions are continuous, we have that $C^k(I) \subseteq C^{k+1}(I)$.

Example 6.6.3. Let $k \geq 2$. The function

$$f(x) = \begin{cases} 0 & x = 0 \\ x^k \sin(1/x) & x \neq 0 \end{cases}$$

is $k-1$ times differentiable, but $f^{(k-1)}$ is discontinuous at 0. Thus f is C^{k-2}.

Theorem 6.6.4 (Taylor's Theorem)**.** *Let $f : [a, b] \to \mathbb{R}$ and $n \in \mathbb{N}$. Suppose that $f^{(n-1)}$ exists and is* *continuous on* $[a, b]$ *and* $f^{(n)}$ *exists on* (a, b)*. For any* $\alpha \in [a, b]$ *define*

$$P(t) = \sum_{k=0}^{n-1} \frac{f^{(k)}(\alpha)}{k!}(t - \alpha)^k.$$

Given any $\beta \in [a, b]$*,* $\beta \neq \alpha$*, there exists* $\xi \in [a, b]$ *between* α *and* β *such that*

$$f(\beta) = P(\beta) + \frac{f^{(n)}(\xi)}{n!} \cdot (\beta - \alpha)^n.$$

Proof. Since $\beta \neq \alpha$, let M be the number satisfying

$$f(\beta) = P(\beta) + M \cdot (\beta - \alpha)^n$$

and define $g : [a, b] \to \mathbb{R}$ via

$$g(t) = f(t) - P(t) - M(t - \alpha)^n.$$

Since $P(t)$ is a polynomial of degree $n - 1$, we have that $P^{(n)}(t) = 0$. Thus

$$g^{(n)}(t) = f^{(n)}(t) - n! \cdot M$$

on (a, b) and so we are done if we may show that $g^{(n)}(\xi) = 0$ for some ξ between α and β.

Since $P^{(k)}(\alpha) = f^{(k)}(\alpha)$ for all $0 \leq k \leq n - 1$ we have that

$$g(\alpha) = g^{(1)}(\alpha) = \cdots = g^{(n-1)}(\alpha) = 0. \tag{6.6.1}$$

Now, by our choice of M we also have that $g(\beta) = 0$. Thus by the Mean Value Theorem we may find ξ_1 between α and β such that $g'(\xi_1) = 0$. But then by (6.6.1) and the Mean Value Theorem, we may find ξ_2 between ξ_1 and α such that $g^{(2)}(\xi_2) = 0$. Continuing like this, we find for all $1 \leq k \leq n$ some ξ_k (living between α and β) such that $g^{(k)}(\xi_k) = 0$. Taking $\xi = \xi_n$ we are done. □

Definition 6.6.5. Let $f : [a, b] \to \mathbb{R}$ and let P be as in the statement of Theorem 6.6.4. We call P the *n Taylor polynomial of f (centered at α)* and denote it by $P_n(x)$.

Taylor's theorem gives us a way for determining, in some cases, that the Taylor polynomials approach our original function f as the next example shows.

Example 6.6.6. Suppose there exists a differentiable function $f : \mathbb{R} \to \mathbb{R}$ satisfying

$$f' = f \tag{6.6.2}$$

$$f(0) = 1. \tag{6.6.3}$$

Inductively applying (6.6.2) we find that

$$f^{(n)} = f$$

for all n and thus the Taylor polynomials of f centered at 0 are given by

$$P_n(x) = \sum_{k=0}^{n-1} \frac{f^{(k)}(0)}{k!} x^k = \sum_{k=0}^{n-1} \frac{f(0)}{k!} x^k = \sum_{k=0}^{n-1} \frac{x^k}{k!}.$$

Now, fix $x \in \mathbb{R}$. By Taylor's theorem, for we may find $\xi \in \mathbb{R}$ with $|\xi| < |x|$ such that

$$|f(x) - P_n(x)| = \frac{|f^{(n)}(\xi)|}{n!}|x|^n = \frac{|f(\xi)|}{n!}|x|^n. \tag{6.6.4}$$

Since f is differentiable it is continuous, so set

$$M = \sup_{|\xi| \leq |x|} |f(\xi)| < \infty.$$

Then (6.6.4) gives that

$$|f(x) - P_n(x)| \leq M \cdot \frac{|x|^n}{n!}.$$

Letting $n \to \infty$ and noting that $|x|^n/n! \to 0$, we find that $P_n(x) \to f(x)$. Thus we must have that

$$f(x) = \lim_{n\to\infty} P_n(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!} \tag{6.6.5}$$

and hence f is uniquely determined.

Note that we do not yet have the tools to differentiate infinite series, so we cannot yet confirm that the function f given by (6.6.5) satisfies $f' = f$.

Example 6.6.7. Even when the Taylor polynomials converge, they may not converge to the value of the original function. Consider the function f defined by

$$f(x) = \begin{cases} 0 & x \leq 0 \\ e^{-1/x} & x > 0. \end{cases}$$

Assuming that $\frac{\mathrm{d}}{\mathrm{d}x}e^x = e^x$ and other basic properties of e^x, one may show that f is infinitely differentiable with

$$f^{(n)}(0) = 0$$

for all $n \geq 0$. Thus the Taylor polynomials of f centered at 0 are given by $P_n(x) = 0$ for all n. It follows that $(P_n(x))_n$ always converges but not to $f(x)$ unless $x \leq 0$.

Exercises

Exercise 6.1. Let $f : I \to \mathbb{R}$, $I \subseteq \mathbb{R}$ an interval, be a differentiable function with $|f'(x)| \leq M$ for all $x \in I$.

(i) Show that f is M-Lipschitz, i.e. that $|f(x) - f(y)| \leq M|x - y|$ for all $x, y \in I$.

(ii) Deduce that f is uniformly continuous on I.

Exercise 6.2. Deduce as a special case of Exercise 6.1 that if $f : I \to \mathbb{R}$ is a differentiable function with $f'(x) = 0$, then f is constant.

Exercise 6.3. Let $f : \mathbb{R} \to \mathbb{R}$ be a differentiable function with $|f'(x)| \leq A$ for some constant $A < 1$. Then, given any $a \in \mathbb{R}$, show that the sequence $(x_n)_n$ defined by

$$x_{n+1} = f(x_n), \quad x_0 = a$$

converges to some $x \in \mathbb{R}$. Moreover, show that this x satisfies $f(x) = x$.

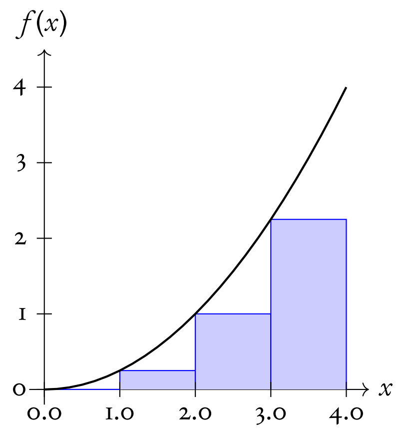
7 Integration

7.1 The Riemann integral

When one first encounters the Riemann integral, they often define it via Riemann sums as

$$\int_a^b f(x)\mathrm{d}x = \lim_{n\to\infty} \sum_{k=0}^{n-1} \frac{b-a}{n} \cdot f\left(a + (b-a) \cdot \frac{k}{n}\right).$$

This corresponds to partitioning the interval $[a, b]$ into n even pieces and approximating the area under f using rectangles, then letting the number of pieces in our partition go to infinity.



This approach is useful in practice for computations, but it is too rigid of a theoretical definition to be useful for proving theorems. To have a more workable definition, we need to consider *all* partitions, not just evenly spaced ones, and we need to eliminate the choices of where we sample our function on each piece.

Definition 7.1.1. Let $[a, b] \subseteq \mathbb{R}$ be a finite interval. A *partition* is a finite sequence of points $x_0, \ldots, x_n$ such that

$$a = x_0 \leq x_1 \leq \cdots \leq x_n = b.$$

Given a partition $P = \{x_0, \ldots, x_n\}$ as above, we write $\Delta x_i = x_i - x_{i-1}$ for $i = 1, \ldots, n$. We define the *mesh of P*, denoted $\text{mesh}(P)$, to be

$$\text{mesh}(P) = \max_i \Delta x_i.$$

Let $f : [a, b] \to \mathbb{R}$ be a bounded function and P a partition of $[a, b]$. We define

$$U(P, f) = \sum_{i=1}^{n} \Delta x_i \cdot \sup_{x \in [x_{i-1}, x_i]} f(x)$$
$$L(P, f) = \sum_{i=1}^{n} \Delta x_i \cdot \inf_{x \in [x_{i-1}, x_i]} f(x)$$

and then set

$$\overline{\int_a^b} f(x) \mathrm{d}x = \inf\{U(P, f) : P \text{ a partition of } [a, b]\}$$
$$\underline{\int_a^b} f(x) \mathrm{d}x = \sup\{L(P, f) : P \text{ a partition of } [a, b]\}.$$

These are referred to as the *upper* and *lower Riemann integrals of f* respectively.

If the upper and lower Riemann integrals of f happen to be equal, we say that f is *(Riemann) integrable* and denote the common value by

$$\int_a^b f\,\mathrm{d}x.$$

Remark 7.1.2. You should think of $U(P,f)$ and $L(P,f)$ as over- and under-estimates respectively of the area under f on the interval $[a,b]$. As the partition gets finer, for nice functions we expect the extent to which these are over and under estimates to decrease. This is why the upper integral of f is defined using and infimum while the lower integral is defined using a supremum.

We ought first make sure that the lower and upper Riemann integrals are well-defined. For this, suppose that $m \leq f \leq M$ on $[a,b]$, since f was assumed to be bounded. Then

$$m(b-a) \leq L(P,f) \leq U(P,f) \leq M(b-a)$$

so the lower and upper Riemann integrals are well-defined and finite.

Example 7.1.3. Consider the function $f(x) = x^2$ and let $a > 0$. For each n consider the partition $P_n = \{0, a/n, 2a/n, \dots, a\}$. Since f is monotone increasing, we have that

$$\begin{aligned} L(P_n,f) &= \sum_{k=0}^{n-1} \left(\frac{ak}{n}\right)^2 \cdot \frac{a}{n} \\ &= \frac{a^3}{n^3} \sum_{k=0}^{n-1} k^2 \\ &= \frac{a^3}{n^3} \cdot \frac{n(n-1)(2n-1)}{6} \end{aligned}$$

and

$$\begin{aligned} U(P_n,f) &= \sum_{k=1}^{n} \left(\frac{ak}{n}\right)^2 \cdot \frac{a}{n} \\ &= \frac{a^3}{n^3} \sum_{k=1}^{n} k^2 \\ &= \frac{a^3}{n^3} \cdot \frac{n(n+1)(2n+1)}{6}. \end{aligned}$$

Then we have that

$$\lim_{n\to\infty} U(P_n,f) = \lim_{n\to\infty} L(P_n,f) = \frac{a^3}{3}.$$

It follows that

$$\frac{a^3}{3} \leq \underline{\int_0^a} x^2 \mathrm{d}x \leq \overline{\int_0^a} x^2 \mathrm{d}x \leq \frac{a^3}{3}$$

so

$$\int_0^a x^2 \mathrm{d}x = \frac{a^3}{3}.$$

Definition 7.1.4. Let P and P^* be partitions. We say that P^* is a *refinement* of P if $P \subseteq P^*$.

Proposition 7.1.5. *Let P^* be a refinement of P. Then $U(P^*, f) \leq U(P, f)$ and $L(P^*, f) \geq L(P, f)$.*

Proof. By induction, we may assume that P^* contains one more point than P. Let $P = \{x_0, \dots, x_n\}$ and $P^* = P \cup \{y\}$ where $x_i < y < x_{i+1}$. Then we have that

$$\begin{aligned} & U(P^*, f) - U(P, f) \\ = \; & (y - x_i) \cdot \sup_{x \in [x_i, y]} f(x) + (x_{i+1} - y) \cdot \sup_{x \in [y, x_{i+1}]} f(x) - (x_{i+1} - x_i) \cdot \sup_{x \in [x_i, x_{i+1}]} f(x). \end{aligned}$$

Using that

$$\sup_{x \in [x_i, y]} f(x) \leq \sup_{x \in [x_i, x_{i+1}]} f(x)$$
$$\sup_{x \in [y, x_{i+1}]} f(x) \leq \sup_{x \in [x_i, x_{i+1}]} f(x)$$

we have that

$$\begin{aligned} & U(P^*, f) - U(P, f) \\ \geq \; & [(y - x_i) + (x_{i+1} - y) - (x_{i+1} - x_i)] \cdot \sup_{x \in [x_i, x_{i+1}]} f(x) \\ = \; & 0 \end{aligned}$$

as required.

The proof for $L(-, f)$ is similar. □

Theorem 7.1.6. *For all bounded $f : [a, b] \to \mathbb{R}$, we have that $\underline{\int_a^b} f \mathrm{d}x \leq \overline{\int_a^b} f \mathrm{d}x$.*

Proof. Let P_1, P_2 be any two partitions and set $P^* = P_1 \cup P_2$ which is a refinement of both P_1 and P_2. Then by Proposition 7.1.5 we have that

$$L(P_1, f) \leq L(P^*, f) \leq U(P^*, f) \leq U(P_2, f).$$

Keeping P_1 fixed and taking the infimum over P_2 we get that

$$L(P_1, f) \leq \overline{\int_a^b} f \, \mathrm{d}x.$$

Now taking the supremum over all P_1 the result follows. □

Theorem 7.1.7. *A bounded function $f : [a, b] \to \mathbb{R}$ is integrable if and only if for all $\varepsilon > 0$ there exists a partition P of $[a, b]$ such that $U(P, f) - L(P, f) < \varepsilon$.*

Proof. Suppose that the assumption holds. Fix $\varepsilon > 0$ and let P be such that $U(P, f) - L(P, f) < \varepsilon$. Then we have

$$L(P, f) \leq \underline{\int_a^b} f \, \mathrm{d}x \leq \overline{\int_a^b} f \, \mathrm{d}x \leq U(P, f)$$

so

$$0 \leq \overline{\int_a^b} f \, \mathrm{d}x - \underline{\int_a^b} f \, \mathrm{d}x \leq U(P, f) - L(P, f) < \varepsilon.$$

Since ε was arbitrary, we find that $\overline{\int_a^b} f \, \mathrm{d}x = \underline{\int_a^b} f \, \mathrm{d}x$ as required.

Now suppose that f is integrable and fix $\varepsilon > 0$. We may then find partitions P_1, P_2 such that

$$0 \leq U(P_1, f) - \int_a^b f \, \mathrm{d}x < \varepsilon/2$$
$$0 \leq \int_a^b f \, \mathrm{d}x - L(P_2, f) < \varepsilon/2.$$

Thus letting $P = P_1 \cup P_2$ we have

$$0 \leq U(P, f) - L(P, f) \leq U(P_1, f) - L(P_2, f) < \varepsilon$$

as required. □

We take account of some related facts before moving on.

Proposition 7.1.8. *Consider the property of a partition P that*

$$U(P, f) - L(P, f) < \varepsilon. \tag{7.1.1}$$

(i) If (7.1.1) *is satisfied for a partition P, then it is satisfied for all refinements of P.*

(ii) If (7.1.1) *holds for* $P = \{x_0, \ldots, x_n\}$ *and* $s_i, t_i \in [x_{i-1}, x_i]$ *then*

$$\sum_{i=1}^{n} |f(s_i) - f(t_i)| \Delta x_i < \varepsilon.$$

(iii) If f is integrable and the hypotheses of (ii) hold, then

$$\left| \sum_{i=1}^{n} f(t_i) \Delta x_i - \int_a^b f \, \mathrm{d}x \right| < \varepsilon.$$

Proof. (i) follows from Proposition 7.1.5. For (ii), notice that

$$|f(s_i) - f(t_i)| \leq \sup_{x \in [x_{i-1}, x_i]} f(x) - \inf_{x \in [x_{i-1}, x_i]} f(x)$$

so

$$\sum_{i=1}^{n} |f(s_i) - f(t_i)| \Delta x_i \leq U(P, f) - L(P, f) < \varepsilon.$$

For (iii), we have that

$$L(P, f) \leq \sum_{i=1}^{n} f(t_i) \Delta x_i \leq U(P, f)$$
$$L(P, f) \leq \int_a^b f \, \mathrm{d}x \leq U(P, f)$$

from which the result follows. □

Proposition 7.1.8(iii) gives us the computational freedom to compute the integral of many integrable functions using left and right Riemann sums. However, the cost is that we must be able to show our function f is integrable with respect to our definition. For this, we have the following useful theorems.

Theorem 7.1.9. *Any continuous function* $f : [a, b] \to \mathbb{R}$ *is integrable.*

In fact, for every $\varepsilon > 0$ *there exists a* $\delta > 0$ *such that* $\mathrm{mesh}(P) < \delta$ *implies* $U(P, f) - L(P, f) < \varepsilon$.

Proof. The second claim implies the first via Theorem 7.1.7. Fix $\varepsilon > 0$. Since $[a, b]$ is compact, f is uniformly continuous by Theorem 5.5.4. Thus we may find some $\delta > 0$ such that

$$|x - y| < \delta \implies |f(x) - f(y)| < \frac{\varepsilon}{b - a}.$$

Now let $P = \{x_0, \ldots, x_n\}$ be any partition of $[a, b]$ such that $\mathrm{mesh}(P) < \delta$. By Theorem 5.3.3, for each i, we may find $s_i, t_i \in [x_{i-1}, x_i]$ such that

$$f(s_i) = \sup_{x \in [x_{i-1}, x_i]} f(x)$$
$$f(t_i) = \inf_{x \in [x_{i-1}, x_i]} f(x).$$

Since $\mathrm{mesh}(P) < \delta$, we have that $|s_i - t_i| < \delta$ so

$$f(s_i) - f(t_i) < \frac{\varepsilon}{b-a}.$$

Thus

$$\begin{aligned} U(P, f) - L(P, f) &= \sum_{i=1}^{n} (f(s_i) - f(t_i)) \Delta x_i \\ &< \frac{\varepsilon}{b-a} \sum_{i=1}^{n} \Delta x_i \\ &= \varepsilon \end{aligned}$$

as required. □

Corollary 7.1.10. *Let $f : [a, b] \to \mathbb{R}$ be continuous, then*

$$\int_a^b f \, \mathrm{d}x = \lim_{n \to \infty} \frac{b-a}{n} \sum_{k=0}^{n-1} f\left(a + (b-a) \cdot \frac{k}{n}\right),$$

i.e. $\int_a^b f \, \mathrm{d}x$ may be computed using the left Riemann sum. Similarly, f may be computed with the right Riemann sum, midpoint sums, etc...

Proof. By Theorem 7.1.9 we know that f is integrable, and that partitions with sufficiently small mesh satisfy (7.1.1). Thus we may compute the integral by taking evenly sized partitions of decreasing width, and may sample any point in each sub-interval by Proposition 7.1.8(iii). □

Theorem 7.1.11. *Monotone functions $f : [a, b] \to \mathbb{R}$ are integrable.*

In fact, for every $\varepsilon > 0$ there exists a $\delta > 0$ such that $\mathrm{mesh}(P) < \delta$ implies $U(P, f) - L(P, f) < \varepsilon$.

Proof. For simplicity, assume that f is monotone increasing. Let P be a partition. Since f is monotone increasing, we have that

$$\sup_{[x_{i-1},x_i]} f(x) = f(x_i)$$

$$\sup_{[x_{i-1},x_i]} f(x) = f(x_{i-1})$$

so

$$\begin{aligned} U(P_n,f) - L(P_n,f) &= \sum_{k=1}^{n}(f(x_k) - f(x_{k-1}))\Delta x_i \\ &\leq \operatorname{mesh}(P)\sum_{k=1}^{n}(f(x_k) - f(x_{k-1})) \\ &= \operatorname{mesh}(P) \cdot (f(b) - f(a)) \end{aligned}$$

Making mesh(P) sufficiently small, we have that $U(P_n,f) - L(P_n,f) < \varepsilon$. □

Corollary 7.1.12. *Integrals of monotone functions may be computed using left Riemann sums, midpoint sums, etc...*

In fact, we may do slightly better than Theorem 7.1.9.

Theorem 7.1.13. *A bounded function $f : [a,b] \to \mathbb{R}$ with finitely many discontinuities is integrable.*

Proof. Suppose that $m \leq f \leq M$ and let $y_1, \dots, y_\ell$ be the discontinuity points of f. Choose disjoint open intervals (v_i, u_i) each of length $< \varepsilon$ such that $y_i \in (u_i, v_i)$.

Then f is continuous on

$$K = [a,b] \setminus \bigcup_{i=1}^{n}(v_i, u_i)$$

which is a compact set. Hence f is uniformly continuous on K so we may find a $\delta > 0$ such that if $|t - s| < \delta$, $t, s \in K$, then $|f(t) - f(s)| < \varepsilon$. Now choose a partition P as follows: every u_i and v_i is in P, and all other endpoints are chosen so that $\Delta x_i < \delta$ provided $x_i \neq v_j$ for some j. Then just as in the proof of Theorem 7.1.9 we have that

$$\begin{aligned} U(P,f) - L(P,f) &\leq \sum_{\substack{k=1 \\ x_k \neq v_j}}^{n}(\sup_{[x_{k-1},x_k]} f(x) - \inf_{[x_{k-1},x_k]} f(x))\Delta x_k + \sum_{k=1}^{\ell}(M-m)\Delta v_k \\ &< \varepsilon \sum_{\substack{k=1 \\ x_k \neq v_j}}^{n} \Delta x_k + \ell(M-m)\varepsilon \\ &< \varepsilon \cdot (b - a + \ell(M-m)) \end{aligned}$$

and the right hand quantity may be made arbitrarily small. Thus f is integrable by Theorem 7.1.7. $\square$

Remark 7.1.14. In fact, a function with only countably many discontinuities is also integrable. It is possible to fully classify which functions are Riemann integrable and the answer turns out to be those whose discontinuity set is "sparse enough" in some precise sense. See Lebesgue's criterion for Riemann integrability.

Theorem 7.1.15. *Let $f : [a, b] \to \mathbb{R}$, $m \le f \le M$, be integrable and $\phi : [m, M] \to \mathbb{R}$ be continuous. Then $\phi \circ f : [a, b] \to \mathbb{R}$ is integrable.*

Proof. Fix $\varepsilon > 0$. Since $[m, M]$ is compact, ϕ is uniformly continuous and we may find $\delta > 0$ such that $|s - t| \le \delta$ implies $|\phi(s) - \phi(t)| < \varepsilon$.

Now, since f is integrable we may find a partition $P = \{x_0, \dots, x_n\}$ with

$$U(P, f) - L(P, f) < \delta^2.$$

Let

$$M_i = \sup_{[x_{i-1}, x_i]} f(x), \qquad m_i = \inf_{[x_{i-1}, x_i]} f(x)$$

and define the analogous quantities M_i^* and m_i^* for $\phi \circ f$. We now break the indices $i = 1, \dots, n$ into two cases. Let A be the set of those i such that $M_i - m_i \le \delta$ and let B be the remaining i.

By construction of δ, if $i \in A$ then $M_i^* - m_i^* < \varepsilon$. Now, we have that

$$\delta^2 > U(P, f) - L(P, f) > \delta \sum_{i \in B} \Delta x_i$$

so

$$\sum_{i \in B} \Delta x_i < \delta.$$

Let $K = \sup |\phi(x)|$. It follows that

$$\begin{aligned} U(P, \phi \circ f) - L(P, \phi \circ f) &= \sum_{i \in A} (M_i^* - m_i^*) \Delta x_i + \sum_{i \in B} (M_i^* - m_i^*) \Delta x_i \\ &< \varepsilon \cdot (b - a) + 2K \cdot \delta. \end{aligned}$$

Since ε was arbitrary and we are free to shrink δ, we see that $\phi \circ f$ is integrable by Theorem 7.1.7. $\square$

7.2 Some properties of the integral

Theorem 7.2.1. *Let $f, g : [a, b] \to \mathbb{R}$ be integrable functions.*

(i) For all $c, d \in \mathbb{R}$, $cf + dg$ is integrable and

$$\int_a^b (cf + dg)\mathrm{d}x = c\int_a^b f\mathrm{d}x + d\int_a^b g\mathrm{d}x.$$

(ii) If $f \leq g$ on $[a, b]$ then

$$\int_a^b f\mathrm{d}x \leq \int_a^b g\mathrm{d}x.$$

(iii) If $a < c < b$ then $f|_{[a,c]}$ and $f|_{[c,d]}$ are both integrable and

$$\int_a^b f\mathrm{d}x = \int_a^c f\mathrm{d}x + \int_c^b f\mathrm{d}x.$$

(iv) If $|f(x)| \leq M$ on $[a, b]$ then

$$\left|\int_a^b f\mathrm{d}x\right| \leq M(b - a).$$

Proof. For (i), it suffices to consider the cases $c = d = 1$ and $d = 0$. The remaining cases follow from these. The case of $d = 0$ is easy from definition, so we omit it.

Thus assume that $c = d = 1$. For this, notice that for any partition P we have that

$$L(P, f) + L(P, g) \leq L(P, f + g) \leq U(P, f + g) \leq U(P, f) + U(P, g). \tag{7.2.1}$$

Let $\varepsilon > 0$. Since f and g are integrable, we may find partitions P_1 and P_2 such that

$$U(P_1, f) - L(P_1, f) < \varepsilon, \quad U(P_2, g) - L(P_2, g) < \varepsilon. \tag{7.2.2}$$

Let $P = P_1 \cup P_2$ so that (7.2.2) holds for P_1, P_2 replaced by P. Then (7.2.1) implies that

$$U(P, f + g) - L(P, f + g) < 2\varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, we see that $f + g$ is integrable.

Moreover, using the same P as above, we have that

$$U(P, h) < \int_a^b h\mathrm{d}x + \varepsilon \quad (h = f, g).$$

Thus (7.2.1) implies that

$$\int_a^b (f+g)\,\mathrm{d}x \le U(P, f+g) \le U(P, f) + U(P, g) < \int_a^b f\,\mathrm{d}x + \int_a^b g\,\mathrm{d}x + 2\varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, we see that

$$\int_a^b (f+g)\,\mathrm{d}x \le \int_a^b f\,\mathrm{d}x + \int_a^b g\,\mathrm{d}x.$$

Replacing f and g with $-f$ and $-g$ we get the reverse inequality, and so we are done with (i).

The proofs of (ii)-(iv) are omitted, but may be proven with similar strategies. □

Example 7.2.2. Let $a \in \mathbb{R}$ and consider the function $\delta_a : \mathbb{R} \to \mathbb{R}$ defined by

$$\delta_a(x) = \begin{cases} 1 & x = a \\ 0 & x \neq a. \end{cases}$$

Then δ_a is integrable over any interval by Theorem 7.1.13 since it has a single discontinuity. Moreover, we can compute its integral as follows. Let $[c, d]$ be any interval containing a. Then for any $\varepsilon > 0$ we have that

$$\begin{aligned} \int_c^d \delta_a \mathrm{d}x &= \int_c^{a-\varepsilon} \delta_a \mathrm{d}x + \int_{a-\varepsilon}^{a+\varepsilon} \delta_a \mathrm{d}x + \int_{a+\varepsilon}^d \delta_a \mathrm{d}x \\ &= \int_{a-\varepsilon}^{a+\varepsilon} \delta_a \mathrm{d}x. \end{aligned} \tag{7.2.3}$$

But $|\delta_a| \le 1$ so we have that

$$\left| \int_{a-\varepsilon}^{a+\varepsilon} \delta_a \mathrm{d}x \right| \le 2\varepsilon \tag{7.2.4}$$

by Theorem 7.2.1(iv). Since $\varepsilon > 0$ was arbitrary, (7.2.3) and (7.2.4) imply that $\int_c^d \delta_a \mathrm{d}x = 0$. Since δ_a is zero away from a, this shows that the integral of δ_a over *any* interval is 0.

Corollary 7.2.3. *Let $f : [a, b] \to \mathbb{R}$ be an integrable function and let $g : [a, b] \to \mathbb{R}$ be a function which differs from f at finitely many points. Then g is integrable and*

$$\int_a^b f\,\mathrm{d}x = \int_a^b g\,\mathrm{d}x.$$

Proof. Let $y_1, \dots, y_m$ be the points where f and g disagree. We may write

$$g = f + \sum_{i=1}^{m} c_i \delta_{y_i}$$

where $c_i = g(y_i) - f(y_i)$. Then by Example 7.2.2 and Theorem 7.2.1, we have that g is integrable and

$$\begin{aligned}\int_a^b g \, \mathrm{d}x &= \int_a^b f \, \mathrm{d}x + \sum_{i=1}^{m} c_i \int_a^b \delta_{y_i} \, \mathrm{d}x \\ &= \int_a^b f \, \mathrm{d}x\end{aligned}$$

as required. □

Example 7.2.4. Define the step function $H_a(x)$ as follows:

$$H_a(x) = \begin{cases} 1 & x \geq a \\ 0 & x < 0 \end{cases}$$

which is integrable as it is monotone. Now consider the following "staircase" function given by

$$f(x) = \sum_{n=1}^{\infty} 2^{-n} H_{1-2^{-n}}(x)$$

which is also integrable as it is monotone. Notice also that

$$0 \leq f \leq \sum_{n=1}^{\infty} 2^{-n} = 1.$$

For each $n \in \mathbb{N}$ we may write

$$\begin{aligned}\int_0^1 f \, \mathrm{d}x &= \int_0^{1-2^{-n}} f \, \mathrm{d}x + \int_{1-2^{-n}}^1 f \, \mathrm{d}x \\ &= \sum_{k=1}^{n-1} 2^{-k} \int_0^{1-2^{-n}} H_{1-2^{-k}}(x) \, \mathrm{d}x + \int_{1-2^{-n}}^1 f \, \mathrm{d}x \\ &= \sum_{k=1}^{n-1} 2^{-k}(2^{-k} - 2^{-n}) + \int_{1-2^{-n}}^1 f \, \mathrm{d}x.\end{aligned} \tag{7.2.5}$$

Since $0 \le f \le 1$ we have that

$$0 \le \int_{1-2^{-n}}^{1} f \, dx \le 2^{-n}$$

which goes to zero as $n \to \infty$. Thus, letting $n \to \infty$ in (7.2.5), we get that

$$\int_0^1 f \, dx = \lim_{n\to\infty} \sum_{k=1}^{n-1} 2^{-k}(2^{-k} - 2^{-n}).$$

Moreover,

$$0 \le 2^{-n} \sum_{k=1}^{n-1} 2^{-k} \le 2^{-n} \sum_{k=1}^{\infty} 2^{-k} = 2^{-n} \longrightarrow 0$$

so

$$\int_0^1 f \, dx = \sum_{k=1}^{\infty} 2^{-2k} = \frac{1}{3}.$$

Theorem 7.2.5. *Let $f, g : [a, b] \to \mathbb{R}$ be integrable. Then*

(i) fg is integrable

(ii) $|f|$ is integrable and

$$\left| \int_a^b f dx \right| \le \int_a^b |f| dx.$$

Proof. For (i), we have that

$$4fg = (f + g)^2 - (f - g)^2. \tag{7.2.6}$$

By Theorem 7.1.15, taking $\phi(t) = t^2$, we see that squares of integrable functions are integrable. Thus the right hand side of (7.2.6) is integrable by Theorem 7.2.1.

For (ii), we have that $\phi(x) = |x|$ is continuous, so $|f|$ is integrable by Theorem 7.1.15. Moreover, $f \le |f|$ and $-f \le |f|$ on $[a, b]$ so (ii) follows from Theorem 7.2.1(ii). □

7.3 The fundamental theorem of calculus

Theorem 7.3.1. *Let $f : [a, b] \to \mathbb{R}$ be integrable. Define $F : [a, b] \to \mathbb{R}$ via*

$$F(x) = \int_a^x f(t) \, dt.$$

Then F is continuous and if f is continuous at a point x_0, then F is differentiable at x_0 with $F'(x_0) = f(x_0)$.

Proof. Since f is integrable, it is bounded. Thus let $|f| \leq M$. Then we have that, for $x > y$,

$$|F(x) - F(y)| = \left| \int_y^x f(t)\,\mathrm{d}t \right| \leq M(x-y).$$

It follows that F is M-Lipschitz, hence continuous.

Now suppose that f is continuous at x_0. Fix $\varepsilon > 0$ and choose $\delta > 0$ such that

$$|x - x_0| < \delta \implies |f(x) - f(x_0)| < \varepsilon.$$

Then for $a \leq s \leq t \leq b$ and $x_0 - \delta < s, t < x_0 + \delta$ we have

$$\left| \frac{F(s) - F(t)}{s - t} - f(x_0) \right| = \left| \frac{1}{s-t} \int_s^t [f(u) - f(x_0)]\,\mathrm{d}u \right| < \varepsilon.$$

This implies that $F'(x_0) = f(x_0)$. □

Theorem 7.3.2 (Fundamental Theorem of Calculus)**.** *Suppose that* $f : [a, b] \to \mathbb{R}$ *is integrable and* $F : [a, b] \to \mathbb{R}$ *is such that* $F' = f$. *Then*

$$\int_a^b f\,\mathrm{d}x = F(b) - F(a).$$

Proof. Let $\varepsilon > 0$ be given. Then there exists a partition $P = \{x_0, \ldots, x_n\}$ such that $U(P, f) - L(P, f) < \varepsilon$. By the mean value theorem, we may find points $t_i \in [x_{i-1}, x_i]$ such that

$$F(x_i) - F(x_{i-1}) = f(t_i)\Delta x_i.$$

It follows that

$$\sum_{i=1}^n f(t_i)\Delta x_i = F(b) - F(a).$$

However, Proposition 7.1.8(iii) then tells us that

$$\left| F(b) - F(a) - \int_a^b f(x)\,\mathrm{d}x \right| < \varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, the result follows. □

Remark 7.3.3. The fundamental theorem of calculus tells us that integrals may be computed using so-called "anti-derivatives" of our integrand f, i.e. functions F such that $F' = f$. It also tells us that if an anti-derivative exists, it *must* be given by integrating.

On the other hand Theorem 7.3.1 tells us that if f happens to be continuous, then integrating f does indeed give an anti-derivative. However, it is not always true that an integrable function has an anti-derivative as Example 6.4.3 shows so the hypothesis that f be continuous is necessary.

Example 7.3.4. We have that

$$\frac{\mathrm{d}}{\mathrm{d}x}\frac{x^3}{3} = x^2.$$

Thus by the fundamental theorem of calculus,

$$\int_0^a x^2 \, \mathrm{d}x = \frac{a^3}{3}$$

which agrees with the computation in Example 7.1.3.

Via the fundamental theorem of calculus, for every differentiation rule we get a corresponding integration rule. The following two corollaries are consequences of the product rule and chain rule for differentiation, respectively.

Theorem 7.3.5 (Integration by parts)**.** *Let $f, g : [a, b] \to \mathbb{R}$ be two differentiable functions with integrable derivatives. Then*

$$\int_a^b f(x)g'(x) \, \mathrm{d}x = f(b)g(b) - f(a)g(a) - \int_a^b f'(x)g(x) \, \mathrm{d}x.$$

Proof. We have that

$$\frac{\mathrm{d}}{\mathrm{d}x} f(x)g(x) = f'(x)g(x) + f(x)g'(x)$$

where the right hand side is integrable as f, g are continuous and their derivatives are assumed to be integrable. Thus by the fundamental theorem of calculus we have that

$$\int_a^b [f'(x)g(x) + f(x)g'(x)] \, \mathrm{d}x = f(b)g(b) - f(a)g(a)$$

from which the result follows. □

Theorem 7.3.6 (Change of variables)**.** *Let $\phi : [a, b] \to [A, B]$ be a differentiable function with integrable derivative and let $f : [A, B] \to \mathbb{R}$ be continuous. Then*

$$\int_a^b f(\phi(x))\phi'(x) \, \mathrm{d}x = \int_{\phi(a)}^{\phi(b)} f \, \mathrm{d}x.$$

Proof. ϕ is continuous, hence integrable, so $f \circ \phi$ is integrable by Theorem 7.1.15 since f is continuous. Moreover, because f is continuous it has an anti-derivative F via Theorem 7.3.1. We then have that

$$\frac{\mathrm{d}}{\mathrm{d}x}F(\phi(x)) = f(\phi(x))\phi'(x)$$

by the chain rule, and the right hand side is integrable. Thus by the fundamental theorem of calculus we have that

$$\begin{aligned}\int_a^b f(\phi(x))\phi'(x)\,\mathrm{d}x &= F(\phi(b)) - F(\phi(b)) \\ &= \int_{\phi(a)}^{\phi(b)} f\,\mathrm{d}x\end{aligned}$$

as required. □

Example 7.3.7. Assume for now that $\frac{\mathrm{d}}{\mathrm{d}x}e^x = e^x$. Then we have that

$$\begin{aligned}\int_0^1 xe^{x^2}\,\mathrm{d}x &= \int_0^1 \frac{1}{2}e^u\,\mathrm{d}u \\ &= \frac{1}{2}[e^1 - e^0] \\ &= (e-1)/2\end{aligned}$$

where we used the substitution $u = x^2$ on the first step.

Exercises

Exercise 7.1. Consider the function

$$f(x) = \begin{cases} 1 & x = 0 \\ \frac{1}{q} & x = p/q,\ \ p, q \in \mathbb{Z},\ \ p/q \text{ in reduced form},\ \ q > 0 \\ 0 & x \in \mathbb{R} \setminus \mathbb{Q}. \end{cases}$$

Show that f is integrable on any interval and compute $\displaystyle\int_a^b f\,\mathrm{d}x$.

Exercise 7.2. For a subset $A \subseteq \mathbb{R}$ let

$$1_A(x) = \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases}$$

and set

$$g_n(x) = n \cdot 1_{(0,1/n)}(x) - (n+1) \cdot 1_{(0,1/(n+1))}(x).$$

Consider the function $f : [0,1] \to \mathbb{R}$ given by

$$f(x) = \sum_{n=1}^{\infty} g_n(x).$$

(i) Show that f is well-defined and $f(x) = 1_{(0,1)}(x)$.

(ii) Show that 1_I for I an interval is integrable.

(iii) Compute

$$\int_0^1 f \, \mathrm{d}x$$

and

$$\sum_{n=1}^{\infty} \int_0^1 g_n \, \mathrm{d}x$$

and show they are not equal. Deduce that integration does not always commute with infinite sums.

Exercise 7.3. Show that there is no differentiable function $f : \mathbb{R} \to \mathbb{R}$ with integrable derivative which satisfies $f(x)f'(x) = 1$ for all x.

Exercise 7.4. Show that if $f : [a,b] \to \mathbb{R}$ is continuous with $f \geq 0$ and $\int_a^b f \, \mathrm{d}x = 0$ then $f = 0$.

Exercise 7.5. Suppose $f : [a,b] \to \mathbb{R}$ is integrable and $f(x) = 0$ for all $x \in [a,b] \cap \mathbb{Q}$. Show that $\int_a^b f \, \mathrm{d}x = 0$.

8 Sequences of functions

8.1 Pointwise convergence

Definition 8.1.1. Let X, Y be metric spaces and $(f_n)_n$ a sequence of functions $X \to Y$ and $f : X \to Y$. We say that f_n *converges pointwise to* f, written $f_n \to f$ *pointwise*, if for all $x \in X$ we have that

$$f(x) = \lim_{n\to\infty} f_n(x).$$

In this case we will also write $\lim_{n\to\infty} f_n = f$.

We'll show through examples that the notion of pointwise convergence is a very weak one—many desirable properties such as continuity and differentiability may hold for each f_n while not holding for their pointwise limit.

Recall that a function f is continuous at a point x if $\lim_{t\to x} f(t) = f(x)$. Thus, if we assume that each f_n is continuous and $f_n \to f$ pointwise, then the question of whether f is continuous at x amounts to whether

$$\lim_{t\to x}\lim_{n\to\infty} f_n(t) = \lim_{n\to\infty}\lim_{t\to x} f_n(t),$$

that is whether the two limits commute. Similarly, if $f_n \to f$ pointwise and each f_n is differentiable, then the question of whether f is differentiable with $f_n'(x) \to f'(x)$ amounts to a question about commuting limits.

In the following examples, we show that many common operations do not commute with pointwise limits.

Example 8.1.2. Consider double indexed set of numbers

$$s_{m,n} = \frac{m}{m+n}.$$

Then we have that

$$\lim_{m\to\infty} s_{m,n} = 1$$

for all n and

$$\lim_{n\to\infty} s_{m,n} = 0$$

for all m. It follows that

$$\lim_{n\to\infty}\lim_{m\to\infty} s_{m,n} = 1 \neq 0 = \lim_{m\to\infty}\lim_{n\to\infty} s_{m,n}$$

and thus the limits do not commute in this case.

Example 8.1.3. We can promote Example 8.1.2 to witness a discontinuous pointwise limit of continuous functions. For this, consider the functions $f_m : [0,\infty) \to \mathbb{R}$ given by

$$f_m(x) = \frac{mx}{1+mx}.$$

Then each f_m is continuous and $f_m \to f$ where

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

is discontinuous at 0.

Example 8.1.4. Consider the functions $f_n : [0,1] \to \mathbb{R}$ given by

$$f_n(x) = nx(1-x^2)^{n-1}.$$

Then for $0 < x \leq 1$ we have that $|1-x^2| < 1$ so $\lim_n f_n(x) = 0$, and for $x = 0$ we have $f_n(0) = 0$. Thus $f_n \to 0$ pointwise.

However, notice that f_n has anti-derivative $-(1-x^2)^n/2$ so

$$\int_0^1 f_n(x)\,\mathrm{d}x = \frac{-(1-x^2)^n}{2}\Big|_0^1 = \frac{1}{2}.$$

Thus

$$\lim_{n\to\infty}\int_0^1 f_n(x)\,\mathrm{d}x = \frac{1}{2} \neq 0 = \int_0^1 \lim_{n\to\infty} f_n(x)\,\mathrm{d}x.$$

Example 8.1.5. Consider

$$f_n(x) = \frac{\sin(nx)}{n}.$$

Then $f_n \to 0$ pointwise but

$$f_n'(x) = \cos(nx)$$

so $(f_n'(x))_n$ does not converge for any x, and certainly not to 0.

8.2 Uniform convergence

The previous section shows that in general one must use extreme caution when swapping the order of limit operations. To be able to commute limiting operations, we need a notion of convergence that is stronger than pointwise convergence.

Definition 8.2.1. Let X, Y be metric spaces and let $(f_n)_n$ be a sequence of functions $X \to Y$ and let $f : X \to Y$. We say that f_n *converges uniformly to* f, written $f_n \to f$ *uniformly*, if for all $\varepsilon > 0$ there exists an $N \in \mathbb{N}$ such that for all $n \geq N$ and $x \in X$ we have $d_Y(f_n(x), f(x)) < \varepsilon$.

If $f_n : X \to \mathbb{C}$ are complex valued functions, then we say that $\sum_n f_n$ converges uniformly if the sequence of partial sums converges uniformly.

It is an easy check to see that if $f_n \to f$ uniformly then $f_n \to f$ pointwise. Indeed, pointwise convergence lets the N occurring in Definition 8.2.1 depend on both ε and $x \in X$, whereas uniform converges requires a single N depending only on ε which works for all $x \in X$.

Alternatively, one may formulate uniform convergence as: $f_n \to f$ uniformly if and only if for all

$\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that for all $n \geq N$ we have

$$\sup_{x \in X} d_Y(f_n(x), f(x)) < \varepsilon.$$

Said differently, we have the following proposition.

Proposition 8.2.2. *Let $f_n : X \to Y$, $n \in \mathbb{N}$, and $f : X \to Y$ be functions and set*

$$M_n = \sup_{x \in X} d_Y(f_n(x), f(x)).$$

Then $f_n \to f$ uniformly if and only if $M_n \to 0$ as $n \to \infty$.

Example 8.2.3. Let $f_n(x) = x^n$. Then $f_n \to 0$ uniformly on $[0, 1 - \delta]$ for any $\delta > 0$ but not on $[0, 1)$. Indeed, fix $\varepsilon > 0$. Then

$$\begin{aligned} |f_n(x)| < \varepsilon &\iff |x|^n < \varepsilon \\ &\iff n > \frac{\ln(\varepsilon)}{\ln(x)}. \end{aligned} \tag{8.2.1}$$

We wish to make this true for all x whenever $n \geq N$ for some sufficiently large N. If $x \in [0, 1 - \delta]$ then we may take

$$N > \frac{\ln(\varepsilon)}{\ln(1 - \delta)}$$

so that $f_n \to 0$ uniformly on $[0, 1 - \delta]$. However,

$$\lim_{x \to 1^-} \frac{\ln(\varepsilon)}{\ln(x)} = \infty$$

so (8.2.1) cannot be made to hold for all $n \geq N$ and $x \in [0, 1)$ for any fixed $N \in \mathbb{N}$. Thus $f_n \to 0$ pointwise but not uniformly on $[0, 1)$.

We also would like a uniform notion of Cauchy-ness, which we define next.

Definition 8.2.4. Let X, Y be metric spaces and let $(f_n)_n$ be a sequence of functions $X \to Y$. We say that $(f_n)_n$ is *uniformly Cauchy* if for all $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that for all $n, m \geq N$ and $x \in X$ we have $d_Y(f_n(x), f_m(x)) < \varepsilon$.

Similar to uniform convergence, $(f_n)_n$ being uniformly Cauchy means that for every $x \in X$ the sequence $(f_n(x))_n$ is Cauchy *and* the $N \in \mathbb{N}$ appearing in the definition of Cauchy-ness may be chosen independent of x.

Theorem 8.2.5. *Let X, Y be metric spaces and $(f_n)_n$ a sequence of functions $X \to Y$.*

(i) If $(f_n)_n$ is uniformly convergent, then it is uniformly Cauchy.

(ii) If $(f_n)_n$ is uniformly Cauchy and pointwise convergent, then it is uniformly convergent.

(iii) If $(f_n)_n$ is uniformly Cauchy and Y is complete, then $(f_n)_n$ is uniformly convergent.

Proof. The proof of (i) is the same as the proof that every convergent sequence is Cauchy using that the expressions appearing may be chosen uniformly to work for every point in X.

For (ii), let $f_n \to f$ pointwise. Fix $\varepsilon > 0$. Then we may find $N \in \mathbb{N}$ such that for $n, m \geq N$ and $x \in X$ we have

$$d_Y(f_n(x), f_m(x)) < \varepsilon/2.$$

Then, taking the limit as $m \to \infty$ and using that d_Y is continuous, we find that for all $n \geq N$ and $x \in X$ we have

$$d_Y(f_n(x), f(x)) \leq \varepsilon/2 < \varepsilon$$

as required.

For (iii), suppose that $(f_n)_n$ is uniformly Cauchy. Then for every $x \in X$ we have that $(f_n(x))_n$ is Cauchy and hence convergent since Y is complete. Thus we may build a function $f : X \to Y$ such that $f_n \to f$ pointwise and so we are done by (ii). $\square$

Theorem 8.2.6 (Weierstrass M-test)**.** *Let $f_n : X \to \mathbb{C}$ be a sequence of functions and suppose that*

$$\sup_{x \in X} |f_n(x)| \leq M_n < \infty$$

for each n. If $\sum_n M_n < \infty$, then $\sum_n f_n$ is uniformly convergent.

Proof. Let $(s_n)_n$ be the sequence of partial sums of $\sum_n f_n$. Then we have that for $n > m$

$$\begin{aligned} \sup_{x \in X} |s_n(x) - s_m(x)| &= \sup_{x \in X} \left| \sum_{k=m+1}^{n} f_k(x) \right| \\ &\leq \sum_{k=m+1}^{n} M_k. \end{aligned}$$

Since $\sum_n M_n$ is convergent, the right hand side can be made arbitrarily small provided n and m are sufficiently large. It follows that $(s_n)_n$ is uniformly Cauchy, hence uniformly convergent by Theorem 8.2.5(iii) since $\mathbb{C}$ is complete. $\square$

Example 8.2.7. Consider the power series

$$\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!} \tag{8.2.2}$$

which has infinite radius of convergence. This series does not converge uniformly on all of $\mathbb{R}$. Indeed, it is not uniformly Cauchy on $\mathbb{R}$. For any fixed $N \in \mathbb{N}$ and $x > 0$ we have

$$\left| \sum_{k=N}^{m} \frac{x^k}{k!} \right| \geq \frac{x^N}{N!}.$$

which may be made arbitrarily large by making x sufficiently large.

However, the next best thing does hold. The series (8.2.2) converges uniformly on $[-R, R]$ for any $R \in \mathbb{R}$. To see this, notice that

$$\sup_{x \in [-R,R]} \left| \frac{x^n}{n!} \right| = \frac{R^n}{n!}$$

and

$$\sum_{n=0}^{\infty} \frac{R^n}{n!} = \exp(R) < \infty.$$

Thus the series (8.2.2) is uniformly convergent on $[-R, R]$ by the Weierstrass M-test.

8.3 Behavior of continuity, integration and differentiation

8.3.1 Continuity

Theorem 8.3.1. *Let $f_n : X \to Y$ and $f : X \to Y$ and suppose $f_n \to f$ uniformly. Fix $x \in X$ and suppose that*

$$\lim_{t \to x} f_n(t) = A_n.$$

Then $(A_n)_n$ is Cauchy and

$$\lim_{t \to x} f(t) = \lim_{n \to \infty} A_n$$

in the sense that one limit exists if and only if the other does, in which case they are equal.

Said differently, if $(f_n)_n$ is uniformly convergent and $x \in X$, then

$$\lim_{t \to x} \lim_{n \to \infty} f_n(t) = \lim_{n \to \infty} \lim_{t \to x} f_n(t).$$

Proof. Fix $\varepsilon > 0$. Since $(f_n)_n$ uniformly convergent, it is uniformly Cauchy. Thus there exists $N \in \mathbb{N}$

such that for $n, m \geq N$ and $t \in X$ we have that

$$d_Y(f_n(t), f_m(t)) < \varepsilon/2.$$

Then taking the limit as $t \to x$ and using that d_Y is continuous, we get that

$$d_Y(A_n, A_m) \leq \varepsilon/2 < \varepsilon.$$

Thus $(A_n)_n$ is Cauchy.

Next, keep $\varepsilon > 0$ fixed. Since $f_n \to f$ uniformly and $(A_n)_n$ is Cauchy, we may find $N \in \mathbb{N}$ such that

$$d_Y(f(t), f_n(t)) < \varepsilon/3$$

for all $t \in X$ and for all $n, m \geq N$ we have

$$d_Y(A_n, A_m) < \varepsilon/3.$$

Then, for $n \geq N$,

$$\begin{aligned} d_Y(f(t), A_n) &\leq d_Y(f(t), f_N(t)) + d_Y(f_N(t), A_N) + d_Y(A_N, A_n) \\ &< \frac{2\varepsilon}{3} + d_Y(f_N(t), A_N). \end{aligned}$$

Now since $f_N(t) \to A_N$ as $t \to x$, we may find a neighborhood V of x such that $d_Y(f_N(t), A_N) < \varepsilon/3$ whenever $t \in V$. It follows that for $t \in V$ and $n \geq N$ we have

$$d_Y(f(t), A_n) < \varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, the result follows. □

Theorem 8.3.2. *Let X, Y be metric spaces and let $f_n : X \to Y$ be continuous functions with $f_n \to f$ uniformly. Then f is continuous.*

Proof. This is an immediate consequence of Theorem 8.3.1. □

The converse to this theorem does not necessarily hold. That is to say, there may exist continuous functions $f_n : X \to Y$ and $f : X \to Y$ such that $f_n \to f$ pointwise but not uniformly. Under certain restrictive situations, however, we do have a converse.

Proposition 8.3.3. *Let $f_1 \geq f_2 \geq f_3 \geq \cdots$ be a decreasing sequence of continuous function $K \to \mathbb{R}$ for K compact. If $f_n \to f$ pointwise for some continuous function $f : K \to \mathbb{R}$, then $f_n \to f$ uniformly.*

Proof. Consider $g_n = f_n - f$. Then g_n is continuous, non-negative and $g_1 \geq g_2 \geq g_3 \geq \cdots$. Moreover, it suffices to show that $g_n \to 0$ uniformly.

Fix $\varepsilon > 0$. Let

$$K_n = \{x \in K : g_n(x) \geq \varepsilon\} = g_n^{-1}([\varepsilon, \infty))$$

which is closed since g_n is continuous. Thus $K_n \subseteq K$ is compact as it is a closed subset of a compact space. Moreover, because $g_1 \geq g_2 \geq g_3 \geq \cdots$ we have that $K_1 \supseteq K_2 \supseteq K_3 \supseteq \cdots$. Finally, we must have that

$$\bigcap_{n \in \mathbb{N}} K_n = \emptyset$$

since if $x \in \bigcap_n K_n$ then $g_n(x) \geq \varepsilon$ for all n contradicting that $g_n(x) \to 0$. However, by Lemma this means that $K_N = \emptyset$ for some $N \in \mathbb{N}$. It follows that $0 \leq g_n(x) < \varepsilon$ for all $x \in X$ and $n \geq N$. Thus $g_n \to 0$ uniformly. □

Example 8.3.4. The compactness assumption in Proposition 8.3.3 is necessary. Indeed,

$$f_n(x) = \frac{1}{nx + 1}$$

decreases pointwise to 0 on $(0, 1)$ but not uniformly.

8.3.2 Integration

Theorem 8.3.5. *Suppose that $f_n : [a, b] \to \mathbb{R}$ are integrable functions and $f_n \to f$ uniformly. Then f is integrable and*

$$\int_a^b f \, dx = \lim_{n \to \infty} \int_a^b f_n \, dx.$$

Proof. Set

$$\varepsilon_n = \sup |f_n(x) - f(x)|$$

so that $\varepsilon_n \to 0$ as $f_n \to f$ uniformly. Then we have that

$$f_n - \varepsilon_n \leq f \leq f_n + \varepsilon_n \tag{8.3.1}$$

so f is bounded as the f_n are. Moreover, (8.3.1) implies that

$$\int_a^b (f_n - \varepsilon_n) \, dx \leq \underline{\int_a^b} f \, dx \leq \overline{\int_a^b} f \, dx \leq \int_a^b (f_n + \varepsilon_n) \, dx \tag{8.3.2}$$

so

$$0 \leq \overline{\int_a^b} f \, \mathrm{d}x - \underline{\int_a^b} f \, \mathrm{d}x \leq 2\varepsilon_n(b-a).$$

Letting $n \to \infty$ we see that $\underline{\int_a^b} f \, \mathrm{d}x = \overline{\int_a^b} f \, \mathrm{d}x$ so f is integrable.

Moreover, (8.3.2) also shows that

$$\left| \int_a^b f \, \mathrm{d}x - \int_a^b f_n \, \mathrm{d}x \right| \leq \varepsilon_n(b-a)$$

so letting $n \to \infty$ the result follows. □

Corollary 8.3.6. *Suppose that $f_n : [a, b] \to \mathbb{R}$ are integrable and*

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

where the series converges uniformly. Then f is integrable and

$$\int_a^b f \, \mathrm{d}x = \sum_{n=0}^{\infty} \int_a^b f_n \, \mathrm{d}x.$$

8.3.3 Differentiation

The sequence of functions

$$f_n(x) = \frac{\sin(nx)}{n}$$

that we encountered in Example 8.1.5 in fact converges to 0 uniformly. However, the sequence $(f_n')_n$ of derivatives does not converge at *any* point as we saws. Thus we need stronger hypotheses to control the derivatives.

Theorem 8.3.7. *Let $(f_n)_n$ be a sequence of differentiable functions $[a, b] \to \mathbb{R}$. Suppose that $(f_n')_n$ converges uniformly and for some $x_0 \in [a, b]$ we have that $(f_n(x_0))_n$ converges. Then $(f_n)_n$ converges uniformly to some $f : [a, b] \to \mathbb{R}$ and*

$$f'(x) = \lim_{n\to\infty} f_n'(x).$$

Proof. Fix $\varepsilon > 0$. $(f_n')_n$ is uniformly Cauchy and $(f_n(x_0))_n$ is Cauchy, so we may find $N \in \mathbb{N}$ such that for all $n, m \geq N$ we have

$$|f_n(x_0) - f_m(x_0)| < \varepsilon/2 \tag{8.3.3}$$

and

$$|f_n'(t) - f_m'(t)| < \frac{\varepsilon}{2(b-a)} \tag{8.3.4}$$

for all $t \in [a, b]$.

Applying the mean value theorem to $f_n - f_m$ we find using (8.3.4) that

$$|f_n(x) - f_m(x) - f_n(t) + f_m(t)| \leq \frac{\varepsilon|t-x|}{2(b-a)} \leq \frac{\varepsilon}{2} \tag{8.3.5}$$

for all $t, x \in [a, b]$, provided $n, m \geq N$. It follows that for all $n, m \geq N$ and $x \in [a, b]$ we have

$$\begin{aligned}|f_n(x) - f_m(x)| &\leq |f_n(x) - f_m(x) - f_n(x_0) + f_m(x_0)| + |f_n(x_0) - f_m(x_0)| \\ &< \varepsilon\end{aligned}$$

by combining (8.3.3) and (8.3.5). Thus $(f_n)_n$ is uniformly Cauchy and therefore uniformly convergent, say with $f_n \to f$.

Now fix $x \in [a, b]$ and define

$$\phi_n(t) = \frac{f_n(t) - f_n(x)}{t-x}, \quad \phi(x) = \frac{f(t) - f(x)}{t-x}$$

on $[a, b] \setminus \{x\}$. Now, the first inequality in (8.3.5) shows that for $n, m \geq N$ we have

$$|\phi_n(t) - \phi_m(t)| \leq \frac{\varepsilon}{2(b-a)}.$$

Thus $(\phi_n)_n$ converges uniformly on $[a, b] \setminus \{x\}$. Moreover, $\phi_n \to \phi$ pointwise since $f_n \to f$, so $\phi_n \to \phi$ uniformly.

By Theorem 8.3.1 we find that

$$f'(x) = \lim_{t\to x} \phi(t) = \lim_{n\to\infty} \lim_{t\to x} \phi_n(t) = \lim_{n\to\infty} f_n'(x). \qquad \square$$

Example 8.3.8. Requiring the convergence of $(f_n)_n$ at at least on point $x_0 \in [a, b]$ in Theorem 8.3.7 is necessary. Indeed, if one were to drop that assumption, then the sequence

$$f_n(x) = n$$

of constant functions becomes a counterexample.

8.4 Power series

We now apply the results of the previous section to study power series.

Theorem 8.4.1. *Let*

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

be a power series with radius of convergence R. Then for all $r < R$, f converges uniformly on $B_r(\alpha)$.

Proof. We have that

$$\sup_{x \in B_r(\alpha)} |c_n(x-\alpha)^n| = |c_n| r^n.$$

Moreover, notice that the power series

$$\sum_{n=0}^{\infty} |c_n| \cdot x^n$$

also has radius of convergence R. Thus

$$\sum_{n=0}^{\infty} |c_n| r^n < \infty$$

since $r < R$, and hence f converges uniformly on $B_r(\alpha)$ by the Weierstrass M-test. □

Lemma 8.4.2. *Let $\sum_n c_n(x-\alpha)^n$ be a power series with radius of convergence R. Then the power series*

$$\sum_{n=1}^{\infty} n c_n (x-\alpha)^{n-1}$$

also has radius of convergence $\geq R$.

Proof. By translating, assume that $\alpha = 0$. Now, take any $x_0 \in B_R(0)$ and choose some $y \in B_R(0)$ with $|x_0| < |y|$. Then for n sufficiently large we have that

$$n \left| \frac{x_0}{y} \right|^n < 1$$

since $|x_0/y| < 1$, and thus for n sufficiently large we have that

$$|n c_n x_0^n| = |c_n y^n| \cdot |n(x_0/y)^n| < |c_n y^n|. \tag{8.4.1}$$

But since $|y| < R$ we have that

$$\sum_{n=0}^{\infty} c_n y^n$$

converges absolutely. Thus by comparison and using (8.4.1) we have that

$$\sum_{n=1}^{\infty} n c_n x_0^n$$

converges absolutely. Since x_0 was arbitrary, we find that $\sum_n n c_n x^{n-1}$ converges on $B_R(0)$ and thus has radius of convergence $\geq R$. □

Theorem 8.4.3. *Let*

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

have radius of convergence R. Then f is differentiable on $B_R(\alpha)$ with

$$f'(x) = \sum_{n=1}^{\infty} n c_n (x - \alpha)^{n-1}.$$

Proof. It suffices to prove the result on balls $B_r(\alpha)$ for any $r < R$. Thus fix $r < R$.

Let $(s_n)_n$ be the sequence of partial sums of f. Then we have that $(s_n')_n$ is the sequence of partial sums for the series

$$\sum_{n=1}^{\infty} n c_n (x - \alpha)^{n-1}.$$

Combining Lemma 8.4.2 and Theorem 8.4.1 we see that $(s_n')_n$ converges uniformly on $B_r(\alpha)$. Moreover, $(s_n)_n$ converges uniformly to f on $B_r(\alpha)$ also by Theorem 8.4.1. It follows then from Theorem 8.3.7 that f is differentiable on $B_r(\alpha)$ with

$$f'(x) = \lim_{n \to \infty} s_n'(x) = \sum_{n=1}^{\infty} n c_n (x - \alpha)^{n-1}$$

as required. □

Example 8.4.4. Consider again

$$\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}.$$

Then by Theorem 8.4.3 we have that

$$\begin{aligned}\exp'(x) &= \sum_{n=1}^{\infty} \frac{nx^{n-1}}{n!} \\ &= \sum_{n=1}^{\infty} \frac{x^{n-1}}{(n-1)!} \\ &= \exp(x).\end{aligned}$$

By the above example, we know that $f(x) = \exp(x)$ satisfies

$$f'(x) = f(x), \quad f(0) = 1. \tag{8.4.2}$$

By Example 6.6.6, we know that $\exp(x)$ is therefore the *unique* function satisfying (8.4.2). Thus any familiar results one may have heard of about the exponential function can be deduced from (8.4.2) alone. We give an example of this in the following theorem.

Theorem 8.4.5. *For all $x, y \in \mathbb{R}$ we have that*

$$\exp(x + y) = \exp(x)\exp(y).$$

Proof. Fix $y \in \mathbb{R}$ and consider the function

$$g(x) = \exp(x + y)\exp(-x).$$

Then we have by the chain rule that

$$\begin{aligned}g'(x) &= \exp(x + y)\exp(-x) - \exp(x + y)\exp(-x) \\ &= 0.\end{aligned}$$

It follows that g is a constant function, so

$$\exp(x + y)\exp(-x) = g(x) = g(0) = \exp(y) \tag{8.4.3}$$

for all $x, y \in \mathbb{R}$.

Letting $y = 0$ in (8.4.3) we get that

$$\exp(x)\exp(-x) = 1.$$

Thus multiplying both sides of (8.4.3) by $\exp(x)$ the result follows. $\square$

8.5 $C(K)$ and alternative perspectives

We can phrase the previous sections in terms of familiar language.

Definition 8.5.1. Let X be a metric space. Let $B(X)$ be the set of bounded functions $X \to \mathbb{R}$ and let $C(X)$ be the set of continuous, bounded functions $X \to \mathbb{R}$.

If $X = I$ is an interval, define $\mathscr{R}(I)$ to be the set of Riemann integrable functions $I \to \mathbb{R}$.

Remark 8.5.2. If K is compact then by Theorem 5.3.3 all continuous functions $K \to \mathbb{R}$ are bounded, so $C(K)$ is just the set of continuous functions $K \to \mathbb{R}$.

For general X, it is clear from definition that we have

$$C(X) \subseteq B(X).$$

For $X = I$ an interval, we have by Theorem 7.1.9 that

$$C(I) \subseteq \mathscr{R}(I) \subseteq B(X).$$

We now define a function $\|\cdot\|_\infty : B(X) \to \mathbb{R}$ called the sup-*norm* which is given by

$$\|f\|_\infty = \sup_{x \in X} |f(x)|.$$

Note that $\|f\|_\infty$ is well-defined since $f \in B(X)$ is bounded.

The sup-norm turns $B(X)$ into a metric space via the metric

$$d(f, g) = \|f - g\|_\infty$$

and thus the subsets $C(X)$ and $\mathscr{R}(I)$ also become metric spaces via the same metric. From here on out, whenever we write $C(X)$, $B(X)$ or $\mathscr{R}(I)$ it will be assumed to have this metric.

Proposition 8.5.3. *Let $(f_n)_n$ be a sequence in $B(X)$ and $f \in B(X)$. Then*

(i) $f_n \to f$ uniformly if and only if $f_n \to f$ in $B(X)$.

(ii) $(f_n)_n$ is uniformly Cauchy if and only if $(f_n)_n$ is Cauchy in $B(X)$.

Proof. (i) is an immediate consequence of Proposition 8.2.2 and (ii) may be proven similarly. $\square$

With the viewpoint, if we consider only bounded functions then Theorem 8.2.5 follows from general results about convergent and Cauchy sequences in metric spaces. Moreover, the theorems in the previous section involving continuity and integrability may be encapsulated as follows.

Theorem 8.5.4. *(i)* $B(X)$ *is complete.*

(ii) $C(X) \subseteq B(X)$ *is a closed subspace.*

(iii) $\mathscr{R}(I) \subseteq B(I)$ *is a closed subspace.*

Proof. (i) follows from Theorem 8.2.5(iii) which tells us that uniformly Cauchy sequences of functions $X \to \mathbb{R}$ are uniformly convergent.

For (ii), suppose that $(f_n)_n$ is a sequence of bounded, continuous functions with $f_n \to f$ in $B(X)$. Then by Theorem 8.3.2 we have that f is continuous, so $f \in C(X)$. Thus $C(X) \subseteq B(X)$ is closed.

For (iii), suppose that $(f_n)_n$ is a sequence of integrable functions $I \to \mathbb{R}$ with $f_n \to f$ in $B(I)$. Then by Theorem 8.3.5 we have that f is integrable, so $f \in \mathscr{R}(I)$. Thus $\mathscr{R}(I) \subseteq B(I)$ is closed as required. □

Theorem 8.5.5. *The integral*

$$\int_a^b : \mathscr{R}([a, b]) \to \mathbb{R}$$

is $(b - a)$*-Lipschitz, in particular uniformly continuous.*

Proof. Let $f, g \in \mathscr{R}([a, b])$. Then we have that

$$\begin{aligned} \left| \int_a^b f \, \mathrm{d}x - \int_a^b g \, \mathrm{d}x \right| &= \left| \int_a^b (f - g) \, \mathrm{d}x \right| \\ &\leq \sup_{x \in [a,b]} |f(x) - g(x)| \cdot (b - a) \\ &= (b - a) \cdot \|f - g\|_\infty \end{aligned}$$

so the integral is $(b - a)$-Lipschitz. □

Combining Theorem 8.5.4(iii) with Theorem 8.5.5 we may deduce the full statement of Theorem 8.3.5 using the sequence definition of continuity.

Exercises

Exercise 8.1. Consider the sequence of functions $f_n : \mathbb{R} \to \mathbb{R}$ given by

$$f_n(x) = 1_{[n,\infty)}(x) = \begin{cases} 1 & \text{if } x \geq n \\ 0 & \text{if } x < n. \end{cases}$$

Find a function $f : \mathbb{R} \to \mathbb{R}$ such that $f_n \to f$ pointwise. Is this convergence uniform?

Exercise 8.2. Suppose that $f_1 \geq f_2 \geq f_3 \geq \cdots$ are continuous functions $[a, b] \to \mathbb{R}$ with $f_n \to 0$ pointwise. Does it follow that

$$\int_a^b f_n \, \mathrm{d}x \to 0?$$

Prove or give a counterexample.

Exercise 8.3. We say that a sequence $(f_n)_n$ of functions $X \to \mathbb{R}$ is *uniformly bounded* if there exists an M such that

$$\sup_{x \in X} |f_n(x)| \leq M$$

for all n. Show that a uniformly convergent sequence of bounded functions is uniformly bounded.

Exercise 8.4. Show that if $(f_n)_n$ and $(g_n)_n$ are sequences of bounded functions $X \to \mathbb{R}$ with $f_n \to f$ and $g_n \to g$ uniformly, then $f_n g_n \to fg$ uniformly. Show that this is not necessarily the case if we drop the assumption that the f_n's and g_n's are bounded.

9 Advanced topics

9.1 The Stone-Weierstrass theorem

In Section 6.6 saw a potential method for approximating infinitely differentiable functions by polynomials using Taylor polynomials. However, for this method to work we need our function (i) to be infinitely differentiable and (ii) have bounds on its higher derivatives so that Taylor's error term vanishes. It turns out, however, that if we allow ourselves the freedom approximate with *any* polynomial, not just Taylor's polynomials, then both of these issues vanish.

Theorem 9.1.1 (Weierstrass)**.** *If f is a continuous function $[a, b] \to \mathbb{R}$, then there exists a sequence of polynomials $(P_n)_n$ with $P_n \to f$ uniformly.*

Put differently, polynomials are dense in $C([a, b])$.

Proof. Assume for convenience that $[a, b] = [0, 1]$, as the same proof works for a general interval with minor modification. Moreover, we may assume that $f(0) = f(1) = 0$ since we may consider the function

$$g(x) = f(x) - f(0) + x(f(0) - f(1))$$

which satisfies $g(0) = g(1) = 0$. If $P_n \to g$ uniformly, then $P_n + f - g \to f$ uniformly and $P_n + f - g$ are still polynomials.

Since $f(0) = f(1) = 1$, we may extend f to a continuous function $\mathbb{R} \to \mathbb{R}$ by setting f to be zero outside of $[0, 1]$. Since f is only non-zero on $[0, 1]$, which is compact, f is then uniformly continuous on $\mathbb{R}$.

For $n \in \mathbb{N}$, set

$$Q_n(x) = c_n(1 - x^2)^n$$

where c_n is chosen such that

$$\int_{-1}^{1} Q_n(x)\,\mathrm{d}x = 1.$$

Now,

$$\begin{aligned}\int_{-1}^{1} (1 - x^2)^n\,\mathrm{d}x = 2\int_{0}^{1} (1 - x^2)^n\,\mathrm{d}x \geq 2\int_{0}^{1/\sqrt{n}} (1 - x^2)^n\,\mathrm{d}x \\ \geq \int_{0}^{1/\sqrt{n}} (1 - nx^2)\,\mathrm{d}x \\ = \frac{4}{3\sqrt{n}} \\ > \frac{1}{\sqrt{n}}\end{aligned} \tag{9.1.1}$$

where we have used that $(1 - x^2)^n \geq 1 - nx^2$ on $(0, 1)$. Thus follows by considering

$$h(x) = (1 - x^2)^n - (1 - nx^2)$$

which satisfies $h(0) = 0$ and $h'(x) \geq 0$ on $(0, 1)$. It follows then from (9.1.1) and the definition of c_n that

$$c_n < \sqrt{n}.$$

Using this bound, for any $\delta > 0$ we have that, on $[\delta, 1]$,

$$0 \leq Q_n(x) \leq \sqrt{n}(1 - \delta^2)^n$$

so $Q_n \to 0$ uniformly on $[\delta, 1]$.

Now set

$$P_n(x) = \int_{-1}^{1} f(x+t)Q_n(t)\,\mathrm{d}t$$

for $x \in [0,1]$. Since f is zero outside $(0,1)$, we have that

$$P_n(x) = \int_{-x}^{1-x} f(x+t)Q_n(t)\,\mathrm{d}t = \int_{0}^{1} f(u)Q_n(u-x)\,\mathrm{d}u$$

where the last inequality comes from making the change of variables $t = u - x$. Since Q_n is a polynomial, the last expression shows that $P_n(x)$ is a polynomial. We will show that $P_n \to f$ uniformly.

Fix $\varepsilon > 0$ and choose $1 > \delta > 0$ such that

$$|x-y| < \delta \implies |f(x) - f(y)| < \varepsilon/2$$

and set $M = \sup |f(x)|$. We then have for any $0 \leq x \leq 1$ that

$$\begin{aligned} |P_n(x) - f(x)| &= \left| \int_{-1}^{1} [f(x+t) - f(x)]Q_n(t)\,\mathrm{d}t \right| \\ &\leq \int_{-1}^{1} |f(x+t) - f(x)| \cdot Q_n(t)\,\mathrm{d}t \\ &\leq 2M \int_{-1}^{-\delta} Q_n(t)\,\mathrm{d}t + \frac{\varepsilon}{2} \int_{-\delta}^{\delta} Q_n(t)\,\mathrm{d}t + 2M \int_{\delta}^{1} Q_n(t)\,\mathrm{d}t \\ &\leq 4M \int_{\delta}^{1} Q_n(t)\,\mathrm{d}t + \frac{\varepsilon}{2}. \end{aligned}$$

Since $Q_n \to 0$ uniformly on $[\delta, 1]$ this expression is $< \varepsilon$ whenever n is sufficiently large. Thus $P_n \to f$ uniformly. □

Our goal for the rest of this section is to extract precisely what properties of polynomials were necessary in proving Theorem 9.1.1 and thereby prove its generalization due to Stone.

Definition 9.1.2. Let $\mathscr{A} \subseteq C(K)$. We say that $\mathscr{A}$ is an *algebra* if for all $f, g \in \mathscr{A}$ and $c \in \mathbb{R}$ we have that (i) $f+g$, (ii) fg and (iii) cf belong to $\mathscr{A}$.

Proposition 9.1.3. *Let $\mathscr{A} \subseteq C(K)$ be an algebra. Then $\overline{\mathscr{A}} \subseteq C(K)$ is also an algebra.*

Proof. Let $f, g \in \overline{\mathscr{A}}$. Then we may find sequences $(f_n)_n$ and $(g_n)_n$ in $\mathscr{A}$ with $f_n \to f$ and $g_n \to g$. However, we then have that

$$f_n + g_n \to f + g, \quad f_n g_n \to fg, \quad cf_n \to cf.$$

Since $\mathscr{A}$ is an algebra, all of these sequences belong to $\mathscr{A}$. Thus $f + g, fg, cf \in \overline{\mathscr{A}}$ as required. $\square$

Definition 9.1.4. Let $\mathscr{A} \subseteq C(K)$ be an algebra. We say that $\mathscr{A}$ *separates points* if for any two distinct points $x_1, x_2 \in K$ we may find some $f \in \mathscr{A}$ with $f(x_1) \neq f(x_2)$.

We say that $\mathscr{A}$ *vanishes at no point of* K if for every $x \in K$ there exists some $f \in \mathscr{A}$ with $f(x) \neq 0$.

Example 9.1.5. The set of polynomials in $C([a, b])$ is an algebra which separates points and vanishes at no point of $[a, b]$. Indeed, for vanishing the constant function 1 never vanishes. For separating points, the polynomial $p(x) = x$ separates any two distinct points.

Lemma 9.1.6. *Let $\mathscr{A} \subseteq C(K)$ be an algebra which separates points and vanishes at no point of K. Then given any two distinct points $x_1, x_2 \in K$ and any two $c_1, c_2 \in \mathbb{R}$, there exists $f \in \mathscr{A}$ with*

$$f(x_1) = c_1, \quad f(x_2) = c_2.$$

Proof. The hypotheses give us functions $g, h, k \in \mathscr{A}$ such that

$$g(x_1) \neq g(x_2), \quad h(x_1) \neq 0, \quad k(x_2) \neq 0.$$

Set

$$u = gk - g(x_1)k, \quad v = gh - g(x_2)h$$

so that $u, v \in \mathscr{A}$ and $u(x_1) = v(x_2) = 0$, $u(x_2) \neq 0$, $v(x_1) \neq 0$. Then let

$$f = \frac{c_1 v}{v(x_1)} + \frac{c_2 u}{u(x_2)} \in \mathscr{A}.$$

$\square$

Theorem 9.1.7 (Stone-Weierstrass)**.** *Let K be compact and $\mathscr{A} \subseteq C(K)$ an algebra which separates points and vanishes on no points of K. Then $\mathscr{A}$ is dense in $C(K)$.*

We will prove this via a series of lemmas.

Lemma 9.1.8. *For every interval $[-a, a]$ there exists a sequence of polynomials P_n with $P_n(0) = 0$ such that $P_n \to |\cdot|$ uniformly on $[-a, a]$.*

Proof. $f(x) = |x|$ is continuous on $[-a, a]$ so by Theorem 9.1.1 there exists a sequence of polynomials $(P_n)_n$ with $P_n \to f$ uniformly on $[-a, a]$. Since $f(0) = 0$ we must have that $P_n(0) \to 0$. Thus set

$$P_n^*(x) = P_n(x) - P_n(0).$$

Then $P_n^*(0) = 0$ and $P_n^* \to f$ uniformly as required. $\square$

Lemma 9.1.9. *Let $\mathscr{A}$ be as in Theorem 9.1.7 and let $\mathscr{B} = \overline{\mathscr{A}}$. If $f \in \mathscr{B}$ then $|f| \in \mathscr{B}$.*

Proof. Let $a = \sup |f(x)|$ and let $\varepsilon > 0$ be fixed. By Lemma 9.1.8 we may find real numbers $c_1, \ldots, c_n$ such that

$$\left| \sum_{i=1}^{n} c_i y^i - |y| \right| < \varepsilon$$

for all $y \in [-a, a]$. Since $\mathscr{B}$ is an algebra by Proposition 9.1.3, we have that

$$g = \sum_{i=1}^{n} c_i f^i \in \mathscr{B}.$$

Moreover, by definition of a, we have that

$$|g(x) - |f(x)|| < \varepsilon$$

for all $x \in K$, i.e. $\|g - |f|\|_\infty < \varepsilon$. Since $\varepsilon > 0$ was arbitrary, we have that $|f| \in \overline{\mathscr{B}} = \mathscr{B}$ as required. □

Lemma 9.1.10. *Let $\mathscr{A}$ be as in Theorem 9.1.7 and let $\mathscr{B} = \overline{\mathscr{A}}$. If $f, g \in \mathscr{B}$ then $\max(f, g) \in \mathscr{B}$ and $\min(f, g) \in \mathscr{B}$.*

Proof. We have that

$$\max(f, g) = \frac{f + g}{2} + \frac{|f - g|}{2}$$

and

$$\min(f, g) = \frac{f + g}{2} - \frac{|f - g|}{2}$$

so the result follows from Lemma 9.1.9. □

Lemma 9.1.11. *Let $\mathscr{A}$ be as in Theorem 9.1.7 and let $\mathscr{B} = \overline{\mathscr{A}}$. Fix $\varepsilon > 0$ and $x \in K$ and let $f \in C(K)$. Then there exists a function $g \in \mathscr{B}$ such that $g(x) = f(x)$ and $g(t) > f(t) - \varepsilon$ for all $t \in K$.*

Proof. Since $\mathscr{A} \subseteq \mathscr{B}$, $\mathscr{B}$ also separates points and does not vanish on any point of K. Thus for every $y \in K$, by Lemma 9.1.6, we may find some $h_y \in \mathscr{B}$ such that

$$h_y(x) = f(x), \quad h_y(y) = f(y).$$

Since h_y is continuous, there exists a neighborhood J_y of y such that

$$h_y(t) > f(t) - \varepsilon \tag{9.1.2}$$

for all $t \in J_y$. Since K is compact, we may write

$$K = J_{y_1} \cup \cdots \cup J_{y_n} \tag{9.1.3}$$

for some $y_1, \ldots, y_n \in K$.

Now set

$$g = \max(h_{y_1}, \ldots, h_{y_n}).$$

By Lemma 9.1.10 we have that $g \in \mathscr{B}$. Moreover, we have that $g(x) = f(x)$ and $g(t) > f(t) - \varepsilon$ for all $t \in K$ by (9.1.3) and (9.1.2). □

Proof of Theorem 9.1.7. Fix $\varepsilon > 0$ and let $f \in C(K)$. By Lemma 9.1.11 we may find for every $x \in K$ a function $g_x \in \mathscr{B} = \overline{\mathscr{A}}$ such that $g_x(x) = f(x)$ and $g_x(t) > f(t) - \varepsilon$ for all $t \in K$.

By the continuity of g_x, there exists a neighborhood V_x of x such that

$$g_x(t) < f(t) + \varepsilon \tag{9.1.4}$$

for all $t \in J_x$. Since K is compact we may write

$$K = V_{x_1} \cup \cdots \cup V_{x_n} \tag{9.1.5}$$

for some $x_1 \ldots, x_n \in K$. Then set

$$g = \min(g_{x_1}, \ldots, g_{x_n}) \in \mathscr{B}.$$

By construction of the g_x's, we have that

$$g(t) > f(t) - \varepsilon$$

for all $t \in K$. Moreover, by (9.1.4) and (9.1.5) we have that

$$g(t) < f(t) + \varepsilon$$

for all $t \in K$. Thus $\|g - f\|_\infty < \varepsilon$.

Since ε was arbitrary, we see that $f \in \overline{\mathscr{B}} = \mathscr{B}$ as required. □

We now discuss some applications of this theorem.

Definition 9.1.12. Define $\sin : \mathbb{R} \to \mathbb{R}$ and $\cos : \mathbb{R} \to \mathbb{R}$ via the power series

$$\sin(x) = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}, \quad \cos(x) = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}.$$

One may check that sin and cos indeed have infinite radius of convergence and by Theorem 8.4.3 we see that

$$\sin'(x) = \cos(x), \quad \cos'(x) = -\sin(x).$$

Notice also that

$$\exp(ix) = \cos(x) + i\sin(x)$$

by regrouping the terms in the power series for $\exp(ix)$, which is valid as $\exp(ix)$ is convergent.

Proposition 9.1.13. *Get* $\alpha, \beta \in \mathbb{R}$ *the following relations hold:*

(i) $\cos(\alpha + \beta) = \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta)$

(ii) $\sin(\alpha + \beta) = \sin(\alpha)\cos(\beta) + \sin(\beta)\cos(\alpha)$.

Proof. By Theorem 8.4.5 we have that

$$\begin{aligned}\cos(\alpha+\beta) + i\sin(\alpha+\beta) &= \exp(i(\alpha+\beta)) \\ &= \exp(i\alpha)\exp(i\beta) \\ &= (\cos(\alpha) + i\sin(\alpha))(\cos(\beta) + i\sin(\beta)).\end{aligned}$$

Performing this multiplication and comparing the real and imaginary parts then gives the result. □

Corollary 9.1.14. *Given* $\alpha, \beta \in \mathbb{R}$ *the following relations hold:*

(i) $\sin(\alpha)\sin(\beta) = \frac{1}{2}\left[\cos(\alpha - \beta) - \cos(\alpha + \beta)\right]$

(ii) $\cos(\alpha)\cos(\beta) = \frac{1}{2}\left[\cos(\alpha + \beta) + \cos(\alpha - \beta)\right]$

(iii) $\sin(\alpha)\cos(\beta) = \frac{1}{2}\left[\sin(\alpha + \beta) + \sin(\alpha - \beta)\right]$.

Proof. A direct check using the sum formulas of Proposition 9.1.13. □

One may then show that sin has a *unique* zero in $(0, 2)$ one we define $\pi \in \mathbb{R}$ to be twice the value of this zero, i.e. so that $\pi/2 \in (0, 2)$ and $\sin(\pi/2) = 0$. One may then confirm using the sum formulas in Proposition 9.1.13 that sin and cos are 2π-periodic, i.e. that

$$\sin(x + 2\pi) = \sin(x), \quad \cos(x + 2\pi) = \cos(x)$$

for all $x \in \mathbb{R}$.

Theorem 9.1.15. *Let*

$$\mathscr{A} = \left\{ \sum_{k=0}^{n} c_k \cos(k\pi x) + \sum_{k=1}^{m} b_k \sin(k\pi x) : n, m \in \mathbb{N},\ c_k, b_k \in \mathbb{R} \right\}$$

be the set of all linear combinations of sin *and* cos *with non-negative integer frequencies. Then $\mathscr{A}$ is dense in $C([0,1])$.*

Proof. We show first that $\mathscr{A}$ is an algebra. It is clear that $\mathscr{A}$ is closed under scalar multiplication and addition. The fact that $\mathscr{A}$ is closed under multiplication follows from Corollary 9.1.14.

By Stone-Weierstrass, it suffices then to show that $\mathscr{A}$ separates points and vanishes at no point of $[a,b]$. The latter condition is true since the constant function $1 = \cos(0)$ belongs to $\mathscr{A}$. Second, we have that $\cos(\pi x) \in \mathscr{A}$ is injective on $[0,1]$ hence $\mathscr{A}$ separates points. □

This theorem is the basis for expecting Fourier analysis to be possible.

Another neat application is the following.

Theorem 9.1.16. *Let $f : [a,b] \to \mathbb{R}$ be continuous and suppose that*

$$\int_a^b f(x)x^n \, \mathrm{d}x = 0$$

for all $n \geq 0$. Then $f = 0$.

Proof. By linearity of the integral, this implies that $\int_a^b fp \, \mathrm{d}x = 0$ for all polynomials p. But polynomials are dense in $C([a,b])$ so let $p_n \to f$ uniformly. Then we have that $fp_n \to f^2$ uniformly and thus by Theorem 8.3.5 we have that

$$0 = \int_a^b fp_n \, \mathrm{d}x \to \int_a^b f^2 \, \mathrm{d}x.$$

Hence we have that $\int_a^b f^2 \, \mathrm{d}x = 0$ and since f^2 is continuous and non-negative, this forces $f^2 = 0$ so that $f = 0$. □

9.2 The Baire category theorem

Definition 9.2.1. We say that a metric space X is a *Baire space* if the countable intersection of open dense sets in X is dense.

Before discussing the main theorem, we give an alternative formulation of what it means to be a Baire space.

Definition 9.2.2. Let X be a metric space and $E \subseteq X$. Define the *interior of Einterior*, denoted by E°, to be the set

$$E^\circ = \{x \in E : \exists \text{ a neighborhood } U \text{ of } x \text{ with } U \subseteq E\}.$$

Remark 9.2.3. One easily checks that E° is open and that it is the *largest* open contained in E. In particular, a set E is open if and only if $E^\circ = E$.

Definition 9.2.4. We say that a set $E \subseteq X$ is *nowhere dense* if $E^\circ = \emptyset$. Alternatively, E is nowhere dense if E^c is dense.

Proposition 9.2.5. *A space X is a Baire space if and only if the countable union of nowhere dense closed sets is nowhere dense.*

Proof. Take complements in the usual definition of a Baire space. □

We now have the main result.

Theorem 9.2.6 (Baire Category Theorem)**.** *Every complete metric space is a Baire space.*

For this we need some preliminary definitions and results.

Definition 9.2.7. Let (X, d) be a metric space and $E \subseteq X$ be a non-empty subset. Define the *diameter of E*, denoted by $\operatorname{diam} E$, to be

$$\operatorname{diam} E = \sup_{x,y \in E} d(x, y).$$

Proposition 9.2.8. *Let X be a complete metric space and suppose that* $E_1 \supseteq E_2 \supseteq E_3 \supseteq \cdots$ *is a nested sequence of non-empty closed sets with* $\operatorname{diam} E_n \to 0$*. Then* $\bigcap_n E_n$ *is non-empty.*

Proof. Let $(x_n)_n$ be any sequence such that $x_n \in E_n$. Fix $\varepsilon > 0$ and let $N \in \mathbb{N}$ be such that $\operatorname{diam} E_N < \varepsilon$. Since the E_n's are nested, for all $n, m \geq N$ we have that $x_n, x_m \in E_N$ and thus

$$d(x_n, x_m) \leq \operatorname{diam} E_N < \varepsilon.$$

Hence $(x_n)_n$ is Cauchy and so $x_n \to x$ for some x since X is complete.

Now, for every $N \in \mathbb{N}$ we have that $(x_n)_{n \geq N}$ is a sequence in E_N converging to x. Since E_N is closed we have that $x \in E_N$ and thus $x \in \bigcap_n E_n$ as required. □

Proof of Theorem 9.2.6. Let $\{U_n\}_n$ be open dense subsets in X. Let U be any open of X. Our goal is to show that

$$U \cap \bigcap_n U_n \neq \emptyset.$$

To do this we construct a sequence of points $(x_n)_n$ and radii $(r_n)_n$ with $r_n \to 0$ such that

$$B_{r_1}(x_1) \supseteq B_{r_2}(x_2) \supseteq B_{r_3}(x_3) \supseteq \cdots$$

with the additional property that $\overline{B_{r_n}(x_n)} \subseteq U_n$ for all n and $\overline{B_{r_1}(x_1)} \subseteq U$. It then follows from Proposition 9.2.8 that

$$\emptyset \neq \bigcap_n \overline{B_{r_n}(x_n)} \subseteq U \cap \bigcap_n U_n$$

as required.

To do this, since U_1 is dense, we may find $x_1 \in U \cap U_1$. Since $U \cap U_1$ is open there exists some $\varepsilon < 2$ such that $B_\varepsilon(x) \subseteq U \cap U_1$. Then take $r_1 = \varepsilon/2 < 1$. By construction we have that

$$\overline{B_{r_1}(x_1)} \subseteq B_\varepsilon(x_1) \subseteq U \cap U_1.$$

Now suppose x_{n-1} and r_{n-1} have been constructed. Then since U_n is dense we have that there exists some $x_n \in B_{r_{n-1}}(x_{n-1}) \cap U_n$. Moreover, since $B_{r_{n-1}}(x_{n-1}) \cap U_n$ is open we may find some $\varepsilon < 2/n$ such that $B_\varepsilon(x_n) \subseteq B_{r_{n-1}}(x_{n-1}) \cap U_n$. Then set $r_n = \varepsilon/2 < 1/n$. It follows that

$$\overline{B_{r_n}(x_n)} \subseteq B_\varepsilon(x_n) \subseteq B_{r_{n-1}}(x_{n-1}) \cap U_n.$$

The sequences $(x_n)_n$ and $(r_n)_n$ then satisfy our desired properties. □

We now need to strengthen this result.

Theorem 9.2.9. *Open subsets of complete metric spaces are Baire spaces.*

Proof. Let X be a complete metric space and $U \subseteq X$ be open. Given a closed set $E \subseteq X$ define

$$\rho_E(x) = \inf_{y \in E} d(x, y).$$

One may show that ρ_E is 1-Lipschitz and $\rho_E(x) = 0$ if and only if $x \in E$ since E is closed.

Now define a map

$$\begin{aligned} \varphi : U &\longrightarrow X \times \mathbb{R} \\ u &\longmapsto (u, 1/\rho_{U^c}(u)). \end{aligned}$$

This map is continuous and injective. The image $\varphi(U)$ is closed as well. Indeed, suppose that

$$(u_n, \rho_{U^c}(u_n)^{-1}) \to (a, b).$$

Now, if $a \in U^c$ then $\rho_{U_n}(u_n) \to \rho_{U_n}(a) = 0$ which contradicts $\rho_{U_n}(u_n)^{-1} \to b$. Thus $a \in U$ and so by continuity $\rho_{U^c}(u_n)^{-1} \to \rho_{U^c}(a)^{-1}$. Hence

$$(a, b) = (a, \rho_{U^c}(a)^{-1}) \in \varphi(U).$$

so $\varphi(U)$ is closed.

It follows that $U \cong \varphi(U)$ where the inverse map is projection $\pi_X : X \times \mathbb{R} \to X$. But $\varphi(U)$ is a closed subset of the complete metric space $X \times \mathbb{R}$, hence complete. Thus $\varphi(U)$ is a Baire space by Theorem 9.2.6 and since being a Baire space is a topological property, we see that U is a Baire space. □

This has some immediate corollaries.

Theorem 9.2.10. *Let X be a complete metric space and let $(F_n)_n$ be a countable collection of closed subsets which cover X. Then $\bigcup_n F_n^\circ$ is a dense open subset of X.*

Proof. The fact that $\bigcup_n F_n^\circ$ is open is clear, so we need only show that it is dense.

Let U be a non-empty open set of X. Then we have that

$$U = \bigcup_n (U \cap F_n). \tag{9.2.1}$$

Since U is a Baire space by Theorem 9.2.9 we know that not every $U \cap F_n$ can be nowhere dense in U, as otherwise the union $\bigcup_n (U \cap F_n)$ would be nowhere dense in U contradicting (9.2.1).

Write $\text{int}_X E$ for the interior of E viewed as a subset of X. Then the above shows that we may find some N such that $\text{int}_U(U \cap F_N) \neq \emptyset$. But U is open in X so $\text{int}_U(U \cap F_N)$ is an open in X contained in both U and F_N. Thus

$$\emptyset \neq \text{int}_U(U \cap F_N) \subseteq U \cap \text{int}_X F_N.$$

It follows that $U \cap \bigcup_n F_n^\circ \neq \emptyset$ so $\bigcup_n F_n^\circ$ is dense in X. □

Corollary 9.2.11. *The set of isolated points in a countable complete metric space is dense. In particular, a complete metric space with no isolated points is uncountable.*

Proof. Let X be a countable complete metric space. Then X is a Baire space by Theorem 9.2.6. Let $I \subseteq X$ be the set of isolated points of X, i.e. the collection of points $x \in X$ such that $\{x\}$ is open. Then for every $x \in X \setminus I$ we have that $X \setminus \{x\}$ is a dense open set. Then as X is a Baire space and countable we have that

$$\bigcap_{x \in X \setminus I} X \setminus \{x\} = I$$

is dense. □

Since $\mathbb{R}$ is has no isolated points, this gives an alternative proof that $\mathbb{R}$ is uncountable.

Theorem 9.2.12 (Baire's Simple Limit Theorem)**.** *Let X, Y be metric spaces with X complete. If $f_n : X \to Y$ are continuous with $f_n \to f$ pointwise, then the continuity points of f are dense in X.*

Proof. For every $n \in \mathbb{N}$ and $\varepsilon > 0$, define

$$F_{n,\varepsilon} = \{x \in X : \forall p \geq n,\ d_Y(f_n(x), f_p(x)) \leq \varepsilon\}.$$

Notice that

$$F_{n,\varepsilon} = \bigcap_{p \geq n} \{x \in X : d_Y(f_n(x), f_p(x)) \leq \varepsilon\}$$

is the intersection of closed sets since the f_n's are continuous, so each $F_{n,\varepsilon}$ is closed. Moreover, since $f_n \to f$ pointwise we have that each sequence $(f_n(x))_n$ is Cauchy so

$$X = \bigcup_n F_{n,\varepsilon}.$$

It follows from Theorem 9.2.10 that the set

$$\Omega_\varepsilon = \bigcup_n F_{n,\varepsilon}^\circ$$

is an open dense subset of X.

Next, we show that for every $p \in \Omega_\varepsilon$ there exists a neighborhood V of p such that

$$x \in V \implies d_Y(f(p), f(x)) \leq 3\varepsilon.$$

For this, suppose that $p \in F_{n,\varepsilon}^\circ$. Since f_n is continuous and $F_{n,\varepsilon}^\circ$ is open, there exists a neighborhood V of p contained in $F_{n,\varepsilon}^\circ$ such that

$$x \in V \implies d_Y(f_n(p), f_n(x)) \leq \varepsilon.$$

Since $V \subseteq F_{n,\varepsilon}$ we have that for all $m \geq n$ that

$$x \in V \implies d_Y(f_n(x), f_m(x)) \leq \varepsilon.$$

Letting $m \to \infty$ we find that

$$x \in V \implies d_Y(f_n(x), f(x)) \leq \varepsilon.$$

Thus by the triangle inequality we find that for all $x \in V$ we have

$$\begin{aligned} d_Y(f(p), f(x)) &\leq d_Y(f(p), f_n(p)) + d_Y(f_n(p), f_n(x)) + d_Y(f_n(x), f(x)) \\ &\leq \varepsilon + \varepsilon + \varepsilon \\ &= 3\varepsilon \end{aligned}$$

as required.

It follows then that f is continuous an the set $\bigcap_n \Omega_{1/n}$ which is dense as each $\Omega_{1/n}$ is an open dense set and X is a Baire space by Theorem 9.2.6. □

Corollary 9.2.13. *Let $f : [a, b] \to \mathbb{R}$ be a differentiable function. Then f' is continuous an a dense subset of $[a, b]$.*

Proof. We have that

$$\frac{f(x + 1/n) - f(x)}{1/n} \to f'(x)$$

pointwise. Moreover, each difference quotient on the left is continuous since f is continuous. Thus by Theorem 9.2.12 f' is continuous on a dense subset of $[a, b]$. □

The following theorem may also be proven using the Baire Category Theorem.

Theorem 9.2.14. *The set of no-where differentiable functions are dense in $C([a, b])$.*

Exercises

Exercise 9.1 (†)**.** Let $f : \mathbb{R} \to \mathbb{R}$ be an infinitely differentiable function such that for every $x \in \mathbb{R}$ there exists an $n \in \mathbb{N}$ with $f^{(n)}(x) = 0$. Show that f is a polynomial.

References

[Rud64] Walter Rudin. *Principles of Mathematical Analysis*, volume 3. McGraw-Hill New York, 1964.

Index