REAL ANALYSIS, 425A, SPRING 2025

STEVEN HEILMAN

Abstract. These notes are mostly copied from those of T. Tao from 2003, available here and here.

Contents

Date: April 23, 2025

1. Introduction, Natural Numbers, Real Numbers

1.1. **Introductory Remarks.**

1.1.1. *A rigorous version of calculus.* Here is a "proof" of Euler which in 1735 found the quantity $1 + 1/4 + 1/9 + 1/16 + \cdots$, thereby solving the Basel problem. Do you agree with the logic? Let x be a real number. Then

$$
\begin{aligned}
1 - \pi^2 x^2/6 + \cdots &= \frac{\sin(\pi x)}{\pi x} \quad \text{, by Taylor series} && (1.1.1)\\
&= (1-x)(1+x)(1-x/2)(1+x/2)(1-x/3)(1+x/3)\cdots && (1.1.2)\\
&\qquad \text{, since a nice function is a product of its zeros} \\
&= (1-x^2)(1-x^2/4)(1-x^2/9)\cdots && (1.1.3)\\
&= 1 - x^2(1 + 1/4 + 1/9 + \cdots) + x^4(\cdots) + \cdots && (1.1.4)
\end{aligned}
$$

So, equation the x^2 terms on both sides, we get

$$1 + 1/4 + 1/9 + 1/16 + \cdots = \pi^2/6. \qquad (1.1.5)$$

It is actually possible to make this argument rigorous, but what problems do you see with the amount of rigor? I see a few:

- In what sense does equality hold in (1.1.1)?
- What is the true meaning of an infinite sum, as in (1.1.1)?
- What is the meaning of the infinite product in (1.1.2)?
- Is every function really the product of its zeros? This seems quite unlikely. (In fact it is false in general (consider e^x), but (1.1.2) actually does hold in an appropriate sense.)
- Can we freely rearrange terms in an infinite sum or an infinite product as in (1.1.3) and (1.1.4)? (In general, we cannot, but sometimes we can.)

Euler was a brilliant mathematician, but he also occasionally made some mistakes by using non-rigorous methods. Using intuition and non-rigorous calculations can be very helpful, though! No one else was able to find (1.1.5) at the time. Yet, in order to be entirely certain of facts, we need to ultimately find rigorous proofs of these facts. The above proof would receive only partial credit as a solution on a homework, since it is no longer 1735.

1.1.2. *What will we be learning?* We will learn a fully rigorous version of calculus. That is, we will learn how to answer many of the questions raised in the previous section. The ultimate goal of the course is to develop an ability to read and write rigorous proofs of mathematics. Also, we would like to learn how to rigorously treat calculus. From the time of Newton and Leibniz in the mid 1600s to the time of Cauchy in the mid 1800s, calculus did not have a truly rigorous foundation. And developing such a foundation turned out to be a fairly difficult problem, which arguably lasted to the time of Cantor in the early 1900s. Such a rigorous foundation has been quite influential in all other areas of mathematics.

More generally, in nearly any vocation or avocation, the process of problem solving and thinking rigorously that we learn in this class can be applicable. There is a reason that Euclid's *Elements* were learned by many students in the past, and there is a reason that this abstract, axiomatic method is still taught in our mathematics classes today.

1.1.3. *How will we be learning analysis?* As in the Euclidean axiomatization of geometry, we will begin with the most basic axioms of arithmetic, and we will slowly build up our understanding of numbers. For example, one question that we did not yet address is:

What is a real number?

We perhaps have a good intuitive idea of what a real number is. But what is a real number, really? Maybe you think of a real number in terms of some infinite decimal. So, are the real numbers the set of infinite decimals? For example,

$$1.000000\ldots$$
$$3.141592653589\ldots$$
$$1.34300344300\ldots$$

This seems reasonable at first, but there are some issues with this definition. For example, the following two decimals should really be the same number, even though they look very different.

$$1.00000\ldots \qquad \text{and} \qquad 0.9999999\ldots$$

If you don't agree that these are the same number, then consider what their difference is.

By adjusting for this issue, it is possible to define the real numbers in terms of infinite decimals. However, there are other, better definitions of the real numbers, which are more instructive and more useful later. We will construct the real numbers soon using so-called Cauchy sequences. In order to adjust to axiomatic thinking, and to review induction, we start at the very beginning and define the natural numbers. We emphasize at the outset that we will treat numbers as abstract mathematical objects that satisfy certain properties. Such a treatment perhaps lacks some intuition, but it seems necessary to provide a rigorous foundation of mathematics that can avoid some of the issues we discussed in Euler's proof above. On the other hand, intuition can be quite useful in proving various facts. So, doing mathematics seems to require two complementary modes of thought: the nonrigorous, creative mode, and the rigorous, logical mode.

In this first chapter, we will begin with the axiomatization of the natural numbers, and we will then move to axiomatizations of the integers, rationals, and reals, respectively. The point of studying the axiomatization of the natural numbers is that it will allow a review of induction, and it will lead naturally to our eventual axiomatization of the real number system. However, a rigorous axiomatization of the real number system is a surprisingly difficult creation.

1.1.4. *Why are we learning this material?* This material lays the foundation for a great deal of further subjects. To give just one example, consider Fourier analysis, which is arguably one of the most seminal areas of mathematics. Every time we use a cell phone, or look at a JPEG, or watch an online video (for example, an MPEG), or when a doctor uses an MRI or CT-Scan, Fourier analysis is involved. In Fourier analysis, we begin with a function, we break this function up into simpler pieces, and we then reassemble these pieces. Sometimes we are allowed to break up the function into pieces, and sometimes we are not. The details become unexpectedly subtle. The rigorous way of thinking and the results of this course play a crucial role in dealing with the details of the subject of Fourier analysis.

Abstract reasoning has some advantages and disadvantages. Since abstract reasoning usually does not come naturally, it can be difficult to learn material that is presented in an abstract way. On the other hand, an abstract approach promises more applicability. For example, there are many different ways to interpret a real-valued function on the real line. Such a function could represent the amplitude of a sound wave over time, the price of a stock over time, the displacement of an object over time, and so on.

1.2. **Natural Numbers.** The natural numbers $\mathbb{N}$ are defined by the following axioms.

Definition 1.2.1 (Peano Axioms).

(1) 0 is a natural number.
(2) Every natural number n has a successor $n++$ which is also a natural number.
(3) 0 is not the successor of any natural number. That is, for any natural number n, $n{+}{+} \neq 0$.
(4) Different natural numbers have difference successors. That is, if n, m are natural numbers with $n \neq m$, then $n++ \neq m++$.
(5) (**Principle of Induction**) Let n be a natural number, and let $P(n)$ be any property that holds for n. Assume that $P(0)$ is true, and whenever $P(n)$ is true for any natural number n, $P(n++)$ is also true. Then $P(n)$ is true for every natural number n.

Assumption 1 (The Natural Numbers). There exists a number system $\mathbb{N}$, whose elements we call **natural numbers**, such that Axioms (1) through (5) of Definition 1.2.1 are true.

Definition 1.2.2. Define $1 := 0++$.

Definition 1.2.3 (Addition of Natural Numbers). Let m be a natural number. Define $0 + m := m$. We now define how to add other natural numbers to m. Let n be a natural number. Suppose we have inductively defined $n{+}m$. Then, define $(n{+}{+}){+}m := (n{+}m){+}{+}$.

Remark 1.2.4. By Axiom (5), we have defined addition on all natural numbers n, m.

Exercise 1.2.5. Show that, from Axioms (1), (2) it follows by induction (using Axiom (5)) that addition of two natural numbers produces a natural number.

Remark 1.2.6. Using only the definitions $0 + m = m$ and $(n++) + m = (n + m) ++$, we will deduce all basic facts of arithmetic.

Lemma 1.2.7. *For any natural number n, $n + 0 = n$.*

Remark 1.2.8. Note that we cannot apply commutativity of addition, since it does not immediately follow from the axioms of Definition 1.2.1.

Proof. From Definition 1.2.3, $0+0 = 0$. So, we induct on n. Suppose $n+0 = n$ for a natural number n. We need to show that $(n++)+0 = n++$. From Definition 1.2.3 with $m = 0$, $(n++)+0 = (n+0)++$. By the inductive hypothesis, we then have $(n++)+0 = n++$, as desired. Having completed the inductive step and the base case, we are done. $\square$

Lemma 1.2.9. *For any natural numbers n, m, we have $n + (m++) = (n + m)++$*

Proof. We fix m and induct on n. In the base case $n = 0$, we need to show $0 + (m++) = (0+m)++$. From Definition 1.2.3, we know that $0+(m++) = m++$ and $(0+m)++ = m++$. We conclude that $0 + (m++) = (0 + m)++$, as desired. We now induct on n. Suppose n satisfies $n + (m++) = (n + m)++$. We need to show that

$$(n++) + (m++) = ((n++) + m)++. \qquad (*)$$

From Definition 1.2.3, $(n++)+(m++) = (n+(m++))++$. From the inductive hypothesis, $(n + (m++))++ = ((n + m)++)++$. From Definition 1.2.3, $((n++) + m)++ = ((n+m)++)++$. We conclude that both sides of $(*)$ are equal, so the inductive step holds, and we deduce the lemma. $\square$

Remark 1.2.10. From Definition 1.2.2, Lemma 1.2.7 and Lemma 1.2.9, $n+1 = n+(0++) = (n + 0)++ = n++$, so $n++ = n + 1$ for all natural numbers n.

Proposition 1.2.11 (Addition is Commutative). *For any natural numbers n, m, we have $n + m = m + n$.*

Proof. We fix m and induct on n. In the base case $n = 0$, we need to show that $0+m = m+0$. From Definition 1.2.3, $0+m = m$. From Lemma 1.2.7, $m+0 = m$. Therefore, $0+m = m+0$, as desired. Now, assume that $n + m = m + n$. We need to show that

$$(n++) + m = m + (n++). \qquad (*)$$

From Definition 1.2.3, $(n++) + m = (n + m)++$. From Lemma 1.2.9, $m + (n++) = (m+n)++$. From the inductive hypothesis, $(m+n)++ = (n+m)++$. Putting everything together $(*)$ holds, and the inductive step is complete. $\square$

Proposition 1.2.12 (Addition is Associative). *For any natural numbers a, b, c, we have $(a + b) + c = a + (b + c)$.*

Exercise 1.2.13. Prove Proposition 1.2.12 by fixing two variables and inducting on the third variable.

Proposition 1.2.14 (Cancellation Law). *Let a, b, c be natural numbers such that $a + b = a + c$. Then $b = c$.*

Remark 1.2.15. We have not defined subtraction, so we cannot subtract a from both sides. In fact, we will use the Cancellation Law to *define* subtraction.

Proof. We induct on a. For the base case $a = 0$, we assume that $0 + b = 0 + c$. From Definition 1.2.3, we conclude that $b = c$, thereby proving the base case. Now, assume that:

if $a+b=a+c$, then $b=c$. We need to show that: if $(a++)+b=(a++)+c$, then $b=c$. From Definition 1.2.3, $(a++)+b=(a+b)++$. Similarly, $(a++)+c=(a+c)++$. So, we know that $(a+b)++=(a+c)++$. From the contrapositive of Axiom (4) of Definition 1.2.1, we conclude that $a+b=a+c$. From the inductive hypothesis, $b=c$. So, the inductive step is complete, and we are done. □

Definition 1.2.16 (Positivity). A natural number n is said to be **positive** if and only if $n \neq 0$.

Proposition 1.2.17. *Let a,b be natural numbers. Assume that a is positive. Then $a+b$ is positive.*

Proof. We induct on b. For the base case, $b=0$, and we see that $a+b=a+0=a$. Since a is positive, we conclude that $a+b$ is positive. We now prove the inductive step. Assume that $a+b$ is positive. We need to show that $a+(b++)$ is positive. But $a+(b++)=(a+b)++$, and $(a+b)++\neq 0$ by Axiom (3) of Definition 1.2.1. We have therefore completed the inductive step. □

The following Corollary is the contrapositive of Proposition 1.2.17.

Corollary 1.2.18. *Let a,b be natural numbers such that $a+b=0$. Then $a=b=0$.*

Definition 1.2.19 (Order). Let n,m be natural numbers. We say that n **is greater than or equal to** m, and we write $n \geq m$ or $m \leq n$, if and only if $n=m+a$ for some natural number a. We say that n **is strictly greater than** m, and we write $n>m$ or $m<n$, if and only if $n \geq m$ and $n \neq m$.

Proposition 1.2.20 (Properties of Order). *Let a,b,c be natural numbers.*

(1) *$a \geq a$.*
(2) *If $a \geq b$ and $b \geq c$, then $a \geq c$.*
(3) *If $a \geq b$ and $b \geq a$, then $a=b$.*
(4) *$a \geq b$ if and only if $a+c \geq b+c$.*
(5) *$a<b$ if and only if $a+c<b+c$.*

Exercise 1.2.21. Prove Proposition 1.2.20.

Proposition 1.2.22 (Trichotomy of Order). *Let a,b be natural numbers. Then exactly one of the following statements is true: $a<b$, $a>b$ or $a=b$.*

1.2.1. *Multiplication.*

Remark 1.2.23. We will now freely use facts about addition of natural numbers, without referencing the above lemmas and propositions.

Definition 1.2.24 (Multiplication). Let m be a natural number. We define multiplication $\times$ as follows. Define $0\times m := 0$. Now, let n be a natural number, and assume we have inductively defined $n\times m$. Then, define $(n++)\times m := (n\times m)+m$.

Remark 1.2.25. One can show by induction that $n\times m$ is a natural number, for any natural numbers n,m.

Exercise 1.2.26. Imitating the proofs of Lemmas 1.2.7 and 1.2.9 and Proposition 1.2.11, show that, for all natural numbers n,m, we have $n\times 0=0$, $n\times(m++)=(n\times m)+n$ and $n\times m=m\times n$.

Remark 1.2.27. Let n, m, r be natural numbers. As is standard, we write nm to denote $n \times m$. Also, $nm + r$ denotes $(n \times m) + r$.

Remark 1.2.28. If a, b are positive natural numbers, than ab is positive. One can prove this using induction and Proposition 1.2.17.

Proposition 1.2.29 (**Distributive Law**)**.** *For any natural numbers a, b, c, we have $a(b + c) = ab + ac$ and $(b + c)a = ba + ca$.*

Proof. From Exercise 1.2.26, multiplication is commutative. So, it suffices to prove $a(b+c) = ab + ac$. Fix a, b. We then induct on c. The base case corresponds to $c = 0$. We need to prove $a(b + 0) = ab + a0$. The left side is ab, and the right side is $ab + 0 = ab$, so the base case is verified. Now, assume that $a(b + c) = ab + ac$ for some natural number c. We need to show that $a(b + (c{+}{+})) = ab + a(c{+}{+})$. The left side is $a((b + c){+}{+}) = a(b + c) + a$, by Definition 1.2.24. So, by the inductive hypothesis, the left side is $ab + ac + a$. Meanwhile, the right side is $ab + ac + a$, by Definition 1.2.24. So, the inductive step has been completed. □

Remark 1.2.30. From Proposition 1.2.29, we can mimic the proof of Proposition 1.2.12 to prove that, for all natural numbers a, b, c, we have $a(bc) = (ab)c$.

Proposition 1.2.31. *Let a, b be natural numbers with $a < b$. If c is a positive natural number, then $ac < bc$.*

Proof. Since $a < b$, there exists a positive natural number d such that $a + d = b$. Multiplying both sides by c and using Proposition 1.2.29, $bc = ac + dc$. Since d, c are positive, dc is positive by Remark 1.2.28. We conclude that $ac < bc$ by the definition of order, as desired. □

Corollary 1.2.32 (**Cancellation Law**)**.** *Let a, b, c be natural numbers such that $ac = bc$ and such that $c \neq 0$. Then $a = b$.*

Proof. From the trichotomy of order (Proposition 1.2.22), either $a < b$, $a > b$ or $a = b$. Since $c \neq 0$, c is positive. So, if $a < b$, then $ac < bc$ by Proposition 1.2.31. Similarly, if $b < a$, then $bc < ac$ by Proposition 1.2.31. So, the cases $a < b$ and $b < a$ cannot occur. We conclude that $a = b$, as desired. □

Remark 1.2.33. From now on, we will write $n{+}{+}$ as $n + 1$, and we will use basic properties of addition and multiplication of natural numbers.

Proposition 1.2.34 (**The Euclidean Algorithm**)**.** *Let n be a natural number and let q be a positive natural number. Then there exist natural numbers m, r such that $0 \leq r < q$ and such that $n = mq + r$.*

Remark 1.2.35. That is, we can divide n by q, leaving a remainder r, where $0 \leq r < q$.

Exercise 1.2.36. Prove Proposition 1.2.34 by fixing q and using induction on n.

1.3. **Integers.** We have dealt with addition and multiplication of natural numbers above. We would now like to deal with subtraction. In order to do this, we need to construct the integers. We will define the integers as the formal difference of two natural numbers. This is not the only way to define the integers, but it ends up being a bit cleaner than other methods.

Definition 1.3.1 (**Integers**). An **integer** is an expression of the form a——b where a, b are natural numbers. We say that two integers a——b and c——d are equal if and only if $a + d = c + b$. We let $\mathbb{Z}$ denote the set of all integers.

Example 1.3.2. So, the integer 5——2 is equal to 4——1 since $5 + 1 = 4 + 2$.

Remark 1.3.3. We need to verify that three axioms hold for this notion of equality. For any natural numbers a, b, c, d, e, f, we need to show:

(1) a——b is equal to a——b.
(2) If a——b is equal to c——d, then c——d is equal to a——b.
(3) If a——b is equal to c——d, and if c——d is equal to e——f, then a——b is equal to e——f.

These three axioms define an equivalence relation on integers. Properties (1) and (2) follow immediately. To show property (3), note that $a+d = b+c$, and $c+f = d+e$. Adding both equations, we get $a + d + c + f = b + c + d + e$. From the Cancellation Law (Proposition 1.2.14), we conclude that $a + f = b + e$, so that a——b is equal to e——f, as desired.

Definition 1.3.4 (**Addition and Multiplication of Integers**). Let a——b and c——d be two integers. We define the sum $(a$——$b) + (c$——$d)$ by

$$(a\text{——}b) + (c\text{——}d) := (a + c)\text{——}(b + d).$$

We define the product $(a$——$b) \times (c$——$d)$ by

$$(a\text{——}b) \times (c\text{——}d) := (ac + bd)\text{——}(ad + bc).$$

One potential problem with these definitions is that, even though 5——2=4——1, it is not clear that $(5$——$2)+(c$——$d) = (4$——$1)+(c$——$d)$, or that $(5$——$2) \times (c$——$d) = (4$——$1) \times (c$——$d)$. Fortunately, this is not a problem at all.

Lemma 1.3.5. *Let a, a', b, b', c, d be natural numbers such that a—b=a'—b'. Then*

(1) $(a$——$b) + (c$——$d) = (a'$——$b') + (c$——$d)$.
(2) $(a$——$b) \times (c$——$d) = (a'$——$b') \times (c$——$d)$.
(3) $(c$——$d) + (a$——$b) = (c$——$d) + (a'$——$b')$.
(4) $(c$——$d) \times (a$——$b) = (c$——$d) \times (a'$——$b')$.

Proof. We first prove (1). Using Definition 1.3.4, we need to show that $(a + c)$——$(b + d) = (a' + c)$——$(b' + d)$. Using Definition 1.3.1, we need to show that $a+c+b'+d = a'+c+b+d$. Since a——b=a'——b', we know that $a + b' = a' + b$. So, adding $c+d$ to both sides proves (1).

We now prove (2). Using Definition 1.3.4, we need to show that $(ac + bd)$——$(bc + ad) = (a'c + b'd)$——$(b'c + a'd)$. Using Definition 1.3.1, we need to show that $ac + bd + b'c + a'd = a'c + b'd + bc + ad$. The left side can be written $c(a + b') + d(a' + b)$, while the right is $c(a' + b) + d(a + b')$. Since a——b=a'——b', we know that $a + b' = a' + b$. So, both sides of (2) are equal. The remaining claims (3), (4) are proven similarly. □

Remark 1.3.6. Let n, m be any natural numbers. Then the set of integers n——0 behave exactly like the natural numbers. For example, $(n$——$0) + (m$——$0) = (n + m)$——0, and $(n$——$0) \times (m$——$0) = (nm)$——0. Also, $(n$——$0) = (m$——$0)$ if and only if $n = m$. So, we may identify the natural numbers as a subset of the integers via the correspondence $n = (n$——$0)$. Note in particular that under this correspondence, $0 = (0$——$0)$ and $1 = (1$——$0)$.

Remark 1.3.7. Then, for any integer x, we define $x{+}{+} := x + 1$.

Definition 1.3.8. Let $(a \text{---} b)$ be an integer. We define the **negation** $-(a \text{---} b)$ of $(a \text{---} b)$ by $-(a \text{---} b) := (b \text{---} a)$.

Remark 1.3.9. Negation is well-defined. That is, if $(a \text{---} b) = (a' \text{---} b')$, then $-(a \text{---} b) = -(a \text{---} b)$.

Definition 1.3.10. Let n be a natural number. We define $-n := -(n \text{---} 0) = (0 \text{---} n)$. If n is a positive natural number, we call $-n$ a **negative integer**.

Lemma 1.3.11. *Let x be an integer. Then exactly one of the following three statements is true.*

(1) *x is zero.*
(2) *There exists a positive natural number n such that $x = n$.*
(3) *There exists a positive natural number n such that $x = -n$.*

We also define the **subtraction** $(a \text{---} b) - (c \text{---} d)$ of integers by

$$(a \text{---} b) - (c \text{---} d) := (a \text{---} b) + (d \text{---} c).$$

Proposition 1.3.12. *Let x, y, z be integers. Then the following laws of algebra hold.*

- $x + y = y + x$ *(Commutativity of addition)*
- $(x + y) + z = x + (y + z)$ *(Associativity of addition)*
- $x + 0 = 0 + x = x$ *(Additive identity element)*
- $x + (-x) = (-x) + x = 0$ *(Additive inverse)*
- $xy = yx$ *(Commutativity of multiplication)*
- $(xy)z = x(yz)$ *(Associativity of multiplication)*
- $x1 = 1x = x$ *(Multiplicative identity element)*
- $x(y + z) = xy + xz$ *(Left Distributivity)*
- $(y + z)x = yx + zx$ *(Right Distributivity)*

Remark 1.3.13. These properties say that the integers form a **commutative ring**. Note that there is no notion of division within the integers. More specifically, there is no multiplicative inverse property. For example, given $2 \in \mathbb{Z}$, there does not exist an $x \in Z$ such that $2x = 1$. In order to have multiplicative inverses, we will need to enlarge the set of integers to the set of rational numbers. We will realize this goal shortly.

Proof of Associativity of addition. Let x, y, z be integers. Then there exist natural numbers a, b, c, d, e, f such that $x = a \text{---} b$, $y = c \text{---} d$ and such that $z = e \text{---} f$. We compute both sides of the purported inequality $(xy)z = x(yz)$, separately.

$$\begin{aligned}(xy)z &= [(a \text{---} b)(c \text{---} d)](e \text{---} f) = [(ac + bd) \text{---} (bc + ad)](e \text{---} f) \\ &= (ace + bde + bcf + adf) \text{---} (acf + bdf + bce + ade).\end{aligned}$$

$$\begin{aligned}x(yz) &= (a \text{---} b)[(c \text{---} d)(e \text{---} f)] = (a \text{---} b)[(ce + df) \text{---} (cf + de)] \\ &= (ace + adf + bcf + bde) \text{---} (bce + bdf + acf + ade).\end{aligned}$$

So, $(xy)z = x(yz)$ for all integers x, y, z, as desired. $\square$

Proposition 1.3.14. *Let a, b be integers such that $ab = 0$. Then at least one of a, b is zero.*

Exercise 1.3.15. Prove Proposition 1.3.14.

Corollary 1.3.16 (Cancellation Law). *Let a, b, c be integers such that $c \neq 0$ and such that $ac = bc$. Then $a = b$.*

Proof. Since $ac = bc$, we have $(a - b)c = ac - bc = 0$. Since $c \neq 0$, Proposition 1.3.14 implies that $a - b = 0$, so that $a = b$. □

We can now define the order on the integers exactly as we did for the natural numbers.

Definition 1.3.17 (Order). Let n, m be integers. We say that n **is greater than or equal to** m, and we write $n \geq m$ or $m \leq n$, if and only if $n = m + a$ for some natural number a. We say that n **is strictly greater than** m, and we write $n > m$ or $m < n$, if and only if $n \geq m$ and $n \neq m$.

Also, using Proposition 1.3.12, we have the following properties of order

Proposition 1.3.18 (Properties of Order). *Let a, b be integers.*

(1) *$a > b$ if and only if $a - b$ is a positive natural number.*
(2) *If $a > b$, then $a + c > b + c$ for any integer c.*
(3) *If $a > b$, then $ac > bc$ for any positive natural number c.*
(4) *If $a > b$, then $-a < -b$.*
(5) *If $a > b$ and $b > c$, then $a > c$.*
(6) *If $a \geq b$ and $b \geq a$, then $a = b$.*

1.4. **Rationals.** As discussed above, there does not exist an integer x such that $2x = 1$. That is, a general integer does not have a multiplicative inverse. In order to get multiplicative inverses for nonzero integers, we need to enlarge this set to the set of rational numbers. As above, we will define the rational numbers axiomatically.

Definition 1.4.1 (Rational Numbers). A **rational number** is an expression of the form $a//b$, where a, b are integers and $b \neq 0$. Two rational numbers $a//b$ and $c//d$ are considered to be equal if and only if $ad = cb$.

Remark 1.4.2. As before, we need to check that this notion of equality of rational numbers is an equivalence relation. It follows readily that $a//b$ is equal to $a//b$, and if $a//b$ is equal to $c//d$, then $c//d$ is equal to $a//b$. To check the third property, suppose $a//b$ is equal to $c//d$, and $c//d$ is equal to $e//f$. Then $ad = bc$ and $cf = de$. Multiplying both of these equations, we get $adcf = debc$. We need to show that $a//b$ is equal to $e//f$. That is, we need to show that $af = eb$. Since $d \neq 0$, from the Cancellation Law (Corollary 1.3.16), the equation $adcf = debc$ becomes $acf = ebc$. If $c \neq 0$, the Cancellation law implies that $af = eb$, as desired. If $c = 0$, then $ad = bc = 0$ and $de = cf = 0$. And since $b \neq 0$ and $d \neq 0$, Proposition 1.3.14 implies that $a = e = 0$. So, $af = 0 = eb$, as desired. In any case, we have proven that our notion of equality of rational numbers is an equivalence relation.

As before, we now define addition, multiplication, and negation of rational numbers. And we then need to check that these definitions are well-defined.

Definition 1.4.3. Let $a//b$ and $c//d$ be rational numbers. Define their **sum** as follows.

$$(a//b) + (c//d) = (ad + bc)//(db).$$

Define their **product** as follows.

$$(a//b) \times (c//d) := (ac)//(bd).$$

Define the **negation** of $a//b$ as follows.

$$-(a//b) := (-a)//b.$$

Lemma 1.4.4. *Let $a//b, a'//b', c//d$ be rational numbers such that $a//b$ is equal to $a'//b'$. Then the sum, product, and negation are unchanged when we replace $a//b$ with $a'//b'$. And similarly for $c//d$.*

Proof. We prove the first property, since the other proofs are similar. We need to show that $(a//b) + (c//d) = (a'//b') + (c//d)$. That is, we need to show that $(ad + bc)//(bd) = (a'd + b'c)//(b'd)$. That is, we need to show that $(ad + bc)(b'd) = (a'd + b'c)(bd)$, i.e. we need $ab'dd + bb'cd = a'bdd + bb'cd$, i.e. we need $ab'dd = a'bdd$. We know that $a//b = a'//b'$. That is, we know that $ab' = a'b$. So, the claim follows by multiplying both sides of this equation by dd, as desired. □

Remark 1.4.5. Let a, b be integers. The rational numbers $a//1, b//1$ behave exactly like the integers, since we have

$$(a//1) + (b//1) = (a + b)//1, \qquad (a//1) \times (b//1) = (ab)//1, \qquad -(a//1) = (-a)//1.$$

Also, $a//1 = b//1$ if and only if $a = b$. We therefore identify the rational numbers $a//1$ with the integers a by the relation $a = a//1$.

Remark 1.4.6. Let $a//b$ be a rational number. Then $a//b = 0//1$ if and only if $a = 0$. Taking the contrapositive, $a//b \neq 0//1$ if and only if $a \neq 0$.

Definition 1.4.7 (Reciprocal). Let $x = a//b$ be a nonzero rational number. From the previous remark and the definition of rational numbers, $a \neq 0$ and $b \neq 0$. We then define the **reciprocal** x^{-1} of x by $x^{-1} := b//a$. Note that if two rational numbers are equal, then their reciprocals are equal. Also, the reciprocal of 0 is left undefined.

Just as in the case of the integers, we can now prove various properties of the rationals. However, as promised, we now have an additional property. Nonzero numbers now have a multiplicative inverse. Whereas the integers were a commutative ring, the rationals are also a commutative ring. And with this additional multiplicative inverse property, the rationals are now referred to as a **field**.

Proposition 1.4.8. *Let x, y, z be rational numbers. Then the following laws of algebra hold.*

- $x + y = y + x$ *(Commutativity of addition)*
- $(x + y) + z = x + (y + z)$ *(Associativity of addition)*
- $x + 0 = 0 + x = x$ *(Additive identity element)*
- $x + (-x) = (-x) + x = 0$ *(Additive inverse)*
- $xy = yx$ *(Commutativity of multiplication)*
- $(xy)z = x(yz)$ *(Associativity of multiplication)*
- $x1 = 1x = x$ *(Multiplicative identity element)*
- $x(y + z) = xy + xz$ *(Left Distributivity)*
- $(y + z)x = yx + zx$ *(Right Distributivity)*

Finally, if x is nonzero, then

- $xx^{-1} = x^{-1}x = 1$ *(Multiplicative Inverse)*

Proof. We will only prove the associativity of addition, since the other proofs have a similar flavor. Write $x = a//b$, $y = c//d$, $z = e//f$. Then

$$\begin{aligned}(x+y)+z &= ((a//b)+(c//d))+e//f = ((ad+bc)//(bd))+e//f \\ &= (adf+bcf+bde)//(bde).\end{aligned}$$

$$\begin{aligned}x+(y+z) &= (a//b)+((c//d)+(e//f)) = (a//b)+((cd+de)//(df)) \\ &= (adf+bcf+bde)//(bde).\end{aligned}$$

So, $(x+y)+z = x+(y+z)$, as desired. □

Definition 1.4.9 (Quotient). Let x, y be rational numbers such that $y \neq 0$. We define the **quotient** x/y of x and y by

$$x/y := x \times y^{-1}.$$

Remark 1.4.10. For any integers a, b with $b \neq 0$, note that $a/b = a//b$, since

$$a/b = ab^{-1} = (a//1) \times (1//b) = a//b.$$

So, from now on, we use the notation a/b instead of $a//b$.

Remark 1.4.11. From now on, we will use the field axioms of Proposition 1.4.8 without explicit reference.

As in the case of integers, we now define positive and negative rational numbers.

Definition 1.4.12. A rational number x is said to be **positive** if and only if $x = a/b$ for some positive integers a, b. A rational number x is said to be **negative** if and only if $x = -y$ for a positive rational number y.

Remark 1.4.13. A positive integer is a positive rational number, and a negative integer is a negative rational number, so our notions of positive and negative are consistent.

Lemma 1.4.14. *Let x be a rational number. Then exactly one of the following three statements is true.*

- *x is equal to 0.*
- *x is a positive rational number.*
- *x is a negative rational number.*

We now define an order on the rationals that extends the notion of order on the integers.

Definition 1.4.15 (Order). Let x, y be rational numbers. We write $x > y$ if and only if $x - y$ is a positive rational number. We write $x < y$ if and only if $y - x$ is a positive rational number. We write $x \geq y$ if and only if either $x > y$ or $x = y$. We write $x \leq y$ if and only if either $x < y$ or $x = y$.

Proposition 1.4.16 (Properties of Order). *Let x, y, z be rational numbers. Then*

(1) *Exactly one of the statements $x = y$, $x < y$, $x > y$ is true.*
(2) *$x < y$ if and only if $y > x$.*
(3) *If $x < y$ and $y < z$, then $x < z$*
(4) *If $x < y$, then $x + z < y + z$.*

(5) *If $x < y$ and if z is positive, then $xz < yz$.*

Remark 1.4.17. The five properties of Proposition 1.4.16 combined with the field axioms of Proposition 1.4.8 say that the set of rational numbers $\mathbb{Q}$ form an **ordered field**.

Unlike the integers, the rationals have the following density property. Given any two rational numbers, there is a third rational number between them.

Proposition 1.4.18. *Given any two rational numbers x, z with $x < z$, there exists a rational number y such that $x < y < z$.*

Proof. Define $y := (x + z)/2$. Since $x < z$ and $1/2$ is positive, Proposition 1.4.16(5) says that $x/2 < z/2$. Adding $z/2$ to both sides and using Proposition 1.4.16(4), we get $x/2 + z/2 < z/2 + z/2 = z$. That is, $y < z$. Adding $x/2$ to both sides of $x/2 < z/2$, we get $x = x/2 + x/2 < x/2 + z/2$. That is, $x < y$. In conclusion, $x < y < z$, as desired. □

Even though the rationals have some density in the sense of Proposition 1.4.18, the set of rational numbers still has many gaps. To illustrate this fact, consider the following classical proposition.

Proposition 1.4.19. *There does not exist a rational number x such that $xx = 2$.*

Proof. We argue by contradiction. Assume that x is rational and $xx = 2$. We may assume that x is positive, since $xx = (-x)(-x)$. Let p, q be integers with $q \neq 0$ such that $x = p/q$. Since x is positive, we may assume that p, q are natural numbers. Since $xx = 2$, we have $pp = 2qq$. Recall that a natural number a is **even** if there exists a natural number b such that $a = 2b$, and a natural number a is **odd** if there exists a natural number b such that $a = 2b + 1$. Note that every natural number is either even or odd, and natural number cannot be both even and odd. Both of these facts follow from Proposition 1.2.34. If a is odd, note that $aa = 4bb + 2b + 2b + 1 = 2(2bb + b + b) + 1$, so aa is odd. So, by taking the contrapositive: if aa is even, then a is even. Since $pp = 2qq$, pp is even, so we conclude that p is even, so there exists a natural number k such that $p = 2k$. Since p is positive, k is positive. Since $pp = 2qq$, we get $pp = 4kk = 2qq$, so $qq = 2kk$. Since $pp = 2qq$, and p, q are positive, we have $q < p$.

In summary, we started with positive natural numbers p, q such that $pp = 2qq$. And we now have positive natural numbers q, k such that $qq = 2kk$, and such that $q < p$. We can therefore iterate this procedure. For any natural number n, suppose inductively we have p_n, q_n positive natural numbers such that $p_n p_n = 2q_n q_n$. Then we have found natural numbers p_{n+1}, q_{n+1} such that $p_{n+1}p_{n+1} = 2q_{n+1}q_{n+1}$, and such that $p_{n+1} < p_n$. The existence of the natural numbers $p_1, p_2, \ldots$ violates the principle of infinite descent (Exercise 1.4.20), so we have obtained a contradiction. We conclude that no rational x satisfies $xx = 2$. □

Exercise 1.4.20. Prove the principle of infinite descent. Let $p_0, p_1, p_2, \ldots$ be an infinite sequence of natural numbers such that $p_0 > p_1 > p_2 > \cdots$. Prove that no such sequence exists. (Hint: Assume by contradiction that such a sequence exists. Then prove by induction that for all natural numbers n, N, we have $p_n \geq N$. Use this fact to obtain a contradiction.)

1.4.1. *Operations on Rationals.* We now introduce a few additional operations on the rationals $\mathbb{Q}$. These operations will help in our construction of the real numbers.

Definition 1.4.21 (Absolute Value). Let x be a rational number. The **absolute value** $|x|$ of x is defined as follows. If $x \geq 0$, then $|x| := x$. If $x < 0$, then $|x| := -x$.

Definition 1.4.22 (Distance). Let x, y be rational numbers. The quantity $|x - y|$ is called the **distance between** x and y. We denote $d(x, y) := |x - y|$.

The following inequalities will be used very often in this course.

Proposition 1.4.23. *Let x, y be rational numbers. Then $|x| \geq 0$, and $|x| = 0$ if and only if $x = 0$. We also have the **triangle inequality***

$$|x + y| \leq |x| + |y|,$$

the bounds

$$-|x| \leq x \leq |x|$$

and the equality

$$|xy| = |x|\,|y|.$$

In particular,

$$|-x| = |x|.$$

Also, the distance $d(x, y)$ satisfies the following properties. Let x, y, z be rational numbers. Then $d(x, y) = 0$ if and only if $x = y$. Also, $d(x, y) = d(y, x)$. Lastly, we have the triangle inequality

$$d(x, z) \leq d(x, y) + d(y, z).$$

Exercise 1.4.24. By breaking into different cases as necessary, prove Proposition 1.4.23.

Exercise 1.4.25. Using the usual triangle inequality, prove the **reverse triangle inequality**: For any rational numbers x, y, we have $|x - y| \geq ||x| - |y||$.

Definition 1.4.26 (Exponentiation). Let x be a rational number. We define $x^0 := 1$. Now, let n be any natural number, and suppose we have inductively defined x^n. Then define $x^{n+1} := x^n \times x$.

The following properties of exponentiation then follow by induction.

Proposition 1.4.27. *Let x, y be rational numbers, and let n, m be natural numbers.*

- *$x^n x^m = x^{n+m}$, $(x^n)^m = x^{nm}$, and $(xy)^n = x^n y^n$.*
- *$x^n = 0$ if and only if $x = 0$ and $n > 0$.*
- *If $x \geq y \geq 0$, then $x^n \geq y^n \geq 0$.*
- *$|x|^n = |x^n|$.*

Definition 1.4.28 (Negative Exponentiation). Let x be a nonzero rational number, and let n be a positive natural number. Define $x^{-n} := 1/x^n$.

Proposition 1.4.29. *Let x, y be nonzero rational numbers, and let n, m be integers.*

- *$x^n x^m = x^{n+m}$, $(x^n)^m = x^{nm}$, and $(xy)^n = x^n y^n$.*
- *If $x \geq y > 0$, then $x^n \geq y^n > 0$ if $n > 0$, and $0 < x^n \leq y^n$ if $n < 0$.*
- *$|x|^n = |x^n|$.*

1.5. **Cauchy Sequences of Rationals.** Having established many properties of the rational numbers, we can finally begin to construct the real number system. As we saw in Proposition 1.4.19, there does not exist a rational number x such that $x^2 = 2$. Nevertheless, we can still find rational numbers x such that x^2 becomes as close as desired to 2. In this sense, the rational numbers have gaps between them. And filling in these gaps will exactly give us the real number system. There are a few different ways to fill in these gaps between the rational numbers. We will discuss the method of Cauchy sequences, since their investigation will lead naturally to further topics of interest.

As a preliminary result, we consider the gaps between the integers.

Proposition 1.5.1. *Let x be a rational number. Then there exists a unique integer n such that $n \leq x < n+1$. In particular, there exists an integer N such that $x < N$.*

Exercise 1.5.2. Using the Euclidean Algorithm (Proposition 1.2.34), prove Proposition 1.5.1.

Proposition 1.5.3. *For any rational number $\varepsilon > 0$, there exists a nonnegative rational number x such that $x^2 < 2 < (x+\varepsilon)^2$.*

Proof. We argue by contradiction. Suppose there exists $\varepsilon > 0$ and there does not exist a nonnegative rational number x such that $x^2 < 2 < (x+\varepsilon)^2$. So, every nonnegative rational number x with $x^2 < 2$ must also satisfy $(x+\varepsilon)^2 \leq 2$. From Proposition 1.4.19, $(x+\varepsilon)^2 \neq 2$, so $(x+\varepsilon)^2 < 2$. Note that $(x+\varepsilon)^2$ is rational and $(x+\varepsilon)^2 < 2$, so using this number in place of x, we see that we must have $(x+2\varepsilon)^2 < 2$ as well. Indeed, an inductive argument shows that, for any natural number n, $(x+n\varepsilon)^2 < 2$. Choosing $x = 0$, we see that $(n\varepsilon)^2 < 2$, for any natural number n. However, since $2/\varepsilon$ is rational, Proposition 1.5.1 says that there exists an integer N such that $N > 2/\varepsilon$. That is, $N\varepsilon > 2$, so $(N\varepsilon)^2 > 4$. This inequality contradicts that $(N\varepsilon)^2 < 2$. Since we have arrived at a contradiction, we conclude that an x exists satisfying the proposition. □

Indeed, we "know" that the sequence of rational numbers

$$1.4, \quad 1.41, \quad 1.414, \quad 1.4142, \quad \ldots$$

becomes arbitrarily close to a number x such that $x^2 = 2$. And this sort of sequential procedure is exactly how we will construct the rational numbers. Note that we define the decimal 1.4142 as the rational number 14142/10000.

Definition 1.5.4 (Sequence of rationals). Let m be an integer. A **sequence** $(a_n)_{n=m}^{\infty}$ **of rationals** is any function from the set $\{n \in \mathbb{N} \colon n \geq m\}$ to $\mathbb{Q}$. Informally, a sequence of rationals is an ordered list of rational numbers.

Example 1.5.5. The sequence $(n^2)_{n=0}^{\infty}$ is the collection $0, 1, 4, 9, 16, \ldots$ of natural numbers.

We will define real numbers as certain limits of sequences of rationals. A general sequence of rationals does not seem to have a sensible limit, so we need to restrict the sequences that we are considering. For example, the sequence $((-1)^n)_{n=0}^{\infty}$ does not seem to have any sensible limit. The following definition states precisely what kind of sequences we would like to focus on. The idea is that, eventually, the sequence elements need to be close to each other. This vague statement is then formalized as follows.

Definition 1.5.6 (**Cauchy sequence**)**.** A sequence $(a_n)_{n=0}^{\infty}$ of rational numbers is said to be a **Cauchy sequence** if and only if, for every rational $\varepsilon > 0$, there exists a natural number $N = N(\varepsilon)$ such that, for all $j, k \geq N$, we have $d(a_j, a_k) < \varepsilon$.

Example 1.5.7. The sequence $(1/n)_{n=1}^{\infty}$ is a Cauchy sequence. To see this, let $\varepsilon > 0$ be a rational number. From Proposition 1.5.1, let N be a natural number such that $N > 2/\varepsilon$. Then $1/N < \varepsilon/2$. Now, let $j, k \geq N$ so that $1/j \leq 1/N$ and $1/k \leq 1/N$. From the triangle inequality, we then have

$$d(1/j, 1/k) = |1/j - 1/k| \leq |1/j| + |1/k| = 1/j + 1/k \leq 2/N < \varepsilon.$$

To get an idea of where we are headed, we are going to *define* the real numbers to be the "limits" of Cauchy sequences. In order to make this statement rigorous, we need to show that a Cauchy sequence has a limit, and we need to discuss when two Cauchy sequences have the same limit. If two Cauchy sequences have the same limit, we will say that they are equal. Before defining the real numbers, we need some preliminary facts about Cauchy sequences.

Definition 1.5.8 (**Bounded Sequence**)**.** Let $M \geq 0$ be rational. A finite sequence of rationals $a_0, \ldots, a_n$ is **bounded by** M if and only if $|a_i| \leq M$ for all $i \in \{0, \ldots, n\}$. An infinite sequence of rationals $(a_i)_{i=0}^{\infty}$ is **bounded by** M if and only if $|a_i| \leq M$ for all $i \in \mathbb{N}$. A sequence $(a_i)_{i=0}^{\infty}$ is **bounded** if and only if there exists a positive rational M such that $(a_i)_{i=0}^{\infty}$ is bounded by M.

Lemma 1.5.9. *Every Cauchy sequence is bounded.*

Exercise 1.5.10. Prove Lemma 1.5.9

Definition 1.5.11 (**Equivalent Cauchy Sequences**)**.** Let $(a_n)_{n=0}^{\infty}, (b_n)_{n=0}^{\infty}$ be Cauchy sequences. We say that these Cauchy sequences are **equivalent** if and only if, for every rational $\varepsilon > 0$, there exists a natural number $N = N(\varepsilon) \geq 0$ such that $|a_n - b_n| < \varepsilon$ for all $n \geq N$.

As with our notations of equivalence of integers and rationals, we need to show that this notion of equivalence is an equivalence relation. That is, we need the following three properties.

Lemma 1.5.12. *Let $(a_n)_{n=0}^{\infty}, (b_n)_{n=0}^{\infty}, (c_n)_{n=0}^{\infty}$ be Cauchy sequences.*
- *$(a_n)_{n=0}^{\infty}$ is equivalent to $(a_n)_{n=0}^{\infty}$.*
- *If $(a_n)_{n=0}^{\infty}$ is equivalent to $(b_n)_{n=0}^{\infty}$, then $(b_n)_{n=0}^{\infty}$ is equivalent to $(a_n)_{n=0}^{\infty}$.*
- *If $(a_n)_{n=0}^{\infty}$ is equivalent to $(b_n)_{n=0}^{\infty}$, and if $(b_n)_{n=0}^{\infty}$ is equivalent to $(c_n)_{n=0}^{\infty}$, then $(a_n)_{n=0}^{\infty}$ is equivalent to $(c_n)_{n=0}^{\infty}$.*

Proof. We prove the third item. Let $\varepsilon > 0$ be a rational number. Note that $\varepsilon/2 > 0$ is a rational number. So, by assumption, there exist $L, M > 0$ such that, for all $n \geq L$, $|a_n - b_n| < \varepsilon/2$, and for all $n \geq M$, $|b_n - c_n| < \varepsilon/2$. Define $N := \max(L, M)$. Then, for all $n \geq N$, we have by the triangle inequality

$$|a_n - c_n| = |a_n - b_n + b_n - c_n| \leq |a_n - b_n| + |b_n - c_n| < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

That is, $(a_n)_{n=0}^{\infty}$ is equivalent to $(c_n)_{n=0}^{\infty}$, as desired. ☐

Remark 1.5.13. The above proof strategy occurs very often in analysis, so it should be ingrained in your memory. The idea is that, in order to prove that two things are close, you add and subtract the same number, and then apply the triangle inequality.

1.6. **Construction of the Real Numbers.** We can now finally give a definition of a real number. As in our construction of the integers and rational numbers, we will begin by using some artificial symbol to designate a real number. However, the construction of the real numbers requires a new ingredient, which is the Cauchy sequence of rational numbers.

Definition 1.6.1 (**Real Number**)**.** A **real number** is an object of the form $\mathrm{LIM}_{n\to\infty}a_n$, where $(a_n)_{n=0}^{\infty}$ is a Cauchy sequence. Two real numbers $\mathrm{LIM}_{n\to\infty}a_n$, $\mathrm{LIM}_{n\to\infty}b_n$ are equal if and only if $(a_n)_{n=0}^{\infty}, (b_n)_{n=0}^{\infty}$ are equivalent Cauchy sequences. The set of all real numbers is denoted by $\mathbb{R}$

Remark 1.6.2. We refer to $\mathrm{LIM}_{n\to\infty}a_n$ as the **formal limit** of the Cauchy sequence $(a_n)_{n=0}^{\infty}$. Later on, we will show that a Cauchy sequence has an actual limit as $n \to \infty$, which explains our use of this notation.

Even though we define real numbers in terms of Cauchy sequences, which allows us to axiomatize the real number system and prove facts about this system, our approach perhaps does not have many direct consequences for other results concerning real numbers and functions. To use an analogy, even though we know that all materials in the world are made of atoms, this fact only marginally affects our material interaction with the physical world. On the other hand, the exact way that we construct and analyze the real numbers *does* influence our understanding of other mathematical objects. To use the same analogy as before, our understanding of atoms *does* allow us to better understand some things that we encounter in the physical world, such as light, the sun, etc.

As in our treatment of the integers and rationals, we now define arithmetic on the real numbers.

Definition 1.6.3 (**Addition of Real Numbers**)**.** Let $x = \mathrm{LIM}_{n\to\infty}a_n$ and let $y = \mathrm{LIM}_{n\to\infty}b_n$ be real numbers. Then define the sum of x and y by $x + y := \mathrm{LIM}_{n\to\infty}(a_n + b_n)$.

We now check that addition of two real numbers give a real number, and that addition is well-defined.

Lemma 1.6.4. *Let $x = \mathrm{LIM}_{n\to\infty}a_n$ and let $y = \mathrm{LIM}_{n\to\infty}b_n$ be real numbers. Then $x + y$ is also a real number.*

Proof. We need to show that $(a_n + b_n)_{n=0}^{\infty}$ is a Cauchy sequence. The proof is similar to that of Lemma 1.5.12. Let $\varepsilon > 0$ be a rational number. Note that $\varepsilon/2 > 0$ is a rational number. By assumption, there exist $L, M > 0$ such that, for all $j, k \geq L$, $|a_j - a_k| < \varepsilon/2$, and for all $j, k \geq M$, $|b_j - b_k| < \varepsilon/2$. Define $N := \max(L, M)$. Then, for all $j, k \geq N$, we have by the triangle inequality

$$|a_j + b_j - a_k - b_k| = |a_j - a_k + b_j - b_k| \leq |a_j - a_k| + |b_j - b_k| < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

That is, $(a_n + b_n)_{n=0}^{\infty}$ is a Cauchy sequence, as desired. □

Lemma 1.6.5. *Let $x = \mathrm{LIM}_{n\to\infty}a_n$ and let $y = \mathrm{LIM}_{n\to\infty}b_n$ be real numbers. Let $x' = \mathrm{LIM}_{n\to\infty}a_n'$ be a real number such that $x = x'$. Then $x + y = x' + y$.*

Proof. Let $\varepsilon > 0$ be a rational number. Since $x = x'$, there exists $N > 0$ such that, for all $n \geq N$, $|a_n - a_n'| < \varepsilon$. Then, for all $n \geq N$,

$$|a_n + b_n - a_n' - b_n| = |a_n - a_n'| < \varepsilon.$$

That is, $(a_n + b_n)_{n=0}^{\infty}$ is equivalent to $(a_n' + b_n)_{n=0}^{\infty}$, as desired. □

Remark 1.6.6. If additionally y' is equivalent to y, then $x + y = x + y'$. To see this, note that addition is commutative for real numbers, which follows from the commutativity of addition for rational numbers.

We now define multiplication.

Definition 1.6.7 (Multiplication of Real Numbers). Let $x = \text{LIM}_{n\to\infty}a_n$ and let $y = \text{LIM}_{n\to\infty}b_n$ be real numbers. Define the product $xy := \text{LIM}_{n\to\infty}(a_n b_n)$.

Proposition 1.6.8. *Let $x = \text{LIM}_{n\to\infty}a_n$ and let $y = \text{LIM}_{n\to\infty}b_n$ be real numbers. Then xy is a real number. Also if $x' = \text{LIM}_{n\to\infty}a'_n$ is a real number such that $x = x'$, then $xy = x'y$.*

Exercise 1.6.9. Prove Proposition 1.6.8.

Remark 1.6.10. We can now realize the rational numbers as a subset of the real numbers. Given a rational number $q \in \mathbb{Q}$, consider the constant Cauchy sequence $q, q, q, q, \ldots$. Then addition and multiplication are identical for $q \in \mathbb{Q}$ and for the Cauchy sequence $q, q, q, q, \ldots$. Moreover, this identification of rational numbers within the real numbers is consistent with our two notions of equality. That is, $p, q \in \mathbb{Q}$ are equal if and only if the Cauchy sequences $p, p, p, \ldots$ and $q, q, q, \ldots$ are equal.

Definition 1.6.11. Since we have defined multiplication of real numbers, we can now define the **negation** of a real number x by

$$-x := (-1) \times x.$$

We therefore see that

$$-(\text{LIM}_{n\to\infty}a_n) = \text{LIM}_{n\to\infty}(-a_n).$$

Also, we define **subtraction** of real numbers x, y by

$$x - y := x + (-y).$$

We therefore see that

$$\text{LIM}_{n\to\infty}a_n - (\text{LIM}_{n\to\infty}b_n) = \text{LIM}_{n\to\infty}(a_n - b_n).$$

We will now show that the real number system satisfies all of the usual algebraic identities with which we are acquainted. That is, the number system $\mathbb{R}$ is a **field**. The final property of the field, the multiplicative inverse, is a bit tricky to verify, so we will deal with that last. That is, we will first only assert that $\mathbb{R}$ is a commutative ring.

Proposition 1.6.12. *Let x, y, z be real numbers. Then the following laws of algebra hold.*

- $x + y = y + x$ *(Commutativity of addition)*
- $(x + y) + z = x + (y + z)$ *(Associativity of addition)*
- $x + 0 = 0 + x = x$ *(Additive identity element)*
- $x + (-x) = (-x) + x = 0$ *(Additive inverse)*
- $xy = yx$ *(Commutativity of multiplication)*
- $(xy)z = x(yz)$ *(Associativity of multiplication)*
- $x1 = 1x = x$ *(Multiplicative identity element)*
- $x(y + z) = xy + xz$ *(Left Distributivity)*
- $(y + z)x = yx + zx$ *(Right Distributivity)*

Proof. We only prove the associativity of a multiplication, the others being similar. As we will see, these properties follow readily from the corresponding properties of the rational numbers. Let x, y, z be real numbers. Write $x = \mathrm{LIM}_{n\to\infty} a_n$, $y = \mathrm{LIM}_{n\to\infty} b_n$, $z = \mathrm{LIM}_{n\to\infty} c_n$. Then $(xy) = \mathrm{LIM}_{n\to\infty}(a_n b_n)$, and $(xy)z = \mathrm{LIM}_{n\to\infty}[(a_n b_n)c_n]$. From associativity of multiplication of rationals, we then have

$$(xy)z = \mathrm{LIM}_{n\to\infty}[a_n(b_n c_n)] = x \times \mathrm{LIM}_{n\to\infty}(b_n c_n) = x(yz),$$

as desired. □

We now need to define the reciprocal. Note that we cannot simply define the reciprocal of a Cauchy sequence $a_0, a_1, \ldots$ to be the sequence a_0^{-1}, a_1^{-1}, since some of the elements of the sequence $a_0, a_1, \ldots$ could be zero. Thankfully, this problem can be circumvented by simply waiting for the Cauchy sequence to be nonzero.

Lemma 1.6.13. *Let x be a nonzero real number. Then there exists a rational number $\varepsilon > 0$ such that, for any Cauchy sequence $(a_n)_{n=0}^{\infty}$ with $x = \mathrm{LIM}_{n\to\infty} a_n$, there exists $N > 0$ such that, for all $n \geq N$, $|a_n| > \varepsilon$. In this statement, note that ε does not depend on the Cauchy sequence, but N does.*

Proof. Since x is nonzero, $(a_n)_{n=0}^{\infty}$ is not equivalent to the Cauchy sequence $0, 0, 0, \ldots$. So, negating the statement "$(a_n)_{n=0}^{\infty}$ is equivalent to $0, 0, 0, \ldots$, " we get the following. There exists a rational $\varepsilon > 0$ such that, for all natural numbers $L > 0$, there exists $\ell > L$ such that $|a_\ell| \geq 3\varepsilon$. Since $(a_n)_{n=0}^{\infty}$ is a Cauchy sequence, there exists $M > 0$ such that, for all $j, k > M$, we have $|a_j - a_k| < \varepsilon$. So, if we choose $L := M$, there exists $\ell > L = M$ such that $|a_\ell| \geq 3\varepsilon$. So, for any $n > \ell > M$, we have by Exercise 1.4.25

$$|a_n| = |a_n - a_\ell + a_\ell| \geq |a_\ell| - |a_n - a_\ell| > 3\varepsilon - \varepsilon = 2\varepsilon.$$

So, the assertion is proven with an ε that may depend on the chosen Cauchy sequence $(a_n)_{n=0}^{\infty}$. To see that we can choose ε to not depend on the particular Cauchy sequence, let $(a_n')_{n=0}^{\infty}$ be any Cauchy sequence equivalent to $(a_n)_{n=0}^{\infty}$. That is, there exists $K > 0$ such that, for all $n > K$, we have $|a_n - a_n'| < \varepsilon$. Finally, define $N := \max(\ell, K)$. Then, for any $n > N$, we have

$$|a_n'| = |a_n' - a_n + a_n| \geq |a_n| - |a_n - a_n'| \geq 2\varepsilon - \varepsilon = \varepsilon.$$

Since $(a_n')_{n=0}^{\infty}$ is any Cauchy sequence equivalent to $(a_n)_{n=0}^{\infty}$, we have shown that the number ε does not depend on the particular Cauchy sequence, as desired. □

With this lemma, we can now define the inverse of a real number.

Definition 1.6.14 (Inverse). Let x be a nonzero real number. Let $(a_n)_{n=0}^{\infty}$ be any Cauchy sequence with $x = \mathrm{LIM}_{n\to\infty} a_n$. From Lemma 1.6.13, there exists a rational $\varepsilon > 0$ and a natural number $N > 0$ such that, for all $n > N$, $|a_n| > \varepsilon > 0$. Consider the equivalent Cauchy sequence b_n where $b_n := a_n$ for all $n > N$, and $b_n := 1$ for all $0 \leq n \leq N$. Then $x = \mathrm{LIM}_{n\to\infty} b_n$, and $|b_n| > \varepsilon$ for all $n \geq 0$. So, we define the **reciprocal** x^{-1} of x as $x^{-1} := \mathrm{LIM}_{n\to\infty}(b_n^{-1})$.

We now need to check that x^{-1} is a real number, and also that x^{-1} is well-defined. That is, we need to show that x^{-1} does not depend on the Cauchy sequence $(a_n)_{n=0}^{\infty}$.

Lemma 1.6.15. *Let $\delta > 0$. Let $(a_n)_{n=0}^{\infty}$ be a Cauchy sequence such that $|a_n| > \delta$ for all $n \geq 0$. Then $(a_n^{-1})_{n=0}^{\infty}$ is a Cauchy sequence.*

Proof. Let $\varepsilon > 0$. Since $|a_n| > \delta > 0$ for all $n \geq 0$, we have $|a_n|^{-1} < 1/\delta$ for all $n \geq 0$. Since $(a_n)_{n=0}^{\infty}$ is a Cauchy sequence, there exists $N > 0$ such that, for all $j, k > N$, we have $|a_j - a_k| < \varepsilon\delta^2$. Then, for all $j, k > N$, we have

$$\left|a_j^{-1} - a_k^{-1}\right| = |a_j|^{-1} |a_k|^{-1} |a_k - a_j| < \delta^{-2}\varepsilon\delta^2 = \varepsilon.$$

That is, the sequence $(a_n^{-1})_{n=0}^{\infty}$ is a Cauchy sequence. □

Lemma 1.6.16. *Let x be a nonzero real number. Let $(a_n)_{n=0}^{\infty}$ and $(a_n')_{n=0}^{\infty}$ be Cauchy sequences such that $x = \mathrm{LIM}_{n\to\infty} a_n$ and such that $x = \mathrm{LIM}_{n\to\infty} a_n'$. Then, after changing a finite number of terms of these Cauchy sequences, we have: $\mathrm{LIM}_{n\to\infty} a_n^{-1}$ is equivalent to $\mathrm{LIM}_{n\to\infty}(a_n')^{-1}$.*

Proof. Let $\varepsilon > 0$. From Lemma 1.6.13, let $\delta > 0$ and let $L > 0$ such that, for all $n > L$, $|a_n| > \delta$ and $|a_n'| > \delta$. Since $(a_n)_{n=0}^{\infty}$ and $(a_n')_{n=0}^{\infty}$ are equivalent, there exists $M > 0$ such that, for all $n > M$, we have $|a_n - a_n'| < \varepsilon\delta^2$. Define $N := \max(L, M)$. Then, for all $n > N$,

$$\left|a_n^{-1} - (a_n')^{-1}\right| = |a_n|^{-1} |a_n'|^{-1} |a_n - a_n'| < \delta^{-2}\varepsilon\delta^2 = \varepsilon.$$

So, if we define $b_n := a_n$ for all $n > N$, $b_n' := a_n'$ for all $n \geq N$, and $b_n = b_n' = 1$ for all $0 \leq n \leq N$, we see that $\mathrm{LIM}_{n\to\infty} b_n^{-1}$ is equivalent to $\mathrm{LIM}_{n\to\infty}(b_n')^{-1}$, as desired. □

Lemma 1.6.15 shows that x^{-1} is a real number whenever x is a nonzero real number. And Lemma 1.6.16 shows that x^{-1} is well-defined.

Remark 1.6.17. If x is a nonzero real number, it follows from Definition 1.6.14 that $xx^{-1} = x^{-1}x = 1$. Combining this fact with Proposition 1.6.12, we conclude that $\mathbb{R}$ is a field, as previously asserted.

Remark 1.6.18. Note that our definition of reciprocal is consistent with the definition of reciprocal of a rational number.

Definition 1.6.19 (Division). Let x, y be real numbers with y nonzero. We then define $x/y := x \times y^{-1}$. We then have the **cancellation law** (which follows from the same property for rational numbers). If x, y, z are real numbers with z nonzero, and if $xz = yz$, then $x = y$.

Remark 1.6.20. We now have all of the usual arithmetic operations on the real numbers. We now turn to the order properties of the reals. Note that we cannot simply say that: a Cauchy sequence is positive if and only if its elements are all positive. For example, the Cauchy sequence $-1, 1, 1, 1, 1, 1, 1, 1, \ldots$ corresponds to the positive real number 1, but it has a negative value in the sequence. For another example, note that the Cauchy sequence $1, 1/2, 1/3, 1/4, 1/5, \ldots$ has all positive elements, but it is equivalent to the sequence $0, 0, 0 \ldots$, which is certainly not positive. So, we need to be careful in defining positivity.

1.6.1. *Ordering of the Reals.*

Definition 1.6.21. A real number x is said to be **positive** if and only if there exists a positive rational $\varepsilon > 0$ such that, for any Cauchy sequence $(a_n)_{n=0}^{\infty}$ with $x = \mathrm{LIM}_{n\to\infty}(a_n)$, there exists a natural number $N > 0$ such that, for all $n > N$, we have $a_n > \varepsilon > 0$. A real number x is said to be **negative** if and only if $-x$ is positive.

Remark 1.6.22. Note that these definitions are consistent with the definitions of positivity and negativity for rational numbers. For example, if $x > 0$ is rational, then Lemma 1.6.13 implies that there exists $\varepsilon > 0$ such that, for any Cauchy sequence $(a_n)_{n=0}^{\infty}$ with $x = \mathrm{LIM}_{n\to\infty} a_n$ there exists $N > 0$ such that for all $n > N$, $a_n > \varepsilon > 0$. (You will investigate the details of this argument in Exercise 1.6.31.)

Proposition 1.6.23. *For every real number x, exactly one of the following statements is true: x is positive, x is negative, or x is zero. If x, y are positive real numbers, then $x + y$ is positive, and xy is positive.*

Exercise 1.6.24. Using Lemma 1.6.13, prove Proposition 1.6.23

We can now define order, since we have just defined positivity and negativity.

Definition 1.6.25. Let x, y be real numbers. We say that x **is greater than** y, and we write $x > y$ if and only if $x - y$ is a positive real number. We say that x is **less than** y, and we write $x < y$ if and only if $y - x$ is a positive real number. We write $x \geq y$ if and only if $x > y$ or $x = y$, and we similarly define $x \leq y$.

Remark 1.6.26. This ordering on the reals is consistent with the ordering we gave for the rational numbers. That is, if a, b are two rational numbers with $a < b$, then the real numbers a, b also satisfy $a < b$. And similarly for the assertion $a > b$.

The real numbers now satisfy all of the same axioms for order than the rational numbers satisfied in Proposition 1.4.16.

Proposition 1.6.27 (Properties of Order). *Let x, y, z be real numbers. Then*

(1) *Exactly one of the statements $x = y$, $x < y$, $x > y$ is true.*
(2) *$x < y$ if and only if $y > x$.*
(3) *If $x < y$ and $y < z$, then $x < z$*
(4) *If $x < y$, then $x + z < y + z$.*
(5) *If $x < y$ and if z is positive, then $xz < yz$.*

Remark 1.6.28. In conclusion, the real numbers form an **ordered field**.

Proof. We only prove (5), since the other proofs similarly follow from Proposition 1.6.23 and basic algebra. Suppose $x < y$ and z is positive. Since $x < y$, $y - x$ is positive. So, from Proposition 1.6.23, $z(y - x)$ is positive, so $xz < yz$, as desired. □

Proposition 1.6.29. *Let x be a positive real number. Then x^{-1} is also a positive real number. If y is a positive real number with $x > y$, then $x^{-1} < y^{-1}$.*

Proof. Let x be a positive real number. Since $xx^{-1} = 1$, the real number x^{-1} is nonzero. (If we had $x^{-1} = 0$, then $xx^{-1} = 0$.) We show that x^{-1} is positive by contradiction. If x^{-1} were not positive, it would be negative, since $x^{-1} \neq 0$. From Proposition 1.6.23, we get that xx^{-1} is negative, contradicting that $xx^{-1} = 1$. We therefore conclude that x^{-1} is positive.

We now show that $x^{-1} < y^{-1}$ by contradiction. Assume that $x^{-1} \geq y^{-1}$. Then from Proposition 1.6.27(5) applied twice, $xx^{-1} \geq xy^{-1} > yy^{-1}$, i.e. $1 > 1$, a contradiction. We conclude that $x^{-1} < y^{-1}$, as desired. □

Proposition 1.6.30. *Let x, y be real numbers. Suppose $(a_n)_{n=0}^{\infty}, (b_n)_{n=0}^{\infty}$ are Cauchy sequences with $x = \mathrm{LIM}_{n\to\infty} a_n$ and $y = \mathrm{LIM}_{n\to\infty} b_n$. Assume that there exists $N > 0$ such that for all $n > N$, we have $a_n \leq b_n$. Then $x \leq y$.*

Proof. We argue by contradiction. Suppose $x > y$. Then $x - y$ is positive. Note that $(a_n - b_n)_{n=0}^{\infty}$ is a Cauchy sequence such that $x - y = \mathrm{LIM}_{n\to\infty}(a_n - b_n)$. So, by Definition 1.6.21, there exists $\delta > 0$ and there exists $M > 0$ such that, for all $n > M$, we have $a_n - b_n > \delta > 0$. In particular, we have $a_{M+1} > b_{M+1}$, a contradiction. Since we have achieved a contradiction, we are done. □

Exercise 1.6.31. Prove the following variant of Lemma 1.6.13: Let x be a positive real number. Then there exists a rational number $\varepsilon > 0$ such that, for any Cauchy sequence $(a_n)_{n=0}^{\infty}$ with $x = \mathrm{LIM}_{n\to\infty}a_n$, there exists $N > 0$ such that, for all $n \geq N$, $a_n > \varepsilon$. In this statement, note that ε does not depend on the Cauchy sequence, but N does. (And similarly, when x is a negative real number.)

Remark 1.6.32. Since we have defined positive and negative real numbers, we can then define the absolute value $|x|$ exactly as in Definition 1.4.21. We then define $d(x, y) := |x - y|$ just as before, but now for real numbers x, y. Note that, if $(a_n)_{n=0}^{\infty}$ is a Cauchy sequence such that $x = \mathrm{LIM}_{n\to\infty}a_n$, then $|a_n|$ is a Cauchy sequence for $|x|$, by Exercise 1.6.31.

Theorem 1.6.33 (**Triangle Inequality for Real Numbers**). *Let x, y be real numbers. Then $|x + y| \leq |x| + |y|$.*

Proof. Suppose $(a_n)_{n=0}^{\infty}, (b_n)_{n=0}^{\infty}$ are Cauchy sequences with $x = \mathrm{LIM}_{n\to\infty}a_n$, $y = \mathrm{LIM}_{n\to\infty}b_n$. From the triangle inequality for rational numbers (Proposition 1.4.23), $|a_n + b_n| \leq |a_n| + |b_n|$ for all $n \in \mathbb{N}$. By Remark 1.6.32, note that $(|a_n|)_{n=0}^{\infty}$ is a Cauchy sequence for $|x|$, and $(|b_n|)_{n=0}^{\infty}$ is a Cauchy sequence for $|y|$, and $(|a_n + b_n|)_{n=0}^{\infty}$ is a Cauchy sequence for $|x + y|$. Since $|a_n + b_n| \leq |a_n| + |b_n|$ for all $n \in \mathbb{N}$, Proposition 1.6.30 implies $|x + y| \leq |x| + |y|$. □

Theorem 1.6.34 (**The Rationals are Dense in the Real Numbers**). *Let x be a real number and let $\varepsilon > 0$ be any rational number. Then there exists a rational number y such that $|x - y| < \varepsilon$.*

Exercise 1.6.35. Prove Theorem 1.6.34.

Theorem 1.6.36 (**Archimedean Property**). *Let x, ε be any positive real numbers. Then there exists a positive integer N such that $N\varepsilon > x$.*

Proof. From Propositions 1.6.29 and 1.6.23, ε/x is a positive real number. Let $(a_n)_{n=0}^{\infty}$ be a Cauchy sequence of rationals such that $\varepsilon/x = \mathrm{LIM}_{n\to\infty}a_n$. From Exercise 1.6.31, there exists a rational number y and there exists a natural number M such that, for all $n > M$, we have $a_n > y > 0$. Write $y = p/q$ with $p, q \in \mathbb{N}$, $p \neq 0$, $q \neq 0$. Then $a_n > y \geq 1/q > 0$, so $\varepsilon/x \geq 1/q$ by Proposition 1.6.30, so $(q + 1)\varepsilon > x$. Setting $N := q + 1$ completes the proof. □

Corollary 1.6.37. *Let x, z be real numbers with $x < z$. Then there exists a rational number y with $x < y < z$.*

Exercise 1.6.38. Using Theorems 1.6.34 and 1.6.36, prove Corollary 1.6.37.

1.7. **The Least Upper Bound Property.** We have constructed the real numbers, defined their arithmetic operations, and proven a few basic properties of the real numbers. We can now finally describe some of the useful properties of the real numbers. The least upper bound property is the first such property. It will give a rigorous statement to the intuition that the real numbers "have no gaps" between them. We will see more rigorous statements of this intuition within our discussion of limits and completeness.

Definition 1.7.1 (Upper bound). Let E be a subset of $\mathbb{R}$, and let M be a real number. We say that M is an **upper bound** for E if and only if for every x in E, we have $x \leq M$.

Example 1.7.2. The set $\{t \in \mathbb{R}\colon 0 \leq t \leq 1\}$ has an upper bound of 1. The set $\{t \in \mathbb{R}\colon t > 0\}$ has no upper bound.

Definition 1.7.3 (Least upper bound). Let E be a subset of $\mathbb{R}$, and let M be a real number. We say that M is a **least upper bound** for E if and only if: M is an upper bound for E, and any other upper bound M' of E satisfies $M \leq M'$.

Example 1.7.4. The set $\{t \in \mathbb{R}\colon 0 \leq t \leq 1\}$ has a least upper bound of 1.

Proposition 1.7.5. *Let E be a subset of $\mathbb{R}$. Then E has at most one least upper bound.*

Proof. Let M, M' be two least upper bounds for E. We will show that $M = M'$. From Definition 1.7.3 applied to M, we have $M \leq M'$. From Definition 1.7.3 applied to M', we have $M' \leq M$. Therefore, $M = M'$. $\square$

The following Theorem is taken as an axiom in the book. However, it can instead be proven from our construction of the real numbers. The proof is a bit long, so it could be skipped on a first reading.

Theorem 1.7.6 (Least Upper Bound Property). *Let E be a nonempty subset of $\mathbb{R}$. If E has some upper bound, then E has exactly one least upper bound.*

Proof. From Proposition 1.7.5, E has at most one least upper bound. We therefore need to show that E has at least one least upper bound. In order to find the least upper bound for E, we will construct a Cauchy sequence of rational numbers which come very close to the least upper bound of E.

Let M be an upper bound for E. Let $x_0 \in E$, and let n be a positive integer. From the Archimedean property (Theorem 1.6.36), there exists $K \in \mathbb{N}$ such that $x_0 + K/n > M$. That is, $x_0 + K/n$ is an upper bound for E. Since $x_0 \in E$, $x_0 - 1/n$ is not an upper bound for E. So, there exists an integer i with $0 \leq i \leq K$ such that $x_0 + i/n$ is an upper bound for E, though $x_0 + (i-1)/n$ is not an upper bound for E. To see that i exists, just let i be the smallest natural number such that $x_0 + i/n$ is an upper bound for E.

Note that $x_0 + (i-1)/n < x_0 + i/n$. From Corollary 1.6.37, there exists a rational number a_n such that

$$x_0 + (i-1)/n < a_n < x_0 + i/n.$$

Therefore, $a_n + 1/n$ is an upper bound for E since $a_n + 1/n > x_0 + i/n$, but $a_n - 1/n$ is not an upper bound for E since $a_n - 1/n < x_0 + (i-1)/n$.

Consider the sequence of rational numbers $(a_n)_{n=0}^{\infty}$. We will show that this sequence is a Cauchy sequence. Let n, m be positive integers. Then $a_n + 1/n$ is always an upper bound for E, while $a_m - 1/m$ is not an upper bound for E. Therefore, $a_n + 1/n > a_m - 1/m$. Similarly, $a_m + 1/m > a_n - 1/n$. Therefore, for all positive integers n, m,

$$-1/n - 1/m < a_n - a_m < 1/n + 1/m.$$

In particular, for any positive integer N, we have for all $n, m \geq N$,

$$-2/N < a_n - a_m < 2/N. \qquad (*)$$

Let $\varepsilon > 0$ be a rational number. From the Archimedean property (Theorem 1.6.36), there exists a positive integer N such that $N\varepsilon > 2$, so that $0 < 2/N < \varepsilon$. So, for any rational number ε, there exists a positive integer N such that, for all $n, m \geq N$, we have

$$-\varepsilon < a_n - a_m < \varepsilon.$$

So, $(a_n)_{n=0}^{\infty}$ is a Cauchy sequence.

Define $x := \mathrm{LIM}_{n\to\infty} a_n$. We will show that x is a least upper bound of E. We first show that x is an upper bound for E. Setting $m = N$ in $(*)$, we get that, for all $n \geq N$,

$$-2/N < a_n - a_N < 2/N.$$

So, from Proposition 1.6.30, for all positive integers N,

$$-2/N \leq x - a_N \leq 2/N. \qquad (**)$$

Let $y \in E$. For each positive integer N, recall that $a_N + 1/N$ is an upper bound for E. So, $y \leq a_N + 1/N$. From $(**)$, $-2/N \leq x - a_N$, so adding these two inequalities, we get $y - 3/N \leq x$. Since $y - 3/N \leq x$ for all positive integers N, we conclude that $y \leq x$. (Note that if we had $y > x$, then there exists a positive integer N such that $N(y - x) > 3$ by the Archimedean property, so $y - x > 3/N$, so $y - 3/N > x$, a contradiction.) In conclusion, x is an upper bound for E.

We now conclude by showing that x is the least upper bound for E. Let z be any other upper bound for E. We need to show that $x \leq z$. For any positive integer N, we know that $a_N - 1/N$ is not an upper bound for E. So, there exists $e \in E$ such that $a_N - 1/N < e \leq z$, so $a_N - 1/N < z$. From $(**)$, $x - a_N \leq 2/N$. Adding these two inequalities, $x < z + 3/N$ for all positive integers N. Therefore, $x \leq z$, as desired. □

Definition 1.7.7 (**Supremum**)**.** Let E be a subset of $\mathbb{R}$ with some upper bound. The least upper bound of E is called the **supremum** of E. The supremum of E, which exists by Theorem 1.7.6, is denoted by $\sup(E)$ or $\sup E$. If E has no upper bound, we use the symbol $+\infty$ and we write $\sup(E) = +\infty$. If E is empty, we write $\sup(E) = -\infty$.

Definition 1.7.8 (**Infimum**)**.** Let E be a subset of $\mathbb{R}$ with some lower bound. The greatest lower bound of E is called the **infimum** of E. The infimum of E, which exists by Theorem 1.7.6, is denoted by $\inf(E)$ or $\inf E$. If E has no lower bound, we write $\inf(E) = -\infty$. If E is empty, we write $\inf(E) = +\infty$.

In Proposition 1.4.19, we saw that there does not exist a rational number x such that $x^2 = 2$. However, Theorem 1.7.6 allows us to show that there exists a real number x such that $x^2 = 2$. In this sense, the real numbers do not have a "gap" here. And indeed, we can always take the square root of a real positive number, and recover another positive real number.

Proposition 1.7.9. *There exists a real number x such that $x^2 = 2$.*

Proof. Let E be the set $E := \{y \in \mathbb{R} \colon y \geq 0 \,\text{and}\, y^2 < 2\}$. Note that E has an upper bound of 2, since $2^2 = 4 > 2$. So, by Theorem 1.7.6, there exists a real number x such that x is the unique least upper bound of E. We will show that $x^2 = 2$. In order to show $x^2 = 2$, we will show that either $x^2 < 2$ or $x^2 > 2$ lead to contradictions.

Assume for the sake of contradiction that $x^2 < 2$. Since 2 is an upper bound for E, and x is the least upper bound of E, we have $x \leq 2$. Let $0 < \varepsilon < 1$ be a real number. Then $\varepsilon^2 < \varepsilon$, so

$$(x+\varepsilon)^2 = x^2 + 2x\varepsilon + \varepsilon^2 \leq x^2 + 4\varepsilon + \varepsilon = x^2 + 5\varepsilon.$$

Since $x^2 < 2$, we can choose $0 < \varepsilon < 1$ such that $x^2 + 5\varepsilon < 2$, by the Archimedean property (more specifically, $2 - x^2 > 0$, so we can choose $\varepsilon > 0$ such that $2 - x^2 > 5\varepsilon$ by the Archimedian property, i.e. $x^2 + 5\varepsilon < 2$). Therefore, $(x+\varepsilon)^2 < 2$. So, $x + \varepsilon \in E$, but $x + \varepsilon > x$, contradicting the fact that x is an upper bound for E. We conclude that $x^2 < 2$ does not hold.

Now, assume for the sake of contradiction that $x^2 > 2$. As before, $1 \leq x \leq 2$. Let $0 < \varepsilon < 1$ be a real number. Then $\varepsilon^2 < \varepsilon$, so

$$(x-\varepsilon)^2 = x^2 - 2x\varepsilon + \varepsilon^2 \geq x^2 - 2x\varepsilon \geq x^2 - 4\varepsilon.$$

Since $x^2 > 2$, we can choose $0 < \varepsilon < 1$ such that $x^2 - 4\varepsilon > 2$, by the Archimedean property. That is, $(x-\varepsilon)^2 > 2$. So, for any $y \in E$, we must have $x - \varepsilon \geq y$. (If not, then $0 < x - \varepsilon < y$, so $(x-\varepsilon)^2 < y^2$, so $y^2 > 2$, contradicting that $y \in E$.) So, $x - \varepsilon$ is an upper bound for E, but $x - \varepsilon < x$, contradicting the fact that x is the least upper bound for E. We conclude that $x^2 > 2$ does not hold.

Finally, we conclude that $x^2 = 2$, as desired. □

2. Cardinality, Sequences, Series, Subsequences

2.1. **Cardinality of Sets.** In the previous sections, we constructed the real numbers, and discussed the completeness of the real numbers. We showed that the real numbers are a set of numbers that are larger than the rational numbers, in the sense that the rational numbers are contained in the real numbers. Also, there are real numbers that are not rational, such as the square root of two. There is even another sense in which the set of real numbers is much larger than the set of rational numbers. But what do we mean by this? There are evidently infinitely many rational numbers, and there are infinitely many real numbers. So how can one infinite thing be larger than another infinite thing? These questions lead us to the notion of cardinality.

The basic question we ask is: what does it mean for two sets to be of the same size? In essentially all cultures of the world, there are two fundamental concepts of numbers. The first concept is the notion of one, two and many. That is, essentially every culture of the world recognizes that the natural numbers exist, in some sense. (This is one reason that we call these numbers the natural numbers, after all.) The second concept of numbers is the notion of a bijective correspondence. What does it mean that I have the same number of apples and oranges? Well, it means that I can put the first apple next to the first orange, and I put the second apple next to the second orange, and so on, until every apple is matched to exactly one orange, and every orange is matched to exactly one apple. This is the notion of bijective correspondence which we use to define cardinality.

Let's now phrase this discussion using mathematical terminology. Let X, Y be sets, and let $f: X \to Y$ be a function.

Definition 2.1.1 (**Bijection**). The function $f: X \to Y$ is said to be **bijective** (or a **one-to-one correspondence**) if and only if: for every $y \in Y$, there exists exactly one $x \in X$ such that $f(x) = y$.

Example 2.1.2. Consider the sets $X = \{0, 1, 2\}$ and $Y = \{1, 2, 4\}$. Define $f: X \to Y$ by $f(0) = 1$, $f(1) = 4$ and $f(2) = 2$. Then f is a bijection.

Example 2.1.3. Consider the sets $X = \mathbb{N} = \{0, 1, 2, \ldots\}$ and $Y = \{1, 2, 3, 4, \ldots\}$. Define $f: X \to Y$ so that, for all $x \in X$, $f(x) := x + 1$. Then f is a bijection.

Remark 2.1.4. A function $f: X \to Y$ is bijective if and only if it is both injective and surjective. Also, if f is a bijection, then f is invertible. That is, there exists a function $f^{-1}: Y \to X$ such that $f(f^{-1}(y)) = y$ for all $y \in Y$, and $f^{-1}(f(x)) = x$ for all $x \in X$.

Definition 2.1.5 (Cardinality). Two sets X, Y are said to have the same **cardinality** if and only if there exists a bijection from X onto Y.

Remark 2.1.6. The important thing to note here is that X and Y may be finite or infinite. At this point, it is not clear whether or not two infinite sets can have different cardinality. However, we will show below that the real numbers and the rational numbers do not have the same cardinality.

Exercise 2.1.7. Show that the notion of two sets having equal cardinality is an equivalence relation. That is, show:

- X has the same cardinality as X.
- If X has the same cardinality as Y, then Y has the same cardinality as X.
- If X has the same cardinality as Y, and if Y has the same cardinality as Z, then X has the same cardinality as Z.

Definition 2.1.8. Let n be a natural number. A set X is said to have **cardinality** n if and only if X has the same cardinality as $\{i \in \mathbb{N}: 1 \leq i \leq n\}$. We also say that X has n **elements** if and only if X has cardinality n.

Proposition 2.1.9. *Let n be a natural number, and suppose X is a set with cardinality n. Let m be any natural number such that $m \neq n$. Then X does not have cardinality m.*

Definition 2.1.10. A set X is **finite** if and only if there exists a natural number n such that X has cardinality n. Otherwise, the set X is called **infinite**.

Theorem 2.1.11. *The set of natural numbers $\mathbb{N}$ is infinite.*

Exercise 2.1.12. Using a proof by contradiction, prove Theorem 2.1.11.

Definition 2.1.13 (Countable Set). A set X is said to be **countably infinite** (or just **countable**) if and only if X has the same cardinality as $\mathbb{N}$. A set X is said to be **at most countable** if X is either finite or countable.

Exercise 2.1.14. Let X be a subset of the natural numbers $\mathbb{N}$. Then X is at most countable.

Exercise 2.1.15. Let X be a subset of a countable set Y. Then X is at most countable.

Exercise 2.1.16. Let $f: \mathbb{N} \to Y$ be a function. Then $f(\mathbb{N})$ is at most countable. (Hint: consider the set $A := \{n \in \mathbb{N}: f(n) \neq f(m) \text{ for all } 0 \leq m < n\}$. Prove that f is a bijection from A onto $f(\mathbb{N})$. Then use Exercise 2.1.14.)

Exercise 2.1.17. Let X be a countable set. Let $f: X \to Y$ be a function. Then $f(X)$ is at most countable.

We will now show that the integers and the rational numbers are countable.

Proposition 2.1.18. *Let X, Y be countable sets. Then $X \cup Y$ is a countable set.*

Exercise 2.1.19. Prove Proposition 2.1.18

Corollary 2.1.20. *The integers $\mathbb{Z}$ are countable.*

Proof. Write $\mathbb{Z} = \{0, 1, 2, \ldots\} \cup \{-1, -2, -3, \ldots\}$. We have therefore written $\mathbb{Z}$ as the union of two countable sets. Applying Proposition 2.1.18, we see that $\mathbb{Z}$ is countable. □

Definition 2.1.21 (Cartesian product). Let X, Y be sets. Define the set $X \times Y$ so that

$$X \times Y := \{(x, y) \colon x \in X \,\text{and}\, y \in Y\}.$$

The following strengthening of Proposition 2.1.18 shows that a countable union of countable sets is still countable.

Lemma 2.1.22. $\mathbb{N} \times \mathbb{N}$ *is countable.*

Proof. We need to construct a bijection $f \colon \mathbb{N} \times \mathbb{N} \to \mathbb{N}$. Let $k \in \mathbb{N}$, and consider the "diagonal"

$$D_k := \{(x, y) \in \mathbb{N} \times \mathbb{N} \colon x + y = k\}.$$

Note that the cardinality of D_k is $k + 1$, and the cardinality of $D_0 \cup D_1 \cup \cdots \cup D_k$ is $1+2+\cdots+k+1 = (k+1)(k+2)/2$. Define $a_k := (k+1)(k+2)/2$. Note that $a_k+k+2 = a_{k+1}$. We define $f(0, 0) := 0$, and we then define f inductively as follows. Suppose we have defined f on $D_0, D_1, \ldots, D_k$ so that f maps $D_0 \cup D_1 \cup \cdots \cup D_k$ onto $\{0, 1, \ldots, a_k - 1\}$. Then, define $f(0, k + 1) := a_k$, $f(1, k) := a_k + 1$, $f(2, k - 1) := a_k + 2$, and so on. In general, for any $0 \leq j \leq k + 1$, define $f(j, k + 1 - j) := a_k + j$. We have therefore defined f so that f maps $D_0 \cup \cdots \cup D_{k+1}$ onto $\{0, 1, \ldots, a_{k+1} - 1\}$. The map f can be visualized in the following way

$$\begin{pmatrix} (0,0) & (0,1) & (0,2) & (0,3) & \cdots \\ (1,0) & (1,1) & (1,2) & \cdots & \\ (2,0) & (2,1) & \ddots & & \\ (3,0) & \vdots & & & \\ \vdots & & & & \end{pmatrix} \longrightarrow \begin{pmatrix} 0 & 1 & 3 & 6 & \cdots \\ 2 & 4 & 7 & \cdots & \\ 5 & 8 & \ddots & & \\ 9 & \vdots & & & \\ \vdots & & & & \end{pmatrix}$$

We now prove that f is a bijection. By the definition of f, if k is any natural number, then f is a bijection from D_k onto $\{a_k, a_k + 1, \ldots, a_{k+1} - 1\}$. We first show that f is injective. Let $(a, b), (c, d) \in \mathbb{N} \times \mathbb{N}$. Assume that $f(a, b) = f(c, d)$. For any natural numbers k, k' with $k \neq k'$, the sets of integers $\{a_k, a_k + 1, \ldots, a_{k+1} - 1\}$ and $\{a_{k'}, a_{k'} + 1, \ldots, a_{k'+1} - 1\}$ are disjoint. So, if $f(a, b) = f(c, d)$, there must exist a natural number k such that $f(a, b)$ and $f(c, d)$ are both contained in $\{a_k, a_k + 1, \ldots, a_{k+1} - 1\}$. Since f is a bijection from D_k onto $\{a_k, a_k + 1, \ldots, a_{k+1} - 1\}$, we conclude that $(a, b) = (c, d)$. Therefore, f is injective.

We now conclude by showing that f is surjective. Let $n \in \mathbb{N}$. We need to find $(a, b) \in \mathbb{N}\times\mathbb{N}$ such that $f(a, b) = n$. Since $\mathbb{N} = \cup_{k\in\mathbb{N}}\{a_k, a_k+1, \ldots, a_{k+1}-1\}$, there exists a natural number k such that n is in the set $\{a_k, a_k + 1, \ldots, a_{k+1} - 1\}$. Since f is a bijection from D_k onto $\{a_k, a_k + 1, \ldots, a_{k+1} - 1\}$, there exists $(a, b) \in D_k$ such that $f(a, b) = n$. Therefore, f is surjective. In conclusion, f is a bijection, as desired. □

Exercise 2.1.23. Using Lemma 2.1.22, prove the following statement. If X, Y are countable sets, then $X \times Y$ is countable.

Corollary 2.1.24. *The rational numbers* $\mathbb{Q}$ *are countable.*

Proof. From Corollary 2.1.20, the integers $\mathbb{Z}$ are countable. So, the nonzero integers $\mathbb{Z} \smallsetminus \{0\}$ are also countable. Define a function $f\colon \mathbb{Z} \times (\mathbb{Z} \smallsetminus \{0\}) \to \mathbb{Q}$ by

$$f(a, b) := a/b.$$

Since $b \neq 0$, f is well-defined. From Exercise 2.1.23, f is then a function from a countable set into the rational numbers $\mathbb{Q}$. Also, from the definition of the rational numbers, $f(\mathbb{Z} \times (\mathbb{Z} \smallsetminus \{0\})) = \mathbb{Q}$. From Exercise 2.1.17, we conclude that $\mathbb{Q}$ is at most countable. Since $\mathbb{Q}$ contains the integers, $\mathbb{Q}$ is not finite. Therefore, $\mathbb{Q}$ is countable, as desired. □

In summary, the natural numbers, integers, and rational numbers are countable. Surprisingly, the real numbers are not countable as we will show further below.

Definition 2.1.25 (Uncountable Set). Let X be a set. We say that X is **uncountable** if and only if X is not finite, and X is not countable.

Definition 2.1.26 (Power Set). Let X be a set. Define the **power set** 2^X as the set of all subsets of X. Equivalently, 2^X is the set of all functions $f\colon X \to \{0, 1\}$.

Remark 2.1.27. To see the equivalence of these two definitions, for any subset A of X, we associate A with the function $f\colon X \to \{0, 1\}$ where $f(x) = 1$ if and only if $x \in A$. In the other direction, given a function $f\colon X \to \{0, 1\}$, we associate f to the set $A = \{x \in X : f(x) = 1\}$. This association gives a bijection between the subset of A, and the set of all functions $f\colon X \to \{0, 1\}$.

Proposition 2.1.28. *Let* X *be a set. Then* X *and* 2^X *do not have the same cardinality.*

Proof. We argue by contradiction. Suppose X and 2^X have the same cardinality. Then there exists a bijection $f\colon X \to 2^X$. Consider the following subset V of X.

$$V := \{x \in X : x \notin f(x)\}.$$

We will achieve a contradiction by showing that V is not in the range of f. Since f is a bijection and $V \in 2^X$, there exists $y \in X$ such that $f(y) = V$. We now consider two cases.

Case 1. $y \in f(y)$. If $y \in f(y)$, then $y \in V$, since $f(y) = V$. However, from the definition of V, if $y \in V$, then $y \notin f(y)$, a contradiction.

Case 2. $y \notin f(y)$. If $y \notin f(y)$, then $y \notin V$, since $f(y) = V$. So, from the definition of V, $y \in f(y)$, a contradiction.

In either case, we get a contradiction. We conclude that X and 2^X do not have the same cardinality. □

Corollary 2.1.29. $\mathbb{N}$ *and* $2^{\mathbb{N}}$ *do not have the same cardinality. In particular,* $2^{\mathbb{N}}$ *is uncountable.*

Corollary 2.1.30. *The set of real numbers* $\mathbb{R}$ *is uncountable.*

Proof. Let $f\colon \mathbb{N} \to \{0, 1\}$ be an element of $2^{\mathbb{N}}$. For any natural number n, define

$$a_n := \sum_{i=1}^{n} 3^{-i} f(i).$$

One can show that $(a_n)_{n=0}^{\infty}$ is a Cauchy sequence of rational numbers. We therefore define a map $F\colon 2^{\mathbb{N}} \to \mathbb{R}$ so that

$$F(f) := (\sum_{i=1}^{n} 3^{-i} f(i))_{n=0}^{\infty}.$$

We will show that F is an injection. Let $f, g\colon \mathbb{N} \to \{0,1\}$ such that $f \neq g$. Then there exists $N \in \mathbb{N}$ such that $f(N) \neq g(N)$. Without loss of generality, N is the smallest element of $\mathbb{N}$ such that $f(N) \neq g(N)$. Also, without loss of generality, $f(N) = 1$ and $g(N) = 0$. By the definition of N, we have $f(i) = g(i)$ for all $1 \leq i \leq N-1$. Therefore,

$$\begin{aligned} F(f) - F(g) &= (\sum_{i=1}^{n} 3^{-i} f(i))_{n=0}^{\infty} - (\sum_{i=1}^{n} 3^{-i} g(i))_{n=0}^{\infty} \\ &= (\sum_{i=1}^{n} 3^{-i} (f(i) - g(i)))_{n=0}^{\infty} = (3^{-N} + \sum_{i=N+1}^{n} 3^{-i}(f(i) - g(i)))_{n=N}^{\infty} \end{aligned}$$

Since $f(i), g(i) \in \{0,1\}$ for all $i \in \mathbb{N}$, we have $|f(i) - g(i)| \leq 1$. So, for any $n \geq N+1$, we have by the triangle inequality

$$\left|\sum_{i=N+1}^{n} 3^{-i}(f(i) - g(i))\right| \leq \sum_{i=N+1}^{n} 3^{-i} \leq (2/3)3^{-N}.$$

So, $3^{-N} + \sum_{i=N+1}^{n} 3^{-i}(f(i) - g(i)) \geq 3^{-N} - (2/3)3^{-N} = 3^{-N-1}$. Therefore, $F(f) - F(g) \geq 3^{-N-1} > 0$. In particular, $F(f) \neq F(g)$.

We conclude that $F\colon 2^{\mathbb{N}} \to \mathbb{R}$ is an injection. From Corollary 2.1.29, $2^{\mathbb{N}}$ is uncountable. Since F is an injection, F is a bijection onto its image $F(2^{\mathbb{N}})$. That is, $F(2^{\mathbb{N}})$ is uncountable. Finally, if $\mathbb{R}$ were countable, then all of its subsets would be at most countable, by Exercise 2.1.15. But we have found an uncountable subset $F(2^{\mathbb{N}})$ of $\mathbb{R}$. We therefore conclude that $\mathbb{R}$ is not countable. We also know that $\mathbb{R}$ is not finite, since it contains $\mathbb{N}$. We conclude that $\mathbb{R}$ is uncountable. □

2.2. **Sequences of Real Numbers.** This course has a few fundamental concepts. One of these fundamental concepts is the Cauchy sequence. We will now introduce another fundamental concept, which is a variation on the Cauchy sequence. We will discuss sequences of real numbers and their limits. This topic is perhaps a bit more familiar, though it will turn out that a sequence of real numbers will have a limit if and only if this sequence is a Cauchy sequence. So, in some sense, we have been working with a familiar topic all along.

Our more general discussion of sequences of real numbers will inform our later investigation of derivatives and integration. More specifically, we can define derivatives and integrals in terms of limits of sequences of real numbers. So, a thorough understanding of limits of sequences of real numbers allows a quick and thorough investigation of derivatives and integrals.

Definition 2.2.1 (**Cauchy Sequence**)**.** Let $(a_n)_{n=0}^{\infty}$ be a sequence of real numbers. We say that $(a_n)_{n=0}^{\infty}$ is a **Cauchy sequence** if and only if, for any real $\varepsilon > 0$, there exists a natural number $N = N(\varepsilon)$ such that, for all $n, m \geq N$, we have $|a_n - a_m| < \varepsilon$.

Remark 2.2.2. Our previous definition of a Cauchy sequence asked for the same condition to hold for all rational $\varepsilon > 0$. So, Definition 2.2.1 may appear to be stricter than our previous definition of a Cauchy sequence. However, given any real $\varepsilon > 0$, Corollary 1.6.37 gives a rational $\varepsilon' > 0$ with $\varepsilon' < \varepsilon$. So, within Definition 2.2.1, it is equivalent to require the definition to hold for all rational $\varepsilon > 0$, or for all real $\varepsilon > 0$. That is, our previous definition and our current definition of a Cauchy sequence both coincide.

Definition 2.2.3 (Convergent Sequence). Let $(a_n)_{n=0}^{\infty}$ be a sequence of real numbers, and let L be a real number. We say that the sequence $(a_n)_{n=0}^{\infty}$ **converges to** L if and only if, for every real $\varepsilon > 0$, there exists a natural number $N = N(\varepsilon)$ such that, for all $n \geq N$, we have $|a_n - L| < \varepsilon$.

Proposition 2.2.4. *Let $(a_n)_{n=0}^{\infty}$ be a sequence of real numbers, and let L, L' be a real numbers with $L \neq L'$. Then $(a_n)_{n=0}^{\infty}$ cannot simultaneously converge to L and converge to L'.*

Proof. We argue by contradiction. Suppose $(a_n)_{n=0}^{\infty}$ converges to L and to L'. Define $\varepsilon := |L - L'|/4 > 0$. Since $(a_n)_{n=0}^{\infty}$ converges to L, there exists N such that, for all $n \geq N$, we have $|a_n - L| < \varepsilon$. Since $(a_n)_{n=0}^{\infty}$ converges to L', there exists N' such that, for all $n \geq N'$, we have $|a_n - L'| < \varepsilon$. Setting $M := \max(N, N')$, we have

$$|a_M - L| < |L - L'|/4, \qquad |a_M - L'| < |L - L'|/4.$$

By the triangle inequality,

$$|L - L'| = |L - a_M + a_M - L'| \leq |a_M - L| + |a_M - L'| < |L - L'|/2.$$

Since $|L - L'| > 0$, we have shown that $2 < 1$, a contradiction. We conclude that it cannot occur that $(a_n)_{n=0}^{\infty}$ converges to L and to L' with $L \neq L'$. □

Definition 2.2.5 (Limit). Let $(a_n)_{n=0}^{\infty}$ be a sequence of real numbers that is converging to a real number L. We then say that the sequence $(a_n)_{n=0}^{\infty}$ is **convergent**, and we write

$$L = \lim_{n\to\infty} a_n.$$

If $(a_n)_{n=0}^{\infty}$ is not convergent, we say that the sequence $(a_n)_{n=0}^{\infty}$ is **divergent**, and we say the limit of L is undefined.

Remark 2.2.6. By Proposition 2.2.4, if $(a_n)_{n=0}^{\infty}$ converges to some limit L, then this limit is unique. So, we call L the **limit** of the sequence $(a_n)_{n=0}^{\infty}$.

Remark 2.2.7. Instead of writing $(a_n)_{n=0}^{\infty}$ converges to L, we will sometimes write $a_n \to L$ as $n \to \infty$.

Proposition 2.2.8. $\lim_{n\to\infty}(1/n) = 0$.

Proof. Let $\varepsilon > 0$ be a real number. By the Archimedian property (Theorem 1.6.36), there exists a positive integer N such that $0 < 1/N < \varepsilon$. So, for all $n \geq N$, we have $|a_n - 0| = |a_n| = 1/n \leq 1/N < \varepsilon$. □

Exercise 2.2.9. Let $(a_n)_{n=m}^{\infty}$ be a sequence of real numbers converging to 0. Show that $(|a_n|)_{n=m}^{\infty}$ also converges to zero.

The following Theorem shows that Cauchy sequences and convergent sequences are the same thing. This Theorem also demonstrates that the real numbers are complete, in that a Cauchy sequence of real numbers converges to a real number. Note that the corresponding statement for the rational numbers is false. That is, a Cauchy sequence of rational numbers does not necessarily converge to a rational number. So, in this sense, the real numbers do not have any "holes," but the rational numbers do.

Theorem 2.2.10 (**Completeness of $\mathbb{R}$**). *Let $(a_n)_{n=0}^{\infty}$ be a sequence of real numbers. Then $(a_n)_{n=0}^{\infty}$ is convergent if and only if $(a_n)_{n=0}^{\infty}$ is a Cauchy sequence.*

Exercise 2.2.11. Prove Theorem 2.2.10. (Hint: Given a Cauchy sequence $(a_n)_{n=0}^{\infty}$, use that the rationals are dense in the real numbers to replace each real a_n by some rational a_n', so that $|a_n - a_n'|$ is small. Then, ensure that the sequence $(a_n')_{n=0}^{\infty}$ is a Cauchy sequence of rationals and that $(a_n')_{n=0}^{\infty}$ defines a real number which is the limit of the original sequence $(a_n)_{n=0}^{\infty}$.)

As a Corollary of Theorem 2.2.10, the formal limits of Cauchy sequences of rationals are actual limits. That is, we used a sensible notation for formal limits during our construction of the real number system.

Corollary 2.2.12. *Let $(a_n)_{n=0}^{\infty}$ be a Cauchy sequence of rational numbers. Then $(a_n)_{n=0}^{\infty}$ converges to* $\mathrm{LIM}_{n\to\infty}a_n$. *That is,*

$$\mathrm{LIM}_{n\to\infty}a_n = \lim_{n\to\infty} a_n.$$

Definition 2.2.13. Let M be a real number. A sequence $(a_n)_{n=0}^{\infty}$ is **bounded by** M if and only if $|a_n| \leq M$ for all $n \in \mathbb{N}$. We say that $(a_n)_{n=0}^{\infty}$ is **bounded** if and only if there exists a real number M such that $(a_n)_{n=0}^{\infty}$ is bounded by M.

Recall that any Cauchy sequence of rational numbers is bounded. The proof of this statement also shows that any Cauchy sequence of real numbers is bounded. So, from Theorem 2.2.10 we get the following.

Corollary 2.2.14. *Every convergent sequence is bounded.*

Theorem 2.2.15 (**Limit Laws**). *Let $(a_n)_{n=0}^{\infty}$, $(b_n)_{n=0}^{\infty}$ be convergent sequences. Let x, y be real numbers such that $x = \lim_{n\to\infty} a_n$, $y = \lim_{n\to\infty} b_n$.*

(i) *The sequence $(a_n + b_n)_{n=0}^{\infty}$ converges to $x + y$. That is,*

$$\lim_{n\to\infty}(a_n + b_n) = (\lim_{n\to\infty} a_n) + (\lim_{n\to\infty} b_n).$$

(ii) *The sequence $(a_n b_n)_{n=0}^{\infty}$ converges to xy. That is,*

$$\lim_{n\to\infty}(a_n b_n) = (\lim_{n\to\infty} a_n)(\lim_{n\to\infty} b_n).$$

(iii) *For any real number c, the sequence $(ca_n)_{n=0}^{\infty}$ converges to cx. That is,*

$$c\lim_{n\to\infty} a_n = \lim_{n\to\infty}(ca_n).$$

(iv) *The sequence $(a_n - b_n)_{n=0}^{\infty}$ converges to $x - y$. That is,*

$$\lim_{n\to\infty}(a_n - b_n) = (\lim_{n\to\infty} a_n) - (\lim_{n\to\infty} b_n).$$

(v) *Suppose $x \neq 0$ and there exists m such that $a_n \neq 0$ for all $n \geq m$. Then $(a_n^{-1})_{n=m}^{\infty}$ converges to x^{-1}. That is,*

$$\lim_{n\to\infty} a_n^{-1} = (\lim_{n\to\infty} a_n)^{-1}.$$

(vi) *Suppose $x \neq 0$ and there exists m such that $a_n \neq 0$ for all $n \geq m$. Then $(b_n/a_n)_{n=m}^{\infty}$ converges to y/x. That is,*

$$\lim_{n\to\infty}(b_n/a_n) = (\lim_{n\to\infty} b_n)/(\lim_{n\to\infty} a_n).$$

(vii) *Suppose* $a_n \geq b_n$ *for all* $n \geq 0$. *Then* $x \geq y$.

Exercise 2.2.16. Prove Theorem 2.2.15.

2.3. **The Extended Real Number System.** Now that we have defined limits, it is slightly more convenient to add two additional symbols to the real number system, namely $+\infty$ and $-\infty$.

Definition 2.3.1 (Extended Real Number System). The **extended real number system** $\mathbb{R}^*$ is the real line $\mathbb{R}$ with two additional elements $+\infty$ and $-\infty$. These two additional elements are distinct from each other, and these two elements are distinct from all other elements of the real line. So, $\mathbb{R}^* = \mathbb{R} \cup \{+\infty\} \cup \{-\infty\}$. An extended real number x is called **finite** if and only if x is a real number, and x is called **infinite** if and only if x is equal to $+\infty$ or $-\infty$. (Note that these notions of finite and infinite are similar to but distinct from our notions of finite and infinite sets.)

Definition 2.3.2 (Negation). The operation of negation is defined for any extended real number x by defining $-(+\infty) := -\infty$, and $-(-\infty) := +\infty$. And for any finite extended real number x, we use the usual operation of negation.

So, $-(-x) = x$ for any $x \in \mathbb{R}^*$. We can also extend the order on $\mathbb{R}$ to an order on $\mathbb{R}^*$.

Definition 2.3.3 (Order). Let x, y be extended real numbers. We say that x is less than or equal to y, and we write $x \leq y$, if and only if one of the following statements holds.

- x, y are real numbers, and $x \leq y$ as real numbers.
- $y = +\infty$.
- $x = -\infty$.

We say that $x < y$ if and only if $x \leq y$ and $x \neq y$. We sometimes write $y > x$ to indicate $x < y$, and we sometimes write $y \geq x$ to indicate $x \leq y$.

Remark 2.3.4. One can then check that this order on $\mathbb{R}^*$ satisfies the usual properties of order. Let $x, y, z \in \mathbb{R}^*$. Then

- $x \leq x$
- If $x \leq y$ and $y \leq x$ then $x = y$.
- If $x \leq y$ and $y \leq z$ then $x \leq z$.
- If $x \leq y$ then $-y \leq -x$.

Remark 2.3.5. It would be nice to extend other operations such as addition and multiplication to the extended real number system. However, doing so could introduce several inconsistencies within the various arithmetic operations. So, we will not extend other operations of arithmetic to $\mathbb{R}^*$. For example, it seems reasonable to define $1 + \infty = \infty$ and $2 + \infty = \infty$, but then $1 + \infty = 2 + \infty$, so the cancellation law no longer holds on $\mathbb{R}^*$.

One convenient property of the extended real number system is that the supremum and infimum operations are a bit easier to handle. In particular, the Theorem below can be stated succinctly, without explicitly reverting to different cases.

Definition 2.3.6 (Supremum). Let E be a subset of $\mathbb{R}^*$. We define the **supremum** $\sup(E)$ or **least upper bound** of E by the following conditions.

- If E is contained in $\mathbb{R}$ (so that $+\infty$ and $-\infty$ are not elements of E), then $\sup(E)$ is already defined.

- If E contains $+\infty$, define $\sup(E) := +\infty$.
- If E does not contain $+\infty$, and if E does contain $-\infty$, then $E \setminus \{-\infty\}$ is a subset of $\mathbb{R}$. So, we define $\sup(E) := \sup(E \setminus \{-\infty\})$.

Definition 2.3.7 (**Infimum**)**.** Let E be a subset of $\mathbb{R}^*$. We define the **infimum** $\inf(E)$ or **greatest lower bound** of E by $\inf(E) := -(\sup(-E))$.

Theorem 2.3.8. *Let E be a subset of $\mathbb{R}^*$. Then the following statements hold.*

- *For every $x \in E$, we have $x \leq \sup(E)$ and $x \geq \inf(E)$.*
- *Let $M \in \mathbb{R}^*$ be an upper bound for E, so that $x \leq M$ for all $x \in E$. Then $\sup(E) \leq M$.*
- *Let $M \in \mathbb{R}^*$ be a lower bound for E, so that $x \geq M$ for all $x \in E$. Then $\inf(E) \geq M$.*

Exercise 2.3.9. Prove Theorem 2.3.8

Remark 2.3.10 (**Limits and Infinity**)**.** Let $(a_n)_{n=0}^{\infty}$ be a sequence. If for all positive integers M, there exists N such that, for all $n \geq N$, we have $a_n > M$, we then write $\lim_{n\to\infty} a_n = +\infty$. In this case, we still say that the limit of the sequence does not exist. If for all negative integers M, there exists N such that, for all $n \geq N$, we have $a_n < M$, we then write $\lim_{n\to\infty} a_n = -\infty$. In this case, we still say that the limit of the sequence does not exist.

2.4. **Suprema and Infima of Sequences.** The extended real number system and Theorem 2.3.8 simplify our notation for suprema and infima of sets. One of the main motivations for suprema and infima is that they will aid our rigorous investigation of sequences of real numbers. That is, given a sequence of real numbers $(a_n)_{n=0}^{\infty}$, we will consider the suprema and infima of the *subset of real numbers*, $\{a_n : n \in \mathbb{N}\} \subseteq \mathbb{R}$.

Definition 2.4.1 (Suprema and infima of a sequence)**.** Let $(a_n)_{n=m}^{\infty}$ be a sequence of real numbers. Define $\sup(a_n)_{n=m}^{\infty}$ to be the supremum of the set $\{a_n : n \geq m, n \in \mathbb{N}\}$. Define $\inf(a_n)_{n=m}^{\infty}$ to be the infimum of the set $\{a_n : n \geq m, n \in \mathbb{N}\}$.

Example 2.4.2. For any $n \in \mathbb{N}$, let $a_n := (-1)^n$. Then $\sup(a_n)_{n=0}^{\infty} = 1$ and $\inf(a_n)_{n=0}^{\infty} = -1$.

Example 2.4.3. For any positive integer n, let $a_n := 1/n$. Then $\sup(a_n)_{n=1}^{\infty} = 1$ and $\inf(a_n)_{n=1}^{\infty} = 0$. Note that the infimum of the sequence $(a_n)_{n=1}^{\infty}$ is not actually a member of the sequence $(a_n)_{n=1}^{\infty}$.

Proposition 2.4.4. *Let $(a_n)_{n=m}^{\infty}$ be a sequence of real numbers. Let x be the extended real number $x := \sup(a_n)_{n=m}^{\infty}$. Then $a_n \leq x$ for all $n \geq m$. Also, for any $M \in \mathbb{R}^*$ which is an upper bound for $(a_n)_{n=m}^{\infty}$ (so that $a_n \leq M$ for all $n \geq m$), we have $x \leq M$. Finally, for any $y \in \mathbb{R}^*$ such that $y < x$, there exists at least one integer n with $n \geq m$ such that $y < a_n \leq x$.*

Exercise 2.4.5. Prove Proposition 2.4.4 using Theorem 2.3.8.

In Corollary 2.2.14, we saw that every convergent sequence is bounded. The converse of this statement is not true. The sequence $a_n = (-1)^n$ is bounded in absolute value by 1, but this sequence is not convergent. However, if we change the statement of the converse slightly, then it does become both true and quite useful. For example, we have the following.

Proposition 2.4.6. *Let $(a_n)_{n=m}^{\infty}$ be a bounded sequence of real numbers. Assume also that $(a_n)_{n=m}^{\infty}$ is monotone increasing. That is, $a_{n+1} \geq a_n$ for all $n \geq m$. Then the sequence $(a_n)_{n=m}^{\infty}$ is convergent. In fact,*

$$\lim_{n\to\infty} a_n = \sup(a_n)_{n=m}^{\infty}.$$

Exercise 2.4.7. Prove Proposition 2.4.6 using Proposition 2.4.4.

Remark 2.4.8. One can similarly show that a bounded monotone decreasing sequence $(a_n)_{n=m}^{\infty}$ (i.e. a sequence with $a_{n+1} \leq a_n$ for all $n \geq m$) is convergent.

Remark 2.4.9. A sequence $(a_n)_{n=m}^{\infty}$ is said to be **monotone** if and only if it is monotone increasing or monotone decreasing. If $(a_n)_{n=m}^{\infty}$ is monotone, then from Proposition 2.4.6 and Corollary 2.2.14, we see that $(a_n)_{n=m}^{\infty}$ converges if and only if $(a_n)_{n=m}^{\infty}$ is bounded.

2.5. **Limsup, Liminf, and Limit Points.** In order to understand the limits of sequences, it is helpful to first generalize our notion of a limit to the notion of a limit point. We then study this slightly generalized notion of a limit. We will use the limsup and liminf as upper and lower bounds on the set of limit points, respectively. Ultimately, if we for example want to prove that the limit of a sequence exists, it will sometimes be much easier to find upper and lower bounds on the set of limit points. Then, if we can show that the upper bound is equal to the lower bound, then we will have shown that the sequence is convergent.

Definition 2.5.1 (Limit Point). Let $(a_n)_{n=m}^{\infty}$ be a sequence of real numbers and let x be a real number. We say that x is a **limit point** of the sequence $(a_n)_{n=m}^{\infty}$ if and only if: for every real $\varepsilon > 0$, for every natural number $N \geq m$, there exists $n \geq N$ such that $|a_n - x| < \varepsilon$.

Proposition 2.5.2. *Let $(a_n)_{n=m}^{\infty}$ be a sequence of real numbers that converges to a real number x. Then x is a limit point of $(a_n)_{n=m}^{\infty}$. Moreover, x is the only limit point of $(a_n)_{n=m}^{\infty}$.*

Exercise 2.5.3. Prove Proposition 2.5.2.

Definition 2.5.4 (Limsup). Let $(a_n)_{n=m}^{\infty}$ be a sequence of real numbers. For any natural number n with $n \geq m$, define $b_n := \sup_{t\geq n} a_t$. Since the set $\{t \in \mathbb{N} \colon t \geq n+1\}$ is contained in the set $\{t \in \mathbb{N} \colon t \geq n\}$, we conclude that $b_{n+1} \leq b_n$ for all $n \geq m$. That is the sequence $(b_n)_{n=m}^{\infty}$ is monotone decreasing. We therefore define the **limit superior** by $\limsup_{n\to\infty} a_n := \lim_{n\to\infty} b_n$. The limit on the right either exists as a real number, or if the limit does not exist, we denote this limit with the extended real number $-\infty$. In summary, the following definition makes sense by Remark 2.4.9.

$$\limsup_{n\to\infty} a_n := \lim_{n\to\infty} \sup_{m\geq n} a_m.$$

Definition 2.5.5 (Liminf). Let $(a_n)_{n=m}^{\infty}$ be a sequence of real numbers. Reasoning as before, if we define $b_n := \inf_{m\geq n} a_m$, then $b_{n+1} \geq b_n$ for all $n \geq m$. So, the following definition of the **limit inferior** makes sense.

$$\liminf_{n\to\infty} a_n := \lim_{n\to\infty} \inf_{m\geq n} a_m.$$

Remark 2.5.6.

$$\limsup_{n\to\infty} a_n = \inf_{n\geq m} \sup_{t\geq n} a_t, \quad \text{and} \quad \liminf_{n\to\infty} a_n = \sup_{n\geq m} \inf_{t\geq n} a_t.$$

These identities follows from the monotonicity in n of the sequences $\sup_{t\geq n} a_t$ and $\inf_{t\geq n} a_t$, and Proposition 2.4.6

Proposition 2.5.7 (**Properties of Limsup/Liminf**). *Let $(a_n)_{n=m}^{\infty}$ be a sequence of real numbers. Let L^+ be the limit superior of this sequence, and let L^- be the limit inferior of this sequence. (Note that $L^+, L^- \in \mathbb{R}^*$.)*

(i) *For every $x > L^+$ there exists $N \geq m$ such that $a_n < x$ for all $n \geq N$. For every $y < L^-$ there exists $N \geq m$ such that $a_n > y$ for all $n \geq N$.*
(ii) *For every $x < L^+$ and for every $N \geq m$ there exists $n \geq N$ such that $a_n > x$. For every $y > L^-$ and for every $N \geq m$ there exists $n \geq N$ such that $a_n < y$.*
(iii) $\inf(a_n)_{n=m}^{\infty} \leq L^- \leq L^+ \leq \sup(a_n)_{n=m}^{\infty}$.
(iv) *If c is any limit point of $(a_n)_{n=m}^{\infty}$, then $L^- \leq c \leq L^+$.*
(v) *If L^+ is finite, then it is a limit point of $(a_n)_{n=m}^{\infty}$. If L^- is finite, then it is a limit point of $(a_n)_{n=m}^{\infty}$.*
(vi) *Let c be a real number. If $(a_n)_{n=m}^{\infty}$ converges to c, then $L^+ = L^- = c$. Conversely, if $L^+ = L^- = c$, then $(a_n)_{n=m}^{\infty}$ converges to c.*

Proof of (i). If $L^+ = +\infty$, there is nothing to prove. So, assume that $L^+ \neq +\infty$. Then $L^+ \in \mathbb{R}\cup\{-\infty\}$. Let $x > L^+$. From Remark 2.5.6, $L^+ = \inf_{n\geq m}\sup_{t\geq n} a_t$. From Proposition 2.4.4, there exists $n \geq m$ such that $x > \sup_{t\geq n} a_t$. Using Proposition 2.4.4 again, for all $t \geq n$, we have $x > a_t$, as desired. The second assertion follows similarly. □

Proof of (ii). If $L^+ = -\infty$, there is nothing to prove. So, assume that $L^+ \neq -\infty$. Then $L^+ \in \mathbb{R}\cup\{+\infty\}$. Let $x < L^+$. From Remark 2.5.6, $L^+ = \inf_{n\geq m}\sup_{t\geq n} a_t$. From Proposition 2.4.4, for all $n \geq m$ we have $x < \sup_{t\geq n} a_t$. Using Proposition 2.4.4 again, there exists $t \geq m$ such that $x < a_t$, as desired. The second assertion follows similarly. □

Exercise 2.5.8. Prove parts (iii)-(vi) of Proposition 2.5.7

Remark 2.5.9. Proposition 2.5.7(iv) and Definitions 2.5.4,2.5.5 say that, if L^+ and L^- are both finite, then they are the largest and smallest limit points of the sequence, respectively. Proposition 2.5.7(vi) shows that, to test whether or not a sequence converges, it suffices to compute the limit superior and limit inferior of the sequence.

Lemma 2.5.10 (**Comparison Principle**). *Let $(a_n)_{n=m}^{\infty}$, $(b_n)_{n=m}^{\infty}$ be sequences of real numbers. Assume that $a_n \leq b_n$ for all $n \geq m$. Then*

- $\sup(a_n)_{n=m}^{\infty} \leq \sup(b_n)_{n=m}^{\infty}$.
- $\inf(a_n)_{n=m}^{\infty} \leq \inf(b_n)_{n=m}^{\infty}$.
- $\limsup_{n\to\infty} a_n \leq \limsup_{n\to\infty} b_n$.
- $\liminf_{n\to\infty} a_n \leq \liminf_{n\to\infty} b_n$.

Exercise 2.5.11. Prove Lemma 2.5.10.

Corollary 2.5.12 (**Squeeze Test/ Squeeze Theorem**). *Let $(a_n)_{n=m}^{\infty}$, $(b_n)_{n=m}^{\infty}$, $(c_n)_{n=m}^{\infty}$ be sequences of real numbers such that there exists a natural number M such that, for all $n \geq M$,*

$$a_n \leq b_n \leq c_n.$$

Assume that $(a_n)_{n=m}^{\infty}$ and $(c_n)_{n=m}^{\infty}$ converge to the same limit L. Then $(b_n)_{n=m}^{\infty}$ converges to L.

Exercise 2.5.13. Prove Corollary 2.5.12 using Lemma 2.5.10.

2.5.1. *Exponentiation by Rationals.* For x, y real numbers, it would be nice to define x^y in some way. In the case that x is negative and y is e.g. 1/3, defining x^y requires complex analysis. In this class, we will only be able to define x^y for positive real numbers x. To this end, in this section, we will let x be a positive real number, and we will define x^y for rational y.

Definition 2.5.14. Let $x > 0$ be a positive real number, and let $n \geq 1$ be a positive integer. We define the n^{th} **root of** x, and write $x^{1/n}$, by the formula

$$x^{1/n} := \sup\{y \in \mathbb{R} \colon y \geq 0 \text{ and } y^n \leq x\}.$$

For x a positive real number and n a positive integer, we now show that $x^{1/n}$ is finite.

Lemma 2.5.15. *Let $x > 0$ be a positive real number, and let $n \geq 1$ be a positive integer. Then the set $E := \sup\{y \in \mathbb{R} \colon y \geq 0$ and $y^n \leq x\}$ is nonempty and bounded from above. Consequently, $x^{1/n}$ is a real number by the Least Upper Bound property (Theorem 1.7.6).*

Proof. Since x is positive, $0 \in E$, so E is nonempty. We now show that E is bounded from above. We consider two cases: $x \leq 1$ and $x > 1$. In the first case, $x \leq 1$, and we claim that 1 is an upper bound for E. That is, if $y \in \mathbb{R}$ and $y \geq 0$ with $y^n \leq x \leq 1$, then $y \leq 1$. We prove this by contradiction. Suppose $y > 1$. Since $y > 1$, it follows by induction on n that $y^n > 1$ as well, contradicting that $y^n \leq 1$. We conclude that E is bounded above by 1 when $x \leq 1$. We now consider the case $x > 1$. We claim that x is an upper bound for E. That is, if $y \in \mathbb{R}$ and $y \geq 0$ with $y^n \leq x$, then $y \leq x$. We prove this by contradiction. Suppose $y > x$. Since $x > 1$, we have $y > x > 1$. If then follows by induction on n that $y^n > x$, contradicting that $y^n \leq x$. We conclude that E is bounded above by x when $x > 1$. Having exhausted all cases for $x > 0$, we are done. $\square$

Lemma 2.5.16. *Let $x, y > 0$ be positive real numbers, and let $n, m \geq 1$ be positive integers.*

(i) *If $y = x^{1/n}$, then $y^n = x$.*
(ii) *If $y^n = x$, then $y = x^{1/n}$.*
(iii) *$x^{1/n}$ is a positive real number.*
(iv) *$x > y$ if and only if $x^{1/n} > y^{1/n}$.*
(v) *If $x > 1$ then $x^{1/n}$ decreases when n increases. If $x < 1$, then $x^{1/n}$ increases when n increases. If $x = 1$, then $x^{1/n} = 1$ for all positive integers n.*
(vi) *$(xy)^{1/n} = x^{1/n}y^{1/n}$.*
(vii) *$(x^{1/n})^{1/m} = x^{1/(nm)}$.*

Exercise 2.5.17. Prove Lemma 2.5.16.

Remark 2.5.18. Note the following cancellation law from Lemma 2.5.16(ii). If x, y are positive real numbers, and if $x^n = y^n$ for a positive integer n, then $x = y$. Note that the positivity of x, y is needed, since $(-3)^2 = 3^2$, but $3 \neq -3$.

Given a positive x and a rational number q, we can now define x^q. Due to the density of rational numbers within the real numbers, we therefore come very close to a general definition of x^y where y is real.

Definition 2.5.19 (Exponentiation to a Rational). Let $x > 0$ be a positive real number, and let q be a rational number. We now define x^q. Write $q = a/b$ where a is an integer, and b is a positive integer. We then define

$$x^q := (x^{1/b})^a.$$

We now show that this definition is well-defined.

Lemma 2.5.20. *Let a, a' be integers and let b, b' be positive integers such that $a/b = a'/b'$. Let x be a positive real number. Then $(x^{1/b})^a = (x^{1/b'})^{a'}$.*

Proof. We consider three cases: $a = 0$, $a < 0$, and $a > 0$. If $a = 0$, then we must have $a' = 0$ since $a/b = a'/b'$, so $(x^{1/b})^0 = 1 = (x^{1/b'})^0$, as desired.

If $a > 0$, then $a' > 0$ since $a/b = a'/b'$, and $a, b, b' > 0$. Define $y := x^{1/(ab')}$. Since $ab' = a'b$, we have $y = x^{1/(a'b)}$. From Lemma 2.5.16(vii), $y = (x^{1/b})^{1/a'} = (x^{1/b'})^{1/a}$. From Lemma 2.5.16(ii), we therefore have $y^{a'} = x^{1/b}$ and $y^a = x^{1/b'}$. So,

$$(x^{1/b'})^{a'} = (y^a)^{a'} = y^{aa'} = (y^{a'})^a = (x^{1/b})^a.$$

So, the case $a > 0$ is done. Finally, suppose $a < 0$. Then $a' < 0$ as well, so $-a$ and $-a'$ are positive. From the previous case, $(x^{1/b})^{-a} = (x^{1/b'})^{-a'}$. Taking the reciprocal of both sides completes the proof. □

Lemma 2.5.21. *Let $x, y > 0$ be positive real numbers, and let q, r be rational numbers.*

(i) *x^q is a positive real number.*
(ii) *$x^{q+r} = x^q x^r$ and $(x^q)^r = x^{qr}$.*
(iii) *$x^{-q} = 1/x^q$.*
(iv) *If $q > 0$, then $x > y$ if and only if $x^q > y^q$.*
(v) *If $x > 1$, then $x^q > x^r$ if and only if $q > r$. If $x < 1$, then $x^q > x^r$ if and only if $q < r$.*

Exercise 2.5.22. Prove Lemma 2.5.21.

2.5.2. *Some Standard Limits.* We can now compute some standard limits.

Remark 2.5.23. Let c be a real number. Then $\lim_{n\to\infty} c = c$.

Proposition 2.2.8 gives us the following.

Corollary 2.5.24. *For any positive integer k, we have $\lim_{n\to\infty} 1/(n^{1/k}) = 0$.*

Proof. From Lemma 2.5.16, $1/(n^{1/k})$ is decreasing in n and bounded below by 0. By Proposition 2.4.6 (for decreasing sequences bounded from below), there exists a real number $L \geq 0$ such that

$$L = \lim_{n\to\infty} 1/(n^{1/k}).$$

Taking both sides to the power k, and applying Theorem 2.2.15(ii) k times,

$$L^k = [\lim_{n\to\infty} 1/(n^{1/k})]^k = \lim_{n\to\infty} 1/(n^{k/k}) = \lim_{n\to\infty} (1/n) = 0.$$

The last equality follows from Proposition 2.2.8. Since $L^k = 0$, we know that L is not positive by Lemma 2.5.21(i). Since $L \geq 0$, we conclude that $L = 0$, as desired. □

Remark 2.5.25. By using the limit laws as in Corollary 2.5.24, it follows that, for any positive rational $q > 0$, we have $\lim_{n\to\infty} 1/(n^q) = 0$. Consequently, n^q does not converge as $n \to \infty$.

Exercise 2.5.26. Let $-1 < x < 1$. Then $\lim_{n\to\infty} x^n = 0$. Using the identity $(1/x^n)x^n = 1$ for $x > 1$, conclude that x^n does not converge as $n \to \infty$ for $x > 1$.

Lemma 2.5.27. *For any $x > 0$, we have $\lim_{n\to\infty} x^{1/n} = 1$.*

Exercise 2.5.28. Prove Lemma 2.5.27. (Hint: first, given any $\varepsilon > 0$, show that $(1+\varepsilon)^n$ has no real upper bound M, as $n \to \infty$. To prove this claim, set $x = 1/(1+\varepsilon)$ and use Exercise 2.5.26. Now, with this preliminary claim, show that for any $\varepsilon > 0$ and for any real M, there exists a positive integer n such that $M^{1/n} < 1+\varepsilon$. Now, use these two claims, and consider the cases $y > 1$ and $y < 1$ separately.)

2.6. **Infinite Series.** We will now begin our discussion of infinite series. One reason to care about infinite series is that Fourier analysis essentially reduces the study of certain functions to the study of infinite series. For another motivation, our study of infinite series is a precursor to the study of sequences of functions, and to the study of integrals. So, the study of infinite series provides a foundation for several other important topics.

We will briefly discuss finite series, and we will then move on to infinite series.

2.6.1. *Finite Series.*

Definition 2.6.1 (**Finite Series/ Finite Sum**)**.** Let m, n be integers. Let $(a_i)_{i=m}^n$ be a finite sequence of real numbers. Define the **finite sum** $\sum_{i=m}^n a_i$ by the recursive formula

$$\sum_{i=m}^{n} a_i := 0 \qquad , \text{ if } n < m$$

$$\sum_{i=m}^{n+1} a_i := (\sum_{i=m}^{n} a_i) + a_{n+1} \quad , \text{ if } n \geq m-1$$

Remark 2.6.2. To clarify the expressions we have used, a series is an expression of the form $\sum_{i=m}^n a_i$, and this series is equal to a real number, which is itself the sum of the series. The distinction between series and sum is not really important.

The following properties of summation can be proven by various inductive arguments.

Lemma 2.6.3. *Let $m \leq n < p$ be integers, and let $(a_i)_{i=m}^n, (b_i)_{i=m}^n$ be a sequences of real numbers, let k be an integer, and let c be a real number. Then*

- $$\sum_{i=m}^{n} a_i + \sum_{i=n+1}^{p} a_i = \sum_{i=m}^{p} a_i.$$
- $$\sum_{i=m}^{n} a_i = \sum_{j=m+k}^{n+k} a_{j-k}.$$
- $$\sum_{i=m}^{n} (a_i + b_i) = (\sum_{i=m}^{n} a_i) + (\sum_{i=m}^{n} b_i).$$
- $$\sum_{i=m}^{n} (c a_i) = c(\sum_{i=m}^{n} a_i).$$
- $$\left|\sum_{i=m}^{n} a_i\right| \leq \sum_{i=m}^{n} |a_i|.$$
- *If $a_i \leq b_i$ for all $m \leq i \leq n$, then* $\displaystyle\sum_{i=m}^{n} a_i \leq \sum_{i=m}^{n} b_i.$

Exercise 2.6.4. Prove Lemma 2.6.3.

We can also define sums over finite sets.

Definition 2.6.5. Let X be a finite set of cardinality $n \in \mathbb{N}$. Let $f\colon X \to \mathbb{R}$ be a function. We define $\sum_{x\in X} f(x)$ as follows. Let $g\colon \{1, 2, \ldots, n\} \to X$ be any bijection, which exists since X has cardinality n. We then define

$$\sum_{x\in X} f(x) := \sum_{i=1}^{n} f(g(i)).$$

Exercise 2.6.6. Show that this definition is well defined. That is, for any two bijections $g, h\colon \{1, 2, \ldots, n\} \to X$, we have $\sum_{i=1}^{n} f(g(i)) = \sum_{i=1}^{n} f(h(i))$.

Lemma 2.6.3 translates readily to sums over finite sets.

Lemma 2.6.7.

(i) *If X is empty and if $f\colon X \to \mathbb{R}$ is a function, then*

$$\sum_{x\in X} f(x) = 0.$$

(ii) *If $X = \{x_0\}$ consists of a single element and if $f\colon X \to \mathbb{R}$ is a function, then*

$$\sum_{x\in X} f(x) = f(x_0).$$

(iii) *If X is a finite set, if $f\colon X \to \mathbb{R}$ is a function, and if $g\colon Y \to X$ is a bijection between sets X, Y, then*

$$\sum_{x\in X} f(x) = \sum_{y\in Y} f(g(y)).$$

(iv) *Let $m \leq n$ be integers, let $(a_i)_{i=m}^{n}$ be a sequence of real numbers, and let $X = \{i \in \mathbb{N}\colon m \leq i \leq n\}$. Then*

$$\sum_{i=m}^{n} a_i = \sum_{i\in X} a_i.$$

(v) *Let X, Y be disjoint finite sets (so $X \cap Y = \emptyset$). Let $f\colon X \cup Y \to \mathbb{R}$ be a function. Then*

$$\sum_{x\in X\cup Y} f(x) = (\sum_{x\in X} f(x)) + (\sum_{y\in Y} f(y)).$$

(vi) *Let X be a finite set, let $f\colon X \to \mathbb{R}$ and let $g\colon X \to \mathbb{R}$ be functions. Then*

$$\sum_{x\in X} (f(x) + g(x)) = (\sum_{x\in X} f(x)) + (\sum_{x\in X} g(x)).$$

(vii) *Let X be a finite set, let $f\colon X \to \mathbb{R}$ be a function, and let $c \in \mathbb{R}$. Then*

$$\sum_{x\in X} (cf(x)) = c(\sum_{x\in X} f(x)).$$

(viii) *Let X be a finite set, let $f\colon X \to \mathbb{R}$ and let $g\colon X \to \mathbb{R}$ be functions such that $f(x) \leq g(x)$ for all $x \in X$. Then*

$$\sum_{x\in X} f(x) \leq \sum_{x\in X} g(x).$$

(ix) *Let X be a finite set, and let $f\colon X \to \mathbb{R}$ be a function. Then*

$$\left|\sum_{x\in X} f(x)\right| \leq \sum_{x\in X} |f(x)|.$$

Exercise 2.6.8. Prove Lemma 2.6.7.

Lemma 2.6.9. *Let X, Y be finite sets. Let $f\colon (X \times Y) \to \mathbb{R}$ be a function. Then*

$$\sum_{x\in X}(\sum_{y\in Y} f(x,y)) = \sum_{(x,y)\in X\times Y} f(x,y).$$

Exercise 2.6.10. Prove Lemma 2.6.9 by induction on the size of X.

Corollary 2.6.11 (Fubini's Theorem for finite sets). *Let X, Y be finite sets, and let $f\colon X \times Y \to \mathbb{R}$ be a function. Then*

$$\sum_{x\in X}(\sum_{y\in Y} f(x,y)) = \sum_{(x,y)\in X\times Y} f(x,y) = \sum_{(y,x)\in Y\times X} f(x,y) = \sum_{y\in Y}(\sum_{x\in X} f(x,y)).$$

Proof. Lemma 2.6.9 gives the first and last equalities. For the remaining middle equality, note that $g\colon X \times Y \to Y \times X$ defined by $g(x,y) := (y,x)$ is a bijection. So, Lemma 2.6.7(iii) completes the proof. □

Remark 2.6.12. As we saw in the first homework, Corollary 2.6.11 is false for infinite sums. So, we can already see that more care is needed when we pass to infinite sums.

2.6.2. *Infinite Series.*

Definition 2.6.13 (Infinite Series). Let $(a_n)_{n=m}^{\infty}$ be a sequence of real numbers. An infinite series is any formal expression of the form

$$\sum_{n=m}^{\infty} a_n.$$

We sometimes write this series as

$$a_m + a_{m+1} + a_{m+2} + \cdots .$$

So far, we have only given a formal definition for the expression $\sum_{n=m}^{\infty} a_n$. The sum only makes sense as a real number via the following definition.

Definition 2.6.14 (Convergent Sum). Let $\sum_{n=m}^{\infty} a_n$ be a formal infinite series. For any integer $N \geq m$, define the N^{th} **partial sum** S_N of this series by $S_N := \sum_{n=m}^{N} a_n$. Note that S_N is a real number. If the sequence $(S_N)_{N=m}^{\infty}$ converges to some limit L as $N \to \infty$, then we say that the infinite series $\sum_{n=m}^{\infty} a_n$ is **convergent**, and this infinite series **converges to** L. We also write $L = \sum_{n=m}^{\infty} a_n$ and say that L is the **sum** of the infinite series $\sum_{n=m}^{\infty} a_n$. If the partial sums diverge, then we say that the infinite series $\sum_{n=m}^{\infty} a_n$ is **divergent**, and we do not assign any real number to the infinite series $\sum_{n=m}^{\infty} a_n$.

Proposition 2.6.15. *Let $\sum_{n=m}^{\infty} a_n$ be a formal series of real numbers. Then $\sum_{n=m}^{\infty} a_n$ converges if and only if: for every real number $\varepsilon > 0$, there exists an integer $N \geq M$ such that, for all $p, q \geq N$,*

$$\left|\sum_{n=p}^{q} a_n\right| < \varepsilon.$$

Exercise 2.6.16. Prove Proposition 2.6.15. (Hint: use Theorem 2.2.10).

Corollary 2.6.17 (Zero Test). *Let $\sum_{n=m}^{\infty} a_n$ be a formal series of real numbers. If $\sum_{n=m}^{\infty} a_n$ converges, then $\lim_{n\to\infty} a_n = 0$. Note that the contrapositive says: if a_n does not converge to zero as $n \to \infty$, then $\sum_{n=m}^{\infty} a_n$ does not converge.*

Exercise 2.6.18. Using Proposition 2.6.15, prove Corollary 2.6.17.

Remark 2.6.19. The converse of Corollary 2.6.17 is false. For example, the series $\sum_{n=1}^{\infty} 1/n$ does not converge. On the other hand, as we will see below, the series $\sum_{n=1}^{\infty}(-1)^n/n$ does converge.

Definition 2.6.20 (Absolute Convergence). Let $\sum_{n=m}^{\infty} a_n$ be a formal series of real numbers. We say that the series $\sum_{n=m}^{\infty} a_n$ is **absolutely convergent** if and only if the series $\sum_{n=m}^{\infty} |a_n|$ is convergent. If a series is not absolutely convergent, then it is absolutely divergent.

Proposition 2.6.21. *Let $\sum_{n=m}^{\infty} a_n$ be a formal series of real numbers. If this series is absolutely convergent, then it is convergent. Moreover,*

$$\left|\sum_{n=m}^{\infty} a_n\right| \leq \sum_{n=m}^{\infty} |a_n|.$$

Exercise 2.6.22. Prove Proposition 2.6.21.

Proposition 2.6.23 (Alternating Series Test). *Let $(a_n)_{n=m}^{\infty}$ be a decreasing sequence of nonnegative real numbers. That is, $a_{n+1} \leq a_n$ and $a_n \geq 0$ for all $n \geq m$. Then the series $\sum_{n=m}^{\infty}(-1)^n a_n$ converges if and only if $a_n \to 0$ as $n \to \infty$.*

Proof. Suppose $\sum_{n=m}^{\infty}(-1)^n a_n$ converges. From the Zero Test (Corollary 2.6.17), we know that $(-1)^n a_n \to 0$ as $n \to \infty$. Therefore, $a_n \to 0$ as $n \to \infty$ as desired.

We now prove the converse. The idea is that looking only at even partial sums (or odd partial sums) reveals a monotonicity of the sequence. Suppose $\lim_{n\to\infty} a_n = 0$. Let $N \geq m$ and define $S_N := \sum_{n=m}^{N}(-1)^n a_n$. Note that

$$S_{N+2} = S_N + (-1)^{N+1}a_{N+1} + (-1)^{N+2}a_{N+2} = S_N + (-1)^{N+1}(a_{N+1} - a_{N+2}).$$

Recall that $a_{N+1} \geq a_{N+2}$. So, if N is odd, then $S_{N+2} \geq S_N$, and if N is even, $S_{N+2} \leq S_N$.

Suppose N is even. Then for any natural number k, $S_{N+2k} \leq S_N$. Also, $S_{N+2k+1} \geq S_{N+1} = S_N - a_{N+1}$, and $S_{N+2k+1} = S_{N+2k} - a_{N+2k+1} \leq S_{N+2k}$ since $a_{N+2k+1} \geq 0$. So, for any natural number k,

$$S_N - a_{N+1} \leq S_{N+2k+1} \leq S_{N+2k} \leq S_N.$$

In summary, for any integer $n \geq N$,

$$S_N - a_{N+1} \leq S_n \leq S_N.$$

Using the assumption $a_n \to 0$, if we are given any $\varepsilon > 0$, there exists a natural number N such that, for all $n > N$, we have $|a_n| < \varepsilon$, so that

$$S_N - \varepsilon \leq S_n \leq S_N.$$

That is, for any $\varepsilon > 0$, there exists a natural number N such that, for all $j, k > N$, we have $|S_j - S_k| < \varepsilon$. So, the sequence $(S_n)_{n=m}^{\infty}$ is a Cauchy sequence, and it therefore converges by Theorem 2.2.10. □

The following Proposition should be contrasted with Lemma 2.6.3. Note in particular the extra assumptions that are needed in the following statements.

Proposition 2.6.24.

- *Let $\sum_{n=m}^{\infty} a_n$ be a series of real numbers converging to x, and let $\sum_{n=m}^{\infty} b_n$ be a series of real numbers converging to y. Then $\sum_{n=m}^{\infty}(a_n + b_n)$ is a convergent series that converges to $x + y$. That is,*

$$\sum_{n=m}^{\infty}(a_n + b_n) = (\sum_{n=m}^{\infty} a_n) + (\sum_{n=m}^{\infty} b_n).$$

- *Let $\sum_{n=m}^{\infty} a_n$ be a series of real numbers converging to x, and let c be a real number. Then $\sum_{n=m}^{\infty}(ca_n)$ is a convergent series that converges to cx. That is,*

$$\sum_{n=m}^{\infty}(ca_n) = c(\sum_{n=m}^{\infty} a_n).$$

- *Let $\sum_{n=m}^{\infty} a_n$ be a series of real numbers, and let k be a natural number. If one of the two series $\sum_{n=m}^{\infty} a_n$ or $\sum_{n=m+k}^{\infty} a_n$ converges, then the other also converges, and we have*

$$\sum_{n=m}^{\infty} a_n = (\sum_{n=m}^{m+k-1} a_n) + (\sum_{n=m+k}^{\infty} a_n).$$

- *Let $\sum_{n=m}^{\infty} a_n$ be a series of real numbers converging to x, and let k be an integer. Then $\sum_{n=m+k}^{\infty} a_{n-k}$ also converges to x.*

Exercise 2.6.25. Prove Proposition 2.6.24.

Remark 2.6.26. From Proposition 2.6.24, changing any finite number of terms of a series does not affect the convergence of the series. We will therefore eventually de-emphasize the starting index of a series.

2.6.3. *Sums of Nonnegative Numbers.* From Proposition 2.6.21, if a series converges absolutely, then it also converges. In practice, we often show that a series converges by showing that it is absolutely convergent. Therefore, it is nice to have several ways to show whether or not a series is absolutely convergent. In other words, given a series of nonnegative numbers, it is desirable to verify its convergence. So, in this section, we will discuss series of nonnegative numbers.

Let $\sum_{n=m}^{\infty} a_n$ be a series of nonnegative real numbers. Since $a_n \geq 0$ for all $n \geq m$, the partial sums $S_N := \sum_{n=m}^{N} a_n$ are increasing. That is, $S_{N+1} \geq S_N$ for all integers $N \geq m$. From Remark 2.4.9, $(S_N)_{N=m}^{\infty}$ is convergent if and only if it has an upper bound M. We summarize this discussion as follows.

Proposition 2.6.27. *Let $\sum_{n=m}^{\infty} a_n$ be a formal series of nonnegative real numbers. Then this series is convergent if and only if there exists a real number M such that, for all integers $N \geq m$, we have*

$$\sum_{n=m}^{N} a_n \leq M.$$

Corollary 2.6.28 (**Comparison Test**). *Let $\sum_{n=m}^{\infty} a_n, \sum_{n=m}^{\infty} b_n$ be formal series of real numbers. Assume that $|a_n| \leq b_n$ for all $n \geq m$. If $\sum_{n=m}^{\infty} b_n$ is convergent, then $\sum_{n=m}^{\infty} a_n$ is absolutely convergent. Moreover,*

$$\left|\sum_{n=m}^{\infty} a_n\right| \leq \sum_{n=m}^{\infty} |a_n| \leq \sum_{n=m}^{\infty} b_n.$$

Exercise 2.6.29. Prove Corollary 2.6.28.

Remark 2.6.30. The contrapositive of Corollary 2.6.28 says: if $|a_n| \leq b_n$ for all $n \geq m$, and if $\sum_{n=m}^{\infty} a_n$ is absolutely divergent, then $\sum_{n=m}^{\infty} b_n$ does not converge.

Example 2.6.31. Let x be a real number and consider the series

$$\sum_{n=0}^{\infty} x^n.$$

If $|x| \geq 1$, then this series diverges by the Zero Test (Corollary 2.6.17). If $|x| < 1$, then we can use induction to show that the partial sums satisfy

$$\sum_{n=0}^{N} x^n = (1 - x^{N+1})/(1 - x). \qquad (*)$$

If $|x| < 1$ then $\lim_{N\to\infty} x^N = 0$ by Exercise 2.5.26. So, using the Limit Laws,

$$\lim_{N\to\infty} (1 - x^{N+1})/(1 - x) = 1/(1 - x).$$

So, $\sum_{n=0}^{\infty} x^n$ converges to $1/(1 - x)$ when $|x| < 1$. Moreover, this convergence is absolute, by Corollary 2.6.28.

Proposition 2.6.32 (**Dyadic Criterion**). *Let $(a_n)_{n=1}^{\infty}$ be a decreasing sequence of nonnegative real numbers. That is, $a_{n+1} \leq a_n$ and $a_n \geq 0$ for all $n \geq 1$. Then the series $\sum_{n=1}^{\infty} a_n$ converges if and only if the following series converges:*

$$\sum_{k=0}^{\infty} 2^k a_{(2^k)} = a_1 + 2a_2 + 4a_4 + 8a_8 + \cdots .$$

Proof. Let N be a positive integer and let K be a natural number. Let $S_N := \sum_{n=1}^{N} a_n$, and let $T_K := \sum_{k=0}^{K} 2^k a_{2^k}$. We claim that

$$S_{2^{K+1}-1} \leq T_K \leq 2S_{2^K}. \qquad (*)$$

We prove this claim by induction on K. In the case $K = 0$, we want to show $S_1 \leq T_0 \leq 2S_1$. Now, $S_1 = a_1$ and $T_0 = a_1$, so $S_1 \leq T_0 \leq 2S_1$ holds.

We now prove the inductive step. Suppose $(*)$ holds for some K. Then, note that

$$S_{2^{K+2}-1} = S_{2^{K+1}-1} + \sum_{n=2^{K+1}}^{2^{K+2}-1} a_n \leq S_{2^{K+1}-1} + \sum_{n=2^{K+1}}^{2^{K+2}-1} a_{2^{K+1}} = S_{2^{K+1}-1} + 2^{K+1}a_{2^{K+1}}.$$

Similarly,

$$S_{2^{K+1}} = S_{2^K} + \sum_{n=2^K+1}^{2^{K+1}} a_n \geq S_{2^K} + \sum_{n=2^K+1}^{2^{K+1}} a_{2^{K+1}} = S_{2^K} + 2^K a_{2^{K+1}}.$$

So, applying the inductive hypothesis,

$$S_{2^{K+2}-1} \leq T_K + 2^{K+1}a_{2^{K+1}} = T_{K+1}.$$

$$2S_{2^{K+1}} \geq T_K + 2^{K+1}a_{2^{K+1}} = T_{K+1}$$

So, we have completed the inductive step for $(*)$, thereby proving $(*)$.

We can now use $(*)$ to complete the proof. If $\sum_{n=1}^{\infty} a_n$ converges, then the partial sums S_{2^K} are bounded as $K \to \infty$ by Proposition 2.6.27. So the right inequality of $(*)$ shows that the partial sums T_K are bounded as $K \to \infty$. So, by Proposition 2.6.27, $\sum_{k=0}^{\infty} 2^k a_{(2^k)}$ converges. Conversely, suppose $\sum_{k=0}^{\infty} 2^k a_{(2^k)}$ converges. Then the partial sums T_K are bounded as $K \to \infty$ by Proposition 2.6.27. By $(*)$, the partial sums S_{2^K} are bounded as $K \to \infty$. Now, for any positive integer n, there exists a natural number N such that $n \leq 2^N$. So, since $S_n \leq S_{n+1}$ for all natural numbers n, we conclude that $S_n \leq S_{2^N}$. So, the partial sums S_n are bounded as $n \to \infty$. That is, $\sum_{n=1}^{\infty} a_n$ converges, by Proposition 2.6.27. □

Corollary 2.6.33. *Let $q > 0$ be a rational number. Then the series $\sum_{n=1}^{\infty} 1/n^q$ is convergent when $q > 1$ and it is divergent when $q \leq 1$.*

Proof. The sequence $(1/n^q)_{n=1}^{\infty}$ is nonnegative and decreasing by Lemma 2.5.21(iv). We can therefore apply the Dyadic Criterion (Theorem 2.6.32). The series $\sum_{n=1}^{\infty} 1/n^q$ is then convergent if and only if the following series is convergent

$$\sum_{k=0}^{\infty} 2^k \frac{1}{(2^k)^q} = \sum_{k=0}^{\infty} (2^{1-q})^k.$$

In the last equality, we used Lemma 2.5.21(ii). In Example 2.6.31, we showed that the geometric series $\sum_{k=0}^{\infty} x^k$ is convergent if and only if $|x| < 1$. So, the series $\sum_{n=1}^{\infty} 1/n^q$ is convergent if and only if $|2^{1-q}| < 1$, i.e. if and only if $q > 1$. (The last claim follows by Lemma 2.5.21.) □

Remark 2.6.34. In particular, the **harmonic series** $\sum_{n=1}^{\infty} 1/n$ diverges.

2.6.4. *Rearrangement of Series.* Let $(a_n)_{n=1}^{N}$ be a sequence of real numbers. From Exercise 2.6.6, any rearrangement of a finite series gives the same sum. That is, for any bijection $g\colon \{1, \ldots, n\} \to \{1, \ldots, n\}$, we have

$$\sum_{n=1}^{N} a_n = \sum_{n=1}^{N} a_{g(n)}.$$

The corresponding statement for infinite series is false. For example, consider the sequence $a_n = (-1)^{n+1}/(n+1)$. Recall that $\sum_{n=0}^{\infty} a_n$ converges by the Alternating Series Test (Proposition 2.6.23). However, there exists a bijection $g\colon \mathbb{N} \to \mathbb{N}$ such that $\sum_{n=0}^{\infty} a_n$ actually *diverges*. So, we cannot rearrange convergent infinite series and expect the sum of the rearranged series to be the same or even to converge at all.

Exercise 2.6.35. For any $n \in \mathbb{N}$, define $a_n := (-1)^{n+1}/(n+1)$. Find a bijection $g\colon \mathbb{N} \to \mathbb{N}$ such that the series $\sum_{n=0}^{\infty} a_{g(n)}$ diverges.

In fact, given any real number L, the series $\sum_{n=1}^{\infty}(-1)^n/n$ can be rearranged so that the rearranged series converges to L.

Theorem 2.6.36. *Let $\sum_{n=0}^{\infty} a_n$ be a convergent series which is not absolutely convergent. Let L be a real number. Then there exists a bijection $g\colon \mathbb{N} \to \mathbb{N}$ such that $\sum_{n=0}^{\infty} a_{g(n)}$ converges to L.*

However, we can rearrange absolutely convergent series.

Proposition 2.6.37. *Let $\sum_{n=0}^{\infty} a_n$ be an absolutely convergent series of real numbers. Let $g\colon \mathbb{N} \to \mathbb{N}$ be a bijection. Then $\sum_{m=0}^{\infty} a_{g(m)}$ is also convergent. Moreover,*

$$\sum_{n=0}^{\infty} a_n = \sum_{m=0}^{\infty} a_{g(m)}.$$

2.7. **Ratio and Root Tests.** The following test for series generalizes our investigation of the convergence of the geometric series from Example 2.6.31.

Theorem 2.7.1 (Root Test). *Let $\sum_{n=m}^{\infty} a_n$ be a series of real numbers. Define $\alpha := \limsup_{n\to\infty} |a_n|^{1/n}$.*

- *If $\alpha < 1$, then the series $\sum_{n=m}^{\infty} a_n$ is absolutely convergent. In particular, the series $\sum_{n=m}^{\infty} a_n$ is convergent.*
- *If $\alpha > 1$, then the series $\sum_{n=m}^{\infty} a_n$ is divergent.*
- *If $\alpha = 1$, no conclusion is asserted.*

Proof. First assume that $\alpha < 1$. Since $|a_n|^{1/n} \geq 0$ for every positive integer n, we know that $\alpha \geq 0$. Let $\varepsilon > 0$ so that $\varepsilon + \alpha < 1$. (For example, we could let $\varepsilon := (1-\alpha)/2$.) By Proposition 2.5.7(i), there exists an integer N such that, for all $n \geq N$, we have $|a_n|^{1/n} \leq (\alpha + \varepsilon)$. That is, $|a_n| \leq (\alpha+\varepsilon)^n$. Since $0 < \alpha + \varepsilon < 1$, the geometric series $\sum_{n=N}^{\infty}(\alpha+\varepsilon)^n$ converges. So, by the Comparison Test (Corollary 2.6.28), $\sum_{n=N}^{\infty} a_n$ converges. Therefore, $\sum_{n=m}^{\infty} a_n$ converges by Lemma 2.6.3, since a finite number of terms do not affect the convergence of the infinite sum.

Now, assume that $\alpha > 1$. By Proposition 2.5.7(ii), for every $N \geq m$ there exists $n \geq N$ such that $|a_n|^{1/n} \geq 1$. That is, $|a_n| \geq 1$. In particular, a_n does not converge to zero as $n \to \infty$. So, by the Zero Test (Corollary 2.6.17), we conclude that $\sum_{n=m}^{\infty} a_n$ does not converge. □

The Root Test is not always easy to use directly, but we can replace the roots by ratios, which are sometimes easier to handle.

Lemma 2.7.2. *Let* $(b_n)_{n=m}^{\infty}$ *be a sequence of positive numbers. Then*

$$\liminf_{n\to\infty} \frac{b_{n+1}}{b_n} \leq \liminf_{n\to\infty} b_n^{1/n} \leq \limsup_{n\to\infty} b_n^{1/n} \leq \limsup_{n\to\infty} \frac{b_{n+1}}{b_n}.$$

Proof. The middle inequality is Proposition 2.5.7(iii). We will only then prove the right inequality.

Let $L := \limsup_{n\to\infty} \frac{b_{n+1}}{b_n}$. If $L = +\infty$ there is nothing to show, so we assume that $L < +\infty$. Since b_n is positive for each $n \geq m$, we know that $L \geq 0$.

Let $\varepsilon > 0$. From Proposition 2.5.7(i), there exists an integer $N \geq m$ such that, for all $n \geq N$, we have $(b_{n+1}/b_n) \leq L + \varepsilon$. That is, $b_{n+1} \leq (L+\varepsilon)b_n$. By induction, we conclude that, for all $n \geq N$,

$$b_n \leq (L+\varepsilon)^{n-N} b_N.$$

That is, for all $n \geq N$,

$$b_n^{1/n} \leq (b_N(L+\varepsilon)^{-N})^{1/n}(L+\varepsilon). \qquad (*)$$

Letting $n \to \infty$ on the right side of $(*)$, and applying the Limit Laws and Lemma 2.5.27,

$$\lim_{n\to\infty} (b_N(L+\varepsilon)^{-N})^{1/n}(L+\varepsilon) = L + \varepsilon.$$

So, applying the Comparison Principle (Lemma 2.5.10 to $(*)$,

$$\limsup_{n\to\infty} b_n^{1/n} \leq L + \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, we conclude that $\limsup_{n\to\infty} b_n^{1/n} \leq L$, as desired. □

Exercise 2.7.3. Prove the left inequality of Lemma 2.7.2.

Combining Theorem 2.7.1 and Lemma 2.7.2 gives the following.

Corollary 2.7.4 (Ratio Test). *Let* $\sum_{n=m}^{\infty} a_n$ *be a series of nonzero numbers. (So,* a_{n+1}/a_n *is defined for any* $n \geq m$*.)*

- *If* $\limsup_{n\to\infty} \frac{|a_{n+1}|}{|a_n|} < 1$*, then the series* $\sum_{n=m}^{\infty} a_n$ *is absolutely convergent. In particular,* $\sum_{n=m}^{\infty} a_n$ *is convergent.*
- *If* $\liminf_{n\to\infty} \frac{|a_{n+1}|}{|a_n|} > 1$*, then the series* $\sum_{n=m}^{\infty} a_n$ *is divergent. In particular,* $\sum_{n=m}^{\infty} a_n$ *is not absolutely convergent.*

2.8. **Subsequences.** Our investigation now shifts attention from series back to sequences. We focus our attention on ways to decompose a sequence into smaller parts which are easier to understand. One popular paradigm in mathematics (and in science more generally) is to take a complicated object and break it into pieces which are simpler to understand. Subsequences are one manifestation of this paradigm.

Definition 2.8.1 (Subsequence). Let $(a_n)_{n=0}^{\infty}, (b_n)_{n=0}^{\infty}$ be sequences of real numbers. We say that $(b_n)_{n=0}^{\infty}$ is a **subsequence** of $(a_n)_{n=0}^{\infty}$ if and only if there exists a function $f\colon \mathbb{N} \to \mathbb{N}$ which is strictly increasing (i.e. $f(n+1) > f(n)$ for all $n \in \mathbb{N}$) such that, for all $n \in \mathbb{N}$,

$$b_n = a_{f(n)}$$

Example 2.8.2. The sequence $(a_{2n})_{n=0}^{\infty}$ is a subsequence of $(a_n)_{n=0}^{\infty}$, since $f(n) := 2n$ is an increasing function from $\mathbb{N}$ to $\mathbb{N}$, and $a_{2n} = a_{f(n)}$.

Here are some basic properties of subsequences.

Lemma 2.8.3. *Let $(a_n)_{n=0}^{\infty}, (b_n)_{n=0}^{\infty}, (c_n)_{n=0}^{\infty}$ be sequences of real numbers. Then $(a_n)_{n=0}^{\infty}$ is a subsequence of $(a_n)_{n=0}^{\infty}$. Also, if $(b_n)_{n=0}^{\infty}$ is a subsequence of $(a_n)_{n=0}^{\infty}$, and if $(c_n)_{n=0}^{\infty}$ is a subsequence of $(b_n)_{n=0}^{\infty}$, then $(c_n)_{n=0}^{\infty}$ is a subsequence of $(a_n)_{n=0}^{\infty}$.*

Exercise 2.8.4. Prove Lemma 2.8.3.

Subsequences and limits are closely related, as we now show.

Proposition 2.8.5. *Let $(a_n)_{n=0}^{\infty}$ be a sequence of real numbers, and let L be a real number.*

- *If the sequence $(a_n)_{n=0}^{\infty}$ converges to L, then every subsequence of $(a_n)_{n=0}^{\infty}$ converges to L.*
- *Conversely, if every subsequence of $(a_n)_{n=0}^{\infty}$ converges to L, then $(a_n)_{n=0}^{\infty}$ itself converges to L.*

Exercise 2.8.6. Prove Proposition 2.8.5.

Proposition 2.8.7. *Let $(a_n)_{n=0}^{\infty}$ be a sequence of real numbers, and let L be a real number.*

- *Suppose L is a limit point of $(a_n)_{n=0}^{\infty}$. Then there exists a subsequence of $(a_n)_{n=0}^{\infty}$ which converges to L.*
- *Conversely, if there exists a sequence of $(a_n)_{n=0}^{\infty}$ which converges to L, then L is a limit point of $(a_n)_{n=0}^{\infty}$.*

Exercise 2.8.8. Prove Proposition 2.8.7.

The following important theorem says: every bounded sequence has a convergent subsequence.

Theorem 2.8.9 (Bolzano-Weierstrass). *Let $(a_n)_{n=0}^{\infty}$ be a bounded sequence. That is, there exists a real number M such that $|a_n| \leq M$ for all $n \in \mathbb{N}$. Then there exists a subsequence of $(a_n)_{n=0}^{\infty}$ which converges.*

Proof. Let $L := \limsup_{n\to\infty} a_n$. From the Comparison Principle (Lemma 2.5.10), $|L| \leq M$. In particular, L is a real number. So, by Proposition 2.5.7(v), L is a limit point of $(a_n)_{n=0}^{\infty}$. By Proposition 2.8.7, there exists a subsequence of $(a_n)_{n=0}^{\infty}$ which converges to L. □

Remark 2.8.10. Note that we could have defined $L := \liminf_{n\to\infty} a_n$ and the proof would have still worked.

3. Real functions, Continuity, Differentiability

3.1. **Functions on the real line.** We now focus our attention on functions on the real line $\mathbb{R}$, rather than functions on $\mathbb{N}$ (i.e. sequences). The properties of the real line $\mathbb{R}$, most notably its completeness property, allow functions on $\mathbb{R}$ to have additional properties that functions on $\mathbb{N}$ do not have. For example, we can define and understand continuity and differentiability.

Definition 3.1.1. Let X, Y be sets and let $f\colon X \to Y$ be a **function**. That is, for every $x \in X$, the function f assigns to x some element $f(x) \in Y$. We say that X is the **domain** of f.

Example 3.1.2. Some common domains for functions on the real line are:

- The positive half-line $\mathbb{R}^+ := \{x \in \mathbb{R} \colon x > 0\}$.
- The negative half-line $\mathbb{R}^- := \{x \in \mathbb{R} \colon x < 0\}$.
- The **closed intervals** $[a, b] := \{x \in \mathbb{R} \colon a \leq x \leq b\}$, $a, b \in \mathbb{R}$.
- The **open intervals** $(a, b) := \{x \in \mathbb{R} \colon a < x < b\}$, $a, b \in \mathbb{R}$.
- The **half-open intervals** $(a, b] := \{x \in \mathbb{R} \colon a < x \leq b\}$ and $[a, b) := \{x \in \mathbb{R} \colon a \leq x < b\}$, $a, b \in \mathbb{R}$.
- $[a, \infty) := \{x \in \mathbb{R} \colon a \leq x < \infty\}$, $(-\infty, a] := \{x \in \mathbb{R} \colon -\infty < x \leq a\}$.
- $(a, \infty) := \{x \in \mathbb{R} \colon a < x < \infty\}$, $(-\infty, a) := \{x \in \mathbb{R} \colon -\infty < x < a\}$.
- The entire real line $\mathbb{R} = (-\infty, \infty)$.

Definition 3.1.3 (Restriction). Given a function $f \colon \mathbb{R} \to \mathbb{R}$ and given a subset $X \subseteq \mathbb{R}$, define the **restriction** $f|_X$ of f to X so that, for any $x \in X$, $f|_X(x) := f(x)$.

Remark 3.1.4. One can similarly restrict the range of a function, if the function only takes values in a smaller range. For example, the function $f(x) := x^2$ is a function $f \colon \mathbb{R} \to \mathbb{R}$, but it can also be considered as a function $f \colon \mathbb{R} \to [0, \infty)$.

Remark 3.1.5. There is a distinction between a function $f \colon \mathbb{R} \to \mathbb{R}$ and its value $f(x)$ for $x \in \mathbb{R}$, but it is not that important. For example, if we use $f(x) := x^2$ with $f \colon \mathbb{R} \to \mathbb{R}$, and we let $g := f|_{[0,1]}$, then $g(x) = f(x)$ for all $x \in [0, 1]$. But f and g are not considered to be the same function, since their domains are different.

Definition 3.1.6 (Composition). Let $f \colon X \to Y$ and let $g \colon Y \to Z$ be functions. We define the **composition** $g \circ f$ by the formula $(g \circ f)(x) := g(f(x))$.

Definition 3.1.7 (Arithmetic of Functions). Real valued functions inherit the arithmetic of the real numbers as follows. Let $f, g \colon X \to \mathbb{R}$. Then the sum $(f + g) \colon X \to \mathbb{R}$ is defined so that, for all $x \in X$,

$$(f + g)(x) := f(x) + g(x).$$

The difference $(f - g) \colon X \to \mathbb{R}$ is defined so that, for all $x \in X$,

$$(f - g)(x) := f(x) - g(x).$$

The product $(fg) \colon X \to \mathbb{R}$ is defined so that, for all $x \in X$,

$$(fg)(x) := f(x)g(x).$$

If $g(x) \neq 0$ for all $x \in X$, then the quotient $(f/g) \colon X \to \mathbb{R}$ is defined so that, for all $x \in X$,

$$(f/g)(x) := f(x)/g(x).$$

If $c \in \mathbb{R}$, then the function $cf \colon X \to \mathbb{R}$ is defined so that, for all $x \in X$,

$$(cf)(x) := c(f(x)).$$

3.1.1. *Limits of Functions.*

Definition 3.1.8 (Adherent Point). Let E be a subset of $\mathbb{R}$, and let x be a real number. We say that x is an **adherent point** of E if and only if, for all $\varepsilon > 0$, there exists $y \in E$ such that $|x - y| < \varepsilon$.

Remark 3.1.9. All points in E are adherent points of E.

Definition 3.1.10 (Closure). Let E be a subset of $\mathbb{R}$. Then the **closure of** E, denoted $\overline{E}$, is defined to be the set of adherent points of E.

Proposition 3.1.11. *Let $a < b$ be real numbers. Let I be any of the four intervals (a, b), $(a, b]$, $[a, b)$ or $[a, b]$. Then the closure of I is $[a, b]$.*

Exercise 3.1.12. Prove Proposition 3.1.11.

Lemma 3.1.13. *Let X be a subset of $\mathbb{R}$, and let x be an element of $\mathbb{R}$. Then x is an adherent point of X if and only if there exists a sequence $(a_n)_{n=0}^{\infty}$ of elements of X such that $\lim_{n\to\infty} a_n = x$.*

Definition 3.1.14 (Convergence of a function). Let X be a subset of $\mathbb{R}$, let $f : X \to \mathbb{R}$ be a function, let E be a subset of X, let x_0 be an adherent point of E, and let L be a real number. We say that f **converges** to L at x_0 in E, and we write $\lim_{x\to x_0; x\in E} f(x) = L$ if and only if: for all $\varepsilon > 0$, there exists $\delta = \delta(\varepsilon)$ such that, for all $x \in E$ with $|x - x_0| < \delta$, we have $|f(x) - L| < \varepsilon$.

If f does not converge to any real number L at x_0, we say that f **diverges** at x_0, and we leave $\lim_{x\to x_0; x\in E} f(x)$ undefined.

Remark 3.1.15. We will often omit the set E from our notation and just write $\lim_{x\to x_0} f(x)$. However, we must be careful when doing this.

We can equivalently talk about convergence of f in terms of sequences in the domain of f, as we now show.

Proposition 3.1.16. *Let X be a subset of $\mathbb{R}$, let $f : X \to \mathbb{R}$ be a function, let E be a subset of X, let x_0 be an adherent point of E, and let L be a real number. Then the following two statements are equivalent. (That is, one statement is true if and only if the other statement is true.)*

- *f converges to L at x_0 in E.*
- *For every sequence $(a_n)_{n=0}^{\infty}$ which consists entirely of elements of E, and which converges to x_0, the sequence $(f(a_n))_{n=0}^{\infty}$ converges to L.*

Exercise 3.1.17. Prove Proposition 3.1.16.

Remark 3.1.18. Due to Proposition 3.1.16, we will sometimes say "$f(x)$ goes to L as $x \to x_0$ in E" or "f has limit L at x_0 in E" instead of "f converges to L at x_0" or "$\lim_{x\to x_0} f(x) = L$".

Corollary 3.1.19. *Let X be a subset of $\mathbb{R}$, let $f : X \to \mathbb{R}$ be a function, let E be a subset of X, let x_0 be an adherent point of E. Then f can have at most one limit at x_0 in E.*

Proof. Suppose f has two limits L, L' at x_0 in E. We will show that $L = L'$. Since x_0 is an adherent point of E, Lemma 3.1.13 says that there exists a sequence $(a_n)_{n=0}^{\infty}$ of elements of E such that $a_n \to x_0$ as $n \to \infty$. By Proposition 3.1.16, the sequence $(f(a_n))_{n=0}^{\infty}$ converges to both L and L' as $n \to \infty$. By Proposition 2.2.4, we conclude that $L = L'$, as desired. □

By Proposition 3.1.16, the Limit Laws for sequences (Theorem 2.2.15) then give analogous limit laws for functions.

Proposition 3.1.20 (Limit Laws for functions). *Let X be a subset of $\mathbb{R}$, let $f, g : X \to \mathbb{R}$ be functions, let E be a subset of X, let x_0 be an adherent point of E, and let c be a real number. Assume that f has limit L at x_0 in E, and g has limit M at x_0 in E. Then $f + g$ has limit $L + M$ at x_0 in E, $f - g$ has limit $L - M$ at x_0 in E, fg has limit LM at x_0 in E, and cf has limit cL at x_0 in E. If additionally $g(x) \neq 0$ for all $x \in E$ and $M \neq 0$, then f/g has limit L/M at x_0 in E.*

Proof. We only prove the first claim, since the others are proven similarly. Since x_0 is an adherent point of E, Lemma 3.1.13 says that there exists a sequence $(a_n)_{n=0}^{\infty}$ of elements of E such that $a_n \to x_0$ as $n \to \infty$. By Proposition 3.1.16, the sequence $(f(a_n))_{n=0}^{\infty}$ converges to L. Similarly, the sequence $(g(a_n))_{n=0}^{\infty}$ converges to M. By the Limit Laws for sequences (Theorem 2.2.15), the sequence $(f(a_n) + g(a_n))_{n=0}^{\infty}$ converges to $L + M$. By Proposition 3.1.16, we conclude that $f + g$ has limit $L + M$ at x_0 in E. $\square$

Remark 3.1.21. Let $c \in \mathbb{R}$. Using Proposition 3.1.16, we can verify the following limits

$$\lim_{x \to x_0; x \in \mathbb{R}} c = c.$$

$$\lim_{x \to x_0; x \in \mathbb{R}} x = x_0.$$

Then, using the limit laws of Proposition 3.1.16, we can e.g. compute

$$\lim_{x \to x_0; x \in \mathbb{R}} x^2 = x_0^2.$$

$$\lim_{x \to x_0; x \in \mathbb{R}} (x^2 + x) = x_0^2 + x_0.$$

Example 3.1.22. Let $f\colon \mathbb{R} \to \mathbb{R}$ so that

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

Then $\lim_{x \to 0; x \in (0,\infty)} f(x) = 1$ and $\lim_{x \to 0; x \in (-\infty,0)} f(x) = 0$. However, $\lim_{x \to 0; x \in [0,\infty)} f(x)$ and $\lim_{x \to 0; x \in \mathbb{R}} f(x)$ are both undefined.

Example 3.1.23. Let $f\colon \mathbb{R} \to \mathbb{R}$ so that

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

Then $\lim_{x \to 0; x \in \mathbb{R} \setminus \{0\}} f(x) = 0$, but $\lim_{x \to 0; x \in \mathbb{R}} f(x)$ is undefined.

Example 3.1.24. Let $f\colon \mathbb{R} \to \mathbb{R}$ so that

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

Then $\lim_{x \to 0; x \in \mathbb{R}} f(x)$ does not exist. To see this, consider the sequences $(1/n)_{n=1}^{\infty}$ and $(\sqrt{2}/n)_{n=1}^{\infty}$. Both sequences converge to zero as $n \to \infty$, though the first sequence consists of rational numbers, and the second sequence consists of irrational numbers. So, $f(1/n) \to 1$ as $n \to \infty$, while $f(\sqrt{2}/n) \to 0$ as $n \to \infty$. Therefore, $\lim_{x \to 0; x \in \mathbb{R}} f(x)$ does not exist.

The following proposition says that the limit of f at x_0 depends only on points near x_0.

Proposition 3.1.25. *Let X be a subset of $\mathbb{R}$, let $f\colon X \to \mathbb{R}$ be a function, let E be a subset of X, let x_0 be an adherent point of E, let L be a real number, and let δ be a positive real number. Then the following two statements are equivalent:*

- $\lim_{x \to x_0; x \in E} f(x) = L.$
- $\lim_{x \to x_0; x \in E \cap (x_0 - \delta, x_0 + \delta)} f(x) = L.$

Exercise 3.1.26. Prove Proposition 3.1.25.

3.2. **Continuous Functions.** As we saw from the examples in the previous section, there are many functions that behave very strangely with respect to limits. However, there are still large classes of functions that behave well with respect to limits. Such functions are called continuous.

When learning a new concept (such as continuous functions), it is often beneficial to consider various examples which satisfy or do not satisfy the properties of the new concept. We will therefore continue our family of examples from the previous section.

Definition 3.2.1 (**Continuous Function**)**.** Let X be a subset of $\mathbb{R}$ and let $f\colon X \to \mathbb{R}$ be a function. Let x_0 be an element of X. We say that f is **continuous** at x_0 if and only if

$$\lim_{x\to x_0; x\in X} f(x) = f(x_0).$$

That is, the limit of f at x_0 in X exists, and this limit is equal to $f(x_0)$. We say that f is **continuous on** X (or we just say that f is **continuous**) if and only if f is continuous at x_0 for every $x_0 \in X$. We say that f is **discontinuous** at x_0 if and only if f is not continuous at x_0.

Example 3.2.2. Let $f\colon \mathbb{R} \to \mathbb{R}$ so that

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

Then f is continuous on $\mathbb{R} \smallsetminus \{0\}$, but f is discontinuous at 0.

Example 3.2.3. Let $f\colon \mathbb{R} \to \mathbb{R}$ so that

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

Then f is continuous on $\mathbb{R} \smallsetminus \{0\}$, but f is discontinuous at 0. However, if we redefine f so that $f(0) := 0$, then f would be continuous on $\mathbb{R}$. We therefore say that f has a removable discontinuity at 0.

Example 3.2.4. Let $f\colon \mathbb{R} \to \mathbb{R}$ so that

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

As we saw previously, f is discontinuous at zero. In fact, f is discontinuous on all of $\mathbb{R}$.

Proposition 3.2.5. *Let X be a subset of $\mathbb{R}$, let $f\colon X \to \mathbb{R}$ be a function, and let $x_0 \in X$. Then the following three statements are equivalent.*

- *f is continuous at x_0*
- *For every sequence $(a_n)_{n=0}^{\infty}$ consisting of elements of X such that $\lim_{n\to\infty} a_n = x_0$, we have $\lim_{n\to\infty} f(a_n) = f(x_0)$.*
- *For every $\varepsilon > 0$, there exists a $\delta = \delta(\varepsilon) > 0$ such that, for all $x \in X$ with $|x - x_0| < \delta$, we have $|f(x) - f(x_0)| < \varepsilon$.*

Exercise 3.2.6. Prove Proposition 3.2.5

Proposition 3.2.7. *Let X be a subset of $\mathbb{R}$, and let $f, g\colon X \to \mathbb{R}$ be functions. Let $x_0 \in X$. If f, g are both continuous at x_0, then $f + g$ and $f \cdot g$ are continuous at x_0. If g is nonzero on X, then f/g is continuous at x_0.*

Proof. Apply the Limit Laws (Proposition 3.1.20) and Definition 3.2.1. □

Remark 3.2.8. Let $x, c \in \mathbb{R}$. Note that the constant function $f(x) := c$ and the function $f(x) := x$ are continuous. Then, Proposition 3.2.7 implies that polynomials are continuous, and rational functions are continuous whenever the denominator is nonzero. For example, the function $(x^2 + 1)/(x - 1)$ is continuous on $\mathbb{R} \smallsetminus \{1\}$.

Proposition 3.2.9. *The function $f(x) := |x|$ is continuous on $\mathbb{R}$.*

Proof. Let $x_0 \in \mathbb{R}$. We split into three cases: $x_0 > 0$, $x_0 < 0$ and $x_0 = 0$. Suppose first that $x_0 > 0$. Define $\delta := |x_0|/2$. We show that f is continuous at x_0. From Proposition 3.1.25, it suffices to show that

$$x_0 = \lim_{x \to x_0; x \in (x_0 - \delta, x_0 + \delta)} f(x) = \lim_{x \to x_0; x \in (x_0/2, 3x_0/2)} f(x).$$

If $x \in (x_0/2, 3x_0/2)$, since $x_0 > 0$, we know that $x > 0$. So, $f(x) = x$. Therefore,

$$\lim_{x \to x_0; x \in (x_0/2, 3x_0/2)} f(x) = \lim_{x \to x_0; x \in (x_0/2, 3x_0/2)} x = x_0,$$

as desired. The case $x_0 < 0$ is similar.

We now conclude with the case $x_0 = 0$. Let $(a_n)_{n=0}^{\infty}$ be a sequence of real numbers converging to zero. From Proposition 3.2.5, it suffices to show that $(f(a_n))_{n=0}^{\infty}$ converges to zero. That is, it suffices to show: if $(a_n)_{n=0}^{\infty}$ converges to zero, then $(|a_n|)_{n=0}^{\infty}$ converges to zero. This follows from Exercise 2.2.9. □

Proposition 3.2.10. *Let X, Y be subsets of $\mathbb{R}$. Let $f\colon X \to Y$ and let $g\colon Y \to \mathbb{R}$ be functions. Let $x_0 \in X$. If f is continuous at x_0, and if g is continuous at $f(x_0)$, then $g \circ f$ is continuous at x_0.*

Exercise 3.2.11. Prove Proposition 3.2.10.

3.2.1. *Left and Right Limits.*

Definition 3.2.12. Let X be a subset of $\mathbb{R}$, let $f\colon X \to \mathbb{R}$ be a function, and let x_0 be a real number. If x_0 is an adherent point of $X \cap (x_0, \infty)$, then we define the **right limit** $f(x_0^+)$ of f at x_0 by the formula

$$f(x_0^+) := \lim_{x \to x_0; x \in X \cap (x_0, \infty)} f(x).$$

If this limit does not exist, or if x_0 is not an adherent point of $X \cap (x_0, \infty)$, we leave this limit undefined. Similarly, if x_0 is an adherent point of $X \cap (-\infty, x_0)$, then we define the **left limit** $f(x_0^-)$ of f at x_0 by the formula

$$f(x_0^-) := \lim_{x \to x_0; x \in X \cap (-\infty, x_0)} f(x).$$

If this limit does not exist, or if x_0 is not an adherent point of $X \cap (x_0, \infty)$, we leave this limit undefined.

Remark 3.2.13. Sometimes, we write $\lim_{x \to x_0^+} f(x)$ instead of $\lim_{x \to x_0; x \in X \cap (x_0, \infty)} f(x)$, and sometimes, we write $\lim_{x \to x_0^-} f(x)$ instead of $\lim_{x \to x_0; x \in X \cap (-\infty, x_0)} f(x)$.

The following proposition shows that, if both the left and right limits of a function exist at a point x_0, and if these limits are equal to $f(x_0)$, then f is continuous at x_0.

Proposition 3.2.14. *Let X be a subset of $\mathbb{R}$ containing a real number x_0. Suppose x_0 is an adherent point of both $X \cap (x_0, \infty)$ and $X \cap (-\infty, x_0)$. Let $f\colon X \to \mathbb{R}$ be a function. If $f(x_0^+)$ and $f(x_0^-)$ both exist, and we have $f(x_0^+) = f(x_0^-) = f(x_0)$, then f is continuous at x_0.*

3.2.2. *The Maximum Principle.* We can now begin to prove some of properties of continuous functions. The Maximum Principle says that a continuous function on a closed interval $[a, b]$ achieves its maximum and minimum values on $[a, b]$.

Definition 3.2.15. Let X be a subset of $\mathbb{R}$, and let $f\colon X \to \mathbb{R}$ be a function. We say that f is **bounded from above** if and only if there exists a real number M such that $f(x) \leq M$ for all $x \in X$. We say that f is **bounded from below** if and only if there exists a real number M such that $f(x) \geq M$ for all $x \in X$. We say that f is **bounded** if and only if there exists a real number M such that $|f(x)| \leq M$ for all $x \in X$.

Remark 3.2.16. A function is bounded if and only if it is bounded from above and from below.

Remark 3.2.17. Some continuous functions are not bounded. For example, the function $f\colon \mathbb{R} \to \mathbb{R}$ defined by $f(x) := x$ is unbounded on $\mathbb{R}$. Also, the function $f(x) := 1/x$ is unbounded on $(0, 1)$.

However, if f is continuous on a closed interval, then it is automatically bounded, as we now show, using the Bolzano-Weierstrass Theorem in an indirect manner.

Lemma 3.2.18. *Let $a < b$ be real numbers. Let $f\colon [a, b] \to \mathbb{R}$ be a continuous function. Then f is bounded.*

Proof. We argue by contradiction. Assume f is not bounded. Then, for every natural number n, there exists a point $x_n \in [a, b]$ such that $|f(x_n)| > n$. Since the sequence $(x_n)_{n=0}^{\infty}$ is contained in the closed interval $[a, b]$, the Bolzano-Weierstrass Theorem (Theorem 2.8.9) shows that there exists a subsequence $(x_{n_j})_{j=0}^{\infty}$ of $(x_n)_{n=0}^{\infty}$ such that $(x_{n_j})_{j=0}^{\infty}$ converges to some real number y as $j \to \infty$. Note that $n_j \geq j$ by the definition of a subsequence. Since $(x_{n_j})_{j=0}^{\infty}$ is a convergent sequence contained in $[a, b]$, we know that y is an adherent point of $[a, b]$. From Proposition 3.1.11, we conclude that y is also in $[a, b]$, so that y is in the domain of f. Now, since f is continuous on $[a, b]$, it is continuous at y so

$$\lim_{j \to \infty} f(x_{n_j}) = f(y). \qquad (*)$$

Since $n_j \geq j$, the definition of the sequence $(x_n)_{n=0}^{\infty}$ shows that $\left|f(x_{n_j})\right| \geq n_j \geq j$. That is, for all natural numbers $j > 1 + |f(y)|$, we have $\left|f(x_{n_j})\right| \geq j > 1 + |f(y)|$. So, $\lim_{j \to \infty} f(x_{n_j}) \neq f(y)$, contradicting $(*)$. Since we have achieved a contradiction, the proof is concluded. □

Definition 3.2.19. Let $f\colon X \to \mathbb{R}$ be a function, and let $x_0 \in X$. We say that f **attains its maximum** at x_0 if and only if $f(x_0) \geq f(x)$ for all $x \in X$. We say that f **attains its minimum** at x_0 if and only if $f(x_0) \leq f(x)$ for all $x \in X$.

We can now modify the proof of Lemma 3.2.18 a bit to give a stronger statement.

Theorem 3.2.20 (The Maximum Principle). *Let $a < b$ be real numbers. Let $f\colon [a,b] \to \mathbb{R}$ be a function that is continuous on $[a,b]$. Then f attains its maximum and minimum on $[a,b]$.*

Proof. We will show that f attains its maximum on $[a,b]$. Such a result applied to $-f$ then implies that f also attains its minimum on $[a,b]$.

From Lemma 3.2.18, there exists a real number M such that $-M \leq f(x) \leq M$ for all $x \in [a,b]$. Define

$$E := f([a,b]) = \{f(x)\colon x \in [a,b]\}.$$

Note that E is a nonempty subset of $\mathbb{R}$ that is bounded from above (and below). From the Least Upper Bound property (Theorem 1.7.6), E has a least upper bound $S := \sup(E)$.

For each positive integer n, the real number $S - 1/n$ is not an upper bound for E, since S is the least upper bound of E. So, there exists some $x_n \in [a,b]$ such that $f(x_n) \geq S - 1/n$. We are now once again in a position to apply the Bolzano-Weierstrass Theorem. Since the sequence $(x_n)_{n=1}^{\infty}$ is contained in the closed interval $[a,b]$, the Bolzano-Weierstrass Theorem (Theorem 2.8.9) shows that there exists a subsequence $(x_{n_j})_{j=1}^{\infty}$ of $(x_n)_{n=1}^{\infty}$ such that $(x_{n_j})_{j=1}^{\infty}$ converges to some real number y as $j \to \infty$. Note that $n_j \geq j$ by the definition of a subsequence, so $-1/n_j \geq -1/j$. Since $(x_{n_j})_{j=1}^{\infty}$ is a convergent sequence contained in $[a,b]$, we know that y is an adherent point of $[a,b]$. From Proposition 3.1.11, we conclude that y is also in $[a,b]$, so that y is in the domain of f. Now, since f is continuous on $[a,b]$, it is continuous at y so

$$\lim_{j\to\infty} f(x_{n_j}) = f(y). \qquad (*)$$

Since $n_j \geq j$, the definition of the sequence $(x_n)_{n=1}^{\infty}$ shows that

$$f(x_{n_j}) \geq S - 1/n_j \geq S - 1/j.$$

Also, since S is the supremum of f, we have $f(x_{n_j}) \leq S$. So, letting $j \to \infty$ and using the Squeeze Theorem (Corollary 2.5.12), we conclude that $S = \lim_{j\to\infty} f(x_{n_j}) = f(y)$, as desired. □

Remark 3.2.21. For a function $f\colon [a,b] \to \mathbb{R}$, we write $\sup_{x\in[a,b]} f(x)$ as shorthand for $\sup\{f(x)\colon x \in [a,b]\}$, and we write $\inf_{x\in[a,b]} f(x)$ as shorthand for $\inf\{f(x)\colon x \in [a,b]\}$

Remark 3.2.22. The assumptions of Theorem 3.2.20 cannot be weakened in general. For example, consider the function $f(x) := x$ on the open interval $(0,1)$. Then $\sup_{x\in(0,1)} f(x) = 1$ and $\inf_{x\in(0,1)} f(x) = 0$, but f does not take the value 1 or 0 on the open interval $(0,1)$, even though f is continuous.

Also, consider the function $f\colon [-1,1] \to \mathbb{R}$ defined by

$$f(x) := \begin{cases} x+1 & \text{, if } x \in [-1,0) \\ 0 & \text{, if } x = 0 \\ x-1 & \text{, if } x \in (0,1] \end{cases}.$$

Then $\sup_{x\in[-1,1]} f(x) = 1$ and $\inf_{x\in[-1,1]} f(x) = -1$, but f does not take the value 1 or -1 on the closed interval $[-1,1]$. Note that f is discontinuous at $x = 0$, so Theorem 3.2.20 does not apply.

3.2.3. *The Intermediate Value Theorem.* From Theorem 3.2.20, we know that a continuous function $f\colon [a,b] \to \mathbb{R}$ attains its minimum and maximum on $[a,b]$. We now show that f also attains all values in between the maximum and minimum.

Theorem 3.2.23 (**Intermediate Value Theorem**). *Let $a < b$ be real numbers. Let $f\colon [a,b] \to \mathbb{R}$ be function that is continuous on $[a,b]$. Let y be a real number between $f(a)$ and $f(b)$, so that either $f(a) \leq y \leq f(b)$ or $f(a) \geq y \geq f(b)$. Then there exists a $c \in [a,b]$ such that $f(c) = y$.*

Proof. Without loss of generality, assume that $f(a) \leq y \leq f(b)$. If $y = f(a)$ or $y = f(b)$, we just set $c = a$ or $c = b$ as needed. We therefore assume that $f(a) < y < f(b)$. Define

$$E := \{x \in [a,b] \colon f(x) < y\}.$$

Since $f(a) < y$, E is nonempty. Since E is contained in $[a,b]$, E is bounded from above. By the Least Upper Bound property (Theorem 1.7.6), E has a least upper bound $c := \sup(E)$. We will prove that $f(c) = y$.

Since b is an upper bound for E, we know that $c \leq b$. Since $a \in E$, we know that $a \leq c$. So, $c \in [a,b]$. By looking to the left of c, we will show that $f(c) \leq y$, and then by looking to the right of c, we will show that $f(c) \geq y$.

We now show that $f(c) \leq y$. Let n be a positive integer. Then $c - 1/n < c = \sup(E)$, so $c - 1/n$ is not an upper bound for E. So, there exists a point $x_n \in E$ such that $x_n > c - 1/n$. Since c is an upper bound for E, $x_n \leq c$. So

$$c - 1/n \leq x_n \leq c.$$

Letting $n \to \infty$, we conclude by the Squeeze Theorem (Corollary 2.5.12) that $\lim_{n\to\infty} x_n = c$. Since f is continuous at c, we have $\lim_{n\to\infty} f(x_n) = f(c)$. Since $x_n \in E$ for every positive integer n, we have $f(x_n) < y$ for every positive integer n. By the Comparison Principle (Lemma 2.5.10), we conclude that

$$f(c) = \lim_{n\to\infty} f(x_n) \leq y.$$

We now show that $f(c) \geq y$. Since $f(c) \leq y < f(b)$, we have $c \neq b$. Since $c \in [a,b]$, we then have $c < b$. So, there exists a positive integer m such that, for all $n \geq m$, $c + 1/n < b$. Then $c + 1/n > c$. Since $c = \sup(E)$, we conclude that $c + 1/n \notin E$. Also, $c + 1/n \in [a,b]$. So, by the definition of E, we have $f(c + 1/n) \geq y$. Since f is continuous at c, we have $\lim_{n\to\infty} f(c + 1/n) = f(c)$. By the Comparison Principle (Lemma 2.5.10), we conclude that

$$f(c) = \lim_{n\to\infty} f(c + 1/n) \geq y.$$

Finally, $y \leq f(c) \leq y$, so $f(c) = y$, as desired. □

Remark 3.2.24. The assumption that f is continuous is necessary for Theorem 3.2.23. For example, consider the function

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

Remark 3.2.25. Theorem 3.2.23 gives another way to prove the existence of n^{th} roots. For example, for $x \in \mathbb{R}$, define $f(x) := x^2$, $f\colon [0,2] \to \mathbb{R}$. Then $f(0) = 0$, $f(2) = 4$, so choosing $y = 2$, there exists at least one $c \in [0,2]$ such that $f(c) = c^2 = 2$.

Corollary 3.2.26. *Let $a < b$ be real numbers. Let $f: [a, b] \to \mathbb{R}$ be a continuous function on $[a, b]$. Let $M := \sup_{x\in[a,b]} f(x)$ be the maximum value of f on $[a, b]$, and let $m := \inf_{x\in[a,b]} f(x)$ be the minimum value of f on $[a, b]$. Let y be a real number such that $m \leq y \leq M$. Then there exists $c \in [a, b]$ such that $f(c) = y$. Moreover, $f([a, b]) = [m, M]$.*

Exercise 3.2.27. Prove Corollary 3.2.26.

3.2.4. *Monotone Functions.*

Definition 3.2.28. Let X be a subset of $\mathbb{R}$ and let $f: X \to \mathbb{R}$ be a function. We say that f is **monotone increasing** if and only if $f(y) \geq f(x)$ for all $x, y \in X$ with $y > x$. We say that f is **strictly monotone increasing** if and only if $f(y) > f(x)$ for all $x, y \in X$ with $y > x$. Similarly, we say that f is **monotone decreasing** if and only if $f(y) \leq f(x)$ for all $x, y \in X$ with $y > x$. We say that f is **strictly monotone decreasing** if and only if $f(y) < f(x)$ for all $x, y \in X$ with $y > x$. We say that f is **monotone** if and only if it is either monotone increasing or monotone decreasing. We say that f is **strictly monotone** if and only if it is either strictly monotone increasing or strictly monotone decreasing.

A strictly monotone and continuous function has a continuous inverse, as we now show.

Proposition 3.2.29. *Let $a < b$ be real numbers, and let $f: [a, b] \to \mathbb{R}$ be a function which is both continuous and strictly monotone increasing. Then f is a bijection from $[a, b]$ to $[f(a), f(b)]$, and the inverse function $f^{-1}: [f(a), f(b)] \to [a, b]$ is also continuous and strictly monotone increasing.*

Exercise 3.2.30. Prove Proposition 3.2.29. (Hint: To prove that f^{-1} is continuous, use the ε–δ definition of continuity.)

3.2.5. *Uniform Continuity.* There is a bit of an odd point in the definition of continuity. A function $f: \mathbb{R} \to \mathbb{R}$ is continuous if and only if it is continuous at every $x \in \mathbb{R}$. That is, given any $x_0 \in \mathbb{R}$ and any $\varepsilon > 0$, there exists a $\delta = \delta(x_0, \varepsilon)$ such that, if $|x - x_0| < \delta$, then $|f(x) - f(x_0)| < \varepsilon$. Note in particular that δ may depend on x_0. For example, the function $f: (0, \infty) \to \mathbb{R}$ defined by $f(x) := 1/x$ is continuous on $(0, \infty)$, but f is not bounded. The problem here is that, if $\varepsilon > 0$ is fixed, then $\delta(x_0, \varepsilon)$ must be chosen to be smaller and smaller as $x_0 \to 0^+$. It would be nicer if we could select δ in a way that does not depend on x_0, as in the following definition.

Definition 3.2.31 (Uniform Continuity). Let X be a subset of $\mathbb{R}$, and let $f: X \to \mathbb{R}$ be a function. We say that f is **uniformly continuous** if and only if, for every $\varepsilon > 0$ there exists $\delta > 0$ such that, if $x, x_0 \in X$ satisfy $|x - x_0| < \delta$, then $|f(x) - f(x_0)| < \varepsilon$.

Remark 3.2.32. A uniformly continuous function is continuous.

Example 3.2.33. The function $f: \mathbb{R} \to \mathbb{R}$ defined by $f(x) := x$ is uniformly continuous. On the other hand, the function $f: (0, \infty) \to \mathbb{R}$ defined by $f(x) := 1/x$ is not uniformly continuous.

Just as in the case of continuity, there is a way to characterize uniform continuity using sequences. We now explore this characterization.

Definition 3.2.34. Let $(a_n)_{n=m}^{\infty}, (b_n)_{n=m}^{\infty}$ be two sequences of real numbers. We say that $(a_n)_{n=m}^{\infty}$ and $(b_n)_{n=m}^{\infty}$ are **equivalent** if and only if for every real $\varepsilon > 0$, there exists an integer $N = N(\varepsilon) > m$ such that, for all $n \geq N$, we have $|a_n - b_n| < \varepsilon$.

Lemma 3.2.35. *Let $(a_n)_{n=m}^{\infty}, (b_n)_{n=m}^{\infty}$ be two sequences of real numbers. Then $(a_n)_{n=m}^{\infty}$ and $(b_n)_{n=m}^{\infty}$ are equivalent if and only if $\lim_{n\to\infty}(a_n - b_n) = 0$.*

Exercise 3.2.36. Prove Lemma 3.2.35.

Note that equivalent sequences need not converge.

Proposition 3.2.37. *Let X be a subset of $\mathbb{R}$ and let $f\colon X \to \mathbb{R}$ be a function. Then the following two statements are equivalent.*

- *f is uniformly continuous on X.*
- *For any two equivalent sequences $(a_n)_{n=m}^{\infty}, (b_n)_{n=m}^{\infty}$, the sequences $(f(a_n))_{n=m}^{\infty}, (f(b_n))_{n=m}^{\infty}$ are also equivalent sequences.*

Exercise 3.2.38. Prove Proposition 3.2.37.

Remark 3.2.39. From Proposition 3.2.5, we saw that continuous functions map convergent sequences to convergent sequences. Proposition 3.2.37 then says that uniformly continuous functions map equivalent sequences to equivalent sequences.

Corollary 3.2.40. *Let X be a subset of $\mathbb{R}$ and let $f\colon X \to \mathbb{R}$ be a uniformly continuous function. Let x_0 be an adherent point of X. Then $\lim_{x\to x_0} f(x)$ exists (and so it is a real number.)*

Exercise 3.2.41. Prove Corollary 3.2.40

Remark 3.2.42. Note that Corollary 3.2.40 is false in general, if f is just continuous. For example, consider again $f(x) := 1/x$, where $f\colon (0,\infty) \to \mathbb{R}$. Then $\lim_{x\to 0^+} f(x)$ does not exist. But also recall that f is not uniformly continuous.

Uniformly continuous functions also map bounded sets to bounded sets.

Proposition 3.2.43. *Let X be a subset of $\mathbb{R}$, and let $f\colon X \to \mathbb{R}$ be a uniformly continuous function. Assume that E is a bounded subset of X. Then $f(E)$ is also bounded.*

Exercise 3.2.44. Prove Proposition 3.2.43.

Since uniformly continuous functions have such nice properties, it is helpful to have some conditions to easily verify uniform continuity, as in the following Theorem.

Theorem 3.2.45. *Let $a < b$ be real numbers, and let $f\colon [a, b] \to \mathbb{R}$ be a function which is continuous on $[a, b]$. Then f is also uniformly continuous on $[a, b]$.*

Proof. We argue by contradiction. Suppose f is not uniformly continuous on $[a, b]$. So, using Proposition 3.2.37, there exist two equivalent sequences $(a_n)_{n=m}^{\infty}, (b_n)_{n=m}^{\infty}$ contained in $[a, b]$ such that $(f(a_n))_{n=m}^{\infty}, (f(b_n))_{n=m}^{\infty}$ are not equivalent. That is, there exists an $\varepsilon > 0$ such that, for all integers $N > m$, there exists an integer $n \geq N$ such that

$$|f(a_n) - f(b_n)| \geq \varepsilon. \qquad (*)$$

In particular, the following set is infinite

$$A := \{n \in \mathbb{N}\colon\ |f(a_n) - f(b_n)| \geq \varepsilon\}.$$

That is, given any set of natural numbers $n_0 < n_1 < \cdots < n_j$ in A, there exists an integer $n_{j+1} > n_j$ so that $\left|f(a_{n_j}) - f(b_{n_j})\right| \geq \varepsilon$. So, consider the sequences $(a_{n_j})_{j=0}^{\infty}, (b_{n_j})_{j=0}^{\infty}$ which are equivalent and contained in $[a,b]$. By the Bolzano-Weierstrass Theorem, there exists a subsequence $(a_{n_{j_k}})_{k=0}^{\infty}$ of $(a_{n_j})_{j=0}^{\infty}$ such that $(a_{n_{j_k}})_{k=0}^{\infty}$ converges as $k \to \infty$. From Lemma 3.2.35, since $(a_{n_{j_k}})_{k=0}^{\infty}$ and $(b_{n_{j_k}})_{k=0}^{\infty}$ are equivalent sequences, we conclude that $(b_{n_{j_k}})_{k=0}^{\infty}$ converges as $k \to \infty$ as well. Using Lemma 3.2.35 again, $(a_{n_{j_k}})_{k=0}^{\infty}$ and $(b_{n_{j_k}})_{k=0}^{\infty}$ converge to the same point $c \in [a,b]$. So, using the Limit Laws (Proposition 3.1.20),

$$\lim_{k\to\infty}(f(a_{n_{j_k}}) - f(b_{n_{j_k}})) = 0$$

Since this violates $(*)$, we have achieved a contradiction, concluding the proof. □

3.2.6. *Limits at Infinity.*

Definition 3.2.46. Let X be a subset of $\mathbb{R}$. We say that $+\infty$ is an adherent point of X if and only if for every $M \in \mathbb{R}$ there exists an $x \in X$ such that $x > M$. We say that $-\infty$ is an adherent point of X if and only if for every $M \in \mathbb{R}$ there exists an $x \in X$ such that $x < M$.

Definition 3.2.47. Let X be a subset of $\mathbb{R}$ such that $+\infty$ is an adherent point of X. Let $f\colon X \to \mathbb{R}$ be a function and let L be a real number. We say that $f(x)$ **converges to** L as $x \to +\infty$ if and only if, for every $\varepsilon > 0$, there exists a real M such that, for all $x \in X$ with $x > M$, we have $|f(x) - L| < \varepsilon$. Similarly, if $-\infty$ is an adherent point of X, then we say that $f(x)$ **converges to** L as $x \to -\infty$ if and only if, for every $\varepsilon > 0$, there exists a real M such that, for all $x \in X$ with $x < M$, we have $|f(x) - L| < \varepsilon$.

Example 3.2.48. Let $f\colon (0,\infty) \to \mathbb{R}$ be defined by $f(x) := 1/x$. Then $\lim_{x\to+\infty} f(x) = 0$.

3.3. **Derivatives.** We will soon define a derivative, but before doing so, we adjust slightly the definition of adherent point.

Definition 3.3.1. Let X be a subset of $\mathbb{R}$ and let x be a real number. We say that x is a **limit point** of X (or x is a **cluster point** of X) if and only if x is an adherent point of $X \smallsetminus \{x\}$.

Remark 3.3.2. That is, x is a limit point of X if and only if, for every real $\varepsilon > 0$, there exists a $y \in X$ with $y \neq x$ such that $|y - x| < \varepsilon$.

Lemma 3.1.13 then implies the following.

Lemma 3.3.3. *Let X be a subset of $\mathbb{R}$, and let x be a real number. Then x is a limit point of X if and only if there exists a sequence $(a_n)_{n=m}^{\infty}$ of elements of $X \smallsetminus \{x\}$ such that $(a_n)_{n=m}^{\infty}$ converges to x.*

Lemma 3.3.4. *Let I be a (possibly infinite) interval. That is, I is equal to a set of the form (a,b), $[a,b]$, $(a,b]$, $[a,b)$, $(a,+\infty)$, $[a,+\infty)$, $(-\infty,b)$, $(-\infty,b]$ or $(-\infty,\infty)$ where $a,b \in \mathbb{R}$ and $a < b$. Then every element of I is a limit point of I.*

Proof. We only prove the case $I = [a,b]$ and leave the rest as exercises.

Suppose $x \in [a,b)$. Then there exists a positive integer N such that, for all $n \geq N$, $x + 1/n < b$. So, the sequence $(x + 1/n)_{n=N}^{\infty}$ is contained in $I \smallsetminus \{x\}$, and this sequence converges to x. Therefore, x is a limit point of $[a,b]$, by Lemma 3.3.3. To deal with the remaining case of $x = b$, we do the same thing but we use the sequence $(x - 1/n)_{n=N}^{\infty}$. □

We can now define derivatives.

Definition 3.3.5. Let X be a subset of $\mathbb{R}$, and let x_0 be an element of X which is also a limit point of X. Let $f\colon X \to \mathbb{R}$ be a function. If the limit

$$\lim_{x\to x_0; x\in X\setminus\{x_0\}} \frac{f(x)-f(x_0)}{x-x_0}.$$

converges to a real number L, then we say that f is **differentiable** at x_0 on X **with derivative** L, and we write $f'(x_0) := L$. If this limit does not exist, or if x_0 is not a limit point of X, we leave $f'(x_0)$ undefined, and we say that f is **not differentiable** at x_0 on X.

Remark 3.3.6. Note that we need x_0 to be a limit point of $X \setminus \{x_0\}$, otherwise the limit in the definition of the derivative would be undefined. Often, the set X will be an interval as in Lemma 3.3.4, so this issue will not arise.

Example 3.3.7. Let $f\colon \mathbb{R} \to \mathbb{R}$ be defined by $f(x) := x$. Then

$$f'(x_0) = \lim_{x\to x_0; x\in\mathbb{R}\setminus\{x_0\}} \frac{x-x_0}{x-x_0} = 1.$$

Let $f\colon \mathbb{R} \to \mathbb{R}$ be defined by $f(x) := x^2$. Then

$$\begin{aligned} f'(x_0) &= \lim_{x\to x_0; x\in\mathbb{R}\setminus\{x_0\}} \frac{x^2-x_0^2}{x-x_0} = \lim_{x\to x_0; x\in\mathbb{R}\setminus\{x_0\}} \frac{(x+x_0)(x-x_0)}{x-x_0} \\ &= \lim_{x\to x_0; x\in\mathbb{R}\setminus\{x_0\}} (x+x_0) = 2x_0. \end{aligned}$$

In general, if k is a positive integer, and if $f(x) := x^k$, $f\colon \mathbb{R} \to \mathbb{R}$, then

$$\begin{aligned} f'(x_0) &= \lim_{x\to x_0; x\in\mathbb{R}\setminus\{x_0\}} \frac{x^k-x_0^k}{x-x_0} = \lim_{x\to x_0; x\in\mathbb{R}\setminus\{x_0\}} \frac{(\sum_{j=1}^k x^{k-j}x_0^{j-1})(x-x_0)}{x-x_0} \\ &= \lim_{x\to x_0; x\in\mathbb{R}\setminus\{x_0\}} \sum_{j=1}^k x^{k-j}x_0^{j-1} = \sum_{j=1}^k x_0^{k-1} = kx_0^{k-1}. \end{aligned}$$

Remark 3.3.8. Sometimes one writes $f'(x)$ as df/dx, but we will not do so here.

We now give an example of a continuous function that is not differentiable at zero.

Example 3.3.9. Define $f(x) := |x|$. For $x_0 \in (-\infty, 0) \cup (0, \infty)$, one can show that f is differentiable. However, f is not differentiable at 0. To see this, observe that

$$\lim_{x\to 0; x\in(0,\infty)} \frac{f(x)-f(0)}{x-0} = \lim_{x\to 0; x\in(0,\infty)} \frac{x-f(0)}{x-0} = 1.$$

$$\lim_{x\to 0; x\in(-\infty,0)} \frac{f(x)-f(0)}{x-0} = \lim_{x\to 0; x\in(-\infty,0)} \frac{-x-f(0)}{x-0} = -1.$$

Therefore, $\lim_{x\to 0; x\in\mathbb{R}\setminus\{0\}} \frac{f(x)-f(0)}{x-0}$ does not exist. So, f is not differentiable at 0.

Even though a function may be continuous but not differentiable at a point, a function that is differentiable at a point is always continuous at that point.

Proposition 3.3.10. *Let X be a subset of $\mathbb{R}$, let x_0 be a limit point of X, and let $f\colon X \to \mathbb{R}$ be a function. If f is differentiable at x_0, then f is also continuous at x_0.*

Exercise 3.3.11. Prove Proposition 3.3.10

If a function is differentiable at x_0, then it is approximately linear at x_0 in the following sense.

Proposition 3.3.12. *Let X be a subset of $\mathbb{R}$, let x_0 be a limit point of X, let $f\colon X \to \mathbb{R}$ be a function, and let L be a real number. Then the following two statements are equivalent.*

- *f is differentiable at x_0 on X with derivative L.*
- *For every $\varepsilon > 0$, there exists a $\delta = \delta(\varepsilon) > 0$ such that, if $x \in X$ satisfies $|x - x_0| < \delta$, then*
$$|f(x) - [f(x_0) + L(x - x_0)]| \leq \varepsilon |x - x_0| .$$

Exercise 3.3.13. Prove Proposition 3.3.12.

Remark 3.3.14. The second item is understood informally as $f(x) \approx f(x_0) + f'(x_0)(x - x_0)$.

Definition 3.3.15. Let X be a subset of $\mathbb{R}$ and let $f\colon X \to \mathbb{R}$ be a function. We say that f is **differentiable** on X if and only if f is differentiable at x_0 for all $x_0 \in X$.

Using this definition and Proposition 3.3.10, we get the following.

Corollary 3.3.16. *Let X be a subset of $\mathbb{R}$ and let $f\colon X \to \mathbb{R}$ be a function that is differentiable on X. Then f is continuous on X.*

Theorem 3.3.17 (Properties of Derivatives). *Let X be a subset of $\mathbb{R}$, let x_0 be a limit point of X, and let $f\colon X \to \mathbb{R}$ and $g\colon X \to \mathbb{R}$ be functions.*

(i) *If f is constant, so that there exists $c \in \mathbb{R}$ such that $f(x) = c$ for all $x \in X$, then f is differentiable at x_0 and $f'(x_0) = 0$.*
(ii) *If f is the identity function, so that $f(x) = x$ for al $x \in X$, then f is differentiable at x_0 and $f'(x_0) = 1$.*
(iii) *If f, g are differentiable at x_0, then $f + g$ is differentiable at x_0, and $(f + g)'(x_0) = f'(x_0) + g'(x_0)$. (**Sum Rule**)*
(iv) *If f, g are differentiable at x_0, then fg is differentiable at x_0, and $(fg)'(x_0) = f'(x_0)g(x_0) + g'(x_0)f(x_0)$. (**Product Rule**)*
(v) *If f is differentiable at x_0, and if $c \in \mathbb{R}$, then cf is differentiable at x_0, and $(cf)'(x_0) = cf'(x_0)$.*
(vi) *If f, g are differentiable at x_0, then $f - g$ is differentiable at x_0, and $(f - g)'(x_0) = f'(x_0) - g'(x_0)$.*
(vii) *If g is differentiable at x_0, and if $g(x) \neq 0$ for all $x \in X$, then $1/g$ is differentiable at x_0, and $(1/g)'(x_0) = -\frac{g'(x_0)}{(g(x_0))^2}$.*
(viii) *If f, g are differentiable at x_0, and if $g(x) \neq 0$ for all $x \in X$, then f/g is differentiable at x_0, and*
$$(f/g)'(x_0) = \frac{g(x_0)f'(x_0) - f(x_0)g'(x_0)}{(g(x_0))^2}. \qquad (\textbf{\textit{Quotient Rule}})$$

Exercise 3.3.18. Prove Theorem 3.3.17. For the product rule, you may need the following identity
$$f(x)g(x) - f(x_0)g(x_0) = f(x)(g(x) - g(x_0)) + g(x_0)(f(x) - f(x_0)).$$

Theorem 3.3.19 (Chain Rule). *Let X, Y be subsets of $\mathbb{R}$, let $x_0 \in X$ be a limit point of X, and let $y_0 \in Y$ be a limit point of Y. Let $f\colon X \to Y$ be a function such that $f(x_0) = y_0$*

and such that f is differentiable at x_0. Let $g\colon Y \to \mathbb{R}$ be a function that is differentiable at y_0. Then the function $g \circ f\colon X \to \mathbb{R}$ is differentiable at x_0, and

$$(g \circ f)'(x_0) = g'(y_0) f'(x_0).$$

Exercise 3.3.20. Prove Theorem 3.3.19. (Hint: using Proposition 3.1.16, it suffices to consider a sequence $(a_n)_{n=0}^{\infty}$ of elements of X converging to x_0. Also, from Proposition 3.3.10, f is continuous, so $(f(a_n))_{n=0}^{\infty}$ converges to $f(x_0)$.)

3.3.1. *Local Extrema.*

Definition 3.3.21. Let $f\colon X \to \mathbb{R}$ be a function, and let $x_0 \in X$. We say that f **attains a local maximum** at x_0 if and only if there exists a $\delta > 0$ such that the restriction $f|_{X \cap (x_0-\delta, x_0+\delta)}$ attains a maximum at x_0. We say that f **attains a local minimum** at x_0 if and only if there exists a $\delta > 0$ such that the restriction $f|_{X \cap (x_0-\delta, x_0+\delta)}$ attains a minimum at x_0.

Remark 3.3.22. If $f\colon X \to \mathbb{R}$ attains a maximum at x_0, then we sometimes say that f attains a **global maximum** at x_0.

Proposition 3.3.23. *Let $a < b$ be real numbers, and let $f\colon (a,b) \to \mathbb{R}$ be a function. If $x_0 \in (a,b)$, if f is differentiable at x_0, and if f attains a local maximum or minimum at x_0, then $f'(x_0) = 0$.*

Exercise 3.3.24. Prove Proposition 3.3.23.

Remark 3.3.25. Note that Proposition 3.3.23 is not true if f we assume that $f\colon [a,b] \to \mathbb{R}$ achieves a local maximum or minimum. For example, the function $f\colon [0,1] \to \mathbb{R}$ defined by $f(x) := x$ satisfies $f'(x) = 1$ for all $x \in [0,1]$, while f achieves a local maximum at $x = 1$ and a local minimum at $x = 0$.

Theorem 3.3.26 (Rolle's Theorem). *Let $a < b$ be real numbers, and let $f\colon [a,b] \to \mathbb{R}$ be a continuous function which is differentiable on (a,b). Assume that $f(a) = f(b)$. Then there exists $x \in (a,b)$ such that $f'(x) = 0$.*

Exercise 3.3.27. Prove Theorem 3.3.26. (Hint: use Proposition 3.3.23 and the Maximum Principle, Theorem 3.2.20.)

Theorem 3.3.26 then has the following useful corollary.

Corollary 3.3.28 (Mean Value Theorem). *Let $a < b$ be real numbers, and let $f\colon [a,b] \to \mathbb{R}$ be a continuous function which is differentiable on (a,b). Then there exists $x \in (a,b)$ such that*

$$f'(x) = \frac{f(b) - f(a)}{b - a}.$$

Proof. Consider the function $g\colon [a,b] \to \mathbb{R}$ defined by

$$g(y) := f(y) - \frac{f(b) - f(a)}{b - a}(y - a). \qquad (*)$$

Note that $g(a) = f(a) = g(b)$, g is continuous on $[a,b]$ by Proposition 3.2.7, and g is differentiable on (a,b) by Theorem 3.3.17(v) and (iii). So by Theorem 3.3.26, there exists $x \in (a,b)$ such that $g'(x) = 0$. Using $(*)$ and Theorem 3.3.17, $g'(x) = 0$ says that

$$0 = f'(x) - \frac{f(b) - f(a)}{b - a}.$$

□

3.3.2. *Monotone Functions and Derivatives.* We now explore the connection between the monotonicity of a function and the sign of its derivative.

Proposition 3.3.29. *Let X be a subset of $\mathbb{R}$, let x_0 be a limit point of X, and let $f\colon X \to \mathbb{R}$ be a function. If f is monotone increasing and if f is differentiable at x_0, then $f'(x_0) \geq 0$. If f is monotone decreasing and if f is differentiable at x_0, then $f'(x_0) \leq 0$.*

Exercise 3.3.30. Prove Proposition 3.3.29.

Remark 3.3.31. Note that we need to assume that f is both monotone and differentiable, since there exist functions that are monotone but not differentiable. Consider for example $f\colon \mathbb{R} \to \mathbb{R}$ defined by

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

A strictly monotone increasing function can have a zero derivative. Consider for example $f\colon \mathbb{R} \to \mathbb{R}$ defined by $f(x) := x^3$, and note that $f'(0) = 0$. However, a converse statement is true, as we now show.

Proposition 3.3.32. *Let $a < b$ be real numbers, and let $f\colon [a,b] \to \mathbb{R}$ be a differentiable function. If $f'(x) > 0$ for all $x \in [a,b]$, then f is strictly monotone increasing. If $f'(x) < 0$ for all $x \in [a,b]$, then f is strictly monotone decreasing. If $f'(x) = 0$ for all $x \in [a,b]$, then f is a constant function.*

Exercise 3.3.33. Prove Proposition 3.3.32. (Hint: for the final statement, use the Mean-Value Theorem.)

3.3.3. *Inverse Functions and Derivatives.* Let X, Y be subsets of $\mathbb{R}$. If we have a bijective function $f\colon X \to Y$ which is differentiable, then the derivative of f^{-1} is related nicely to the derivative of f, as we now show.

Lemma 3.3.34. *Let X, Y be subsets of $\mathbb{R}$. Let $f\colon X \to Y$ be a bijection, so that $f^{-1}\colon Y \to X$ is a function. Let $x_0 \in X$ and $y_0 \in Y$ such that $f(x_0) = y_0$. (Consequently, $x_0 = f^{-1}(y_0)$.) If f is differentiable at x_0 and if f^{-1} is differentiable at y_0, then $f'(x_0) \neq 0$ and*

$$(f^{-1})'(y_0) = \frac{1}{f'(x_0)}.$$

Proof. Note that $(f^{-1} \circ f)(x) = x$ for all $x \in X$. So, from the Theorem 3.3.17(ii) and the Chain Rule (Theorem 3.3.19),

$$1 = (f^{-1} \circ f)'(x_0) = (f^{-1})'(y_0) f'(x_0).$$

Since $(f^{-1})'(y_0) f'(x_0) = 1$, we know that $f'(x_0) \neq 0$, and $(f^{-1})'(y_0) = 1/f'(x_0)$ □

Remark 3.3.35. As a consequence of Lemma 3.3.34, we see that if f is differentiable at x_0 with $f'(x_0) = 0$, then f^{-1} is not differentiable at $y_0 = f(x_0)$. For example, consider the function $f(x) := x^n$, where n is a positive integer and $f\colon [0,\infty) \to [0,\infty)$. Then $f^{-1}(x) = x^{1/n}$, $f^{-1}\colon [0,\infty) \to [0,\infty)$. And if $n \geq 2$, then $f'(0) = 0$, so f^{-1} is not differentiable at 0.

Lemma 3.3.34 is deficient, in that we need to assume that f^{-1} is differentiable at $f(x_0)$. It would be more preferable to know that f^{-1} is differentiable by only using information about f. Such a goal is accomplished in the following theorem.

Theorem 3.3.36 (Inverse Function Theorem). *Let X, Y be subsets of $\mathbb{R}$. Let $f\colon X \to Y$ be bijection, so that $f^{-1}\colon Y \to X$ is a function. Let $x_0 \in X$ and $y_0 \in Y$ such that $f(x_0) = y_0$. If f is differentiable at x_0, if f^{-1} is continuous at y_0, and if $f'(x_0) \neq 0$, then f^{-1} is differentiable at y_0 with*

$$(f^{-1})'(y_0) = \frac{1}{f'(x_0)}.$$

Proof. We are required to show that

$$\lim_{y \to y_0; y \in Y \setminus \{y_0\}} \frac{f^{-1}(y) - f^{-1}(y_0)}{y - y_0} = \frac{1}{f'(x_0)}.$$

By Proposition 3.1.16, given any sequence $(y_n)_{n=1}^\infty$ of elements in $Y \setminus \{y_0\}$ that converges to y_0, it suffices to show that

$$\lim_{n \to \infty} \frac{f^{-1}(y_n) - f^{-1}(y_0)}{y_n - y_0} = \frac{1}{f'(x_0)}. \qquad (*)$$

Note that f is a bijection, so there exists a sequence of elements $(x_n)_{n=1}^\infty$ such that $f(x_n) = y_n$ for all $n \geq 1$. Moreover, since $(y_n)_{n=1}^\infty$ is contained in $Y \setminus \{y_0\}$, since $f(x_0) = y_0$, and since f is a bijection, the sequence $(x_n)_{n=1}^\infty$ is contained in $X \setminus \{x_0\}$. Since $y_n \to y_0$ as $n \to \infty$, and since f^{-1} is continuous at y_0 by assumption, we have $f^{-1}(y_n) = x_n \to x_0 = f^{-1}(y_0)$ as $n \to \infty$. So, since f is differentiable at x_0, we have by Proposition 3.1.16 that

$$\lim_{n \to \infty} \frac{f(x_n) - f(x_0)}{x_n - x_0} = f'(x_0).$$

That is,

$$\lim_{n \to \infty} \frac{y_n - y_0}{f^{-1}(y_n) - f^{-1}(y_0)} = f'(x_0). \qquad (**)$$

Since $y_n \neq y_0$ for all $n \geq 1$, the numerator on the left of $(**)$ is nonzero. Also, by hypothesis, $f'(x_0) \neq 0$. So, we can invert both sides of $(**)$ and apply the limit laws (Theorem 2.2.15(v)) to conclude that $(*)$ holds, as desired. □

4. Riemann Sums, Riemann Integrals, Fundamental Theorem of Calculus

4.1. **Riemann Sums.** Within calculus, the two most fundamental concepts are differentiation and integration. We have covered differentiation already, and we now move on to integration. Defining an integral is fairly delicate. In the case of the derivative, we created one limit, and the existence of this limit dictated whether or not the function in question was differentiable. In the case of the Riemann integral, there is also a limit to discuss, but it is much more complicated than in the case of differentiation.

We should mention that there is more than one way to construct an integral, and the Riemann integral is only one such example. Within this course, we will only be discussing the Riemann integral. The Riemann integral has some deficiencies which are improved upon by other integration theories. However, those other integration theories are more involved, so we focus for now only on the Riemann integral.

Our starting point will be partitions of intervals into smaller intervals, which will form the backbone of the Riemann sum. The Riemann sum will then be used to create the Riemann integral through a limiting constructing.

Definition 4.1.1 (**Partition**)**.** Let $a < b$ be real numbers. A **partition** P of the interval $[a, b]$ is a finite subset of real numbers $x_0, \ldots, x_n$ such that

$$a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b.$$

We write $P = \{x_0, x_1, \ldots, x_n\}$.

Remark 4.1.2. Let P, P' be partitions of $[a, b]$. Then the union $P \cup P'$ of P and P' is also a partition of $[a, b]$.

Definition 4.1.3 (**Upper and Lower Riemann Sums**)**.** Let $a < b$ be real numbers, let $f\colon [a, b] \to \mathbb{R}$ be a bounded function, and let $P = \{x_0, \ldots, x_n\}$ be a partition of $[a, b]$. For every integer $1 \leq i \leq n$, the function $f|_{[x_{i-1},x_i]}$ is also a bounded function. So, $\sup_{x\in[x_{i-1},x_i]} f(x)$ and $\inf_{x\in[x_{i-1},x_i]} f(x)$ exist by the Least Upper Bound property (Theorem 1.7.6). We therefore define the **upper Riemann sum** $U(f, P)$ by

$$U(f, P) := \sum_{i=1}^{n} \Big(\sup_{x\in[x_{i-1},x_i]} f(x)\Big)(x_i - x_{i-1}).$$

We also define the **lower Riemann sum** $L(f, P)$ by

$$U(f, P) := \sum_{i=1}^{n} \Big(\inf_{x\in[x_{i-1},x_i]} f(x)\Big)(x_i - x_{i-1}).$$

Remark 4.1.4. For each integer $1 \leq i \leq n$, we define a function $g\colon [a, b] \to \mathbb{R}$ such that $g(x) := \sup_{y\in[x_{i-1},x_i]} f(y)$ for all $x_{i-1} \leq x < x_i$, with $g(b) := f(b)$. Then g is constant on $[x_{i-1}, x_i)$ for all $1 \leq i \leq n$, and $f(x) \leq g(x)$ for all $x \in [a, b]$. The upper Riemann sum $U(f, P)$ then represents the area under the function g, which is meant to upper bound the area under the function f. Similarly, for each integer $1 \leq i \leq n$, we define a function $h\colon [a, b] \to \mathbb{R}$ such that $h(x) := \inf_{y\in[x_{i-1},x_i]} f(y)$ for all $x_{i-1} \leq x < x_i$, with $h(b) := f(b)$. Then h is constant on $[x_{i-1}, x_i)$ for all $1 \leq i \leq n$, and $h(x) \leq g(x)$ for all $x \in [a, b]$. The lower Riemann sum $L(f, P)$ then represents the area under the function g, which is meant to lower bound the area under the function f.

Definition 4.1.5 (**Upper and Lower Integrals**)**.** Let $a < b$ be real numbers, let $f\colon [a, b] \to \mathbb{R}$ be a bounded function. We define the **upper Riemann integral** $\overline{\int_a^b} f$ of f on $[a, b]$ by

$$\overline{\int_a^b} f := \inf\{U(f, P)\colon P \text{ is a partition of } [a, b]\}.$$

We also define the **lower Riemann integral** $\underline{\int_a^b} f$ of f on $[a, b]$ by

$$\underline{\int_a^b} f := \sup\{L(f, P)\colon P \text{ is a partition of } [a, b]\}.$$

Lemma 4.1.6. *Let $f\colon [a,b]\to\mathbb{R}$ be a bounded function, so that there exists a real number M such that $-M\leq f(x)\leq M$ for all $x\in[a,b]$. Then*

$$-M(b-a)\leq \underline{\int_a^b} f\leq \overline{\int_a^b} f\leq M(b-a)$$

In particular, $\overline{\int_a^b}$ and $\underline{\int_a^b}$ exist as real numbers.

Proof. If we choose P to be the partition $P=\{a,b\}$, then $U(f,P)=(b-a)\sup_{x\in[a,b]}f(x)$ and $L(f,P)=(b-a)\inf_{x\in[a,b]}f(x)$. So, $U(f,P)\leq(b-a)M$ and $L(f,P)\geq(b-a)(-M)$. So, $-M(b-a)\leq\underline{\int_a^b}f$ and $\overline{\int_a^b}f\leq M(b-a)$ by the definition of supremum and infimum, respectively.

We now show that $\underline{\int_a^b}f\leq\overline{\int_a^b}f$. Let P,Q be any partitions of $[a,b]$. By the definition of $L(f,P)$ and $U(f,Q)$, we have

$$-\infty<L(f,P)\leq L(f,P\cup Q)\leq U(f,P\cup Q)\leq U(f,Q)<\infty.$$

So, we know that the set $\{U(f,Q)\colon Q \text{ is a partition of } [a,b]\}$ is nonempty and bounded from below. Similarly, the set $\{L(f,P)\colon P \text{ is a partition of } [a,b]\}$ is nonempty and bounded from above. Then, by the least upper bound property (Theorem 1.7.6), $\overline{\int_a^b}f$ and $\underline{\int_a^b}f$ exist as real numbers. Since $L(f,P)\leq U(f,Q)$ for any partitions P,Q, if we take the supremum over P of both dies, we get $\underline{\int_a^b}f\leq U(f,Q)$ for all partitions Q of $[a,b]$. Then, if we take the infimum over P of both sides, we get the desired inequality $\underline{\int_a^b}f\leq\overline{\int_a^b}f$.

□

4.2. **Riemann Integral.**

Definition 4.2.1 (**Riemann Integral**)**.** Let $a<b$ be real numbers, let $f\colon[a,b]\to\mathbb{R}$ be a bounded function. If $\overline{\int_a^b}f=\underline{\int_a^b}f$ we say that f is **Riemann integrable** on $[a,b]$, and we define

$$\int_a^b f:=\overline{\int_a^b}f=\underline{\int_a^b}f.$$

Remark 4.2.2. Defining the Riemann integral of an unbounded function takes more care, and we defer this issue to later courses.

Theorem 4.2.3 (**Laws of integration**)**.** *Let $a<b$ be real numbers, and let $f,g\colon[a,b]\to\mathbb{R}$ be Riemann integrable functions on $[a,b]$. Then*

(i) *The function $f+g$ is Riemann integrable on $[a,b]$, and $\int_a^b(f+g)=(\int_a^b f)+(\int_a^b g)$.*
(ii) *For any real number c, cf is Riemann integrable on $[a,b]$, and $\int_a^b(cf)=c(\int_a^b f)$.*
(iii) *The function $f-g$ is Riemann integrable on $[a,b]$, and $\int_a^b(f-g)=(\int_a^b f)-(\int_a^b g)$.*
(iv) *If $f(x)\geq 0$ for all $x\in[a,b]$, then $\int_a^b f\geq 0$.*
(v) *If $f(x)\geq g(x)$ for all $x\in[a,b]$, then $\int_a^b f\geq\int_a^b g$.*
(vi) *If there exists a real number c such that $f(x)=c$ for $x\in[a,b]$, then $\int_a^b f=c(b-a)$.*

(vii) *Let c, d be real numbers such that $c \leq a < b \leq d$. Then $[c, d]$ contains $[a, b]$. Define $F(x) := f(x)$ for all $x \in [a, b]$ and $F(x) := 0$ otherwise. Then F is Riemann integrable on $[c, d]$, and $\int_c^d F = \int_a^b f$.*

(viii) *Let c be a real number such that $a < c < b$. Then $f|_{[a,c]}$ and $f|_{[c,b]}$ are Riemann integrable on $[a, c]$ and $[c, b]$ respectively, and*

$$\int_a^b f = \int_a^c f|_{[a,c]} + \int_c^b f|_{[c,b]}.$$

Exercise 4.2.4. Prove Theorem 4.2.3.

Remark 4.2.5. Concerning Theorem 4.2.3(viii), we often write $\int_a^c f$ instead of $\int_a^c f|_{[a,c]}$.

4.2.1. *Riemann integrability of continuous functions.* So far we have discussed some properties of Riemann integrable functions, but we have not shown many functions that are actually Riemann integrable. In this section, we show that a continuous function on a closed interval is Riemann integrable.

Theorem 4.2.6. *Let $a < b$ be real numbers, and let $f: [a, b] \to \mathbb{R}$ be a continuous function on $[a, b]$. Then f is Riemann integrable.*

Proof. We will produce a family of partitions of the interval $[a, b]$ such that the upper and lower Riemann integrals of f are arbitrarily close to each other.

From Theorem 3.2.45, f is uniformly continuous on $[a, b]$. Let $\varepsilon > 0$. Then, by uniform continuity of f, there exists $\delta = \delta(\varepsilon) > 0$ such that, if $x, y \in [a, b]$ satisfy $|x - y| < \delta$, then $|f(x) - f(y)| < \varepsilon$. By the Archimedean property, there exists a positive integer N such that $(b - a)/N < \delta$.

Consider the partition P of the interval $[a, b]$ of the form

$$P = \{x_0, \ldots, x_N\} = \{a, a+(b-a)/N, a+2(b-a)/N, a+3(b-a)/N, \ldots, a+(N-1)(b-a)/N, b\}.$$

Note that $x_i - x_{i-1} = (b - a)/N$ for all $1 \leq i \leq N$. Since f is continuous on $[a, b]$, f is also continuous on $[x_{i-1}, x_i]$ for each $1 \leq i \leq N$. In particular, $f|_{[x_{i-1},x_i]}$ achieves its maximum and minimum for all $1 \leq i \leq N$. So, for each $1 \leq i \leq N$, there exist $m_i, M_i \in [x_{i-1}, x_i]$ such that

$$\inf_{x \in [x_{i-1},x_i]} f(x) = f(m_i), \quad \text{and} \quad \sup_{x \in [x_{i-1},x_i]} f(x) = f(M_i).$$

Since $x_i - x_{i-1} = (b - a)/N < \delta$, we have $|m_i - M_i| < \delta$ for each $1 \leq i \leq n$. Since f is uniformly continuous, we conclude that

$$\inf_{x \in [x_{i-1},x_i]} f(x) = f(m_i) > f(M_i) - \varepsilon = (\sup_{x \in [x_{i-1},x_i]} f(x)) - \varepsilon, \quad \forall\, 1 \leq i \leq n. \qquad (*)$$

We now estimate $U(f, P)$ and $L(f, P)$. By the definition of $U(f, P)$ and $L(f, P)$, we have

$$L(f, P) \leq U(f, P).$$

However, $L(f, P)$ is also close to $U(f, P)$ by $(*)$:

$$L(f, P) = \frac{b-a}{N} \sum_{i=1}^{N} (\inf_{x \in [x_{i-1},x_i]} f(x)) > \frac{b-a}{N} \sum_{i=1}^{N} [(\sup_{x \in [x_{i-1},x_i]} f(x)) - \varepsilon] = -(b - a)\varepsilon + U(f, P).$$

By the definition of $\underline{\int_a^b} f$, we conclude that

$$\underline{\int_a^b} f > -(b-a)\varepsilon + U(f,P).$$

By the definition of $\overline{\int_a^b} f$, we conclude that

$$\underline{\int_a^b} f > -(b-a)\varepsilon + \overline{\int_a^b} f.$$

Since $\varepsilon > 0$ is arbitrary, we get

$$\underline{\int_a^b} f \geq \overline{\int_a^b} f.$$

Combining this inequality with Lemma 4.1.6, we conclude that $\underline{\int_a^b} f = \overline{\int_a^b} f$. That is, f is Riemann integrable. □

Exercise 4.2.7. Let $a < b$ be real numbers. Let $f: [a,b] \to \mathbb{R}$ be a bounded function. Let $c \in [a,b]$. Assume that, for each $\delta > 0$, we know that f is Riemann integrable on the set $\{x \in [a,b]: |x-c| \geq \delta\}$. Then f is Riemann integrable on $[a,b]$.

4.2.2. *Piecewise Continuous Functions.* We can now expand a bit more the family of functions that are Riemann integrable.

Proposition 4.2.8. *Let $a < b$ be real numbers. Assume that $f: [a,b] \to \mathbb{R}$ is continuous at every point of $[a,b]$, except for a finite number of points. Assume also that f is bounded. Then f is Riemann integrable.*

Proof. By Theorem 4.2.3(viii) and an inductive argument, it suffices to consider the case that f is discontinuous at a single point $c \in [a,b]$. Let $\delta > 0$. Then f is continuous on the set $E := \{x \in [a,b]: |x-c| \geq \delta\}$. Note that E consists of either one or two closed intervals. Since $f|_E$ is continuous, we then conclude that $f|_E$ Riemann integrable by Theorem 4.2.6. Then Exercise 4.2.7 says that f is Riemann integrable on $[a,b]$, as desired. □

4.2.3. *Monotone Functions.* It turns out that monotone functions are Riemann integrable as well. There exist monotone functions that are not piecewise continuous, so the current section is not subsumed by the previous one.

Proposition 4.2.9. *Let $a < b$ be real numbers, and let $f: [a,b] \to \mathbb{R}$ be a monotone function. Then f is Riemann integrable.*

Proof. Let $\varepsilon > 0$. Without loss of generality, f is monotone increasing. Then $f(a) \leq f(x) \leq f(b)$ for all $x \in [a,b]$, so f is bounded. By the Archimedean property, there exists a positive integer N such that $(b-a)(f(b)-f(a))/N < \varepsilon$.

Consider the partition P of the interval $[a,b]$ of the form

$$P = \{x_0, \ldots, x_N\} = \{a, a+(b-a)/N, a+2(b-a)/N, a+3(b-a)/N, \ldots, a+(N-1)(b-a)/N, b\}.$$

Note that $x_i - x_{i-1} = (b-a)/N$ for all $1 \leq i \leq N$. We now estimate $U(f,P)$ and $L(f,P)$. By the definition of $U(f,P)$ and $L(f,P)$, we have

$$L(f,P) \leq U(f,P). \qquad (*)$$

However, since f is monotonically increasing,

$$\begin{aligned}
L(f,P) &= \frac{b-a}{N}\sum_{i=1}^{N}(\inf_{x\in[x_{i-1},x_i]} f(x)) \geq \frac{b-a}{N}\sum_{i=1}^{N} f(x_{i-1}) = \frac{b-a}{N}\left(f(x_0)+\sum_{i=1}^{N-1} f(x_i)\right)\\
&\geq \frac{b-a}{N}\left(f(x_0)+\sum_{i=1}^{N-1}(\sup_{x\in[x_{i-1},x_i]} f(x))\right) = \frac{b-a}{N}(f(x_0)-\sup_{x\in[x_{N-1},x_N]} f(x)) + U(f,P)\\
&\geq \frac{b-a}{N}(f(a)-f(b)) + U(f,P) \geq -\varepsilon + U(f,P).
\end{aligned}$$

By the definition of $\underline{\int_a^b} f$, we conclude that

$$\underline{\int_a^b} f \geq -\varepsilon + U(f,P).$$

By the definition of $\overline{\int_a^b} f$, we conclude that

$$\underline{\int_a^b} f \geq -\varepsilon + \overline{\int_a^b} f.$$

Since $\varepsilon > 0$ is arbitrary, we get

$$\underline{\int_a^b} f \geq \overline{\int_a^b} f.$$

Combining this inequality with Lemma 4.1.6, we conclude that $\underline{\int_a^b} f = \overline{\int_a^b} f$. That is, f is Riemann integrable. $\square$

4.2.4. *A Non-Riemann Integrable Function.* Unfortunately, not every function is Riemann integrable. We have seen that unbounded functions cause some difficulty in our definition of the Riemann integral, since their Riemann sums can be $+\infty$ or $-\infty$. However, there are even bounded functions that are not Riemann integrable.

Consider the following function $f\colon \mathbb{R} \to [0,1]$, which we encountered in our investigation of limits.

$$f(x) := \begin{cases} 1 & \text{, if } x \in \mathbb{Q} \\ 0 & \text{, if } x \notin \mathbb{Q} \end{cases}.$$

For any partition P of $[0,1]$, we automatically have $L(f,P) = 0$ and $U(f,P) = 1$. (Justify this statement.) Therefore, $\underline{\int_0^1} f = 0$ and $\overline{\int_0^1} f = 1$, so that this function f is not Riemann integrable on $[0,1]$.

4.3. **Fundamental Theorem of Calculus.** The Fundamental Theorem of Calculus says, roughly speaking, that differentiation and integration negate each other. This fact is remarkable on its own, but it will also allow us to actually compute a wide range of integrals. (Note that we have not yet been able to compute any integrals.)

Theorem 4.3.1 (First Fundamental Theorem of Calculus). *Let $a < b$ be real numbers. Let $f\colon [a,b] \to \mathbb{R}$ be a continuous function on $[a,b]$. Assume that f is also differentiable on $[a,b]$, and f' is Riemann integrable on $[a,b]$. Then*

$$\int_a^b f' = f(b) - f(a).$$

Proof. Let $P = \{x_0, \ldots, x_n\}$ be a partition of $[a,b]$. Then

$$f(b) - f(a) = f(x_n) - f(x_0) = \sum_{i=1}^{n} (f(x_i) - f(x_{i-1})). \qquad (*)$$

By the Mean Value Theorem (Corollary 3.3.28), for each $1 \leq i \leq n$ there exists $y_i \in [x_{i-1}, x_i]$ such that

$$(x_i - x_{i-1}) f'(y_i) = f(x_i) - f(x_{i-1}).$$

Substituting these equalities into $(*)$, we get

$$f(b) - f(a) = \sum_{i=1}^{n} (x_i - x_{i-1}) f'(y_i).$$

Applying the definitions of $L(f', P)$ and $U(f', Q)$, we have: for all partitions P, Q of $[a,b]$,

$$L(f', P) \leq f(b) - f(a) \leq U(f', Q).$$

From Definition 4.1.5, we then get

$$\underline{\int_a^b} f' \leq f(b) - f(a) \leq \overline{\int_a^b} f'. \qquad (**)$$

Since f' is Riemann integrable, $\underline{\int_a^b} f' = \overline{\int_a^b} f' = \int_a^b f'$. So, $(**)$ implies that $\int_a^b f' = f(b) - f(a)$, as desired. □

Theorem 4.3.2 (Second Fundamental Theorem of Calculus). *Let $a < b$ be real numbers. Let $f\colon [a,b] \to \mathbb{R}$ be a Riemann integrable function. Define a function $F\colon [a,b] \to \mathbb{R}$ by*

$$F(x) := \int_a^x f.$$

Then F is continuous. Moreover, if $x_0 \in [a,b]$ and if f is continuous at x_0, then F is differentiable at x_0 and $F'(x_0) = f(x_0)$.

Proof. Since $f\colon [a,b] \to \mathbb{R}$ is Riemann integrable, f is bounded by the definition of Riemann integrability. So, there exists a real number M such that $-M \leq f(x) \leq M$ for all $x \in [a,b]$. Let $x, y \in [a,b]$. Without loss of generality, $x \leq y$. Then, by Theorem 4.2.3(viii)

$$F(y) - F(x) = \int_a^y f - \int_a^x f = \int_x^y f. \qquad (*)$$

So, by Theorem 4.2.3(v),

$$-M(y - x) \leq \int_x^y f = F(y) - F(x) = \int_x^y f \leq M(y - x).$$

That is, $|F(y) - F(x)| \leq M\,|y - x|$. Interchanging the roles of x and y leaves this statement unchanged, so for any $x, y \in [a, b]$, we have

$$|F(y) - F(x)| \leq M\,|y - x|\,.$$

In particular, F is uniformly continuous, so F is continuous.

Now, suppose f is continuous at x_0. Using Proposition 3.3.12, it suffices to show: there exists a real number L such that, for any $\varepsilon > 0$, there exists $\delta > 0$ such that, if $y \in [a, b]$ satisfies $|y - x_0| < \delta$, then

$$|F(y) - [F(x_0) + L(y - x_0)]| \leq \varepsilon\,|y - x_0|\,. \qquad (**)$$

We set $L := f(x_0)$. Let $\varepsilon > 0$. Applying the continuity of f at x_0, there exists $\delta > 0$ such that if $y \in [a, b]$ satisfies $|y - x_0| < \delta$, then

$$f(x_0) - \varepsilon \leq f(y) \leq f(x_0) + \varepsilon.$$

Assume first that y satisfies $y > x_0$. Then integrating and applying Theorem 4.2.3(v),

$$(f(x_0) - \varepsilon)(y - x_0) \leq \int_{x_0}^{y} f \leq (f(x_0) + \varepsilon)(y - x_0).$$

So, using $(*)$,

$$|F(y) - F(x_0) - f(x_0)(y - x_0)| = \left|\left(\int_{x_0}^{y} f\right) - f(x_0)(y - x_0)\right| \leq \varepsilon\,|y - x_0|\,.$$

That is, we proved $(**)$ holds for $y > x_0$. The case $y < x_0$ is proven similarly, and the case $y = x_0$ follows since then both sides of $(**)$ are zero. □

4.3.1. *Consequences of the Fundamental Theorem.* One of the consequences of the Fundamental Theorem of Calculus is that we can now actually compute some integrals. For example, if $\alpha \in \mathbb{Q}$, $\alpha \neq -1$, and if $0 < a < b$ are real numbers, then $f(x) := (\alpha + 1)^{-1} x^{\alpha+1}$ satisfies $f'(x) = x^{\alpha}$. So, by Theorem 4.3.1,

$$\int_a^b x^{\alpha} = \frac{1}{\alpha + 1}(b^{\alpha+1} - a^{\alpha+1}).$$

Remark 4.3.3. Let $\beta \in \mathbb{Q}$, let $x > 0$ and let $f(x) := x^{\beta}$. Let's justify the formula $f'(x) = \beta x^{\beta-1}$. Write $\beta = p/q$ with $p \in \mathbb{Z}$ and $q \in \mathbb{N}$ with $q \neq 0$. Then $f(x) = (x^p)^{1/q}$. Recall that the function $h(x) = x^{1/q}$ is differentiable for $x > 0$ by the Inverse Function Theorem, Theorem 3.3.36. If $p \geq 0$, then we have already verified by explicit calculation that $g(x) := x^p$ is differentiable in Example 3.3.7. If $p < 0$, then $g(x) := x^p = 1/x^{-p}$ is differentiable by the Quotient Rule, or Theorem 3.3.17(vii). In summary, we can write $f(x) = h(g(x))$, where h is differentiable when $g(x) > 0$, g is differentiable when $x > 0$, and $g(x) > 0$ when $x > 0$. So, f is differentiable when $x > 0$, by the Chain Rule, Theorem 3.3.19.

Theorem 4.3.4. *Let $a < b$ be real numbers. Let $f, g\colon [a, b] \to \mathbb{R}$ be Riemann integrable functions. Then the product fg is Riemann integrable.*

Theorem 4.3.5 (Integration by Parts). *Let $a < b$ be real numbers. Let $f, g\colon [a, b] \to \mathbb{R}$ be differentiable functions such that f' and g' are Riemann integrable. Then*

$$\int_a^b f g' = f(b)g(b) - f(a)g(a) - \int_a^b f' g$$

Proof. Since f is differentiable on $[a, b]$ it is continuous on $[a, b]$ by Proposition 3.3.10. So, f is Riemann integrable by Theorem 4.2.6, and then fg' is Riemann integrable by Theorem 4.3.4. Similarly, $f'g$ is Riemann integrable.

Since f, g are differentiable, Theorem 3.3.17(iv) says that fg is differentiable and $(fg)' = f'g + fg'$. Since $f'g$ and fg' are Riemann integrable, $f'g + fg'$ is Riemann integrable by Theorem 4.2.3(i). So, applying Theorem 4.3.1,

$$\int_a^b (f'g + g'f) = \int_a^b (fg)' = f(b)g(b) - f(a)g(a).$$

□

Theorem 4.3.6 (**Change of Variables, version 1**). *Let $a < b$ be real numbers. Let $\phi\colon [a, b] \to [\phi(a), \phi(b)]$ be a differentiable function such that $\phi(a) < \phi(b)$ and such that ϕ' is Riemann integrable. Let $f\colon [\phi(a), \phi(b)] \to \mathbb{R}$ be continuous on $[\phi(a), \phi(b)]$. Then $(f \circ \phi)\phi'$ is Riemann integrable on $[a, b]$, and*

$$\int_a^b (f \circ \phi)\phi' = \int_{\phi(a)}^{\phi(b)} f.$$

Proof. Since ϕ is differentiable, ϕ is continuous. Then $f \circ \phi$ is continuous, since it is the composition of two continuous functions. For $t \in [\phi(a), \phi(b)]$, define $F(t) := \int_{\phi(a)}^t f$. Recall that f is Riemann integrable by Theorem 4.2.6. Now, $F'(t) = f(t)$ for all $t \in [\phi(a), \phi(b)]$ by the second fundamental theorem of calculus, Theorem 4.3.2. For any $x \in [a, b]$, define $g(x) := F \circ \phi(x)$. Then, by the Chain Rule, we have

$$g'(x) = F'(\phi(x))\phi'(x) = f(\phi(x))\phi'(x).$$

Note that g' is the product of two Riemann integrable functions, so g' is Riemann integrable (from Theorem 4.3.4). So, applying the first fundamental theorem of calculus, Theorem 4.3.1, we get

$$\int_a^b (f \circ \phi)\phi' = \int_a^b g' = g(b) - g(a) = F(\phi(b)) - F(\phi(a)) = F(\phi(b)) = \int_{\phi(a)}^{\phi(b)} f.$$

□

The following theorem is more difficult to prove, but it allows a change of variables for any Riemann integrable function f.

Theorem 4.3.7 (**Change of Variables, version 2**). *Let $a < b$ be real numbers. Let $\phi\colon [a, b] \to [\phi(a), \phi(b)]$ be differentiable, strictly monotone increasing function. Assume that ϕ' is Riemann integrable on $[a, b]$. Let $f\colon [\phi(a), \phi(b)] \to \mathbb{R}$ be Riemann integrable on $[\phi(a), \phi(b)]$. Then $(f \circ \phi)\phi'$ is Riemann integrable on $[a, b]$, and*

$$\int_a^b (f \circ \phi)\phi' = \int_{\phi(a)}^{\phi(b)} f.$$

5. Metric Spaces, Topology, Continuity, Compactness

We begin this section by presenting an abstract approach to analysis, which generalizes the things we did for analysis on the real line. Specifically, we will first generalize a lot of the arguments from the real line to the setting of metric spaces. We will then apply this general theory in our discussion of analysis on Euclidean spaces of any dimension, power series, and trigonometric functions.

5.1. Metric Spaces.

Recall that a sequence of real numbers $(x_n)_{n=0}^{\infty}$ converges to a real number x if and only if, for every real $\varepsilon > 0$, there exists an integer $N = N(\varepsilon)$ such that, for all $n > N$, we have $|x_n - x| < \varepsilon$. That is, eventually, the sequence $(x_n)_{n=0}^{\infty}$ is within a distance ε of x. And this is true for any $\varepsilon > 0$. So, whenever we have a space of points, and we can define some notion of distance between two points, then we should be able to make a similar definition of convergence of sequences. We are therefore led to consider the following question. What are the crucial properties of the distance function $d\colon \mathbb{R} \times \mathbb{R} \to [0, \infty)$ where $d(x, y) := |x - y|$ that allow us to consider convergence of sequences? The following properties suffice, as we shall see further below.

Definition 5.1.1 (Metric Space). A **metric space** (X, d) is a set X together with a function $d\colon X \times X \to [0, \infty)$ which satisfies the following properties.

- For all $x \in X$, we have $d(x, x) = 0$.
- For all $x, y \in X$ with $x \neq y$, we have $d(x, y) > 0$. (Positivity)
- For all $x, y \in X$, we have $d(x, y) = d(y, x)$. (Symmetry)
- For all $x, y, z \in X$, we have $d(x, z) \leq d(x, y) + d(y, z)$. (Triangle inequality)

Example 5.1.2 (The real line). As mentioned above, define the function $d\colon \mathbb{R}\times\mathbb{R} \to [0, \infty)$ by $d(x, y) := |x - y|$, where $x, y \in \mathbb{R}$. Then $(\mathbb{R}, d)$ is a metric space.

Example 5.1.3 (Euclidean space). Let n be a positive integer. Define $\mathbb{R}^n$ by

$$\mathbb{R}^n = \{(x_1, \ldots, x_n)\colon x_i \in \mathbb{R}\ \forall\, i \in \{1, \ldots, n\}\}.$$

Define the **Euclidean metric** (or ℓ_2 metric) $d\colon \mathbb{R}^n \times \mathbb{R}^n \to [0, \infty)$ by

$$d_{\ell_2}((x_1, \ldots, x_n), (y_1, \ldots, y_n)) := \left(\sum_{i=1}^{n} (x_i - y_i)^2\right)^{1/2}.$$

Then $(\mathbb{R}^n, d_{\ell_2})$ is a metric space.

There are actually many interesting metrics to consider on $\mathbb{R}^n$. Here is another one.

Example 5.1.4. Let n be a positive integer. Define the ℓ_1 metric $d_{\ell_1}\colon \mathbb{R}^n \times \mathbb{R}^n \to [0, \infty)$ by

$$d_{\ell_1}((x_1, \ldots, x_n), (y_1, \ldots, y_n)) := \sum_{i=1}^{n} |x_i - y_i|.$$

Then $(\mathbb{R}^n, d_{\ell_1})$ is a metric space.

The last two examples actually satisfy a few additional properties which you may recall from your linear algebra class.

Definition 5.1.5 (**Normed linear space**)**.** Let X be a vector space over $\mathbb{R}$. A **normed linear space** $(X, \|\cdot\|)$ is a vector space X over $\mathbb{R}$ together with a norm function $\|\cdot\| : X \to [0, \infty)$ which satisfies the following properties.

- $\|0\| = 0$.
- For all $x \in X$ with $x \neq 0$, we have $\|x\| > 0$. (Positivity)
- For all $x \in X$ and for all $\alpha \in \mathbb{R}$, we $\|\alpha x\| = |\alpha| \|x\|$. (Homogeneity)
- For all $x, y \in X$, we have $\|x + y\| \leq \|x\| + \|y\|$. (Triangle inequality)

Exercise 5.1.6. Let $(X, \|\cdot\|)$ be a normed linear space. Define $d \colon X \times X \to \mathbb{R}$ by $d(x, y) := \|x - y\|$. Show that (X, d) is a metric space.

Example 5.1.7. Let n be a positive integer. Define the ℓ_2 norm on $\mathbb{R}^n$ by

$$\|(x_1, \ldots, x_n)\|_{\ell_2} := \left(\sum_{i=1}^{n} x_i^2\right)^{1/2}.$$

Then $(\mathbb{R}^n, \|\cdot\|_{\ell_2})$ is a normed linear space. From Exercise 5.1.6, $(\mathbb{R}^n, d_{\ell_2})$ is a metric space, which we saw in Example 5.1.3.

Example 5.1.8. Let n be a positive integer. Define the ℓ_1 norm on $\mathbb{R}^n$ by

$$\|(x_1, \ldots, x_n)\|_{\ell_1} := \sum_{i=1}^{n} |x_i|.$$

Then $(\mathbb{R}^n, \|\cdot\|_{\ell_1})$ is a normed linear space. From Exercise 5.1.6, $(\mathbb{R}^n, d_{\ell_1})$ is a metric space, which we saw in Example 5.1.4.

Example 5.1.9. Let n be a positive integer. Define the ℓ_∞ norm on $\mathbb{R}^n$ by

$$\|(x_1, \ldots, x_n)\|_{\ell_\infty} := \max_{i=1,\ldots,n} |x_i|.$$

Then $(\mathbb{R}^n, \|\cdot\|_{\ell_\infty})$ is a normed linear space. From Exercise 5.1.6, $(\mathbb{R}^n, d_{\ell_\infty})$ is a metric space, where $d_{\ell_\infty}((x_1, \ldots, x_n), (y_1, \ldots, y_n)) := \max_{i=1}^{n} |x_i - y_i|$.

Example 5.1.10. Let n be a positive integer and let $1 \leq p < \infty$ be a real number. Define the ℓ_p norm on $\mathbb{R}^n$ by

$$\|(x_1, \ldots, x_n)\|_{\ell_p} := \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}.$$

Then $(\mathbb{R}^n, \|\cdot\|_{\ell_p})$ is a normed linear space, though the triangle inequality is a bit more difficult to prove. From Exercise 5.1.6, $(\mathbb{R}^n, d_{\ell_p})$ is a metric space.

Exercise 5.1.11. Let n be a positive integer and let $x \in \mathbb{R}^n$. Show: $\|x\|_{\ell_\infty} = \lim_{p \to \infty} \|x\|_{\ell_p}$.

Euclidean space is actually even more special than a normed linear space, which we also learned in linear algebra class. Specifically, $\mathbb{R}^n$ is an inner product space.

Definition 5.1.12 (**Real Inner product space**)**.** Let X be a vector space over $\mathbb{R}$. A **real inner product space** $(X, \langle \cdot, \cdot \rangle)$ is a vector space X over $\mathbb{R}$ together with an inner product function $\langle \cdot, \cdot \rangle \colon X \times X \to \mathbb{R}$ which satisfies the following properties.

- $\langle 0, 0 \rangle = 0$.

- For all $x \in X$ with $x \neq 0$, we have $\langle x, x\rangle > 0$.
- For all $x, y \in X$, we have $\langle x, y\rangle = \langle y, x\rangle$. (Symmetry)
- For all $x \in X$ and for all $\alpha \in \mathbb{R}$, we $\langle \alpha x, y\rangle = \alpha\langle x, y\rangle$. (Homogeneity)
- For all $x, y, z \in X$, we have $\langle x + y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$. (Linearity)

Exercise 5.1.13. Let $(X, \langle\cdot,\cdot\rangle)$ be a real inner product space. Define $\|\cdot\| : X \to [0,\infty)$ by $\|x\| := \sqrt{\langle x, x\rangle}$. Show that $(X, \|\cdot\|)$ is a normed linear space. Consequently, from Exercise 5.1.6, if we define $d\colon X \times X \to [0,\infty)$ by $d(x, y) := \sqrt{\langle (x - y), (x - y)\rangle}$, then (X, d) is a metric space.

In order to prove Exercise 5.1.13, the following inequality is useful.

Theorem 5.1.14 (Cauchy-Schwarz Inequality). *Let $(X, \langle\cdot,\cdot\rangle)$ be a real inner product space. Then, for all $x, y \in X$, we have*

$$|\langle x, y\rangle| \leq \langle x, x\rangle^{1/2}\langle y, y\rangle^{1/2}.$$

Proof. It follows from Definition 5.1.12 that, if $x = 0$, then $\langle x, y\rangle = 0$ for any $y \in X$. (You should have proven this in your linear algebra class.) So, if $x = 0$, then both sides of the Cauchy-Schwarz inequality are zero, and the inequality therefore holds. Similarly, if $y = 0$, then both sides of the inequality are zero. We therefore assume that $x \neq 0$ and $y \neq 0$. For any $x \in X$, define $\|x\| := \sqrt{\langle x, x\rangle}$.

Starting from x, we subtract the projection of x onto y. Define $\delta := -\langle x, y\rangle / \|y\|^2$. We then have

$$0 \leq \|x + \delta y\|^2 = \|x\|^2 + 2\delta\langle x, y\rangle + |\delta|^2 \|y\|^2 = \|x\|^2 - |\langle x, y\rangle|^2 / \|y\|^2 .$$

□

Remark 5.1.15. For any $(x_1, \ldots, x_n), (y_1, \ldots, y_n) \in \mathbb{R}^n$, define

$$\langle (x_1, \ldots, x_n), (y_1, \ldots, y_n)\rangle := \sum_{i=1}^{n} x_i y_i.$$

Then $(\mathbb{R}^n, \langle\cdot,\cdot\rangle)$ is an inner product space. Note that

$$\|(x_1, \ldots, x_n)\|_{\ell_2} = \langle (x_1, \ldots, x_n), (x_1, \ldots, x_n)\rangle^{1/2}.$$

However, there does not exist an inner product $\langle\cdot,\cdot\rangle'$ on $\mathbb{R}^n$ such that $\|(x_1, \ldots, x_n)\|_{\ell_1}$ is equal to $(\langle (x_1, \ldots, x_n), (x_1, \ldots, x_n)\rangle')^{1/2}$. The last statement is more difficult to prove.

In the first part of this course, we will mostly focus on metric spaces. Once we deal with Fourier series, the subject of inner product spaces will reappear. However, we will need to deal with complex inner product spaces, so we now recall their definition. Recall that, for $\alpha, \beta \in \mathbb{R}$ we define

$$\overline{\alpha + \beta\sqrt{-1}} := \alpha - \beta\sqrt{-1}.$$

Definition 5.1.16 (Complex Inner product space). Let X be a vector space over $\mathbb{C}$. A **complex inner product space** $(X, \langle\cdot,\cdot\rangle)$ is a vector space X together with an inner product function $\langle\cdot,\cdot\rangle\colon X \times X \to \mathbb{C}$ which satisfies the following properties.

- $\langle 0, 0\rangle = 0$.
- For all $x \in X$ with $x \neq 0$, we have $\langle x, x\rangle > 0$.

- For all $x, y \in X$, we have $\langle x, y\rangle = \overline{\langle y, x\rangle}$. (Conjugate Symmetry)
- For all $x \in X$ and for all $\alpha \in \mathbb{C}$, we $\langle \alpha x, y\rangle = \alpha\langle x, y\rangle$. (Homogeneity)
- For all $x, y, z \in X$, we have $\langle x + y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$. (Linearity)

Remark 5.1.17. Let $(X, \langle\cdot,\cdot\rangle)$ be a complex inner product space. Then, for any $x, y, z \in X$ and for any $\alpha \in \mathbb{C}$, it follows from Definition 5.1.16 that

- $\langle x, y + z\rangle = \langle x, y\rangle + \langle x, z\rangle$.
- $\langle x, \alpha y\rangle = \overline{\alpha}\langle x, y\rangle$.

Example 5.1.18. Let n be a positive integer. Let $(z_1, \ldots, z_n), (w_1, \ldots, x_n) \in \mathbb{C}^n$. The standard inner product on $\mathbb{C}^n$ is defined by

$$\langle (z_1, \ldots, z_n), (w_1, \ldots, w_n)\rangle := \sum_{i=1}^{n} z_i\overline{w_i}.$$

Then $(\mathbb{C}^n, \langle\cdot,\cdot\rangle)$ is a complex inner product space.

Exercise 5.1.19 (Cauchy-Schwarz Inequality). Let $(X, \langle\cdot,\cdot\rangle)$ be a complex inner product space. Let $x, y \in X$. Modify the proof of Theorem 5.1.14 to prove the Cauchy-Schwarz inequality for complex inner product spaces:

$$|\langle x, y\rangle| \leq \langle x, x\rangle^{1/2}\langle y, y\rangle^{1/2}.$$

Exercise 5.1.20. Let $(X, \langle\cdot,\cdot\rangle)$ be a complex inner product space. Let $x, y \in X$. As usual, let $\|x\| := \sqrt{\langle x, x\rangle}$. Prove **Pythagoras's theorem**: if $\langle x, y\rangle = 0$, then $\|x + y\|^2 = \|x\|^2 + \|y\|^2$

5.1.1. *Convergence of Sequences.* We now define convergence of sequences in a metric space in a way which imitates the convergence of sequences of real numbers.

Definition 5.1.21. Let (X, d) be a metric space. Let $(x^{(j)})_{j=k}^{\infty}$ be a sequence of elements of X. Let x be an element of X. We say that the sequence $(x^{(j)})_{j=k}^{\infty}$ **converges to x with respect to the metric** d if and only if, for every $\varepsilon > 0$, there exists an integer $J = J(\varepsilon)$ such that, for all $j > J$, we have $d(x^{(j)}, x) < \varepsilon$.

Proposition 5.1.22. *Let n be a positive integer. Let $x \in \mathbb{R}^n$. Let $(x^{(j)})_{j=k}^{\infty}$ be a sequence of elements of $\mathbb{R}^n$. We write $x^{(j)} = (x_1^{(j)}, \ldots, x_n^{(j)})$, so that for each $1 \leq i \leq n$, we have $x_i^{(j)} \in \mathbb{R}$, that is, $x_i^{(j)}$ is the i^{th} coordinate of $x^{(j)}$. Then the following three statements are equivalent.*

- *$(x^{(j)})_{j=k}^{\infty}$ converges to x with respect to d_{ℓ_1}.*
- *$(x^{(j)})_{j=k}^{\infty}$ converges to x with respect to d_{ℓ_2}.*
- *$(x^{(j)})_{j=k}^{\infty}$ converges to x with respect to d_{ℓ_∞}.*

Exercise 5.1.23. Prove Proposition 5.1.22.

Due to Proposition 5.1.22, we say that the metrics d_{ℓ_1}, d_{ℓ_2} and d_{ℓ_∞} are equivalent on $\mathbb{R}^n$. In fact, for any $p, p' \in \mathbb{R}$ with $1 \leq p, p' \leq \infty$, the metrics d_{ℓ_p} and $d_{\ell_{p'}}$ are equivalent on $\mathbb{R}^n$. In fact, something stronger is true. Let $\|\cdot\|_a, \|\cdot\|_b$ be any two norms on $\mathbb{R}^n$. Define the metrics d_a, d_b so that, for any $x, y \in \mathbb{R}^n$, we have $d_a(x, y) := \|x - y\|_a$ and $d_b(x, y) := \|x - y\|_b$. Then d_a and d_b are equivalent on $\mathbb{R}^n$.

As in the case of the real line, a sequence cannot converge to two distinct points.

Proposition 5.1.24. *Let (X, d) be a metric space. Let $(x^{(j)})_{j=k}^{\infty}$ be a sequence of elements of X. Let x, x' be elements of X. Assume that the sequence $(x^{(j)})_{j=k}^{\infty}$ converges to x with respect to d. Assume also that the sequence $(x^{(j)})_{j=k}^{\infty}$ converges to x' with respect to d. Then $x = x'$.*

Exercise 5.1.25. Prove Proposition 5.1.24.

Due to Proposition 5.1.24, if $(x^{(j)})_{j=k}^{\infty}$ is a sequence of elements of X which converges to x with respect to d, we then write $\lim_{j\to\infty} x^{(j)} = x$. Although the latter notation has no indication of the metric d, confusion should often not arise as to what metric the convergence is using.

5.2. **Topology of Metric Spaces.** The open intervals such as $(0, 1)$ and closed intervals such as $[1, 2]$ played a central role in analysis on the real line. The open interval is a special case of an open set, and the closed interval is a special case of a closed set. There is a way to generalize the notions of open and closed set to general metric spaces, so we pursue these notions now. We begin by generalizing the notion of an open interval to the notion of a metric ball. The language of topology is used everywhere throughout mathematics, so it is quite useful even just to learn the terminology.

Definition 5.2.1 (Metric Ball). Let (X, d) be a metric space, let x_0 be a point in X, and let $r > 0$ be a positive real number. We define the **ball** $B_{(X,d)}(x_0, r)$ in X, centered at x_0 with radius r to be the set

$$B_{(X,d)}(x_0, r) := \{x \in X : d(x, x_0) < r\}.$$

When the metric space (X, d) is apparent, we abbreviate $B_{(X,d)}(x_0, r)$ as $B(x_0, r)$. If $(X, \|\cdot\|)$ is a normed linear space, we write $B_{(X,\|\cdot\|)}(x_0, r)$ to denote the set $\{x \in X : \|x - x_0\| < r\}$.

Example 5.2.2. Let $x, y \in \mathbb{R}$. In $\mathbb{R}$ with the metric $d(x, y) := |x - y|$, note that $B_{(\mathbb{R},d)}(x_0, r)$ is the open interval $(x_0 - r, x_0 + r)$.

Example 5.2.3. In $\mathbb{R}^2$ with the metric d_{ℓ_2}, we have

$$B_{(\mathbb{R}^2,d_{\ell_2})}((0, 0), 1) = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 < 1\}.$$

However, with the metric d_{ℓ_1}, we have

$$B_{(\mathbb{R}^2,d_{\ell_1})}((0, 0), 1) = \{(x, y) \in \mathbb{R}^2 : |x| + |y| < 1\}.$$

Also, with the metric d_{ℓ_∞}, we have

$$B_{(\mathbb{R}^2,d_{\ell_\infty})}((0, 0), 1) = \{(x, y) \in \mathbb{R}^2 : \max\{|x| , |y|\} < 1\}.$$

So, $B_{(\mathbb{R}^2,d_{\ell_1})}((0, 0), 1)$ is a diamond, $B_{(\mathbb{R}^2,d_{\ell_2})}((0, 0), 1)$ is a disc, and $B_{(\mathbb{R}^2,d_{\ell_\infty})}((0, 0), 1)$ is a square. Moreover,

$$B_{(\mathbb{R}^2,d_{\ell_1})}((0, 0), 1) \subseteq B_{(\mathbb{R}^2,d_{\ell_2})}((0, 0), 1) \subseteq B_{(\mathbb{R}^2,d_{\ell_\infty})}((0, 0), 1).$$

Remark 5.2.4. Note that if (X, d) is a nonempty metric space, and if $x_0 \in X$ with $r > 0$, then $B_{(X,d)}(x_0, r)$ is nonempty, since it contains x_0. Moreover, if $0 < r < r'$, we have the containment $B_{(X,d)}(x_0, r) \subseteq B_{(X,d)}(x_0, r')$.

Definition 5.2.5. Let (X, d) be a metric space, let E be a subset of X, and let x_0 be a point in X. We say that x_0 is an **interior point** of E if and only if there exists $r > 0$ such that $B(x_0, r) \subseteq E$. We say that x_0 is an **exterior point** of E if and only if there exists $r > 0$ such that $B(x_0, r) \cap E = \emptyset$. We say that x_0 is a **boundary point** of E if and only if x_0 is neither an interior point nor an exterior point of E.

Remark 5.2.6. The set of all interior points of E is called the **interior** of E, and it is denoted as $\text{int}(E)$. The set of all exterior points of E is called the **exterior** of E, and it is denoted as $\text{ext}(E)$. The set of all boundary points of E is called the **boundary** of E, and it is denoted as ∂E.

Remark 5.2.7. If x_0 is an interior point of E, then x_0 is an element of E. If x_0 is an exterior point of E, then x_0 is not an element of E. If x_0 is a boundary point of E, then x_0 may or may not be an element of E.

Example 5.2.8. Consider the real line $\mathbb{R}$ with the usual metric. The open interval $(0, 1)$ has interior $(0, 1)$, it has exterior $(-\infty, 0) \cup (1, \infty)$, and it has boundary $\{0, 1\}$. The closed interval $[0, 1]$ has interior $(0, 1)$, it has exterior $(-\infty, 0) \cup (1, \infty)$, and it has boundary $\{0, 1\}$. The half-open interval $(0, 1]$ has interior $(0, 1)$, it has exterior $(-\infty, 0) \cup (1, \infty)$, and it has boundary $\{0, 1\}$.

Definition 5.2.9. Let (X, d) be a metric space, let E be a subset of X, and let x_0 be a point in X. We say that x_0 is an **adherent point** of E if and only if for every real $r > 0$, we have $B(x_0, r) \cap E \neq \emptyset$. The set of all adherent points of E is called the **closure** of E and is denoted by $\overline{E}$.

The definitions of interior, exterior, boundary and closure are related by the following proposition.

Proposition 5.2.10. *Let (X, d) be a metric space, let E be a subset of X, and let x_0 be a point in X. Then the following statements are equivalent.*

- *x_0 is an adherent point of E.*
- *x_0 is either an interior point of E or a boundary point of E.*
- *There exists a sequence $(x^{(j)})_{j=1}^{\infty}$ of elements of E which converges to x_0 with respect to the metric d.*

Exercise 5.2.11. Prove Proposition 5.2.10.

We now define open and closed sets in terms of the boundary points of a set. As we will see below, we can equivalently define open and closed sets using only open balls.

Definition 5.2.12 (Open and Closed Sets). Let (X, d) be a metric space, let E be a subset of X. We say that E is **closed** if and only if E contains all of its boundary points, i.e. when $\partial E \subseteq E$. We say that E is **open** if and only if E contains none of its boundary points, i.e. when $(\partial E) \cap E = \emptyset$. If E contains some of its boundary points but not others, then E is neither open nor closed.

Remark 5.2.13. If a set E has no boundary, then E is simultaneously open and closed. For example, if $E = X$, then E is both open and closed. Also, if $E = \emptyset$, then E is both open and closed.

Example 5.2.14. We continue Example 5.2.8. Consider the real line $\mathbb{R}$ with the usual metric. The open interval $(0,1)$ has boundary $\{0,1\}$, so the open interval is open. The closed interval $[0,1]$ has boundary $\{0,1\}$, so the closed interval is closed. The half-open interval $(0,1]$ has boundary $\{0,1\}$, so the half-open interval is neither open nor closed.

As promised, we now show some equivalent definitions of open and closed sets.

Proposition 5.2.15 (Properties of Open and Closed Sets). *Let (X,d) be a metric space.*

(i) *Let E be a subset of X. Then E is open if and only if $E = \mathrm{int}(E)$. That is, E is open if and only if, for every $x \in E$, there exists an $r > 0$ such that $B(x,r) \subseteq E$.*

(ii) *Let E be a subset of X. Then E is closed if and only if E contains all of its adherent points, i.e. when $E = \overline{E}$. That is, E is closed if and only if, for every convergent sequence $(x^{(j)})_{j=0}^{\infty}$ consisting of elements of E, the limit $\lim_{n\to\infty} x_n$ of the sequence also lies in E.*

(iii) *For any $x_0 \in X$, for any $r > 0$, the open ball $B(x_0,r)$ is an open set. The set $\{x \in X : d(x,x_0) \leq r\}$ is a closed set. The latter set is sometimes called the **closed ball** of radius r centered at x_0.*

(iv) *Let $x_0 \in X$. Then the singleton set $\{x_0\}$ is closed.*

(v) *If E is a subset of X, then E is open if and only if $X \setminus E$ is closed. Here we have denoted $X \setminus E := \{x \in X : x \notin E\}$ as the complement of E in X.*

(vi) *If $E_1, \ldots, E_n$ is a finite collection of open sets, then $E_1 \cap \cdots \cap E_n$ is an open set. If $F_1, \ldots, F_n$ is a finite collection of closed sets, then $F_1 \cup \cdots \cup F_n$ is a closed set.*

(vii) *If $\{E_\alpha\}_{\alpha\in I}$ is collection of open sets, (where the index set I can be finite, countable, or uncountable), then $\cup_{\alpha\in I} E_\alpha$ is an open set. If $\{F_\alpha\}_{\alpha\in I}$ is collection of closed sets, (where the index set I can be finite, countable, or uncountable), then $\cap_{\alpha\in I} F_\alpha$ is a closed set.*

(viii) *If E is any subset of X, then $\mathrm{int}(E)$ is the largest open set contained in E. That is, $\mathrm{int}(E)$ is open, and if V is any open set such that $V \subseteq E$, then $V \subseteq \mathrm{int}(E)$ also. Similarly, $\overline{E}$ is the smallest closed set containing E. That is, $\overline{E}$ is closed, and if V is any closed set such that $V \supseteq E$, then $V \supseteq \overline{E}$ also.*

Exercise 5.2.16. Prove Proposition 5.2.15.

Remark 5.2.17. Proposition 5.2.15(vi) does not hold for countable collections of sets, as we can see from the following examples which involve open and closed intervals on the real line.

$$\bigcap_{n\in\mathbb{N}} \left(-\frac{1}{n+1}, 1+\frac{1}{n+1}\right) = [0,1].$$

$$\bigcup_{n\in\mathbb{N}} \left[\frac{1}{n+1}, 1-\frac{1}{n+1}\right] = (0,1).$$

As we see from the following example, it is natural to consider the open and closed sets of a subset of a metric space. Such notions are formalized by the relative topology.

Example 5.2.18. Consider $\mathbb{R}^2$, and define $Y := \{(x,0) : x \in \mathbb{R}\}$, so that Y is a subset of $\mathbb{R}^2$. Let d_{ℓ_2} denote the ℓ_2 metric on $\mathbb{R}^2$. If we restrict d_{ℓ_2} to Y resulting in $d_{\ell_2}|_{Y\times Y}$, then

$(Y, d_{\ell_2}|_{Y\times Y})$ is a metric space. In fact, we can identify $(Y, d_{\ell_2}|_{Y\times Y})$ with the real line $\mathbb{R}$ with its usual metric. Now, consider the set

$$E := \{(x, 0)\colon\ -1 < x < 1\}.$$

Then E is a subset of Y, and E is also a subset of $\mathbb{R}^2$. When we consider E as a subset of Y, then E is an open set, since E is equal to the ball $B_{(Y,d_{\ell_2}|_{Y\times Y})}((0,0),1)$. However, when we consider E as a subset of $\mathbb{R}^2$, then E is no longer an open set. To see this, note that for any $r > 0$, the ball $B_{(\mathbb{R}^2,d_{\ell_2})}((0,0),r)$ is not contained in E. So, by Proposition 5.2.15(i), E is not open in $(\mathbb{R}^2, d_{\ell_2})$.

To summarize the above example: there is a sensible way to discuss open sets of a subset of a metric space, and it involves restricting the metric.

Definition 5.2.19 (**Relative Topology**). Let (X, d) be a metric space, let Y be a subset of X, and let E be a subset of Y. We say that E is **relatively open with respect to** Y if and only if E is open in the metric space $(Y, d|_{Y\times Y})$. Similarly, we say that E is **relatively closed with respect to** Y if and only if E is closed in the metric space $(Y, d|_{Y\times Y})$.

The definitions of relatively open and relatively closed sets are consistent with set intersection in the following way.

Proposition 5.2.20. *Let (X, d) be a metric space, let Y be a subset of X, and let E be a subset of Y.*

- *E is relatively open with respect to Y if and only if $E = V \cap Y$ for some set $V \subseteq X$ which is open in X.*
- *E is relatively closed with respect to Y if and only if $E = V \cap Y$ for some set $V \subseteq X$ which is closed in X.*

Proof. We only prove the first assertion, since the second assertion is proven similarly. First, assume that E is relatively open with respect to Y. Then E is open in the metric space $(Y, d|_{Y\times Y})$. So, by Proposition 5.2.15(i), for any $e \in E$ there exists $r = r(e) > 0$ such that $B_{(Y,d|_{Y\times Y})}(e, r(e)) \subseteq E$. Note that

$$E = \bigcup_{e\in E} B_{(Y,d|_{Y\times Y})}(e, r(e)). \tag{5.2.1}$$

Specifically, every set on the right is contained in E, so the union of the sets on the right is contained in E. And conversely, every element of E appears on the right side, so the right side contains the left side. We therefore deduce the equality (5.2.1). With this equality in mind, define

$$V := \bigcup_{e\in E} B_{(X,d)}(e, r(e)). \tag{5.2.2}$$

From Proposition 5.2.15(i), V is open in (X, d). For any $e \in E$, we have

$$Y \cap B_{(X,d)}(e, r(e)) = B_{(Y,d|_{Y\times Y})}(e, r(e)). \tag{5.2.3}$$

To see this, note that if $y \in B_{(Y,d|_{Y\times Y})}(e, r(e))$, then $d|_{Y\times Y}(e, y) < r(e)$. Since e and y are in Y, we have $d|_{Y\times Y}(e, y) = d(e, y)$. So, $d(e, y) < r(e)$, so that $y \in B_{(X,d)}(e, r(e))$. Conversely, let $y \in Y \cap B_{(X,d)}(e, r(e))$. then $d(e, y) < r(e)$. Since e and y are in Y, we again have

$d(e,y) = d|_{Y\times Y}(e,y)$. So, $d|_{Y\times Y}(e,y) < r(e)$, so that $y \in B_{(Y,d|_{Y\times Y})}(e,r(e))$. In conclusion, (5.2.3) holds. Combining (5.2.1), (5.2.2) and (5.2.3), we conclude that

$$V\cap Y = Y\cap\left(\bigcup_{e\in E} B_{(X,d)}(e,r(e))\right) = \bigcup_{e\in E}\left(Y\cap B_{(X,d)}(e,r(e))\right) = \bigcup_{e\in E} B_{(Y,d|_{Y\times Y})}(e,r(e)) = E.$$

Conversely, assume that there exists $V\subseteq X$ which is open in X, such that $V\cap Y = E$. We will show that E is relatively open in Y. By Proposition 5.2.15(i), for any $v\in V$ there exists $r = r(v) > 0$ such that $B_{(X,d)}(e,r(e))\subseteq V$. Note that

$$V = \bigcup_{v\in V} B_{(X,d)}(v,r(v)). \tag{5.2.4}$$

Now, consider the set

$$E' := \bigcup_{v\in V\cap Y} B_{(Y,d|_{Y\times Y})}(v,r(v)). \tag{5.2.5}$$

From Proposition 5.2.15(i), E' is open in $(Y,d|_{Y\times Y})$. We will then conclude by showing that $E = E'$. As before, for any $v\in V\cap Y$, we have the equality

$$Y\cap B_{(X,d)}(v,r(v)) = B_{(Y,d|_{Y\times Y})}(v,r(v)). \tag{5.2.6}$$

Combining (5.2.4), (5.2.5) and (5.2.6), we get $E = V\cap Y = E'$, completing the proof. □

5.3. **Cauchy sequences and Completeness.** Recall the following definition of a subsequence. Let $(x^{(j)})_{j=k}^{\infty}$ be a sequence of elements of a metric space (X,d). Let $j_1, j_2, \ldots$ be an increasing sequence of integers such that

$$k\leq j_1 < j_2 < j_3 < \cdots$$

We then say that $(x^{(j_m)})_{m=1}^{\infty}$ is a **subsequence** of the sequence $(x^{(j)})_{j=k}^{\infty}$.

Lemma 5.3.1. *Let $(x^{(j)})_{j=k}^{\infty}$ be a sequence of elements of a metric space (X,d) which converges to some limit $x\in X$. Then every subsequence of $(x^{(j)})_{j=k}^{\infty}$ also converges to x.*

Exercise 5.3.2. Prove Lemma 5.3.1.

Definition 5.3.3 (Cauchy sequence). Let $(x^{(j)})_{j=k}^{\infty}$ be a sequence of elements of a metric space (X,d). We say that the sequence $(x^{(j)})_{j=k}^{\infty}$ is a **Cauchy sequence** if and only if, for every $\varepsilon > 0$, there exists an integer $J = J(\varepsilon)$ such that, for all $j,\ell > J$, we have $d(x^{(j)},x^{(\ell)}) < \varepsilon$.

Lemma 5.3.4. *Let $(x^{(j)})_{j=k}^{\infty}$ be a sequence of elements of a metric space (X,d) which converges to some limit $x\in X$. Then $(x^{(j)})_{j=k}^{\infty}$ is also a Cauchy sequence.*

Exercise 5.3.5. Prove Lemma 5.3.1.

The converse is false sometimes, as we saw in Corollary 2.2.12. If $(x^{(j)})_{j=k}^{\infty}$ is a Cauchy sequence in X, then the sequence $(x^{(j)})_{j=k}^{\infty}$ may not converge to an element of X. For example, we saw that a Cauchy sequence of rationals numbers may converge to a real number which is not itself a rational number.

However, a Cauchy sequence of real numbers does converge to a real number, as we learned in Theorem 2.2.10, which we restate:

Theorem 5.3.6. *Let $(\mathbb{R}, d)$ be the real line with the usual metric, so that for any $x, y \in \mathbb{R}$, we have $d(x, y) := |x - y|$. If $(x^{(j)})_{j=k}^{\infty}$ is a Cauchy sequence of elements of $\mathbb{R}$, then there exists some $x \in \mathbb{R}$ such that $(x^{(j)})_{j=k}^{\infty}$ converges to x with respect to d.*

We cast the latter property into the following definition.

Definition 5.3.7 (**Completeness**)**.** Let (X, d) be a metric space. We say that (X, d) is **complete** if and only if the following property holds. For any Cauchy sequence $(x^{(j)})_{j=k}^{\infty}$ of elements of X, then there exists some $x \in X$ such that $(x^{(j)})_{j=k}^{\infty}$ converges to x with respect to d.

So, $(\mathbb{R}, d)$ is a complete metric space by Theorem 5.3.6. However, the metric space $(\mathbb{Q}, d)$ is not complete. For example, we can construct a sequence of rational numbers that converges to $\sqrt{2}$, but $\sqrt{2}$ is not a rational number.

Complete metric spaces are always closed when they are considered as subsets of other metric spaces, as we now show.

Proposition 5.3.8.

- *Let (X, d) be a metric space, and let Y be a subset of X, so that $(Y, d|_{Y \times Y})$ is a metric space. If $(Y, d|_{Y \times Y})$ is complete, then Y is closed in (X, d).*
- *Conversely, assume that (X, d) is a complete metric space and that Y is a closed subset of X. Then $(Y, d|_{Y \times Y})$ is complete.*

Exercise 5.3.9. Prove Proposition 5.3.8.

A metric space (X, d) which is not complete may or may not be closed, when considered as a subset of another metric space. For example, if d is the standard metric on $\mathbb{R}$, then $\mathbb{Q}$ is closed in $(\mathbb{Q}, d|_{\mathbb{Q} \times \mathbb{Q}})$, but $\mathbb{Q}$ is not closed in $(\mathbb{R}, d)$. We briefly mention that, given any metric space (X, d), there is a way to form the **completion** $(\overline{X}, \overline{d})$ of (X, d), so that $(\overline{X}, \overline{d})$ is a complete metric space that contains (X, d). This procedure imitates our construction of the real numbers using Cauchy sequences of rational numbers.

5.4. **Compactness.** We have now arrived at the extremely useful concept of compactness. Compactness expresses the exact properties that are needed to obtain the conclusion of the Bolzano-Weierstrass Theorem. As we recall, the Bolzano-Weierstrass Theorem 2.8.9 is very useful, and likewise compactness is very useful.

Definition 5.4.1 (**Boundedness**)**.** A sequence $(x^{(j)})_{j=k}^{\infty}$ in a metric space (X, d) is said to be **bounded** if and only if there exists $x \in X$ and there exists $r > 0$ such that $x^{(j)} \in B(x, r)$ for all $j \geq k$. Similarly, a subset E of a metric space (X, d) is said to be **bounded** if and only if there exists $x \in X$ and there exists $r > 0$ such that $E \subseteq B(x, r)$.

Theorem 5.4.2 (**Bolzano-Weierstrass**)**.** *Let $(\mathbb{R}, d)$ be the real line with the standard metric. Let $(x^{(j)})_{j=k}^{\infty}$ be a bounded sequence in $\mathbb{R}$. Then there exists a subsequence of $(x^{(j)})_{j=k}^{\infty}$ that converges in $(\mathbb{R}, d)$.*

Corollary 5.4.3. *Let n be a positive integer. Let $(\mathbb{R}^n, d)$ denote Euclidean space with either of the metrics $d = d_{\ell_2}$ or $d = d_{\ell_1}$. Let $(x^{(j)})_{j=k}^{\infty}$ be a bounded sequence in $\mathbb{R}^n$. Then there exists a subsequence of $(x^{(j)})_{j=k}^{\infty}$ that converges in $(\mathbb{R}^n, d)$.*

This convergent subsequence property is called compactness.

Definition 5.4.4 (Compactness). A metric space (X, d) is said to be **compact** if and only if every sequence in (X, d) has at least one convergent subsequence.

A compact metric space satisfies the following two special properties.

Proposition 5.4.5. *Let (X, d) be a compact metric space. Then (X, d) is both complete and bounded.*

Exercise 5.4.6. Prove Proposition 5.4.5. (Hint: prove each property separately, and use argument by contradiction.)

We often talk about compact sets rather than compact metric spaces, so we make the following definition.

Definition 5.4.7 (Compactness of a Set). Let (X, d) be a metric space, and let Y be a subset of X. We say that Y is **compact** if and only if the metric space $(Y, d|_{Y\times Y})$ is compact.

Corollary 5.4.8. *Let (X, d) be a metric space, and let Y be a compact subset of X. Then Y is closed and bounded.*

Proof. Apply Proposition 5.4.5 and then Proposition 5.3.8. □

In Euclidean space, the converse of Corollary 5.4.8 is true. The following Theorem therefore gives a useful characterization of compact subsets of Euclidean space.

Theorem 5.4.9. *Let n be a positive integer. Let $(\mathbb{R}^n, d)$ denote Euclidean space with the metric $d = d_{\ell_2}$ or $d = d_{\ell_1}$. Let E be a subset of $\mathbb{R}^n$. Then E is compact if and only if E is both closed and bounded.*

Exercise 5.4.10. Prove Theorem 5.4.9 using Corollary 5.4.3.

Compact sets of metric spaces can be equivalently characterized using open covers. This open cover property will become useful in our discussion of continuous functions. The property says: any (possibly uncountable) open cover of a compact set has a finite subcover. The following proof is a bit lengthy, so it can be skipped on a first reading.

Theorem 5.4.11 (Open Cover Characterization of Compactness). *Let (X, d) be a metric space and let Y be a compact subset of X. Let I be an index set. Let $\{V_\alpha\}_{\alpha\in I}$ be a collection of open sets in X. Assume that*

$$Y \subseteq \bigcup_{\alpha\in I} V_\alpha.$$

*(That is, the collection $\{V_\alpha\}_{\alpha\in I}$ **covers** Y.) Then, there exists a finite set $A \subseteq I$ such that*

$$Y \subseteq \bigcup_{\alpha\in A} V_\alpha.$$

Proof. Let $y \in Y$. Then there exists $\alpha \in I$ such that $y \in V_\alpha$. Since V_α is open, there exists $r > 0$ such that $B(y, r) \subseteq V_\alpha$. For each $y \in Y$, define $r(y) \in \mathbb{R}$ by

$$r(y) := \sup\{r \in (0, \infty) \colon \exists\, \alpha \in I \,\text{such that}\, B(y, r) \subseteq V_\alpha\}.$$

We showed that for every $y \in Y$, we have $r(y) > 0$. Define $r_0 \in \mathbb{R}$ by

$$r_0 := \inf\{r(y) \colon y \in Y\}.$$

Since $r(y) > 0$ for all $y \in Y$, we have $r_0 \geq 0$. We now consider the cases $r_0 = 0$ and $r_0 > 0$ separately.

Case 1. $r_0 = 0$. In this case, for every positive integer j, there exists $y \in Y$ such that $r(y) < 1/j$. So, for every positive integer j, let $y^{(j)} \in Y$ satisfy $r(y^{(j)}) < 1/j$. (We can do this by the countable axiom of choice.) By the Squeeze Theorem, $\lim_{j\to\infty} r(y^{(j)}) = 0$. Since $(y^{(j)})_{j=1}^{\infty}$ is a sequence in Y, and since Y is compact, there exists a subsequence $(y^{(j_k)})_{k=1}^{\infty}$ that converges to some point $y_0 \in Y$.

Since $y_0 \in Y$, we know as above that there exists $\alpha_0 \in I$ such that $y_0 \in V_{\alpha_0}$. And since V_{α_0} is open, there exists $\varepsilon > 0$ such that $B(y_0, \varepsilon) \subseteq V_{\alpha_0}$. Since $y^{(j_k)}$ converges to y_0 as $k \to \infty$, there exists a positive integer K such that, for all $k > K$, we have $y^{(j_k)} \in B(y_0, \varepsilon/2)$. By the triangle inequality, if $k > K$ and if $z \in B(y^{(j_k)}, \varepsilon/2)$, then $d(z, y_0) \leq d(z, y^{(j_k)}) + d(y^{(j_k)}, y_0) < \varepsilon/2 + \varepsilon/2 = \varepsilon$. That is, if $k > K$, then $B(y^{(j_k)}, \varepsilon/2) \subseteq B(y_0, \varepsilon)$. Since $B(y_0, \varepsilon) \subseteq V_{\alpha_0}$, we conclude that $r(y^{(j_k)}) \geq \varepsilon/2$ for all $k > K$. The last condition implies that $\lim_{k\to\infty} r(y^{(j_k)}) \neq 0$, a contradiction. We conclude that Case 1 does not occur, i.e. we must have $r_0 > 0$.

Case 2. $r_0 > 0$. In this case, for every $y \in Y$, we have $r(y) > r_0/2$. So, for all $y \in Y$, there exists $\alpha = \alpha(y) \in I$ such that $B(y, r_0/2) \in V_\alpha$. We now argue by contradiction. Suppose there does not exist a finite collection $\{V_\alpha\}_{\alpha \in A}$ that covers Y. Let $y^{(1)}$ be any point in Y. We construct a sequence of points in Y recursively. Suppose we are given $y^{(1)}, \ldots, y^{(j)}$ a sequence of points in Y. Given these points, the union $B(y^{(1)}, r_0/2) \cup \cdots \cup B(y^{(j)}, r_0/2)$ is contained in the union $V_{\alpha(y^{(1)})} \cup \cdots \cup V_{\alpha(y^{(j)})}$ for some $\alpha(y^{(1)}), \ldots, \alpha(y^{(j)}) \in I$. By our contradictory assumption, the latter set does not cover Y, so the set $B(y^{(1)}, r_0/2) \cup \cdots \cup B(y^{(j)}, r_0/2)$ does not cover Y. That is, there exists some $y^{(j+1)} \in Y$ such that $y^{(j+1)} \notin B(y^{(i)}, r_0/2)$ for all $1 \leq i \leq j$. That is, $d(y^{(j+1)}, y^{(i)}) \geq r_0/2$ for all $1 \leq i \leq j$. From the latter property, the sequence $(y^{(j)})_{j=1}^{\infty}$ is a sequence that has no convergent subsequence. (If a subsequence $(y^{(j_k)})_{k=1}^{\infty}$ converged to some $z \in Y$, then there would exist a positive integer K such that, for all $k > K$, we would have $d(y^{(j_k)}, z) < r_0/4$, so that $d(y^{(j_{K+2})}, y^{(j_{K+1})}) \leq d(y^{(j_{K+2})}, z) + d(z, y^{(j_{K+1})}) < r_0/2$, a contradiction.) We have therefore contradicted the compactness of Y. Since we have achieved a contradiction, the proof is done. □

Remark 5.4.12. The converse is also true. If a set Y has the property that every open cover of Y has a finite subcover, then Y is compact.

Theorem 5.4.11 has the following useful corollary.

Corollary 5.4.13. *Let (X, d) be a metric space, and let $K_1, K_2, \ldots$ be a sequence of nonempty compact subsets of X such that*

$$K_1 \supseteq K_2 \supseteq K_3 \supseteq \cdots .$$

Then the intersection $\cap_{j=1}^{\infty} K_j$ is nonempty.

Exercise 5.4.14. Prove Corollary 5.4.13. (Hint: first, work in the compact metric space $(K_1, d|_{K_1 \times K_1})$. Then, consider the sets $K_1 \smallsetminus K_j$ which are open in K_1. Assume for the sake of contradiction that $\cap_{j=1}^{\infty} K_j = \emptyset$. Then apply Theorem 5.4.11.)

Theorem 5.4.15. *Let (X, d) be a metric space.*

(i) *Let Y be a compact subset of X, and let Z be a subset of Y. Then Z is compact if and only if Z is closed.*
(ii) *Let $Y_1, \ldots, Y_n$ be compact subsets of X. Then $Y_1 \cup \cdots \cup Y_n$ is compact.*

(iii) *Every finite subset of X is compact.*

5.5. **Continuity.** We can readily generalize the notion of continuity of a function $f\colon \mathbb{R} \to \mathbb{R}$ to the setting of a function between metric spaces $f\colon X \to Y$. We just take the usual definition and then we replace the absolute values with the required metric, as follows.

Definition 5.5.1 (Continuity). Let (X, d_X) and let (Y, d_Y) be metric spaces. Let $f\colon X \to Y$ be a function. Let $x_0 \in X$. We say that f is **continuous** at x_0 if and only if, for all $\varepsilon > 0$, there exists $\delta = \delta(\varepsilon) > 0$ such that, if $x \in X$ satisfies $d_X(x, x_0) < \delta$, then $d_Y(f(x), f(x_0)) < \varepsilon$. We say that the function f is **continuous** if and only if it is continuous at every point in X.

Remark 5.5.2. Suppose $f\colon X \to Y$ is continuous and K is a subset of X. Then the restriction of f to K, $f|_K\colon K \to Y$ is also continuous.

As on the real line, continuous functions maps convergent sequences to convergent sequences.

Theorem 5.5.3. *Let (X, d_X) and let (Y, d_Y) be metric spaces. Let $f\colon X \to Y$ be a function. Then the following two statements are equivalent.*

- *f is continuous at x_0.*
- *If we have a sequence $(x^{(j)})_{j=1}^{\infty}$ in X which converges to x_0 with respect to d_X, then the sequence $(f(x^{(j)}))_{j=1}^{\infty}$ converges to $f(x_0)$ with respect to the metric d_Y.*

Exercise 5.5.4. Prove Theorem 5.5.3.

In fact, there is even a way to characterize continuous functions using the inverse images of open and closed sets.

Theorem 5.5.5. *Let (X, d_X) and let (Y, d_Y) be metric spaces. Let $f\colon X \to Y$ be a function. Then the following four statements are equivalent.*

- *f is continuous at x_0, for all $x_0 \in X$.*
- *For all $x_0 \in X$, if we have a sequence $(x^{(j)})_{j=1}^{\infty}$ in X which converges to x_0 with respect to d_X, then the sequence $(f(x^{(j)}))_{j=1}^{\infty}$ converges to $f(x_0)$ with respect to the metric d_Y.*
- *For all open sets W in Y, the set $f^{-1}(W) = \{x \in X : f(x) \in W\}$ is an open set in X.*
- *For all closed sets V in Y, the set $f^{-1}(V)$ is a closed set in X.*

Exercise 5.5.6. Prove Theorem 5.5.5.

Remark 5.5.7. For a continuous function, it is not always true that the image of an open set is open, and it is not always true that the image of a closed set is closed.

We can now quickly show that the composition of continuous functions is continuous.

Corollary 5.5.8. *Let (X, d_X), (Y, d_Y) and (Z, d_Z) be metric spaces. Let $f\colon (X, d_X) \to (Y, d_Y)$ be a continuous function and let $g\colon (Y, d_Y) \to (Z, d_Z)$ be a continuous function. Then $g \circ f\colon (X, d_X) \to (Z, d_Z)$ is a continuous function.*

Exercise 5.5.9. Prove Corollary 5.5.8.

5.6. **Continuity and Compactness.**

Remark 5.6.1. From now on, unless otherwise specified, $\mathbb{R}^n$ refers to Euclidean space $\mathbb{R}^n$ with $n \geq 1$ a positive integer, and where we use the metric d_{ℓ_2} on $\mathbb{R}^n$. In particular, $\mathbb{R}$ refers to the metric space $\mathbb{R}$ equipped with the metric $d(x,y) = |x-y|$.

On the real line, we learned from the Extreme Value Theorem that the continuous image of a closed interval is another closed interval. The appropriate generalization of this statement to metric spaces now follows.

Theorem 5.6.2. *Let (X, d_X) and (Y, d_Y) be metric spaces. Let $f\colon (X, d_X) \to (Y, d_Y)$ be a continuous function. Suppose $K \subseteq X$ is a compact set. Then $f(K) = \{f(x)\colon x \in K\}$ is also a compact set.*

Exercise 5.6.3. Prove Theorem 5.6.2

Combining this Theorem with the characterization of compactness in Euclidean spaces (i.e. Heine-Borel, Theorem 5.4.9), we get the following statement.

Corollary 5.6.4. *Let K be a closed and bounded subset of $\mathbb{R}^n$. Let $f\colon K \to \mathbb{R}^m$ be a continuous function. Then the set $f(K)$ is also closed and bounded. In particular, the function f is bounded on K.*

This Corollary allows us to state our generalization of the Extreme Value Theorem, which we now refer to as the Maximum Principle.

Definition 5.6.5. Let X be a set. Let $f\colon X \to \mathbb{R}$ be a function. We say that f **attains its maximum** at $x_0 \in X$ if and only if $f(x_0) \geq f(x)$ for all $x \in X$. We say that f **attains its minimum** at $x_0 \in X$ if and only if $f(x_0) \leq f(x)$ for all $x \in X$.

Theorem 5.6.6 (The Maximum Principle). *Let K be a closed and bounded subset of $\mathbb{R}^n$, and let $f\colon K \to \mathbb{R}$ be a continuous function. Then there exist points $a, b \in K$ such that f attains its maximum at a and f attains its minimum at b.*

Exercise 5.6.7. Prove Theorem 5.6.6. (Hint: use Corollary 5.6.4 and then consider the numbers $\sup_{x \in K} f(x)$ and $\inf_{x \in K} f(x)$.)

5.7. **Continuity and Connectedness.** Recall that the Intermediate Value Theorem says that a continuous function on an interval has an interval as its range. The appropriate generalization of this statement to metric spaces involves the concept of connectedness.

Definition 5.7.1 (Connectedness). Let (X, d) be a metric space. We say that X is **disconnected** if and only if there exist disjoint nonempty open sets V, W in X such that $V \cup W = X$. (Equivalently, X is disconnected if and only if X contains a proper non-empty subset which is both open and closed.) We say that X is **connected** if and only if X is not disconnected.

Example 5.7.2. The set $X = [0,1] \cup [2,3]$ with the metric $d(x,y) = |x-y|$ is disconnected, since the sets $[0,1]$ and $[2,3]$ are both open in X.

Definition 5.7.3. Let (X, d) be a metric space and let Y be a subset of X. We say that Y is **connected** if and only if the metric space $(Y, d|_{Y \times Y})$ is connected. We say that Y is **disconnected** if and only if the metric space $(Y, d|_{Y \times Y})$ is disconnected.

For the sake of examples, we now identify the connected subsets of the real line.

Theorem 5.7.4. *Let X be a subset of the real line $\mathbb{R}$. Then the following statements are equivalent.*

- *X is connected.*
- *For any $x, y \in X$ with $x < y$, the closed interval $[x, y]$ is also contained in X.*

Proof. We first show the forward implication. Suppose X is connected. We argue by contradiction. Let $x, y \in X$ with $x < y$ such that $[x, y]$ is not contained in X. Then there exists $x < z < y$ such that $z \notin X$. Then the sets $(-\infty, z) \cap X$ and $(z, \infty) \cap X$ are both disjoint, nonempty, relatively open sets whose union in X. Therefore, X is disconnected, a contradiction. We conclude that the forward implication holds.

We now prove the more involved reverse implication. Suppose for any $x, y \in X$ with $x < y$, the closed interval $[x, y]$ is also contained in X. We need to show that X is connected. We argue by contradiction. Suppose that X is disconnected. Then there exist two disjoint, nonempty, relatively open sets V, W such that $V \cup W = X$. Since V, W are nonempty, let $v \in V$ and let $w \in W$. Without loss of generality, $v < w$. By assumption, the closed interval $[v, w]$ is contained in X. Consider the real number

$$x = \sup([v, w] \cap V).$$

By the definition of x, we have $x \in [v, w]$. We will derive a contradiction by trying to determine whether or not $x \in V$.

Suppose $x \in V$. Since $w \notin V$, we have $x \neq w$, so $x \in [v, w)$. Since V is relatively open in X, there exists $\varepsilon > 0$ such that $(x - \varepsilon, x + \varepsilon) \cap X \subseteq V$. Since $x \in [v, w) \subseteq X$ as well, there exists $\delta > 0$ such that $[x, x + \delta) \subseteq V$. But then x is not the least upper bound of $[v, w] \cap V$, a contradiction.

We must therefore have $x \notin V$. Since $x \in [v, w] \subseteq X$, and since V, W are disjoint, we must have $x \in W$. Since $v \in V$, we have $x \neq v$, so $x \in (v, w]$. Since W is relatively open in X, there exists $\varepsilon > 0$ such that $(x - \varepsilon, x + \varepsilon) \cap X \subseteq W$. Since $x \in (v, w] \subseteq X$ as well, there exists $\delta > 0$ such that $(x - \delta, x] \subseteq W$. Since V and W are disjoint, we once again conclude that x is not the least upper bound of $[v, w] \cap V$. In any case, we have achieved a contradiction. We finally conclude that X is connected, as desired. □

Remark 5.7.5. So, $\mathbb{R}$ is connected, and so are the intervals $(a, b]$, $[a, b)$, (a, b), $[a, b]$, $(-\infty, b)$, $(-\infty, b]$, (a, ∞), $[a, \infty)$. Additionally, the empty set $\emptyset$ and singleton sets $\{a\}$ are connected. We therefore have a complete list of connected subsets of $\mathbb{R}$.

It turns out that connected sets are mapped to connected sets by continuous functions. This fact particularly implies the Intermediate Value Theorem.

Theorem 5.7.6. *Let (X, d_X) and (Y, d_Y) be metric spaces. Let $f \colon X \to Y$ be a continuous function. Let E be a connected subset of X. Then $f(E)$ is connected.*

Exercise 5.7.7. Prove Theorem 5.7.6.

Theorem 5.7.8 (Intermediate Value Theorem). *Let (X, d) be a metric space. Let $f \colon X \to \mathbb{R}$ be a continuous function. Let E be a connected subset of X and let a, b be any two elements of E. Let y be a real number between $f(a)$ and $f(b)$, so that either $f(a) \leq y \leq f(b)$ or $f(b) \leq y \leq f(a)$. Then there exists $c \in E$ such that $f(c) = y$.*

Exercise 5.7.9. Prove Theorem 5.7.8 using Theorem 5.7.6.

6. Sequences and Series of Functions, Convergence

6.1. **Sequences of Functions.** As we have seen in analysis, it is often desirable to discuss sequences of points that converge. Below, we will see that it is similarly desirable to discuss sequences of functions that converge in various senses. There are many distinct ways of discussing the convergence of sequences of functions. We will only discuss two such modes of convergence, namely pointwise and uniform convergence. Before beginning this discussion, we discuss the limiting values of functions between metric spaces, which should generalize our notion of limiting values of functions on the real line.

6.1.1. *Limiting Values of Functions.*

Definition 6.1.1. Let (X, d_X) and (Y, d_Y) be metric spaces, let E be a subset of X, let $f\colon X \to Y$ be a function, let $x_0 \in X$ be an adherent point of E, and let $L \in Y$. We say that $f(x)$ **converges to** L **in** Y **as** x **converges to** x_0 **in** E, and we write $\lim_{x\to x_0; x\in E} f(x) = L$, if and only if, for every $\varepsilon > 0$, there exists $\delta = \delta(\varepsilon) > 0$ such that, if $x \in E$ satisfies $d_X(x, x_0) < \delta$, then $d_Y(f(x), L) < \varepsilon$.

Remark 6.1.2. So, f is continuous at x_0 if and only if

$$\lim_{x\to x_0; x\in X} f(x) = f(x_0). \qquad (*)$$

And f is continuous on X if and only if, for all $x_0 \in X$, $(*)$ holds.

Remark 6.1.3. When the domain of x of the limit $\lim_{x\to x_0; x\in X} f(x)$ is clear, we will often instead write $\lim_{x\to x_0} f(x)$.

The following equivalence is generalized from its analogue on the real line.

Proposition 6.1.4. *Let (X, d_X) and (Y, d_Y) be metric spaces, let E be a subset of X, let $f\colon X \to Y$ be a function, let $x_0 \in X$ be an adherent point of E, and let $L \in Y$. Then the following statements are equivalent.*

- $\lim_{x\to x_0; x\in E} f(x) = L$.
- *For any sequence $(x^{(j)})_{j=1}^\infty$ in E which converges to x_0 with respect to the metric d_X, the sequence $(f(x^{(j)}))_{j=1}^\infty$ converges to L with respect to the metric d_Y.*

Exercise 6.1.5. Prove Proposition 6.1.4.

Remark 6.1.6. From Propositions 6.1.4 and 5.1.24, the function f can converge to at most one limit L as x converges to x_0.

Remark 6.1.7. The notation $\lim_{x\to x_0; x\in E} f(x)$ implicitly refers to a convergence of the function values $f(x)$ in the metric space (Y, d_Y). Strictly speaking, it would be better to write d_Y somewhere next to the notation $\lim_{x\to x_0; x\in E} f(x)$. However, this omission of notation should not cause confusion.

6.1.2. *Pointwise Convergence and Uniform Convergence.*

Definition 6.1.8 (**Pointwise Convergence**). Let (X, d_X) and (Y, d_Y) be metric spaces. Let $(f_j)_{j=1}^\infty$ be a sequence of functions from X to Y. Let $f\colon X \to Y$ be another function. We say that $(f_j)_{j=1}^\infty$ **converges pointwise** to f on X if and only if, for every $x \in X$, we have

$$\lim_{j\to\infty} f_j(x) = f(x).$$

That is, for all $x \in X$, we have

$$\lim_{j\to\infty} d_Y(f_j(x), f(x)) = 0.$$

That is, for every $x \in X$ and for every $\varepsilon > 0$, there exists $J > 0$ such that, for all $j > J$, we have $d_Y(f_j(x), f(x)) < \varepsilon$.

Remark 6.1.9. Note that, if we change the point x, then the limiting behavior of $f_j(x)$ can change quite a bit. For example, let j be a positive integer, and consider the functions $f_j\colon [0,1] \to \mathbb{R}$ where $f_j(x) = j$ for all $x \in (0, 1/j)$, and $f_j(x) = 0$ otherwise. Let $f\colon [0,1] \to \mathbb{R}$ be the zero function. Then f_j converges pointwise to zero, since for any $x \in (0, 1]$, we have $f_j(x) = 0$ for all $j > 1/x$. (And $f_j(0) = 0$ for all positive integers j.) However, given any fixed positive integer j, there exists an x such that $f_j(x) = j$. Moreover, $\int_0^1 f_j = 1$ for all positive integers j, but $\int_0^1 f = 0$. So, we see that pointwise convergence does not preserve the integral of a function.

Remark 6.1.10. Pointwise convergence also does not preserve continuity. For example, consider $f_j\colon [0,1] \to \mathbb{R}$ defined by $f_j(x) = x^j$, where $j \in \mathbb{N}$ and $x \in [0,1]$. Define $f\colon [0,1] \to \mathbb{R}$ so that $f(1) = 1$ and so that $f(x) = 0$ for $x \in [0,1)$. Then f_j converges pointwise to f as $j \to \infty$, and each f_j is continuous, but f is not continuous.

In summary, pointwise convergence doesn't really preserve any useful analytic quantities. The above remarks show that some points are changing at much different rates than other points as $j \to \infty$. A stronger notion of convergence will then fix these issues, where all points in the domain are controlled simultaneously.

Definition 6.1.11 (Uniform Convergence). Let (X, d_X) and (Y, d_Y) be metric spaces. Let $(f_j)_{j=1}^\infty$ be a sequence of functions from X to Y. Let $f\colon X \to Y$ be another function. We say that $(f_j)_{j=1}^\infty$ **converges uniformly** to f on X if and only if, for every $\varepsilon > 0$, there exists $J > 0$ such that, for all $j > J$ and for all $x \in X$ we have $d_Y(f_j(x), f(x)) < \varepsilon$.

Remark 6.1.12. Note that the difference between uniform and pointwise convergence is that we simply moved the quantifier "for all $x \in X$" within the statement. This change means that the integer J does not depend on x in the case of uniform convergence.

Remark 6.1.13. The sequences of functions from Remarks 6.1.9 and 6.1.10 do not converge uniformly. So, pointwise convergence does not imply uniform convergence. However, uniform convergence does imply pointwise convergence.

6.2. **Uniform Convergence and Continuity.** We saw that pointwise convergence does not preserve continuity. However, uniform convergence does preserve continuity.

Theorem 6.2.1. *Let (X, d_X) and (Y, d_Y) be metric spaces. Let $(f_j)_{j=1}^\infty$ be a sequence of functions from X to Y. Let $f\colon X \to Y$ be another function. Let $x_0 \in X$. Suppose f_j converges uniformly to f on X. Suppose that, for each $j \geq 1$, we know that f_j is continuous at x_0. Then f is also continuous at x_0.*

Exercise 6.2.2. Prove Theorem 6.2.1. Hint: it is probably easiest to use the $\varepsilon - \delta$ definition of continuity. Once you do this, you may require the triangle inequality in the form

$$d_Y(f(x), f(x_0)) \leq d_Y(f(x), f_j(x)) + d_Y(f_j(x), f_j(x_0)) + d_Y(f_j(x_0), f(x_0)).$$

Corollary 6.2.3. *Let (X, d_X) and (Y, d_Y) be metric spaces. Let $(f_j)_{j=1}^\infty$ be a sequence of functions from X to Y. Let $f\colon X \to Y$ be another function. Suppose $(f_j)_{j=1}^\infty$ converges uniformly to f on X. Suppose that, for each $j \geq 1$, we know that f_j is continuous on X. Then f is also continuous on X.*

Uniform limits of bounded functions are also bounded. Recall that a function $f\colon X \to Y$ between metric spaces (X, d_X) and (Y, d_Y) is **bounded** if and only if there exists a radius $R > 0$ and a point $y_0 \in Y$ such that $f(x) \in B_{(Y,d_Y)}(y_0, R)$ for all $x \in X$.

Proposition 6.2.4. *Let (X, d_X) and (Y, d_Y) be metric spaces. Let $(f_j)_{j=1}^\infty$ be a sequence of functions from X to Y. Let $f\colon X \to Y$ be another function. Suppose $(f_j)_{j=1}^\infty$ converges uniformly to f on X. Suppose also that, for each $j \geq 1$, we know that f_j is bounded. Then f is also bounded.*

Exercise 6.2.5. Prove Proposition 6.2.4.

6.2.1. *The Metric of Uniform Convergence.* We will now see one advantage to our abstract approach to analysis on metric spaces. We can in fact talk about uniform convergence in terms of a metric on a space of functions, as follows.

Definition 6.2.6. Let (X, d_X) and (Y, d_Y) be metric spaces. Let $B(X;Y)$ denote the set of functions $f\colon X \to Y$ that are bounded. Let $f, g \in B(X;Y)$. We define the metric $d_\infty\colon B(X;Y) \times B(X;Y) \to [0,\infty)$ by

$$d_\infty(f,g) := \sup_{x\in X} d_Y(f(x), g(x)).$$

This metric is known as the **sup norm metric** or the L_∞ **metric**. We also use $d_{B(X;Y)}$ as a synonym for d_∞. Note that $d_\infty(f,g) < \infty$ since f, g are assumed to be bounded.

Exercise 6.2.7. Show that the space $(B(X;Y), d_\infty)$ is a metric space.

Example 6.2.8. Let $X = [0,1]$ and let $Y = \mathbb{R}$. Consider the functions $f(x) = x$ and $g(x) = 2x$ where $x \in [0,1]$. Then f, g are bounded, and

$$d_\infty(f,g) = \sup_{x\in[0,1]} |x - 2x| = \sup_{x\in[0,1]} |x| = 1.$$

Here is our promised characterization of uniform convergence in terms of the metric d_∞.

Proposition 6.2.9. *Let (X, d_X) and (Y, d_Y) be metric spaces. Let $(f_j)_{j=1}^\infty$ be a sequence of functions in $B(X;Y)$. Let $f \in B(X;Y)$. Then $(f_j)_{j=1}^\infty$ converges uniformly to f on X if and only if $(f_j)_{j=1}^\infty$ converges to f in the metric $d_{B(X;Y)}$.*

Exercise 6.2.10. Prove Proposition 6.2.9.

Definition 6.2.11. Let (X, d_X) and (Y, d_Y) be metric spaces. Define the set of bounded continuous functions from X to Y as

$$C(X;Y) := \{f \in B(X;Y)\colon f \text{ is continuous}\}.$$

Note that $C(X;Y) \subseteq B(X;Y)$ by the definition of $C(X;Y)$. Also, by Corollary 6.2.3, $C(X;Y)$ is closed in $B(X;Y)$ with respect to the metric d_∞. In fact, more is true.

Theorem 6.2.12. *Let (X, d_X) be a metric space, and let (Y, d_Y) be a complete metric space. Then the space $(C(X;Y), d_{B(X;Y)}|_{C(X;Y)\times C(X;Y)})$ is a complete subspace of $B(X;Y)$. That is, every Cauchy sequence of functions in $C(X;Y)$ converges to a function in $C(X;Y)$.*

Exercise 6.2.13. Prove Theorem 6.2.12

6.3. **Series of Functions and the Weierstrass M-test.** For each positive integer j, let $f_j\colon X \to \mathbb{R}$ be a function. We will now consider infinite series of the form $\sum_{j=1}^{\infty} f_j$. The most natural thing to do now is to determine in what sense the series $\sum_{j=1}^{\infty} f_j$ is a function, and if it is a function, determine if it is continuous. Note that we have restricted the range to be $\mathbb{R}$ since it does not make sense to add elements in a general metric space. Power series and Fourier series perhaps give the most studied examples of series of functions. If $x \in [0,1]$ and if a_j are real numbers for all $j \geq 1$, we want to make sense of the series $\sum_{j=1}^{\infty} a_j \cos(2\pi j x)$. We want to know in what sense this infinite series is a function, and if it is a function, do the partial sums converge in any reasonable manner? We will return to these issues later on.

Definition 6.3.1. Let (X, d_X) be a metric space. For each positive integer j, let $f_j\colon X \to \mathbb{R}$ be a function, and let $f\colon X \to \mathbb{R}$ be another function. If the partial sums $\sum_{j=1}^{J} f_j$ converge pointwise to f as $J \to \infty$, then we say that the infinite series $\sum_{j=1}^{\infty} f_j$ **converge pointwise** to f, and we write $f = \sum_{j=1}^{\infty} f_j$. If the partial sums $\sum_{j=1}^{J} f_j$ converge uniformly to f as $J \to \infty$, then we say that the infinite series $\sum_{j=1}^{\infty} f_j$ **converge uniformly** to f, and we write $f = \sum_{j=1}^{\infty} f_j$. (In particular, the notation $f = \sum_{j=1}^{\infty} f_j$ is ambiguous, since the nature of the convergence of the series is not specified.)

Remark 6.3.2. If a series converges uniformly then it converges pointwise. However, the converse is false in general.

Exercise 6.3.3. Let $x \in (-1,1)$. For each integer $j \geq 1$, define $f_j(x) := x^j$. Show that the series $\sum_{j=1}^{\infty} f_j$ converges pointwise, but not uniformly, on $(-1,1)$ to the function $f(x) = x/(1-x)$. Also, for any $0 < t < 1$, show that the series $\sum_{j=1}^{\infty} f_j$ converges uniformly to f on $[-t,t]$.

Definition 6.3.4. Let $f\colon X \to \mathbb{R}$ be a bounded real-valued function. We define the **sup-norm** $\|f\|_{\infty}$ of f to be the real number

$$\|f\|_{\infty} := \sup_{x \in X} |f(x)|.$$

Exercise 6.3.5. Let X be a set. Show that $\|\cdot\|_{\infty}$ is a norm on the space $B(X;\mathbb{R})$.

Theorem 6.3.6 (Weierstrass M-test). *Let (X,d) be a metric space and let $(f_j)_{j=1}^{\infty}$ be a sequence of bounded real-valued continuous functions on X such that the series (of real numbers) $\sum_{j=1}^{\infty} \|f_j\|_{\infty}$ is absolutely convergent. Then the series $\sum_{j=1}^{\infty} f_j$ converges uniformly to some continuous function $f\colon X \to \mathbb{R}$.*

Exercise 6.3.7. Prove Theorem 6.3.6. (Hint: first, show that the partial sums $\sum_{j=1}^{J} f_j$ form a Cauchy sequence in $C(X;\mathbb{R})$. Then, use Theorem 6.2.12 and the completeness of the real line $\mathbb{R}$.)

Remark 6.3.8. The Weierstrass M-test will be useful in our investigation of power series.

6.4. **Uniform Convergence and Integration.**

Theorem 6.4.1. *Let $a < b$ be real numbers. For each integer $j \geq 1$, let $f_j\colon [a,b] \to \mathbb{R}$ be a Riemann integrable function on $[a,b]$. Suppose f_j converges uniformly on $[a,b]$ to a function $f\colon [a,b] \to \mathbb{R}$, as $j \to \infty$. Then f is also Riemann integrable, and*

$$\lim_{j \to \infty} \int_a^b f_j = \int_a^b f.$$

Remark 6.4.2. Before we begin, recall that we require any Riemann integrable function g to be bounded. Also, for a Riemann integrable function g, we denote $\underline{\int_a^b} g$ as the supremum of all lower Riemann sums of g over all partitions of $[a, b]$. And we denote $\overline{\int_a^b} g$ as the infimum of all upper Riemann sums of g over all partitions of $[a, b]$. Recall also that a function g is defined to be Riemann integrable if and only if $\underline{\int_a^b} g = \overline{\int_a^b} g$.

Proof. We first show that f is Riemann integrable. First, note that f_j is bounded for all $j \geq 1$, since that is part of the definition of being Riemann integrable. So, f is bounded by Proposition 6.2.4. Now, let $\varepsilon > 0$. Since f_j converges uniformly to f on $[a, b]$, there exists $J > 0$ such that, for all $j > J$, we have

$$f_j(x) - \varepsilon \leq f(x) \leq f_j(x) + \varepsilon, \qquad \forall\, x \in [a, b].$$

Integrating this inequality on $[a, b]$, we have

$$\underline{\int_a^b} (f_j(x) - \varepsilon) \leq \underline{\int_a^b} f \leq \overline{\int_a^b} f \leq \overline{\int_a^b} (f_j(x) + \varepsilon).$$

Since f_j is Riemann integrable for all $j \geq 1$, we therefore have

$$-(b-a)\varepsilon + \int_a^b f_j \leq \underline{\int_a^b} f \leq \overline{\int_a^b} f \leq (b-a)\varepsilon + \int_a^b f_j. \qquad (*)$$

In particular, we get

$$0 \leq \overline{\int_a^b} f - \underline{\int_a^b} f \leq 2(b-a)\varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, we conclude that $\overline{\int_a^b} f = \underline{\int_a^b} f$, so f is Riemann integrable.

Now, from $(*)$, we have: for any $\varepsilon > 0$, there exists J such that, for all $j > J$, we have

$$\left| \int_a^b f - \int_a^b f_j \right| \leq (b-a)\varepsilon.$$

Since this holds for any $\varepsilon > 0$, we conclude that $\lim_{j\to\infty} \int_a^b f_j = \int_a^b f$, as desired. □

Remark 6.4.3. In summary, if a sequence of Riemann integrable functions $(f_j)_{j=1}^{\infty}$ converges to f uniformly, then we can interchange limits and integrals

$$\lim_{j\to\infty} \int f_j = \int \lim_{j\to\infty} f_j.$$

Recall that this equality does not hold if we only assume that the functions converge pointwise.

An analogous statement holds for series.

Theorem 6.4.4. *Let $a < b$ be real numbers. For each integer $j \geq 1$, let $f_j \colon [a, b] \to \mathbb{R}$ be a Riemann integrable function on $[a, b]$. Suppose $\sum_{j=1}^{\infty} f_j$ converges uniformly on $[a, b]$. Then $\sum_{j=1}^{\infty} f_j$ is also Riemann integrable, and*

$$\sum_{j=1}^{\infty} \int_a^b f_j = \int_a^b \sum_{j=1}^{\infty} f_j.$$

Exercise 6.4.5. Prove Theorem 6.4.4.

Example 6.4.6. Let $x \in (-1,1)$. We know that $\sum_{j=1}^{\infty} x^j = x/(1-x)$, and the convergence is uniform on $[-r,r]$ for any $0<r<1$. Adding 1 to both sides, we get

$$\sum_{j=0}^{\infty} x^j = \frac{1}{1-x}.$$

And this sum also converges uniformly on $[-r,r]$ for any $0<r<1$. Applying Theorem 6.4.4 and integrating on $[0,r]$, we get

$$\sum_{j=0}^{\infty} \frac{r^{j+1}}{j+1} = \sum_{j=0}^{\infty} \int_0^r x^j = \int_0^r \frac{1}{1-x}.$$

The last function is equal to $-\log(1-r)$, though we technically have not defined the logarithm function yet. We will define the logarithm further below.

6.5. **Uniform Convergence and Differentiation.** We now investigate the relation between uniform convergence and differentiation.

Remark 6.5.1. Suppose a sequence of differentiable functions $(f_j)_{j=1}^{\infty}$ converges uniformly to a function f. We first show that f need not be differentiable. Consider the functions $f_j(x) := \sqrt{x^2+1/j}$, where $x \in [-1,1]$. Let $f(x) = |x|$. Note that

$$|x| \leq \sqrt{x^2+1/j} \leq |x| + 1/\sqrt{j}.$$

These inequalities follow by taking the square root of $x^2 \leq x^2+1/j \leq x^2+1/j+2|x|/\sqrt{j}$. So, by the Squeeze Theorem, $(f_j)_{j=1}^{\infty}$ converges uniformly to f on $[-1,1]$. However, f is not differentiable at 0. In conclusion, uniform convergence does not preserve differentiability.

Remark 6.5.2. Suppose a sequence of differentiable functions $(f_j)_{j=1}^{\infty}$ converge uniformly to a function f. Even if f is assumed to be differentiable, we show that $(f_j)'$ may not converge to f'. Consider the functions $f_j(x) := j^{-1/2}\sin(j\pi x)$, where $x \in [-1,1]$. (We will assume some basic properties of trigonometric functions which we will prove later on. Since we are only providing a motivating example, we will not introduce any circular reasoning.) Let f be the zero function. Since $|\sin(j\pi x)| \leq 1$, we have $d_\infty(f_j,f) \leq j^{-1/2}$, so $(f_j)_{j=1}^{\infty}$ converges uniformly on $[-1,1]$. However, $f_j'(x) = j^{1/2}\pi\cos(j\pi x)$. So, $f_j'(0) = j^{1/2}\pi$. That is, $(f_j')_{j=1}^{\infty}$ does not converge pointwise to f. So, $(f_j')_{j=1}^{\infty}$ does not converge uniformly to $f'=0$. In conclusion, uniform convergence does not imply uniform convergence of derivatives.

However, the converse statement is true, as long as the sequence of functions converges at one point.

Theorem 6.5.3. *Let $a<b$. For every integer $j \geq 1$, let $f_j\colon [a,b] \to \mathbb{R}$ be a differentiable function whose derivative $(f_j)'\colon [a,b] \to \mathbb{R}$ is continuous. Assume that the derivatives $(f_j)'$ converge uniformly to a function $g\colon [a,b] \to \mathbb{R}$ as $j \to \infty$. Assume also that there exists a point $x_0 \in [a,b]$ such that $\lim_{j\to\infty} f_j(x_0)$ exists. Then the functions f_j converge uniformly to a differentiable function f as $j \to \infty$, and $f' = g$.*

Proof. Let $x \in [a,b]$. From the Fundamental Theorem of Calculus, for each $j \geq 1$,

$$f_j(x) - f_j(x_0) = \int_{x_0}^{x} f_j'. \qquad (*)$$

By assumption, $L := \lim_{j\to\infty} f_j(x_0)$ exists. From Theorem 6.2.1, g is continuous, and in particular, g is Riemann integrable on $[a,b]$. Also, by Theorem 6.4.1, $\lim_{j\to\infty}\int_{x_0}^{x} f_j'$ exists and is equal to $\int_{x_0}^{x} g$. We conclude by $(*)$ that $\lim_{j\to\infty} f_j(x)$ exists, and

$$\lim_{j\to\infty} f_j(x) = L + \int_{x_0}^{x} g.$$

Define the function f on $[a,b]$ so that

$$f(x) = L + \int_{x_0}^{x} g.$$

We know so far that $(f_j)_{j=1}^{\infty}$ converges pointwise to f. We now need to show that this convergence is in fact uniform. We defer this part to the exercises. □

Exercise 6.5.4. Complete the proof of Theorem 6.5.3.

Corollary 6.5.5. *Let $a < b$. For every integer $j \geq 1$, let $f_j\colon [a,b] \to \mathbb{R}$ be a differentiable function whose derivative $f_j'\colon [a,b] \to \mathbb{R}$ is continuous. Assume that the series of real numbers $\sum_{j=1}^{\infty} \|f_j'\|_{\infty}$ is absolutely convergent. Assume also that there exists $x_0 \in [a,b]$ such that the series of real numbers $\sum_{j=1}^{\infty} f_j(x_0)$ converges. Then the series $\sum_{j=1}^{\infty} f_j$ converges uniformly on $[a,b]$ to a differentiable function. Moreover, for all $x \in [a,b]$,*

$$\frac{d}{dx}\sum_{j=1}^{\infty} f_j(x) = \sum_{j=1}^{\infty} \frac{d}{dx} f_j(x)$$

Exercise 6.5.6. Prove Corollary 6.5.5.

The following exercise is a nice counterexample to keep in mind, and it also shows the necessity of the assumptions of Corollary 6.5.5.

Exercise 6.5.7. (For this exercise, you can freely use facts about trigonometry that you learned in your previous courses.) Let $x \in \mathbb{R}$ and let $f\colon \mathbb{R} \to \mathbb{R}$ be the function $f(x) := \sum_{j=1}^{\infty} 4^{-j}\cos(32^j \pi x)$. Note that this series is uniformly convergent by the Weierstrass M-test (Theorem 6.3.6). So, f is a continuous function. However, at every point $x \in \mathbb{R}$, f is not differentiable, as we now discuss.

- Show that, for all positive integers j, m, we have
$$|f((j+1)/32^m) - f(j/32^m)| \geq 4^{-m}.$$
(Hint: for certain sequences of numbers $(a_j)_{j=1}^{\infty}$, use the identity
$$\sum_{j=1}^{\infty} a_j = (\sum_{j=1}^{m-1} a_j) + a_m + \sum_{j=m+1}^{\infty} a_j.$$
Also, use the fact that the cosine function is periodic with period 2π, and the summation $\sum_{j=0}^{\infty} r^j = 1/(1-r)$ for all $-1 < r < 1$. Finally, you should require the inequality: for all real numbers x, y, we have $|\cos(x) - \cos(y)| \leq |x - y|$. This inequality follows from the Mean Value Theorem or the Fundamental Theorem of Calculus.)
- Using the previous result, show that, for ever $x \in \mathbb{R}$, f is not differentiable at x. (Hint: for every $x \in \mathbb{R}$ and for every positive integer m, there exists an integer j such that $j \leq 32^m x \leq j+1$.)

- Explain briefly why this result does not contradict Corollary 6.5.5.

6.6. **Uniform Approximation by Polynomials.**

Definition 6.6.1 (**Polynomial**). Let $a < b$ be real numbers and let $x \in [a, b]$. A **polynomial** on $[a, b]$ is a function $f : [a, b] \to \mathbb{R}$ of the form $f(x) = \sum_{j=0}^{k} a_j x^j$, where k is a natural number and $a_0, \ldots, a_k$ are real numbers. If $a_k \neq 0$, then k is called the **degree of** f.

From the previous exercise, we have seen that general continuous functions can behave rather poorly, in that they may never be differentiable. Polynomials on the other hand are infinitely differentiable. And it is often beneficial to deal with polynomials instead of general functions. So, we mention below a result of Weierstrass which says: any continuous function on an interval $[a, b]$ can be uniformly approximated by polynomials.

This fact seems to be related to power series, but it is something much different. It may seem possible to take a general (infinitely differentiable) function, take a high degree Taylor polynomial of this function, and then claim that this polynomial approximates our original function well. There are two problems with this approach. First of all, the continuous function that we start with may not even be differentiable. Second of all, even if we have an infinitely differentiable function, its power series may not actually approximate that function well. Recall that the function $f(x) = e^{-1/x^2}$ (where $f(0) := 0$) is infinitely differentiable, but its Taylor polynomial is identically zero at $x = 0$. In conclusion, we need to use something other than Taylor series to approximate a general continuous function by polynomials.

The proof of the Weierstrass approximation theorem introduces several useful ideas, but it is typically only proven in the honors class. However, later on, we will prove a version of this theorem for trigonometric polynomials, and this proof will be analogous to the proof of the current theorem.

Theorem 6.6.2 (**Weierstrass approximation**). *Let $a < b$ be real numbers. Let $f : [a, b] \to \mathbb{R}$ be a continuous function, and let $\varepsilon > 0$. Then there exists a polynomial P on $[a, b]$ such that $d_\infty(P, f) < \varepsilon$. (That is, $|f(x) - P(x)| < \varepsilon$ for all $x \in [a, b]$.)*

Remark 6.6.3. We can also state this Theorem using metric space terminology. Recall that $C([a, b]; \mathbb{R})$ is the space of continuous functions from $[a, b]$ to $\mathbb{R}$, equipped with the sup-norm metric d_∞. Let $P([a, b]; \mathbb{R})$ be the space of all polynomials on $[a, b]$, so that $P([a, b]; \mathbb{R})$ is a subspace of $C([a, b]; \mathbb{R})$, since polynomials are continuous. Then the Weierstrass approximation theorem says that every continuous function is an adherent point of $P([a, b]; \mathbb{R})$. Put another way, the closure of $P([a, b]; \mathbb{R})$ is $C([a, b]; \mathbb{R})$.

$$\overline{P([a, b]; \mathbb{R})} = C([a, b]; \mathbb{R}).$$

Put another way, every continuous function on $[a, b]$ is the uniform limit of polynomials.

6.7. **Power Series.** We now focus our discussion of series to power series.

Definition 6.7.1 (**Power Series**). Let a be a real number, let $(a_j)_{j=0}^{\infty}$ be a sequence of real numbers, and let $x \in \mathbb{R}$. A **formal power series centered at** a is a series of the form

$$\sum_{j=0}^{\infty} a_j (x - a)^j,$$

For a natural number j, we refer to a_j as the j^{th} **coefficient** of the power series.

Remark 6.7.2. We refer to these power series as formal since their convergence is not guaranteed. Note however that any formal power series centered at a converges at $x = a$. It turns out that we can precisely identify where a formal power series converges just from the asymptotic behavior of the coefficients.

Definition 6.7.3 (**Radius of Convergence**)**.** Let $\sum_{j=0}^{\infty} a_j(x-a)^j$ be a formal power series. The **radius of convergence** $R \geq 0$ of this series is defined to be

$$R := \frac{1}{\limsup_{j\to\infty} |a_j|^{1/j}}.$$

In the definition of R, we use the convention that $1/0 = +\infty$ and $1/(+\infty) = 0$. Note that it is possible for R to then take any value between and including 0 and $+\infty$. Note also that R always exists as a nonnegative real number, or as $+\infty$, since the limit superior of a positive sequence always exists as a nonegative number, or $+\infty$.

Example 6.7.4. The radius of convergence of the series $\sum_{j=0}^{\infty} j(-2)^j(x-3)^j$ is

$$\frac{1}{\limsup_{j\to\infty} |j(-2)^j|^{1/j}} = \frac{1}{\limsup_{j\to\infty} 2j^{1/j}} = \frac{1}{2}.$$

The radius of convergence of the series $\sum_{j=0}^{\infty} 2^{j^2}(x+2)^j$ is

$$\frac{1}{\limsup_{j\to\infty} 2^j} = \frac{1}{+\infty} = 0.$$

The radius of convergence of the series $\sum_{j=0}^{\infty} 2^{-j^2}(x+2)^j$ is

$$\frac{1}{\limsup_{j\to\infty} 2^{-j}} = \frac{1}{0} = +\infty.$$

As we now show, the radius of convergence tells us exactly where the power series converges.

Theorem 6.7.5. *Let $\sum_{j=0}^{\infty} a_j(x-a)^j$ be a formal power series, and let R be its radius of convergence.*

- (a) *(Divergence outside of the radius of convergence) If $x \in \mathbb{R}$ satisfies $|x-a| > R$, then the series $\sum_{j=0}^{\infty} a_j(x-a)^j$ is divergent at x.*
- (b) *(Convergence inside the radius of convergence) If $x \in \mathbb{R}$ satisfies $|x-a| < R$, then the series $\sum_{j=0}^{\infty} a_j(x-a)^j$ is convergent at x.*
- *For the following items (c),(d) and (e), we assume that $R > 0$. Then, let $f\colon (a-R, a+R)$ be the function $f(x) = \sum_{j=0}^{\infty} a_j(x-a)^j$, which exists by part (b).*
- (c) *(Uniform convergence on compact intervals) For any $0 < r < R$, we know that the series $\sum_{j=0}^{\infty} a_j(x-a)^j$ converges uniformly to f on $[a-r, a+r]$. In particular, f is continuous on $(a-R, a+R)$ (by Theorem 6.2.1.)*
- (d) *(Differentiation of power series) The function f is differentiable on $(a-R, a+R)$. For any $0 < r < R$, the series $\sum_{j=0}^{\infty} ja_j(x-a)^{j-1}$ converges uniformly to f' on the interval $[a-r, a+r]$.*

(e) *(Integration of power series) For any closed interval $[y,z]$ contained in $(a-R,a+R)$, we have*

$$\int_y^z f = \sum_{j=0}^{\infty} a_j \frac{(z-a)^{j+1} - (y-a)^{j+1}}{j+1}.$$

Exercise 6.7.6. Prove Theorem 6.7.5. (Hints: for parts (a),(b), use the root test. For part (c), use the Weierstrass M-test. For part (d), use Theorem 6.5.3. For part (e), use Theorem 6.4.4.)

Remark 6.7.7. A power series may converge or diverge when $|x-a| = R$.

Exercise 6.7.8. Give examples of formal power series centered at 0 with radius of convergence $R = 1$ such that

- The series diverges at $x = 1$ and at $x = -1$.
- The series diverges at $x = 1$ and converges at $x = -1$.
- The series converges at $x = 1$ and diverges at $x = 1$.
- The series converges at $x = 1$ and at $x = -1$.

We now discuss functions that are equal to convergent power series.

Definition 6.7.9. Let $a \in \mathbb{R}$ and let $r > 0$. Let E be a subset of $\mathbb{R}$ such that $(a-r,a+r) \subseteq E$. Let $f\colon E \to \mathbb{R}$. We say that the function f is **real analytic on** $(a-r,a+r)$ if and only if there exists a power series $\sum_{j=0}^{\infty} a_j(x-a)^j$ centered at a with radius of convergence R such that $R \geq r$ and such that this power series converges to f on $(a-r,a+r)$.

Example 6.7.10. The function $f\colon (0,2) \to \mathbb{R}$ defined by $f(x) = \sum_{j=0}^{\infty} j(x-1)^j$ is real analytic on $(0,2)$.

From Theorem 6.7.5, if a function f is real analytic on $(a-r,a+r)$, then f is continuous and differentiable. In fact, f is can be differentiated any number of times, as we now show.

Definition 6.7.11. Let E be a subset of $\mathbb{R}$. We say that a function $f\colon E \to \mathbb{R}$ is **once differentiable on** E if and only if f is differentiable on E. More generally, for any integer $k \geq 2$, we say that $f\colon E \to \mathbb{R}$ is k **times differentiable on** E, or just k **times differentiable**, if and only if f is differentiable and f' is $k-1$ times differentiable. If f is k times differentiable, we define the k^{th} derivative $f^{(k)}\colon E \to \mathbb{R}$ by the recursive rule $f^{(1)} := f'$ and $f^{(k)} := (f^{(k-1)})'$, for all $k \geq 2$. We also define $f^{(0)} := f$. A function is said to be **infinitely differentiable** if and only if f is k times differentiable for every $k \geq 0$.

Example 6.7.12. The function $f(x) = |x|^3$ is twice differentiable on $\mathbb{R}$, but not three times differentiable on $\mathbb{R}$. Note that $f''(x) = 6\,|x|$, which is not differentiable at $x = 0$.

Proposition 6.7.13. *Let $a \in \mathbb{R}$ and let $r > 0$. Let f be a function that is real analytic on $(a-r,a+r)$, with the power series expansion*

$$f(x) = \sum_{j=0}^{\infty} a_j(x-a)^j, \quad \forall\, x \in (a-r,a+r).$$

Then, for any integer $k \geq 0$, the function f is k times differentiable on $(a-r, a+r)$, and the k^{th} derivative is given by

$$f^{(k)}(x) = \sum_{j=0}^{\infty} a_{j+k}(j+1)(j+2)\cdots(j+k)(x-a)^j, \quad \forall\, x \in (a-r, a+r).$$

Exercise 6.7.14. Prove Proposition 6.7.13.

Corollary 6.7.15 (Taylor's formula). *Let $a \in \mathbb{R}$ and let $r > 0$. Let f be a function that is real analytic on $(a-r, a+r)$, with the power series expansion*

$$f(x) = \sum_{j=0}^{\infty} a_j(x-a)^j, \quad \forall\, x \in (a-r, a+r).$$

Then, for any integer $k \geq 0$, we have

$$f^{(k)}(a) = k!a_k,$$

*where $k! = 1 \times 2 \times \cdots \times k$, and we denote $0! := 1$. In particular, we have **Taylor's formula***

$$f(x) = \sum_{j=0}^{\infty} \frac{f^{(j)}(a)}{j!}(x-a)^j, \quad \forall\, x \in (a-r, a+r).$$

Exercise 6.7.16. Prove Corollary 6.7.15 using Proposition 6.7.13.

Remark 6.7.17. The series $\sum_{j=0}^{\infty} \frac{f^{(j)}(a)}{j!}(x-a)^j$ is sometimes called the **Taylor series** of f around a. Taylor's formula says that if f is real analytic, then f is equal to its Taylor series. In the following exercise, we see that even if f is infinitely differentiable, it may not be equal to its Taylor series.

Exercise 6.7.18. Define a function $f\colon \mathbb{R} \to \mathbb{R}$ by $f(0) := 0$ and $f(x) := e^{-1/x^2}$ for $x \neq 0$. Show that f is infinitely differentiable, but $f^{(k)}(0) = 0$ for all $k \geq 0$. So, being infinitely differentiable does not imply that f is equal to its Taylor series. (You may freely use properties of the exponential function that you have learned before.)

Corollary 6.7.19 (Uniqueness of power series). *Let $a \in \mathbb{R}$ and let $r > 0$. Let f be a function that is real analytic on $(a-r, a+r)$, with two power series expansions*

$$f(x) = \sum_{j=0}^{\infty} a_j(x-a)^j, \quad \forall\, x \in (a-r, a+r).$$

$$f(x) = \sum_{j=0}^{\infty} b_j(x-a)^j, \quad \forall\, x \in (a-r, a+r).$$

Then $a_j = b_j$ for all $j \geq 0$.

Proof. By Corollary 6.7.15, we have $k!a_k = f^{(k)}(a) = k!b_k$ for all $k \geq 0$. Since $k! \neq 0$ for all $k \geq 0$, we divide by $k!$ to get $a_k = b_k$ for all $k \geq 0$. $\square$

Remark 6.7.20. Note however that a power series can have very different expansions if we change the center of the expansion. For example, the function $f(x) = 1/(1-x)$ satisfies

$$f(x) = \sum_{j=0}^{\infty} x^j, \quad \forall\, x \in (-1,1).$$

However, at the point $1/2$, we have the different expansion

$$f(x) = \frac{1}{1-x} = \frac{2}{1-2(x-1/2)} = \sum_{j=0}^{\infty} 2(2(x-1/2))^j = \sum_{j=0}^{\infty} 2^{j+1}(x-1/2)^j, \quad \forall x \in (0,1).$$

Note also that the first series has radius of convergence 1 and the second series has radius of convergence $1/2$.

6.7.1. *Multiplication of Power Series.*

Lemma 6.7.21 (Fubini's Theorem for Series). *Let $f\colon \mathbb{N}\times\mathbb{N} \to \mathbb{R}$ be a function such that $\sum_{(j,k)\in\mathbb{N}\times\mathbb{N}} f(j,k)$ is absolutely convergent. (That is, for any bijection $g\colon \mathbb{N} \to \mathbb{N}\times\mathbb{N}$, the sum $\sum_{\ell=0}^{\infty} f(g(\ell))$ is absolutely convergent.) Then*

$$\sum_{j=1}^{\infty}\Big(\sum_{k=1}^{\infty} f(j,k)\Big) = \sum_{(j,k)\in\mathbb{N}\times\mathbb{N}} f(j,k) = \sum_{k=1}^{\infty}\Big(\sum_{j=1}^{\infty} f(j,k)\Big).$$

Proof Sketch. We only consider the case $f(j,k) \geq 0$ for all $(j,k)\in\mathbb{N}$. The general case then follows by writing $f = \max(f,0) - \min(f,0)$, and applying this special case to $\max(f,0)$ and $\min(f,0)$, separately.

Let $L := \sum_{(j,k)\in\mathbb{N}\times\mathbb{N}} f(j,k)$. For any $J,K>0$, we have $\sum_{j=1}^{J}\sum_{k=1}^{K} f(j,k) \leq L$. Letting $J,K\to\infty$, we conclude that $\sum_{j=1}^{\infty}\sum_{k=1}^{\infty} f(j,k) \leq L$. Let $\varepsilon>0$. It remains to find J,K such that $\sum_{j=1}^{J}\sum_{k=1}^{K} > L-\varepsilon$. Since $\sum_{(j,k)\in\mathbb{N}\times\mathbb{N}} f(j,k)$ converges absolutely, there exists a finite set $X\subseteq \mathbb{N}\times\mathbb{N}$ such that $\sum_{(j,k)\in X} f(j,k) > L-\varepsilon$. But then we can choose J,K sufficiently large such that $\{(j,k)\in X\} \subseteq \{(j,k)\colon 1\leq j\leq J,\, 1\leq k\leq K\}$. Therefore, $\sum_{j=1}^{J}\sum_{k=1}^{K} f(j,k) \geq \sum_{(j,k)\in X} f(j,k) > L-\varepsilon$, as desired. $\square$

Theorem 6.7.22. *Let $a\in\mathbb{R}$ and let $r>0$. Let f and g be functions that are real analytic on $(a-r,a+r)$, with power series expansions*

$$f(x) = \sum_{j=0}^{\infty} a_j (x-a)^j, \quad \forall\, x\in(a-r,a+r).$$

$$g(x) = \sum_{j=0}^{\infty} b_j (x-a)^j, \quad \forall\, x\in(a-r,a+r).$$

Then the function fg is also real analytic on $(a-r,a+r)$. For each $j\geq 0$, define $c_j := \sum_{k=0}^{j} a_k b_{j-k}$. Then fg has the power series expansion

$$f(x)g(x) = \sum_{j=0}^{\infty} c_j (x-a)^j, \quad \forall\, x\in(a-r,a+r).$$

Proof. Fix $x \in (a-r, a+r)$. By Theorem 6.7.5, both f and g have radius of convergence $R \geq r$. So, both $\sum_{j=0}^{\infty} a_j(x-a)^j$ and $\sum_{j=0}^{\infty} b_j(x-a)^j$ are absolutely convergent. Define

$$C := \sum_{j=0}^{\infty} \left|a_j(x-a)^j\right|, \qquad D := \sum_{j=0}^{\infty} \left|b_j(x-a)^j\right|.$$

Then both C, D are finite.

For any $N \geq 0$, consider the partial sum

$$\sum_{j=0}^{N} \sum_{k=0}^{N} \left|a_j(x-a)^j b_k(x-a)^k\right|.$$

We can re-write this sum as

$$\sum_{j=0}^{N} \left|a_j(x-a)^j\right| \sum_{k=0}^{N} \left|b_k(x-a)^k\right| \leq \sum_{j=0}^{N} \left|a_j(x-a)^j\right| D \leq CD.$$

Since this inequality holds for all $N \geq 0$, the series

$$\sum_{(j,k)\in\mathbb{N}\times\mathbb{N}} \left|a_j(x-a)^j b_k(x-a)^k\right|$$

is convergent. That is, the following series is absolutely convergent.

$$\sum_{(j,k)\in\mathbb{N}\times\mathbb{N}} a_j(x-a)^j b_k(x-a)^k.$$

Now, using Lemma 6.7.21,

$$\sum_{(j,k)\in\mathbb{N}\times\mathbb{N}} a_j(x-a)^j b_k(x-a)^k = \sum_{j=0}^{\infty} a_j(x-a)^j \sum_{k=0}^{\infty} b_k(x-a)^k = \sum_{j=0}^{\infty} a_j(x-a)^j g(x) = f(x)g(x).$$

Rewriting this equality,

$$f(x)g(x) = \sum_{(j,k)\in\mathbb{N}\times\mathbb{N}} a_j(x-a)^j b_k(x-a)^k = \sum_{(j,k)\in\mathbb{N}\times\mathbb{N}} a_j b_k (x-a)^{j+k}.$$

Since the sum is absolutely convergent, we can rearrange the order of summation. For any fixed positive integer ℓ, we sum over all positive integers j, k such that $j + k = \ell$. That is, we have

$$f(x)g(x) = \sum_{\ell=0}^{\infty} \sum_{(j,k)\in\mathbb{N}\times\mathbb{N}\colon\, j+k=\ell} a_j b_k (x-a)^{\ell} = \sum_{\ell=0}^{\infty} (x-a)^{\ell} \sum_{s=0}^{\ell} a_s b_{s-\ell}.$$

□

6.8. **The Exponential and Logarithm.** We can now use the material from the previous sections to define and investigate various special functions.

Definition 6.8.1. For every real number x, we define the **exponential function** $\exp(x)$ to be the real number

$$\exp(x) := \sum_{j=0}^{\infty} \frac{x^j}{j!}.$$

Theorem 6.8.2 (**Properties of the Exponential Function**).

(a) *For every real number* x*, the series* $\sum_{j=0}^{\infty} \frac{x^j}{j!}$ *is absolutely convergent. So,* $\exp(x)$ *exists and is a real number for every* $x \in \mathbb{R}$*, the power series* $\sum_{j=0}^{\infty} \frac{x^j}{j!}$ *has radius of convergence* $R = +\infty$*, and* $\exp$ *is an analytic function on* $(-\infty, +\infty)$*.*

(b) $\exp$ *is differentiable on* $\mathbb{R}$*, and for every* $x \in \mathbb{R}$*, we have* $\exp'(x) = \exp(x)$*.*

(c) $\exp$ *is continuous on* $\mathbb{R}$*, and for all real numbers* $a < b$*, we have* $\int_a^b \exp = \exp(b) - \exp(a)$*.*

(d) *For every* $x, y \in \mathbb{R}$*, we have* $\exp(x + y) = \exp(x)\exp(y)$*.*

(e) $\exp(0) = 1$*. Also, for every* $x \in \mathbb{R}$*, we have* $\exp(x) > 0$*, and* $\exp(-x) = 1/\exp(x)$*.*

(f) $\exp$ *is strictly monotone increasing. That is, whenever* x, y *are real numbers with* $x < y$*, we have* $\exp(x) < \exp(y)$*.*

Exercise 6.8.3. Prove Theorem 6.8.2. (Hints: for part (a), use the ratio test. For parts (b) and (c), use Theorem 6.7.5. For part (d), you may need the binomial formula $(x + y)^k = \sum_{j=0}^{k} \frac{k!}{j!(k-j)!} x^j y^{k-j}$. For part (e), use part (d). For part (f), use part (d) and show that $\exp(x) > 1$ for all $x > 0$.)

Definition 6.8.4. We define the real number e by

$$e := \exp(1) = \sum_{j=0}^{\infty} \frac{1}{j!}$$

Proposition 6.8.5. *For every real number* x*, we have*

$$\exp(x) = e^x.$$

Exercise 6.8.6. Prove Proposition 6.8.5. (Hint: first prove the proposition for natural numbers x. Then, prove the proposition for integers. Then, prove the proposition for rational numbers. Finally, use the density of the rationals to prove the proposition for real numbers. You should find useful identifies for exponentiation by rational numbers.)

From now on, we use $\exp(x)$ and e^x interchangeably.

Remark 6.8.7. Since $e > 1$ by the definition of e, we have $e^x \to +\infty$ as $x \to +\infty$ and $e^x \to 0$ as $x \to -\infty$. So, from the Intermediate Value Theorem, the range of exp is $(0, \infty)$. Since exp is strictly increasing on $\mathbb{R}$, exp is therefore injective on $\mathbb{R}$, so exp is a bijection from $\mathbb{R}$ to $(0, \infty)$. Therefore, exp has an inverse function from $(0, \infty)$ to $\mathbb{R}$.

Definition 6.8.8. We define the **natural logarithm function** $\log \colon (0, \infty) \to \mathbb{R}$ (which is also called ln) to be the inverse of the exponential function. So, $\exp(\log(x)) = x$ for every $x \in (0, \infty)$, and $\log(\exp(x)) = x$ for every $x \in \mathbb{R}$.

Remark 6.8.9. Since exp is continuous and strictly monotone increasing, log is also continuous and strictly monotone increasing by Proposition 3.2.29. Since exp is differentiable and its derivative is never zero, the Inverse Function Theorem (Theorem 3.3.36) implies that log is also differentiable.

Theorem 6.8.10.

(a) *For every* $x \in (0, \infty)$*, we have* $\log'(x) = 1/x$*. So, by the Fundamental Theorem of Calculus, for any* $0 < a < b$*, we have* $\int_a^b (1/t)dt = \log(b) - \log(a)$*.*

(b) *For all* $x, y \in (0, \infty)$*, we have* $\log(x) + \log(y) = \log(xy)$*.*
(c) *For all* $x \in (0, \infty)$*, we have* $\log(1/x) = -\log x$*. In particular,* $\log(1) = 0$*.*
(d) *For any* $x \in (0, \infty)$ *and* $y \in \mathbb{R}$*, we have* $\log(x^y) = y \log x$*.*
(e) *For any* $x \in (-1, 1)$*, we have*

$$-\log(1-x) = \sum_{j=1}^{\infty} \frac{x^j}{j}.$$

In particular, $\log$ *is analytic on* $(0, 2)$ *with the power series expansion*

$$\log(x) = \sum_{j=1}^{\infty} \frac{(-1)^{j+1}}{j}(x-1)^j, \quad \forall\, x \in (0, 2).$$

Exercise 6.8.11. Prove Theorem 6.8.10. (Hints: for part (a), use the Inverse Function Theorem or Chain Rule. For parts (b),(c) and (d), use Theorem 6.8.2 and the laws of exponentiation. For part (e), let $x \in (-1, 1)$, use the geometric series formula $1/(1-x) = \sum_{j=0}^{\infty} x^j$ and integrate using Theorem 6.7.5.)

6.8.1. *A Digression concerning Complex Numbers.* Our investigation of trigonometric functions below is significantly improved by the introduction of the complex number system. We will also use the complex exponential in our discussion of Fourier series.

Definition 6.8.12 (**Complex Numbers**)**.** A **complex number** is any expression of the form $a + bi$ where a, b are real numbers. Formally, the symbol i is a placeholder with no intrinsic meaning. Two complex numbers $a+bi$ and $c+di$ are said to be equal if and only if $a = c$ and $b = d$. Every real number x is considered a complex number, with the identification $x = x+0i$. The sum of two complex numbers is defined by $(a+bi)+(c+di) := (a+c)+(b+d)i$. The difference of two complex numbers is defined by $(a+bi)-(c+di) := (a-c)+(b-d)i$. The product of two complex numbers is defined by $(a+bi)(c+di) := (ac-bd)+(ad+bc)i$. If $c+di \neq 0$, the quotient of two complex numbers is defined by $(a+bi)/(c+di) := (a+bi)(\frac{c}{c^2+d^2} - \frac{d}{c^2+d^2}i)$. The **complex conjugate** of a complex number $a + bi$ is defined by $\overline{a+bi} := a - bi$. The absolute value of a complex number $a + bi$ is defined by $|a + bi| := \sqrt{a^2 + b^2}$. The space of all complex numbers is called $\mathbb{C}$.

Remark 6.8.13. We write i as shorthand for $0 + i$. Note that $i^2 = -1$.

Remark 6.8.14. The complex numbers obey all of the usual rules of algebra. For example, if v, w, z are complex numbers, then $v(w + z) = vw + vz$, $v(wz) = (vw)z$, and so on. Specifically, the complex numbers $\mathbb{C}$ form a **field**. Also, the rules of complex arithmetic are consistent with the rules of real arithmetic. That is, $3+5 = 8$ whether or not we use addition in $\mathbb{R}$ or addition in $\mathbb{C}$.

The operation of complex conjugation preserves all of the arithmetic operations. If w, z are complex numbers, then $\overline{w+z} = \overline{w} + \overline{z}$, $\overline{w-z} = \overline{w} - \overline{z}$, $\overline{w \cdot z} = \overline{w} \cdot \overline{z}$, and $\overline{w/z} = \overline{w}/\overline{z}$ for $z \neq 0$. The complex conjugate and absolute value satisfy $|z|^2 = z\overline{z}$.

Remark 6.8.15. If $z \in \mathbb{C}$, then $|z| = 0$ if and only if $z = 0$. If $z, w \in \mathbb{C}$, then it can be shown that $|zw| = |z|\,|w|$, and if $w \neq 0$, then $|z/w| = |z|\,/\,|w|$. Also, the triangle inequality holds: $|z + w| \leq |z| + |w|$. So, $\mathbb{C}$ is a metric space if we use the metric $d(z, w) := |z - w|$. Moreover, $\mathbb{C}$ is a complete metric space.

The theory we have developed to deal with series of real functions also covers complex-valued functions, with almost no change to the proofs. For example, we can define the exponential function of a complex number z by

$$\exp(z) := \sum_{j=0}^{\infty} \frac{z^j}{j!}.$$

The ratio test then can be proven in exactly the same manner for complex series, and it follows that $\exp(z)$ converges for every $z \in \mathbb{C}$. Many of the properties of Theorem 6.8.2 still hold, though we cannot deal with all of these properties in this class. However, the following identity is proven in the exact same way as in the setting of real numbers: for any $z, w \in \mathbb{C}$, we have

$$\exp(z + w) = \exp(z)\exp(w).$$

Also, we should note that $\overline{\exp(z)} = \exp(\overline{z})$, which follows by conjugating the partial sums $\sum_{j=0}^{J} z^j/j!$, and then letting $J \to \infty$.

We briefly mention that the complex logarithm is more difficult to define, mainly because the exponential function is not invertible on $\mathbb{C}$. This topic is deferred to the complex analysis class.

6.9. **Trigonometric Functions.** Besides the exponential and logarithmic functions, there are many different kinds of special functions. Here, we will only mention the sine and cosine functions. One's first encounter with the sine and cosine functions probably involved their definition in terms of the edge lengths of right triangles. However, we will show below an analytic definition of these functions, which will also facilitate the investigation of the properties that they possess. The complex exponential plays a crucial role in this development.

Definition 6.9.1. Let x be a real number. We then define

$$\cos(x) := \frac{e^{ix} + e^{-ix}}{2}.$$

$$\sin(x) := \frac{e^{ix} - e^{-ix}}{2i}.$$

We refer to cos as the **cosine** function, and we refer to sin as the **sine** function.

Remark 6.9.2. Using the power series expansion for the exponential, we can then derive power series expansions for sine and cosine as follows. Let $x \in \mathbb{R}$. Then

$$e^{ix} = 1 + ix - x^2/2! - ix^3/3! + x^4/4! + \cdots$$

$$e^{-ix} = 1 - ix - x^2/2! + ix^3/3! + x^4/4! - \cdots$$

Therefore, using the definitions of sine and cosine,

$$\cos(x) = 1 - x^2/2! + x^4/4! - \cdots = \sum_{j=0}^{\infty} \frac{(-1)^j x^{2j}}{(2j)!}.$$

$$\sin(x) = x - x^3/3! + x^5/5! - \cdots = \sum_{j=0}^{\infty} \frac{(-1)^j x^{2j+1}}{(2j+1)!}.$$

So, if $x \in \mathbb{R}$ then $\cos(x) \in \mathbb{R}$ and $\sin(x) \in \mathbb{R}$. Also, sine and cosine real analytic on $(-\infty, \infty)$, e.g. since their power series converge on $(-\infty, \infty)$ by the ratio test. In particular, the sine and cosine functions are continuous and infinitely differentiable.

Theorem 6.9.3 (Properties of Sine and Cosine).

(a) *For any real number x we have $\cos(x)^2 + \sin(x)^2 = 1$. In particular, $\sin(x) \in [-1, 1]$ and $\cos(x) \in [-1, 1]$ for all real numbers x.*
(b) *For any real number x, we have $\sin'(x) = \cos(x)$, and $\cos'(x) = -\sin(x)$.*
(c) *For any real number x, we have $\sin(-x) = -\sin(x)$ and $\cos(-x) = \cos(x)$.*
(d) *For any real numbers x, y we have $\cos(x + y) = \cos(x)\cos(y) - \sin(x)\sin(y)$ and $\sin(x + y) = \sin(x)\cos(y) + \cos(x)\sin(y)$.*
(e) *$\sin(0) = 0$ and $\cos(0) = 1$.*
(f) *For every real number x, we have $e^{ix} = \cos(x) + i\sin(x)$ and $e^{-ix} = \cos(x) - i\sin(x)$.*

Exercise 6.9.4. Prove Theorem 6.9.3. (Hints: whenever possible, write everything in terms of exponentials.)

Lemma 6.9.5. *There exists a positive real number x such that $\sin(x) = 0$.*

Proof. We argue by contradiction. Suppose $\sin(x) \neq 0$ for all $x > 0$. We conclude that $\cos(x) \neq 0$ for all $x > 0$, since $\cos(x) = 0$ implies that $\sin(2x) = 0$, by Theorem 6.9.3(d). Since $\cos(0) = 1$, we conclude that $\cos(x) > 0$ for all $x > 0$ by the Intermediate Value Theorem. Since $\sin(0) = 0$ and $\sin'(0) = 1 > 0$, we know that sin is positive for small positive x. Therefore, $\sin(x) > 0$ for all $x > 0$ by the Intermediate Value Theorem.

Define $\cot(x) := \cos(x)/\sin(x)$. Then cot is positive on $(0, \infty)$, and cot is differentiable for $x > 0$. From the quotient rule and Theorem 6.9.3(a), we have $\cot'(x) = -1/\sin^2(x)$. So, $\cot'(x) \leq -1$ for all $x > 0$. Then, by the Fundamental Theorem of Calculus, for all $x, s > 0$, we have $\cot(x + s) \leq \cot(x) - s$. Letting $s \to \infty$ shows that cot eventually becomes negative on $(0, \infty)$, a contradiction. □

Let E be the set $E := \{x \in (0, \infty) \colon \sin(x) = 0\}$, so that E is the set of zeros of the sine function. By Lemma 6.9.5, E is nonempty. Also, since sin is continuous, E is a closed set. (Note that $E = \sin^{-1}(0)$.) In particular, E contains all of its adherent points, so E contains $\inf(E)$.

Definition 6.9.6. We define π to be the number

$$\pi := \inf\{x \in (0, \infty) \colon \sin(x) = 0\}.$$

Then $\pi > 0$ and $\sin(\pi) = 0$. Since sin is nonzero on $(0, \pi)$ and $\sin'(0) = 1 > 0$, we conclude that sin is positive on $(0, \pi)$. Since $\cos'(x) = -\sin(x)$, we see that cos is decreasing on $(0, \pi)$. Since $\cos(0) = 1$, we therefore have $\cos(\pi) < 1$. Since $\sin^2(\pi) + \cos^2(\pi) = 1$ and $\sin(\pi) = 0$, we conclude that $\cos(\pi) = -1$.

We therefore deduce Euler's famous formula

$$e^{i\pi} = \cos(\pi) + i\sin(\pi) = -1.$$

Here are some more properties of sine and cosine.

Theorem 6.9.7.

(a) *For any real x we have $\cos(x+\pi) = -\cos(x)$ and $\sin(x+\pi) = -\sin(x)$. In particular, we have $\cos(x + 2\pi) = \cos(x)$ and $\sin(x + 2\pi) = \sin(x)$, so that $\sin$ and $\cos$ are 2π-periodic.*
(b) *If x is real, then $\sin(x) = 0$ if and only if x/π is an integer.*
(c) *If x is real, then $\cos(x) = 0$ if and only if x/π is an integer plus $1/2$.*

Exercise 6.9.8. Prove Theorem 6.9.7.

7. Appendix: Notation

Let A, B be sets in a space X. Let m, n be a nonnegative integers.

$$
\begin{aligned}
\mathbb{Z} &:= \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}, \text{ the integers}\\
\mathbb{N} &:= \{0, 1, 2, 3, 4, 5, \ldots\}, \text{ the natural numbers}\\
\mathbb{Z}_+ &:= \{1, 2, 3, 4, \ldots\}, \text{ the positive integers}\\
\mathbb{Q} &:= \{m/n \colon m, n \in \mathbb{Z}, n \neq 0\}, \text{ the rationals}\\
\mathbb{R} &\text{ denotes the set of real numbers}\\
\mathbb{R}^* &= \mathbb{R} \cup \{-\infty\} \cup \{+\infty\} \text{ denotes the set of extended real numbers}\\
\mathbb{C} &:= \{x + y\sqrt{-1} : x, y \in \mathbb{R}\}, \text{ the complex numbers}\\
\emptyset &\text{ denotes the empty set, the set consisting of zero elements}\\
\in &\text{ means "is an element of." For example, } 2 \in \mathbb{Z} \text{ is read as "2 is an element of } \mathbb{Z}.\text{"}\\
\forall &\text{ means "for all"}\\
\exists &\text{ means "there exists"}\\
\mathbb{F}^n &:= \{(x_1, \ldots, x_n) \colon x_i \in \mathbb{F}, \forall\, i \in \{1, \ldots, n\}\}\\
A \subseteq B &\text{ means } \forall\, a \in A, \text{ we have } a \in B, \text{ so } A \text{ is contained in } B\\
A \setminus B &:= \{x \in A \colon x \notin B\}\\
A^c &:= X \setminus A, \text{ the complement of } A\\
A \cap B &\text{ denotes the intersection of } A \text{ and } B\\
A \cup B &\text{ denotes the union of } A \text{ and } B
\end{aligned}
$$

Let E be a subset of $\mathbb{R} \cup \{-\infty\} \cup \{+\infty\}$. Let $(a_n)_{n=0}^{\infty}$ be a sequence of real numbers.

$$
\begin{aligned}
\sup(E) &\text{ denotes the smallest upper bound of } E\\
\inf(E) &\text{ denotes the largest lower bound of } E\\
\limsup (a_n)_{n=0}^{\infty} &:= \lim_{n\to\infty} \sup_{m\geq n} (a_n)_{n=m}^{\infty}\\
\liminf (a_n)_{n=0}^{\infty} &:= \lim_{n\to\infty} \inf_{m\geq n} (a_n)_{n=m}^{\infty}
\end{aligned}
$$

Let (X,d) be a metric space, let $x_0 \in X$, let $r > 0$ be a real number, and let E be a subset of X. Let $(x_1,\ldots,x_n)$ be an element of $\mathbb{R}^n$, and let $p \geq 1$ be a real number.

$$B_{(X,d)}(x_0,r) = B(x_0,r) := \{x \in X \colon d(x,x_0) < r\}.$$

$$\overline{E} \text{ denotes the closure of } E$$

$$\mathrm{int}(E) \text{ denotes the interior of } E$$

$$\partial E \text{ denotes the boundary of } E$$

$$\|(x_1,\ldots,x_n)\|_{\ell_p} := (\sum_{i=1}^{n} |x_i|^p)^{1/p}$$

$$\|(x_1,\ldots,x_n)\|_{\ell_\infty} := \max_{i=1,\ldots,n} |x_i|$$

Let $f,g\colon (X,d_X) \to (Y,d_Y)$ be maps between metric spaces. Let $V \subseteq X$, and let $W \subseteq Y$.

$$f(V) := \{f(v) \in Y \colon v \in V\}.$$

$$f^{-1}(W) := \{x \in X \colon f(x) \in W\}.$$

$$d_\infty(f,g) := \sup_{x \in X} d_Y(f(x),g(x)).$$

$B(X;Y)$ denotes the set of functions $f\colon X \to Y$ that are bounded.

$$C(X;Y) := \{f \in B(X;Y) \colon f \text{ is continuous}\}.$$

7.0.1. *Set Theory.* Let X,Y be sets, and let $f\colon X \to Y$ be a function. The function $f\colon X \to Y$ is said to be **injective** (or **one-to-one**) if and only if: for every $x,x' \in V$, if $f(x) = f(x')$, then $x = x'$.

The function $f\colon X \to Y$ is said to be **surjective** (or **onto**) if and only if: for every $y \in Y$, there exists $x \in X$ such that $f(x) = y$.

The function $f\colon X \to Y$ is said to be **bijective** (or a **one-to-one correspondence**) if and only if: for every $y \in Y$, there exists exactly one $x \in X$ such that $f(x) = y$. A function $f\colon X \to Y$ is bijective if and only if it is both injective and surjective.

Two sets X,Y are said to have the same **cardinality** if and only if there exists a bijection from X onto Y.

USC Mathematics, Los Angeles, CA
Email address: stevenmheilman@gmail.com