Eidgenössische Technische Hochschule Zürich

Analysis I: One Variable

Lecture Notes 2023

Alessio Figalli

February 19, 2024

Preface

Welcome to ETH Zürich and to your exploration of these lecture notes. Originally crafted in German for the academic year 2016/2017 by Manfred Einsiedler and Andreas Wieser, these notes were designed for the Analysis I and II courses in the Interdisciplinary Natural Sciences, Physics, and Mathematics Bachelor programs. In the academic year 2019/2020, a substantial revision was undertaken by Peter Jossen.

For the academic year 2023/2024, Alessio Figalli has developed this English version. It differs from the German original in several aspects: reorganization and alternative proofs of some materials, extensive rewriting and expansion in certain areas, and a more concise presentation. This version strictly aligns with the material presented in class, offering a streamlined educational experience.

The courses Analysis I/II and Linear Algebra I/II are fundamental to the mathematics curriculum at ETH and other universities worldwide. They lay the groundwork upon which most future studies in mathematics and physics are built.

Throughout Analysis I/II, we will delve into various aspects of differential and integral calculus. Although some topics might be familiar from high school, our approach requires minimal prior knowledge beyond an intuitive understanding of variables and basic algebraic skills. Contrary to high-school methods, our lectures emphasize the development of mathematical theory over algorithmic practice. Understanding and exploring topics such as differential equations and multidimensional integral theorems is our primary goal. However, students are encouraged to engage with numerous exercises from these notes and other resources to deepen their understanding and proficiency in these new mathematical concepts.

Contents

Chapter 1

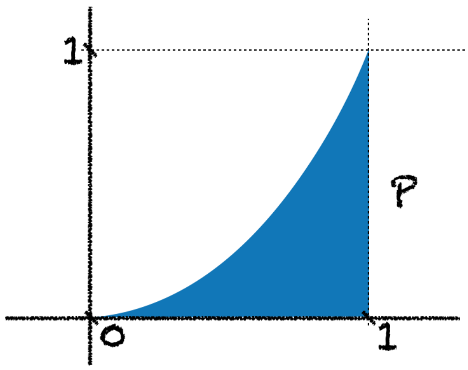
Introduction

1.1 Quadrature of the Parabola

Before we begin our journey in the world of mathematical analysis, as an example of how we want to think and proceed here but also as an introduction to integral calculus, consider the set

$$P = \{(x, y) \in \mathbb{R}^2 \mid 0 \leq x \leq 1,\ 0 \leq y \leq x^2\}. \tag{1.1}$$

Our goal is to calculate its area.



1

This area was already determined as the first curvilinear bounded area by Archimedes (ca. 287–ca. 212 BCE) in the 3rd century BCE. For the area calculation, let us assume that we know what the symbols in the definition in equation (1.1) mean and that P describes the area in the following figure. In particular, we assume for the moment that we already know the set of real numbers $\mathbb{R}$.

Of course, calculating the area of P is not a challenge if we use integrals and the associated calculation rules. However, we do not want to assume we know the integral calculus. Strictly speaking, we must ask ourselves the following fundamental question before calculating:

What is an area?

If we cannot answer this question exactly, then we cannot know what it means to calculate the area of P. Therefore, we qualify our goal in the following way:

Proposition 1.1. — *Suppose there is a notion of area in* $\mathbb{R}^2$ *that satisfies the following properties:*

1. *The area of the rectangle* $[a,b]\times[c,d]=\{(x,y)\in\mathbb{R}^2 \mid a\le x\le b,\ c\le y\le d\}$ *is equal to* $(b-a)(d-c)$*, where* a,b,c,d *are real numbers with* $a\le b, c\le d$.
2. *If* G *is a domain in* $\mathbb{R}^2$ *and* F *is a domain contained in* G*, then the area of* F *is less than or equal to the area of* G.
3. *For sets* F,G *in* $\mathbb{R}^2$ *without common points, the area of the union* $F\cup G$ *is the sum of the areas of* F *and* G.

Then the area of P *as in equation* (1.1) *(if defined at all) is equal to* $\frac{1}{3}$.

In other words, we have left open the question of whether there is a notion of area and for what areas it is defined, but we want to show that $\frac{1}{3}$ is the only reasonable value for the area of P.

For the proof of Proposition 1.1 we need a lemma (also called an “auxiliary theorem”):

Lemma 1.2. — *Let* $n\ge 1$ *be a natural number. Then*

$$1^2+2^2+\cdots+(n-1)^2+n^2=\frac{n^3}{3}+\frac{n^2}{2}+\frac{n}{6}. \tag{1.2}$$

1 *Proof.* We perform the proof using induction. For $n=1$, the left-hand side of equation (1.2) is equal to 1 and the right-hand side is equal to $\frac{1}{3}+\frac{1}{2}+\frac{1}{6}=1$. So equation (1.2) is true for $n=1$. *This part of the proof is called the beginning of induction.*

Suppose we already know that equation (1.2) holds for the natural number n. We now want to show that it follows that equation (1.2) also holds for $n+1$. The left-hand side of equation (1.2), for $(n+1)$ instead of n, is given by

$$1^2+2^2+\cdots+n^2+(n+1)^2=\frac{n^3}{3}+\frac{n^2}{2}+\frac{n}{6}+(n+1)^2=\frac{n^3}{3}+\frac{3n^2}{2}+\frac{13n}{6}+1$$

where, in the first equality, we have used the validity of equation (1.2) for the number n. The right-hand side of equation (1.2), for $(n+1)$ instead of n, is given by

$$\begin{aligned}\frac{(n+1)^3}{3}+\frac{(n+1)^2}{2}+\frac{n+1}{6}&=\frac{n^3+3n^2+3n+1}{3}+\frac{n^2+2n+1}{2}+\frac{n+1}{6}\\&=\frac{n^3}{3}+\frac{3n^2}{2}+\frac{13n}{6}+1\end{aligned}$$

This shows that the left and right sides of equation (1.2) also agree for $n+1$. *This part of the proof is called the induction step.*

It follows that equation (1.2) is true for $n=1$ due to the validity of the beginning of the induction. Therefore, it is also true for $n=2$ due to the induction step, and for $n=3$ again due to the induction step. Continuing in this way, we obtain (1.2) for any natural number.

We say that equation (1.2) follows by means of induction for all natural numbers $n \geq 1$. Furthermore, we indicate the end of the proof with a small square. □

Proof of Proposition 1.1. We assume that there is a notion of area content with the properties in the proposition and that it is defined for P. Suppose that I is the area content of P. We cover P for a given natural number $n \geq 1$ with rectangles whose base has length $\frac{1}{n}$, as in Figure 1.1 on the left.

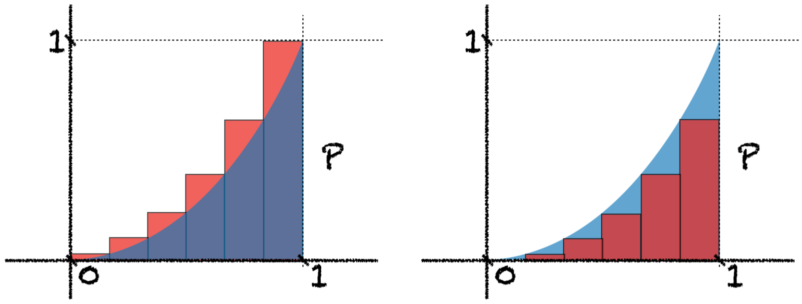


Figure 1.1: Approximation of P with $n = 6$ rectangles.

Note that the first rectangle from the left has height $\left(\frac{1}{n}\right)^2$, the second one $\left(\frac{2}{n}\right)^2$, and so on. Hence, from the assumed properties of the area, thanks to Lemma 1.2 the following inequality holds:

1

$$\begin{aligned} I &\leq \frac{1}{n}\frac{1}{n^2} + \frac{1}{n}\frac{2^2}{n^2} + \cdots + \frac{1}{n}\frac{n^2}{n^2} \\ &= \frac{1}{n^3}(1^2 + 2^2 + \cdots + n^2) \\ &= \frac{1}{n^3}\left(\frac{n^3}{3} + \frac{n^2}{2} + \frac{n}{6}\right) \\ &= \frac{1}{3} + \frac{1}{2n} + \frac{1}{6n^2}. \end{aligned}$$

Note that the straight-line segments where the rectangles touch have area 0, and we may ignore them.

If, on the other hand, we use rectangles as in Figure 1.1 on the right, we also get

$$\begin{aligned} I &\geq \frac{1}{n}\frac{0}{n^2} + \frac{1}{n}\frac{1^2}{n^2} + \cdots + \frac{1}{n}\frac{(n-1)^2}{n^2} \\ &= \frac{1}{n^3}(1^2 + \cdots + (n-1)^2) \\ &= \frac{1}{n^3}(1^2 + \cdots + (n-1)^2 + n^2 - n^2) \\ &= \frac{1}{n^3}\left(\frac{n^3}{3} + \frac{n^2}{2} + \frac{n}{6} - n^2\right) \\ &= \frac{1}{3} - \frac{1}{2n} + \frac{1}{6n^2} \end{aligned}$$

So in summary

$$-\frac{1}{2n}+\frac{1}{6n^2}\le I-\frac{1}{3}\le\frac{1}{2n}+\frac{1}{6n^2} \tag{1.3}$$

for all natural numbers $n \ge 1$. The only number that satisfies this for all natural numbers $n \ge 1$ is 0. Therefore, $I - \frac{1}{3} = 0$ and the proposition follows. □

To rigorously prove the statement above, one needs to show that the only real number satisfying (1.3) for all $n \ge 1$ is 0. This is intuitively clear: indeed, taking n larger and larger, the two expressions $\frac{1}{2n}+\frac{1}{6n^2}$ and $-\frac{1}{2n}+\frac{1}{6n^2}$ get smaller and smaller. However, we cannot give a proof of it at this point, as we lack a rigorous definition of the real numbers.

As already mentioned, we have not answered the question of whether there is a notion of area for sets in $\mathbb{R}^2$. Nor have we described precisely what domains in $\mathbb{R}^2$ are, but we have implicitly assumed that domains are those subsets of $\mathbb{R}^2$ to which we can assign an area. The notions of the Riemann and Lebesgue integrals and measurable sets answer these fundamental questions.

The idea of the proof is illustrated in the following applet.

Applet 1.3 (Estimating an area). *We use up to 1000 rectangles to estimate the area from below and from above. In the proof below, however, we will use an unlimited number of rectangles and can thus determine the area exactly without any fuzziness.*

1 Note that in the previous examples, we informally used the notion of "set". Here, for completeness, we give a more precise definition.

Interlude: Naive Set Theory

The central assumptions of naive set theory are the following postulates (1), (2), and (4). Postulate (3), which we state here in addition, does not usually belong to them, but is a consequence of the so-called *axiom of regularity* in Zermelo-Frenkel set theory.

(1) A **set** consists of distinguishable elements.

(2) A set is distinctively determined by its **elements**.

(3) A set is not an element of itself.

(4) Every statement A about elements of a set X defines the set of elements in X for which the statement A is true; one writes $\{x \in X \mid A \text{ is true for } x\}$.

The **empty set**, written as $\varnothing$ (or sometimes also $\{\}$), is the set containing no elements.

We write "$x \in X$" if x is an element of the set X. If x is not an element of the set X, we write $x \notin X$. Sometimes we describe a set by a concrete list of its elements, for example, $X = \{x_1, x_2, \ldots, x_n\}$, but often it is more convenient to use postulate (4). For example, one can write

$$\{n \in \mathbb{Z} \mid \exists\, m \in \mathbb{Z} : n = 2m\}$$

1 to describe the set of even numbers. The symbol ∃ means "there exists", while the symbol ∀ means "for all". The symbols " | " and ":" in the formula above both mean "such that". With time, this mathematical terminology will become familiar, one just needs time and practice.

1.2 Tips on Studying

"All beginnings are difficult"

You will notice a big difference between school mathematics and university mathematics. The latter also uses its own language, which you will have to learn. The sooner you take this on, the more you will take away from the lectures. This brings us to the next tip.

"There is no silver bullet to mathematics"
(handed down quotation from Euclid to the Egyptian king Ptolemy I)

You cannot learn mathematics by watching it; just as you cannot learn tennis or skiing by watching all available tournaments or world championships on television. Rather, you should learn mathematics like a language, and a language is taught by using it. Discuss the topics of the lectures with colleagues. Explain to each other the proofs from the lecture or the solution of the exercise examples. Above all, solve as many exercises as possible; this is the only way to be sure that you have mastered the topics.

It is fine to work on the exercises in small groups. This even has the advantage that the group discussions make the objects of the lectures more lively. However, you should ensure that you fully understand the solutions, explain them and subsequently solve similar problems on your own.

"He who asks is a fool for a minute. He who does not ask is a fool all his life."
Confucius

Ask as many questions as you can and then ask them when they come up. Probably many of your colleagues have the same question, or have not even noticed the problem. This allows the lecturer or teaching assistant to simultaneously fix a problem for many and identify problems in students where she or he thought none existed. Furthermore, good question formulation needs to be practiced; the first year is the ideal time to do this.

Chapter 2

The Real Numbers: Maximum, Supremum, and Sequences

2.1 The Axioms of the Real Numbers

2.1.1 Ordered Fields

Although we all have an intuitive notion of real numbers, our aim is to introduce them properly.
2 To do so, we shall first introduce the notion of a field and then the concept of a relation.

We start by introducing some fundamental notions that will be covered in much more depth in the algebra lectures. Loosely speaking, a group is a set equipped with an “operation” that satisfies a list of properties. For our purposes, it is enough to know that an operation is something that takes two elements of a set and gives back a third element of the set. We now specify which properties the operation has to satisfy so that we can speak of a group.

Interlude: Groups

A **group** is a (non-empty) set G endowed with an operation "$\star$" that satisfy the following properties:

- (Associativity) for all $a, b, c \in G$, we have

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

- (Neutral element) there exists a neutral element, i.e., $e \in G$ such that for all $a \in G$ we have

$$a \star e = e \star a = a.$$

- (Inverse element) for each element $a \in G$ there is an inverse element, i.e., $a^{-1} \in G$ such that

$$a \star a^{-1} = a^{-1} \star a = e.$$

Note that, in general, we do not require that $a \star b = b \star a$ for arbitrary $a, b \in G$. If this property holds, the group is called **commutative** or **abelian**.

Example 2.1. — To make these concepts easier to access, let us assume for the moment
2 that we already know the natural numbers $\mathbb{N}$ and the integers $\mathbb{Z}$. We check whether they
satisfy the above properties if we replace $\star$ with the operations you are already familiar with.

1. Consider the natural numbers $\mathbb{N} = \{0, 1, 2, 3, ...\}$ with the usual addition $+$ that you probably know since primary school. This is

   - associative, since for any natural numbers $k, l, m \in \mathbb{N}$ we have

   $$(k + l) + m = k + (l + m),$$

   - has a neutral element $0 \in \mathbb{N}$, since for any natural number $n \in \mathbb{N}$ we have

   $$0 + n = n + 0 = n,$$

   but

   - no element apart from 0 has an inverse element. In fact, the inverse element of $n \in \mathbb{N} \backslash \{0\}$ would be $-n$, which is not included in the natural numbers $\mathbb{N}$.

2. The same arguments show that the integers $\mathbb{Z} = \{..., -3, -2, -1, 0, 1, 2, 3, ...\}$, again with the addition, form a group. Moreover, this is a commutative group, since for all integers $n, m \in \mathbb{Z}$ we have

$$n + m = m + n.$$

3. As a different example, consider the set of nonzero rational numbers

$$\mathbb{Q}^\times = \left\{ \frac{p}{q} \mid p, q \in \mathbb{Z},\ p, q \neq 0 \right\}$$

with the usual multiplication $\cdot$ between numbers. In this case, one can check that the multiplication is associative and commutative, the neutral element is 1, and the inverse of $\frac{p}{q}$ is $\frac{q}{p}$. Hence, this is a commutative group.

2.2. — It follows directly from the definition of the neutral element that it is unique. Indeed, assume that additionally to $e \in G$, we have a second element e' with the property such that $e' \star a = a \star e' = a$ for all elements $a \in G$. Then, we can choose $a = e$ and obtain

2

$$e = e \star e' = e',$$

where the first equality follows from the fact that e' is neutral, while in the second equality we used that that e is neutral.

We can thus speak of *the* neutral element of a group.

In the same spirit, assume that for an element $a \in G$, there exist two inverse elements a^{-1} and $\tilde{a}^{-1}$. Then, using associativity, we observe that

$$a^{-1} = a^{-1} \star e = a^{-1} \star (a \star \tilde{a}^{-1}) = (a^{-1} \star a) \star \tilde{a}^{-1} = e \star \tilde{a}^{-1} = \tilde{a}^{-1}.$$

So also for inverse elements, we might speak of *the* inverse element. In particular, since $a \star a^{-1} = e$, we deduce that a is the inverse of a^{-1}, thus

$$(a^{-1})^{-1} = a. \tag{2.1}$$

Interlude: Rings and Fields

A **ring** is a (non-empty) set R in which we can "add" and "multiply" elements in a compatible way. More precisely, a ring is a commutative group R whose first operation "+" is called "addition", and it is also equipped with an additional operation called "multiplication" and denoted by "·".

The neutral element for the addition is denoted by 0. The multiplication $\cdot$ is associative, has a neutral element (usually denoted by 1), and satisfies the following property:

- (Distributivity) for all $a, b, c \in R$, we have

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

  and

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

A ring in which the operation $\cdot$ is also commutative is called a **commutative ring**. Note that, in a ring, we do not require that elements have an inverse for the multiplication.

A commutative ring in which every "non-zero" element (that is, any element other than the neutral element 0 for the addition) has an inverse element for the multiplication $\cdot$
2 is called a **field**. In other words, a ring R is a field if $R \backslash \{0\}$ is a commutative group for the operation $\cdot$.

Fields will usually be denoted by the letter K for the corresponding German word "Körper". We write $K^\times = K \setminus \{0\}$ for the set of invertible elements in K.

Example 2.3. — Let us continue with our examples. We have already established that the integers $\mathbb{Z}$ form a commutative group, but are they also a ring with the usual multiplication? We must check:

- Associativity of the multiplication: For all integers $k, l, m \in \mathbb{Z}$, we have

$$(k \cdot l) \cdot m = k \cdot (l \cdot m).$$

- Neutral element for the multiplication: The neutral element for the multiplication is $1 \in \mathbb{Z}$ as, for all integers $k \in \mathbb{Z}$, we have

$$1 \cdot k = k \cdot 1 = k.$$

- Distributivity: For all $k, l, m \in \mathbb{Z}$ we have

$$k \cdot (l + m) = k \cdot l + k \cdot m$$

and

$$(k+l)\cdot m = k\cdot m + l\cdot m.$$

Hence, we conclude that $\mathbb{Z}$ is a ring. Moreover, since the multiplication is commutative (namely, $k\cdot l = l\cdot k$), it is a commutative ring. However, it is not a field since no element other than 1 and -1 has a multiplicative inverse. For example, the multiplicative inverse of 2 is $\frac{1}{2}$, which is not an integer.

Example 2.4. — The set of rational numbers $\mathbb{Q} = \left\{\frac{p}{q} \mid p, q \in \mathbb{Z},\ q \neq 0\right\}$ with the usual addition and multiplication is a field.

2.5. — Before going on, we look at some immediate consequences of the definition of field. In the current notation, $-a$ denotes the inverse of a with respect to the addition, while a^{-1} is the inverse of a with respect to the multiplication. Note that, in the current context, (2.1) implies that

$$-(-a) = a, \qquad \text{and} \quad (a^{-1})^{-1} = a \quad \text{whenever} \quad a \neq 0. \tag{2.2}$$

Let K be a field, and $a, b \in K$. Then the following holds:

(i) $0\cdot a = 0$ *and* $a\cdot 0 = 0$.
Proof: Since 0 is the neutral element for the addition, we have $0 = 0+0$. Hence, using
2 distributivity, we get

$$0\cdot a = (0+0)\cdot a = (0\cdot a) + (0\cdot a).$$

Adding $-0\cdot a$ (i.e., the inverse of $0\cdot a$ for the addition), we deduce that $0\cdot a = 0$. The case of $a\cdot 0$ is analogous.

(ii) $a\cdot(-b) = -(a\cdot b) = (-a)\cdot b$. *In particular, we have* $(-1)\cdot a = -a$.
Proof: By the distributive law, we have

$$a\cdot b + a\cdot(-b) = a\cdot(b+(-b)) = a\cdot 0 = 0.$$

So $a\cdot(-b)$ is the additive inverse of $a\cdot b$, i.e., $-(a\cdot b) = a\cdot(-b)$. Taking $b = 1$ gives $-a = (-1)\cdot a$.

The validity of $(-a)\cdot b = -(a\cdot b)$ follows exchanging a and b in the argument above.

(iii) $(-a)\cdot(-b) = a\cdot b$. *In particular, we have* $(-a)^{-1} = -(a^{-1})$.
Proof: By (ii) we know that $-(a\cdot b) = a\cdot(-b)$. Hence, recalling (2.2),

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

On the other hand, applying (ii) with $(-b)$ instead of b, we also have

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

Combining the two identities above, we conclude that $(-a)\cdot(-b) = a\cdot b$. Finally, taking $b = a^{-1}$ yields $(-a)\cdot(-(a^{-1})) = a \cdot a^{-1} = 1$, which gives the second assertion.

Remark 2.6. — A natural question one may ask is the following: *Is it possible to construct a field K where 0 (i.e., the neutral element for $+$) and 1 (i.e., the one for $\cdot$) are equal?* Assume that $0 = 1$. Then, using (i) above and the fact that 1 is the neutral element for multiplication, we get

$$0 = a \cdot 0 = a \cdot 1 = a$$

for every $a \in K$. So, the only possibility for having $0 = 1$ is that K consists of the single element 0. From now on, we shall assume that K always contains at least two elements, so in particular, 0 and 1 cannot coincide.

Next, we introduce the second ingredient of an ordered field, the order relation. Again we will do so in steps:

Interlude: Cartesian Product

Let X and Y be two sets. The **cartesian product** $X \times Y$ is the set of ordered pairs of elements in X and Y:

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

2

Example 2.7. — The cartesian product $X \times Y$ of $X = \{A, B, C, D, E, F, G, H\}$ and $Y = \{1, 2, 3, 4, 5, 6, 7, 8\}$ is what we use to write down positions on a chess board: To each pair we can associate a unique square on the chess board and vice versa. For instance, the black king starts the game on the square corresponding to $(E, 8) \in X \times Y$.

Interlude: Subsets

Let P and Q be sets.

- We say that P is **subset** of Q, and write $P \subset Q$ (or $P \subseteq Q$), if for all $x \in P$ also $x \in Q$ holds.
- We say that P is a **proper subset** of Q, and write $P \subsetneq Q$, if P is a subset of Q but not equal to Q.
- We write $P \not\subset Q$ (or $P \not\subseteq Q$) if P is not a subset of Q.

Equivalent formulations for “P is a subset of Q” are “P is **contained** in Q” and “Q is a **supset** of P”, which we also write as “$Q \supset P$”. The meaning of the statement “Q is a **proper superset** of P”, written $Q \supsetneq P$, is now implicit. Because of the second assumption of naive set theory, two sets P and Q are equal exactly if both $P \subset Q$ and $Q \subset P$ hold.

For instance, $\{x, y\} = \{z\}$ holds if $x = y = z$. Note that there are no "multiplicities" for elements of a set (for instance, $\{x, x, x\} = \{x\}$).

Interlude: Relations

Let X be a set. A **relation** on X is a subset $\mathcal{R} \subset X \times X$, that is, a list of ordered pairs of elements of X. We also write $x\mathcal{R}y$ if $(x, y) \in \mathcal{R}$ and often use symbols such as $<, \ll, \leq, \cong, \equiv, \sim$ for the relations.
If $\sim$ is a relation, we write "$x \not\sim y$" if "$x \sim y$" does not hold. A relation $\sim$ is called:

1. (Reflexive) for all $x \in X : x \sim x$.
2. (Transitive) for all $x, y, z \in X : x \sim y$ and $y \sim z \implies x \sim z$.
3. (Symmetric) for all $x, y \in X : x \sim y \implies y \sim x$.
4. (Antisymmetric) for all $x, y \in X : x \sim y$ and $y \sim x \implies x = y$.

A relation is called **equivalence relation** if it is reflexive, transitive, and symmetric.
A relation is called **order relation** if it is reflexive, transitive, and antisymmetric.

Example 2.8. — We look again at the integers $\mathbb{Z}$ and two examples of relations on them. Let $m, n, p \in \mathbb{Z}$ be integers.

2

- Consider the relation $\leq$ of being "less than or equal to", i.e., we write $n \leq m$ if n is less than or equal to m. Then we see that this relation is:

  1. reflexive because of the trivial observation that $n = n$,
  2. transitive since $n \leq m$ and $m \leq p$ implies $n \leq p$,
  3. not symmetric since for instance $7 \leq 8$ but $8 \not\leq 7$,
  4. antisymmetric since if n is less than or equal to m and vice versa, the only possibility is that they are equal.

  We conclude that $\leq$ is an order relation.

- Consider next the relation $<$ of being "strictly smaller than", i.e., we write $n < m$ if n and m are distinct integers and n is less than m. This relation is:

  1. not reflexive since no integer is strictly smaller than itself,
  2. transitive since $n < m$ and $m < p$ implies $n < p$,
  3. not symmetric since $3 < 5$ but $5 \not< 3$,
  4. antisymmetric. This point is a bit subtle: We need to check for all n, m such that $n < m$ and $m < n$, we have $n = m$. But there are no such n, m and hence the condition of antisymmetry is fulfilled because there is nothing to check.

We conclude that $<$ is neither an equivalence relation nor an order relation, since it does not satisfy the reflexivity property.

Definition 2.9: Ordered Field

Let K be a field, and let $\leq$ be an order relation on the set K. We call $(K, \leq)$, or K for short, an **ordered field** if the following conditions are satisfied:

1. (Linearity of the order) for all $x, y \in K$,

$$\text{at least one between } x \leq y \text{ and } y \leq x \text{ holds.}$$

2. (Compatibility of order and addition) for all $x, y, z \in K$, it holds

$$x \leq y \implies x + z \leq y + z.$$

3. (Compatibility of order and multiplication) for all $x, y \in K$, it holds

$$0 \leq x \text{ and } 0 \leq y \implies 0 \leq x \cdot y.$$

The following terminology is standard and will be used throughout these lecture notes:

- We pronounce $x \leq y$ as "x is less than or equal to y".

2

- For $x, y \in K$ we define $y \geq x$ by $x \leq y$, and pronounce this as "y is greater than or equal to x".
- We say that an element $x \in K$ is non-negative if $x \geq 0$, and non-positive if $x \leq 0$.
- Further, we define $x < y$ (pronounced as "x is smaller than y" or "x is strictly smaller than y") whenever $x \leq y$ and $x \neq y$.
- Analogously, we define $x > y$ when $y < x$, and say "x is greater than y" or "x is strictly greater than y".
- An element $x \in K$ is positive if $x > 0$ holds, and negative if $x < 0$ holds.

We often use these symbols in "equidirectional chains", for example, $x \leq y < z = a$ stands for

$$x \leq y \quad \text{and} \quad y < z \quad \text{and} \quad z = a.$$

Example 2.10. — A well-known example of an ordered field is the one of rational numbers $\mathbb{Q}$, together with the usual order relation given by

$$\frac{p}{q} \leq \frac{p'}{q'} \iff pq' \leq p'q, \qquad p, p' \in \mathbb{Z},\ q, q' \in \mathbb{N}.$$

Here on the right-hand side is the order on the integers, which we assume to be known.

2.11. — Let $(K, \leq)$ be an ordered field, and x, y, z, w denote elements of K. We want to prove a series of properties that follow from the definitions. To simplify the notation, it is customary to write $\cdot$ for multiplication only if it would otherwise be confusing. This is why, in proofs, $\cdot$ may disappear. For example, we may write xy instead of $x \cdot y$.

(a) (Trichotomy) *Either* $x < y$ *or* $x = y$ *or* $x > y$.
Proof: This follows directly from the linearity of the order relation $\leq$.

(b) *If* $x < y$ *and* $y \leq z$*, then* $x < z$ *also holds.*
Proof: First of all, since our assumption implies in particular that $x \leq y$ and $y \leq z$, we have $x \leq z$ according to the transitivity of the order relation.
We now want to prove $x < z$. Assume by contradiction that this is not the case. Then, since $x \leq z$, it must be $x = z$. Recalling that $y \leq z$, this proves that $y \leq x$. However, this contradicts $x < y$. Hence, we conclude that $x = z$ is impossible and, therefore, $x < z$.
Analogously, we see that $x \leq y$ *and* $y < z$ *imply that* $x < z$.

(c) (Addition of Inequalities) *If* $x \leq y$ *and* $z \leq w$ *hold, then* $x + z \leq y + w$ *also holds.*
Proof: Indeed, $x \leq y$ implies $x + z \leq y + z$ according to the additive compatibility in Definition 2.9, and $z \leq w$ implies $y + z \leq y + w$ for the same reason. Transitivity of the order relation implies $x + z \leq y + w$.
Analogously, using inference (b)*, one sees that* $x < y$ *and* $z \leq w$ *imply that* $y + z < y + w$.

3

(d) $x \leq y$ *is equivalent to* $0 \leq y - x$ *holds.*
Proof: If $x \leq y$, then adding $-x$ to both sides, we obtain $0 \leq y - x$.
Vice versa, if $0 \leq y - x$ then adding x to both sides we obtain $x \leq y$.

(e) $x \leq 0$ *is equivalent to* $0 \leq -x$.
Proof: This follows from (d) with $y = 0$.

(f) $x^2 \geq 0$*, and* $x^2 > 0$ *if* $x \neq 0$.
Proof: If $x \geq 0$, the first statement follows from the multiplicative compatibility in Definition 2.9.
If $x \leq 0$, then $-x \geq 0$ by inference (e) and hence $x^2 = (-x)^2 \geq 0$ (the fact that $x^2 = x \cdot x = (-x) \cdot (-x) = (-x)^2$ follows from (iii) in Paragraph 2.5).
Finally, we want to prove that $x^2 > 0$ whenever $x \neq 0$. Assume by contradiction that there exists $x \neq 0$ such that $x^2 = 0$. Since K is a field and $x \neq 0$, x^{-1} (the inverse of x for multiplication) exists. Therefore, recalling (i) in Paragraph 2.5 and using that $x^2 = 0$, we get

$$0 = 0 \cdot x^{-1} = x^2 \cdot x^{-1} = x \cdot x \cdot x^{-1} = x \cdot 1 = x,$$

hence $x = 0$. This contradicts our assumption $x \neq 0$ and concludes the proof.

(g) *It holds* $0 < 1$.
Proof: $1 = 1^2 \geq 0$ by inference (f) and $1 \neq 0$ (see Remark 2.6).

(h) *If* $0 \leq x$ *and* $y \leq z$*, then* $xy \leq xz$*.*
Proof: Using inference (d), according to which $z - y \geq 0$, and the multiplicative compatibility in Definition 2.9, $xz - xy = x(z - y) \geq 0$ holds, and thus the statement follows in turn from inference (d).

(i) *If* $x \leq 0$ *and* $y \leq z$*, then* $xy \geq xz$*.*
Proof: Note that $-x \geq 0$ by inference (e), and $z - y \geq 0$ by inference (d). Thus, by the multiplicative compatibility in Definition 2.9, we get

$$xy - xz = x(y - z) = (-x)(-(y - z)) = (-x)(z - y) \geq 0.$$

(j) *If* $0 < x \leq y$*, then* $0 < y^{-1} \leq x^{-1}$*.*
Proof: We first assert that $x^{-1} > 0$ ($y^{-1} > 0$ follows analogously). Indeed, if not, because of $x^{-1} \neq 0$ then trichotomy in inference (a) implies that $x^{-1} < 0$. Accordingly, $1 = xx^{-1} < 0$ would hold by inference (h), which contradicts (g).
Since $x^{-1} > 0$ and $y^{-1} > 0$, we deduce that $x^{-1}y^{-1} > 0$. Therefore, using (h),

$$y^{-1} = xx^{-1}y^{-1} \leq yx^{-1}y^{-1} = x^{-1}.$$

(k) *If* $0 \leq x \leq y$ *and* $0 \leq z \leq w$*, then* $0 \leq xz \leq yw$*.*
3 Proof: Exercise 2.12.

(l) *If* $x + y \leq x + z$*, then* $y \leq z$*.*
Proof: Exercise 2.12.

(m) *If* $xy \leq xz$ *and* $x > 0$*, then* $y \leq z$*.*
Proof: Exercise 2.12.

Exercise 2.12. — Prove the inferences (k),(l),(m). What happens in (m) when you drop the condition $x > 0$, that is, when $x < 0$ or $x = 0$? For some of the above inferences, formulate and prove similar versions for the strict relation "$<$".

2.13. — Let $(K, \leq)$ be an ordered field. As usual, we write $2, 3, 4, \ldots$ for the elements of K given by $2 = 1 + 1$, $3 = 2 + 1$, et cetera. By the compatibility of $+$ and $\leq$ in Definition 2.9, and recalling property (g) in Paragraph 2.11, the inequalities

$$\ldots < -2 < -1 < 0 < 1 < 2 < 3 < 4 < \ldots$$

hold in K. In particular, the elements $\ldots, -2, -1, 0, 1, 2, 3, \ldots$ of K are all distinct. We identify the set $\mathbb{Z}$ of integers with a subset of K. That is, we call the elements $\{\ldots, -2, -1, 0, 1, 2, 3, \ldots\}$ of K "integers". Consequently, we call elements $\{pq^{-1} \mid p, q \in \mathbb{Z}, q \neq 0\}$ in K rational numbers and thus identify $\mathbb{Q}$ with a subfield of K:

$$\mathbb{Z} \subsetneq \mathbb{Q} \subseteq K.$$

In other words, if $(K, \leq)$ is an ordered field, then it always includes a copy of the rationals inside it.

The above axioms, inferences, and statements in the exercises represent the usual properties for inequalities. We can also use them to solve problems like the one in the following exercise:

EXERCISE 2.14. — Show that

$$\{x \in \mathbb{R} \setminus \{0\} \mid x + \tfrac{3}{x} + 4 \geq 0\} = \{x \in \mathbb{R} \setminus \{0\} \mid -3 \leq x \leq -1 \text{ or } x > 0\}.$$

Hint: note that $x + \frac{3}{x} + 4 = \frac{(x+3)(x+1)}{x}$.

INTERLUDE: FUNCTIONS

A **function** f from a set X to a set Y is an assignment of an element of Y to each element of X. The element $y \in Y$ to which $x \in X$ is assigned is denoted $f(x)$. We write $f : X \to Y$ for a function from X to Y and sometimes also speak of a **map**, **mapping** or **transformation**.

3 We refer to the set X as **domain**, and the set Y as **domain of values** or **codomain**. The set $F = \{(x, f(x)) \mid x \in X\} \subset X \cdot Y$ is called the **graph** of f. In the context of a function $f : X \to Y$, an element x of the domain of definition is also called **argument**, and an element $y = f(x) \in Y$ assumed by the function is also called **value** of the function. If $f : X \to Y$ is a function, one also writes

$$\begin{aligned} f : X &\to Y \\ x &\mapsto f(x), \end{aligned}$$

where $f(x)$ could be a concrete formula. For example, $f : \mathbb{R} \to \mathbb{R}$ with $x \mapsto x^2$ is a fully defined function in this notation. We pronounce "$\mapsto$" as "is mapped to". Two functions $f_1 : X_1 \to Y_1$ and $f_2 : X_2 \to Y_2$ are said to be equal if $X_1 = X_2$, $Y_1 = Y_2$, and $f_1(x) = f_2(x)$ for all $x \in X_1$.

DEFINITION 2.15. — Let $(K, \leq)$ be an ordered field. The **absolute value** or **modulus** on K is the function $|\cdot| : K \to K$ given by

$$|x| = \begin{cases} x & \text{if } x \geq 0 \\ -x & \text{if } x < 0. \end{cases}$$

The **sign** is the function $\operatorname{sgn} : K \to \{-1, 0, 1\}$ given by

$$\operatorname{sgn}(x) = \begin{cases} -1 & \text{if } x < 0 \\ 0 & \text{if } x = 0 \\ 1 & \text{if } x > 0 \end{cases}.$$

2.16. — In what follows, let $(K, \leq)$ always be an ordered field, and x, y, z, w denote elements from K.

(a) *It holds* $x = \operatorname{sgn}(x) \cdot |x|$, *as well as* $|-x| = |x|$ *and* $\operatorname{sgn}(-x) = -\operatorname{sgn}(x)$.

(b) *It holds* $|x| \geq 0$, *and* $|x| = 0$ *exactly when* $x = 0$.
Proof: This follows from the trichotomy property.

(c) (Multiplicativity) *It holds* $\operatorname{sgn}(xy) = \operatorname{sgn}(x)\operatorname{sgn}(y)$ *and* $|xy| = |x||y|$.
To prove this, check all possible four cases (depending on whether x, y are negative or not).

(d) *If* $x \neq 0$, *then* $|x^{-1}| = |x|^{-1}$ *holds.*
Proof: This follows from (c) because of $|x^{-1}||x| = 1$.

(e) $|x| \leq y$ *is equivalent to* $-y \leq x \leq y$.
3 Proof: First we note that, in both cases, y is a non-negative number.
Suppose first $|x| \leq y$. If $x \geq 0$ then $-y \leq 0 \leq x = |x| \leq y$.
If $x < 0$ then $-y \leq -|x| = x < 0 \leq y$ and so again $-y \leq x \leq y$.
Suppose now $-y \leq x \leq y$. If $x \geq 0$ then $|x| = x \leq y$.
If $x < 0$ then we observe that the inequality $-y \leq x$ is equivalent to $-x \leq y$. Hence $|x| = -x \leq y$ also in this case.

(f) $|x| < y$ *is equivalent to* $-y < x < y$.
This is proved by arguing as in (e).

(g) (Triangle inequality) *It holds that*

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

Proof: Note that, by (e), we have $-|x| \leq x \leq |x|$ and $-|y| \leq y \leq |y|$. Adding these two inequalities, we get

$$-(|x| + |y|) \leq x + y \leq |x| + |y|.$$

So, by property (e), $|x + y| \leq |x| + |y|$.

(h) (Inverse triangle inequality) *It holds that* $\big||x| - |y|\big| \leq |x - y|$.
Proof: The triangle inequality in (g) shows that $|x| \leq |x - y + y| \leq |x - y| + |y|$, which

leads to $|x| - |y| \leq |x - y|$. Exchanging x and y, we get $|y| - |x| \leq |x - y|$. So, by property (e), $\big||x| - |y|\big| \leq |x - y|$, as desired.

3

EXERCISE 2.17. — For which $x, y \in \mathbb{R}$ does equality hold in the triangle inequality? And in the inverse triangle inequality?

2.1.2 Axiom of Completeness

To do calculus, ordered fields are generally unsuitable because you can have "gaps". Indeed, think of the rational numbers $\mathbb{Q}$ in $\mathbb{R}$. The field of rational numbers with the usual order is ordered. So we need more properties, or another axiom, to do calculus in an ordered field. This axiom is the so-called "completeness axiom". To a certain extent, the search for this axiom began with the work of Greeks such as Pythagoras, Euclid, and Archimedes, but it did not become successful until the 19th century in the work of numerous mathematicians, including Weierstrass, Heine, Cantor, and Dedekind (see also this link).

DEFINITION 2.18: COMPLETENESS AXIOM

Let $(K, \leq)$ be an ordered field. We say $(K, \leq)$ is **complete** or a **completely ordered field** if the statement (V) is true.

(V) Let X, Y be non-empty subsets of K such that for all $x \in X$ and $y \in Y$ the inequality $x \leq y$ holds. Then there exists $c \in K$ lying between X and Y, in the sense that for all $x \in X$ and $y \in Y$ the inequality $x \leq c \leq y$ holds.

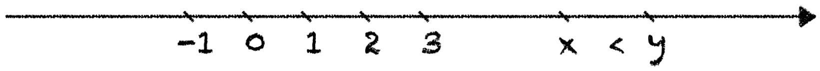
4

We call statement (V) the **completeness axiom**.

DEFINITION 2.19: REAL NUMBERS

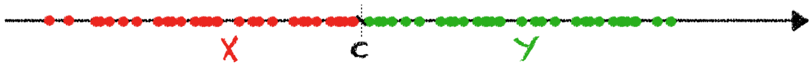
We call **field of real numbers** any completely ordered field. Such a field is denoted with the symbol $\mathbb{R}$.

2.20. — We will often visualise the real numbers as the points on a straight line, which is why we also call it the **number line**.



We interpret the relation $x < y$ for $x, y \in \mathbb{R}$ as "on the straight line, the point y lies to the right of the point x". What does the completeness axiom mean in this picture?

Let X, Y be non-empty subsets of $\mathbb{R}$ such that for all $x \in X$ and all $y \in Y$ the inequality $x \leq y$ holds. Then all elements of X are to the left of all elements of Y as in the following figure.



So, according to the completeness axiom, there exists a number c that lies in between. The existence of the number c is, in a sense, an assurance that $\mathbb{R}$ has no "gaps". It is advisable to visualize definitions, statements, and their proofs on the number line. However, the number line should always be used only as a motivation and to develop a good intuition, but not for rigorous proof.

Interlude: Injective, surjective and bijective functions

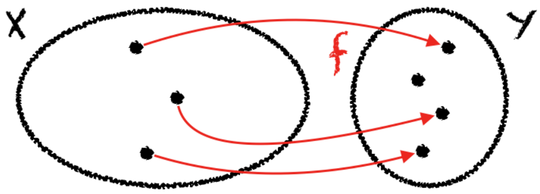
Let $f : X \to Y$ be a function. We call f:

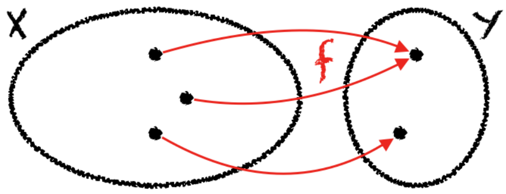
1. **injective** or an **injection** if $f(x_1) = f(x_2) \implies x_1 = x_2$ for all $x_1, x_2 \in X$,
2. **surjective** or a **surjection** if for every $y \in Y$ there exists an $x \in X$ with $f(x) = y$ and
3. **bijective** or a **bijection** if it is both surjective and injective.

Thus, a function $f : X \to Y$ is not injective if there exist two distinct elements $x_1 \neq x_2 \in X$ with $f(x_1) = f(x_2)$, and not surjective if there exists a $y \in Y$ such that $f(x) \neq y$ holds for all $x \in X$.

4

In the following image, an injective function that is not surjective is shown on the left, and a surjective function that is not injective is shown on the right.





Interlude: Image and Preimage of a Function

Definition 2.21. — For a function $f : X \to Y$ and a subset $A \subset X$ we write

$$f(A) = \{y \in Y \mid \exists\, x \in A : f(x) = y\}$$

and call this subset of Y the **image** of A under the function f. For a subset $B \subset Y$ we write

$$f^{-1}(B) = \{x \in X \mid \exists\, y \in B : f(x) = y\}$$

and call this subset of X the **preimage** of B under the function f.

REMARK 2.22. — Saying that a function $f : X \to Y$ is surjective is equivalent to saying that $f(X) = Y$ (i.e., every element of Y is in the image of X under f).

EXAMPLE 2.23. — Let $f : \mathbb{R} \to \mathbb{R}$ be the constant function $x \mapsto 0$ for every $x \in \mathbb{R}$. Then $f^{-1}(\{0\}) = \mathbb{R}$, while $f^{-1}(\{y\}) = \emptyset$ for every $y \neq 0$.

EXAMPLE 2.24. — Let X, Y be two finite sets with the same number of elements (for example, X and Y could be the same set). Then, for a function $f : X \to Y$, injectivity and surjectivity are equivalent.
To show this, assume that X and Y have n elements and write $X = \{x_1, \ldots, x_n\}$. Suppose first that f is injective. Then all the elements $f(x_i)$ are distinct, which means that the set $f(X) = \{f(x_1), \ldots, f(x_n)\}$ also has n elements. Being $f(X)$ a subset of Y and Y has n elements, the only option is that $f(X) = Y$. This proves that injectivity implies surjectivity.
Conversely, to show that surjectivity implies injectivity, we prove that if f is not injective then f is not surjective. So, assume there exist at least two elements $x_i \neq x_j$ such that $f(x_i) = f(x_j)$. This means that $f(X)$ has at most $n - 1$ elements, so f cannot be surjective.

REMARK 2.25. — For infinite sets, injectivity and surjectivity are not necessarily equivalent. Consider for instance the functions $f_1, f_2 : \mathbb{N} \to \mathbb{N}$ defined as

4

$$f_1(n) = n + 1, \qquad f_2(n) = \begin{cases} 0 & \text{if } n = 0, \\ n - 1 & \text{if } n \geq 1. \end{cases}$$

Then f_1 is injective but not surjective, while f_2 is surjective but not injective.

EXERCISE 2.26. — Reformulate the definitions of injectivity, subjectivity, and bijectivity using the notions of image and preimage of a function.

We conclude this section by introducing the **square root function** on $\mathbb{R}_{\geq 0} = \{x \in \mathbb{R} : x \geq 0\}$ as an application of the completeness axiom. We formulate this as an exercise for the reader.

EXERCISE 2.27. — In this exercise, we show the existence and uniqueness of a bijective function $\sqrt{\cdot} : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ with property $(\sqrt{a})^2 = a$ for all $a \in \mathbb{R}_{\geq 0}$.

1. Show that, for all $x, y \in \mathbb{R}_{\geq 0}$: $x < y$ is equivalent to $x^2 < y^2$.

2. Use Step 1 to deduce that, for every $a \in \mathbb{R}_{\geq 0}$, there can exist at most one element $c \in \mathbb{R}_{\geq 0}$ satisfying $c^2 = a$.

3. For a real number $a \in \mathbb{R}_{\geq 0}$ consider the non-empty subsets

$$X = \{x \in \mathbb{R}_{\geq 0} \mid x^2 \leq a\}, \qquad Y = \{y \in \mathbb{R}_{\geq 0} \mid y^2 \geq a\},$$

and apply the completeness axiom to find $c \in \mathbb{R}$ with $x \leq c \leq y$ for all $x \in X$ and $y \in Y$. Prove that $c \in X$ and $c \in Y$ to conclude that both $c^2 \leq a$ and $c^2 \geq a$ hold, thus $c^2 = a$.
Hint: If by contradiction $c \notin X$ (that is, $c^2 > a$), then one can find a suitably small real number $\varepsilon > 0$ such that $(c - \varepsilon)^2 \geq a$. Thus $c - \varepsilon \in Y$, which contradicts $y \geq c$ for every $y \in Y$. The case of $c \notin Y$ is analogous.

We call **square root function** the function $\sqrt{\cdot} : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ that assigns to each $a \in \mathbb{R}_{\geq 0}$ the number $c \in \mathbb{R}_{\geq 0}$ uniquely determined by the above construction. We note that $c^2 = a$, and we call $c = \sqrt{a}$ the **square root** of a. Show that:

4. The function $\sqrt{\cdot}$ is increasing: for $x, y \in \mathbb{R}_{\geq 0}$ with $x < y$, the inequality $\sqrt{x} < \sqrt{y}$ holds.

5. The function $\sqrt{\cdot} : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is bijective.

6. For all $x, y \in \mathbb{R}_{\geq 0}$, $\sqrt{xy} = \sqrt{x}\sqrt{y}$.

4 Exercise 2.28. — For all $x \in \mathbb{R}$, show that $x^2 = |x|^2$ and $\sqrt{x^2} = |x|$.

2.29. — In summary, in a field of real numbers as defined in Definition 2.19, the usual arithmetic rules and equation transformations work, although (as usual) division by zero is not defined. Furthermore, the relations $\leq$ and $<$ satisfy the usual transformation laws for inequalities. In particular, when multiplying by negative numbers, the inequalities must be reversed. We will use these laws in the following without reference. We will see the deep meaning of the completeness axiom when we use it for further statements. In particular, until further notice, we will always refer to it when we use it.

It is not clear for the moment that there is indeed a field of real numbers as in Definition 2.19. The fact that we occasionally even speak of *the* real numbers stems from the fact that, except for certain identifications, there is only one completely ordered field. In this course, we assume, in agreement with your high school experience, that a field of real numbers exists and is unique.

2.1.3 Intervals

Definition 2.30: Intervals

Let $a, b \in \mathbb{R}$. We define:

- the **closed interval** $[a, b]$ as
$$[a, b] = \{x \in \mathbb{R} \mid a \leq x \leq b\}.$$
- The **open interval** (a, b) as
$$(a, b) = \{x \in \mathbb{R} \mid a < x < b\}.$$
- The **half-open intervals** $[a, b)$ and $(a, b]$:
$$[a, b) = \{x \in \mathbb{R} \mid a \leq x < b\} \qquad \text{and} \qquad (a, b] = \{x \in \mathbb{R} \mid a < x \leq b\}.$$

4

- The **unbounded closed intervals**
$$[a, \infty) = \{x \in \mathbb{R} \mid a \leq x\} \qquad \text{and} \qquad (-\infty, b] = \{x \in \mathbb{R} \mid x \leq b\}$$
as well as the **unbounded open intervals**
$$(a, \infty) = \{x \in \mathbb{R} \mid a < x\} \qquad \text{and} \qquad (-\infty, b) = \{x \in \mathbb{R} \mid x < b\}.$$

2.31. — The intervals $(a, b], [a, b), (a, b)$ for $a, b \in \mathbb{R}$ are non-empty exactly when $a < b$, and $[a, b]$ is non-empty exactly when $a \leq b$. If the interval is non-empty, a is called the **left endpoint**, b is called the **right endpoint** and $b - a$ is called the **length of the interval**. Intervals of the kind $[a, b]$, $(a, b]$, $[a, b)$, (a, b) for $a, b \in \mathbb{R}$ are also called **bounded intervals** if we want to distinguish them from the unbounded intervals.
Instead of round brackets, inverted square brackets are sometimes used to denote open and half-open intervals. For example, instead of (a, b) for $a, b \in \mathbb{R}$, one can also find $]a, b[$ in the literature. We will always use round brackets here.

Interlude: Set operations

Let P and Q be sets. The **intersection** $P \cap Q$, the **union** $P \cup Q$, the **relative complement** $P \setminus Q$, and the **symmetric difference** $P \triangle Q$ are defined as by

$$
\begin{aligned}
P \cap Q &= \{x \mid x \in P \text{ and } x \in Q\} \\
P \cup Q &= \{x \mid x \in P \text{ or } x \in Q\} \\
P \setminus Q &= \{x \mid x \in P \text{ and } x \notin Q\} \\
P \triangle Q &= (P \cup Q) \setminus (P \cap Q).
\end{aligned}
$$

These definitions are illustrated in the following pictures. Sketches of this kind are called **Venn diagrams**.

5

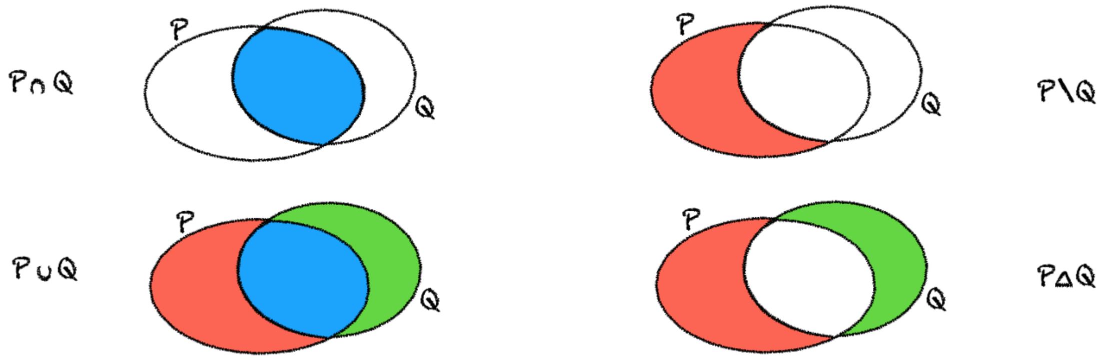


Figure 2.1: Illustration of set operations.

If it is clear from the context that all sets under consideration are subsets of a given basic set X, then the **complement** P^c of P is defined by $P^c = X \setminus P$.

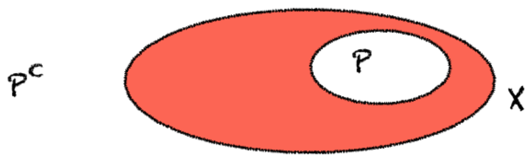


Figure 2.2: The complement $P^c = X \setminus P$ of P in X.

Interlude: Union and Intersection of several sets

Let $\mathcal{A}$ be a family of sets, that is, a set whose elements are themselves sets. Then we define the **union** and the **intersection** of the sets in $\mathcal{A}$ as

$$\bigcup_{A\in\mathcal{A}} A = \{x \mid \exists\, A \in \mathcal{A} : x \in A\}\ , \qquad \bigcap_{A\in\mathcal{A}} A = \{x \mid \forall\, A \in \mathcal{A} : x \in A\}.$$

If $\mathcal{A} = \{A_1, A_2, \ldots\}$, then we also write

$$\bigcup_{n=1}^{\infty} A_n = \{x \mid \exists\, n \geq 1 : x \in A_n\}\ , \qquad \bigcap_{n=1}^{\infty} A_n = \{x \mid \forall\, n \geq 1 : x \in A_n\}$$

for the union and the intersection of the sets in $\mathcal{A}$.

Exercise 2.32. — 1. Show that a finite intersection of intervals is again an interval. Can you describe the endpoints of a non-empty intersection using the endpoints of the original intervals?

2. When is a union of two intervals an interval again? In this case, what happens when you unite two intervals of the same type (open, closed, half-open)?

Definition 2.33: Neighborhoods

5 Let $x \in \mathbb{R}$. A neighbourhood of x is a set containing an open interval I such that $x \in I$. Given $\delta > 0$, the open interval $(x-\delta, x+\delta)$ is called the **δ-neighbourhood** of x.

2.34. — For example, both $[-1, 1]$ and $\mathbb{Q} \cup [-1, 1]$ are neighborhoods of $0 \in \mathbb{R}$ (since they both contain, for instance, $(-1/2, 1/2)$), but $[0, 1]$ is not a neighbourhood of 0.
We note further that, for $\delta > 0$ and $x \in \mathbb{R}$, the δ-neighbourhood of x is given by $\{y \in \mathbb{R} \mid |x-y| < \delta\}$. We will interpret $|x-y|$ as the **distance** from x to y. In terms of "distance", a few of the above inferences can be re-expressed more intuitively. For example, property (a) in Paragraph 2.16 implies that, for $x, y \in \mathbb{R}$, the equality $|x-y| = |-(x-y)| = |y-x|$ holds. In other words, the distance from x to y is equal to the distance from y to x.

Definition 2.35: Open and Closed Sets

A subset $U \subseteq \mathbb{R}$ is called **open** in $\mathbb{R}$ if for every $x \in U$ there exists an open interval I such that $x \in I$ and $I \subset U$.
A subset $F \subseteq \mathbb{R}$ is called **closed** in $\mathbb{R}$ if its complement $\mathbb{R} \setminus F$ is open.

2.36. — Open intervals are open, closed intervals are closed. Intuitively, a subset is open if, for any point x in the set, all points close enough to x are also in the set. Contrary to conventional usage, "open" is not the opposite of "closed".

The sets $\emptyset$ and $\mathbb{R}$ are both open in $\mathbb{R}$. Hence, they are also closed since $\emptyset = \mathbb{R}^c$ and $\mathbb{R} = \emptyset^c$. We note that $\mathbb{Q} \subseteq \mathbb{R}$ and $[a, b) \subset \mathbb{R}$ are neither open nor closed.

Exercise 2.37. — Show that a subset $U \subseteq \mathbb{R}$ is open exactly if, for every element $x \in U$, there exists $\delta > 0$ such that $(x - \delta, x + \delta) \subseteq U$.

5

Exercise 2.38. — Let $\mathcal{U}$ be a family of open sets, and $\mathcal{F}$ be a family of closed subsets of $\mathbb{R}$. Show that the union and the intersection

$$\bigcup_{U \in \mathcal{U}} U \qquad \text{and} \qquad \bigcap_{F \in \mathcal{F}} F$$

are open and closed, respectively.

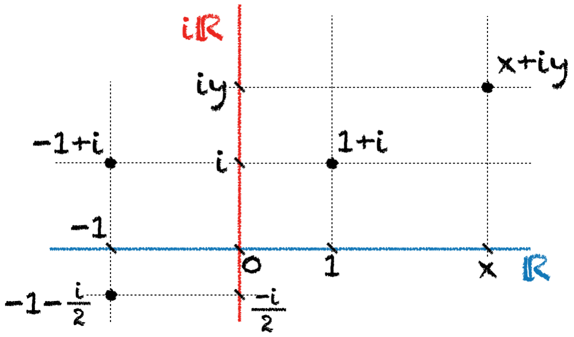
2.2 Complex Numbers

2.2.1 Definition of Complex Numbers

Using a field of real numbers $\mathbb{R}$, we can define the set of complex numbers as

$$\mathbb{C} = \mathbb{R}^2 = \{(x, y) \mid x, y \in \mathbb{R}\}.$$

We call elements $z = (x, y) \in \mathbb{C}$ complex numbers, and will write them in the form $z = x + iy$, where the symbol i is called the **imaginary unit**. Note that in this identification, the symbol $+$ is, for the time being, to be understood as a substitute for the comma. The number $x \in \mathbb{R}$ is called the **real part** of z and one writes $x = \mathrm{Re}(z)$; the number $y \in \mathbb{R}$ is the **imaginary part** of z and one writes $y = \mathrm{Im}(z)$. The elements of $\mathbb{C}$ with imaginary part 0 are also called **real**, and the elements with real part 0 are called **purely imaginary**. Via the injective mapping $x \in \mathbb{R} \mapsto x + i0 \in \mathbb{C}$ we identify $\mathbb{R}$ with the subset of real elements of $\mathbb{C}$.



5

The graphical representation of the set $\mathbb{C}$ is called the **complex plane** or also **Gaussian number plane**. From this geometric point of view, the set of real points is called the **real axis** and the set of purely imaginary points is called the **imaginary axis**.

As you might expect from previous knowledge, i should corresponds to a square root of -1. Hence, we want to define an addition and a multiplication on the set $\mathbb{C}$ so that the set $\mathbb{C}$ together with these operations is a field in which $i^2 = -1$ holds. Note that, if $i^2 = -1$, then it follows from commutativity and distributivity that

$$(x_1 + iy_1) \cdot (x_2 + iy_2) = x_1x_2 + ix_1y_2 + iy_1x_2 + i^2y_1y_2 = (x_1x_2 - y_1y_2) + i(x_1y_2 + y_1x_2).$$

This leads us to the following definition:

Definition 2.39. — We call **addition** and **multiplication** on the set $\mathbb{C} = \mathbb{R} \times \mathbb{R}$ the following operations:

$$\begin{aligned} (x_1, y_1) + (x_2, y_2) &= (x_1 + x_2, y_1 + y_2) && \text{(addition)} \\ (x_1, y_1) \cdot (x_2, y_2) &= (x_1x_2 - y_1y_2, x_1y_2 + x_2y_1) && \text{(multiplication)} \end{aligned}$$

PROPOSITION 2.40: $\mathbb{C}$ IS A FIELD

The set $\mathbb{C} = \mathbb{R} \times \mathbb{R}$, together with the zero element $(0,0)$, the one element $(1,0)$, and the operations introduced in Definition 2.39, is a field.

Proof. We review the axioms of fields. The associativity and commutativity of the addition, and the fact that $(0,0)$ is a neutral element for the addition, are direct consequences of the corresponding properties of the addition of real numbers. The inverse element of (x,y) for the addition is given by $(-x,-y)$, since

$$(x,y) + (-x,-y) = (x-x, y-y) = (0,0).$$

Proving the properties of multiplication requires a little more effort. We start with associativity of multiplication: let $(x_1,y_1),(x_2,y_2)$, and (x_3,y_3) be elements of $\mathbb{C}$. Now calculate

$$\begin{aligned}\big((x_1,y_1)\cdot(x_2,y_2)\big)\cdot(x_3,y_3) &= (x_1x_2 - y_1y_2, x_1y_2 + y_1x_2)\cdot(x_3,y_3)\\ &= (x_1x_2x_3 - y_1y_2x_3 - x_1y_2y_3 - y_1x_2y_3, x_1y_2x_3 + y_1x_2x_3 + x_1x_2y_3 - y_1y_2y_3).\end{aligned}$$

Analogously, we calculate

5

$$\begin{aligned}(x_1,y_1)\cdot\big((x_2,y_2)\cdot(x_3,y_3)\big) &= (x_1,y_1)\cdot(x_2x_3 - y_2y_3, x_2y_3 + y_2x_3)\\ &= (x_1x_2x_3 - y_1y_2x_3 - x_1y_2y_3 - y_1x_2y_3, x_1y_2x_3 + y_1x_2x_3 + x_1x_2y_3 - y_1y_2y_3).\end{aligned}$$

This proves that the multiplication in Definition 2.39 is associative.

The commutativity of multiplication can also be shown by direct computation:

$$(x_1,y_1)\cdot(x_2,y_2) = (x_1x_2 - y_1y_2, x_1y_2 + x_2y_1) = (x_2,y_2)\cdot(x_1,y_1).$$

Also, in the same way, we check that $(1,0)$ is the neutral element for multiplication:

$$(x,y)\cdot(1,0) = (x\cdot 1 - y\cdot 0, x\cdot 0 + y\cdot 1) = (x,y).$$

Next, we check the distributivity law: Let again $(x_1,y_1),(x_2,y_2)$, and (x_3,y_3) be elements of $\mathbb{C}$. Then

$$\begin{aligned}(x_1,y_1)\cdot\big((x_2,y_2)+(x_3,y_3)\big) &= (x_1,y_1)\cdot(x_2+x_3, y_2+y_3)\\ &= (x_1x_2 + x_1x_3 - y_1y_2 - y_1y_3, y_1x_2 + y_1x_3 + x_1y_2 + x_1y_3)\\ &= (x_1x_2 - y_1y_2, y_1x_2 + x_1y_2) + (x_1x_3 - y_1y_3, y_1x_3 + x_1y_3)\\ &= (x_1,y_1)\cdot(x_2,y_2) + (x_1,y_1)\cdot(x_3,y_3),\end{aligned}$$

which shows that $\mathbb{C}$ is a ring when endowed with the addition and multiplication given in Definition 2.39.

To finish the proof, we still need to show the existence of multiplicative inverses. Let $(x,y) \in \mathbb{C}$ be such that $(x,y) \neq (0,0)$. So either $x \neq 0$ or $y \neq 0$ holds, and therefore $x^2+y^2 > 0$. Then the multiplicative inverse of (x,y) is given by $\left(\frac{x}{x^2+y^2}, \frac{-y}{x^2+y^2}\right)$, because

$$\begin{aligned}(x,y)\cdot\left(\tfrac{x}{x^2+y^2}, \tfrac{-y}{x^2+y^2}\right) &= \left(x\cdot\tfrac{x}{x^2+y^2} - y\cdot\tfrac{-y}{x^2+y^2}, y\cdot\tfrac{x}{x^2+y^2} + x\cdot\tfrac{-y}{x^2+y^2}\right)\\ &= \left(\tfrac{x^2+y^2}{x^2+y^2}, \tfrac{yx-xy}{x^2+y^2}\right) = (1,0).\end{aligned}$$

□

Applet 2.41 (Complex numbers)**.** *We consider the field operations (addition, multiplication, multiplicative inverse) on the complex numbers. The true geometric meaning of multiplication and of the multiplicative inverse will be discussed later.*

As already explained, we do not write (x,y) for complex numbers, but $x+iy$. Instead of $x+i0$ we also simply write x, and instead of $0+iy$ we write iy, and finally we also write i for $i1$. By construction, $i^2 = -1$ holds. According to this notation, take $\mathbb{R}$ to be a subset of $\mathbb{C}$. This makes sense since addition and multiplication on $\mathbb{C}$ restricted to $\mathbb{R}$ is equivalent to addition and multiplication on $\mathbb{R}$. Also, given $z,w \in \mathbb{C}$, we write zw for $z\cdot w$.

Given $z \neq 0$, we shall use both z^{-1} and $\frac{1}{z}$ to denote the inverse of z for the multiplication. For instance, $i^{-1} = \frac{1}{i} = -i$ (since $(-i)\cdot i = -i^2 = 1$).

5 **Definition 2.42: Complex Conjugation**

Let $z = x+iy$ be a complex number. We call $\overline{z} = x-iy$ the complex number **conjugate** to z. The mapping $\mathbb{C} \to \mathbb{C}$ given by $z \mapsto \overline{z}$ is called **complex conjugation**.

Lemma 2.43: Properties of Complex Conjugation

The complex conjugation satisfies the following properties:

1. *For all $z \in \mathbb{C}$, $z\overline{z} \in \mathbb{R}$ and $z\overline{z} \geq 0$. Furthermore, for all $z \in \mathbb{C}$, $z\overline{z} = 0$ holds exactly when $z = 0$.*
2. *For all $z,w \in \mathbb{C}$, $\overline{z+w} = \overline{z} + \overline{w}$ holds.*
3. *For all $z,w \in \mathbb{C}$, $\overline{z\cdot w} = \overline{z}\cdot\overline{w}$ holds.*

Proof. Part (1) follows from the fact that, for $z = x+iy$,

$$z\overline{z} = (x+iy)(x-iy) = x^2+y^2.$$

To show parts (2) and (3), we write $z = x_1+iy_1$ and $w = x_2+iy_2$. Then $z+w = (x_1+x_2)+i(y_1+y_2)$ and we get

$$\overline{z+w} = (x_1+x_2) - i(y_1+y_2) = (x_1-iy_1)+(x_2-iy_2) = \overline{z}+\overline{w}.$$

Analogously, since $z \cdot w = (x_1x_2 - y_1y_2) + i(x_1y_2 + y_1x_2)$ we have

$$\overline{z \cdot w} = (x_1x_2 - y_1y_2) - i(x_1y_2 + y_1x_2) = (x_1 - iy_1) \cdot (x_2 - iy_2) = \overline{z} \cdot \overline{w},$$

as desired. □

Exercise 2.44. — Show the identities

$$\mathrm{Re}(z) = \frac{z + \overline{z}}{2} \qquad \text{and} \qquad \mathrm{Im}(z) = \frac{z - \overline{z}}{2i}$$

for all $z \in \mathbb{C}$. In particular, conclude that $\mathbb{R} = \{z \in \mathbb{C} \mid z = \overline{z}\}$. Can you interpret these equalities geometrically?

2.45. — Since $i^2 = -1 < 0$, property (f) in Paragraph 2.11 implies that no order compatible with addition and multiplication can be defined on $\mathbb{C}$. Nevertheless, calculus can be performed on the complex numbers, which is partly addressed in this course but mainly in the course on *complex analysis* in the second year of the study of mathematics and physics. The reason for this is that $\mathbb{C}$ satisfies a generalisation of the completeness axiom, which we can only discuss after some more theory.

5

2.2.2 The Absolute Value on the Complex Numbers

Since there is no ordering relation on the field of complex numbers that would make it an ordered field, we cannot use Definition 2.15 to define an absolute value on $\mathbb{C}$. We still want to define an absolute value on $\mathbb{C}$ in such a way that it has as many properties as possible of the absolute value on $\mathbb{R}$ and is compatible with it. To do this, we use the root function introduced in Exercise 2.27.

Definition 2.46: Absolute Value

The **absolute value** or **norm** on $\mathbb{C}$ is the function $|\cdot| : \mathbb{C} \to \mathbb{R}$ given by

$$|z| = \sqrt{z\overline{z}} = \sqrt{x^2 + y^2}$$

for $z = x + iy \in \mathbb{C}$.

At this point we note that, given $x \in \mathbb{R}$, the absolute value $|x| = \mathrm{sgn}(x) \cdot x$ and the absolute value of x as an element of $\mathbb{C}$ coincide, since $\sqrt{x\overline{x}} = \sqrt{x^2} = |x|$ holds. In particular, the newly introduced notation is consistent, and we have extended the absolute value of $\mathbb{R}$ to $\mathbb{C}$.

Note that $|z| \geq 0$ for all $z \in \mathbb{C}$, and $|z| = 0$ exactly when $z = 0$. Also, the absolute value on $\mathbb{C}$ is multiplicative, namely

$$|zw| = \sqrt{zw\overline{zw}} = \sqrt{z\bar{z}w\overline{w}} = \sqrt{z\bar{z}}\sqrt{w\overline{w}} = |z||w| \qquad \text{for all } z, w \in \mathbb{C}.$$

Furthermore,

$$z^{-1} = \frac{\bar{z}}{|z|^2} \qquad \text{for all } z \neq 0.$$

These are essential consequences of Lemma 2.43. Finally, the triangle inequality holds, as shown in the next proposition.

Proposition 2.47: Triangle Inequality

For all $z, w \in \mathbb{C}$, $|z + w| \leq |z| + |w|$ holds.

To prove this result, we first need the following:

Lemma 2.48: Cauchy-Schwarz Inequality

For all $z = x_1 + iy_1, w = x_2 + iy_2 \in \mathbb{C}$,

$$x_1x_2 + y_1y_2 \leq |z||w|. \tag{2.3}$$

6

Proof. We begin by observing that

$$\begin{aligned} |z|^2|w|^2 - (x_1x_2 + y_1y_2)^2 &= (x_1^2 + y_1^2)(x_2^2 + y_2^2) - (x_1x_2 + y_1y_2)^2 \\ &= x_1^2x_2^2 + y_1^2y_2^2 + y_1^2x_2^2 + x_1^2y_2^2 - (x_1^2x_2^2 + y_1^2y_2^2 + 2x_1x_2y_1y_2) \\ &= y_1^2x_2^2 + x_1^2y_2^2 - 2x_1x_2y_1y_2 \\ &= (y_1x_2 - x_1y_2)^2 \geq 0. \end{aligned}$$

This proves that

$$(x_1x_2 + y_1y_2)^2 \leq |z|^2|w|^2.$$

Taking the square root on both sides and recalling Exercise 2.28, we get $|x_1x_2 + y_1y_2| \leq |z||w|$ which implies (2.3) (recall that $x \leq |x|$ for any $x \in \mathbb{R}$). □

Proof of Proposition 2.47. We write $z = x_1 + iy_1$ and $w = x_2 + iy_2$. Then, since $z + w = (x_1 + x_2) + i(y_1 + y_2)$, we have

$$\begin{aligned} |z + w|^2 &= (x_1 + x_2)^2 + (y_1 + y_2)^2 \\ &= x_1^2 + x_2^2 + y_1^2 + y_2^2 + 2(x_1x_2 + y_1y_2) \\ &= |z|^2 + |w|^2 + 2(x_1x_2 + y_1y_2). \end{aligned}$$

Hence, using (2.3), we get

$$|z+w|^2 = |z|^2 + |w|^2 + 2(x_1x_2 + y_1y_2) \leq |z|^2 + |w|^2 + 2|z||w| = (|z| + |w|)^2.$$

Taking the square root on both sides concludes the proof. □

2.49. — The absolute value of the complex number $z = x + iy$ is the square root of $x^2 + y^2$ and, in the geometric notion of complex numbers, is equivalent to the length of the straight line from the origin $0 + i0$ to z. In the same way, for two complex numbers z and w, we interpret $|z - w|$ as the **distance** from z to w.

Definition 2.50: Circular Disks

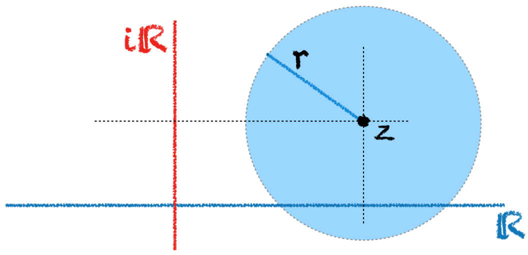
The **open circular disk** with radius $r > 0$ around a point $z \in \mathbb{C}$ is the set

$$B(z, r) = \{w \in \mathbb{C} \mid \ |z - w| < r\}.$$

The **closed circular disk** with radius $r > 0$ around $z \in \mathbb{C}$ is the set

$$\overline{B(z, r)} = \{w \in \mathbb{C} \mid \ |z - w| \leq r\}.$$

6 2.51. — The open circular disk $B(z, r)$ thus consists precisely of those points that have distance strictly less than r from z. Open circular disks in $\mathbb{C}$ and open intervals in $\mathbb{R}$ are compatible in the following sense: If $x \in \mathbb{R}$ and $r > 0$, then the intersection of the open circular disk $B(x, r) \subseteq \mathbb{C}$ with $\mathbb{R}$ is just the open interval $(x - r, x + r)$ liyng symmetrically about x.



Exercise 2.52. — Show the following property of open circular disks: let $z_1, z_2 \in \mathbb{C}$, $r_1 > 0$ and $r_2 > 0$. For each point $z \in B(z_1, r_1) \cap B(z_2, r_2)$ there exists a radius $r > 0$ such that

$$B(z, r) \subseteq B(z_1, r_1) \cap B(z_2, r_2).$$

Illustrate your choice of radius r in a picture.

The definition of open set in $\mathbb{C}$ given below generalizes that in $\mathbb{R}$ from Exercise 2.37.

Definition 2.53: Open and Closed Sets

A subset $U \subseteq \mathbb{C}$ is called **open** in $\mathbb{C}$ if for every point in U there exists an open circular disk around that point contained in U. More formally, for all $z \in U$ there exists a radius $r > 0$ such that $B(z, r) \subseteq U$.

6 A subset $A \subseteq \mathbb{C}$ is called **closed** in $\mathbb{C}$ if its complement $\mathbb{C} \setminus A$ is open.

For example, thanks to Exercise 2.52, all open circular disks are open. In addition to open circular disks, there are many other subsets of $\mathbb{C}$. For example, every union of open subsets is open. We will return to studying open sets and related notions in much greater generality in the second semester.

2.3 Maximum and Supremum

2.3.1 Existence of the Supremum

Definition 2.54: Bounded Sets, Maxima and Minima

Let $X \subseteq \mathbb{R}$ be a subset of the real numbers.

- X is **bounded from above** if there exists $s \in \mathbb{R}$ such that $x \leq s$ for all $x \in X$. Any such $s \in \mathbb{R}$ is called an **upper bound** of X. An upper bound that is itself contained in X is called a **maximum**. We write

$$s = \max(X)$$

  if such a maximum of X exists and is equal to s.
- The terms **bounded from below**, **lower bound**, and **minimum** are defined analogously.
- $X \subset \mathbb{R}$ is called **bounded** if it is bounded from above and from below.

6 2.55. — If a subset $X \subseteq \mathbb{R}$ has a maximum, then it is unique. Indeed, if $x_1 \in X$ and $x_2 \in X$ both satisfy the properties of a maximum, it follows that $x_1 \leq x_2$ because x_2 is a maximum, and $x_2 \leq x_1$ because x_1 is a maximum, and hence $x_1 = x_2$. So we may speak of *the* maximum of a subset $X \subset \mathbb{R}$ if one exists.

A closed interval $[a, b]$ with endpoints $a < b$ in $\mathbb{R}$ has $a = \min([a, b])$ as minimum and $b = \max([a, b])$ as maximum. However, there are also sets that do not have a maximum. For example, the open interval (a, b) with endpoints $a < b$ in $\mathbb{R}$ has no maximum. Although the endpoint b would lend itself as a maximum, it does not lie in the set (a, b) and is therefore not a candidate for the maximum. Likewise $\mathbb{R}$ as a subset of $\mathbb{R}$, or also unbounded intervals of the form $[a, \infty), (a, \infty)$ for $a \in \mathbb{R}$, do not have a maximum.

Definition 2.56: Supremum

Let $X \subseteq \mathbb{R}$ be a subset and let $A := \{a \in \mathbb{R} \mid x \leq a \text{ for all } x \in X\}$ be the set of all upper bounds of X. If the minimum $s = \min(A)$ exists, then we call this minimum the **supremum** of X, and write

$$s = \sup(X).$$

In other words, if it exists, the supremum of X is the smallest upper bound of X. We can describe the supremum $s = \sup(X)$ of X directly by

$$x \leq s \text{ for all } x \in X, \qquad \text{and} \qquad x \leq t \text{ for all } x \in X \implies s \leq t. \tag{2.4}$$

Equivalently, the supremum of X could also be characterised by the fact that no real number x_1 strictly smaller than $s = \sup(X)$ is an upper bound of X:

$$x \le s \text{ for all } x \in X, \qquad \text{and} \qquad t < s \implies \exists\, x \in X \text{ with } t < x. \tag{2.5}$$

The statements (2.4) and (2.5) are equivalent.

We have not yet addressed the question of the existence of the supremum. If X is the empty set or if X is unbounded from above, then the supremum cannot exist. In all other cases it exists, as the following theorem shows.

Theorem 2.57: Existence of Supremum

Let $X \subset \mathbb{R}$ be a nonempty subset bounded from above. Then the supremum of X exists.

Proof. By assumption, X is non-empty, and the set of upper bounds $A := \{a \in \mathbb{R} \mid x \le a \text{ for all } x \in X\}$ is also non-empty. Furthermore, for all $x \in X$ and $a \in A$, the inequality $x \le a$ holds. Therefore, according to the completeness axiom in Definition 2.18, it follows that there exists a $c \in \mathbb{R}$ such that

$$x \le c \le a$$

holds for all $x \in X$ and $a \in A$. From the first inequality it follows that c is an upper bound of X, so $c \in A$. From the second inequality it follows that c is the minimum of the set A. □

6 **Applet 2.58** (Supremum of a bounded non-empty set). *We consider a bounded non-empty subset of $\mathbb{R}$ and two equivalent characterizations of the supremum of this set.*

Proposition 2.59: Supremum vs Set Operations

Let X and Y be nonempty subsets of $\mathbb{R}$ bounded from above, and write

$$X + Y := \{x + y \mid x \in X, y \in Y\} \qquad \text{and} \qquad XY := \{xy \mid x \in X, y \in Y\}.$$

The sets $X \cup Y$, $X \cap Y$, and $X + Y$ are bounded from above. Also, if $x \ge 0$ and $y \ge 0$ for all $x \in X$ and $y \in Y$, then XY is bounded from above. Furthermore:

(1) $\sup(X \cup Y) = \max\{\sup(X), \sup(Y)\}$,

(2) if $X \cap Y$ is not empty, then $\sup(X \cap Y) \le \min\{\sup(X), \sup(Y)\}$,

(3) $\sup(X + Y) = \sup(X) + \sup(Y)$ *and*

(4) if $x \ge 0$ and $y \ge 0$ for all $x \in X$ and $y \in Y$, then $\sup(XY) = \sup(X)\sup(Y)$.

Proof. The proof of (1) and (2) is left to the reader.

To prove (3) we write $x_0 = \sup(X)$ and $y_0 = \sup(Y)$. Let $z \in X + Y$. There are $x \in X$ and $y \in Y$ with $z = x + y$, and because of $x \le x_0$ and $y \le y_0$, then $z \le x_0 + y_0$ holds. That is, $x_0 + y_0$ is an upper bound for $X + Y$. We want to show that $x_0 + y_0$ is the supremum of $X + Y$, i.e., the smallest upper bound of $X + Y$.

Let $z_0 = \sup(X + Y)$ (the supremum exists, since we just proved that $X + Y$ is bounded from above) and assume by contradiction that

$$\varepsilon = x_0 + y_0 - z_0 > 0.$$

Since x_0 is the supremum of X, it follows from (2.5) that there exists $x \in X$ with $x > x_0 - \varepsilon/2$, and similarly there exists $y \in Y$ with $y > y_0 - \varepsilon/2$. Set $z = x + y$. It follows that

$$z = x + y > x_0 - \varepsilon/2 + y_0 - \varepsilon/2 = z_0,$$

which contradicts the fact that z_0 is an upper bound for $X + Y$. Thus $z_0 = x_0 + y_0$, as desired.

The proof of (4) is done in a similar way. □

2.60. — For a non-empty subset $X \subseteq \mathbb{R}$ bounded from below, the largest lower bound of X will also be called the **infimum** $\inf(X)$ of X. An existence statement analogous to Theorem 2.57 holds for the infimum. Alternatively, the infimum of X can be written as

$$-\sup\{-x \mid x \in X\}.$$

6

In this way, practically all statements about infima can be traced back to statements about suprema.

2.3.2 Two-point Compactification

2.61. — In this section we want to extend the notions of **supremum** and **infimum** to arbitrary subsets of $\mathbb{R}$. To do this, we use the symbols $\infty = +\infty$ and $-\infty$, which are not real numbers. We define the **extended number line**, which is also called **two-point compactification** of $\mathbb{R}$, by

$$\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$$

and imagine this as the number line

Here we have added the point $+\infty$ to the right of $\mathbb{R}$ and the point $-\infty$ to the left of $\mathbb{R}$ to the straight line. We extend the order relation of the real numbers $\leq$ to $\overline{\mathbb{R}}$ by requiring $-\infty < x < +\infty$ for all $x \in \overline{\mathbb{R}}$.

To simplify the notation, we also write ∞ in place of $+\infty$. Often used calculation rules for the symbols $-\infty$ and ∞, such as the following, are standard conventions:

$$\infty + x = \infty + \infty = \infty \qquad \text{and} \qquad -\infty + x = -\infty - \infty = -\infty$$

for all $x \in \mathbb{R}$. Also, for $x > 0$,

$$x \cdot \infty = \infty \cdot \infty = \infty \qquad \text{and} \qquad x \cdot (-\infty) = \infty \cdot (-\infty) = -\infty.$$

One should use such conventions as sparingly as possible and be careful with them. The expressions $\infty - \infty$ and $0 \cdot \infty$ or similar remain undefined.

6 **Definition 2.62: Indefinite values**

Let X be a subset of $\mathbb{R}$. If $X \subset \mathbb{R}$ is not bounded from above, then we define $\sup(X) = \infty$. If X is empty, we set $\sup(\emptyset) = -\infty$. In this context, we call ∞ and $-\infty$ **indefinite values**.

In other words, the statement $\sup(X) = \infty$ is equivalent, by definition, to the statement that X is not bounded from above, that is,

$$\forall\, x_0 \in \mathbb{R} \;\; \exists\, x \in X \; : \; x > x_0.$$

In the same way $\sup(X) = -\infty$ is equivalent, by definition, to the statement $X = \emptyset$.

Similarly, we define $\inf(\emptyset) = +\infty$, and $\inf(X) = -\infty$ if $X \subseteq \mathbb{R}$ is not bounded from below.

2.4 Consequences of Completeness

We introduced the root function in Section 2.1 using the completeness axiom, and used the completeness axiom in Section 2.3 to prove the existence of the supremum. In this section we will discuss some further consequences of the completeness axiom. We choose for the whole Section 2.4 an arbitrary set of real numbers $\mathbb{R}$, and call the elements of $\mathbb{R}$ real numbers.

2.4.1 The Archimedean Principle

Archimedes' principle is the statement that for every real number $x \in \mathbb{R}$ there is an integer n greater than x. The following theorem, which we prove using the existence of the supremum, is a slightly more precise formulation of Archimedes' Principle.

Theorem 2.63: Archimedean Principle

For every $x \in \mathbb{R}$ there exists exactly one $n \in \mathbb{Z}$ with $n \leq x < n+1$.

Proof. First, let $x \geq 0$ be a non-negative real number. Then $E = \{n \in \mathbb{Z} \mid n \leq x\}$ is an upper bounded subset of $\mathbb{R}$ which is non-empty since $0 \in E$. By Theorem 2.57, the supremum $s_0 = \sup(E)$ exists. Since s_0 is the smallest upper bound of E, then:
7 (i) $s_0 \leq x$ holds;
(ii) there exists $n_0 \in E$ with $s_0 - 1 < n_0$.
It follows from (ii) that $s_0 < n_0 + 1$, and therefore $m \leq s_0 < n_0 + 1$ for every $m \in E$. This proves that $m \leq n_0$ for every $m \in E$, which means that $n_0 = s_0$ is the maximum of E.
Also, using (ii) again, $n_0 + 1 \notin E$, so $n_0 + 1 > x$ by the definition of E.
Recalling (i), this proves that

$$n_0 = s_0 \leq x < n_0 + 1,$$

which shows the validity of Theorem 2.63 in the case $x \geq 0$.

If $x < 0$, then we can apply the above case to $-x$ and conclude the existence of $m \in \mathbb{Z}$ such that $m \leq -x < m+1$, or equivalently $-m-1 < x \leq -m$. If $x = -m$, then the result follows with $n = -m$. If $x < -m$, then the result follows with $n = -m-1$.

For the proof of uniqueness, we assume that for $n_1, n_2 \in \mathbb{Z}$ both inequalities $n_1 \leq x < n_1+1$ and $n_2 \leq x < n_2 + 1$ hold. It follows that $n_1 \leq x < n_2 + 1$ and thus $n_1 \leq n_2$. Similarly $n_2 \leq n_1$, implying $n_1 = n_2$. □

Definition 2.64: Integer and Fractional Parts

The **integer part** $\lfloor x \rfloor$ of a number $x \in \mathbb{R}$ is the integer $n \in \mathbb{Z}$ uniquely determined by Theorem 2.63 with $n \leq x < n+1$. The function from $\mathbb{R}$ to $\mathbb{Z}$ given by $x \mapsto \lfloor x \rfloor$ is called **rounding function**. The **fractional part** of a real number x is $\{x\} = x - \lfloor x \rfloor \in [0, 1)$.

Corollary 2.65: $\frac{1}{n}$ is arbitrarily small

For every $\varepsilon > 0$ there exists an integer $n \geq 1$ such that $\frac{1}{n} < \varepsilon$ holds.

Proof. If $\varepsilon > 1$ then the result holds with $n = 1$.

If $\varepsilon \leq 1$ then $\frac{1}{\varepsilon} \geq 1$ and, thanks to Theorem, 2.63 there exists $n \geq 1$ with $n > \frac{1}{\varepsilon}$. Then, for such n, the inequality $\frac{1}{n} < \varepsilon$ holds. □

Corollary 2.66: Density of $\mathbb{Q}$

For every two real numbers $a, b \in \mathbb{R}$ with $a < b$, there is one $r \in \mathbb{Q}$ with $a < r < b$.

Proof. Set $\varepsilon = b - a$. According to Archimedes' principle in the form of Corollary 2.65, there exists $m \in \mathbb{N}$ with $\frac{1}{m} < \varepsilon$. Similarly, according to Archimedes' principle from Theorem 2.63, there exists $n \in \mathbb{Z}$ with $n \leq ma < n + 1$, or equivalently $\frac{n}{m} \leq a < \frac{n+1}{m}$.

Hence, since $\frac{1}{m} < \varepsilon$ and $a + \varepsilon = b$, we get

$$a < \frac{n+1}{m} = \frac{n}{m} + \frac{1}{m} \leq a + \frac{1}{m} < a + \varepsilon = b.$$

This proves the corollary with $r = \frac{n+1}{m}$. □

Stated differently, the above corollary shows that $\mathbb{Q}$ intersects any open non-empty interval
7 I, that is, $I \cap \mathbb{Q} \neq \emptyset$. A subset X of $\mathbb{R}$ is called **dense** in $\mathbb{R}$ if every open non-empty interval of $\mathbb{R}$ contains an element of X. Corollary 2.66 thus states: $\mathbb{Q}$ *is dense in* $\mathbb{R}$.

The following two exercises establish a generalization of Archimedes' principle, which will be used in the section about decimal fractions.

Exercise 2.67. — Let $A := \{m_0, m_1, m_2, \ldots\} \subset \mathbb{Z}$ be a strictly increasing sequence of integers, i.e., $m_{i+1} > m_i$ for every $i \geq 1$. Then the following analog of Theorem 2.63 holds with A in place of $\mathbb{Z}$: *Given $x \in \mathbb{R}$ with $x \geq m_0$, there exists exactly one element $m_i \in A$ such that $m_i \leq x < m_{i+1}$.*

Exercise 2.68. — Let $A := \{m_0, m_1, m_2, \ldots\} \subset \mathbb{Z}$ be a strictly increasing sequence of positive integers. Then Corollary 2.65 holds with A in place of $\mathbb{N}$. In other words, for every $\varepsilon > 0$ there exists an element $m_i \in A$ such that $\frac{1}{m_i} < \varepsilon$ holds.

2.4.2 Decimal Fraction Expansion and Uncountability (Extra material)

2.69. — A common notion of the real numbers is given by non-terminating decimal fractions. Formally, we define a **decimal fraction** as a sequence of integers

$$a_0, a_1, a_2, a_3, \ldots$$

with $a_0 \in \mathbb{Z}$, and $0 \leq a_n \leq 9$ for all $n \geq 1$. Thus, given a decimal fraction, we can assign to it a unique element of $\mathbb{R}$ as follows.

Suppose $a_0 \geq 0$. We set

$$x_n = \sum_{k=0}^{n} a_k \cdot 10^{-k} \qquad \text{and} \qquad y_n = 10^{-n} + \sum_{k=0}^{n} a_k \cdot 10^{-k}. \tag{2.6}$$

One can check that

$$x_0 \leq x_1 \leq \ldots \leq x_n \leq x_{n+1} \leq \ldots \leq y_{n+1} \leq y_n \leq \ldots \leq y_1 \leq y_0.$$

Thus, if we consider the sets $X = \{x_0, x_1, x_2, \ldots\}$ and $Y = \{y_0, y_1, y_2, \ldots\}$, according to the completeness axiom we can conclude the existence of $c \in \mathbb{R}$ with the property that

$$x_n \leq c \leq y_n \tag{2.7}$$

for all $n \in \mathbb{N}$. Archimedes' principle in the form of Corollary 2.65 shows that there is precisely one real number c that satisfies the inequality $x_n \leq c \leq y_n$ for all $n \geq 0$. If there were two different such numbers, say c and d with $c < d$, it follows from

$$x_n \leq c < d \leq y_n$$

7 and the definition of x_n and y_n the inequality

$$0 < d - c < 10^{-n}$$

for all $n \geq 0$, which contradicts Exercise 2.68 with $m_i = 10^i$. We call the element $c \in \mathbb{R}$ uniquely determined by (2.7) the real number with decimal expansion $a_0, a_1a_2a_3a_4\ldots$. We note that two possible alternative definitions would be

$$c = \sup\{x_n \mid n \geq 0\} \qquad \text{or} \qquad c = \inf\{y_n \mid n \geq 0\}.$$

If a_0 is negative, we first consider the real number c with decimal expansion $-a_0, a_1a_2a_3a_4\ldots$ and then define the real number with decimal expansion $a_0, a_1a_2a_3a_4\ldots$ as $-c$.

2.70. — Now, the following question arises: Can every element of $\mathbb{R}$ be written as a decimal fraction? This is indeed the case: Let $c \in \mathbb{R}$ and $c \geq 0$. Then we can write $a_0 := \lfloor c \rfloor$ and define

$$a_n := \lfloor 10^n c \rfloor - 10\lfloor 10^{n-1} c \rfloor. \tag{2.8}$$

One verifies that the sequence $a_0, a_1, a_2, \ldots$ is indeed a decimal fraction, since $0 \leq a_n \leq 9$ is satisfied for all $n \geq 1$. Also, it is not difficult to check the inequalities (2.7), from which it follows that c is just the real number with decimal expansion $a_0, a_1a_2a_3\ldots$. The case $c \leq 0$ is obtained in the same way, by changing the sign of a_0 in the decimal expansion of $-c$.

2.71. — It is important to remark that two different decimal fractions can represent the same real number. For example, the two decimal fractions

$$0.199999999999\ldots \qquad \text{and} \qquad 0.20000000000000\ldots$$

both represent the real number $\frac{1}{5}$. However, the problem only occurs when a decimal fraction becomes a constant sequence $\ldots 9999 \ldots$ after a certain point. To rule this out, we can consider the following definition: We call **real decimal fraction** any sequence of integers

$$a_0, a_1, a_2, a_3, \ldots$$

with $0 \leq a_n \leq 9$ for all $n \geq 1$, and with the property that for every $n_0 \geq 1$ there exists a $n \geq n_0$ with $a_n \neq 9$.

7 Exercise 2.72. — Let $c \geq 0$ be a real number. Verify that the sequence $a_0, a_1, a_2, \ldots$ defined by (2.8) is a real decimal fraction. Then show that this gives rise to a bijection between $\mathbb{R}$ and the set of all real decimal fractions.

Interlude: Power Set

Let X be a set. The **power set** $\mathcal{P}(X)$ of X is the set of all subsets of X, that is,

$$\mathcal{P}(X) = \{Q \mid Q \text{ is a set and } Q \subset X\}.$$

Example 2.73. — If $X = \{0, 1, 2\}$ then

$$\mathcal{P}(X) = \big\{\emptyset, \{0\}, \{1\}, \{2\}, \{0,1\}, \{0,2\}, \{1,2\}, \{0,1,2\}\big\}.$$

Interlude: Cardinality

Let X and Y be two sets. We say that X and Y have the **same cardinality** or the **same number of elements**, written $X \sim Y$, if there is a bijection $f : X \to Y$. We say that Y is **larger** than X, and write $X \lesssim Y$, if there is an injection $f : X \to Y$.

- We say that the cardinality of the empty set is zero, and write $|\emptyset| = 0$.
- Let X be a set and $n \geq 1$ a natural number. We say the set X has cardinality n, and write $|X| = n$, if there is a bijection from X to $\{1, \dots, n\}$. In this case we call X a **finite set** and write $|X| < \infty$.
- If X is not finite, we call X an **infinite set**.
- A set is called **countable** if there is a bijection to $\mathbb{N}$. The cardinality of $\mathbb{N}$ is also called $\aleph_0$, pronounced **Aleph-**0.
- A set X is called **uncountable** if it is infinite and not countable.
- The cardinality of $\mathcal{P}(\mathbb{N})$ is written $\mathfrak{c}$ and called the **continuum**.

Remark 2.74. — A nontrivial fact is that

7

$$X \lesssim Y \quad \text{and} \quad Y \lesssim X \implies X \sim Y.$$

In other words, if there exist an injective map $f : X \to Y$ and an injective map $g : Y \to X$, then one can find a bijective map $h : X \to Y$. This fact is known as the Schröder–Bernstein Theorem.

Theorem 2.75: Cantor's Theorem

Let X be a set. Then the power set $\mathcal{P}(X)$ is larger than X and not equal to X.

Proof. The function $i : X \to \mathcal{P}(X)$ given by $i(x) = \{x\}$ is injective. So $\mathcal{P}(X)$ is larger than X. It remains to show that there is no bijection from X to $\mathcal{P}(X)$. To show this, we assume that there is a bijection $f : X \to \mathcal{P}(X)$ and take this to a contradiction. For this, we define the set

$$A = \{x \in X \mid x \notin f(x)\}.$$

In other words, $A \in \mathcal{P}(X)$ consists of all elements x in X for which x is not an element of the subset $f(x) \subset X$.

Since, by assumption, $f : X \to \mathcal{P}(X)$ is a bijection, there exists $a \in X$ such that $A = f(a)$. We now ask ourselves: does a belong to A or not?

If $a \in A$ then, by the definition of A, $a \notin f(a)$. However, this is impossible since $f(a) = A$. Vice versa, if $a \notin A$ it means that $a \in f(a)$, that again is impossible since $f(a) = A$

This proves that there exists no $a \in X$ with $f(a) = A$, which contradicts the surjectivity of f. So there can be no bijection $f : X \to \mathcal{P}(X)$. □

Proposition 2.76: Uncountability of $\mathbb{R}$

The set $\mathbb{R}$ is uncountable.

Proof. By Theorem 2.75, $\mathcal{P}(\mathbb{N})$ is strictly larger than $\mathbb{N}$, and thus uncountable. Thus, to show that $\mathbb{R}$ is uncountable, it suffices to prove the existence of an injection

$$\varphi : \mathcal{P}(\mathbb{N}) \to \mathbb{R}.$$

7 We construct such an injection by assigning to each subset A the real number $\varphi(A)$ whose decimal fraction expansion is given by $a_0, a_1 a_2 a_3 a_4 \ldots$ with

$$a_n = \begin{cases} 0 & \text{if } n \notin A \\ 1 & \text{if } n \in A \end{cases}$$

Injectivity of the function φ can be proved in two ways. As a first option, one can simply apply Problem 2.72. Alternatively, one can argue as follows. Let A and B be distinct subsets of $\mathbb{N}$, and let n be the smallest element of $A \triangle B$ (recall Definition 2). If $n \in A$ and $n \notin B$ then $\varphi(A) > \varphi(B)$ holds. On the other hand, if $n \notin A$ and $n \in B$ then $\varphi(A) < \varphi(B)$ holds. Therefore $\varphi(A) \neq \varphi(B)$ holds in all cases, which proves the injectivity. □

2.5 Sequences of Real Numbers

2.5.1 Convergence of Sequences

Let X be a set. Intuitively, a sequence in X is a non-terminating sequence of elements $x_0, x_1, x_2, x_3, \ldots$ of X indexed by the natural numbers. Here, we are interested in sequences and their properties in $\mathbb{R}$. We give now a precise definition, although we have already used the concept of sequence before in an intuitive way.

Definition 2.77: Sequences

A **sequence** in $\mathbb{R}$ is a function $a : \mathbb{N} \to \mathbb{R}$. The image $a(n)$ of $n \in \mathbb{N}$ is also written as a_n and is called the n-th **element** of a. Instead of $a : \mathbb{N} \to \mathbb{R}$ one often writes $(a_n)_{n\in\mathbb{N}}$, $(a_n)_{n=0}^{\infty}$, or $(a_n)_{n\geq 0}$.

Since we primarily use the letter x to denote a real number, for sequences of real numbers we shall mostly use the notation $(x_n)_{n\in\mathbb{N}}$, $(x_n)_{n=0}^{\infty}$, or $(x_n)_{n\geq 0}$.

Definition 2.78: (Eventually) Constant Sequences

7 A sequence $(x_n)_{n=0}^{\infty}$ is called **constant** if $x_n = x_m$ for all $m, n \in \mathbb{N}$, and **eventually constant** if there exists $N \in \mathbb{N}$ with $x_n = x_m$ for all $m, n \in \mathbb{N}$ with $m, n \geq N$.

Example 2.79. — A sequence in $\mathbb{R}$ is, for example, $(x_n)_{n=0}^{\infty}$ given by $x_0 = 1$, $x_n = \frac{1}{n}$ for $n \geq 1$.

Definition 2.80: Convergence for Sequences

Let $(x_n)_{n=0}^{\infty}$ be a sequence in $\mathbb{R}$. We say that $(x_n)_{n=0}^{\infty}$ is **convergent** or **converges** if there exists an element $A \in \mathbb{R}$ with the following property:

$$\forall \varepsilon > 0 \ \exists N \in \mathbb{N} \text{ such that } \forall n \in \mathbb{N} \ : \ n \geq N \implies |x_n - A| < \varepsilon.$$

In this case, we write

$$\lim_{n\to\infty} x_n = A \tag{2.9}$$

and call the point A the **limit** of the sequence $(x_n)_{n=0}^{\infty}$.

8 It is a priori not clear that a converging sequence has only one limit. In the following lemma we show that the limit is indeed unique, so the notation (2.9) is justified.

Lemma 2.81: Uniqueness of the Limit

A convergent sequence $(x_n)_{n=0}^{\infty}$ *in* $\mathbb{R}$ *has exactly one limit.*

Proof. Let $A \in \mathbb{R}$ and $B \in \mathbb{R}$ be limits of the sequence $(x_n)_{n=0}^{\infty}$. Let $\varepsilon > 0$. Then we can find $N_A, N_B \in \mathbb{N}$ such that for all $n \geq N_A$ it holds $|x_n - A| < \frac{\varepsilon}{2}$, and for all $n \geq N_B$ it holds $|x_n - B| < \frac{\varepsilon}{2}$. Set $N = \max\{N_A, N_B\}$. Then

$$|A - B| \leq |A - x_N| + |x_N - B| < \tfrac{\varepsilon}{2} + \tfrac{\varepsilon}{2} = \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary it follows that $|A - B| = 0$, and so $A = B$. □

Example 2.82. — A constant sequence $(x_n)_{n=0}^{\infty}$ with $x_n = A \in \mathbb{R}$ for all $n \in \mathbb{N}$ converges to A. Similarly, eventually constant sequences converge to the value they eventually take.

Example 2.83. — The sequence of real numbers $\left(\frac{1}{n}\right)_{n=1}^{\infty}$ converges to zero, i.e., $\lim_{n\to\infty} \frac{1}{n} = 0$. Indeed, given $\varepsilon > 0$, by Archimedes' principle (Theorem 2.63) there exists $N \in \mathbb{N}$ with $\frac{1}{N} < \varepsilon$. Therefore, for every $n \in \mathbb{N}$ with $n \geq N$, we have $0 \leq \frac{1}{n} \leq \frac{1}{N} < \varepsilon$.

8 Example 2.84. — The sequence of real numbers $(y_n)_{n=0}^{\infty}$ given by $y_n = (-1)^n$ for $n \in \mathbb{N}$ is not convergent, since the sequence members $1, -1, 1, -1, 1, -1, \ldots$ alternate between 1 and 1 and, in particular, do not approach any real number.

2.5.2 Convergent Subsequences and Accumulation Points

Let $(x_n)_{n=0}^{\infty}$ be a sequence in $\mathbb{R}$. We call a subsequence of $(x_n)_{n=0}^{\infty}$ any sequence obtained by keeping only certain elements of the sequence $(x_n)_{n=0}^{\infty}$ and ignoring all others. For example

$$x_0, x_1, x_4, x_9, x_{16}, x_{25}, \ldots$$

is a subsequence. The formal definition is as follows:

Definition 2.85: Subsequences

Let $(x_n)_{n=0}^{\infty}$ be a sequence in $\mathbb{R}$. A **subsequence** of $(x_n)_{n=0}^{\infty}$ is a sequence of the form $(x_{n_k})_{k=0}^{\infty}$, where $(n_k)_{k=0}^{\infty}$ is a sequence of nonnegative integers such that $n_{k+1} > n_k$ for all $k \in \mathbb{N}$.

Remark 2.86. — In the previous definition, since $n_{k+1} > n_k$ for all $k \in \mathbb{N}$, it follows that $n_k \geq k$ for every $k \in \mathbb{N}$. (Exercise: Prove this fact by induction on $k \in \mathbb{N}$.)

Lemma 2.87: Subsequences of Convergent Sequences are Convergent

Let $(x_n)_{n=0}^{\infty}$ be a sequence in $\mathbb{R}$ converging to $A \in \mathbb{R}$. Then, each subsequence of $(x_n)_{n=0}^{\infty}$ also converges to A.

Proof. We leave the proof as an exercise. □

2.88. — A sequence can have convergent subsequences without itself converging. For example, the sequence of real numbers given by $x_n = (-1)^n$ is not convergence, while the subsequences

$$(x_{2n})_{n=0}^{\infty} \quad \text{and} \quad (x_{2n+1})_{n=0}^{\infty}$$

are constant, and so in particular they are convergent.

Definition 2.89: Accumulation Points of Sequences

Let $(x_n)_{n=0}^{\infty}$ be a sequence in $\mathbb{R}$. A point $A \in \mathbb{R}$ is called an **accumulation point** of the sequence $(x_n)_{n=0}^{\infty}$ if for every $\varepsilon > 0$ and every $N \in \mathbb{N}$ there exists a natural number $n \geq N$ with $|x_n - A| < \varepsilon$.

Proposition 2.90: Subsequences and Accumulation Points

Let $(x_n)_{n=0}^{\infty}$ be a sequence in $\mathbb{R}$. An element $A \in \mathbb{R}$ is an accumulation point of $(x_n)_{n=0}^{\infty}$
8 *if and only if there exists a convergent subsequence of $(x_n)_{n=0}^{\infty}$ with limit A.*

Proof. Suppose $A \in \mathbb{R}$ is an accumulation point of the given sequence. We recursively construct a subsequence $(x_{n_k})_{k=0}^{\infty}$ that converges to A.
First, we apply the definition of accumulation point with $\varepsilon = 1$ to find an integer n_0 such that $|x_{n_0} - A| \leq 1$ holds.
Then, we apply the definition of accumulation point with $\varepsilon = 2^{-1}$ and $N = n_0 + 1$ to find $n_1 > n_0$ such that $|x_{n_1} - A| \leq 2^{-1}$.
Then, we apply the definition of accumulation point with $\varepsilon = 2^{-2}$ and $N = n_1 + 1$ to find $n_2 > n_1$ such that $|x_{n_2} - A| \leq 2^{-2}$.
In general, we apply the definition of accumulation point with $\varepsilon = 2^{-k}$ and $N = n_{k-1} + 1$ to find $n_k > n_{k-1}$ such that $|x_{n_k} - A| \leq 2^{-k}$.
We claim that the subsequence $(x_{n_k})_{k=0}^{\infty}$ of $(x_n)_{n=0}^{\infty}$ converges to A. Indeed, for every $\varepsilon > 0$ there exists $N \in \mathbb{N}$ with $2^{-N} < \varepsilon$. Then, for every $k \geq N$ it holds

$$|x_{n_k} - A| \leq 2^{-k} \leq 2^{-K} < \varepsilon,$$

which shows that A is the limit of the sequence $(x_{n_k})_{k=0}^{\infty}$.

Conversely, suppose A is the limit of a converging subsequence $(x_{n_k})_{k=0}^{\infty}$. Fix $\varepsilon > 0$ and $N \in \mathbb{N}$. Since x_{n_k} converges to A, there exists $N_0 \in \mathbb{N}$ such that $|x_{n_k} - A| < \varepsilon$ for all $k \geq N_0$. Hence, to satisfy the definition of an accumulation point, we want to choose $k \geq N_0$ so that

$n_k \geq N$. Recalling that $n_k \geq k$ (see Remark 2.86), we define $N_1 = \max\{N_0, N\}$ so that

$$|x_{n_{N_1}} - A| < \varepsilon \qquad \text{and} \qquad n_{N_1} \geq N.$$

So, A is an accumulation point of $(x_n)_{n=0}^{\infty}$. □

Corollary 2.91: Accumulation Points have Infinite Elements nearby

Let $A \in \mathbb{R}$ be an accumulation point of the sequence $(x_n)_{n=0}^{\infty}$. Then, for every $\varepsilon > 0$, there are infinitely many elements of the sequence $(x_n)_{n=0}^{\infty}$ inside $(A - \varepsilon, A + \varepsilon)$.

Proof. By Proposition 2.90 there exists a sequence $(x_{n_k})_{k \geq 0}$ such that $\lim_{k\to\infty} x_{n_k} = A$.
8 In particular, given $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that all elements $(x_{n_k})_{k \geq N}$ are inside $(A - \varepsilon, A + \varepsilon)$. □

Corollary 2.92: Accumulation Points of Convergent Sequences

A converging sequence has exactly one accumulation point, which coincides with its limit.

Proof. This follows from Lemma 2.87 and Proposition 2.90. □

Exercise 2.93. — Let $(x_n)_{n=0}^{\infty}$ be a sequence in $\mathbb{R}$, and let $F \subseteq \mathbb{R}$ be the set of accumulation points of the sequence $(x_n)_{n=0}^{\infty}$. Show that F is closed.

2.5.3 Addition, Multiplication, and Inequalities

2.94. — Given sequences $(x_n)_{n=0}^\infty$ and $(y_n)_{n=0}^\infty$, and a number $\alpha \in \mathbb{R}$, sums and scalar multiples of sequences are given by

$$\begin{aligned}(x_n)_{n=0}^\infty + (y_n)_{n=0}^\infty &= (x_n + y_n)_{n=0}^\infty \\ \alpha \cdot (x_n)_{n=0}^\infty &= (\alpha x_n)_{n=0}^\infty.\end{aligned}$$

Sequences can also be multiplied as follows:

$$(x_n)_{n=0}^\infty \cdot (y_n)_{n=0}^\infty = (x_n y_n)_{n=0}^\infty.$$

Remark 2.95. — With the addition and multiplication defined above, the set of sequences forms a commutative ring, where the zero element is the constant sequence $(0)_{n=0}^\infty$, and the neutral element for multiplication is the constant sequence $(1)_{n=0}^\infty$.

Proposition 2.96: Limits and Operations

Let $(x_n)_{n=0}^\infty$ and $(y_n)_{n=0}^\infty$ be sequences converging to $A \in \mathbb{R}$ and $B \in \mathbb{R}$, respectively.

1. *The sequence $(x_n + y_n)_{n=0}^\infty$ converges to $A + B$.*
8
2. *The sequence $(x_n y_n)_{n=0}^\infty$ converges to AB, and the sequence $(\alpha x_n)_{n=0}^\infty$ converges to αA for every $\alpha \in \mathbb{R}$.*
3. *Suppose $x_n \neq 0$ for all $n \in \mathbb{N}$ and $A \neq 0$. Then the sequence $(x_n^{-1})_{n=0}^\infty$ converges to A^{-1}.*

Proof. 1. Let $\varepsilon > 0$, and let $N_A, N_B \in \mathbb{N}$ be such that

$$|x_n - A| < \tfrac{\varepsilon}{2} \quad \forall\, n \geq N_A, \qquad \text{and} \qquad |y_n - B| < \tfrac{\varepsilon}{2} \quad \forall\, n \geq N_B.$$

Then, with $N = \max\{N_A, N_B\}$, for all $n \geq N$ it holds

$$|(x_n + y_n) - (A + B)| \leq |x_n - A| + |y_n - B| < \tfrac{\varepsilon}{2} + \tfrac{\varepsilon}{2} = \varepsilon,$$

which shows statement (1).

2. We refer to Exercise 2.103.

3. Let A be the limit of the sequence $(x_n)_{n=0}^\infty$. Since $A \neq 0$, choosing $\varepsilon = \frac{|A|}{2}$ in the definition of limit, there exists $N_0 \in \mathbb{N}$ such that, for all $n \geq N_0$,

$$|x_n - A| \leq \frac{|A|}{2}.$$

In particular, by inverse triangle inequality (see property (h) in Paragraph 2.16),

$$|x_n| = |A + (x_n - A)| \geq |A| - |x_n - A| \geq \frac{|A|}{2} \qquad \forall\, n \geq N_0.$$

which implies that

$$\frac{1}{|x_n|} \leq \frac{2}{|A|} \qquad \forall\, n \geq N_0.$$

In particular,

8 $$\left|x_n^{-1} - A^{-1}\right| = \frac{|x_n - A|}{|x_n|\,|A|} \leq 2\frac{|x_n - A|}{|A|^2} \qquad \forall\, n \geq N_0.$$

Now, given $\varepsilon > 0$ we choose N_1 so that $|x_n - A| < \frac{|A|^2}{2}\varepsilon$ for all $n \geq N_1$. Then, for $N = \max\{N_0, N_1\}$ we get

$$\left|x_n^{-1} - A^{-1}\right| \leq 2\frac{|x_n - A|}{|A|^2} < \varepsilon \qquad \forall\, n \geq N.$$

This proves that the sequence $(x_n^{-1})_{n=0}^{\infty}$ converges to A^{-1}, as desired. □

Proposition 2.97: Limits and Inequalities

Let $(x_n)_{n=0}^{\infty}$ and $(y_n)_{n=0}^{\infty}$ be sequences of real numbers with limits $A = \lim_{n\to\infty} x_n$ and $B = \lim_{n\to\infty} y_n$.

1. *If $A < B$, then there exists $N \in \mathbb{N}$ such that $x_n < y_n$ for all $n \geq N$.*

2. *Assume that there exists $N \in \mathbb{N}$ such that $x_n \leq y_n$ for all $n \geq N$. Then $A \leq B$.*

Proof. Suppose $A < B$, and let $\varepsilon = \frac{1}{3}(B - A) > 0$. Then there exist $N_A, N_B \in \mathbb{N}$ such that

$$\begin{aligned} n \geq N_A &\implies A - \varepsilon < x_n < A + \varepsilon \\ n \geq N_B &\implies B - \varepsilon < y_n < B + \varepsilon \end{aligned}$$

9

holds. Note now that by the choice of ε it holds $2\varepsilon < B - A$, therefore

$$A + \varepsilon < B - \varepsilon.$$

Hence, for $N = \max\{N_A, N_B\}$, we have

$$n \geq N \implies x_n < A + \varepsilon < B - \varepsilon < y_n.$$

This shows the first assertion.

For the second assertion, assume by contradiction that $A > B$. Then the first assertion implies that there exists N such that $x_N > y_N$. This contradicts the assumption and proves the result. □

Remark 2.98. — In Proposition 2.97(2), even if one assumes that $x_n < y_n$, one can only deduce that $A \leq B$. Indeed, consider the sequences

$$x_n = -\frac{1}{n}, \qquad y_n = \frac{1}{n} \qquad \forall\, n \geq 1.$$

Then $x_n < y_n$ but both sequences converge to 0.

Lemma 2.99: Sandwich Lemma

Let $(x_n)_{n=0}^{\infty}$, $(y_n)_{n=0}^{\infty}$, and $(z_n)_{n=0}^{\infty}$ be sequences of real numbers such that for some $N \in \mathbb{N}$ the inequalities $x_n \leq y_n \leq z_n$ hold for all $n \geq N$. Suppose that $(x_n)_{n=0}^{\infty}$ and $(z_n)_{n=0}^{\infty}$ are convergent and have the same limit. Then the sequence $(y_n)_{n=0}^{\infty}$ is also convergent, and it holds that

$$\lim_{n\to\infty} x_n = \lim_{n\to\infty} y_n = \lim_{n\to\infty} z_n.$$

Proof. The proof is left as an exercise for the reader. □

Exercise 2.100. — Calculate the following limits, if they exist:

$$\lim_{n\to\infty} \frac{7n^4 + 15}{3n^4 + n^3 + n - 1}, \qquad \lim_{n\to\infty} \frac{n^2 + 5}{n^3 + n + 1}, \qquad \lim_{n\to\infty} \frac{n^5 - 10}{n^2 + 1}.$$

9

2.5.4 Bounded Sequences

In this section, we study bounded sequences of real numbers.

Definition 2.101: Bounded Sequences

A sequence $(x_n)_{n=0}^{\infty}$ in $\mathbb{R}$ is called **bounded** if there exists a real number $M \geq 0$ such that $|x_n| \leq M$ for all $n \in \mathbb{N}$.

Lemma 2.102: Convergent Sequences are Bounded

Every convergent sequence is bounded.

Proof. Let $(x_n)_{n=0}^{\infty}$ be a convergent sequence with limit $A \in \mathbb{R}$. Choosing $\varepsilon = 1$ in the definition of limit, there exists $N \in \mathbb{N}$ such that $|x_n - A| \leq 1$ for all $n \geq N$. In particular, by triangle inequality (see property (g) in Paragraph 2.16),

$$|x_n| \leq |(x_n - A) + A| \leq |x_n - A| + |A| \leq 1 + |A| \qquad \forall\, n \geq N.$$

Set

$$M = \max\{1 + |A|, |x_0|, |x_1|, \ldots, |x_{N-1}|\}.$$

Then $|x_n| \leq M$ holds for all $n \in \mathbb{N}$, as required. □

Exercise 2.103. — Prove statement (2) in Proposition 2.96.
Hint: Note that

$$|x_n y_n - AB| = |x_n y_n - x_n B + x_n B - AB| \leq |y_n - B|, |x_n| + |x_n - A|, |B|$$

and note that, since $(x_n)_{n=0}^{\infty}$ converges, then $|x_n| \leq M$ for some $M > 0$.

2.104. — As we will show, bounded sequences of real numbers always have at least one accumulation point, or equivalently, a convergent subsequence. This fact gives rise to the important notion of superior limit and inferior limit.

Definition 2.105: Monotone sequence

A sequence $(x_n)_{n=0}^{\infty}$ is called:

- **(monotonically) increasing** if for all $m, n \in \mathbb{N}$: $m > n \implies x_m \geq x_n$;
- **strictly (monotonically) increasing** if for all $m, n \in \mathbb{N}$: $m > n \implies x_m > x_n$;
- **(monotonically) decreasing** if for all $m, n \in \mathbb{N}$: $m > n \implies x_m \leq x_n$;
- **strictly (monotonically) decreasing** if for all $m, n \in \mathbb{N}$: $m > n \implies x_m < x_n$.

If the sequence is increasing or decreasing, we say that it is **monotone**. If it is strictly
9 increasing or strictly decreasing, we say that it is **strictly monotone**.

Remark 2.106. — An equivalent definition of increasing/decreasing sequence can be given as follows:

- **(monotonically) increasing** if for all $n \in \mathbb{N}$ it holds $x_{n+1} \geq x_n$;
- **strictly (monotonically) increasing** if for all $n \in \mathbb{N}$ it holds $x_{n+1} > x_n$;
- **(monotonically) decreasing** if for all $n \in \mathbb{N}$ it holds $x_{n+1} \leq x_n$;
- **strictly (monotonically) decreasing** if for all $n \in \mathbb{N}$ it holds $x_{n+1} < x_n$.

2.107. — Monotone bounded sequences are always convergent. We illustrate this with a picture:

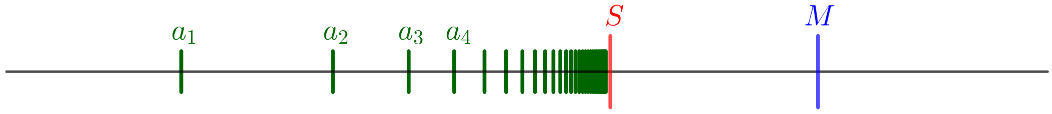


Figure 2.3

The monotonically increasing sequence $a_1, a_2, a_3, \dots$ in the figure, bounded by $M > 0$, has no choice but to converge to the supremum S of the sequence members.

Theorem 2.108: Convergence for Monotone Sequences

A monotone sequence of real numbers $(x_n)_{n=0}^{\infty}$ *converges if and only if it is bounded. If the sequence* $(x_n)_{n=0}^{\infty}$ *is monotonically increasing, then*

$$\lim_{n\to\infty} x_n = \sup\{x_n \mid n \in \mathbb{N}\}.$$

If the sequence $(x_n)_{n=0}^{\infty}$ *is monotonically decreasing, then* $\lim\limits_{n\to\infty} x_n = \inf\{x_n \mid n \in \mathbb{N}\}$.

Proof. If $(x_n)_{n=0}^{\infty}$ is convergent, it follows by Lemma 2.102 that $(x_n)_{n=0}^{\infty}$ is bounded.

Vice versa, let $(x_n)_{n=0}^{\infty}$ be a monotone bounded sequence of real numbers. To show that $(x_n)_{n=0}^{\infty}$ converges, we can assume without loss of generality that $(x_n)_{n=0}^{\infty}$ is monotonically increasing (otherwise we simply replace $(x_n)_{n=0}^{\infty}$ by $(-x_n)_{n=0}^{\infty}$). Since x_n is bounded, there exists $M > 0$ such that $x_n \leq M$ for all $n \in \mathbb{N}$. This means that the set $\{x_n \mid n \in \mathbb{N}\}$ is bounded from above, so the supremum $A = \sup\{x_n \mid n \in \mathbb{N}\}$ exists. We now want to prove that x_n converges to A.
By the definition of supremum, we have:
(i) $x_n \leq A$ for all $n \in \mathbb{N}$;
(ii) for every $\varepsilon > 0$ there exists $N \in \mathbb{N}$ with $x_N > A - \varepsilon$.
Thus, for $n \geq N$, it follows from (i), (ii), and the monotonicity of $(x_n)_{n=0}^{\infty}$, that

9 $$A - \varepsilon < x_N \leq x_n \leq A < A + \varepsilon$$

holds. This proves that $\lim_{n\to\infty} x_n = A$, as desired. □

Remark 2.109. — If $(x_n)_{n=0}^{\infty}$ is monotone and there exists a bounded subsequence $(x_{n_k})_{k=0}^{\infty}$, then the whole sequence is bounded (and therefore converges, thanks to Theorem 2.108).

Indeed, assume for instance that $(x_n)_{n=0}^{\infty}$ is increasing and the subsequence $(x_{n_k})_{k=0}^{\infty}$ is bounded from above by a number M. Then, recalling Remark 2.86, by monotonicity we have

$$x_0 \leq x_k \leq x_{n_k} \leq M \qquad \forall\, k \in \mathbb{N}.$$

So, $(x_n)_{n=0}^{\infty}$ is bounded. The case when $(x_n)_{n=0}^{\infty}$ is increasing is analogous.

Exercise 2.110. — Let $(x_n)_{n=0}^{\infty}$ be the sequence given by $x_0 = 1$ and

$$x_n = \frac{2}{3}\left(x_{n-1} + \frac{1}{x_{n-1}}\right)$$

for $n \geq 1$. Show that $(x_n)_{n=0}^{\infty}$ converges and determine the limit.
Hint: First, prove that the sequence converges to a nonnegative limit. Second, show that if $A \geq 0$ is the limit, then it satisfies $A = \frac{2}{3}\left(A + \frac{1}{A}\right)$. Use this relation to identify A.

EXERCISE 2.111. — Let $(x_n)_{n=0}^{\infty}$ be a monotonically increasing sequence and $(y_n)_{n=0}^{\infty}$ be a monotonically decreasing sequence with $x_n \leq y_n$ for all $n \in \mathbb{N}$. Show that both sequences converge and that $\lim_{n\to\infty} x_n \leq \lim_{n\to\infty} y_n$ holds. Illustrate your argument with a picture similar to the one in Figure 2.3.

2.112. — Let $(x_n)_{n=0}^{\infty}$ be a bounded sequence of real numbers. For the definition of limits and accumulation points of the sequence $(x_n)_{n=0}^{\infty}$, only its long-term behavior is relevant, or more precisely, the end $(x_k)_{n=N}^{\infty}$ for arbitrarily large $N \in \mathbb{N}$. Following this observation, for each $n \in \mathbb{N}$ we define the supremum

$$s_n = \sup \{x_k \mid k \geq n\}$$

over the final part $\{x_k \mid k \geq n\}$ of the sequence. Since $\{x_k \mid k \geq m\} \subset \{x_k \mid k \geq n\}$ for $m > n$, it follows that $s_m \leq s_n$ for $m > n$. The sequence $(s_n)_{n=0}^{\infty}$ is therefore monotonically decreasing.

Since $(x_n)_{n=0}^{\infty}$ is bounded by assumption, $(s_n)_{n=0}^{\infty}$ is also bounded, and so it is a monotonically decreasing bounded sequence. Therefore, the sequence $(s_n)_{n=0}^{\infty}$ converges to the infimum of the set $\{s_n \mid n \in \mathbb{N}\}$ by Theorem 2.108. This infimum is called the **superior limit** of the given sequence $(x_n)_{n=0}^{\infty}$.

Analogously, one can define the **inferior limit** of $(x_n)_{n=0}^{\infty}$ considering

9

$$i_n = \inf \{x_k \mid k \geq n\}.$$

In this case $(i_n)_{n=0}^{\infty}$ is a monotonically increasing bounded sequence, so its limit exists and it is called inferior limit.

We remark that, by the definition of i_n and s_n, the following inequality holds:

$$i_n \leq x_n \leq s_n \qquad \forall\, n \in \mathbb{N}. \tag{2.10}$$

DEFINITION 2.113: SUPERIOR AND INFERIOR LIMITS

Let $(x_n)_{n=0}^{\infty}$ be a bounded sequence of real numbers. The real numbers defined by

$$\limsup_{n\to\infty} x_n = \lim_{n\to\infty} \Big(\sup\{x_k \mid k \geq n\}\Big)$$

and

$$\liminf_{n\to\infty} x_n = \lim_{n\to\infty} \Big(\inf\{x_k \mid k \geq n\}\Big)$$

are called **superior limit**, respectively **inferior limit** of the sequence $(x_n)_{n=0}^{\infty}$. Note that, as a consequence of (2.10) and Proposition 2.97,

$$\liminf_{n\to\infty} x_n \leq \limsup_{n\to\infty} x_n.$$

EXAMPLE 2.114. — Let $(x_n)_n$ be the sequence defined by $x_n = (-1)^n + \frac{1}{n}$ for $n \in \mathbb{N}$. We represent x_n, s_n, and i_n in the following table.

n	1	2	3	4	5	6	7	8	...
x_n	0	$\frac{3}{2}$	$-\frac{2}{3}$	$\frac{5}{4}$	$-\frac{4}{5}$	$\frac{7}{6}$	$-\frac{6}{7}$	$\frac{9}{8}$	...
s_n	$\frac{3}{2}$	$\frac{3}{2}$	$\frac{5}{4}$	$\frac{5}{4}$	$\frac{7}{6}$	$\frac{7}{6}$	$\frac{9}{8}$	$\frac{9}{8}$	...
i_n	-1	-1	-1	-1	-1	-1	-1	-1	...

9 Note here that $s_n = x_n$ when n is even, and $s_n = x_{n+1}$ otherwise. Therefore $\limsup_{n\to\infty}((-1)^n + \frac{1}{n}) = \lim_{n\to\infty}((-1)^{2n} + \frac{1}{2n}) = 1$.
Because $x_n \geq -1$ for every n and $\lim_{n\to\infty} x_{2n+1} = -1$, then $i_n = -1$ for all $n \in \mathbb{N}$. Therefore $\liminf_{n\to\infty}((-1)^n + \frac{1}{n}) = -1$.

LEMMA 2.115: CONVERGENCE VS SUPERIOR AND INFERIOR LIMITS

A bounded sequence $(x_n)_{n=0}^\infty$ *in* $\mathbb{R}$ *converges if and only if* $\limsup_{n\to\infty} x_n = \liminf_{n\to\infty} x_n$.

Proof. Define

$$i_n = \inf\{x_n \mid n \geq N\}, \qquad s_n = \sup\{x_n \mid n \geq N\},$$

$$I = \lim_{n\to\infty} i_n = \liminf_{n\to\infty} x_n, \qquad S = \lim_{n\to\infty} s_n = \limsup_{n\to\infty} x_n.$$

We now prove the result.

10 Suppose first that $I = S$ holds. Recalling (2.10), since by assumption i_n and s_n converge to the same limit, the Sandwich Lemma 2.99 implies that x_n converges to $I = S$.

Conversely, assume that $(x_n)_{n=0}^\infty$ converges to $A \in \mathbb{R}$, and let $\varepsilon > 0$. Then there exists $N \in \mathbb{N}$ such that $A - \varepsilon < x_n < A + \varepsilon$ for all $n \geq N$. In particular

$$A - \varepsilon \leq i_n \leq s_n \leq A + \varepsilon,$$

and therefore it follows from Proposition 2.97 that

$$A - \varepsilon \leq I \leq S \leq A + \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, this implies that $A \leq I \leq S \leq A$, so $A = I = S$. □

Theorem 2.116: Superior and Inferior Limits are Accumulation Points

Let $(x_n)_{n=0}^{\infty}$ be a bounded sequence and $A = \limsup_{n\to\infty} x_n$. Then A is an accumulation point and, for every $\varepsilon > 0$, the following hold:
(1) there are only finitely many elements satisfying $x_n \geq A + \varepsilon$;
(2) infinitely many elements satisfy $A - \varepsilon < x_n < A + \varepsilon$.
An analogous statement holds for the inferior limit.

Proof. Let $\varepsilon > 0$, and write $s_n = \sup\{x_k \mid k \geq n\}$. The sequence $(s_n)_{n=0}^{\infty}$ is monotonically decreasing and converges to A. So there exists $N_0 \in \mathbb{N}$ such that

$$A \leq s_n < A + \varepsilon \qquad \forall\, n \geq N_0. \tag{2.11}$$

Therefore, given $N \in \mathbb{N}$, if we define $N_1 := \max\{N, N_0\}$, then s_{N_1} satisfies (2.11).
Now, since $s_{N_1} = \sup\{x_k \mid k \geq N_1\}$, there exists $n_1 \geq N_1$ satisfying $x_{n_1} > s_{N_1} - \varepsilon$, therefore $x_{n_1} > A - \varepsilon$ (since $s_{N_1} \geq A$ by (2.11)). On the other hand, since $s_{n_1} \leq s_{N_1}$ (recall that s_n is decreasing), (2.10) and (2.11) imply that $x_{n_1} < A + \varepsilon$.
In other words, given $N \in \mathbb{N}$ we proved the existence of an index $n_1 \geq N_1 \geq N$ such that

10

$$A - \varepsilon < x_{n_1} < A + \varepsilon.$$

This shows that A is an accumulation point.

To prove (1), it follows from (2.10) and (2.11) that $x_n < A + \varepsilon$ for all $n \geq N_0$. This implies that only finitely many elements satisfy $x_n \geq A + \varepsilon$.

Since A is an accumulation point, (2) follows from Corollary 2.91. □

Corollary 2.117: Bounded Sequences have Convergent Subsequences

Every bounded sequence of real numbers has an accumulation point and has a convergent subsequence.

Proof. By Theorem 2.116, the limsup (and analogously for the liminf) is always an accumulation point. Also, by Proposition 2.90, every accumulation point is the limit of a converging subsequence. So, a convergent subsequence always exists. □

Exercise 2.118. — Let $(x_n)_{n=0}^{\infty}$ be a bounded sequence in $\mathbb{R}$, and let $E \subseteq \mathbb{R}$ be the set of accumulation points of the sequence $(x_n)_{n=0}^{\infty}$. Show that

$$\limsup_{n\to\infty} x_n = \max E \qquad \text{and} \qquad \liminf_{n\to\infty} x_n = \min E.$$

Exercise 2.119. — Let $(a_n)_{n=0}^{\infty}$, $(b_n)_{n=0}^{\infty}$ and $(c_n)_{n=0}^{\infty}$ be convergent sequences of real numbers, with limits A, B, and C respectively. Let $(x_n)_{n=0}^{\infty}$ be the sequence defined by

$$x_n = \begin{cases} a_n & \text{if } n = 3k,\ k \in \mathbb{N} \\ b_n & \text{if } n = 3k+1,\ k \in \mathbb{N} \\ c_n & \text{if } n = 3k+2,\ k \in \mathbb{N} \end{cases}.$$

Calculate $\limsup_{n\to\infty} x_n$, $\liminf_{n\to\infty} x_n$, and the set of accumulation points of the sequence $(x_n)_{n=0}^{\infty}$.

Exercise 2.120. — Let $(x_n)_{n=0}^{\infty}$ be a bounded sequence of real numbers such that $(x_{n+1} - x_n)_{n=0}^{\infty}$ converges to 0. Set

$$A = \liminf_{n\to\infty} x_n \qquad \text{and} \qquad B = \limsup_{n\to\infty} x_n.$$

Show that the set of accumulation points of the sequence $(x_n)_{n=0}^{\infty}$ is the interval $[A, B]$. Construct an example of such a sequence with $[A, B] = [0, 1]$.

2.5.5 Cauchy-Sequences

Definition 2.121: Cauchy-sequences

10 A sequence $(x_n)_{n=0}^{\infty}$ in $\mathbb{R}$ is a **Cauchy sequence** if for every $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that $|x_n - x_m| < \varepsilon$ for all $m, n \geq N$.

Lemma 2.122: Boundedness of Cauchy Sequences

Cauchy sequences are bounded.

Proof. The proof is very similar to the one of Lemma 2.102. Choosing $\varepsilon = 1$ in the definition of Cauchy sequence, there exists $N \in \mathbb{N}$ such that $|x_n - x_N| \leq 1$ for all $n \geq N$. In particular $|x_n| \leq 1 + |x_N|$ for all $n \geq N$, and therefore

$$|x_n| \leq M = \max\{1 + |x_N|, |x_0|, |x_1|, \ldots, |x_{N-1}|\} \qquad \forall\, n \in \mathbb{N}.$$

□

Exercise 2.123. — Show that a Cauchy sequence converges if and only if it has a convergent subsequence.

Theorem 2.124: Convergence vs Cauchy Sequences

A sequence of real numbers converges if and only if it is a Cauchy sequence.

Proof. Let $(x_n)_{n=0}^{\infty}$ be sequence converging to some limit A, and fix $\varepsilon > 0$. Then there exists N such that for all $n \geq N$ it holds $|x_n - A| < \frac{\varepsilon}{2}$. This implies that

$$|x_n - x_m| \leq |x_n - A| + |x_m - A| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \qquad \forall\, n, m \geq N,$$

hence $(x_n)_{n=0}^{\infty}$ is a Cauchy sequence.

Vice versa, let $(x_n)_{n=0}^{\infty}$ be a Cauchy sequence. Since it is bounded (see Lemma 2.122), Corollary 2.117 implies that $(x_n)_{n=0}^{\infty}$ has a subsequence $(x_{n_k})_{k=0}^{\infty}$ converging to some limit $A \in \mathbb{R}$. Then, given $\varepsilon > 0$, consider N_0 such that

$$|x_n - x_m| < \frac{\varepsilon}{2} \qquad \forall\, m, n \geq N_0,$$

and consider N_1 such that

$$|x_{n_k} - A| < \frac{\varepsilon}{2} \qquad \forall\, k \geq N_1.$$

Then, choosing $N = \max\{N_0, N_1\}$ and recalling that $n_N \geq N$ (see Remark 2.86), it follows
10 that

$$|x_n - A| \leq |x_n - x_{n_N}| + |x_{n_N} - A| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \qquad \forall\, n \geq N,$$

so the whole sequence $(x_n)_{n=0}^{\infty}$ converges to A. □

EXAMPLE 2.125. — Let $(x_n)_{n=0}^{\infty}$ be a sequence of real numbers. We claim that the condition

$$\forall\, \varepsilon > 0\ \exists\, N \in \mathbb{N} \text{ such that } \forall\, n \geq N : |x_n - x_{n+1}| < \varepsilon$$

is not equivalent to the convergence of the sequence. As a counterexample, consider the sequence

$$0, 1, 1 + \tfrac{1}{2}, 2, 2 + \tfrac{1}{3}, 2 + \tfrac{2}{3}, 3, 3 + \tfrac{1}{4}, 3 + \tfrac{2}{4}, 3 + \tfrac{3}{4}, 4, 4 + \tfrac{1}{5}, 4 + \tfrac{2}{5}, 4 + \tfrac{3}{5}, 4 + \tfrac{4}{5}, 5, 5 + \tfrac{1}{6} \ldots$$

that progresses between $n-1$ and n in steps of length $\frac{1}{n}$. This sequence is unbounded and thus not convergent. On the other hand, the distance between two successive elements decreases as the sequence progresses, and becomes arbitrarily small.

2.5.6 Improper Limits

11 We also introduce the **improper limit values** $+\infty$ (often abbreviated to ∞) and $-\infty$ for sequences.

Definition 2.126: Improper Limits

Let $(x_n)_{n=0}^{\infty}$ be a sequence of real numbers. We say $(x_n)_{n=0}^{\infty}$ **diverges to** ∞, and we write

$$\lim_{n\to\infty} x_n = \infty.$$

if for every $M > 0$ there exists $N \in \mathbb{N}$ such that $x_n > M$ for all $n \geq N$.
Similarly, we say that $(x_n)_{n=0}^{\infty}$ **diverges to** $-\infty$ if for every real number $M > 0$ there exists $N \in \mathbb{N}$ such that $x_n < -M$ for all $n \geq N$.
In both cases, we speak of an **improper limit**.

2.127. — An unbounded sequence may not diverge to ∞ or $-\infty$. For example, the sequence

$$0, -1, 2, -3, 4, -5, 6, -7, 8, -9, \dots,$$

i.e., $x_n = (-1)^n n$ is unbounded, but it neither diverges to ∞, nor diverges to $-\infty$.

Exercise 2.128. — Let $(x_n)_{n=0}^{\infty}$ be an unbounded sequence of real numbers. Show that there exists a subsequence that diverges to ∞ or to $-\infty$.

11 2.129. — We can use improper limits to define the superior and inferior limits for unbounded sequences in $\mathbb{R}$. If the sequence $(x_n)_{n=0}^{\infty}$ is not bounded from above, then $\sup\{x_n \,|\, k \geq n\} = \infty$ for all $n \in \mathbb{N}$ and we write

$$\limsup_{n\to\infty} x_n = \infty.$$

If $(x_n)_{n=0}^{\infty}$ is bounded from above but not from below, then we write

$$\limsup_{n\to\infty} x_n = \lim_{n\to\infty} \big(\sup\{x_k \mid k \geq n\}\big)$$

where the right-hand side is a real limit if the monotonically decreasing sequence $\sup\{x_k \,|\, k \geq n\}$ is bounded, and the improper limit $-\infty$ otherwise. We use this terminology analogously for the inferior limit.

Exercise 2.130. — Prove the following version of the sandwich lemma for improper limits. For two sequences of real numbers $(x_n)_{n=0}^{\infty}$ and $(x_n)_{n=0}^{\infty}$ with $x_n \leq y_n$ for all $n \in \mathbb{N}$ holds:

$$\lim_{n\to\infty} x_n = \infty \implies \lim_{n\to\infty} y_n = \infty$$
$$\lim_{n\to\infty} y_n = -\infty \implies \lim_{n\to\infty} x_n = -\infty.$$

2.6 Sequences of Complex Numbers

To study sequences in $\mathbb{C}$ it is often sufficient to consider the corresponding sequences of real and imaginary parts in $\mathbb{R}$.

Definition 2.131: Sequences of Complex Numbers

A sequence of complex numbers $(z_n)_{n=0}^{\infty} = (x_n + iy_n)_{n=0}^{\infty}$ is convergent with limit $A + iB \in \mathbb{C}$ if the two sequences of real numbers $(x_n)_{n=0}^{\infty}$ and $(y_n)_{n=0}^{\infty}$ are convergent, with limits A and B, respectively.
We say that $(z_n)_{n=0}^{\infty} = (x_n + iy_n)_{n=0}^{\infty}$ diverges to ∞ if the sequence $(|z_n|)_{n=0}^{\infty}$ diverges toward ∞, that is,

$$\lim_{n\to\infty} |z_n| = \lim_{n\to\infty} \sqrt{x_n^2 + y_n^2} = \infty.$$

11

Remark 2.132. — As we did for sequences of real numbers, also for sequences of complex numbers we can consider subsequences. This corresponds to consider $(z_{n_k})_{k=0}^{\infty} = (x_{n_k} + iy_{n_k})_{k=0}^{\infty}$ for some strictly increasing sequence of nonnegative integers $(n_k)_{k=0}^{\infty}$.

Exercise 2.133. — Let $(z_n)_{n=0}^{\infty}$ be a convergent sequence in $\mathbb{C}$. Show that $(|z_n|)_{n=0}^{\infty}$ converges, and find the limit. Conversely, does the convergence of $(|z_n|)_{n=0}^{\infty}$ imply the convergence of $(z_n)_{n=0}^{\infty}$?

Exercise 2.134. — Given a complex number $z \in \mathbb{C}$, consider the so-called **geometric sequence** $z_n = z^n$. Determine the set of all complex numbers z for which the sequence $(z^n)_{n=0}^{\infty}$ converges.

Chapter 3

Functions of one Real Variable

In this chapter, we discuss real-valued functions defined on a subset of $\mathbb{R}$, typically an interval. The central concept of this chapter is the one of continuity.

3.1 Real-valued Functions

3.1.1 Boundedness and Monotonicity

In this section, we discuss two elementary properties of real-valued functions that we have already encountered for sequences: boundedness and monotonicity. A **real-valued** function is any function with values in $\mathbb{R}$. As soon as we talk about monotonicity, we will assume that the domain of the functions we consider is a non-empty subset of $\mathbb{R}$. Informally, we then speak of functions in one real **variable**.

3.1. — For an arbitrary non-empty set $D \subset \mathbb{R}$, we define the set of **real-valued** functions on D as

11 $$\mathcal{F}(D) = \{f \mid f : D \to \mathbb{R}\}.$$

One can define addition and scalar multiplication as

$$(f_1 + f_2)(x) = f_1(x) + f_2(x) \qquad \text{and} \qquad (\alpha f_1)(x) = \alpha f_1(x)$$

for $f_1, f_2 \in \mathcal{F}(D)$, $\alpha \in \mathbb{R}$, and $x \in D$. Also, functions in $\mathcal{F}(D)$ can be multiplied as follows:

$$(f_1 f_2)(x) = f_1(x) f_2(x)$$

for all $f_1, f_2 \in \mathcal{F}(D)$ and $x \in D$.

Given $a \in \mathbb{R}$, we use the notation $f \equiv a$ to denote the constant function $f(x) = a$ for all $x \in D$.

Remark 3.2. — The interested reader may notice that, with the addition and multiplication defined above, $\mathcal{F}(D)$ is a commutative ring (the neutral element for addition is the constant function $f \equiv 0$, the neutral element for multiplication is the constant function $f \equiv 1$).

We say that $x \in D$ is a **zero** of $f \in \mathcal{F}(D)$ if $f(x) = 0$ holds. The **zero set** of f is defined by $\{x \in D \mid f(x) = 0\}$. Finally, we define an order relation on $\mathcal{F}(D)$: given $f_1, f_2 \in \mathcal{F}(D)$ we say

$$f_1 \leq f_2 \iff f_1(x) \leq f_2(x) \quad \forall\, x \in D,$$

$$f_1 < f_2 \iff f_1(x) < f_2(x) \quad \forall\, x \in D,$$

We say that $f \in \mathcal{F}(D)$ is **non-negative** if $f \geq 0$, and $f \in \mathcal{F}(D)$ is **positive** if $f > 0$.

Exercise 3.3. — Let $N_1, N_2 \subset D$ be the set of zeros of $f_1 \in \mathcal{F}(D)$ and $f_2 \in \mathcal{F}(D)$ respectively. What is the set of zeros of $f_1 f_2$?

11 Exercise 3.4. — Verify that the relation $\leq$ defined above on $\mathcal{F}(D)$ is indeed an order relation.

Definition 3.5: Bounded Functions

Let D be a non-empty set, and $f : D \to \mathbb{R}$ a function. We say that f is **bounded from above** if there exists $M > 0$ such that

$$f(x) \leq M \qquad \forall\, x \in D.$$

We say that f is **bounded from below** if there exists $M > 0$ such that

$$f(x) \geq -M \qquad \forall\, x \in D.$$

Finally, we say that f is **bounded** if f is bounded from above and from below. Equivalently, f is bounded if there exists $M > 0$ such that

$$|f(x)| \leq M \qquad \forall\, x \in D.$$

Definition 3.6: Monotone Functions

Let $D \subset \mathbb{R}$ and $f : D \to \mathbb{R}$. The function f is called:

1. **(monotonically) increasing** if for all $x, y \in D$:

$$x < y \implies f(x) \leq f(y);$$

2. **strictly (monotonically) increasing** if for all $x, y \in D$:

$$x < y \implies f(x) < f(y);$$

3. **(monotonically) decreasing** if for all $x, y \in D$:

$$x < y \implies f(x) \geq f(y);$$

4. **strictly (monotonically) decreasing** if for all $x, y \in D$:

$$x < y \implies f(x) > f(y).$$

We call a function $f : D \to \mathbb{R}$ **monotone** if it is monotonically increasing or monoton-
11 ically decreasing, and **strictly monotone** if it is strictly monotonically increasing or strictly monotonically decreasing.

Example 3.7. —

- Let $D = [a, b]$ be an interval, and $f : D \to \mathbb{R}$ be the function $f(x) = x^2$. The function f is strictly increasing if $a \geq 0$, and strictly decreasing if $b \leq 0$. If $a < 0 < b$ holds, then f is not monotone.
- For any subset $D \subset \mathbb{R}$ and any odd integer number $n \geq 0$, the function $x \mapsto x^n$ on D is strictly monotonically increasing.
- The rounding function $\lfloor \cdot \rfloor : \mathbb{R} \mapsto \mathbb{R}$ (recall Definition 2.64) is increasing, but not strictly increasing.
- A constant function is both decreasing and monotonically increasing. Conversely, a function on $D \subseteq \mathbb{R}$ that is both monotonically decreasing and monotonically increasing is necessarily constant.

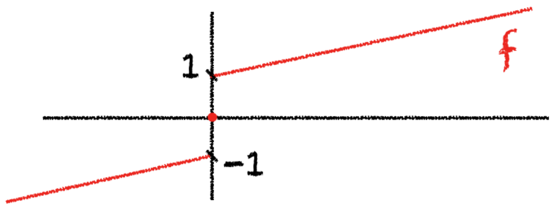


Figure 3.1: A strictly monotone function is always injective. However, it need not be surjective. For example, the function $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = \frac{1}{8}x + \operatorname{sgn}(x)$ is strictly monotone increasing but not subjective, (e.g., $\frac{1}{2}$ is not in the image of f).

Exercise 3.8. — Let $D \subseteq \mathbb{R}$, and let $f_1, f_2 \in \mathcal{F}(D)$ be strictly increasing. Show that:

(i) $f_1 + f_2 \in \mathcal{F}(D)$ is strictly increasing;

(ii) given $a \in \mathbb{R}$, the function $af \in \mathcal{F}(D)$ is strictly increasing if $a > 0$, and strictly decreasing if $a < 0$;

(iii) if $f_1 > 0$ and $f_2 > 0$, then $f_1 f_2$ is strictly increasing.

3.1.2 Continuity

Definition 3.9: Continuous Functions

Let $D \subseteq \mathbb{R}$ be a subset and let $f : D \to \mathbb{R}$ be a function. We say that f is **continuous at a point** $x_0 \in D$ if for all $\varepsilon > 0$ there exists $\delta > 0$ such that the following implication holds:

11

$$\forall\, x \in D, \qquad |x - x_0| < \delta \implies |f(x) - f(x_0)| < \varepsilon.$$

The function f is **continuous in** D if it is continuous at every point of D.

Remark 3.10. — In the definition of continuity, it is only important to check the implication for $\varepsilon > 0$ small. Indeed, assume that a function f satisfies the following:
There exists $\varepsilon_0 > 0$ such that, for all $\varepsilon \in (0, \varepsilon_0]$ there exists $\delta > 0$ such that

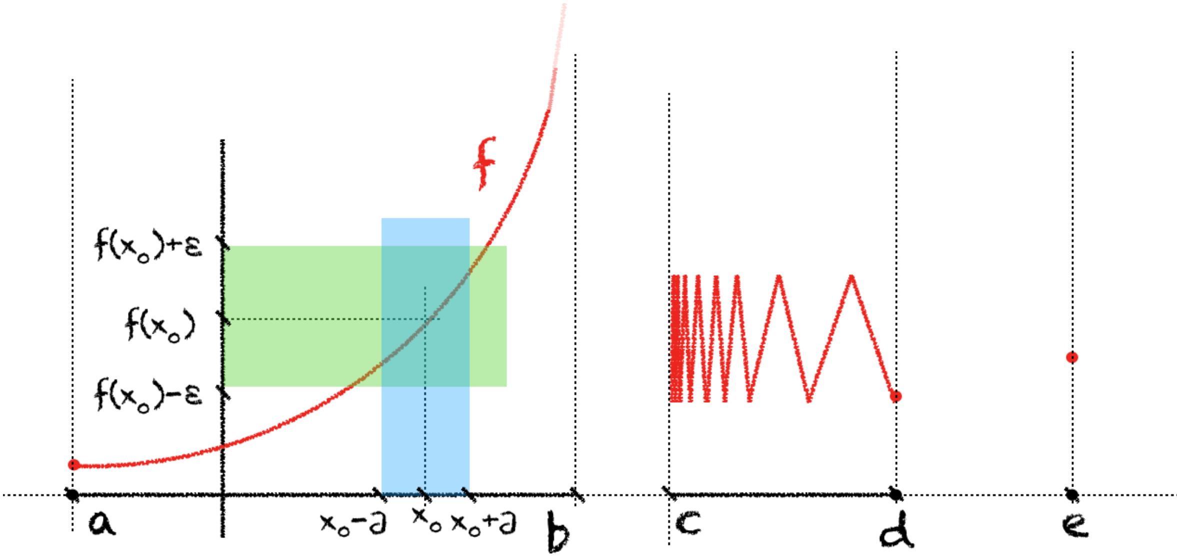
$$\forall\, x \in D, \qquad |x - x_0| < \delta \implies |f(x) - f(x_0)| < \varepsilon.$$

Then f is continuous at x_0. Indeed, for $\varepsilon > \varepsilon_0$ we can choose the number $\delta > 0$ corresponding to ε_0 to get

$$\forall\, x \in D, \qquad |x - x_0| < \delta \implies |f(x) - f(x_0)| < \varepsilon_0 < \varepsilon.$$

In other words, if δ works for ε_0, then it works for all $\varepsilon > \varepsilon_0$.

3.11. — The following illustration shows a continuous function on $D = [a, b) \cup (c, d] \cup \{e\}$. We see that f is continuous at every point x_0: no matter how small one chooses $\varepsilon > 0$, for a suitable $\delta > 0$ we have that for all x that are δ-close to x_0, $f(x)$ is also ε-close to $f(x_0)$.



Applet 3.12 (Continuity). *We consider a function that is continuous at most (but not all) points in the domain of definition.*

Example 3.13. — • Let a and b be real numbers. The affine function $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = ax + b$ is continuous.

11 Indeed, if $a = 0$, then the function is constant and therefore continuous. So, let $a \neq 0$. Given $x_0 \in \mathbb{R}$ and $\varepsilon > 0$, note that $|f(x) - f(x_0)| = |a||x - x_0|$ holds for all $x \in \mathbb{R}$. Thus, considering the choice $\delta = \frac{\varepsilon}{|a|}$, for any $x \in \mathbb{R}$ with $|x - x_0| < \delta$ we have

$$|f(x) - f(x_0)| = |a||x - x_0| < |a|\delta = \varepsilon$$

and thus f is continuous.

- The function $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = |x|$ is also continuous. Indeed, let $x_0 \in \mathbb{R}$ and $\varepsilon > 0$. Choosing $\delta = \varepsilon$ we notice that if $|x - x_0| < \delta$ then

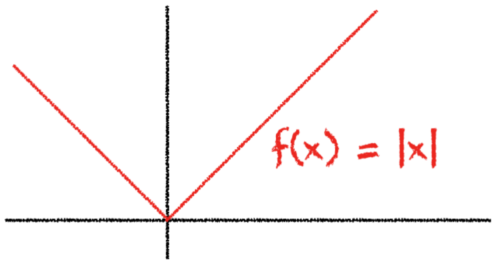
$$|f(x) - f(x_0)| = \big||x| - |x_0|\big| \leq |x - x_0| < \delta = \varepsilon$$

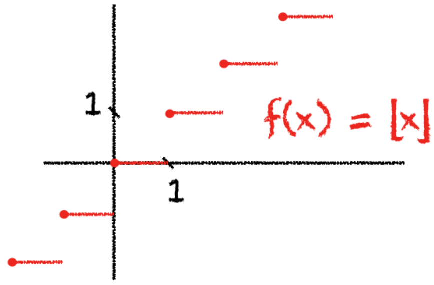
by the inverse triangle inequality.

- The rounding function $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = \lfloor x \rfloor$ (recall Definition 2.64) is not continuous at points in $\mathbb{Z}$. Indeed, if $x_0 \in \mathbb{Z}$, then for any $\delta > 0$ small it holds that $|(x_0 - \frac{1}{2}\delta) - x_0| < \delta$ and

$$\big|\lfloor x_0 - \tfrac{1}{2}\delta \rfloor - \lfloor x_0 \rfloor\big| = \lfloor x_0 \rfloor - \lfloor x_0 - \tfrac{1}{2}\delta \rfloor = x_0 - (x_0 - 1) = 1.$$

Thus, the continuity condition is not satisfied at $x_0 \in \mathbb{Z}$ for any $\varepsilon < 1$.





Exercise 3.14. — Show that the functions $f : \mathbb{R} \to \mathbb{R}$ and $g : [0, \infty) \to \mathbb{R}$ given by
11 $f(x) = x^2$ and $g(x) = \sqrt{x}$ are both continuous.

Interlude: Restriction

Given $D \subseteq \mathbb{R}$ and a function $f : D \to \mathbb{R}$ defined on D, for any $D' \subset D$ one can consider the **restriction** of f to D': this is the function $f|_{D'} : D' \to \mathbb{R}$ defined as

$$f|_{D'}(x) = f(x) \qquad \forall\, x \in D'.$$

Note that we consider $f|_{D'}$ and f as different functions, since their domains of definition are not the same – except, of course, when $D' = D$.

Exercise 3.15. — Let $D \subseteq \mathbb{R}$ and $f : D \to \mathbb{R}$ be a continuous function. Let $D' \subset D$. Show that the restriction $f|_{D'}$ is continuous.

Proposition 3.16: Combination of Continuous Functions

Let $D \subseteq \mathbb{R}$, and let $f_1, f_2 : D \to \mathbb{R}$ be functions continuous at a point $x_0 \in D$. Then also the functions $f_1 + f_2$, $f_1 f_2$, and $a f_1$ for $a \in \mathbb{R}$ are continuous at x_0.

Proof. Let $\varepsilon > 0$. Since f_1 and f_2 are continuous at x_0, there exist $\delta_1, \delta_2 > 0$ such that, for all $x \in D$, it holds

12
$$|x - x_0| < \delta_1 \implies |f_1(x) - f_1(x_0)| < \frac{\varepsilon}{2}$$
$$|x - x_0| < \delta_2 \implies |f_2(x) - f_2(x_0)| < \frac{\varepsilon}{2}.$$

Hence, if we set $\delta = \min\{\delta_1, \delta_2\} > 0$, then

$$|x - x_0| < \delta \implies |(f_1+f_2)(x) - (f_1+f_2)(x_0)| \leq |f_1(x) - f_1(x_0)| + |f_2(x) - f_2(x_0)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, we obtain that $f_1 + f_2$ is continuous at $x_0 \in D$.

The argument for $f_1 f_2$ is similar but a little more complicated. Given $x \in D$, using the triangle inequality we have the estimate

$$\begin{aligned}|f_1(x)f_2(x) - f_1(x_0)f_2(x_0)| &= |f_1(x)f_2(x) - f_1(x_0)f_2(x) + f_1(x_0)f_2(x) - f_1(x_0)f_2(x_0)| \\ &\leq |f_1(x)f_2(x) - f_1(x_0)f_2(x)| + |f_1(x_0)f_2(x) - f_1(x_0)f_2(x_0)| \\ &= |f_1(x) - f_1(x_0)||f_2(x)| + |f_1(x_0)||f_2(x) - f_2(x_0)|.\end{aligned}$$

Let $\varepsilon > 0$ and choose $\delta_1, \delta_2 > 0$ such that, for $x \in D$,

$$\begin{aligned}|x - x_0| < \delta_1 &\implies |f_1(x) - f_1(x_0)| < \frac{\varepsilon}{2(|f_2(x_0)| + \varepsilon)}, \\ |x - x_0| < \delta_2 &\implies |f_2(x) - f_2(x_0)| < \frac{\varepsilon}{2(|f_1(x_0)| + 1)}.\end{aligned}$$

12 Then, for $x \in D$ with $|x - x_0| < \delta = \min\{\delta_1, \delta_2\}$, it holds

$$\begin{aligned}|f_2(x)| &= |f_2(x) - f_2(x_0) + f_2(x_0)| \leq |f_2(x) - f_2(x_0)| + |f_2(x_0)| \\ &< \frac{\varepsilon}{2(|f_1(x_0)| + 1)} + |f_2(x_0)| < \varepsilon + |f_2(x_0)|,\end{aligned}$$

therefore

$$|f_1(x) - f_1(x_0)||f_2(x)| < \frac{\varepsilon}{2(|f_2(x_0)| + \varepsilon)}(\varepsilon + |f_2(x_0)|) = \frac{\varepsilon}{2}.$$

Analogously, for the second term, for $x \in D$ with $|x - x_0| < \delta$ we have

$$|f_1(x_0)||f_2(x) - f_2(x_0)| < |f_1(x_0)|\frac{\varepsilon}{2(|f_1(x_0)| + \varepsilon)} < \frac{\varepsilon}{2}.$$

Combining the inequalities above, we obtain $|f_1(x)f_2(x) - f_1(x_0)f_2(x_0)| < \varepsilon$ as desired.

The statement about af_1 for $a \in \mathbb{R}$ follows choosing f_2 equal to the constant function $f_2 \equiv a$ and applying the previous result. □

Interlude: Sum Notation

Let $n \geq 1$ be an integer, and let $a_1, \ldots, a_n$ be real numbers. In the following, we want to use the commonly used sum notation

$$\sum_{j=1}^{n} a_j = a_1 + a_2 + \cdots + a_n,$$

We will refer to a_j as the **summand** and j as the **index** or the **running variable** of the sum. If J is a finite set, and if for each $j \in J$ a number a_j is given, we write

$$\sum_{j \in J} a_j$$

for the sum of all numbers in the set $\{a_j \mid j \in J\}$. Finally, we will use the convention $\sum_{j \in \emptyset} a_j = 0$ for the sum over the empty index set.

Example 3.17. — Polynomial functions are functions of the form

$$p(x) = a_0 + a_1 x + \ldots + a_n x^n = \sum_{j=0}^{n} a_j x^j$$

12 with $a_j \in \mathbb{R}$. We claim that these functions are continuous. Indeed, since $x \mapsto x$ is continuous, also $x \mapsto x^2 = x \cdot x$ is continuous, $x \mapsto x^3 = x \cdot x^2$ is continuous, and so on. This means that all the functions $\{a_0, a_1 x, \ldots, a_n x^n\}$ are continuous, and therefore their sum is continuous.

Remark 3.18. — It is also common to use the notation

$$\prod_{j=1}^{n} a_j = a_1 a_2 \cdots a_n$$

for **products**. In this context, a_j are referred to as **factors**. Also, by convention $\prod_{j \in \emptyset} a_j = 1$.

Interlude: Composition of Functions

Let $f : X \to Y$ and $g : Y \to Z$ be functions. Then the **composition** of f with g is the function

$$g \circ f : X \to Z$$

which is defined by $g \circ f(x) = g(f(x))$ for all $x \in X$.

Remark 3.19. — Let $f : W \to X$, $g : X \to Y$, $h : Y \to Z$ be functions. Then, we can consider both the function $h \circ (g \circ f) : W \to Z$ and the function $(h \circ g) \circ f : W \to Z$. However, the parentheses are irrelevant because the following is true for all $w \in W$:

$$h \circ (g \circ f)(w) = h(g \circ f(w)) = h(g(f(w)) = h \circ g(f(w)) = (h \circ g) \circ f(w).$$

In other words, $h \circ (g \circ f) = (h \circ g) \circ f$ and we say that the composition of functions is **associative**. Therefore, we write $h \circ g \circ f : W \to Z$.

Proposition 3.20: Composition of Continuous Functions

Let D_1 and D_2 be subsets of $\mathbb{R}$ and let $x_0 \in D_1$. Let $f : D_1 \to D_2$ be a continuous function at x_0 and let $g : D_2 \to \mathbb{R}$ be a continuous function at $f(x_0)$. Then $g \circ f : D_1 \to \mathbb{R}$ is continuous at x_0. In particular, the composition of continuous functions is again continuous.

Proof. Let $\varepsilon > 0$. Then, due to the continuity of g, there exists $\eta > 0$ at $f(x_0)$, so that for all $y \in D_2$

$$|y - f(x_0)| < \eta \implies |g(y) - g(f(x_0))| < \varepsilon.$$

Since $\eta > 0$ and f is continuous at x_0, there exists $\delta > 0$ such that for all $x \in D_1$.

$$|x - x_0| < \delta \implies |f(x) - f(x_0)| < \eta.$$

12 Combining these two implications together, it follows that for all $x \in D_1$.

$$|x - x_0| < \delta \implies |f(x) - f(x_0)| < \eta \implies |g(f(x)) - g(f(x_0))| < \varepsilon$$

holds. This shows that $g \circ f$ is continuous at the point x_0. □

Remark 3.21. — Applying Proposition 3.20 with $g(x) = |x|$ (see Example 3.13), we deduce that if $f : D \to \mathbb{R}$ is continuous then also the function $x \mapsto |f(x)|$ is continuous.

Exercise 3.22. — Show that the function $f : \mathbb{R}^\times \to \mathbb{R}$ given by $x \mapsto \frac{1}{x}$ is continuous. Deduce that if $g : D \to \mathbb{R}$ is a continuous function that has no zeros inside D, then the function $x \mapsto \frac{1}{g(x)}$ is continuous in D.
Conclude that functions of the form $x \mapsto \frac{h(x)}{g(x)}$ are continuous in D whenever $h : D \to \mathbb{R}$ and $g : D \to \mathbb{R}$ are continuous and g does not have zeros in D.

Exercise 3.23. — Let $a < b < c$ be real numbers, and $f_1 : [a, b] \to \mathbb{R}$ and $f_2 : [b, c] \to \mathbb{R}$ be continuous functions. Show that the function $f : [a, c] \to \mathbb{R}$ given by

$$f(x) = \begin{cases} f_1(x) & \text{if } x \in [a, b) \\ f_2(x) & \text{if } x \in [b, c] \end{cases}$$

is continuous exactly when $f_1(b) = f_2(b)$ holds.

Exercise 3.24. — Let $I \subset \mathbb{R}$ be an open interval and let $f : I \to \mathbb{R}$ a function. Show that f is continuous if and only if, for every open set $U \subset \mathbb{R}$, $f^{-1}(U)$ is also open.

Exercise 3.25. — Let $D \subseteq \mathbb{R}$ be a subset and $f : D \to \mathbb{R}$ be a function. Show the following statements.

1. If f is continuous at $x_0 \in D$, then there exists an open neighborhood U of x_0 and a real number $M > 0$ such that $|f(x)| \le M$ for all $x \in D \cap U$.
2. If f is continuous at $x_0 \in D$ and $f(x_0) \ne 0$, then there is an open neighborhood U of x_0 such that $f(x)f(x_0) > 0$ for all $x \in D \cap U$, that is, $f(x)$ and $f(x_0)$ have the same sign.

3.1.3 Sequential Continuity

Continuity can also be characterized using sequences, as shown in the following theorem. Roughly speaking, the content of this theorem is that a function $f : D \to \mathbb{R}$ is continuous if and only if it maps convergent sequences to convergent sequences, with the correct limit. This concept is also called **sequential continuity**.

12 Theorem 3.26: Continuity = Sequential Continuity

Let $D \subseteq \mathbb{R}$ be a subset, $f : D \to \mathbb{R}$ be a function, and $\bar{x} \in D$. The function f is continuous at $\bar{x}$ if and only if, for every sequence $(x_n)_{n=0}^\infty \subset D$ converging to $\bar{x}$, the sequence $(f(x_n))_{n=0}^\infty$ converges to $f(\bar{x})$.

Proof. Suppose f is continuous at $\bar{x}$ and that $(x_n)_{n=0}^\infty$ is a convergent sequence in D with $\lim_{n\to\infty} x_n = \bar{x}$. Let $\varepsilon > 0$. Due to the continuity of f at the point $\bar{x}$, there exists $\delta > 0$ with

$$|x - \bar{x}| < \delta \implies |f(x) - f(\bar{x})| < \varepsilon$$

for all $x \in D$. Due to the convergence of $(x_n)_{n=0}^\infty$ to $\bar{x}$, there exists $N \in \mathbb{N}$ such that

$$n \ge N \implies |x_n - \bar{x}| < \delta.$$

Thus

$$n \ge N \implies |f(x_n) - f(\bar{x})| < \varepsilon,$$

which implies that the sequence $(f(x_n))_{n=0}^\infty$ converges to $f(x_0)$.

For the converse, we assume that f is not continuous at $\bar{x}$. Then there exists $\varepsilon > 0$ such that, for all $\delta > 0$, one can find $x \in D$ with

$$|x - \bar{x}| < \delta \qquad \text{and} \qquad |f(x) - f(\bar{x})| \ge \varepsilon.$$

Using this for $n \in \mathbb{N}$ and $\delta = 2^{-n} > 0$, for each $n \in \mathbb{N}$ we find $x_n \in D$ with

$$|x_n - \bar{x}| < 2^{-n} \qquad \text{and} \qquad |f(x_n) - f(\bar{x})| \geq \varepsilon.$$

From these inequalities we conclude that the sequence $(x_n)_{n=0}^{\infty}$ converges to $\bar{x}$, but $(f(x_n))_{n=0}^{\infty}$ does not converge to $f(\bar{x})$. □

12 Remark 3.27. — As the proof above shows, if a function $f : D \to \mathbb{R}$ is not continuous at $\bar{x}$, then one can find $\varepsilon > 0$ and sequence of points $(x_n)_{n=0}^{\infty} \subset D$ converging to $\bar{x}$ such that $|f(x_n) - f(\bar{x})| \geq \varepsilon$.

Exercise 3.28. — Let $D \subseteq \mathbb{R}$ be a subset and $f : D \to \mathbb{R}$ be a continuous function. Suppose that $(x_n)_{n=0}^{\infty}$ is a sequence in D such that $(f(x_n))_{n=0}^{\infty}$ converges. Must $(x_n)_{n=0}^{\infty}$ also converge?

3.2 Continuous Functions

3.2.1 The Intermediate Value Theorem

In this section we prove a fundamental theorem that formalises the idea that the graph of a continuous function on an interval is a continuous curve and so does not make any jumps. Here we show that a continuous function f on an interval $[a, b]$ contained in the domain of definition takes all values between $f(a)$ and $f(b)$. As we shall see, the proof uses the existence of the supremum, and thus indirectly the completeness axiom.

Theorem 3.29: Intermediate Value Theorem

Let $f : [a, b] \to \mathbb{R}$ be a continuous function with $f(a) \leq f(b)$. Then, for every real number c with $f(a) \leq c \leq f(b)$ there exists $\bar{x} \in [a, b]$ such that $f(\bar{x}) = c$.

Proof. Fix $c \in [f(a), f(b)]$, and define the set

$$X = \{x \in [a, b] \mid f(x) \leq c\}\,.$$

Since $a \in X$ and $X \subseteq [a, b]$, the set X is non-empty and bounded from above. Therefore, by Theorem 2.57, the supremum $\bar{x} = \sup(X) \in [a, b]$ exists. We will now use the continuity of f
13 at $\bar{x}$ to show that $f(\bar{x}) = c$ holds.

First of all, since $\bar{x}$ is the supremum of X, for any $n \geq 0$ we can find an element $x_n \in X \cap [\bar{x} - 2^{-n}, \bar{x}]$. Since $x_n \in X$ it follows that $f(x_n) \leq c$. Then, since the sequence $(x_n)_{n=0}^{\infty}$ converges to $\bar{x}$ (because $|x_n - \bar{x}| \leq 2^{-n}$) and f is continuous, Theorem 3.26 implies that

$$f(\bar{x}) = \lim_{n\to\infty} f(x_n),$$

while Proposition 2.97 implies that

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

Hence, $f(\bar{x}) \leq c$.

Suppose now by contradiction that $f(\bar{x}) < c$. Since $c \leq f(b)$, it follows that $\bar{x} < b$. Then, by continuity, given $\varepsilon = c - f(\bar{x}) > 0$ there exists $\delta > 0$ such that, for $x \in [a, b]$,

$$|x - \bar{x}| < \delta \implies |f(x) - f(\bar{x})| < \varepsilon \implies f(x) < f(\bar{x}) + \varepsilon = c.$$

This implies in particular that $(\bar{x} - \delta, \bar{x} + \delta) \cap [a, b] \subset X$. In particular X contains the set $(\bar{x}, \bar{x} + \delta) \cap (\bar{x}, b]$, a contradiction to the fact that $\bar{x} = \sup(X)$. In conclusion $f(\bar{x}) = c$, as desired. □

REMARK 3.30. — If $f : [a, b] \to \mathbb{R}$ is a continuous function such that $f(a) > f(b)$, Theorem 3.29 still holds in the following way:
For every real number c with $f(a) \geq c \geq f(b)$ there exists $\bar{x} \in [a, b]$ such that $f(\bar{x}) = c$.
To prove this statement, there are two possible ways:
(1) repeat the proof of Theorem 3.29 defining $X = \{x \in [a, b] \mid f(x) \geq c\}$;
(2) apply Theorem 3.29 to the function $g = -f$.

EXERCISE 3.31. — Let I be a non-empty interval and $f : I \to \mathbb{R}$ a continuous injective function. Show that f is strictly monotone.

3.2.2 Inverse Function Theorem

INTERLUDE: IDENTITY AND INVERSE FUNCTION

Given a set X, the **identity function** $\mathrm{id}_X : X \to X$ is defined as

$$\mathrm{id}_X(x) = x \qquad \text{for every } x \in X.$$

Given a bijective function $f : X \to Y$, one can uniquely construct a function $g : Y \to X$
13 as follows: for every $y \in Y$, $g(y)$ is the uniquely determined element $x \in X$ satisfying $f(x) = y$. With this definition, it follows

$$g \circ f = \mathrm{id}_X \qquad \text{and} \qquad f \circ g = \mathrm{id}_Y .$$

The function g is called **inverse function** (or **inverse mapping**) of f, and is often denoted by f^{-1}.

REMARK 3.32. — The existence of an inverse function is characteristic of bijective functions: A function $f : X \to Y$ is bijective if and only if there exists a function $g : Y \to X$ with $g \circ f = \mathrm{id}_X$ and $f \circ g = \mathrm{id}_Y$.

3.33. — In this subsection, we show that every continuous strictly monotone function has an inverse function that is also continuous.

THEOREM 3.34: INVERSE FUNCTION THEOREM

Let I be an interval and $f : I \to \mathbb{R}$ be a continuous strictly monotone function. Then $f(I) \subset \mathbb{R}$ is an interval, and the mapping $f : I \to f(I)$ has a continuous strictly monotone inverse function $f^{-1} : f(I) \to I$.

Proof. Without loss of generality, we can assume that I is non-empty and not a single point. Also, we can assume that f is strictly increasing (otherwise replace f with $-f$).

We write $J = f(I)$, and first notice that the function $f : I \to J$ is bijective, since it is surjective by definition, and due to strict monotonicity it is also injective. Thus, there exists a uniquely determined inverse $g = f^{-1} : J \to I$.

We note that the function g is strictly increasing: since f is strictly increasing,

$$x_1 < x_2 \iff f(x_1) < f(x_2) \qquad \forall\, x_1, x_2 \in I,$$

which leads to

$$g(y_1) < g(y_2) \iff y_1 < y_2 \qquad \forall\, y_1, y_2 \in J.$$

We show that J is an interval. Indeed, let $y_1 = f(x_1)$ and $y_2 = f(x_2)$ be elements of J, with $y_1 < y_2$. Then $x_1 < x_2$, and it follows from the Intermediate Value Theorem 3.29 applies to the closed interval $[x_1, x_2]$ that $[y_1, y_2] \subset f([x_1, x_2]) \subset J$. In other words, given any two element in J, the closed interval between them is contained inside J. Thus, J is an interval.

It remains to show that $g = f^{-1}$ is continuous. Let $\bar{y} \in J$, and assume by contradiction that g is not continuous at $\bar{y}$. Because of Remark 3.27, we can find $\varepsilon > 0$ and a sequence of points $(y_n)_{n=0}^{\infty} \subset J$ such that

$$\lim_{n\to\infty} y_n = \bar{y} \qquad \text{but} \qquad |g(y_n) - g(\bar{y})| \geq \varepsilon.$$

13 Define $x_n = g(y_n)$ and $\bar{x} = g(\bar{y})$. The property above tells us that, for every $n \in \mathbb{N}$,

$$\text{either } x_n \leq \bar{x} - \varepsilon, \qquad \text{or } x_n \geq \bar{x} + \varepsilon.$$

In particular, there are infinitely many indices n's for which one of the above options holds. Without loss of generality, let us assume that there are infinitely many indices n's for which $x_n \leq \bar{x} - \varepsilon$, and define a subsequence x_{n_k} using such indices, so that

$$x_{n_k} = g(y_{n_k}) \leq \bar{x} - \varepsilon \qquad \text{for all } k \in \mathbb{N}.$$

Since f is strictly increasing and $\bar{x} - \varepsilon \in I$ (since both $x_{n_k} = g(y_{n_k})$ and $\bar{x} = g(\bar{y})$ belong to I, and I is an interval), we deduce that

$$y_{n_k} = f(x_{n_k}) \leq f(\bar{x} - \varepsilon) < f(\bar{x}) = \bar{y}.$$

Hence, using Proposition 2.97 we conclude that

$$\bar{y} = \lim_{k\to\infty} y_{n_k} \leq f(\bar{x} - \varepsilon) < f(\bar{x}) = \bar{y},$$

a contradiction. □

Example 3.35. — Let $n \in \mathbb{N}$, $n \geq 1$. The function $f : [0, \infty) \to [0, \infty)$ given by $x \mapsto x^n$

is continuous, strictly increasing, and surjective. According to the Inverse Function Theorem, there exists a continuous strictly increasing inverse $[0,\infty) \to [0,\infty)$, that we express as

$$x \mapsto \sqrt[n]{x}$$

and is called **n-th root**. Furthermore, for $m, n \in \mathbb{N}$ with $n \geq 1$, we can define

$$x^{\frac{m}{n}} = \underbrace{\sqrt[n]{x} \cdot \ldots \cdot \sqrt[n]{x}}_{m \text{ times}} \qquad \text{for } x \in [0,\infty).$$

Also, thanks to Exercise 3.22, we can define the continuous functions

13

$$x^{-\frac{m}{n}} = \frac{1}{x^{\frac{m}{n}}} \qquad \text{for } x \in (0,\infty).$$

EXERCISE 3.36. — For any real number $a > 0$, we define the sequence of real numbers $(x_n)_{n=0}^{\infty}$ by $x_n = \sqrt[n]{a}$. Show that the sequence $(x_n)_{n=0}^{\infty}$ converges, and that

$$\lim_{n\to\infty} \sqrt[n]{a} = 1.$$

EXERCISE 3.37. — We define a sequence of real numbers $(x_n)_{n=0}^{\infty}$ by $x_n = \sqrt[n]{n}$. Show that this sequence converges, with limiting value

$$\lim_{n\to\infty} \sqrt[n]{n} = 1.$$

3.3 Continuous Functions on Compact Intervals

In this section we show that continuous functions on **bounded closed** intervals – so-called **compact** intervals – have special properties.

3.3.1 Boundedness and Extrema

The key property of compact intervals, that will be used several times, is the following:

Lemma 3.38: Compactness

Let $[a, b]$ be a compact interval, and $(x_n)_{n=0}^{\infty}$ a sequence contained in $[a, b]$. Then there exists a subsequence $(x_{n_k})_{k=0}^{\infty}$ such that $\lim_{k\to\infty} x_{n_k} = \bar{x}$ with $\bar{x} \in [a, b]$.

Proof. Since the sequence $(x_n)_{n=0}^{\infty}$ is bounded (being contained inside $[a, b]$), Corollary 2.117 implies the existence of a subsequence $(x_{n_k})_{k=0}^{\infty}$ such that $\lim_{k\to\infty} x_{n_k} = \bar{x}$. Then, since $a \le x_{n_k} \le b$ for every $k \in \mathbb{N}$, Proposition 2.97 implies that $a \le \bar{x} \le b$. □

Theorem 3.39: Boundedness

Let $[a, b]$ be a compact interval, and $f : [a, b] \to \mathbb{R}$ a continuous function. Then f is bounded.

14 *Proof.* Assume by contradiction that f is not bounded. Then, for every $n \in \mathbb{N}$ there exists a point $x_n \in [a, b]$ such that $|f(x_n)| \ge n$. Applying Lemma 3.38, we can find a subsequence $(x_{n_k})_{k=0}^{\infty}$ such that $\lim_{k\to\infty} x_{n_k} = \bar{x} \in [a, b]$.

Hence, by the continuity of $|f|$ (recall Remark 3.21) we deduce that the sequence $\{|f(x_{n_k})|\}_{k=0}^{\infty}$ converges to the real number $|f(\bar{x})|$, a contradiction since

$$|f(x_{n_k})| \ge n_k \to \infty \qquad \text{as } k \to \infty.$$

□

Exercise 3.40. — Find examples of ...

1. ...an unbounded continuous function on a bounded open interval.
2. ...an unbounded continuous function on an unbounded closed interval.
3. ...an unbounded function on a compact interval that is discontinuous at only one point.

Definition 3.41: Extreme Values

Let $D \subset \mathbb{R}$ and $f : D \to \mathbb{R}$ be a real-valued function on D. We say that the function f takes its **maximal value** at a point $x_0 \in D$ if $f(x) \le f(x_0)$ holds for all $x \in D$. We call $f(x_0)$ the **maximum** of f. Similarly, f takes its **minimal value** at $x_0 \in D$ if $f(x) \ge f(x_0)$ holds for all $x \in D$. In that case we call $f(x_0)$ the **minimum** of f. Maxima and minima are summarily called **extreme values** or **extrema**.

Theorem 3.42: Extrema

Let $[a, b]$ be a compact interval, and $f : [a, b] \to \mathbb{R}$ a continuous function. Then f attains both its maximum and minimum.

Proof. Since f is bounded (thanks to Theorem 3.39), Theorem 2.57 implies that the supremum $S = \sup(f([a, b]))$ exists. By definition of supremum, for any $n \in \mathbb{N}$ there exists $y_n = f(x_n)$ with $x_n \in [a, b]$, such that $S - 2^{-n} \le y_n \le S$. Hence, $\lim_{n\to\infty} y_n = S$.

Thanks to Lemma 3.38, there exists a subsequence $(x_{n_k})_{k=0}^{\infty}$ such that $\lim_{k\to\infty} x_{n_k} = \bar{x} \in [a, b]$. Hence, by the continuity of f and Theorem 3.26, this implies that

$$f(\bar{x}) = \lim_{k\to\infty} f(x_{n_k}) = \lim_{k\to\infty} y_{n_k} = S,$$

14 where the convergence of $(y_{n_k})_{k=0}^{\infty}$ to S follows from Lemma 2.87. Thus, f attains its maximum at $\bar{x} \in [a, b]$.

Applying the same argument to $-f$, it follows that f attains also its minimum. □

Exercise 3.43. — Does any continuous function f on the open interval $(0, 1)$ attain its maximum?

3.3.2 Uniform Continuity

Definition 3.44: Uniform Continuity

Let $D \subseteq \mathbb{R}$. A function $f : D \to \mathbb{R}$ is called **uniformly continuous** if for all $\varepsilon > 0$ there exists $\delta > 0$ such that the implication

$$|x - y| < \delta \implies |f(x) - f(y)| < \varepsilon$$

holds for all $x, y \in D$.

Exercise 3.45. — Show that the polynomial function $f(x) = x^2$ is continuous on $\mathbb{R}$ but not uniformly continuous. Also, the restriction of f to $[0, 1]$ is uniformly continuous.

Theorem 3.46: Compact Intervals and Uniform Continuity

Let $[a,b]$ be a compact interval, and $f:[a,b]\to\mathbb{R}$ a continuous function. Then f is uniformly continuous.

Proof. Assume by contradiction that f is not uniformly continuous. This means that there exists $\varepsilon>0$ such that, for all $\delta>0$, one can find $x,y\in[a,b]$ with

$$|x-y|<\delta \qquad \text{and} \qquad |f(x)-f(y)|\geq\varepsilon.$$

Using this with $\delta=2^{-n}>0$, for each $n\in\mathbb{N}$ we can find $x_n,y_n\in[a,b]$ satisfying

$$|x_n-y_n|<2^{-n} \qquad \text{and} \qquad |f(x_n)-f(y_n)|\geq\varepsilon. \tag{3.1}$$

Thanks to Lemma 3.38, there exists a subsequence $(x_{n_k})_{k=0}^{\infty}$ such that $\lim_{k\to\infty}x_{n_k}=\bar{x}\in[a,b]$. In addition, since $|x_{n_k}-y_{n_k}|<2^{-n_k}$ we have

$$|y_{n_k}-\bar{x}|\leq|y_{n_k}-x_{n_k}|+|x_{n_k}-\bar{x}|<2^{-n_k}+|x_{n_k}-\bar{x}|.$$

Since the right-hand side tends to 0 as $k\to\infty$, also the subsequence $(y_{n_k})_{k=0}^{\infty}$ converges to $\bar{x}$.

Now, since f is continuous at $\bar{x}$ and $\lim_{k\to\infty}x_{n_k}=\lim_{k\to\infty}y_{n_k}=\bar{x}$, it follows from
14 Theorem 3.26 that

$$\lim_{k\to\infty}f(x_{n_k})=\lim_{k\to\infty}f(y_{n_k})=f(\bar{x}).$$

This means that there exist $N_1,N_2\in\mathbb{N}$ such that

$$k\geq N_1 \implies |f(x_{n_k})-f(\bar{x})|<\frac{\varepsilon}{2}, \qquad \text{and} \qquad k\geq N_2 \implies |f(y_{n_k})-f(\bar{x})|<\frac{\varepsilon}{2}.$$

This implies that, for $k\geq N=\max\{N_1,N_2\}$ we have

$$|f(x_{n_k})-f(y_{n_k})|\leq|f(x_{n_k})-f(\bar{x})|+|f(y_{n_k})-f(\bar{x})|<\frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon,$$

a contradiction to (3.1) which proves the result. □

Exercise 3.47. — Does the statement of Theorem 3.46 hold for continuous functions on the open interval $(0,1)$?

Exercise 3.48. — In this exercise we consider another continuity term.

1. Let $D\subset\mathbb{R}$ be a subset. We call a real-valued function f on D **Lipschitz continuous** if there exists $L\geq0$ such that

$$|f(x)-f(y)|\leq L|x-y| \qquad \text{for all } x,y\in D.$$

Give examples of Lipschitz continuous functions, and show that a Lipschitz continuous function is also uniformly continuous.

14 2. Let $f : \mathbb{R}_{\geq 0} \to \mathbb{R}$ be the root function, $f(x) = \sqrt{x}$. Show that:
(i) $f_{[0,1]} : [0, 1] \to \mathbb{R}$ is not Lipschitz continuous;
(ii) $f_{[1,\infty)} : [1, \infty) \to \mathbb{R}$ is Lipschitz continuous;
(iii) $f : [0, \infty) \to \mathbb{R}$ is uniformly continuous.

3.4 Example: Exponential and Logarithmic Functions

In this section we will use the notion of convergence and limits of sequences to define the exponential function, and to prove some of its properties such as continuity.

3.4.1 Definition of the Exponential Function

We start with a preliminary lemma, that we will use repeatedly in this section.

Lemma 3.49: Bernoulli's inequality

For all $a \in \mathbb{R}$ with $a \geq -1$ and $n \in \mathbb{N}$ with $n \geq 1$, $(1+a)^n \geq 1 + na$ holds.

Proof. We use induction. For $n = 1$ we have $(1+a)^n = 1 + a = 1 + na$.

Now, suppose that the inequality $(1+a)^n \geq 1 + na$ holds for some $n \geq 1$. Since $1 + a \geq 0$ by assumption, it follows that

$$(1+a)^{n+1} = (1+a)^n(1+a) \geq (1+na)(1+a) = 1 + na + a + na^2 \geq 1 + (n+1)a,$$

yielding the induction step. The lemma follows. □

Proposition 3.50: Existence of the Exponential

15 *Let $x \in \mathbb{R}$ be a real number. The sequence of real numbers $(a_n)_{n=1}^{\infty}$ given by*

$$a_n = \left(1 + \frac{x}{n}\right)^n$$

is convergent, and its limit is a positive real number.

To prove the result, we first show the following:

Lemma 3.51: Monotonicity

Given $x \in \mathbb{R}$, let $n_0 \in \mathbb{N}$ be an integer satisfying $n_0 \geq 1$ and $n_0 > -x$. Then the sequence $(a_n)_{n=n_0}^{\infty}$ defined in Proposition 3.50 is increasing.

Proof. First, we note that for $n \geq n_0$ (so that $x + n > 0$) it holds

$$\frac{x}{(n+1)(n+x)} \leq \frac{x+n}{(n+1)(n+x)} = \frac{1}{n+1} \leq 1,$$

that is

$$-\frac{x}{(n+1)(n+x)} \geq -1 \qquad \forall n \geq n_0.$$

Hence, for $n \geq n_0$ we can use Bernoulli inequality in Lemma 3.49 to obtain

$$\frac{\left(1+\frac{x}{n+1}\right)^{n+1}}{\left(1+\frac{x}{n}\right)^{n}} = \left(1+\frac{x}{n}\right)\left(\frac{1+\frac{x}{n+1}}{1+\frac{x}{n}}\right)^{n+1} = \frac{n+x}{n}\left(\frac{n(n+1+x)}{(n+1)(n+x)}\right)^{n+1}$$
$$= \frac{n+x}{n}\left(1-\frac{x}{(n+1)(n+x)}\right)^{n+1} \geq \frac{n+x}{n}\left(1-\frac{x}{n+x}\right) = 1.$$

Thus $a_n \leq a_{n+1}$, as desired. □

Proof of Proposition 3.50. Fix $x \in \mathbb{R}$, and let $n_0 \in \mathbb{N}$ be an integer satisfying $n_0 \geq 1$ and $n_0 > -x$. As shown in Lemma 3.51, the sequence $(a_n)_{n=n_0}^{\infty}$ is increasing. So, if we prove that it is bounded, Theorem 2.108 will show that the sequence $(a_n)_{n=n_0}^{\infty}$ (and hence the sequence $(a_n)_{n=0}^{\infty}$) converges.

To show boundedness, we first consider the case $x \leq 0$. In this case we note that

$$0 < 1 + \frac{x}{n} \leq 1 \qquad \forall\, n \geq n_0,$$

thus $0 < \left(1+\frac{x}{n}\right)^n \leq 1$ holds. So 1 is an upper bound for the increasing positive sequence $(a_n)_{n=n_0}^{\infty}$. Therefore

$$\lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n = \sup\left\{\left(1+\tfrac{x}{n}\right)^n \mid n \geq n_0\right\} > 0.$$

15 This proves the proposition for $x \leq 0$.

For $x \geq 0$ and $n > x$ we use that

$$\left(1+\frac{x}{n}\right)^n\left(1-\frac{x}{n}\right)^n = \left(1-\frac{x^2}{n^2}\right)^n \leq 1,$$

which implies the estimate

$$1 \leq \left(1+\frac{x}{n}\right)^n \leq \left(1-\frac{x}{n}\right)^{-n} = \frac{1}{\left(1+\frac{(-x)}{n}\right)^n} \qquad \forall\, n > x.$$

Since the sequence $\left(\left(1+\frac{(-x)}{n}\right)^n\right)_{n=1}^{\infty}$ converges to a positive number (by the case $x \leq 0$ above), Proposition 2.96(3) implies that also the sequence $\left(\left(1+\frac{(-x)}{n}\right)^{-n}\right)_{n=1}^{\infty}$ converges, and in particular it is bounded (see Lemma 2.102). This implies that the monotonically increasing sequence $\left(\left(1+\frac{x}{n}\right)^n\right)_{n=1}^{\infty}$ is also bounded, so it converges. □

Definition 3.52: Exponential Function

The **exponential function** $\exp : \mathbb{R} \to \mathbb{R}_{>0}$ is the function defined by

$$\exp(x) = \lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n$$

for all $x \in \mathbb{R}$.

Euler's number $e \in \mathbb{R}$ is defined as

$$e = \exp(1) = \lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n. \tag{3.2}$$

Its numerical value is

$$e = 2.71828182845904523536028747135266249775724709369995\ldots$$

A useful consequence of Lemma 3.51 is the following lower bound on $\exp(x)$.

Corollary 3.53: Growth of Exponential

Given $n \in \mathbb{N}$ with $n \geq 1$, the exponential function satisfies

$$\exp(x) \geq \left(1 + \frac{x}{n}\right)^n \qquad \forall\, x > -n.$$

Proof. It suffices to observe that, given $x \in \mathbb{R}$ with $x > -n$, Lemma 3.51 and Definition 3.52 imply that $a_n \leq a_{n+1} \leq \ldots \leq \exp(x)$. □

3.4.2 Properties of the Exponential Function

15 **Theorem 3.54: Properties of Exponential Map**

The exponential function $\exp : \mathbb{R} \to \mathbb{R}_{>0}$ *is bijective, strictly monotonically increasing, and continuous. Furthermore*

$$\exp(0) = 1, \tag{3.3}$$

$$\exp(-x) = \exp(x)^{-1} \qquad \text{for all } x \in \mathbb{R}, \tag{3.4}$$

$$\exp(x+y) = \exp(x)\exp(y) \qquad \text{for all } x, y \in \mathbb{R}, \tag{3.5}$$

Proof. (Extra material)

We start by checking (3.3), (3.4) and (3.5).

1. The identity (3.3) follows directly from the definition of the exponential function.
2. For (3.4), we apply Proposition 2.96(2) to say that

$$\exp(x)\exp(-x) = \lim_{n\to\infty}\left(1 + \frac{x}{n}\right)^n \lim_{n\to\infty}\left(1 - \frac{x}{n}\right)^n = \lim_{n\to\infty}\left(1 - \frac{x^2}{n^2}\right)^n.$$

   Now, for $n \geq |x|$ we have $-\frac{x^2}{n^2} \geq -1$. Thus, using Bernoulli's inequality, we obtain

$$1 - \tfrac{x^2}{n} \leq \left(1 - \tfrac{x^2}{n^2}\right)^n \leq 1 \qquad \forall\, n \geq |x|$$

Since the left-hand side converges to 1 as $n \to \infty$, the Sandwich Lemma 2.99 implies that $\lim_{n\to\infty} \left(1 - \frac{x^2}{n^2}\right)^n = 1$, so (3.4) holds.

3. We prove (3.5) by an argument similar to the one above. As a preparation, we first calculate for $n \geq 1$ the product

$$\left(1 - \frac{x}{n}\right)\left(1 - \frac{y}{n}\right)\left(1 + \frac{x+y}{n}\right) = 1 - \frac{(x+y)^2}{n^2} + \frac{xy}{n^2}\left(1 + \frac{x+y}{n}\right) = 1 + \frac{c_n}{n^2},$$

where c_n stands for

$$c_n = -(x^2 + y^2) - xy + xy\frac{x+y}{n}.$$

Note that, since $\left|\frac{x+y}{n}\right| \leq \frac{|x|+|y|}{n} \leq 1$ for $n \geq |x| + |y|$, we have

$$-2|xy| \leq -xy + xy\frac{x+y}{n} \leq 2|xy| \qquad \forall\, n \geq |x| + |y|.$$

Also, since $(|x| - |y|)^2 \geq 0$, developing the square we get

$$2|xy| \leq x^2 + y^2.$$

Hence, for $n \geq |x| + |y|$ we obtain

15 $$0 \geq -(x^2 + y^2) + 2|xy| \geq c_n \geq -(x^2 + y^2) - 2|xy| \geq -2(x^2 + y^2).$$

In particular, $\frac{c_n}{n^2} \geq -1$ for $n \geq |x|+|y|$. Therefore, by the Bernoulli inequality, it follows

$$1 - 2\frac{x^2 + y^2}{n} \leq 1 + \frac{c_n}{n} \leq \left(1 + \frac{c_n}{n^2}\right)^n \leq 1 \qquad \forall\, n \geq |x| + |y|,$$

so $\lim_{n\to\infty} \left(1 + \frac{c_n}{n^2}\right)^n = 1$ by the Sandwich Lemma 2.99. Using (3.4) and Proposition 2.96(2), we get

$$\frac{\exp(x+y)}{\exp(x)\exp(y)} = \lim_{n\to\infty} \left(1 - \frac{x}{n}\right)^n \left(1 - \frac{y}{n}\right)^n \left(1 + \frac{x+y}{n}\right)^n = \lim_{n\to\infty} \left(1 + \frac{c_n}{n^2}\right)^n = 1$$

which proves (3.5).

It remains to prove the continuity, monotonicity, and bijectivity of the exponential function. We shall first prove some useful estimates. First of all, we claim that

$$\exp(x) \geq 1 + x \qquad \forall\, x \in \mathbb{R} \tag{3.6}$$

Indeed, for $x \leq -1$ this is clear since $\exp(x) > 0$. For $x > -1$, it follows from Corollary 3.53.

Combining (3.6) and (3.4), we also deduce that

$$\exp(x) = \frac{1}{\exp(-x)} \leq \frac{1}{1-x} \qquad \forall\, x < 1. \tag{3.7}$$

Furthermore, as a consequence of (3.6), we can also show that

$$\exp\left(-\frac{1}{x}\right) < x < \exp(x) \qquad \forall\, x \in \mathbb{R}. \tag{3.8}$$

Indeed, (3.6) implies that $x < 1 + x \leq \exp(x)$. Applying this inequality to $\frac{1}{x}$ we obtain

$$\frac{1}{x} < \exp\left(\frac{1}{x}\right) \implies x > \frac{1}{\exp\left(\frac{1}{x}\right)} = \exp\left(-\frac{1}{x}\right).$$

We can now prove the desired properties of exp.

1. *The function* exp *is continuous.* We first prove the continuity of exp at 0.

   For this, given $x \in (-\delta, \delta)$ with $\delta \in (0, 1)$, we apply (3.7) to deduce that

   $$x \in [0, \delta) \implies |\exp(x) - 1| = \exp(x) - 1 \leq \frac{1}{1-x} - 1 \leq \frac{1}{1-\delta} - 1 = \frac{\delta}{1-\delta},$$

   and (3.6) to prove that

   $$x \in (-\delta, 0] \implies |\exp(x) - 1| = 1 - \exp(x) \leq 1 - (1 + x) = -x \leq \delta \leq \frac{\delta}{1-\delta}.$$

   15 Thus, for $x \in (-\delta, \delta)$ with $\delta \in (0, 1)$ we have

   $$|\exp(x) - \exp(0)| \leq \frac{\delta}{1-\delta}.$$

   Recalling Remark 3.10, to prove the continuity of exp at 0 it suffices to consider $\varepsilon \in (0, 1]$. So, let $\varepsilon \in (0, 1]$ and choose $\delta = \frac{\varepsilon}{2+\varepsilon}$. With this choice we see that $\delta < 1$ and that

   $$|\exp(x) - \exp(0)| \leq \frac{\delta}{1-\delta} = \frac{\varepsilon}{2} < \varepsilon.$$

   It follows that exp is continuous at 0.

   To show continuity at an arbitrary point $x_0 \in \mathbb{R}$, we write

   $$\exp(x) = \exp(x - x_0 + x_0) = \exp(x - x_0)\exp(x_0).$$

   We can thus write the exponential function as a composition, namely $\exp = \mu \circ \exp \circ \tau$, with $\tau : \mathbb{R} \to \mathbb{R}$ and $\mu : \mathbb{R} \to \mathbb{R}$ given by

   $$\tau(x) = x - x_0 \qquad \text{and} \qquad \mu(x) = \exp(x_0)x$$

   Note that the functions τ and μ are continuous. In particular, τ is continuous at x_0, and exp is continuous at $0 = \tau(x_0)$. It follows that exp is continuous at x_0 from Proposition 3.20.

2. *The function* exp *is strictly increasing.* For all real numbers $x < y$ it follows from (3.6) that $\exp(y-x) \geq 1 + (y-x) > 1$, therefore

$$\exp(x) < \exp(x)\exp(y-x) = \exp(y).$$

   Thus $\exp : \mathbb{R} \to \mathbb{R}_{>0}$ strictly increasing.

3. *The function* $\exp : \mathbb{R} \to \mathbb{R}_{>0}$ *is bijective.* The exponential function is strictly increasing and, therefore, injective. To show surjectivity, choose $a \in \mathbb{R}_{>0}$ arbitrarily. If we set $x_0 = -a^{-1}$ and $x_1 = a$, it follows from (3.8) that

$$\exp(x_0) < a < \exp(x_1).$$

   Since exp is continuous on all $\mathbb{R}$, it follows from the Intermediate Value Theorem 3.29, that there exists $x \in [x_0, x_1]$ such that $\exp(x) = a$. This shows the assertion and finishes the proof. □

15

3.4.3 The Natural Logarithm

Definition 3.55: Logarithm

The unique inverse function

$$\log : \mathbb{R}_{>0} \to \mathbb{R}$$

of the bijective mapping $\exp : \mathbb{R} \to \mathbb{R}_{>0}$ is called **logarithm**.

Corollary 3.56: Properties of Logarithm

The logarithm $\log : \mathbb{R}_{>0} \to \mathbb{R}$ *is a strictly monotonically increasing, continuous and bijective function. Furthermore*

$$\log(1) = 0, \tag{3.9}$$

$$\log(a^{-1}) = -\log(a) \qquad \textit{for all } a \in \mathbb{R}_{>0}, \tag{3.10}$$

$$\log(ab) = \log(a) + \log(b) \qquad \textit{for all } a, b \in \mathbb{R}_{>0}, \tag{3.11}$$

Proof. This follows directly from Theorem 3.54 and the Inverse Function Theorem 3.34. The equations (3.9), (3.10), and (3.11) follow from the corresponding properties of the exponential, choosing $x = \log a$ and $y = \log b$. □

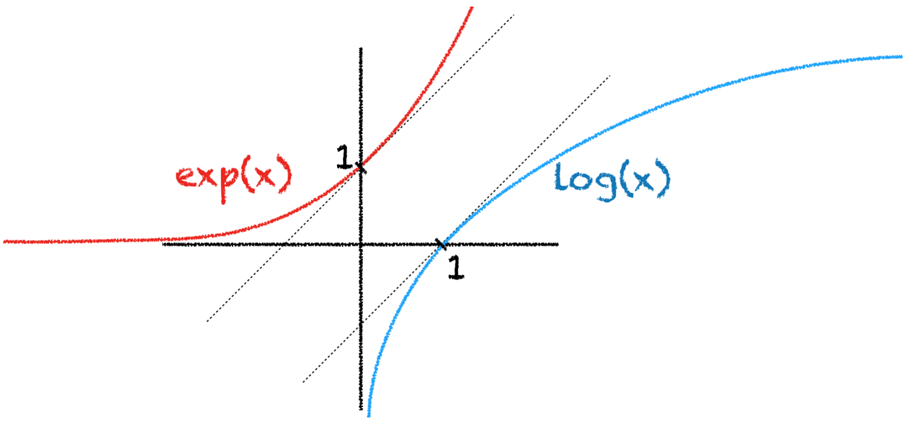


Figure 3.2: The graphs of the exponential function and the logarithm. The auxiliary lines show $\exp(x) \geq x+1$ and $\log(x) \leq x-1$.

3.57. — The logarithm function defined here is also called the **natural logarithm** to
distinguish it from the logarithm with another **base** $a > 1$, typically $a = 10$ or $a = 2$. Let
15 $a > 1$ be a real number. We can define the logarithm $\log_a : \mathbb{R}_{>0} \to \mathbb{R}$ in base a: the notation

$$\log_a(x) = \frac{\log x}{\log a} \qquad \forall\, x \in \mathbb{R}_{>0}.$$

Verify that $\log_{10}(10^n) = n$ holds for all $n \in \mathbb{Z}$. However, we will not use this definition, not even for $a = 10$ or $a = 2$, and $\log(x)$ will always denote the natural logarithm of $x \in \mathbb{R}_{>0}$ to base e.

3.58. — We can use the logarithm and exponential mapping to define more general powers. For a positive number $a > 0$ and arbitrary exponents $x \in \mathbb{R}$, we write

$$a^x = \exp(x \log(a)).$$

In particular, we write $e^x = \exp(x \log(e)) = \exp(x)$ for all $x \in \mathbb{R}$.

Also, for $x > 0$ and $a \in \mathbb{R}$, we define

$$x^a = \exp(a \log(x)).$$

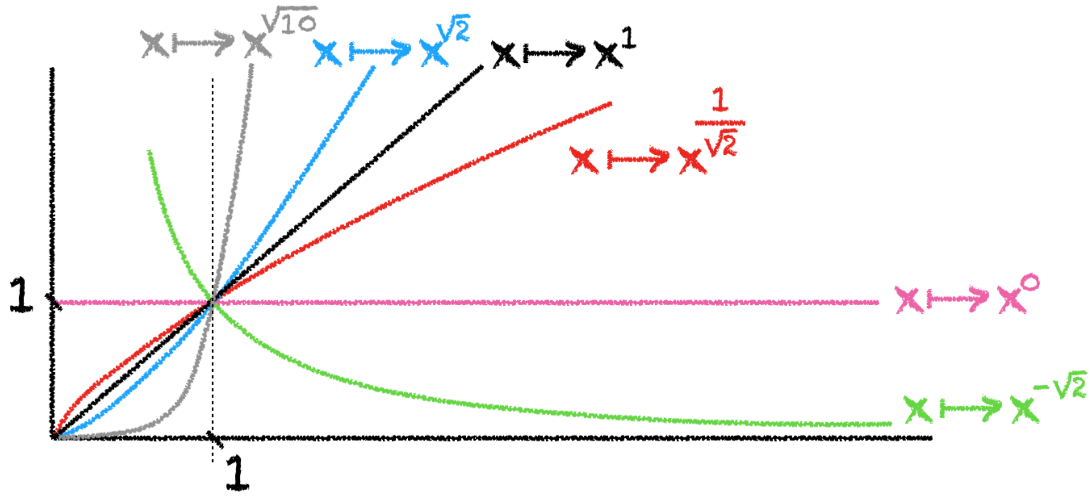


Figure 3.3: Graphs of $a \mapsto a^x$ for different exponents $x \in \mathbb{R}$.

Exercise 3.59. — Show that for $x \in \mathbb{Q}$ and $a > 0$ this definition agrees with the definition of rational powers from Example 3.35. Furthermore, check the calculation rules

$$\log(a^x) = x \log(a), \quad a^x a^y = a^{x+y}, \quad (a^x)^y = a^{xy}$$

15 for $a > 0$ and $x, y \in \mathbb{R}$.

Exercise 3.60. — Let $a > 0$ be a positive number. Show that there exists a real number $C_a > 0$ such that $\log(x) \leq C_a x^a$ holds for all $x > 0$.

Exercise 3.61. — Show that for all real numbers $x \geq -1$ and $p \geq 1$, the continuous Bernoulli inequality

$$(1 + x)^p \geq 1 + px.$$

holds.

Exercise 3.62. — In this exercise, we consider another continuity term (compare with Exercise 3.48).

1. Let $D \subset \mathbb{R}$ be a subset and $\alpha \in (0, 1]$. We call a real-valued function f on D α-**Hölder continuous** if there exists $L \geq 0$ such that

   $$|f(x) - f(y)| \leq L|x - y|^\alpha \qquad \text{for all } x, y \in D.$$

   Show that a α-Hölder continuous function is also uniformly continuous.
   (Note that α-Hölder continuous functions with $\alpha = 1$ correspond to Lipschitz functions.)

2. Given $\alpha \in (0, 1]$, consider the function $f : [0, \infty) \to \mathbb{R}$ given by $f(x) = x^\alpha$. Show that f is α-Hölder continuous.
   Hint: use the inequality

15

$$(x + y)^\alpha \leq x^\alpha + y^\alpha \qquad \forall\, x, y \geq 0,\ \alpha \in (0, 1]. \tag{3.12}$$

Applet 3.63 (Slide rule)**.** *Using the slide rule, calculate some products and quotients. Recall the properties of the logarithm to see how to do these calculations. Before the introduction of electronic calculators, these mechanical aids were widely used.*

3.5 Limits of Functions

We consider functions $f : D \to \mathbb{R}$ on a subset $D \subset \mathbb{R}$ and want to define limits of $f(x)$ for the case when $x \in D$ tends to a point $x_0 \in \mathbb{R}$. Typical situations are $D = \mathbb{R}$, $D = [0, 1]$, or $D = (0, 1)$, and $x_0 = 0$ in all cases.

3.5.1 Limit in the Vicinity of a Point

3.64. — We specify for this section a non-empty subset $D \subset \mathbb{R}$, and an element $x_0 \in \mathbb{R}$ such that

$$D \cap (x_0 - \delta, x_0 + \delta) \neq \emptyset \tag{3.13}$$

holds for all $\delta > 0$. Whenever (3.13) holds, we say that x_0 is an **accumulation point** of D. Note that when $x_0 \in D$, (3.13) is always satisfied.

We remark that (3.13) implies that there exists a sequence of points in D converging to x_0.

Definition 3.65: Limit of a Function

Let $f : D \to \mathbb{R}$, and assume that $D \cap (x_0 - \delta, x_0 + \delta) \neq \emptyset$ for every $\delta > 0$.
A real number $L \in \mathbb{R}$ is called **limit of $f(x)$ as $x \to x_0$** if for every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$x \in D \cap (x_0 - \delta, x_0 + \delta) \implies |f(x) - L| < \varepsilon.$$

15 In general, the limit of $f(x)$ as $x \to x_0$ may not exist. But if a limit exists, then it is uniquely determined. Therefore, from now on, we speak of *the* limit and write

$$\lim_{x \to x_0} f(x) = L$$

if the limit of $f(x)$ as $x \to x_0$ exists and is equal to L. Informally, this means that the function values of f are arbitrarily close to L if $x \in D$ is close to x_0.

3.66. — The limit of a function satisfies properties analogous to Proposition 2.96. If f and g are functions on D such that the limits

$$\lim_{x \to x_0} f(x) = L_1 \qquad \text{and} \qquad \lim_{x \to x_0} g(x) = L_2$$

exist, so do the limits

$$\lim_{x \to x_0} f(x) + g(x) = L_1 + L_2 \qquad \text{and} \qquad \lim_{x \to x_0} f(x)g(x) = L_1 L_2.$$

The inequality $f \leq g$ implies $L_1 \leq L_2$. Finally, the sandwich lemma holds: if h is another function on D with $f \leq h \leq g$ and $L_1 = L_2$, then $\lim_{x \to x_0} h(x) = L_1 = L_2$.

Recalling the definition of continuity, one has the following:

Remark 3.67. — Let $f : D \to \mathbb{R}$ be a function. If $x_0 \in D$, then f is continuous at x_0 if and only if $\lim_{x\to x_0} f(x) = f(x_0)$ holds.

3.68. — Suppose that $x_0 \in D$ is an accumulation point of $D \setminus \{x_0\}$. Let $f : D \to \mathbb{R}$ be a function, and consider $f|_{D\setminus\{x_0\}} : D \setminus \{x_0\} \to \mathbb{R}$ for the restriction of f to $D \setminus \{x_0\}$. It is possible that f may be discontinuous at the point x_0, but the limiting value

$$L = \lim_{x\to x_0} f|_{D\setminus\{x_0\}}(x) \tag{3.14}$$

nevertheless exists. Under these circumstances, the point $x_0 \in D$ is called a **removable discontinuity** of f, and one also writes

$$L = \lim_{\substack{x\to x_0 \\ x\neq x_0}} f(x) \tag{3.15}$$

in place of (3.14). Note that, in the situation above, the function $\widetilde{f} : D \to \mathbb{R}$ defined by

15

$$\widetilde{f}(x) = \begin{cases} f(x) & \text{if } x \in D \setminus \{x_0\} \\ L & \text{if } x = x_0 \end{cases} \tag{3.16}$$

is continuous at the point x_0. In other words, we have removed the discontinuity of the function $\widetilde{f}$ by replacing the value of the function f at the location x_0 with L.

3.69. — Suppose instead that $x_0 \notin D$ but the limit in (3.15) exists. In this situation, we call the function $\widetilde{f}$ defined in (3.16) the **continuous extension** of f on $D \cup \{x_0\}$.

Arguing exactly as in the proof of Theorem 3.26, we also have the validity of the following:

Lemma 3.70: Limit vs Sequences

Let $f : D \to \mathbb{R}$. Then $L = \lim_{x\to\bar{x}} f(x)$ if and only if, for every sequence $(x_n)_{n=0}^{\infty} \subset D$ converging to $\bar{x}$, $\lim_{n\to\infty} f(x_n) = L$ also holds.

Finally, we state a result on the limit of the composition with a continuous function.

Proposition 3.71: Limit and Composition

Let $E \subset \mathbb{R}$, and let $f : D \to E$ be such that the limit $\bar{y} = \lim_{x\to\bar{x}} f(x)$ exists and belongs to E. Let $g : E \to \mathbb{R}$ be continuous at $\bar{y}$. Then $\lim_{x\to\bar{x}} g(f(x)) = g(\bar{y})$ holds.

Proof. Let $(x_n)_{n=0}^{\infty} \subset D$ be a sequence converging to $\bar{x}$. By Lemma 3.70, $\lim_{n\to\infty} f(x_n) = \bar{y}$
15 holds, and therefore, since g is continuous at $\bar{y}$, Theorem 3.26 yields $\lim_{n\to\infty} g(f(x_n)) = g(\bar{y})$. Since $(x_n)_{n=0}^{\infty}$ is an arbitrary sequence converging to $\bar{x}$, it follows from Lemma 3.70 that $\lim_{x\to\bar{x}} g(f(x)) = g(\bar{y})$, as desired. □

3.72. — We can introduce conventions for improper limits of functions, as we have already done for sequences.

Definition 3.73: Improper limits

Let $f : D \to \mathbb{R}$, and assume that $D \cap (x_0 - \delta, x_0 + \delta) \neq \emptyset$ for every $\delta > 0$.
We say that f **diverges to** ∞ **as** $x \to x_0$**,** and write $\lim_{x\to x_0} f(x) = \infty$, if for every real number $M > 0$ there exists $\delta > 0$ with the property that

$$x \in D \cap (x_0 - \delta, x_0 + \delta) \implies f(x) \geq M \qquad \forall\, x \in D.$$

Analogously, $\lim_{x\to x_0} f(x) = -\infty$, if for every real number $M > 0$ there exists $\delta > 0$ with the property that

$$x \in D \cap (x_0 - \delta, x_0 + \delta) \implies f(x) \leq -M \qquad \forall\, x \in D.$$

16 3.5.2 One-sided Limits

In addition to what was discussed above, several other limit terms are useful: In this subsection, we introduce one which is sensitive to which direction the limit point x_0 is approached, and where the limit point x_0 may also be one of the symbols ∞ or $-\infty$.

Definition 3.74: One-sided Limit

Let $f : D \to \mathbb{R}$, and assume that $D \cap [x_0, x_0 + \delta) \neq \emptyset$ for every $\delta > 0$.
A real number L is called **limit from the right** of f at x_0 if for every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$x \in D \cap [x_0, x_0 + \delta) \implies |f(x) - L| < \varepsilon.$$

If the limit from the right of f exists at x_0, we use the notation

$$L = \lim_{\substack{x\to x_0 \\ x\geq x_0}} f(x). \tag{3.17}$$

If x_0 is an accumulation point of $D \cap (x_0, \infty)$ then, as in Paragraph 3.68, we can also use a version of the limit from the right where $x \neq x_0$ and we write

$$L = \lim_{\substack{x \to x_0 \\ x > x_0}} f(x), \qquad \text{or also} \qquad L = \lim_{x \to x_0^+} f(x).$$

Analogous to limits from the right, we can also define limits from the left, with the notation

$$\lim_{\substack{x \to x_0 \\ x \leq x_0}} f(x) \qquad \text{and} \qquad \lim_{x \to x_0^-} f(x) = \lim_{\substack{x \to x_0 \\ x < x_0}} f(x).$$

Finally, we can also allow the symbols $+\infty$ and $-\infty$ for one-sided limits, as in Paragraph 3.72.

Definition 3.75: Limits at Infinity

Let $f : D \to \mathbb{R}$, and assume that $D \cap (R, \infty) \neq \emptyset$ for every $R > 0$.
A real number L is called **limit of f as $x \to \infty$** if for every $\varepsilon > 0$ there exists $R > 0$ such that

$$x \in D \cap (R, \infty) \implies |f(x) - L| < \varepsilon.$$

Instead, we say that f **diverges to ∞ as $x \to \infty$** if for every $M > 0$ there exists $R > 0$ such that

$$x \in D \cap (R, \infty) \implies f(x) \geq M.$$

16

If the limit of f as $x \to \infty$ exists, then it is unique and we write

$$L = \lim_{x \to \infty} f(x).$$

If f diverges to ∞ as as $x \to \infty$, we write

$$\lim_{x \to \infty} f(x) = \infty.$$

Of course, also limits at $-\infty$ and/or the case when f diverges to $-\infty$ can be considered, and the definitions are analogous.

3.76. — Limits as $x \to \infty$ can be transformed into limits from the right as $x \to 0$. Indeed, if D and f are given as in Definition 3.75, consider the set E and the function $g : E \to \mathbb{R}$ defined as

$$E = \{x \in \mathbb{R}_{>0} \mid x^{-1} \in D\}, \qquad g(x) = f(x^{-1}).$$

Then it holds

$$\lim_{x \to \infty} f(x) = \lim_{x \to 0^+} g(x).$$

This means, in particular, that one limit exists if and only if the other limit exists.

Definition 3.77: One-sided Continuity and Jumps

Let $f : D \to \mathbb{R}$, and let $x_0 \in D$. If the limit from the right

$$\lim_{x \to x_0^+} f(x)$$

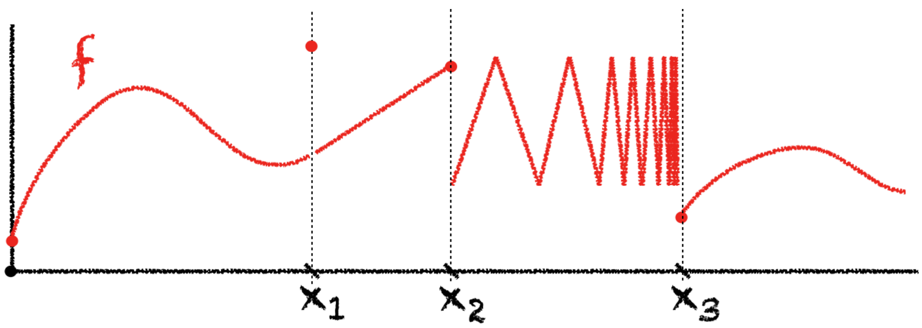
exists and it is equal to $f(x_0)$, then we say that f is **continuous from the right** at x_0. Similarly, we define **continuous from the left**.
We call $x_0 \in D$ a **jump point** if the one-sided limits

$$\lim_{x \to x_0^-} f(x) \qquad \text{and} \qquad \lim_{x \to x_0^+} f(x) \tag{3.18}$$

both exist but are different.

3.78. — The following graph represents a function with three points of discontinuity x_1, x_2, x_3.

16



The discontinuity x_1 is a removable discontinuity. At x_1, the function is neither continuous from the left nor from the right. The point x_1 is also not a jump point. Although both limit values (3.18) exist for $x \to x_1$, they are equal.

At the point x_2, the function f is continuous from the left but not continuous from the right. The point x_2 is a jump point. At the point x_3, f is continuous from the right. Finally, x_3 is not a jump point because the limit from the left does not exist.

Example 3.79. — The domain of definition for all functions in this example is $D = \mathbb{R}_{>0}$. We want to study the limit as $x \to 0^+$ of the function $f : D \to \mathbb{R}$ given by

$$f(x) = x^x = \exp(x \log(x)).$$

To do this, let us first calculate two other limits.

1. We claim that

$$\lim_{y \to \infty} y \exp(-y) = 0 \tag{3.19}$$

holds. Indeed, Corollary 3.53 implies that $\exp(y) \geq (1+\frac{y}{2})^2$ holds for $y > 0$. This gives $0 \leq y\exp(-y) \leq \frac{y}{(1+\frac{y}{2})^2} \leq \frac{4}{y}$, which implies (3.19) because of the sandwich lemma.

2. Next we want to show that
$$\lim_{x\to 0^+} x\log x = 0. \tag{3.20}$$
So let $\varepsilon > 0$. Because of (3.19) there exists $R > 0$ such that $|y\exp(-y)| < \varepsilon$ for all $y > R$. Set $\delta = \exp(-R)$ and consider $x \in (0,\delta)$. Then $y = -\log x$ satisfies $y > R$ due to the strict monotonicity of the logarithm, therefore $|x\log x| = |\exp(-y)y| < \varepsilon$, which shows (3.20).

Because of Proposition 3.71 and since the exponential mapping is continuous, (3.20) yields

$$\lim_{x\to 0^+} x^x = \lim_{x\to 0^+} \exp(x\log x) = \exp(0) = 1.$$

Exercise 3.80. — Let $a \in \mathbb{R}$. Calculate the following limits, if they exist:

$$\lim_{x\to 2} \frac{x^3 - x^2 - x - 2}{x-2}, \qquad \lim_{x\to\infty} \frac{3e^{2x} + e^x + 1}{2e^{2x} - 1}, \qquad \lim_{x\to\infty} \frac{e^x}{x^a}, \qquad \lim_{x\to\infty} \frac{\log(x)}{x^a}.$$

In each case, choose a suitable domain on which the given formulas define a function.

16

3.5.3 Landau Notation

We now introduce two common notations that relate the asymptotic behaviour of a function to the asymptotic behaviour of another function – that is, describe relative asymptotic behaviour. These notations are named after the German-Jewish mathematician Edmund Georg Hermann Landau (1877 - 1938).

Definition 3.81: Landau Big-O

Let $f, g : D \to \mathbb{R}$. We write

$$f(x) = O(g(x)) \quad \text{as } x \to x_0$$

if there exist $M > 0$ and $\delta > 0$ such that

$$x \in D \cap (x_0 - \delta, x_0 + \delta) \implies |f(x)| \leq M|g(x)|.$$

Then we say that f is a **big-O** of g as $x \to x_0$.

If $g(x) \neq 0$ for all $x \in D \setminus \{x_0\}$ sufficiently close to x_0, then $f(x) = O(g(x))$ is equivalent to $\frac{f(x)}{g(x)}$ being bounded in a neighbourhood of x_0.

We can also allow for x_0 the elements ∞ and $-\infty$ of the extended number line, as discussed in the next definition. We define only the case at ∞, the definition for $-\infty$ is analogous.

Definition 3.82: Landau Big-O at Infinity

Let $f, g : D \to \mathbb{R}$. We write

$$f(x) = O(g(x)) \quad \text{as } x \to \infty$$

if there exist $M > 0$ and $R > 0$ such that

$$x \in D \cap (R, \infty) \implies |f(x)| \leq M|g(x)|.$$

The advantage of this notation is that we do not need to introduce the name for the upper bound M. If we are not particularly interested in this constant, then we can concentrate on the essentials in calculations. In this context, one also speaks of **implicit constants**.

Example 3.83. —
- If f and g are bounded and continuous in a neighborhood of x_0 with $g(x_0) \neq 0$, then $f(x) = O(g(x))$ holds.
- It holds
$$x^2 = O(x) \quad \text{as } x \to 0$$
but not $x = O(x^2)$, since $\frac{x}{x^2}$ is not bounded near 0.
- It holds

16
$$\frac{3x^3}{x^3+3} = O(1) \quad \text{as } x \to \infty$$
but not $\frac{3x^3}{x^3+3} = O(x^\alpha)$ for $\alpha < 0$.

As discussed above, the big-O means that f is bounded by a multiple of g. One may also consider a stronger condition, namely that f is asymptotically negligible with respect to g. This leads to the following definition.

Definition 3.84: Little-o

Let $f, g : D \to \mathbb{R}$. We write

$$f(x) = o(g(x)) \quad \text{as } x \to x_0$$

if for every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$x \in D \cap (x_0 - \delta, x_0 + \delta) \implies |f(x)| \leq \varepsilon|g(x)|.$$

Then we say that f is a **little-o** of g as $x \to x_0$.

If $g(x) \neq 0$ for all x in a neighbourhood of x_0, then $f(x) = o(g(x))$ as $x \to x_0$ is equivalent to

$$\lim_{x \to x_0} \frac{f(x)}{g(x)} = 0.$$

REMARK 3.85. — If $f(x) = o(g(x))$ then $f(x) = O(g(x))$.

Again, we can consider the little-o condition at infinity.

DEFINITION 3.86: LANDAU LITTLE-O AT INFINITY

Let $f, g : D \to \mathbb{R}$. We write

$$f(x) = o(g(x)) \quad \text{as } x \to \infty$$

if for every $\varepsilon > 0$ there exists $R > 0$ such that

$$x \in D \cap (R, \infty) \implies |f(x)| \le \varepsilon |g(x)|.$$

EXAMPLE 3.87. —

- $x = o(x^2)$ as $x \to \infty$, and $x^2 = o(x)$ as $x \to 0$.
- For any $\alpha < 1$, the following holds true

$$\frac{3x^3}{2x^2 + x^{10}} = o(|x|^\alpha) \quad \text{as } x \to 0$$

but not for $\alpha \ge 1$. In fact, for any $\alpha < 1$ the limit exists.

16

$$\lim_{x\to 0} \left| \frac{3x^3}{|x|^\alpha (2x^2 + x^{10})} \right| = \lim_{x\to 0} |x|^{1-\alpha} \frac{3}{(2 + x^8)} = \frac{3}{2} \lim_{x\to 0} |x|^{1-\alpha}$$

and is equal to 0.

EXERCISE 3.88. — Let $p > 1$, $a \in \mathbb{R}$, and $b > 0$. Show the following:

1. $x^p = o(x) \quad \text{as } x \to 0$;
2. $x = o(x^p) \quad \text{as } x \to \infty$;
3. $x^a = o(e^x) \quad \text{as } x \to \infty$;
4. $\log(x) = o(x^b) \quad \text{as } x \to \infty$.

EXERCISE 3.89. — Let f_1, f_2, g be real-valued functions on D. Show that if $f_1(x) = o(g(x))$ and $f_2(x) = o(g(x))$ as $x \to x_0$, then

$$\alpha_1 f_1(x) + \alpha_2 f_2(x) = o(g(x)) \quad \text{as } x \to x_0$$

for every $\alpha_1, \alpha_2 \in \mathbb{R}$. Formulate and show the analogous statement for big-O.

EXERCISE 3.90. — Let f_1, f_2, g_1, g_2 be real-valued functions on D. Show that:
- If $f_1(x) = o(g_1(x))$ and $f_2(x) = o(g_2(x))$ as $x \to x_0$, then $f_1(x) f_2(x) = o(g_1(x) g_2(x))$ as

$x \to x_0$;

• If $f_1(x) = o(g_1(x))$ and $f_2(x) = O(g_2(x))$ as $x \to x_0$, then $f_1(x)f_2(x) = o(g_1(x)g_2(x))$ as $x \to x_0$;

• If $f_1(x) = O(g_1(x))$ and $f_2(x) = O(g_2(x))$ as $x \to x_0$, then $f_1(x)f_2(x) = O(g_1(x)g_2(x))$ as $x \to x_0$.

Example 3.91. — Let $f(x) = x+x^3+4x^4+x^7$ and $g(x) = x+\frac{3x^2}{1+x}$. Then $f(x) = x+o(x^2)$ and $g(x) = x + O(x^2)$ as $x \to 0$. In particular, their product satisfies the following:

$$\begin{aligned} f(x)g(x) &= \big(x + o(x^2)\big)\big(x + O(x^2)\big) = x^2 + o(x^2)x + xO(x^2) + o(x^2) \cdot O(x^2) \\ &= x^2 + o(x^3) + O(x^3) + o(x^4) = x^2 + O(x^3) \qquad \text{as } x \to 0. \end{aligned}$$

3.92. — Landau notation is often used as a placeholder, for example, to express that one term in a sum is increasing or decreasing faster than the others. In an expression of the form

$$f(x) + o(g(x)) \quad \text{as } x \to x_0$$

16 the term $o(g(x))$ stands for a function $h : D \to \mathbb{R}$ with the property that

$$h(x) = o(g(x)) \quad \text{as } x \to x_0.$$

This applies analogously to the big-O notation.

Example 3.93. — One writes

$$\begin{aligned} \frac{x^3 - 7x^2 + 6x + 2}{x^2} &= x - 7 + O\left(\frac{1}{x}\right) && \text{as } x \to \infty \\ &= x - 7 + o(1) && \text{as } x \to \infty \\ &= x + O(1) && \text{as } x \to \infty \\ &= x + o(x) && \text{as } x \to \infty \end{aligned}$$

and thus remembers on the right-hand side only those terms that make up the bulk of the term as $x \to \infty$. It may perhaps come as a surprise that, in the above example, all four formulas could be true or useful. The assertions all follow directly from polynomial division with remainder, and depending on the context, one might want to use the slightly more precise assertion with error $-7 + O\left(\frac{1}{x}\right)$ or the coarser assertion using error $o(x)$.

3.6 Sequences of Functions

3.6.1 Pointwise Convergence

Definition 3.94: Sequences of Functions

A **sequence** of real-valued functions on a subset $D \subset \mathbb{R}$ is a family of functions $f_n : D \to \mathbb{R}$ indexed by $\mathbb{N}$. The function f_n is called the n-th **element** of the sequence. One often writes $(f_n)_{n\in\mathbb{N}}$ or also $(f_n)_{n=0}^{\infty}$ for a sequence of functions.

Definition 3.95: Pointwise Convergence

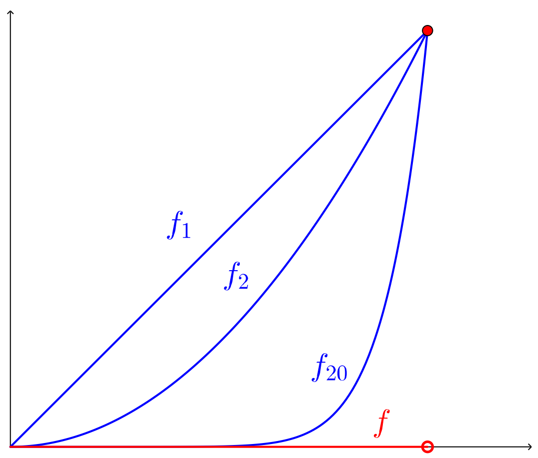
Let $D \subset \mathbb{R}$, let $(f_n)_{n=0}^{\infty}$ be a sequence of functions $f_n : D \to \mathbb{R}$, and let $f : D \to \mathbb{R}$ be a function. We say the sequence $(f_n)_{n=0}^{\infty}$ converges **pointwise** to f if for every $x \in D$ the sequence of real numbers $(f_n(x))_{n=0}^{\infty}$ converges to $f(x)$. In this case we call f the **pointwise limit** of the sequence $(f_n)_{n=0}^{\infty}$.

Exercise 3.96. — Show that the pointwise limit of a sequence of functions is uniquely determined if it exists.

3.97. — In the following example we show that in general the continuity property is not
17 preserved under pointwise convergence.

Example 3.98. — Let $D = [0,1]$ and let $f_n : D \to \mathbb{R}$ be given by $f_n(x) = x^n$. Then the sequence of continuous functions $(f_n)_{n=0}^{\infty}$ converge pointwise to the non-continuous function $f : D \to \mathbb{R}$ given by

$$f(x) = \lim_{n\to\infty} f_n(x) = \lim_{n\to\infty} x^n = \begin{cases} 0 & \text{for } x < 1 \\ 1 & \text{for } x = 1. \end{cases}$$



3.6.2 Uniform Convergence

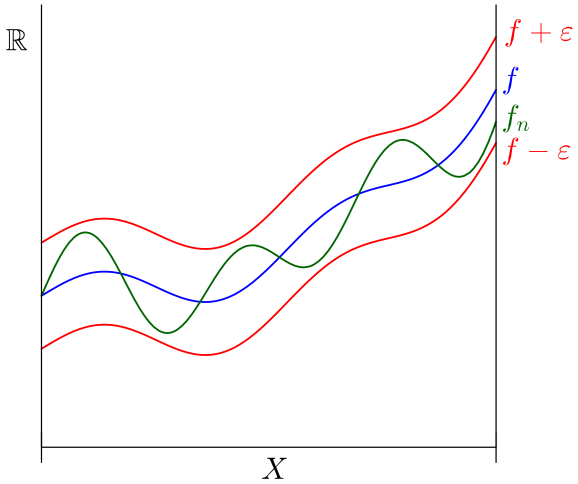
Definition 3.99: Uniform Convergence

Let $D \subset \mathbb{R}$, let $(f_n)_{n=0}^{\infty}$ be a sequence of functions $f_n : D \to \mathbb{R}$, and let $f : D \to \mathbb{R}$ be a function. We say the sequence $(f_n)_{n=0}^{\infty}$ **converges uniformly** to f if for every $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that for all $n \geq N$ and all $x \in D$

$$|f_n(x) - f(x)| < \varepsilon.$$

3.100. — The estimate $|f_n(x) - f(x)| < \varepsilon$ is equivalent to $f(x) - \varepsilon \leq f_n(x) \leq f(x) + \varepsilon$. Thus, uniform convergence can also be described by the graph of a function sequence and its limit function, as the following figure shows: The function sequence f_n converges uniformly to f if for every $\varepsilon > 0$ the graph of f_n lies in the "ε-tube" around f for all sufficiently large n.

17



Exercise 3.101. — Let D be a set and let $(f_n)_{n=0}^{\infty}$ be a sequence of functions $f_n : D \to \mathbb{R}$. Show that if $(f_n)_{n=0}^{\infty}$ converges uniformly to a function f, then $(f_n)_{n=0}^{\infty}$ also converges pointwise to f.

Theorem 3.102: Uniform Convergence Preserves Continuity

Let $D \subset \mathbb{R}$ and let $(f_n)_{n=0}^{\infty}$ be a sequence of continuous functions $(f_n)_{n=0}^{\infty}$ converging uniformly to $f : D \to \mathbb{R}$. Then f is continuous.

Proof. Let $\bar{x} \in D$ and $\varepsilon > 0$. First, by the uniform convergence of f_n to f, there exists $N \in \mathbb{N}$ such that $|f_N(y) - f(y)| < \frac{\varepsilon}{3}$ for all $y \in D$. Then, since f_n is continuous at $\bar{x}$, there exists $\delta > 0$ such that

$$|x - \bar{x}| < \delta \implies |f_N(x) - f_N(\bar{x})| < \frac{\varepsilon}{3}$$

holds for all $x \in D$. Then, for all $x \in D$ with $|x - \bar{x}| < \delta$ it follows that

$$\begin{aligned} |f(x) - f(\bar{x})| &= |f(x) - f_N(x) + f_N(x) - f_N(\bar{x}) + f_N(\bar{x}) - f(\bar{x})| \\ &\leq |f(x) - f_N(x)| + |f_N(x) - f_N(\bar{x})| + |f_N(\bar{x}) - f(\bar{x})| \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon, \end{aligned}$$

which proves that f is continuous at $\bar{x}$. Since $\bar{x}$ is arbitrary in D, the theorem follows. □

Exercise 3.103. — Let $(f_n)_{n=0}^{\infty}$ be a sequence of bounded real-valued function on a set D, and let $f : D \to \mathbb{R}$ be another real-valued function on D. Suppose that $D = D_1 \cup D_2$ for two subsets, $(f_n|_{D_1})_{n=0}^{\infty}$ tends uniformly towards $f|_{D_1}$ and $(f_n|_{D_2})_{n=0}^{\infty}$ tends uniformly towards $f|_{D_2}$. Show that then $(f_n)_{n=0}^{\infty}$ also tends uniformly towards f.

Exercise 3.104. — Let $(f_n)_{n=0}^{\infty}$ be a sequence of bounded real-valued functions on a set
17 D. Show that if $(f_n)_{n=0}^{\infty}$ converges uniformly to a function $f : D \to \mathbb{R}$, then f is also bounded. Find also an example in which a sequence $(f_n)_{n=0}^{\infty}$ of bounded functions converges pointwise to an unbounded function.

Exercise 3.105. — Let $D \subset \mathbb{R}$ and $(f_n)_{n=0}^{\infty}$ be a sequence of uniformly continuous real-valued functions on D that tends uniformly to $f : D \to \mathbb{R}$. Let $(x_n)_{n=0}^{\infty}$ be a sequence in D that converges towards $\bar{x} \in D$. Show that

$$\lim_{n \to \infty} f_n(x_n) = f(\bar{x}). \tag{3.21}$$

Find an example that shows that pointwise convergence of $(f_n)_{n=0}^{\infty}$ to f is not sufficient to infer (3.21).

Exercise 3.106. — Let $D \subset \mathbb{R}$ and $(f_n)_{n=0}^{\infty}$ be a sequence of uniformly continuous real-valued functions on D, uniformly converging to $f : D \to \mathbb{R}$. Show that f is uniformly continuous.

Chapter 4

Series and Power Series

In this chapter we will consider so-called series, i.e., "infinite sums", which will lead us to the definitions of known functions, in particular to the definitions of trigonometric functions.

4.1 Series of Real Numbers

Definition 4.1: Convergent Series

Let $(a_n)_{n=0}^{\infty}$ be a sequence of real numbers, and let $A \in \mathbb{R}$. We say that the series $\sum_{k=0}^{\infty} a_k$ converges to $A \in \mathbb{R}$ if

$$A = \lim_{n\to\infty} \sum_{k=0}^{n} a_k.$$

In other words, computing the infinite sum $\sum_{k=0}^{\infty} a_k$ corresponds to finding (if it exists) the limit of the sequence $(s_n)_{n=0}^{\infty}$ given by the **partial sums**

17

$$s_n = \sum_{k=0}^{n} a_k.$$

We call a_n the **n-th element** or the **n-th summand** of the series. We call the series $\sum_{k=0}^{\infty} a_k$ **convergent** if the limit exists, in which case we call it the **value of the series**. Otherwise, the series is **not convergent**.
If the sequence s_n diverges to ∞ (respectively, $-\infty$), then we call the series $\sum_{k=0}^{\infty} a_k$ **divergent** to ∞ (respectively, $-\infty$).

Remark 4.2. — Unless otherwise specified, all sequences always consist of real numbers.

Proposition 4.3: Necessary Condition for Convergence

If the series $\sum_{k=0}^{\infty} a_k$ *converges, then the sequence* $(a_n)_{n=0}^{\infty}$ *is a* ***null sequence****, that is,* $\lim_{n\to\infty} a_n = 0$.

Proof. By assumption, the partial sums $s_n = \sum_{k=0}^{n} a_k$ for $n \in \mathbb{N}$ have a limit $\lim_{n\to\infty} s_n = S$. Thus, since $a_n = s_n - s_{n-1}$, the following holds:

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} (s_n - s_{n-1}) = S - S = 0. \qquad \square$$

Example 4.4 (Geometric Series). — The **geometric series** $\sum_{n=0}^{\infty} q^n$ to $q \in \mathbb{R}$ converges exactly when $|q| < 1$ and, in this case,

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

Indeed, by Proposition 4.3, convergence of the series implies that $q^k \to 0$, hence $|q| < 1$.

Conversely, for $|q| < 1$, first one prove by induction on $n \in \mathbb{N}$ the validity of the "geometric sum formula"

17

$$\sum_{k=0}^{n} q^k = \frac{1-q^{n+1}}{1-q} \qquad \forall\, n \in \mathbb{N},\ q \neq 1.$$

Then, thanks to Example 2.134, it holds that

$$\sum_{k=0}^{n} q^k = \frac{1-q^{n+1}}{1-q} \to \frac{1}{1-q} \qquad \text{as } n \to \infty.$$

Example 4.5 (Harmonic Series). — The converse of Proposition 4.3 does not hold. For example, the **harmonic series** $\sum_{k=1}^{\infty} \frac{1}{k}$ is divergent. We prove the divergence with a concrete estimate.

Let $\ell \in \mathbb{N}$ and consider $n = 2^\ell$. Then the partial sum of the harmonic series for n satisfies the estimate

$$\begin{aligned}\sum_{k=1}^{2^\ell} \frac{1}{k} &= 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} + \frac{1}{9} + \cdots + \frac{1}{2^{\ell-1}+1} + \cdots + \frac{1}{2^\ell} \\ &\geq 1 + \frac{1}{2} + \underbrace{\frac{1}{4} + \frac{1}{4}}_{=\frac{1}{2}} + \underbrace{\frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8}}_{=\frac{1}{2}} + \frac{1}{16} + \cdots + \underbrace{\frac{1}{2^\ell} + \cdots + \frac{1}{2^\ell}}_{=\frac{1}{2}} = 1 + \frac{\ell}{2}.\end{aligned}$$

Since $\ell \in \mathbb{N}$ is arbitrary we see that the partial sums are not bounded, and therefore, the harmonic series is divergent.

Exercise 4.6. — Let $\sum_{k=0}^{\infty} a_k$, $\sum_{k=0}^{\infty} b_k$ be convergent series, and let $\alpha, \beta \in \mathbb{R}$. Prove that the series $\sum_{k=0}^{\infty}(\alpha a_k + \beta b_k)$ converge, with

$$\sum_{k=0}^{\infty}(\alpha a_k + \beta b_k) = \alpha \sum_{k=0}^{\infty} a_k + \beta \sum_{k=0}^{\infty} b_k.$$

Lemma 4.7: Convergence of the Tail

Let $\sum_{k=0}^{\infty} a_k$ be a series. For each $N \in \mathbb{N}$, the series $\sum_{k=N}^{\infty} a_k$ is convergent if and only if the series $\sum_{k=0}^{\infty} a_k$ is convergent, and in this case

$$\sum_{k=0}^{\infty} a_k = \sum_{k=0}^{N-1} a_k + \sum_{k=N}^{\infty} a_k.$$

The same statement holds replacing convergent with divergent.

Proof. For every $n \geq N$, it holds

$$\sum_{k=0}^{n} a_k = \sum_{k=0}^{N-1} a_k + \sum_{k=N}^{n} a_k.$$

17 In particular, the partial sums of $\sum_{k=N}^{\infty} a_k$ converge exactly when the partial sums of $\sum_{k=0}^{\infty} a_k$ converge. The case of a divergent sequence is completely analogous. □

4.1.1 Series with Non-negative Elements

Proposition 4.8: Non-negative Series either Converge or Diverge

Given a series $\sum_{k=0}^{\infty} a_k$ with non-negative elements $a_k \geq 0$ for all $k \in \mathbb{N}$, the partial sums $s_n = \sum_{k=0}^{n} a_k$ form a monotonically increasing sequence. If the sequence $(s_n)_{n=0}^{\infty}$ is bounded, then the series $\sum_{k=0}^{\infty} a_k$ converges; otherwise, it diverges to ∞.

Proof. From $a_{n+1} \geq 0$ it follows that $s_{n+1} = s_n + a_{n+1} \geq s_n$ for all $n \in \mathbb{N}$, so the sequence $(s_n)_{n=0}^{\infty}$ is increasing. If the partial sums $\{s_n \mid n \in \mathbb{N}\}$ are bounded, then they converge according to Theorem 2.108. □

Remark 4.9. — If $\sum_{k=0}^{\infty} a_k$ is a series with non-negative elements, then the sequence of partial sums $(s_n)_{n=0}^{\infty}$ is bounded if and only if there exists a subsequence $(s_{n_k})_{k=0}^{\infty}$ which is bounded (see Remark 2.109).

Corollary 4.10: Majorant and Minorant Criterion

Let $\sum_{k=0}^{\infty} a_k$, $\sum_{k=0}^{\infty} b_k$ *be two series such that* $0 \leq a_k \leq b_k$ *for all* $k \in \mathbb{N}$. *Then* $0 \leq \sum_{k=0}^{\infty} a_k \leq \sum_{k=0}^{\infty} b_k$ *holds and, in particular,*

$$\sum_{k=0}^{\infty} b_k \text{ convergent} \implies \sum_{k=0}^{\infty} a_k \text{ convergent}$$
$$\sum_{k=0}^{\infty} a_k \text{ divergent to } \infty \implies \sum_{k=0}^{\infty} b_k \text{ divergent to } \infty.$$

These implications hold even if $0 \leq a_n \leq b_n$ *holds only for* $n \geq N$, *for some* $N \in \mathbb{N}$.

Proof. From $a_k \leq b_k$ it follows $\sum_{k=0}^{n} a_k \leq \sum_{k=0}^{n} b_k$ for all $n \in \mathbb{N}$. Thus, according to the monotonicity of the sequence of partial sums,

$$\sum_{k=0}^{\infty} a_k = \sup\left\{\sum_{k=0}^{n} a_k \mid n \in \mathbb{N}\right\} \leq \sup\left\{\sum_{k=0}^{n} b_k \mid n \in \mathbb{N}\right\} = \sum_{k=0}^{\infty} b_k.$$

The last statement of the proposition is a consequence of Lemma 4.7. □

Under the assumptions of the corollary, one calls the series $\sum_{k=0}^{\infty} b_k$ a majorant of the series $\sum_{k=0}^{\infty} a_k$, and the latter is a minorant of the series $\sum_{k=0}^{\infty} b_k$. This is why one refers to
17 the above result as the **majorant** and the **minorant criterion**.

Example 4.11. — The series $\sum_{k=0}^{\infty} \frac{1}{k^2}$ is convergent. In fact, $a_k = \frac{1}{k^2} \leq \frac{1}{k(k-1)} = b_k$ holds for $k \geq 2$. Also, since $\frac{1}{k(k-1)} = \frac{1}{k-1} - \frac{1}{k}$, the series $\sum_{k=2}^{\infty} b_k$ is convergent since its partial sums can be computed by solving a telescopic sum:

$$\begin{aligned}\sum_{k=2}^{n} \frac{1}{k(k-1)} &= \sum_{k=2}^{n}\left(\frac{1}{k-1} - \frac{1}{k}\right)\\ &= \left(1 - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \left(\frac{1}{3} - \frac{1}{4}\right) + \ldots + \left(\frac{1}{n-1} - \frac{1}{n}\right) = 1 - \frac{1}{n}.\end{aligned}$$

In particular $\sum_{k=2}^{\infty} \frac{1}{k^2} \leq \sum_{k=2}^{\infty} b_k = 1$.

Example 4.12. — We want to show that the series

$$\sum_{n=0}^{\infty} a_n \qquad \text{with} \qquad a_n = \frac{2n - 10}{n^3 - 10n + 100}$$

converges. Note that $\lim_{n\to\infty} n^2 a_n = 2$, therefore there exists $N \in \mathbb{N}$ such that $0 \leq n^2 a_n \leq 3$ for all $n \geq N$, or equivalently $0 \leq a_n \leq 3n^{-2}$ for all $n \geq N$. Thus, Corollary 4.10 and Example 4.11 imply that the series converges.

PROPOSITION 4.13: CAUCHY CONDENSATION TEST

Let $(a_k)_{k=0}^{\infty}$ be a monotonically decreasing sequence of non-negative real numbers. Then

$$\sum_{k=0}^{\infty} a_k \quad \textit{converges} \iff \sum_{k=0}^{\infty} 2^k a_{2^k} \quad \textit{converges.}$$

Proof. (Extra material) Due to the monotonicity of the sequence $(a_k)_{k=0}^{\infty}$, the following inequalities hold:

$$1 \cdot a_1 \geq a_2 \geq 1 \cdot a_2, \qquad 2 \cdot a_2 \geq a_3 + a_4 \geq 2 \cdot a_4,$$

$$4 \cdot a_4 \geq a_5 + a_6 + a_7 + a_8 \geq 4 \cdot a_8, \qquad 8 \cdot a_8 \geq a_9 + \ldots + a_{16} \geq 8 \cdot a_{16},$$

and, more in general,

$$2^k a_{2^k} \geq a_{2^k+1} + \ldots + a_{2^{k+1}} \geq 2^k a_{2^{k+1}} \qquad \forall\, k \in \mathbb{N}.$$

Summing all these inequalities for $k = 0, \ldots, n$ gives

$$\sum_{k=0}^{n} 2^k a_{2^k} \geq \sum_{k=0}^{n} \left(a_{2^k+1} + \ldots + a_{2^{k+1}}\right) = \sum_{j=2}^{2^{n+1}} a_j$$

17 and

$$\sum_{j=2}^{2^{n+1}} a_j = \sum_{k=0}^{n} \left(a_{2^k+1} + \ldots + a_{2^{k+1}}\right) \geq \sum_{k=0}^{n} 2^k a_{2^{k+1}} = \frac{1}{2}\sum_{k=0}^{n} 2^{k+1} a_{2^{k+1}} = \frac{1}{2}\sum_{k=1}^{n+1} 2^k a_{2^k}.$$

This shows that

$$\sum_{k=0}^{n} 2^k a_{2^k} \geq \sum_{j=2}^{2^{n+1}} a_j \geq \frac{1}{2}\sum_{k=1}^{n+1} 2^k a_{2^k}.$$

Because of Remark 4.9 and Corollary 4.10, it follows that the series $\sum_{k=0}^{\infty} 2^k a_{2^k}$ is bounded (and therefore converge) if and only if the series $\sum_{k=0}^{\infty} a_k$ is bounded (and therefore converge). □

EXAMPLE 4.14. — Given $p \in \mathbb{R}$, the series $\sum_{n=1}^{\infty} \frac{1}{n^p}$ converges exactly when $p > 1$. Indeed:

- For $p \leq 0$ we see that $\frac{1}{n^p} \geq 1$, so the series diverges according to Proposition 4.3.
- For $p > 0$, since the sequence $\left(\frac{1}{n^p}\right)_{n=1}^{\infty}$ is decreasing, Proposition 4.13 implies that $\sum_{n=1}^{\infty} \frac{1}{n^p}$ converges if and only if

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

converges. According to Example 4.4, this series converges exactly when $2^{1-p} < 1$, that is, $p > 1$.

Remark 4.15. — The argument in Example 4.14 gives an alternative proof that the harmonic series diverges (see Example 4.5).

17

Exercise 4.16. — Given $p \in \mathbb{R}$, the series $\sum_{n=2}^{\infty} \frac{1}{n \log(n)^p}$ converges exactly when $p > 1$. *Hint:* for $p \leq 0$, compare the series above with the harmonic series; for $p > 0$, use Proposition 4.13 and Example 4.14.

Exercise 4.17. — Is the series $\sum_{n=3}^{\infty} \frac{1}{n \log(n) \log\log(n)}$ convergent or divergent?

4.1.2 Conditional Convergence

Definition 4.18: Absolute and Conditional Convergence

We say that a series $\sum_{k=0}^{\infty} a_k$ is **absolutely convergent** if the series $\sum_{k=0}^{\infty} |a_k|$ converges.
The series $\sum_{k=0}^{\infty} a_k$ is **conditionally convergent** if it converges but is not absolutely convergent.

The critical property of a conditionally convergent sequence is that one can rearrange the terms to obtain any possible limit!

Theorem 4.19: Riemann's Rearrangement Theorem

18 *Let $\sum_{n=0}^{\infty} a_n$ be a conditionally convergent series with real members, and let $A \in \mathbb{R}$. There exists a bijection $\varphi : \mathbb{N} \to \mathbb{N}$ such that*

$$A = \sum_{n=0}^{\infty} a_{\varphi(n)}. \tag{4.1}$$

Proof. (Extra material) Let $\sum_{n=0}^{\infty} a_n$ be a conditionally convergent series. Then $a_n \to 0$ as $n \to \infty$ (by Proposition 4.3) and $\sum_{n=0}^{\infty} |a_n| = \infty$ by assumption.

We divide the natural numbers $\mathbb{N}$ into two parts:

$$P = \{n \in \mathbb{N} \mid a_n \geq 0\}, \qquad N = \{n \in \mathbb{N} \mid a_n < 0\}$$

depending on the sign of the corresponding a_n, and we enumerate the elements in P and N in ordered sequence, i.e., $P = \{p_0, p_1, \ldots\}$ and $N = \{n_0, n_1, \ldots\}$, with $p_0 < p_1 < p_2 < \cdots$ and

$n_0 < n_1 < \cdots$. Note that we must have

$$\sum_{k=0}^{\infty} a_{p_k} = \infty \qquad \text{and} \qquad \sum_{k=0}^{\infty} -a_{n_k} = \infty. \tag{4.2}$$

Indeed, if for instance $\sum_{k=0}^{\infty} a_{p_k} = \infty$ but $\sum_{k=0}^{\infty}(-a_{n_k}) < \infty$, then $\sum_{n=1}^{\infty} a_n = \infty$. Or if both $\sum_{k=0}^{\infty} a_{p_k} < \infty$ and $\sum_{k=0}^{\infty}(-a_{n_k}) < \infty$, then $\sum_{n=1}^{\infty} |a_n| < \infty$. In any case, this would give a contradiction, so (4.2) must hold.

Now, given $A \in \mathbb{R}$, we construct the bijective mapping $\varphi : \mathbb{N} \to \mathbb{N}$ recursively in the following way: If $A < 0$, set $\varphi(0) = n_1$, and if $A \geq 0$, set $\varphi(0) = p_1$. Then, assuming we have already defined $\varphi(0), \varphi(1), \ldots, \varphi(n)$, we consider $s_n = a_{\varphi(0)} + a_{\varphi(1)} + \cdots + a_{\varphi(n)}$ and define

$$\varphi(n+1) = \begin{cases} \min(P \setminus \{\varphi(0), \varphi(1), \ldots, \varphi(n)\}) & \text{if } s_n < A \\ \min(N \setminus \{\varphi(0), \varphi(1), \ldots, \varphi(n)\}) & \text{if } s_n \geq A. \end{cases}$$

The mapping $\varphi : \mathbb{N} \to \mathbb{N}$ thus defined is injective by construction, and surjective due to the divergence of the series (4.2). Since $a_n \to 0$ as $n \to \infty$, the sequence of partial sums $(s_n)_{n=0}^{\infty}$ converges to A, which shows (4.1). □

Exercise 4.20. — Complete the details omitted from the proof of Theorem 4.19. Also, show that for A one can also take one of the symbols $-\infty$ or ∞.

18

4.1.3 Convergence Criteria of Leibniz and Cauchy

Definition 4.21. — For a sequence $(a_k)_{k=0}^{\infty}$ of non-negative numbers, we call the series $\sum_{k=0}^{\infty}(-1)^k a_k$ the corresponding **alternating series**.

Proposition 4.22: Leibniz criterion

Let $(a_k)_{k=0}^{\infty}$ be a monotonically decreasing sequence of non-negative real numbers converging to zero. Then the alternating series $\sum_{k=0}^{\infty}(-1)^k a_k$ converges and

$$\sum_{k=0}^{2n+1} (-1)^k a_k \leq \sum_{k=0}^{\infty} (-1)^k a_k \leq \sum_{k=0}^{2n} (-1)^k a_k \tag{4.3}$$

holds for all $n \in \mathbb{N}$.

Proof. For $n \in \mathbb{N}$, let $s_n = \sum_{k=0}^{n}(-1)^k a_k$. Since the sequence $(a_n)_{n=0}^{\infty}$ is decreasing and non-negative, we have

$$s_{2n+2} = s_{2n} - a_{2n+1} + a_{2n+2} \leq s_{2n},$$
$$s_{2n+1} = s_{2n-1} + a_{2n} - a_{2n+1} \geq s_{2n-1},$$

$$s_{2n+2} = s_{2n+1} + a_{2n+2} \geq s_{2n+1}.$$

for all $n \in \mathbb{N}$. In other words,

$$s_1 \leq s_3 \leq \ldots \leq s_{2n-1} \leq s_{2n+1} \leq \ldots \leq s_{2n+2} \leq s_{2n} \leq \ldots \leq s_2 \leq s_0.$$

Hence, the sequence $(s_{2n})_{n=0}^{\infty}$ is decreasing and bounded from below, while the sequence $(s_{2n+1})_{n=0}^{\infty}$ is increasing and bounded from above, so the limits $A = \lim_{n\to\infty} s_{2n+1}$ and $B = \lim_{n\to\infty} s_{2n}$ exist and satisfy

$$s_1 \leq s_3 \leq \ldots \leq s_{2n-1} \leq s_{2n+1} \leq A \leq B \leq s_{2n+2} \leq s_{2n} \leq \ldots \leq s_2 \leq s_0.$$

However, since $s_{2n+2} - s_{2n+1} = a_{2n+2}$ converges to zero, then $A = B$ and the result follows. □

Example 4.23 (Alternating Harmonic Series). — Consider the **alternating harmonic series**

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} = -\sum_{n=1}^{\infty} \frac{(-1)^n}{n} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \ldots$$

Proposition 4.22 guarantees that this series converges, while the series of its absolute values

18
$$\sum_{n=1}^{\infty} \frac{1}{n} = \sum_{n=1}^{\infty} \frac{|(-1)^{n+1}|}{n}$$

diverges to infinity (see Example 4.5). So this series is only conditionally convergent.

Theorem 4.24: Cauchy criterion

The series $\sum_{k=0}^{\infty} a_k$ converges if and only if for every $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that, for $n \geq m \geq N$,

$$\left|\sum_{k=m+1}^{n} a_k\right| < \varepsilon.$$

Proof. By definition, the series $\sum_{k=0}^{\infty} a_k$ converges if and only if the sequence of partial sums $s_n = \sum_{k=0}^{n} a_k$ converges. By Theorem 2.124, the sequence $(s_n)_{n=0}^{\infty}$ converges if and only if it is a Cauchy sequence, namely, for every $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that

$$|s_n - s_m| < \varepsilon \qquad \forall\, n, m \geq N.$$

By the definition of s_n, this is equivalent to saying that

$$\left|\sum_{k=m+1}^{n} a_k\right| < \varepsilon \qquad \forall\, n \geq m \geq N,$$

which proves the result. □

Example 4.25. — To see the divergence of the harmonic series, we can also use the Cauchy criterion. We do this by setting $\varepsilon = \frac{1}{2}$ and noticing that, for $N \in \mathbb{N}$, the following holds:

18

$$\sum_{k=N}^{2N} \frac{1}{k} = \underbrace{\frac{1}{N} + \frac{1}{N+1} + \ldots + \frac{1}{2N}}_{N+1 \text{ terms}} \geq \frac{N+1}{2N} > \frac{1}{2}.$$

Hence the harmonic series cannot converge, since it does not satisfy the Cauchy criterion in Theorem 4.24.

4.2 Absolute Convergence

In this section we will look at absolutely convergent series and prove some convergence criteria. As before, unless otherwise specified, all sequences consist of real numbers.

4.2.1 Criteria for Absolute Convergence

Proposition 4.26: Absolute Convergence implies Convergence

An absolutely convergent series $\sum_{n=0}^{\infty} a_n$ *is also convergent and satisfies the generalized triangle inequality*

$$\left|\sum_{n=0}^{\infty} a_n\right| \leq \sum_{n=0}^{\infty} |a_n|.$$

Proof. Since the series $\sum_{n=0}^{\infty} |a_n|$ converges, according to the Cauchy criterion in Theorem 4.24, for every $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that, for $n \geq m \geq N$,

$$\sum_{k=m+1}^{n} |a_k| < \varepsilon.$$

18 From this and the triangle inequality it follows that

$$\left|\sum_{k=m+1}^{n} a_k\right| \leq \sum_{k=m+1}^{n} |a_k| < \varepsilon.$$

This proves that also the series $\sum_{n=0}^{\infty} a_n$ satisfies the Cauchy criterion, so it converges.

The second part now follows from the inequality

$$\left|\sum_{k=0}^{n} a_k\right| \leq \sum_{k=0}^{n} |a_k| \leq \sum_{k=0}^{\infty} |a_k| \qquad \forall\, n \in \mathbb{N},$$

taking the limit as $n \to \infty$. □

We now prove two important criteria to guarantee the absolute convergence of a series. In their proofs, we will implicitly use the following fact:

Assume that a sequence $(x_n)_{n=0}^{\infty}$ converges to a limit α. Then, given $q > \alpha$ (respectively, $q < \alpha$) then there exists $N \in \mathbb{N}$ such that $x_n < q$ (respectively, $x_n > q$) for every $n \geq N$.

This fact is a consequence of Proposition 2.97.

Corollary 4.27: Cauchy Root Criterion

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

$$\alpha = \limsup_{n\to\infty} \sqrt[n]{|a_n|} \in \mathbb{R} \cup \{\infty\}.$$

Then

$$\alpha < 1 \implies \sum_{n=0}^{\infty} a_n \text{ converges absolutely}$$

and

$$\alpha > 1 \implies \sum_{n=0}^{\infty} a_n \text{ does not converge.}$$

Proof. Suppose $\alpha < 1$, and define $q = \frac{1+\alpha}{2} \in (\alpha, 1)$. Recalling that by definition

$$\limsup_{n\to\infty} \sqrt[n]{|a_n|} = \lim_{n\to\infty} \sup_{k\geq n} \sqrt[k]{|a_k|},$$

we have that $x_n = \sup_{k\geq n} \sqrt[k]{|a_k|} \to \alpha$ as $n \to \infty$. Hence, since $\alpha < q$, there exists $N \in \mathbb{N}$ such that

$$x_N = \sup_{k\geq N} \sqrt[k]{|a_k|} < q.$$

18

This yields $|a_k| < q^k$ for all $k \geq N$, so the series $\sum_{k=1}^{\infty} |a_k|$ converges according to Corollary 4.10 and the convergence of the geometric series in Example 2.134 (recall that $q < 1$).

If $\alpha > 1$ holds, since the limsup is an accumulation point (see Theorem 2.116), Proposition 2.90 implies the existence of a subsequence $(a_{n_k})_{k=0}^{\infty}$ such that $\lim_{k\to\infty} \sqrt[n_k]{|a_{n_k}|} > 1$. In particular, there exists $N \in \mathbb{N}$ such that

$$\sqrt[n_k]{|a_{n_k}|} > 1 \qquad \text{for all } k \geq N,$$

or equivalently, $|a_{n_k}| > 1$ for all $k \geq N$. In particular, the sequence $(a_n)_{n=0}^{\infty}$ does not converge to zero and therefore, according to Proposition 4.3, $\sum_{n=0}^{\infty} a_n$ does not converge. □

4.28. — Let $(a_n)_{n=0}^{\infty}$ be a sequence of real numbers and $\alpha = \limsup_{n\to\infty} \sqrt[n]{|a_n|}$ as in the root criterion. If $\alpha = 1$ holds, then no decision about convergence or divergence of the series $\sum_{n=0}^{\infty} a_n$ can be made using the root criterion:

- According to Example 3.37, $\sqrt[n]{1/n} \to 1$ as $n \to \infty$, and by Example 4.5 the series $\sum_{n=0}^{\infty} \frac{1}{n}$ diverges.
- On the other hand, $\sqrt[n]{1/n^2} \to 1$ as $n \to \infty$, but the series $\sum_{n=0}^{\infty} \frac{1}{n^2}$ converges according to Example 4.11.

Corollary 4.29: D'Alembert's quotient criterion

Let $(a_n)_{n=0}^{\infty}$ be a sequence of real numbers with $a_n \neq 0$ for all $n \in \mathbb{N}$, and assume that

$$\lim_{n\to\infty} \frac{|a_{n+1}|}{|a_n|} \quad \textit{exists.}$$

Let $\alpha \geq 0$ denote such a limit. Then

$$\alpha < 1 \implies \sum_{n=0}^{\infty} a_n \textit{ converges absolutely}$$

and

$$\alpha > 1 \implies \sum_{n=0}^{\infty} a_n \textit{ does not converge.}$$

Proof. The proof is similar to the one of Corollary 4.27.

Suppose first $\alpha < 1$, and define $q = \frac{1+\alpha}{2} \in (\alpha, 1)$. Since $\lim_{n\to\infty} \frac{|a_{n+1}|}{|a_n|} = \alpha < q$, there exists $N \in \mathbb{N}$ such that

$$\frac{|a_{k+1}|}{|a_k|} < q \qquad \forall\, k \geq N.$$

18 This implies that

$$|a_k| = \underbrace{\frac{|a_k|}{|a_{k-1}|}}_{<q} \cdot \underbrace{\frac{|a_{k-1}|}{|a_{k-2}|}}_{<q} \cdot \ldots \cdot \underbrace{\frac{|a_{N+1}|}{|a_N|}}_{<q} \cdot |a_N| < q^{k-N}|a_N| = \frac{|a_N|}{q^N} q^k \qquad \forall\, k > n.$$

Since $q < 1$, the series $\sum_{k=N+1}^{\infty} |a_k|$ converges absolutely.

If instead $\alpha > 1$ holds, there exists $N \in \mathbb{N}$ such that

$$\frac{|a_{k+1}|}{|a_k|} > 1 \qquad \forall\, k \geq N,$$

therefore

$$|a_k| = \frac{|a_k|}{|a_{k-1}|} \cdot \frac{|a_{k-1}|}{|a_{k-2}|} \cdot \ldots \cdot \frac{|a_{N+1}|}{|a_N|} \cdot |a_N| > |a_N| \qquad \forall\, k > n.$$

This implies that the sequence $(a_n)_{n\geq 0}$ does not converge to zero. Hence, according to Proposition 4.3, $\sum_{n=0}^{\infty} a_n$ does not converge. □

Exercise 4.30 (A Generalization of Corollary 4.29). — Let $(a_n)_{n=0}^{\infty}$ be a sequence of real numbers with $a_n \neq 0$ for all $n \in \mathbb{N}$, and define

$$\alpha_+ = \limsup_{n\to\infty} \frac{|a_{n+1}|}{|a_n|}, \qquad \alpha_- = \liminf_{n\to\infty} \frac{|a_{n+1}|}{|a_n|}.$$

Prove the following implications:

18 $$\alpha_+ < 1 \implies \sum_{n=0}^{\infty} a_n \text{ converges absolutely,} \qquad \alpha_- > 1 \implies \sum_{n=0}^{\infty} a_n \text{ does not converge.}$$

Is the second implication still true if one replaces α_- with α_+?

4.2.2 Reordering Series

Theorem 4.31: Rearrangement for Absolutely Convergent Series

Let $\sum_{n=0}^{\infty} a_n$ be an absolutely convergent series of real numbers. Let $\varphi : \mathbb{N} \to \mathbb{N}$ be a bijection. Then the series $\sum_{n=0}^{\infty} a_{\varphi(n)}$ is absolutely convergent, and

$$\sum_{n=0}^{\infty} a_n = \sum_{n=0}^{\infty} a_{\varphi(n)}. \tag{4.4}$$

Proof. Let $\varphi : \mathbb{N} \to \mathbb{N}$ be a bijection, and fix $\varepsilon > 0$. By the convergence of $\sum_{n=0}^{\infty} |a_n|$, there exists $N \in \mathbb{N}$ such that

$$\sum_{k=N+1}^{\infty} |a_k| < \frac{\varepsilon}{2}.$$

Let M be the maximum of the finite set $\{\varphi^{-1}(k) \mid 0 \le k \le N\}$. Equivalently, $M \in \mathbb{N}$ is the
19 smallest number such that

$$\{a_0, \ldots, a_N\} \subset \{a_{\varphi(0)}, \ldots, a_{\varphi(M)}\}.$$

Then, since $\{a_0, \ldots, a_N\} \subset \{a_{\varphi(0)}, \ldots, a_{\varphi(n)}\}$ for $n \ge M$,

$$\sum_{\ell=0}^{n} a_{\varphi(\ell)} - \sum_{k=0}^{N} a_k = \sum_{\substack{0 \le \ell \le n,\\ \varphi(\ell) > N}} a_{\varphi(\ell)}.$$

This implies that, for $n \ge M$,

$$\begin{aligned}\left|\sum_{\ell=0}^{n} a_{\varphi(\ell)} - \sum_{k=0}^{\infty} a_k\right| &= \left|\sum_{\ell=0}^{n} a_{\varphi(\ell)} - \sum_{k=0}^{N} a_k - \sum_{k=N+1}^{\infty} a_k\right| = \left|\sum_{\substack{0 \le \ell \le n,\\ \varphi(\ell) > N}} a_{\varphi(\ell)} - \sum_{k=N+1}^{\infty} a_k\right| \\ &\le \left|\sum_{\substack{0 \le \ell \le n,\\ \varphi(\ell) > N}} a_{\varphi(\ell)}\right| + \left|\sum_{k=N+1}^{\infty} a_k\right| \le \sum_{\substack{0 \le \ell \le n,\\ \varphi(\ell) > N}} |a_{\varphi(\ell)}| + \sum_{k=N+1}^{\infty} |a_k|.\end{aligned}$$

Note now that all terms of the form $a_{\varphi(\ell)}$ with $\varphi(\ell) > N$ and contained inside the infinite set $\{a_k \mid k > N\}$, therefore

$$\sum_{\substack{0\le\ell\le n,\\ \varphi(\ell)>N}} |a_{\varphi(\ell)}| \le \sum_{k=N+1}^{\infty} |a_k|.$$

Combining these two inequalities together we obtain

$$\left|\sum_{\ell=0}^{n} a_{\varphi(\ell)} - \sum_{k=0}^{\infty} a_k\right| \le 2\sum_{k=N+1}^{\infty} |a_k| < \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, this shows the validity of (4.4).

If we apply the same argument as above to the series $\sum_{n=0}^{\infty} |a_n|$ we obtain that $\sum_{n=0}^{\infty} |a_{\varphi(n)}| = \sum_{n=0}^{\infty} |a_n| < \infty$, so $\sum_{n=0}^{\infty} a_{\varphi(n)}$ is absolutely convergent. □

4.2.3 Products of Series

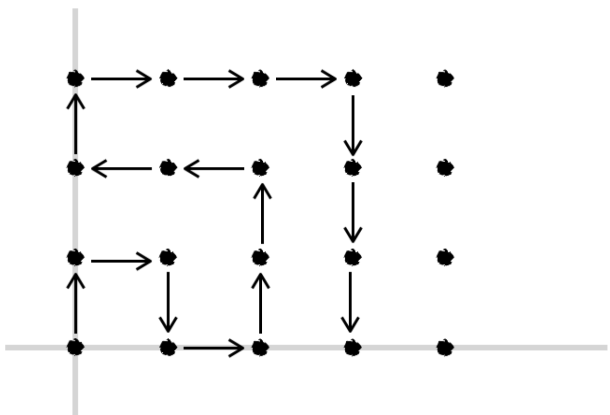
Theorem 4.32: Product Theorem

Let $\sum_{n=0}^{\infty} a_n$ and $\sum_{n=0}^{\infty} b_n$ be absolutely convergent series, and let $\alpha : \mathbb{N} \to \mathbb{N} \times \mathbb{N}$ be a
19 *bijection. Write $\alpha(n) = (\alpha_1(n), \alpha_2(n)) \in \mathbb{N} \times \mathbb{N}$ for any $n \in \mathbb{N}$. Then*

$$\left(\sum_{n=0}^{\infty} a_n\right)\left(\sum_{n=0}^{\infty} b_n\right) = \sum_{n=0}^{\infty} a_{\alpha_1(n)} b_{\alpha_2(n)}, \tag{4.5}$$

and the series on the left converges absolutely.

Proof. Consider first a bijection $\alpha : \mathbb{N} \to \mathbb{N}^2$ written as $\alpha(n) = (\alpha_1(n), \alpha_2(n))$ such that $\{\alpha(k) \mid 0 \le k < n^2\} = \{0, 1, \dots, n-1\}^2$ for all $n \in \mathbb{N}$. For example, $(\alpha(n))_{n=0}^{\infty}$ could pass through the set $\mathbb{N}^2$ as in the following figure.



Then, for each $n \in \mathbb{N}$,

$$\sum_{k=0}^{n^2-1} |a_{\alpha_1(k)}||b_{\alpha_2(k)}| = \left(\sum_{\ell=0}^{n-1} |a_\ell|\right)\left(\sum_{m=0}^{n-1} |b_m|\right).$$

Since the right-hand side is bounded by $\left(\sum_{\ell=0}^{\infty} |a_\ell|\right)\left(\sum_{m=0}^{\infty} |b_m|\right)$, it follows that

$$\sup_{n\in\mathbb{N}} \sum_{k=0}^{n^2-1} |a_{\alpha_1(k)}||b_{\alpha_1(k)}| \le \left(\sum_{\ell=0}^{\infty} |a_\ell|\right)\left(\sum_{m=0}^{\infty} |b_m|\right) < \infty,$$

so the series $\sum_{k=1}^{\infty} a_{\alpha_1(k)} b_{\alpha_2(k)}$ is absolutely convergent, and in particular converges.

Considering now the identity

$$\sum_{k=0}^{n^2-1} a_{\alpha_1(k)} b_{\alpha_2(k)} = \left(\sum_{\ell=0}^{n-1} a_\ell\right)\left(\sum_{m=0}^{n-1} b_m\right)$$

and taking the limit as $n \to \infty$, recalling Proposition 2.96(2) we obtain

$$\lim_{n\to\infty} \sum_{k=0}^{n^2-1} a_{\alpha_1(k)} b_{\alpha_2(k)} = \left(\lim_{n\to\infty} \sum_{\ell=0}^{n-1} a_\ell\right)\left(\lim_{n\to\infty} \sum_{m=0}^{n-1} b_m\right).$$

19

This proves the validity of (4.5) for the specific bijection α constructed above.

For an arbitrary bijection $\beta : \mathbb{N} \to \mathbb{N}^2$, consider the bijection $\varphi = \alpha^{-1} \circ \beta : \mathbb{N} \to \mathbb{N}$ with α as above, so that $\beta = \alpha \circ \varphi$. Then, if we write $\beta(n) = (\beta_1(n), \beta_2(n)) = (\alpha_1(\varphi(n)), \alpha_2(\varphi(n)))$, the Rearrangement Theorem 4.31 and the validity of (4.5) for α imply that

$$\sum_{n=0}^{\infty} a_{\beta_1(n)} b_{\beta_2(n)} = \sum_{n=0}^{\infty} a_{\alpha_1(\varphi(n))} b_{\alpha_2(\varphi(n))} = \sum_{n=0}^{\infty} a_{\alpha_1(n)} b_{\alpha_2(n)} = \left(\sum_{n=0}^{\infty} a_n\right)\left(\sum_{n=0}^{\infty} b_n\right).$$

□

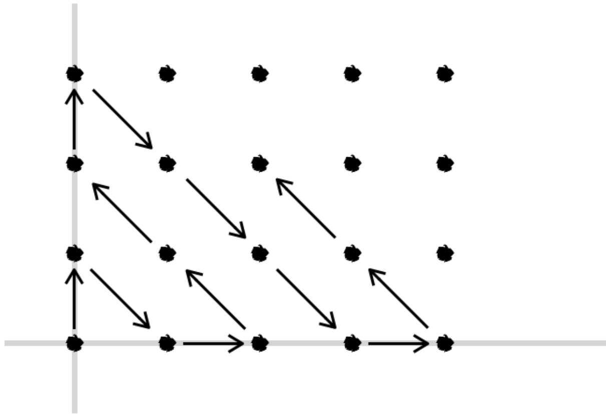
Corollary 4.33: Cauchy Product

If $\sum_{n=0}^{\infty} a_n$ and $\sum_{n=0}^{\infty} b_n$ are absolutely convergent series, then

$$\left(\sum_{n=0}^{\infty} a_n\right)\left(\sum_{n=0}^{\infty} b_n\right) = \sum_{n=0}^{\infty}\left(\sum_{k=0}^{n} a_{n-k} b_k\right),$$

where the series $\sum_{n=0}^{\infty}\left(\sum_{k=0}^{n} a_{n-k} b_k\right)$ is absolutely convergent.

Proof. Consider the bijection $\alpha : \mathbb{N} \to \mathbb{N} \times \mathbb{N}$, $\alpha(n) = (\alpha_1(n), \alpha_2(n))$, as in the picture below.



Then it follows from Theorem 4.32 that

$$\left(\sum_{n=0}^{\infty} a_n\right)\left(\sum_{n=0}^{\infty} b_n\right) = \sum_{n=0}^{\infty} a_{\alpha_1(n)} b_{\alpha_2(n)}.$$

Now, if we write explicitly the terms appearing in the sum and we group them in blocks of length $1, 2, 3, 4, \ldots$ (geometrically, this corresponds to grouping terms that belong to the same diagonal in the figure above), we see that

19

$$\begin{aligned}\sum_{n=0}^{\infty} a_{\alpha_1(n)} b_{\alpha_2(n)} &= a_0b_0 + (a_0b_1 + a_1b_0) + (a_2b_0 + a_1b_1 + a_0b_2) \\ &\quad + (a_0b_3 + a_1b_2 + a_2b_1 + a_3b_0) + \ldots \\ &= \sum_{n=0}^{\infty}\left(\sum_{k=0}^{n} a_{n-k}b_k\right).\end{aligned}$$

Finally, the absolute convergence follows from the triangle inequality and Theorem 4.32:

$$\sum_{n=0}^{\infty}\left|\sum_{k=0}^{n} a_{n-k}b_k\right| \le \sum_{n=0}^{\infty}\sum_{k=0}^{n} |a_{n-k}b_k| = \sum_{n=0}^{\infty} |a_{\alpha_1(n)}||b_{\alpha_2(n)}| < \infty.$$ □

EXAMPLE 4.34. — Let $q \in \mathbb{R}$ be such that $|q| < 1$. Then $\sum_{n=0}^{\infty} q^n$ converges absolutely according to Example 4.4. If we apply the Cauchy product to this series with itself, we get

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

In this way, we obtain an explicit formula for $\sum_{n=0}^{\infty} nq^n$:

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

4.3 Series of Complex Numbers

To study series in $\mathbb{C}$, it is often sufficient to consider the corresponding series of real and imaginary parts in $\mathbb{R}$.

Definition 4.35: Series of Complex Numbers

Let $(z_n)_{n=0}^{\infty} = (x_n + iy_n)_{n=0}^{\infty}$ be a sequence of complex numbers and let $Z = A + iB \in \mathbb{C}$. The series $\sum_{n=0}^{\infty} z_n$ is convergent with limit Z if the two series of real numbers $\sum_{n=0}^{\infty} x_n$ and $\sum_{n=0}^{\infty} y_n$ are convergent, with limits A and B, respectively.
19 We say that $\sum_{n=0}^{\infty} z_n$ converges absolutely if the series $\sum_{n=0}^{\infty} |z_n|$ converges.

Remark 4.36. — Let $(z_n)_{n=0}^{\infty} = (x_n + iy_n)_{n=0}^{\infty}$ be a sequence of complex numbers and assume that the series $\sum_{n=0}^{\infty} |z_n|$ converges. Since

$$0 \le |x_n| \le |z_n|, \qquad 0 \le |y_n| \le |z_n| \qquad \forall\, n \in \mathbb{N},$$

the Majorant Criterion in Corollary 4.10 implies that the series $\sum_{n=0}^{\infty} |x_n|$ and $\sum_{n=0}^{\infty} |y_n|$ converge, i.e., the series of the real and imaginary part are absolutely convergent.

4.4 Power Series

Our next goal is to investigate power series. These are series where the elements are powers of the variable $x \in \mathbb{R}$ (or $z \in \mathbb{C}$, if one wants to consider complex power series) multiplied by a coefficient.

4.4.1 Radius of Convergence

Definition 4.37: Power Series

A **power series** with real coefficients is a series of the form

$$\sum_{n=0}^{\infty} a_n x^n,$$

where $(a_n)_{n=0}^{\infty}$ is a sequence in $\mathbb{R}$, and $x \in \mathbb{R}$. Here, x is called **variable**, and the element $a_n \in \mathbb{R}$ is called the **coefficient** of x^n.

Addition and multiplication of power series are given by

19

$$\begin{aligned} \sum_{n=0}^{\infty} a_n x^n + \sum_{n=0}^{\infty} b_n x^n &= \sum_{n=0}^{\infty} (a_n + b_n) x^n, \\ \left(\sum_{n=0}^{\infty} a_n x^n\right)\left(\sum_{n=0}^{\infty} b_n x^n\right) &= \sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} a_{n-k} b_k\right) x^n, \end{aligned}$$

where the product formula follows from Corollary 4.33:

$$\left(\sum_{n=0}^{\infty} a_n x^n\right)\left(\sum_{n=0}^{\infty} b_n x^n\right) = \sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} a_{n-k} x^{n-k} b_k x^k\right) = \sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} a_{n-k} b_k\right) x^n.$$

4.38. — A power series is a polynomial if only finitely many of its coefficients are zero. The convergence of a power series depends heavily on the coefficients a_n, and is answered in Theorem 4.41.

Definition 4.39: Radius of Convergence

Let $\sum_{n=0}^{\infty} a_n x^n$. The **radius of convergence** of the series is the number $R \in \mathbb{R}_{\geq 0}$ or the symbol $R = \infty$, defined by

$$\rho = \limsup_{n \to \infty} \sqrt[n]{|a_n|} \qquad \text{and} \qquad R = \begin{cases} 0 & \text{if } \rho = \infty \\ \rho^{-1} & \text{if } 0 < \rho < \infty \\ \infty & \text{if } \rho = 0. \end{cases}$$

Exercise 4.40. — Find, for each $R \in [0, \infty) \cup \{\infty\}$, a power series $\sum_{n=0}^{\infty} a_n x^n$ with radius of convergence R.

Theorem 4.41: Convergence of Power Series

Let $\sum_{n=0}^{\infty} a_n x^n$ be a power series with radius of convergence $R \in (0, \infty) \cup \{\infty\}$. The series $\sum_{n=0}^{\infty} a_n x^n$ converges absolutely for all $x \in \mathbb{R}$ with $|x| < R$, and does not converge for all $x \in \mathbb{R}$ with $|x| > R$. In particular, for $x \in (-R, R)$, we can define the function $f(x) = \sum_{n=0}^{\infty} a_n x^n$.

Proof. Let $x \in \mathbb{R}$, and write $\rho = \limsup_{n\to\infty} \sqrt[n]{|a_n|}$ as in Definition 4.39. It holds

$$\limsup_{n\to\infty} \sqrt[n]{|a_n x^n|} = \limsup_{n\to\infty} \sqrt[n]{|a_n|}|x| = \rho|x|$$

According to the root criterion applied to the series $\sum_{n=0}^{\infty} b_n$ with $b_n = a_n x^n$, the series converges absolutely for all $x \in \mathbb{R}$ with $\rho|x| < 1$, and does not converge if $\rho|x| > 1$ (in particular, if $\rho = 0$, then the series converges absolutely for all $x \in \mathbb{R}$). □

Theorem 4.42: Continuity of Power Series

Let $\sum_{n=0}^{\infty} a_n x^n$ be a power series with radius of convergence $R \in (0, \infty) \cup \{\infty\}$, and define the polynomials $f_n(x) = \sum_{k=0}^{n} a_k x^k$. For any $r < R$, the sequence of polynomials
19 *$(f_n)_{n=0}^{\infty}$ converges uniformly to f on $(-r, r)$. In particular, the power series defines a continuous function $f : (-R, R) \to \mathbb{R}$.*

Proof. To prove the result, we note that Theorem 4.41 applied with $x = r$ implies that $\sum_{n=0}^{\infty} |a_n| r^n < \infty$ holds. Therefore, for every $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that $\sum_{n=N+1}^{\infty} |a_n| r^n < \varepsilon$. Thus, for all $x \in (-r, r)$ and $n \geq N$,

$$|f_n(x) - f(x)| = \left|\sum_{k=0}^{n} a_k x^k - \sum_{k=0}^{\infty} a_k x^k\right| = \left|\sum_{k=n+1}^{\infty} a_k x^k\right| \leq \sum_{k=N+1}^{\infty} |a_k||x|^k \leq \sum_{k=N}^{\infty} |a_k| r^k < \varepsilon.$$

This proves the uniform convergence inside $(-r, r)$ of the sequence of continuous functions $(f_n)_{n=0}^{\infty}$ to f so, by Theorem 3.102, f is continuous inside $(-r, r)$. Since $r < R$ is arbitrary, $f : (-R, R) \to \mathbb{R}$ is continuous. □

Example 4.43. — In general, it is not true that the partial sums $f_n(x) = \sum_{k=0}^{n} a_k x^k$ tend uniformly to the function $f(x) = \sum_{k=0}^{\infty} a_k x^k$ in the whole open interval $(-R, R)$. We illustrate this with the geometric series.

The radius of convergence of the series $\sum_{n=0}^{\infty} x^n$ is $R = 1$, and the function defined by this power series is equal to $f(x) = \frac{1}{1-x}$ on $(-1, 1)$ (see Example 4.4). In particular, if the convergence of the sequence of partial sums on $(-1, 1)$ were uniform, then applying the notion of uniform convergence with $\varepsilon = 1$ we would find $N \in \mathbb{N}$ such that, for all $n \geq N$ and

$x \in (-1, 1)$, the estimate

$$\left| \sum_{k=0}^{n} x^k - \frac{1}{1-x} \right| < 1$$

holds. Choosing $n = N$, thanks to the triangle inequality we obtain

$$\left| \frac{1}{1-x} \right| < 1 + \left| \sum_{k=0}^{N} x^k \right| \leq 1 + \sum_{k=0}^{N} |x|^k \leq 2 + N \qquad \forall\, x \in (-1, 1).$$

However, this is a contradiction, since $\lim\limits_{x \to 1^-} \frac{1}{1-x} = \infty$.

19 Exercise 4.44. — Calculate the radius of convergence R of the power series

$$\sum_{n=1}^{\infty} \frac{(\sqrt{n^2+n} - \sqrt{n^2+1})^n}{n^2} x^n,$$

and show the convergence of the power series also at $x = -R$ and $x = R$.

Exercise 4.45. — Let $\sum_{n=0}^{\infty} a_n x^n$ be a power series with $a_n \neq 0$ for all $n \in \mathbb{N}$, and assume that the limit $\lim_{n \to \infty} \frac{|a_n|}{|a_{n+1}|}$ exists. Then, the radius of convergence R is given by $R = \lim_{n \to \infty} \frac{|a_n|}{|a_{n+1}|}$.
Hint: Combine Corollary 4.29 and Proposition 2.96(3).

Proposition 4.46: Radius of Convergence of Sum and Product

Let $R \geq 0$, and let $\sum_{n=0}^{\infty} a_n x^n$ and $\sum_{n=0}^{\infty} b_n x^n$ be power series with radius of convergence at least R. Then their sum and product also have radii of convergence at least R.

Proof. Due to the linearity of the limit and Corollary 4.33, the absolute convergence of the
20 power series $\sum_{n=0}^{\infty} a_n x^n$ and $\sum_{n=0}^{\infty} b_n x^n$ for $|x| < R$ implies that both power series

$$\sum_{n=0}^{\infty} (a_n + b_n) x^n \quad \text{and} \quad \left(\sum_{n=0}^{\infty} a_n x^n \right) \left(\sum_{n=0}^{\infty} b_n x^n \right) = \sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} a_{n-k} b_k \right) x^n$$

absolutely converge for $|x| < R$. Since a power series does not converge for $|x|$ larger than its radius of converges, this implies that both power series have radii of convergence at least R. $\square$

Example 4.47. — If $\sum_{n=0}^{\infty} a_n x^n$ has at least radius of convergence 1, then

$$\frac{1}{1-x}\sum_{n=0}^{\infty} a_n x^n = \sum_{n=0}^{\infty}(a_0 + \ldots + a_n)x^n \qquad \forall\, x \in (-1,1) \tag{4.6}$$

Indeed, the power series $\sum_{n=0}^{\infty} x^n$ has radius of convergence 1, and for $x \in (-1,1)$ we have $\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}$, so (4.6) follows from Proposition 4.46.

Exercise 4.48. — Calculate $\sum_{n=1}^{\infty} n2^{-n}$.

4.4.2 Complex Power Series

Analogously to the real case, one can consider series with complex coefficients and a complex variable $z \in \mathbb{C}$. This generalization will be useful later.

Definition 4.49: Complex Power Series

A **complex power series** with complex coefficients is a series of the form

$$\sum_{n=0}^{\infty} a_n z^n,$$

20 where $(a_n)_{n=0}^{\infty}$ is a sequence in $\mathbb{C}$, and $z \in \mathbb{C}$.

Addition and multiplication of power series are the same as in the real case, namely

$$\sum_{n=0}^{\infty} a_n z^n + \sum_{n=0}^{\infty} b_n z^n = \sum_{n=0}^{\infty}(a_n + b_n)z^n, \tag{4.7}$$

$$\left(\sum_{n=0}^{\infty} a_n z^n\right)\left(\sum_{n=0}^{\infty} b_n z^n\right) = \sum_{n=0}^{\infty}\left(\sum_{k=0}^{n} a_{n-k}b_k\right)z^n. \tag{4.8}$$

Also, the radius of convergence is still defined as in Definition 4.39. Several results that are true for real power series, hold also in the complex case.

We note that, with the very same proof as in the real case, the analogue of Theorem 4.41 holds:

Theorem 4.50: Convergence of Power Series

Let $\sum_{n=0}^{\infty} a_n z^n$ be a power series with radius of convergence $R \in (0,\infty) \cup \{\infty\}$. The series $\sum_{n=0}^{\infty} a_n z^n$ converges absolutely for all $z \in \mathbb{C}$ with $|z| < R$, and does not converge for all $z \in \mathbb{C}$ with $|z| > R$. In particular, for $z \in B(0,R)$, we can define the function $f(z) = \sum_{n=0}^{\infty} a_n z^n$.

Continuity of Complex Power Series (Extra material)

Also in the complex case, the analog of Theorem 4.42 holds. To prove that, one defines continuous functions exactly as in Definition 3.9, and uniform convergence as in Definition 3.99 (with the only warning that $|\cdot|$ now denotes the absolute value on $\mathbb{C}$, see Definition 2.46). In this way, one can prove that Theorem 3.102 also holds in the complex case, namely, uniform limit of continuous functions is continuous (in the courses of Analysis II or Complex Analysis,
20 this will be proved in full detail), and we get the following:

Theorem 4.51: Continuity of Power Series

Let $\sum_{n=0}^{\infty} a_n z^n$ be a power series with radius of convergence $R \in (0, \infty) \cup \{\infty\}$, and define the (complex) polynomials $f_n(z) = \sum_{k=0}^{n} a_k z^k$. For any $r < R$, the sequence of polynomials $(f_n)_{n=0}^{\infty}$ converges uniformly to f on $B(0, r)$. In particular, the power series defines a continuous function $f : B(0, R) \to \mathbb{C}$.

4.5 Example: Exponential and Trigonometric Functions

4.5.1 The Exponential Map as Power Series

In section 3.4, we have seen the real exponential mapping and shown its main properties. We now show that we can alternatively define the exponential map by the **exponential series**

$$\exp(x) = \sum_{k=0}^{\infty} \frac{1}{k!} x^k, \tag{4.9}$$

where

$$0! = 1, \qquad k! = 1 \cdot 2 \cdot \ldots \cdot k.$$

It follows directly from the quotient criterion (see Exercise 4.45) that this series has infinite radius of convergence, so in particular the right-hand side of (4.9) is well-defined for all $x \in \mathbb{R}$. Alternatively, one can note that, given $N \in \mathbb{N}$

$$n! \geq \underbrace{n \cdot (n-1) \cdot \ldots \cdot N}_{n-N+1 \text{ terms}} \geq N^{n-N+1},$$

hence

20

$$\limsup_{n\to\infty} \sqrt[n]{|a_n|} \leq \limsup_{n\to\infty} \frac{1}{N^{\frac{n-N+1}{n}}} = \frac{1}{N}.$$

Since $N \in \mathbb{N}$ can be chosen arbitrarily large, this implies that $\limsup_{n\to\infty} \sqrt[n]{|a_n|} = 0$ and, therefore, $R = \infty$.

The representation of the exponential mapping as a power series is, in many aspects, more flexible than the representation as a limit. In addition, as we shall see, its complex version is related to sine and cosine in a very practical way.

We first show that our new definition of exponential coincides with the one in Definition 3.52.

Proposition 4.52: Exponential Map as Power Series

For every $x \in \mathbb{R}$,

$$\sum_{k=0}^{\infty} \frac{1}{k!} x^k = \lim_{n\to\infty} \left(1 + \frac{x}{n}\right)^n.$$

Proof. (Extra material) We first observe that, for any $n \geq 0$, the identity

$$\left(1 + \frac{x}{n}\right)^n = \sum_{k=0}^{n} \frac{n!}{k!(n-k!)} \frac{x^k}{n^k} = \sum_{k=0}^{n} \frac{x^k}{k!} \frac{1}{n^k} \prod_{l=0}^{k-1} (n-l) = \sum_{k=0}^{n} \frac{x^k}{k!} \prod_{l=0}^{k-1} \left(1 - \frac{l}{n}\right)$$

hold. Now, given $x \in \mathbb{R}$ and $\varepsilon > 0$, since $\sum_{k=0}^{\infty} \frac{1}{k!}|x|^k < \infty$ we can find $N \in \mathbb{N}$ such that

$$\sum_{k=N+1}^{\infty} \frac{1}{k!}|x|^k < \varepsilon.$$

Then, for this N, we also have

$$\left|\sum_{k=0}^{N} \frac{1}{k!}x^k - \sum_{k=0}^{\infty} \frac{1}{k!}x^k\right| \le \left|\sum_{k=N+1}^{\infty} \frac{1}{k!}x^k\right| \le \sum_{k=N+1}^{\infty} \frac{1}{k!}|x|^k < \varepsilon \tag{4.10}$$

and furthermore, for $n \ge N$,

$$\begin{aligned}
&\left|\sum_{k=0}^{N} \frac{1}{k!}x^k - \sum_{k=0}^{n} \frac{1}{k!}x^k \prod_{\ell=0}^{k-1}\left(1-\frac{\ell}{n}\right)\right| \\
&\qquad \le \sum_{k=0}^{N} \frac{1}{k!}|x|^k\left(1-\prod_{\ell=0}^{k-1}\left(1-\frac{\ell}{n}\right)\right) + \sum_{k=N+1}^{n} \frac{1}{k!}|x|^k \underbrace{\prod_{\ell=0}^{k-1}\left(1-\frac{\ell}{n}\right)}_{\le 1} \\
&\qquad \le \sum_{k=0}^{N} \frac{1}{k!}|x|^k\left(1-\prod_{\ell=0}^{k-1}\left(1-\frac{\ell}{n}\right)\right) + \sum_{k=N+1}^{\infty} \frac{1}{k!}|x|^k.
\end{aligned}$$

20 This proves that

$$\left|\sum_{k=0}^{N} \frac{1}{k!}x^k - \left(1+\frac{x}{n}\right)^n\right| \le \sum_{k=0}^{N} \frac{1}{k!}|x|^k\left(1-\prod_{\ell=0}^{k-1}\left(1-\frac{\ell}{n}\right)\right) + \sum_{k=N+1}^{\infty} \frac{1}{k!}|x|^k.$$

Since

$$\lim_{n\to\infty}\left(1-\prod_{\ell=0}^{k-1}\left(1-\frac{\ell}{n}\right)\right) = 0 \qquad \forall\, k \in \{0,\ldots,N\},$$

letting $n \to \infty$ in the latter formula yields

$$\left|\sum_{k=0}^{N} \frac{1}{k!}x^k - \lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n\right| \le \sum_{k=N+1}^{\infty} \frac{1}{k!}|x|^k < \varepsilon.$$

Combining this estimate with (4.10) proves that

$$\begin{aligned}
\left|\sum_{k=0}^{\infty} \frac{1}{k!}x^k - \lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n\right| &\le \left|\sum_{k=N+1}^{\infty} \frac{1}{k!}x^k\right| + \left|\sum_{k=0}^{N} \frac{1}{k!}x^k - \lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n\right| \\
&\le \sum_{k=N+1}^{\infty} \frac{1}{k!}|x|^k + \left|\sum_{k=0}^{N} \frac{1}{k!}x^k - \lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n\right| < 2\varepsilon.
\end{aligned}$$

Since $\varepsilon > 0$ is arbitrary, (4.9) follows. □

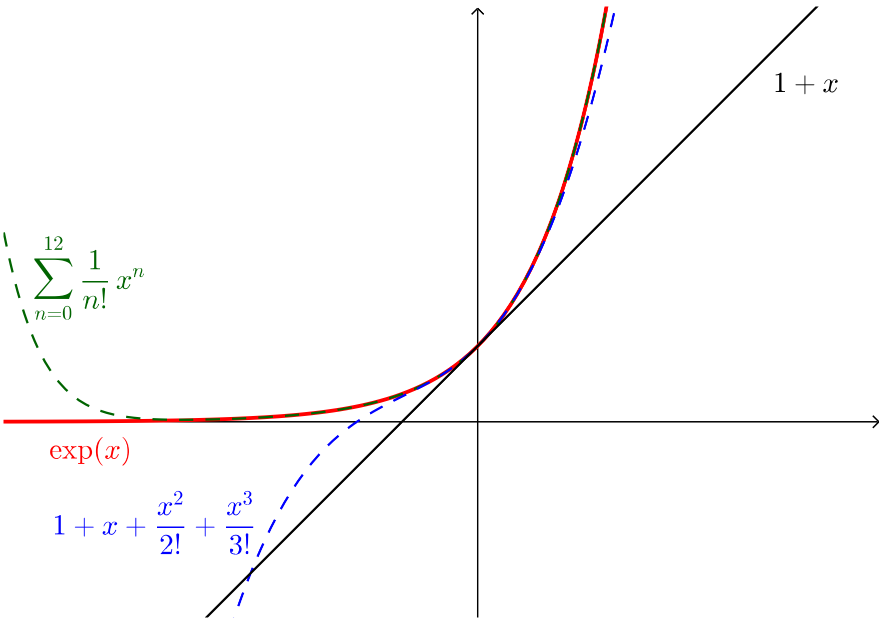


Figure 4.1: The graph of the exponential mapping, and the graphs of some partial sums of the exponential series.

Definition 4.53: The Complex Exponential Map

20 The **complex exponential map** is the function $\exp : \mathbb{C} \to \mathbb{C}$ given by

$$\exp(z) = \sum_{k=0}^{\infty} \frac{1}{n!} z^n$$

for all $z \in \mathbb{C}$.

For a positive real number $a \in \mathbb{R}_{>0}$ and $z \in \mathbb{C}$ we write $a^z = \exp(z \log(a))$, and in particular also $e^z = \exp(z)$ for all $z \in \mathbb{C}$.

Before stating the main properties of the exponential, we recall the binomial formula: given $z, w \in \mathbb{C}$ and $n \in \mathbb{N}$,

$$(z+w)^n = \sum_{k=0}^{n} \binom{n}{k} z^k w^{n-k}, \qquad \binom{n}{k} = \frac{n!}{k!(n-k)!}. \tag{4.11}$$

Theorem 4.54: Properties of the Complex Exponential

The complex exponential mapping $\exp : \mathbb{C} \to \mathbb{C}$ *is continuous. Furthermore*

$$e^{z+w} = e^z e^w \qquad \textit{and} \qquad |e^z| = e^{\mathrm{Re}(z)} \tag{4.12}$$

for all $z, w \in \mathbb{C}$. *In particular,* $|e^{ix}| = 1$ *holds for all* $x \in \mathbb{R}$.

Proof. As noted before, the radius of convergence of the series $\sum_{n=0}^{\infty} \frac{1}{n!} z^n$ is infinite. Therefore, by Theorems 4.50 and 4.51, $\exp : \mathbb{C} \to \mathbb{C}$ is a continuous function.

Given $z, w \in \mathbb{C}$, it follows from (4.7) and the binomial formula (4.11) that

$$\begin{aligned} e^z e^w &= \left(\sum_{n=0}^{\infty} \frac{1}{n!} z^n \right) \left(\sum_{n=0}^{\infty} \frac{1}{n!} w^n \right) = \sum_{n=0}^{\infty} \sum_{k=0}^{n} \frac{1}{k!} \frac{1}{(n-k)!} z^k w^{n-k} \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{k=0}^{n} \binom{n}{k} z^k w^{n-k} = \sum_{n=0}^{\infty} \frac{1}{n!} (z+w)^n = e^{z+w}. \end{aligned}$$

It remains to prove the formula for the absolute value. Since the conjugation $\mathbb{C} \to \mathbb{C}$ is a continuous function, the following holds:

$$\overline{e^z} = \overline{\lim_{n\to\infty} \sum_{k=0}^{n} \frac{1}{k!} z^k} = \lim_{n\to\infty} \overline{\sum_{k=0}^{n} \frac{1}{k!} z^k} = \lim_{n\to\infty} \sum_{k=0}^{n} \frac{1}{k!} \overline{z}^k = e^{\overline{z}},$$

where the equality $\overline{z^k} = \overline{z}^k$ follows from Lemma 2.43(3). Recalling that for a complex number w it holds $|w|^2 = w\overline{w}$ and $w + \overline{w} = 2\operatorname{Re}(w)$, we get

$$|e^z|^2 = e^z \overline{e^z} = e^z e^{\overline{z}} = e^{z+\overline{z}} = e^{2\operatorname{Re}(z)} = (e^{\operatorname{Re}(z)})^2,$$

therefore $|e^z| = e^{\operatorname{Re}(z)}$. In particular $|e^{ix}| = e^0 = 1$. □

20

Exercise 4.55. — Show that $e^z = \lim_{n\to\infty} \left(1 + \frac{z}{n}\right)^n$ holds for all $z \in \mathbb{C}$.

4.5.2 Sine and Cosine

4.56. — Given $x \in \mathbb{R}$, we split the power series of e^{ix} into its odd and even terms:

$$e^{ix} = \sum_{n=0}^{\infty} \frac{i^n}{n!} x^n = \sum_{n=0}^{\infty} \frac{i^{2n}}{(2n)!} x^{2n} + \sum_{n=0}^{\infty} \frac{i^{2n+1}}{(2n+1)!} x^{2n+1}.$$

Noticing that $i^{2n} = (-1)^n$, it follows that

$$e^{ix} = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n} + i \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1}.$$

Starting from this formula, we define the **sine function** and the **cosine function** as

$$\sin(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1} \qquad \text{and} \qquad \cos(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n}, \tag{4.13}$$

so that the identity

$$e^{ix} = \cos(x) + i \sin(x)$$

holds.

As for the exponential, the radius of convergence of these series is infinite, and by Theorems 4.50 and 4.51 they define continuous functions. Since $(-x)^{2n+1} = -x^{2n+1}$ and $(-x)^{2n} = x^{2n}$ for every $n \in \mathbb{N}$, it follows directly from the definition as power series that the sine function is **odd**, i.e., $\sin(-x) = -\sin(x)$, and the cosine function is **even**, i.e., $\cos(-x) = \cos(x)$ for all $x \in \mathbb{R}$.

Theorem 4.57: From Complex Exponential to Sine and Cosine

For all $x \in \mathbb{R}$, the following relations between exponential, sine, and cosine functions hold:

$$e^{ix} = \cos(x) + i\sin(x), \quad \sin(x) = \frac{e^{ix} - e^{-ix}}{2i}, \quad \cos(x) = \frac{e^{ix} + e^{-ix}}{2}.$$

For all $x, y \in \mathbb{R}$, the trigonometric addition formulae apply:

$$\begin{aligned} \sin(x+y) &= \sin(x)\cos(y) + \cos(x)\sin(y) \\ \cos(x+y) &= \cos(x)\cos(y) - \sin(x)\sin(y). \end{aligned} \tag{4.14}$$

Proof. For $x \in \mathbb{R}$ we have

$$e^{ix} = \cos(x) + i\sin(x), \qquad e^{-ix} = \cos(-x) + i\sin(-x) = \cos(x) - i\sin(x).$$

20 If we add (respectively subtract) these two identities, we obtain the formulae for $\cos(x)$ and $\sin(x)$.

To prove the addition formulas, we multiply e^{ix} by e^{iy} and using (4.12) we get

$$\begin{aligned} \cos(x+y) + i\sin(x+y) = e^{i(x+y)} &= e^{ix}e^{iy} \\ &= (\cos(x) + i\sin(x))(\cos(y) + i\sin(y)) \\ &= \cos(x)\cos(y) - \sin(x)\sin(y) + i\big(\sin(x)\cos(y) + \sin(y)\cos(x)\big). \end{aligned}$$

From this equality, the formulae in (4.14) follow. □

In particular, in the case $x = y \in \mathbb{R}$, we obtain the **angle-doubling formulae**

$$\sin(2x) = 2\sin(x)\cos(x) \qquad \text{and} \qquad \cos(2x) = \cos(x)^2 - \sin(x)^2. \tag{4.15}$$

Recalling that $|e^{ix}|^2 = 1$, we also get the **circle equation** for sine and cosine:

$$\cos(x)^2 + \sin(x)^2 = 1 \qquad \forall\, x \in \mathbb{R}.$$

Applet 4.58 (Power Series)**.** *We consider the first partial sums of the power series defining* exp, sin *and* cos *(respectively* sinh, cosh *from the next section). By zooming in and out, you can get a feeling for the quality of the approximations of the various partial sums. In the case* *of the trigonometric functions, you can also clearly see in the picture that the power series form alternating series.*

4.5.3 The Circle Number

Theorem 4.59: Existence of π as First Positive Zero of Sine

There is exactly one number $\pi \in (0,4)$ with $\sin(\pi) = 0$. For this number it holds

$$e^{i\frac{\pi}{2}} = i, \qquad e^{i\pi} = -1, \qquad e^{i2\pi} = 1.$$

Proof. The sequence of real numbers $(\frac{x^n}{n!})_{n=2}^{\infty}$ is monotonically decreasing for all $x \in (0,2]$. Therefore, from the Leibniz criterion for alternating series (see Proposition 4.22), the following estimates hold for every $x \in (0,2]$:

$$x - \frac{x^3}{3!} \leq \sin(x) \leq x - \frac{x^3}{3!} + \frac{x^5}{5!} \qquad \text{and} \qquad 1 - \frac{x^2}{2} \leq \cos(x) \leq 1 - \frac{x^2}{2} + \frac{x^4}{24}.$$

Note that $\sin(0) = 0$ and that, for $x = 1$, we have

20 $$\sin(1) \geq 1 - \frac{1}{6} > \frac{1}{\sqrt{2}}.$$

Therefore, since the sine function is continuous, it follows from the Intermediate Value Theorem 3.29 that there exists a number $p \in (0,1)$ such that $\sin(p) = \frac{1}{\sqrt{2}}$.

Because $\sin^2(p) + \cos^2(p) = 1$ and $\cos(x) \geq 1 - \frac{1}{2}x^2 > 0$ for $x \in [0,1]$, it also follows $\cos(p) = \sqrt{1 - \sin^2(p)} = \frac{1}{\sqrt{2}}$. In other words,

$$e^{ip} = \cos(p) + i\sin(p) = \frac{1+i}{\sqrt{2}}.$$

Hence, if we set $\pi = 4p$, we see that

$$e^{i\frac{\pi}{2}} = e^{i2p} = \left(e^{ip}\right)^2 = \frac{(1+i)^2}{2} = i, \qquad e^{i\pi} = \left(e^{i\frac{\pi}{2}}\right)^2 = i^2 = -1, \qquad e^{i2\pi} = (-1)^2 = 1.$$

In particular, from the identity $\cos(\pi) + i\sin(\pi) = e^{i\pi} = -1$ we deduce that $\sin(\pi) = 0$ and $\cos(\pi) = -1$.

It remains to show the uniqueness of π as in the theorem. From the estimate

$$\sin(x) \geq x - \frac{x^3}{3!} = x\left(1 - \frac{x^2}{6}\right) > 0 \qquad \text{for } x \in (0,2]$$

it follows that the sine function has no zeros in $(0, 2]$. In particular, $\pi \in (2, 4)$. Suppose now, by contradiction, that there exists another value $s \in (2, 4)$ satisfying $\sin(s) = 0$, and define

$$r = \begin{cases} \pi - s & \text{if } 2 < s < \pi \\ s - \pi & \text{if } 2 < \pi < s. \end{cases}.$$

Then $r \in (0, 2)$ and using (4.14) we get (the sign $\pm$ below depends on whether $\pi - s$ is positive or negative)

$$\sin(r) = \pm \sin(\pi - s) = \pm\big(\underbrace{\sin(\pi)}_{=0} \cos(s) - \cos(\pi) \underbrace{\sin(s)}_{=0}\big) = \pm(0 - 0) = 0.$$

This is a contradiction since sin never vanishes on $(0, 2)$. This proves that $\pi \in (0, 4)$ is uniquely determined by the equation $\sin(\pi) = 0$. □

Corollary 4.60: Periodicity of Sine and Cosine

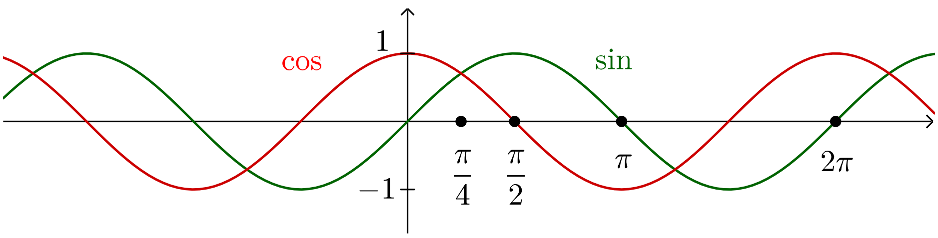
$$\begin{aligned} \sin(x + \tfrac{\pi}{2}) &= \cos(x) & \qquad \cos(x + \tfrac{\pi}{2}) &= -\sin(x) \\ \sin(x + \pi) &= -\sin(x) & \cos(x + \pi) &= -\cos(x) \\ \sin(x + 2\pi) &= \sin(x) & \cos(x + 2\pi) &= \cos(x). \end{aligned}$$

20 *Proof.* Rewriting the formulas in Theorem 4.59 in terms of sine and cosine, we see that

$$\sin(\tfrac{\pi}{2}) = 1, \quad \cos(\tfrac{\pi}{2}) = 0, \quad \sin(\pi) = 0, \quad \cos(\pi) = -1, \quad \sin(2\pi) = 0, \quad \cos(2\pi) = 1.$$

Using these identities and (4.14), the result follows. □

4.61. — From Corollary 4.60 it follows that sine and cosine are both periodic functions with period length 2π. To know the numerical value of $\sin(x)$ or $\cos(x)$ for a given real number x, it is sufficient to know the values of the sine on the interval $[0, \frac{\pi}{2}]$.



Exercise 4.62. — Show that the zeros of $\sin : \mathbb{R} \to \mathbb{R}$ are exactly the points in $\pi\mathbb{Z} \subset \mathbb{R}$, and the zeros of $\cos : \mathbb{R} \to \mathbb{R}$ are exactly the points in $\pi\mathbb{Z} + \frac{\pi}{2}$. Also, $\cos(x) = 1$ only when $x = 2n\pi$ with $n \in \mathbb{Z}$.

Exercise 4.63. — Show the formula

$$\sin(x) - \sin(y) = 2\cos\left(\frac{x+y}{2}\right)\sin\left(\frac{x-y}{2}\right)$$

for all $x, y \in \mathbb{R}$. Use this to show that $\sin : \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \to [-1, 1]$ is strictly monotonically increasing and hence bijective.

Exercise 4.64. — Show that $3.1 < \pi < 3.2$ holds. Using an electronic tool to calculate certain rational numbers may be helpful.

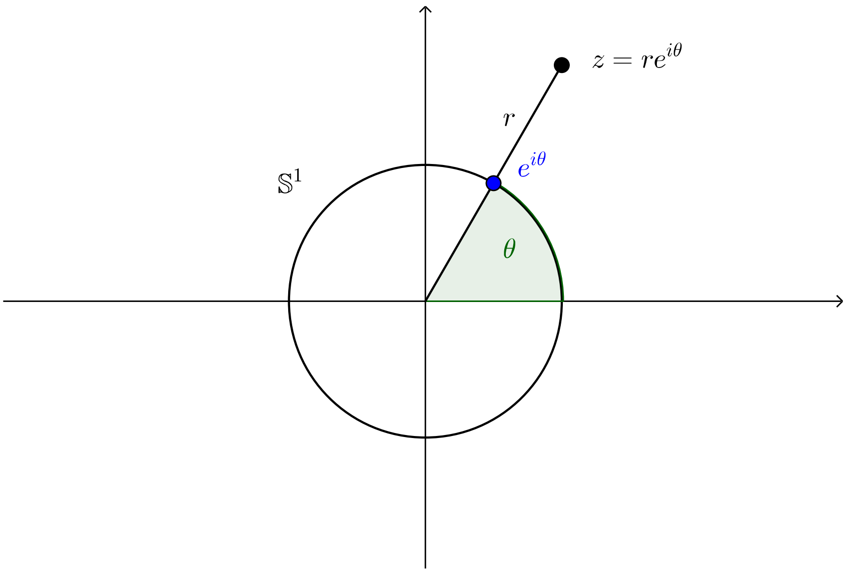
4.5.4 Polar Coordinates and Multiplication of Complex Numbers

Using the complex exponential function, we can express complex numbers in **polar coordinates**, that is, in the form

$$z = re^{i\theta} = r\cos(\theta) + ir\sin(\theta)$$

where r is the distance from the origin $0 \in \mathbb{C}$ to z, i.e., the absolute value $r = |z|$ of z, and θ is the angle enclosed between the half-lines $\mathbb{R}_{\geq 0}$ and $z\mathbb{R}_{\geq 0}$.

20



If $z \neq 0$, then the angle θ is uniquely determined, and is called the **argument** of z and written as $\theta = \arg(z)$. The set of complex numbers with absolute value one is thus

$$\mathbb{S}^1 = \{z \in \mathbb{C} \mid |z| = 1\} = \{e^{i\theta} \mid \theta \in [0, 2\pi)\}$$

and is called the **unit circle** in $\mathbb{C}$.

Proposition 4.65: Existence of Polar Coordinates

For all $z \in \mathbb{C} \setminus \{0\}$ there exist uniquely determined real numbers $r > 0$ and $\theta \in [0, 2\pi)$ with $z = re^{i\theta}$.

Proof. (Extra material) Let $r = |z|$, and consider the complex number $w = \frac{z}{r}$. Note that $|w| = \frac{|z|}{r} = 1$. We want to prove that there exists a unique $\theta \in [0, 2\pi)$ such that $w = e^{i\theta}$.

Assume first that $\mathrm{Im}(w) \geq 0$ and recall that $\mathrm{Re}(w) \in [-1, 1]$ (since $\mathrm{Re}(w)^2 + \mathrm{Im}(w)^2 = 1$). Hence, since $\cos(0) = 1$ and $\cos(\pi) = -1$, according to the Intermediate Value Theorem 3.29 there exists $\theta \in [0, \pi]$ such that $\mathrm{Re}(w) = \cos(\theta)$. Since $\mathrm{Im}(w) \geq 0$ by assumption and $\sin(\theta) \geq 0$ (since $\theta \in [0, \pi]$), this implies that

$$\sin(\theta) = \sqrt{1 - \cos^2(\theta)} = \sqrt{1 - \mathrm{Re}(w)^2} = \mathrm{Im}(w),$$

thus $w = e^{i\theta}$.

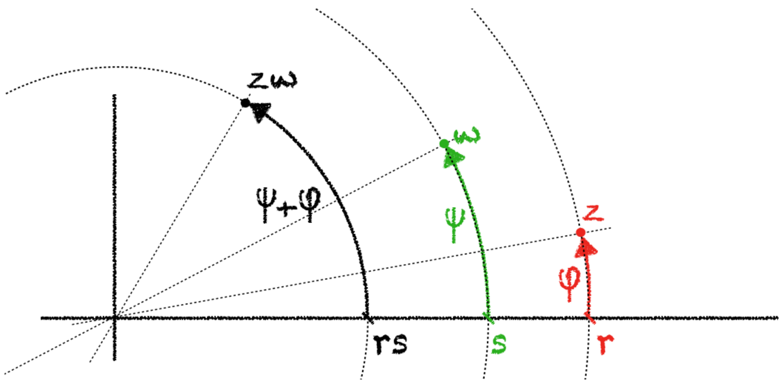
If $\mathrm{Im}(w) < 0$, then we apply the above argument to $-w$ to find $\vartheta \in (0, \pi)$ such that $-w = e^{i\vartheta}$ (note that ϑ must be different from 0 and π, since $\mathrm{Im}(e^0) = \mathrm{Im}(e^{i\pi}) = 0$). Recalling that $e^{i\pi} = -1$, it follows that $w = e^{i\pi}e^{i\vartheta} = e^{i\theta}$ with $\theta = \pi + \vartheta \in (\pi, 2\pi)$.

It remains to show the uniqueness of θ. If $\theta, \theta' \in [0, 2\pi)$ satisfy $w = e^{i\theta} = e^{i\theta'}$, then $e^{i(\theta - \theta')} = 1$, that is,

$$\sin(\theta - \theta') = 0, \qquad \cos(\theta - \theta') = 1.$$

20 Note that $\theta - \theta' \in (-2\pi, 2\pi)$. Hence, from the uniqueness of π in Theorem 4.59 and the formula $\sin(x + \pi) = -\sin(x)$ (see Corollary 4.60) it follows that $\theta - \theta' \in \{-\pi, 0, \pi\}$. But if $\theta - \theta' \in \{-\pi, \pi\}$ then $\cos(\theta - \theta') = -1$, so the only possibility is $\theta - \theta' = 0$, as desired. □

4.66. — In polar coordinates, the multiplication in $\mathbb{C}$ can be reinterpreted as follows: If $z = re^{i\varphi}$ and $w = se^{i\psi}$ are complex numbers, then $zw = rse^{i(\varphi+\psi)}$. In other words, when multiplying complex numbers, the lengths of the vectors multiply and the angles add.



Applet 4.67 (Geometric Meaning of Complex Numbers). *We can see from the polar coordinate lines drawn in the geometrical meaning of the multiplication of complex numbers and the inverses and roots of a given number.*

Exercise 4.68. — Let $w = re^{i\theta}$ be non-zero. Show that the n-th roots of w (namely, the solutions $z \in \mathbb{C}$ to the equation $z^n = w$) are given by the n numbers

$$\left\{ \sqrt[n]{r}\, e^{i\left(2\pi\alpha + \frac{\theta}{n}\right)} \,\middle|\, \alpha = 0, \tfrac{1}{n}, \tfrac{2}{n}, \ldots, \tfrac{n-1}{n} \right\}.$$

When $w = 1$, its n-th roots are given by

$$\left\{ e^{i2\pi\alpha} \,\middle|\, \alpha = 0, \tfrac{1}{n}, \tfrac{2}{n}, \ldots, \tfrac{n-1}{n} \right\}$$

and are called the n-th **roots of unity**.

Exercise 4.69. — For all natural numbers $n \geq 2$, show the identity $\sum_{k=0}^{n-1} e^{i2\pi\frac{k}{n}} = 0$.

4.5.5 The Complex Logarithm

We have defined the real logarithm as the inverse mapping of the bijective mapping $\exp :$
20 $\mathbb{R} \to \mathbb{R}_{>0}$. We now would like to define the logarithm for complex numbers. Unfortunately, however, there is a fundamental problem here: the complex exponential mapping $\exp : \mathbb{C} \to \mathbb{C}$ is not injective, since, for example, $\exp(ix) = 1$ holds for all $x = 2n\pi$ with $n \in \mathbb{Z}$. For this reason, we need to restrict the exponential mapping to a suitable subset D of $\mathbb{C}$ to ensure that the restricted mapping $\exp|_D : D \to \mathbb{C}^\times$ is bijective. This can be achieved by many different subsets. We refer to the lecture on Complex Analysis for concrete choices.

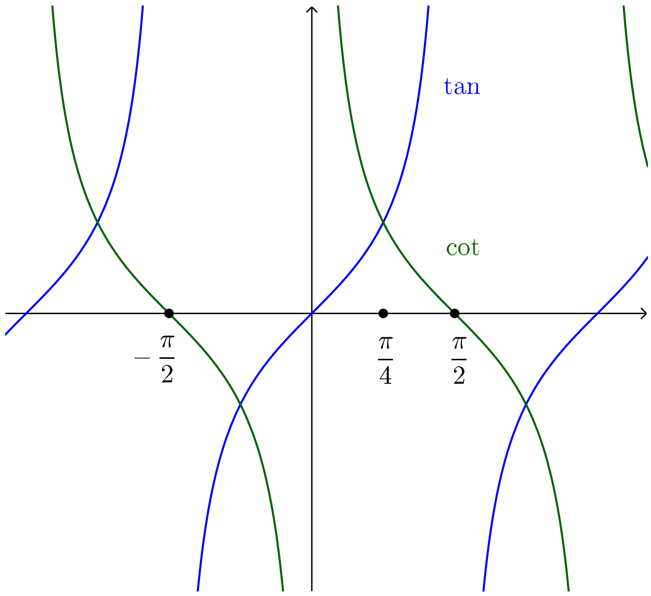
4.5.6 Other Trigonometric Functions

In addition to the exponential function, the sine and the cosine, there are other related functions called trigonometric functions.

4.70. — The **tangent function** and the **cotangent function** are given by.

$$\tan(x) = \frac{\sin(x)}{\cos(x)} \qquad \text{and} \qquad \cot(x) = \frac{\cos(x)}{\sin(x)}$$

defined for all $x \in \mathbb{R}$ with $\cos(x) \neq 0$, respectively with $\sin(x) \neq 0$.



Exercise 4.71. — Show that, for $x, y \in \mathbb{C}$, the addition formula

$$\tan(x+y) = \frac{\tan(x)+\tan(y)}{1-\tan(x)\tan(y)}$$

holds where defined. Find and prove an analogous addition formula for the cotangent.

4.72. — The **hyperbolic sine** and the **hyperbolic cosine** are the functions given by the power series

$$\sinh(x) = \sum_{k=0}^{\infty} \frac{1}{(2k+1)!} x^{2k+1} \qquad \text{and} \qquad \cosh(x) = \sum_{k=0}^{\infty} \frac{1}{(2k)!} x^{2k}.$$

It holds

$$\sinh(x) = \frac{e^x - e^{-x}}{2} \qquad \text{and} \qquad \cosh(x) = \frac{e^x + e^{-x}}{2},$$

and so $e^x = \cosh(x) + \sinh(x)$ for all $x \in \mathbb{R}$. The **hyperbolic tangent** and the **hyperbolic cotangent** are given by

$$\tanh(x) = \frac{\sinh(x)}{\cosh(x)} = \frac{e^x - e^{-x}}{e^x + e^{-x}} \qquad \text{and} \qquad \coth(x) = \frac{\cosh(x)}{\sinh(x)} = \frac{e^x + e^{-x}}{e^x - e^{-x}},$$

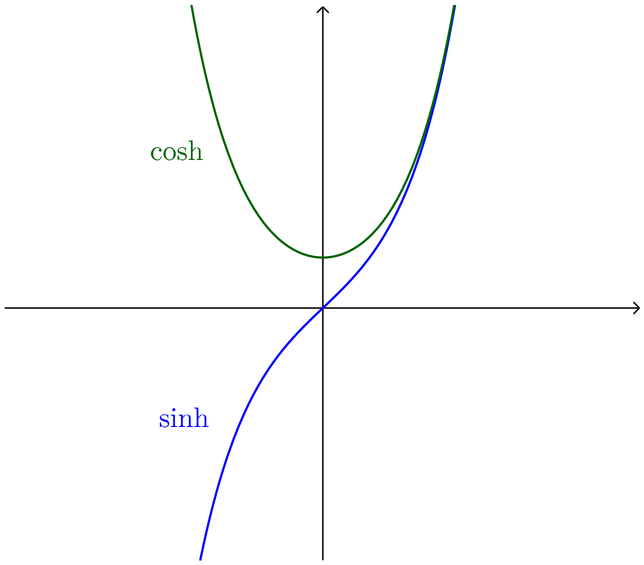
and the hyperbolic cotangent is defined for all $x \in \mathbb{R} \setminus \{0\}$ (since $\sinh(x) \neq 0$ for $x \neq 0$). The functions sinh and tanh are odd, and cosh is even. The addition formulae

$$\begin{aligned} \sinh(x+y) &= \sinh(x)\cosh(y) + \cosh(x)\sinh(y), \\ \cosh(x+y) &= \cosh(x)\cosh(y) + \sinh(x)\sinh(t) \end{aligned}$$

apply for all $x, y \in \mathbb{R}$, as well as the hyperbolic equation

$$\cosh^2(x) - \sinh^2(x) = 1 \qquad \forall\, x \in \mathbb{R}.$$

20



Exercise 4.73. — Starting from the definitions of hyperbolic sine and hyperbolic cosine, prove the above formulae.

Chapter 5

Differential Calculus

We deal with differential calculus in one variable. This is of fundamental importance for understanding functions on $\mathbb{R}$.

5.1 The Derivative

5.1.1 Definition and Geometrical Interpretation

In this section $D \subseteq \mathbb{R}$ will denote a non-empty set with no isolated points, that is, every element $x \in D$ is an accumulation point for $D \setminus \{x\}$. The typical example of such a subset is an interval that is not empty and does not consist of only one point.

Definition 5.1: Derivative

Let $f : D \to \mathbb{R}$ be a function and $x_0 \in D$. We say that f is **differentiable** at x_0 if the limit

$$f'(x_0) = \lim_{\substack{x \to x_0 \\ x \neq x_0}} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{\substack{h \to 0 \\ h \neq 0}} \frac{f(x_0 + h) - f(x_0)}{h} \tag{5.1}$$

21

exists. In this case we call $f'(x_0)$ the **derivative** of f at x_0. If f is differentiable at every point of D, then we also say that f is **differentiable** on D, and we call the resulting function $f' : D \to \mathbb{R}$ the **derivative** of f.

To simplify the notation, we shall often write

$$f'(x_0) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h},$$

without emphasizing that $x \neq x_0$ or $h \neq 0$.

REMARK 5.2. — If $f : D \to \mathbb{R}$ is differentiable at x_0, then it is also continuous at x_0. Indeed, the condition $f'(x_0) = \lim_{x \to x_0} \frac{f(x)-f(x_0)}{x-x_0}$ can be rewritten as

$$\lim_{x \to x_0} \frac{f(x) - f(x_0) - f'(x_0)(x - x_0)}{x - x_0} = 0,$$

or equivalently, recalling the little-o notation in Definition 3.84,

$$f(x) - f(x_0) - f'(x_0)(x - x_0) = o(x - x_0).$$

Therefore

$$\lim_{x \to x_0} f(x) = \lim_{x \to x_0} \Big(f(x_0) + f'(x_0)(x - x_0) + o(x - x_0) \Big) = f(x_0),$$

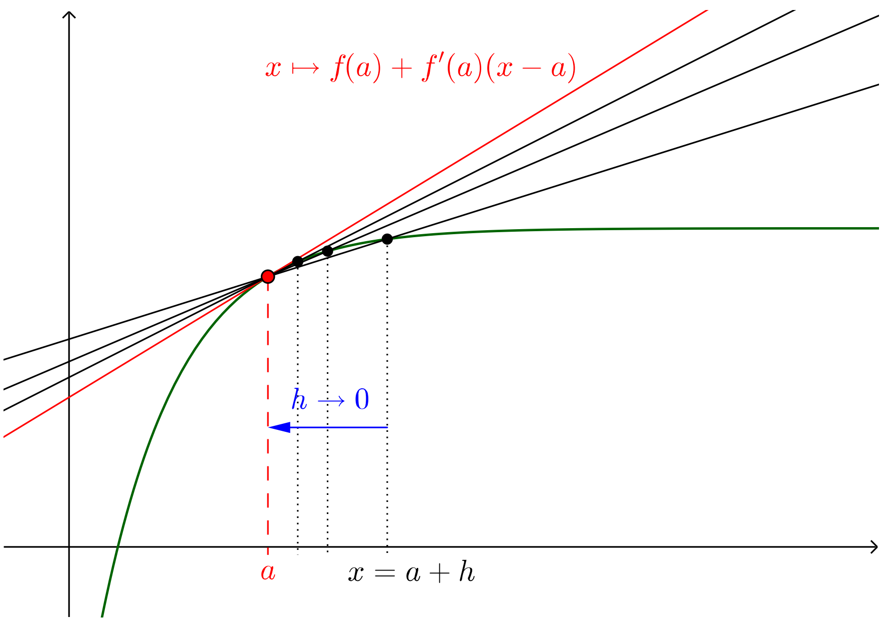
hence f is continuous at x_0.

5.3. — An alternative notation for the derivative of f is $\frac{df}{dx}$. If $x_0 \in D$ is an accumulation point from the right of D, then f is **differentiable from the right** at x_0 if the **derivative from the right**

$$f'_+(x_0) = \lim_{x \to x_0^+} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{h \to 0+} \frac{f(x_0 + h) - f(x_0)}{h}$$

exists. **Differentiability from the left** and the **derivative from the left** $f'_-(x_0)$ are
21 defined analogously considering the limit $x \to x_0^-$.

5.4. — An **affine function** is a function of the form $x \mapsto sx + r$, for real numbers s and r. The graph of an affine function is a non-vertical **line** in $\mathbb{R}^2$. The parameter s in the equation $y = sx + r$ is called the **slope** of the straight line. If $f : D \to \mathbb{R}$ is differentiable at a point $a \in D$, the function $x \mapsto f(a) + f'(a)(x - a)$ is called **affine approximation** of f at a.



The geometric interpretation of the derivative of a real-valued function f at a is the slope of the tangents of the graph at a. This is because when x tends towards a, the secant going through $(a, f(a))$ and $(x, f(x))$ and having the difference quotient as its slope approaches the tangent of the graph at a.

EXAMPLE 5.5. —
- Constant functions are everywhere differentiable and have the zero function as their derivative.
- The identity function $f(x) = x$ is differentiable, and its derivative is the constant function 1. Indeed
$$f'(x_0) = \lim_{x \to x_0} \frac{x - x_0}{x - x_0} = 1 \qquad \forall\, x_0 \in \mathbb{R}.$$

EXAMPLE 5.6. — The exponential function $\exp : \mathbb{R} \to \mathbb{R}_{>0}$ is differentiable and its derivative is again the exponential function. Indeed, for $x \in \mathbb{R}$, since $e^{x+h} = e^x e^h$ we get

$$\begin{aligned}(e^x)' &= \lim_{h \to 0} \frac{e^{x+h} - e^x}{h} = e^x \lim_{h \to 0} \frac{e^h - 1}{h} \\ &= e^x \lim_{h \to 0} \sum_{k=1}^{\infty} \frac{1}{k!} h^{k-1} = e^x \lim_{h \to 0} \sum_{n=0}^{\infty} \frac{1}{(n+1)!} h^n.\end{aligned}$$

Note now that the power series $h \mapsto \sum_{n=0}^{\infty} \frac{1}{(n+1)!} h^n$ has infinite radius of convergence, as it
21 follows for instance from Exercise 4.45. In particular the function $g(h) = \sum_{n=0}^{\infty} \frac{1}{(n+1)!} h^n$ is continuous on $\mathbb{R}$, and $g(0) = \frac{1}{1!} = 1$. Therefore

$$(e^x)' = e^x \lim_{h \to 0} g(h) = e^x g(0) = e^x.$$

More in general, let α be a complex number and let $f : \mathbb{R} \to \mathbb{C}$ be the complex-valued function given by $f(x) = e^{\alpha x}$. The derivative of f can still be defined as the limit of the incremental ratios, namely

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

and one gets

$$\begin{aligned}(e^{\alpha x})' &= \lim_{h \to 0} \frac{e^{\alpha x + \alpha h} - e^{\alpha x}}{h} = e^{\alpha x} \lim_{h \to 0} \frac{e^{\alpha h} - 1}{h} \\ &= e^{\alpha x} \lim_{h \to 0} \sum_{k=1}^{\infty} \frac{1}{k!} \alpha^k h^{k-1} = \alpha e^{\alpha x}.\end{aligned}$$

EXAMPLE 5.7. — Let $f : \mathbb{R} \setminus \{0\} \to \mathbb{R}$ given by $f(x) = \frac{1}{x}$. Then f is differentiable and it holds $f'(x) = -\frac{1}{x^2}$ for all $x \in \mathbb{R} \setminus \{0\}$. In fact

$$f'(x) = \lim_{h \to 0} \frac{\frac{1}{x+h} - \frac{1}{x}}{h} = \lim_{h \to 0} \frac{x - (x+h)}{(x+h)xh}$$

$$= -\lim_{h\to 0}\frac{1}{(x+h)x} = -\frac{1}{\lim_{h\to 0}(x+h)x} = -\frac{1}{x^2}.$$

Definition 5.8: Higher Derivatives

Let $f : D \to \mathbb{R}$ be a function. We define the **higher derivatives** of f, if they exist, by

$$f^{(0)} = f, \qquad f^{(1)} = f', \qquad f^{(2)} = f'', \qquad \ldots, \qquad f^{(n+1)} = (f^{(n)})'$$

for all $n \in \mathbb{N}$. If $f^{(n)}$ exists for any $n \in \mathbb{N}$, f is called **n-times differentiable**. If the n-th derivative $f^{(n)}$ is also continuous, f is called n-times **continuously differentiable**. We denote the set of n-times continuously differentiable functions on D by $C^n(D)$.

Differently put, $C^0(D)$ denotes the set of real-valued continuous functions on D, and $C^1(D)$ denotes the set of all differentiable functions whose derivative is continuous. We call such functions **continuously differentiable** or of **class** C^1. Recursively, for $n \geq 1$ we define

$$C^n(D) = \{f : D \to \mathbb{R} \mid f \text{ is differentiable and } f' \in C^{n-1}(D)\}$$

and say $f \in C^n(D)$ is of **class** C^n.

21 Example 5.9. — The function $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = \operatorname{sgn}(x)x^2$ is differentiable, and the derivative of f is given by $f'(x) = 2|x|$. This shows that f is continuously differentiable, i.e., of class C^1. Since the continuous function f' is not differentiable at 0, f is not of class C^2.

Definition 5.10: Smooth Functions

We define

$$C^\infty(D) = \bigcap_{n=0}^{\infty} C^n(D)$$

and call functions $f \in C^\infty(D)$ **smooth** or of **class** C^∞.

Example 5.11. — The exponential function $\exp : \mathbb{R} \to \mathbb{R}$ is smooth. Polynomial functions are smooth.

5.1.2 Differentiation Rules

As with continuous functions, we do not always want to show by hand that a given function is differentiable. Instead, we want to prove the general rules by which the differentiability of various functions can be traced.

Proposition 5.12: Derivative of Sum and Product

Let $D \subseteq \mathbb{R}$ and $x_0 \in D$ be an accumulation point of $D \setminus \{x_0\}$. Let $f, g : D \to \mathbb{R}$ be differentiable at x_0. Then $f + g$ and $f \cdot g$ are differentiable at x_0, and the following holds:

$$(f+g)'(x_0) = f'(x_0) + g'(x_0), \tag{5.2}$$

$$(fg)'(x_0) = f'(x_0)g(x_0) + f(x_0)g'(x_0). \tag{5.3}$$

In particular, any scalar multiple of f is differentiable at x_0 and $(\alpha f)'(x_0) = \alpha f'(x_0)$ for all $\alpha \in \mathbb{R}$.

Proof. We compute using the properties of the limit introduced in Section 3.5.1:

$$\begin{aligned}\lim_{x\to x_0} \frac{(f+g)(x) - (f+g)(x_0)}{x - x_0} &= \lim_{x\to x_0} \left(\frac{f(x) - f(x_0)}{x - x_0} + \frac{g(x) - g(x_0)}{x - x_0} \right) \\ &= \lim_{x\to x_0} \frac{f(x) - f(x_0)}{x - x_0} + \lim_{x\to x_0} \frac{g(x) - g(x_0)}{x - x_0} \\ &= f'(x_0) + g'(x_0),\end{aligned}$$

and

21

$$\begin{aligned}\lim_{x\to x_0} \frac{(f\cdot g)(x) - (f\cdot g)(x_0)}{x - x_0} &= \lim_{x\to x_0} \frac{(f(x) - f(x_0))g(x) + f(x_0)(g(x) - g(x_0))}{x - x_0} \\ &= \lim_{x\to x_0} \left(\frac{f(x) - f(x_0)}{x - x_0} g(x) \right) + f(x_0) \lim_{x\to x_0} \frac{g(x) - g(x_0)}{x - x_0} \\ &= \lim_{x\to x_0} \frac{f(x) - f(x_0)}{x - x_0} \cdot \lim_{x\to x_0} g(x) + f(x_0) \lim_{x\to x_0} \frac{g(x) - g(x_0)}{x - x_0} \\ &= f'(x_0)g(x_0) + f(x_0)g'(x_0),\end{aligned}$$

where we used that g is continuous at x_0 to say that $\lim_{x\to x_0} g(x) = g(x_0)$ (recall Remark 5.2). □

Corollary 5.13: Higher Order Derivatives of the Product

Let $f, g : D \to \mathbb{R}$ be n-times differentiable. Then $f + g$ and $f \cdot g$ are also n-times differentiable and it holds $f^{(n)} + g^{(n)} = (f+g)^{(n)}$ as well as

$$(fg)^{(n)} = \sum_{k=0}^{n} \binom{n}{k} f^{(k)} g^{(n-k)}.$$

In particular, every scalar multiple $\alpha \in \mathbb{R}$, $(\alpha f)^{(n)} = \alpha f^{(n)}$ for all $\alpha \in \mathbb{R}$.

Proof. For $n = 1$ this corresponds to Proposition 5.12. The general case follows by induction on $n \geq 1$. □

Corollary 5.14: Derivatives of Polynomials

Polynomial functions are differentiable on all $\mathbb{R}$. It holds $(1)' = 0$ and $(x^n)' = nx^{n-1}$ for all $n \geq 1$.

Proof. The cases $n = 0$ and $n = 1$ have already been discussed in Example 5.5. We now prove by induction the case $n > 1$. Assume that, for some $n \geq 1$, $(x^n)' = nx^{n-1}$ holds. Then it follows from (5.3) that $x^{n+1} = xx^n$ is differentiable and

$$(x^{n+1})' = (xx^n)' = 1x^n + x(nx^{n-1}) = (n+1)x^n.$$

This proves the inductive step and concludes the proof.

Differentiability of any polynomial function now follows from the linearity of the derivative, see (5.2). □

Example 5.15. — Thanks to Example 5.6 in the special cases $\alpha = \pm 1$ and $\alpha = \pm i$ in Example 5.6, we deduce that

$$(e^x)' = e^x, \quad (e^{-x})' = -e^{-x}, \quad (e^{ix})' = ie^{ix}, \quad (e^{-ix})' = -ie^{-ix}.$$

As a consequence, recalling Theorem 4.57 we see that

21

$$\sin'(x) = \frac{(e^{ix})' - (e^{-ix})'}{2i} = \frac{e^{ix} + e^{-ix}}{2} = \cos(x),$$

and analogously $\cos'(x) = -\sin(x)$. Similarly, $\sinh'(x) = \cosh(x)$ and $\cosh'(x) = \sinh(x)$.

Theorem 5.16: Chain Rule

Let $D, E \subseteq \mathbb{R}$ be subsets, and let $x_0 \in D$ be an accumulation point of $D \setminus \{x_0\}$. Let $f : D \to E$ be differentiable at x_0 such that $y_0 = f(x_0)$ is an accumulation point of $E \setminus \{y_0\}$, and let $g : E \to \mathbb{R}$ be differentiable at y_0. Then $g \circ f : D \to \mathbb{R}$ is differentiable at x_0 and

$$(g \circ f)'(x_0) = g'(f(x_0))f'(x_0).$$

Proof. Note that one can write $g(y) = g(y_0) + g'(y_0)(y - y_0) + \varepsilon_g(y)(y - y_0)$ with

$$\varepsilon_g(y) = \begin{cases} \frac{g(y)-g(y_0)}{y-y_0} - g'(y_0) & \text{if } y \in E \setminus \{y_0\} \\ 0 & \text{if } y = y_0. \end{cases}$$

Also, since g is differentiable at y_0, the function ε_g is continuous at y_0. Substituting $y = f(x)$ and recalling that $y_0 = f(x_0)$ we get

$$g(f(x)) = g(f(x_0)) + g'(f(x_0))[f(x) - f(x_0))] + \varepsilon_g(f(x))[f(x) - f(x_0)]$$

for all $x \in D$, therefore

$$\begin{aligned}\lim_{x\to x_0}\frac{g(f(x))-g(f(x_0))}{x-x_0} &= \lim_{x\to x_0}\left(g'(f(x_0))\frac{f(x)-f(x_0)}{x-x_0}+\varepsilon_g(f(x))\frac{f(x)-f(x_0)}{x-x_0}\right)\\ &= g'(f(x_0))f'(x_0)+\underbrace{\varepsilon_g(f(x_0))}_{=0}f'(x_0)=g'(f(x_0))f'(x_0),\end{aligned}$$

where we used Proposition 3.71 and the continuity of ε_g at $y_0 = f(x_0)$ to deduce that $\varepsilon_g(f(x)) \to \varepsilon_g(f(x_0))$ as $x \to x_0$. □

Remark 5.17. — By a nontrivial induction argument, one can prove that if $f : D \to E$ and $g : E \to \mathbb{R}$ are n-times differentiable, then also $g \circ f : D \to \mathbb{R}$ is n-times differentiable and one can express the n-th derivative of $g \circ f$ in terms of sums and products of $g' \circ f$, $g^{(2)} \circ f$, $\ldots$, $g^{(n)} \circ f$, f', $f^{(2)}$, $\ldots$, $f^{(n)}$.

Corollary 5.18: Quotient Rule

Let $D \subseteq \mathbb{R}$, $x_0 \in D$ an accumulation point of $D \setminus \{x_0\}$, and $f, g : D \to \mathbb{R}$ differentiable at x_0. If $g(x_0) \neq 0$, then $\frac{f}{g}$ is also differentiable at x_0 and it holds

21

$$\left(\frac{f}{g}\right)'(x_0) = \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g(x_0)^2}.$$

Proof. Let $\psi : \mathbb{R} \setminus \{0\} \to \mathbb{R}$ be the function given by $\psi(y) = \frac{1}{y}$, which is differentiable by Example 5.7. We combine this with the chain rule (Theorem 5.16) and obtain that the function $\frac{1}{g} = \psi \circ g$ is differentiable at x_0, with derivative

$$\left(\frac{1}{g}\right)'(x_0) = -\frac{1}{g(x_0)^2}g'(x_0).$$

If we now use the product rule in Proposition 5.12, it follows that $\frac{f}{g} = f \cdot \frac{1}{g}$ is differentiable at x_0, and

$$\left(\frac{f}{g}\right)'(x_0) = \left(f \cdot \frac{1}{g}\right)'(x_0) = f'(x_0)\frac{1}{g(x_0)} - f(x_0)\frac{g'(x_0)}{g(x_0)^2} = \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g(x_0)^2}.$$

□

Example 5.19. — We determine the derivative of the function $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = \exp\big(\sin(\sin(x^2))\big)$ by applying the chain rule several times. Since $\exp' = \exp$ we obtain

$$f'(x) = \exp(g(x))g'(x),$$

where $g(x) = \sin(\sin(x^2))$. Similarly, because $\sin' = \cos$,

$$g'(x) = \cos(h(x))h'(x),$$

where $h(x) = \sin(x^2)$. Finally,

21 $$h'(x) = \cos(k(x))k'(x)$$

with $k(x) = x^2$. Since $k'(x) = 2x$, this gives

$$f'(x) = \exp\big(\sin(\sin(x^2))\big)\cos(\sin(x^2))\cos(x^2)2x \qquad \forall\, x \in \mathbb{R}.$$

Exercise 5.20. — Determine the derivative of the function $x \mapsto \cos\big(\sin^3(\exp(x))\big)$.

Theorem 5.21: Derivative of the Inverse

Let $D, E \subseteq \mathbb{R}$, and let $f : D \to E$ be a continuous bijective mapping whose inverse $f^{-1} : E \to D$ is also continuous. Let $\bar{x} \in D$ be an accumulation point of $D \setminus \{\bar{x}\}$, and assume that f is differentiable at $\bar{x} \in D$ with $f'(\bar{x}) \neq 0$. Then f^{-1} is differentiable at $\bar{y} = f(\bar{x})$ and

$$(f^{-1})'(\bar{y}) = \frac{1}{f'(\bar{x})}$$

Proof. We first show that $\bar{y}$ is an accumulation point of $E \setminus \{\bar{y}\}$, which allows us to speak of differentiability at $\bar{y}$. In fact, since by assumption $\bar{x}$ is an accumulation point of $D \setminus \{\bar{x}\}$,
22 there exists a sequence $(x_n)_{n=0}^{\infty}$ in $D \setminus \{\bar{x}\}$ with $x_n \to \bar{x}$ as $n \to \infty$. Since f is continuous and bijective, the sequence $(f(x_n))_{n=0}^{\infty}$ converges to $\bar{y} = f(\bar{x})$.

Now, to compute the derivative, let $(y_n)_{n=0}^{\infty}$ be an arbitrary sequence in $E \setminus \{\bar{y}\}$ converging to $\bar{y}$. Then $x_n = f^{-1}(y_n)$ tends towards $\bar{x}$ (since f^{-1} is continuous by assumption), and the following holds:

$$\lim_{n\to\infty} \frac{f^{-1}(y_n) - f^{-1}(\bar{y})}{y_n - \bar{y}} = \lim_{n\to\infty} \frac{x_n - \bar{x}}{f(x_n) - f(\bar{x})} = \lim_{n\to\infty} \left(\frac{f(x_n) - f(\bar{x})}{x_n - \bar{x}}\right)^{-1} = \frac{1}{f'(\bar{x})}.$$

Hence, if we set $g(y) = \frac{f^{-1}(y) - f^{-1}(\bar{y})}{y - \bar{y}}$, this proves that for every sequence y_n converging to $\bar{y}$ it holds $g(y_n) \to \frac{1}{f'(\bar{x})}$. Recalling Lemma 3.70, this proves that $\lim_{y\to\bar{y}} g(y) = \frac{1}{f'(\bar{x})}$, as desired. □

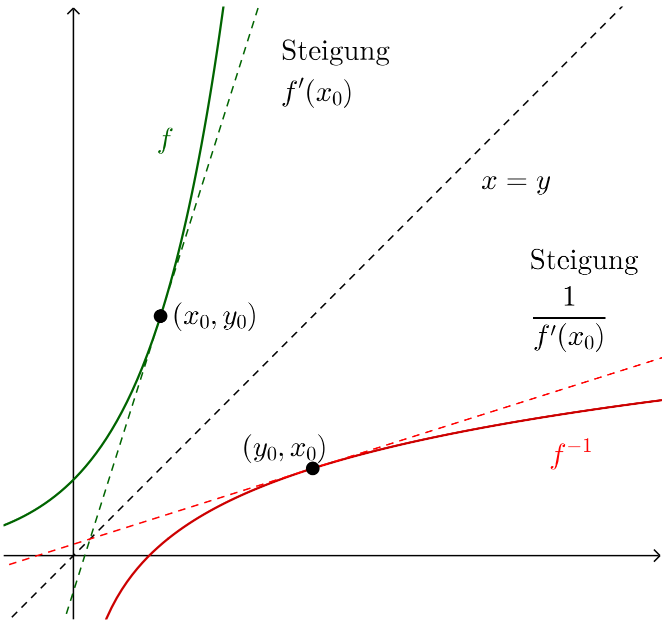


Figure 5.1: An intuitive representation of Theorem 5.21. Mirroring the graph of f and the tangent line at the point (x_0, y_0) around the straight line $x = y$ in $\mathbb{R}^2$, we get the graph of f^{-1} and, this is the assertion, the tangent line at (y_0, x_0). A short calculation shows that the reflection of a straight line with slope m around $x = y$ has slope $\frac{1}{m}$.

22 Example 5.22. — The function $g : \mathbb{R} \setminus \{0\} \to \mathbb{R}$ given by $g(y) = \log(|y|)$ is differentiable, with $g'(y) = \frac{1}{y}$ for all $y \in \mathbb{R} \setminus \{0\}$. Indeed, since the map $\log : \mathbb{R}_{>0} \to \mathbb{R}$ is the inverse of $\exp : \mathbb{R} \to \mathbb{R}_{>0}$, it follows from Theorem 5.21 that g is differentiable at all points $y > 0$ with $g'(y) = \frac{1}{f'(x)}$, where $x = g(y) = \log(y)$. Since $\exp' = \exp$, this gives

$$g'(y) = \log'(y) = \frac{1}{\exp(x)} = \frac{1}{\exp(\log(y))} = \frac{1}{y}.$$

For $y < 0$, since $g(y) = \log(-y)$, we apply the chain rule (Theorem 5.16) to get $g'(y) = -\log'(-y) = -\frac{1}{-y} = \frac{1}{y}$.

Example 5.23. — Given $x > 0$ and $\alpha \in \mathbb{R}$, we can compute the derivative of x^α as follows:

$$x^\alpha = \exp(\alpha \log(x)) \implies (x^\alpha)' = \exp'(\alpha \log(x))\alpha \log'(x) = \exp(\alpha \log(x))\frac{\alpha}{x} = \alpha x^{\alpha-1}.$$

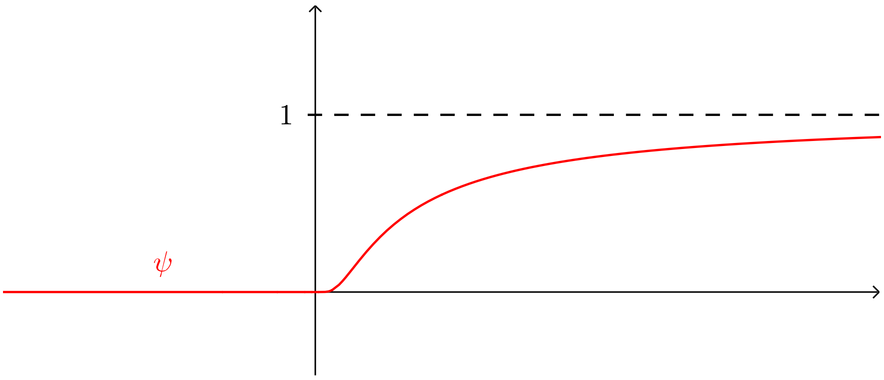
This formula generalizes Corollary 5.14.

Example 5.24. — The logarithm $f = \log : \mathbb{R}_{>0} \to \mathbb{R},\ x \to \log(x)$ is smooth. Indeed, by the example above, $f'(x) = \frac{1}{x}$. Then, by induction and the Leibniz rule, one proves that $f''(x) = -\frac{1}{x^2}$, $f^{(3)}(x) = \frac{2}{x^3}$, and, in general, $f^{(n)}(x) = (-1)^{n-1}\frac{(n-1)!}{x^n}$.

Exercise 5.25. — Consider the function $\psi : \mathbb{R} \to \mathbb{R}$ defined by

$$\psi(x) = \begin{cases} \exp\left(-\frac{1}{x}\right) & \text{if } x > 0 \\ 0 & \text{if } x \le 0 \end{cases}.$$

Show that ψ is smooth on $\mathbb{R}$, and that all its derivatives at 0 vanish.



22

Hint: Show first by induction that, for all $n \in \mathbb{N}$,

$$\psi^{(n)}(x) = \exp\left(-\frac{1}{x}\right) f_n\left(\frac{1}{x}\right) \qquad \forall\, x > 0, \tag{5.4}$$

where f_n is a polynomial. Then, using that the exponential function grows faster than any polynomial, prove by induction that

$$\lim_{x \to 0^+} \frac{\psi(x)}{x^k} = 0 \qquad \forall\, k \in \mathbb{N}.$$

Conclude (again by induction on $n \in \mathbb{N}$) that

$$\psi^{(n+1)}(0) = \lim_{x \to 0} \frac{\psi^{(n)}(x) - \psi^{(n)}(0)}{x - 0} = \lim_{x \to 0^+} \frac{\psi(x) f_n\left(\frac{1}{x}\right) - 0}{x} = \lim_{x \to 0^+} \psi(x) f_n\left(\frac{1}{x}\right) \frac{1}{x} = 0.$$

5.2 Main Theorems of Differential Calculus

5.2.1 Local Extrema

Definition 5.26: Local extrema

Let $D \subseteq \mathbb{R}$ and $x_0 \in D$. We say that a function $f : D \to \mathbb{R}$ has a **local maximum** at x_0 if there exists $\delta > 0$ such that $f(x) \leq f(x_0)$ for all $x \in D \cap (x_0 - \delta, x_0 + \delta)$.
If the inequality is strict, namely that $f(x) < f(x_0)$ for all $x \in D \cap (x_0 - \delta, x_0 + \delta) \setminus \{x_0\}$, then f in x_0 has an **isolated local maximum**. The value $f(x_0)$ is called a **local maximal value** of f.
A **local minimum**, an **isolated local minimum**, and a **local minimal value** of f are defined analogously.
Furthermore, we call x_0 a **local extremum** of f, and $f(x_0)$ a **local extremal value** of f, if f has a local minimum or a local maximum in x_0.

Proposition 5.27: Local Extrema vs First Derivative

Let $D \subseteq \mathbb{R}$ and $f : D \to \mathbb{R}$. Suppose f is differentiable at a local extremum $x_0 \in D$, and
22 *x_0 is both a right-hand side and a left-hand side accumulation point of D (namely, x_0 is an accumulation point for $D \cap (x_0, \infty)$ and $D \cap (-\infty, x_0)$). Then $f'(x_0) = 0$ holds.*

Proof. Without restriction of generality, we assume that f has a local maximum in $x_0 \in D$ (otherwise replace f by $-f$). Since f is differentiable at x_0 and x_0 can be approximated both from the left and right, we see that the following holds.

On the one hand, we have

$$f'_+(x_0) = \lim_{x \to x_0^+} \frac{f(x) - f(x_0)}{x - x_0} \leq 0$$

since, for x close to x_0 to the right of x_0, $f(x) - f(x_0) \leq 0$ and $x - x_0 > 0$.

On the other hand,

$$f'_-(x_0) = \lim_{x \to x_0^-} \frac{f(x) - f(x_0)}{x - x_0} \geq 0$$

since now, for x close to x_0 to the left of x_0, $f(x) - f(x_0) \leq 0$ and $x - x_0 < 0$.

Since f is differentiable at x_0 then $f'(x_0) = f'_+(x_0) = f'_-(x_0)$, therefore $f'(x_0) = 0$. □

Corollary 5.28: Local Extrema in an Interval

Let $I \subseteq \mathbb{R}$ be an interval, and $f : I \to \mathbb{R}$. Let $x_0 \in I$ be a local extremum of f. Then at least one of the following statements is true.

1. *$x_0 \in I$ is an endpoint of I;*
2. *f is not differentiable at x_0;*
3. *f is differentiable at x_0 and $f'(x_0) = 0$.*

In particular, all local extrema of a differentiable function on an open interval are zeros of the derivative.

Exercise 5.29. — Let $f : \mathbb{R} \to \mathbb{R}$ be the polynomial function $f(x) = x^3 - x$. Find all local extrema of f. Find all local extrema of the function $|f|$ on $[-3, 3]$.

5.2.2 The Mean Value Theorem

We now turn to general theorems of differential calculus and their consequences. Our first
22 question will be whether the derivative of a differentiable function takes the slope of certain
secants, and the following theorem will be our starting point.

Theorem 5.30: Rolle's Theorem

Let $a < b$, and $f : [a, b] \to \mathbb{R}$ be a continuous function, differentiable on the open interval (a, b). If $f(a) = f(b)$ holds, then there exists $\xi \in (a, b)$ with $f'(\xi) = 0$.

Proof. According to Theorem 3.42, the minimum and maximum of f exist in $[a, b]$. That is, there exist $x_0, x_1 \in [a, b]$ with $f(x_0) \leq f(x) \leq f(x_1)$ for all $x \in [a, b]$. According to Proposition 5.27, the derivative of f must be zero for all extrema in (a, b). So if either $x_0 \in (a, b)$ or $x_1 \in (a, b)$ holds, then we have already found a $\xi \in (a, b)$ with $f'(\xi) = 0$. Instead, if both x_0 and x_1 are endpoints of the interval, because $f(a) = f(b)$ it follows that f is constant, hence $f'(x) = 0$ holds for all $x \in (a, b)$. □

Theorem 5.31: Mean Value Theorem

Let $a < b$, and $f : [a, b] \to \mathbb{R}$ be a continuous function, differentiable on the open interval (a, b). Then there exists $\xi \in (a, b)$ with

$$f'(\xi) = \frac{f(b) - f(a)}{b - a}.$$

Proof. We define a function $g : [a,b] \to \mathbb{R}$ by

$$g(x) = f(x) - \frac{f(b) - f(a)}{b - a}(x - a)$$

for all $x \in [a,b]$. Note that g is continuous, and satisfies

$$g(a) = f(a), \qquad g(b) = f(b) - (f(b) - f(a)) = f(a).$$

Furthermore, since both functions

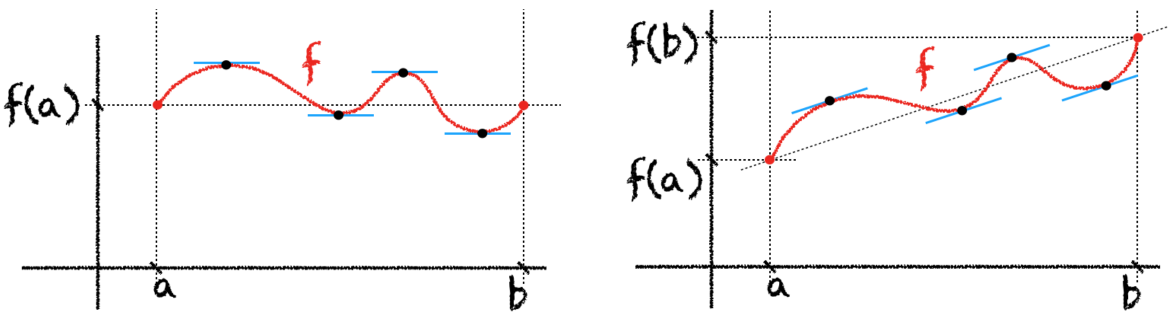
$$x \mapsto f(x) \qquad \text{and} \qquad x \mapsto \frac{f(b) - f(a)}{b - a}(x - a)$$

are differentiable in (a,b), it follows from Proposition 5.12 that g is differentiable in (a,b). Thus, according to Rolle's theorem, there exists $\xi \in (a,b)$ such that

$$0 = g'(\xi) = f'(\xi) - \frac{f(b) - f(a)}{b - a}$$

as desired. □

5.32. — So, in words, Rolle's theorem states that if a differentiable function on an inter-
22 val takes the same value at the endpoints, the slope must be zero somewhere between the
endpoints. We illustrate this in the following picture on the left.



Instead, according to the Mean Value Theorem, for every differentiable function on an interval, there is at least one point where the slope of the function is exactly the average slope. This can be seen in the picture on the right. You can also see here how the proof of the mean value theorem can be traced back to Rolle's theorem by modifying the graph of the function f on the right by shearing in such a way that $f(a) = f(b)$ applies afterwards.

Exercise 5.33. — Let $[a,b]$ be a compact interval and $f : [a,b] \to \mathbb{R}$ be continuously differentiable. Show that f is Lipschitz continuous. What happens if compactness is dropped as a hypothesis?

Example 5.34. — Let $f : [0, 2\pi] \to \mathbb{C}$ be the complex-valued function given by $f(t) = \exp(it) = \cos(t) + i\sin(t)$. At the endpoints of the interval $[0, 2\pi]$, $f(0) = f(2\pi) = 1$ holds. However, the derivative of f never takes the value zero because, according to Example 5.5,

$$f'(t) = i\exp(it) \neq 0$$

for all $t \in [0, 2\pi]$. Thus, the statements of Rolle's Theorem and the Mean Value Theorem for complex-valued functions are false in this generality.

Theorem 5.35: Cauchy Mean Value Theorem

Let $a < b$, and let $f, g : [a, b] \to \mathbb{R}$ be continuous functions, both of them differentiable on (a, b). Then there exists $\xi \in (a, b)$ such that

$$g'(\xi)(f(b) - f(a)) = f'(\xi)(g(b) - g(a)). \tag{5.5}$$

If in addition $g'(x) \neq 0$ holds for all $x \in (a, b)$, then $g(a) \neq g(b)$ and

$$\frac{f'(\xi)}{g'(\xi)} = \frac{f(b) - f(a)}{g(b) - g(a)}.$$

Proof. We define the function $F : [a, b] \to \mathbb{R}$ by

22

$$F(x) = g(x)\big(f(b) - f(a)\big) - f(x)\big(g(b) - g(a)\big).$$

Then $F(a) = F(b)$ since

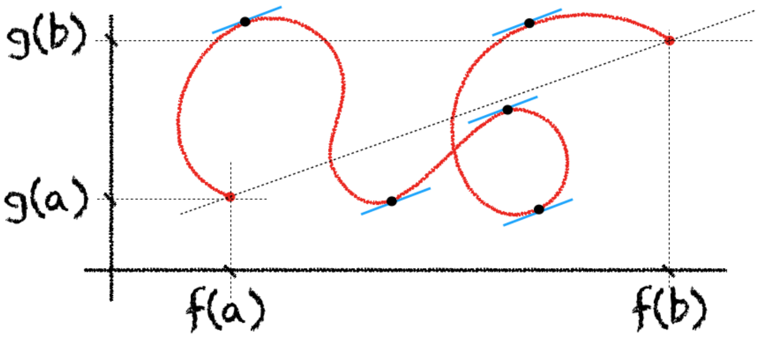
$$\begin{aligned} F(a) &= g(a)\big(f(b) - f(a)\big) - f(a)\big(g(b) - g(a)\big) = g(a)f(b) - f(a)g(b), \\ F(b) &= g(b)\big(f(b) - f(a)\big) - f(b)\big(g(b) - g(a)\big) = g(a)f(b) - f(a)g(b). \end{aligned}$$

Thus, according to Rolle's Theorem 5.30, there exists $\xi \in (a, b)$ such that

$$F'(\xi) = g'(\xi)(f(b) - f(a)) - f'(\xi)(g(b) - g(a)) = 0.$$

This proves (5.5). If in addition $g'(x) \neq 0$ for all $x \in (a, b)$, then it follows from Rolle's theorem that $g(b) \neq g(a)$ holds. After dividing (5.5) by $g'(\xi)(g(b) - g(a))$, we get the second assertion of the theorem. □

5.36. — Just like for the Mean Value Theorem 5.31, Cauchy's Mean Value Theorem has a geometrical interpretation, only this time you have to look in the two-dimensional plane. There, Cauchy's mean value theorem states, under the assumptions made, that the curve $t \mapsto (f(t), g(t))$ has a tangent that is parallel to the straight line through the points $(f(a), g(a))$, $(f(b), g(b))$.



5.2.3 L'Hopital's Rule

The family of results known collectively as *Rule of l'Hôpital (or l'Hospital)* is named after the French mathematician and nobleman Guillaume François Antoine, Marquis de l'Hôpital (1661–1704). It probably goes back to Johann Bernoulli, but was published by de l'Hôpital in his textbook *Analyse des Infiniment Petits pour l'Intelligence des Lignes Courbes.* De l'Hôpital's influential book was the first systematic treatment of infinitesimal calculus. His approach and argumentation are thoroughly geometrical - de l'Hôpital did not know neither a strict notion of limit nor a notion of differentiability.

22

Theorem 5.37: L'Hôpital's Rule

Let $a < b$, and $f, g : (a, b) \to \mathbb{R}$ differentiable functions. Suppose the following hold:

1. $g(x) \neq 0$ *and* $g'(x) \neq 0$ *for all* $x \in (a, b)$.
2. $\lim_{x \to a^+} f(x) = \lim_{x \to a^+} g(x) = 0$.
3. *The limit* $L = \lim_{x \to a^+} \frac{f'(x)}{g'(x)}$ *exists.*

Then $\lim_{x \to a^+} \frac{f(x)}{g(x)}$ *also exists and is equal to* L.

Proof. By assumption (2), we can extend f and g continuously on $[a, b)$ by setting $f(a) = g(a) = 0$. Fix $\varepsilon > 0$. According to assumption (3), there exists $\delta > 0$ such that

$$\frac{f'(\xi)}{g'(\xi)} \in (L - \varepsilon, L + \varepsilon) \qquad \forall \xi \in (a, a + \delta).$$

Now, given $x \in (a, a + \delta)$, we apply the Mean Value Theorem 5.35 on the interval $[a, x]$ to find $\xi_x \in (a, x)$ such that

$$\frac{f(x)}{g(x)} = \frac{f(x) - f(a)}{g(x) - g(a)} = \frac{f'(\xi_x)}{g'(\xi_x)}.$$

Since $\xi_x \in (a,x) \subseteq (a,a+\delta)$, it follows that

22
$$\frac{f(x)}{g(x)} = \frac{f'(\xi_x)}{g'(\xi_x)} \in (L-\varepsilon, L+\varepsilon) \qquad \forall\, x \in (a,a+\delta).$$

Since $\varepsilon > 0$ is arbitrary, this proves that $\lim_{x\to a^+} \frac{f(x)}{g(x)} = L$, as desired. □

Theorem 5.37 is one of several versions of the rule of l'Hôpital For instance, one can assume that the limits in (2) are both improper limits, i.e., that $\lim_{x\to a^+} g(x) = \pm\infty$ and $\lim_{x\to a^+} f(x) = \pm\infty$ hold with arbitrary signs. More precisely, the following holds:

Theorem 5.38: L'Hôpital's Rule for Improper Limits

Let $a < b$, and $f, g : (a,b) \to \mathbb{R}$ differentiable functions. Suppose the following hold:

1. *$g(x) \neq 0$ and $g'(x) \neq 0$ for all $x \in (a,b)$.*
2. $\lim\limits_{x\to a^+} |f(x)| = \lim\limits_{x\to a^+} |g(x)| = \infty$.
3. *The limit $L = \lim\limits_{x\to a^+} \frac{f'(x)}{g'(x)}$ exists.*

Then $\lim\limits_{x\to a^+} \frac{f(x)}{g(x)}$ also exists and is equal to L.

23
Proof. (Extra material) Fix $\varepsilon > 0$. According to assumption (3), there exists $\delta > 0$ such that

$$\frac{f'(\xi)}{g'(\xi)} \in (L-\varepsilon, L+\varepsilon) \qquad \forall\, \xi \in (a,a+\delta).$$

Now, given $x \in (a,a+\delta)$, we apply the Mean Value Theorem 5.35 on the interval $[x,a+\delta]$ to find $\xi_x \in (x,a+\delta)$ such that

$$\frac{f(x)-f(a+\delta)}{g(x)-g(a+\delta)} = \frac{f'(\xi_x)}{g'(\xi_x)}.$$

Since $\xi_x \in (x,a+\delta) \subseteq (a,a+\delta)$, it follows that

$$\frac{f(x)-f(a+\delta)}{g(x)-g(a+\delta)} = \frac{f'(\xi_x)}{g'(\xi_x)} \in (L-\varepsilon, L+\varepsilon) \qquad \forall\, x \in (a,a+\delta). \tag{5.6}$$

Note now that

$$\frac{f(x)}{g(x)} = \frac{f(x)-f(a+\delta)}{g(x)-g(a+\delta)} \cdot \frac{1-\frac{g(a+\delta)}{g(x)}}{1-\frac{f(a+\delta)}{f(x)}}. \tag{5.7}$$

Also, since $|f(x)|, |g(x)| \to \infty$ as $x \to a^+$, $\lim_{x\to a^+} \frac{1-\frac{g(a+\delta)}{g(x)}}{1-\frac{f(a+\delta)}{f(x)}} = 1$. Hence, recalling (5.6) and (5.7), there exists $\eta \in (0, \delta)$ such that

$$\frac{f(x)}{g(x)} \in (L - 2\varepsilon, L + 2\varepsilon) \qquad \forall\, x \in (a, a + \eta).$$

Since $\varepsilon > 0$ is arbitrary, this proves that $\lim_{x\to a^+} \frac{f(x)}{g(x)} = L$, as desired. □

Instead of the limit $x \to a^+$ or $x \to b^-$, one can also consider the limit as $x \to -\infty$ or $x \to \infty$. For instance, the following holds:

Theorem 5.39: L'Hôpital's Rule at Infinity

Let $R > 0$ and $f, g : (R, \infty) \to \mathbb{R}$ differentiable functions. Suppose the following hold:

1. *$g(x) \neq 0$ and $g'(x) \neq 0$ for all $x \in (R, \infty)$.*
2. *either $\lim\limits_{x\to\infty} f(x) = \lim\limits_{x\to\infty} g(x) = 0$ or $\lim\limits_{x\to\infty} |f(x)| = \lim\limits_{x\to\infty} |g(x)| = \infty$.*
3. *The limit $L = \lim\limits_{x\to\infty} \frac{f'(x)}{g'(x)}$ exists.*

Then $\lim\limits_{x\to\infty} \frac{f(x)}{g(x)}$ also exists and is equal to L.

23 *Proof.* (Extra material) If $\lim\limits_{x\to\infty} f(x) = \lim\limits_{x\to\infty} g(x) = 0$, apply Theorem 5.37 in the interval $\left(0, \frac{1}{R}\right)$ to the functions $x \mapsto f\left(\frac{1}{x}\right)$ and $x \mapsto g\left(\frac{1}{x}\right)$.

If $\lim\limits_{x\to\infty} |f(x)| = \lim\limits_{x\to\infty} |g(x)| = \infty$, apply instead Theorem 5.38 in the interval $\left(0, \frac{1}{R}\right)$ to the functions $x \mapsto f\left(\frac{1}{x}\right)$ and $x \mapsto g\left(\frac{1}{x}\right)$. □

Exercise 5.40. — Prove Theorem 5.39 in detail.

Exercise 5.41. — Calculate the following limits using l'Hôpital's rule:

$$\text{(a)} \lim_{x\to 0^+} \frac{\sin(x) - x}{x^2 \sin(x)}; \quad \text{(b)} \lim_{x\to 0} \frac{e^x - x - 1}{\cos x - 1}; \quad \text{(c)} \lim_{x\to 2} \frac{x^4 - 4^x}{\sin(\pi x)}; \quad \text{(d)} \lim_{x\to -\infty} x^3 e^x.$$

Exercise 5.42. — Let $a < b$ be real numbers and $f : [a, b] \to \mathbb{R}$ a continuous function. Suppose $x_0 \in [a, b]$ is a point such that f is differentiable on $[a, b] \setminus \{x_0\}$ and suppose that the limit $\lim_{x\to x_0} f'(x)$ exists. Show that f is differentiable at x_0 and that f' is continuous at x_0.

Exercise 5.43. — Let $I = (a, b) \subseteq \mathbb{R}$ be an interval, and let $f : I \to \mathbb{R}$ be twice differentiable. Show the formula

$$f''(x) = \lim_{h\to 0} \frac{f(x + h) - 2f(x) + f(x - h)}{h^2}$$

for all $x \in I$. Using the sign function $x \mapsto \operatorname{sgn}(x)$, verify that the existence of the above limit does not imply twice differentiability.
Hint: Apply l'Hôpital's Rule twice.

5.2.4 Monotonicity and Convexity via Differential Calculus

The mean value theorem allows us to characterize known properties of functions using the derivative. In the following, when we write $f' \geq 0$, we mean that $f'(x) \geq 0$ for all $x \in I$.

Also, I shall always denote a "non-trivial" interval, that is, a non-empty interval that does not consist of a single point.

Proposition 5.44: Monotonicity vs First Derivative

Let $I \subseteq \mathbb{R}$ be an interval, and let $f : I \to \mathbb{R}$ be a differentiable function. Then

$$f' \geq 0 \iff f \text{ is increasing.}$$

Proof. Suppose f is increasing. Then we can note that $f(x+h) - f(x) \geq 0$ for $h > 0$, and $f(x+h) - f(x) \leq 0$ for $h < 0$. Therefore, in either case $\frac{f(x+h)-f(x)}{h} \geq 0$, and we get

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} \geq 0.$$

23

To prove the converse implication, assume that f is not increasing. Then there exist two points $x_1 < x_2$ in I with $f(x_2) > f(x_1)$, and according to the Mean Value Theorem 5.31 there exists $\xi \in [x_1, x_2]$ with $f'(\xi) = \frac{f(x_2)-f(x_1)}{x_2-x_1} < 0$. So $f' \not\geq 0$ on I. □

Remark 5.45. — If $f' > 0$, the above argument can be used to show that f is strictly increasing. However, the converse is false: the function $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto x^3$ is strictly increasing, but $f'(0) = 0$.

Corollary 5.46: Constant Functions vs First Derivative

Let $I \subseteq \mathbb{R}$ be an interval, and let $f : I \to \mathbb{R}$. Then f is constant if and only if f is differentiable and $f'(x) = 0$ holds for all $x \in I$.

Proof. On the one hand, the derivative of a constant function is the zero function.

Conversely, if f is differentiable and $f' = 0$, then $f' \geq 0$ and $-f' \geq 0$. Hence, Proposition 5.44 implies that both f and $-f$ are increasing, so f is constant. □

Exercise 5.47. — Let $I \subseteq \mathbb{R}$ be an interval, and let $f : I \to \mathbb{R}$. Show that f is a polynomial if and only if f is smooth and there exists $n \in \mathbb{N}$ such that $f^{(n)} = 0$.

Definition 5.48: Convex Functions

Let $I \subseteq \mathbb{R}$ be an interval, and let $f : I \to \mathbb{R}$. Then f is called **convex** if, for all $a, b \in I$ with $a < b$ and all $t \in (0,1)$, the inequality

$$f\big((1-t)a + tb\big) \le (1-t)f(a) + tf(b) \tag{5.8}$$

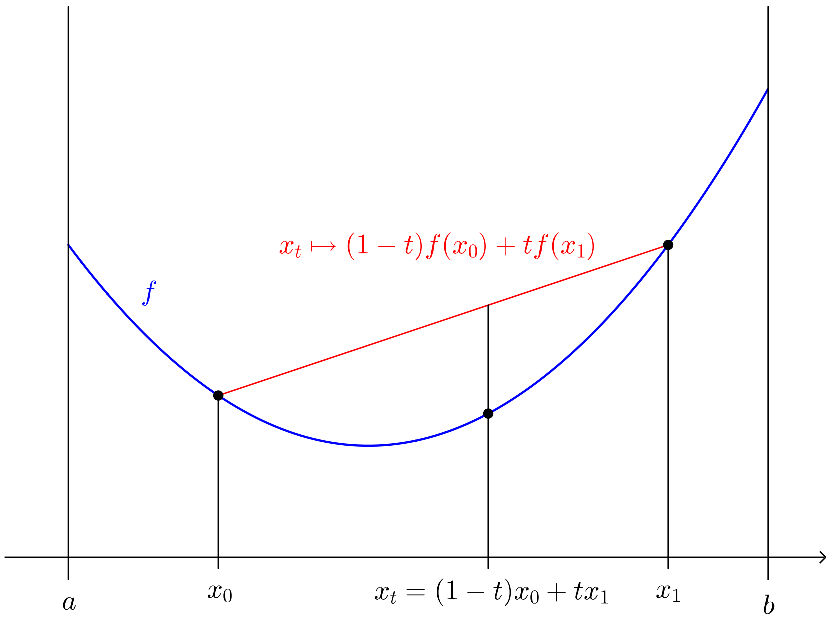
holds. We say that f **strictly convex** if the inequality in (5.8) is strict. A function $g : I \to \mathbb{R}$ is called (**strictly**) **concave** if $-g$ is (strictly) convex.

5.49. — The inequality (5.8) can be understood geometrically as follows: If $a < b$ are points in the domain of definition of f, then the graph of f in the interval $[a, b]$ lies below the secant through the points $(a, f(a))$ and $(b, f(b))$. Convexity can also be characterized by means of slopes of secants. Namely, $f : I \to \mathbb{R}$ is convex if for all $x \in (a, b) \subset I$ the inequality

$$\frac{f(x) - f(a)}{x - a} \le \frac{f(b) - f(x)}{b - x} \tag{5.9}$$

holds, and strictly convex if the inequality above is strict. Geometrically, this means that the slope of the lines through the points $(a, f(a))$ and $(x, f(x))$ is smaller than the slope of the lines through the points $(x, f(x))$ and $(b, f(b))$.

23



Exercise 5.50. — Show that the inequality (5.8) for all $t \in (0,1)$ is equivalent to the inequality (5.9) for all $x \in (a, b)$.

Proposition 5.51: Convexity vs Monotonicity of First Derivative

Let $I \subseteq \mathbb{R}$ be an interval and $f : I \to \mathbb{R}$ a differentiable function. Then f is convex if and only if f' is increasing.

Proof. Suppose first that f' is increasing. Fix $a, b \in I$ with $a < b$, and consider $x \in (a, b)$. According to the Mean Value Theorem 5.31, there exist $\xi \in (a, x)$ and $\zeta \in (x, b)$ such that

$$f'(\xi) = \frac{f(x) - f(a)}{x - a}, \qquad f'(\zeta) = \frac{f(b) - f(x)}{b - x}.$$

Thus, from the monotonicity of f' we deduce (5.9), so f is convex.

Vice versa, suppose that f is convex. Let $a < b$ be two points in I, and consider $h > 0$ small enough so that $a + h < b - h$. Applying (5.9) on the interval $(a, b - h)$ with $x = a + h$ it follows that

$$\frac{f(a+h) - f(a)}{h} \leq \frac{f(b-h) - f(a+h)}{(b-h) - (a+h)}.$$

Also, from (5.9) applied on the interval $(a + h, b)$ with $x = b - h$, we get

$$\frac{f(b-h) - f(a+h)}{(b-h) - (a+h)} \leq \frac{f(b) - f(b-h)}{h}.$$

Combining the two inequalities above, we deduce that for all $h > 0$ sufficiently small,

$$\frac{f(a+h) - f(a)}{h} \leq \frac{f(b) - f(b-h)}{h}. \tag{5.10}$$

Taking the limit as $h \to 0^+$, we obtain $f'(a) \leq f'(b)$. Since $a < b$ are arbitrary, this proves that f' is increasing. □

23

Exercise 5.52. — With the same assumptions as in Proposition 5.51, prove that f is strictly convex if and only if f' is strictly increasing.

Corollary 5.53: Convexity vs Second Derivative

Let $I \subseteq \mathbb{R}$ be an interval, and $f : I \to \mathbb{R}$ a twice differentiable function. Then f is convex if and only if $f'' \geq 0$.

Proof. By Proposition 5.51, f is convex if and only if f' is increasing. Then, by Proposition 5.44 applied to f', we see that f' is increasing if and only if $f'' \geq 0$. The result follows. □

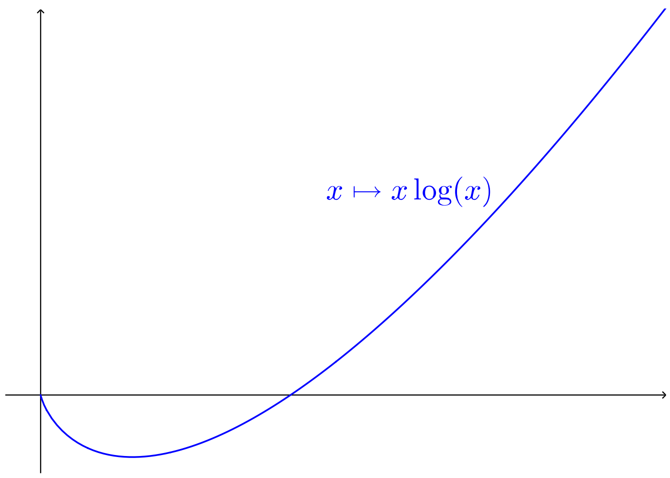
Exercise 5.54. — Under the same assumptions as in Corollary 5.53, prove that if $f''(x) > 0$ for all $x \in I$, then f is strictly convex. Is the converse true?

Exercise 5.55. — The function $f : (0, \infty) \to \mathbb{R}$, $x \mapsto x \log(x)$, is strictly convex. This follows from Corollary 5.53, since f is smooth and

$$f'(x) = \log(x) + x\frac{1}{x} = \log(x) + 1, \quad f''(x) = \frac{1}{x} > 0$$

for all $x > 0$. Furthermore, we already know from Example 3.79 that $\lim_{x\to 0} x \log(x) = 0$. Lastly, we note that $\lim_{x\to 0} f'(x) = -\infty$, all of which can be seen in the graph of f.

23



Exercise 5.56 (Minima of Convex Functions). — Let $I \subseteq \mathbb{R}$ be an interval, and $f : I \to \mathbb{R}$ a convex function. Show that every local minimum of f is a (global) minimum.

Exercise 5.57. — Given $\alpha \in (0, 1]$, prove ˙
Hint: First, setting $t = \frac{x}{x+y}$, show that (3.12) is equivalent to the inequality

$$1 \leq t^\alpha + (1-t)^\alpha \qquad \forall\, t \in [0, 1]. \tag{5.11}$$

Then, prove that the function $x \mapsto x^\alpha$ is concave for $x \in (0, \infty)$ and conclude the validity of (5.11).

5.3 Example: Differentiation of Trigonometric Functions

In this section, we study the derivative and monotonicity properties of trigonometric functions.

5.3.1 Sine and Arc Sine

5.58. — Recalling Exercise 5.15, the functions $\sin : \mathbb{R} \to \mathbb{R}$ and $\cos : \mathbb{R} \to \mathbb{R}$ are smooth, and

$$\sin'(x) = \cos(x), \quad \cos'(x) = -\sin(x).$$

By Theorem 4.59 and Exercise 4.62, the zeros of $\cos : \mathbb{R} \to \mathbb{R}$ are the set $\{\frac{\pi}{2} + k\pi \mid k \in \mathbb{Z}\}$, and $\cos(0) = 1$. From the Intermediate Value Theorem 3.29 it follows that $\sin'(x) = \cos(x) > 0$ for all $x \in (-\frac{\pi}{2}, \frac{\pi}{2})$. Thus, by Remark 5.45, the function

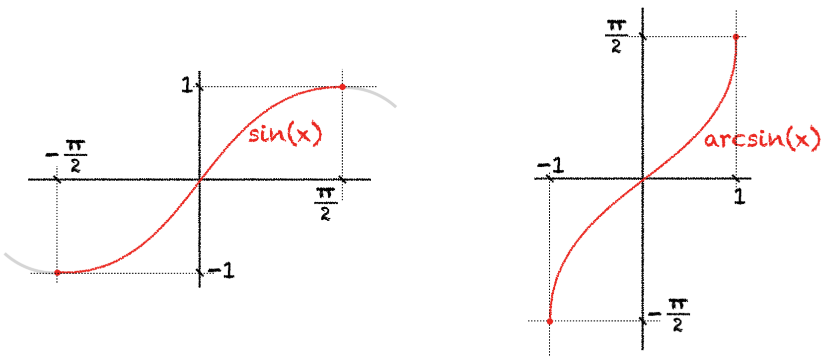
$$\sin : [-\tfrac{\pi}{2}, \tfrac{\pi}{2}] \to [-1, 1] \tag{5.12}$$

is strictly increasing and bijective (recall that $\sin(-\frac{\pi}{2}) = -1$ and $\sin(\frac{\pi}{2}) = 1$). Consequently, the sine function restricted to $[-\frac{\pi}{2}, \frac{\pi}{2}]$ has an inverse, which we express as

$$\arcsin : [-1, 1] \to [-\tfrac{\pi}{2}, \tfrac{\pi}{2}]$$

and call **arcsine**.

23 The following figure shows the graph of the sine function on $[-\frac{\pi}{2}, \frac{\pi}{2}]$, and of its inverse.



Remark 5.59. — Since $\sin'' = -\sin$, it follows that sin is convex in the interval $\left(-\frac{\pi}{2}, 0\right)$, and is concave in the interval $\left(0, \frac{\pi}{2}\right)$.

5.60. — According to Theorem 5.21, the arcsine is differentiable at s if the derivative of sine at $x = \arcsin(s)$ is not zero. Since the derivative of the sine $\sin' = \cos$ vanishes exactly at the boundary points of $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$, for $x \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ and $s = \sin(x)$, Theorem 5.21 yields

23

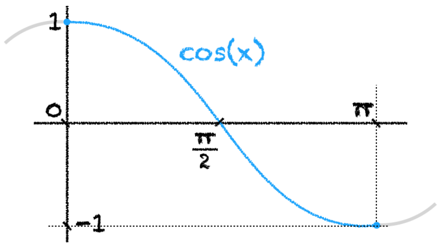
$$\arcsin'(s) = \frac{1}{\cos(x)} = \frac{1}{\sqrt{1-\sin^2(x)}} = \frac{1}{\sqrt{1-s^2}}$$

(recall that $\cos(x)$ is positive for $x \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$, thus $\cos(x) = \sqrt{1-\sin(x)^2}$).

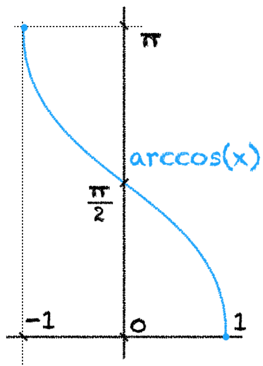
5.61. — The discussion above can be done analogously for the cosine. The cosine function is strictly monotonically decreasing in the interval $[0, \pi]$ and satisfies $\cos(0) = 1$ and $\cos(\pi) = -1$. In particular, the restricted cosine function

$$\cos : [0, \pi] \to [-1, 1]$$

is bijective.



24



The inverse figure is called **arccosine** and is written as

$$\arccos : [-1, 1] \to [0, \pi]$$

Just as for the arcsine, we can apply the differentiation rules for the inverse and get $s = \cos(x)$ for $x \in (0, \pi)$

$$\arccos'(s) = \frac{1}{-\sin(x)} = -\frac{1}{\sqrt{1-\cos^2(x)}} = -\frac{1}{\sqrt{1-s^2}}.$$

Remark 5.62. — Since $\cos'' = -\cos$, it follows that cos is concave in the interval $\left(0, \frac{\pi}{2}\right)$, and is convex in the interval $\left(\frac{\pi}{2}, \pi\right)$.

5.3.2 Tangent and Arc Tangent

5.63. — We consider the restriction $\tan : (-\frac{\pi}{2}, \frac{\pi}{2}) \to \mathbb{R}$ of the tangent function. Using the quotient rule in Corollary 5.18, we get

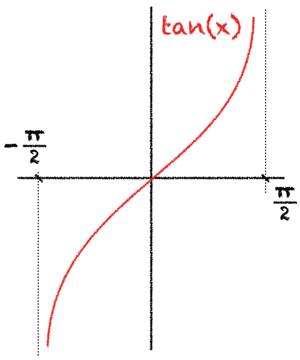
$$\tan'(x) = \left(\frac{\sin(x)}{\cos(x)}\right)' = \frac{\cos(x)\cos(x) - \sin(x)(-\sin(x))}{\cos(x)^2} = \frac{1}{\cos(x)^2}$$

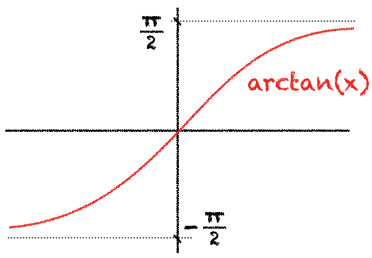
for all $x \in (-\frac{\pi}{2}, \frac{\pi}{2})$. In particular, $\tan : (-\frac{\pi}{2}, \frac{\pi}{2}) \to \mathbb{R}$ is strictly monotonically increasing. Furthermore

$$\lim_{x \to \frac{\pi}{2}-} \tan(x) = \lim_{x \to \frac{\pi}{2}-} \frac{\sin(x)}{\cos(x)} = +\infty \qquad \text{and} \qquad \lim_{x \to -\frac{\pi}{2}+} \tan(x) = \lim_{x \to -\frac{\pi}{2}+} \frac{\sin(x)}{\cos(x)} = -\infty$$

Thus, it follows from the Intermediate Value Theorem 3.29 that the tangent function $\tan : (-\frac{\pi}{2}, \frac{\pi}{2}) \to \mathbb{R}$ is bijective.

24





The inverse

$$\arctan : \mathbb{R} \to \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$$

is called **arctangent**. By Theorem 5.21 the arc tangent is differentiable, and for $x \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ and $s = \tan(x)$ it holds

$$\arctan'(s) = \frac{1}{\frac{1}{\cos^2(x)}} = \cos^2(x).$$

Since

$$s^2 = \frac{\sin^2(x)}{\cos^2(x)} = \frac{\sin^2(x) + \cos^2(x) - \cos^2(x)}{\cos^2(x)} = \frac{1}{\cos^2(x)} - 1,$$

it follows that $1 + s^2 = \frac{1}{\cos^2(x)}$, and therefore

$$\arctan'(s) = \frac{1}{1+s^2} \qquad \forall\, s \in \mathbb{R}.$$

5.64. — The cotangent and its inverse function behave similarly. The restriction $\cot|_{(0,\pi)} : (0,\pi) \to \mathbb{R}$ is strictly monotonically decreasing and bijective. The inverse

$$\operatorname{arccot} : \mathbb{R} \to (0, \pi)$$

is called **arccotangent** and has the derivative

$$\operatorname{arccot}'(s) = -\frac{1}{1+s^2} \qquad \forall\, s \in \mathbb{R}.$$

5.3.3 Hyperbolic Functions

We carry out the analogous discussion for the hyperbolic trigonometric functions

$$\sinh(x) = \frac{e^x - e^{-x}}{2}, \quad \cosh(x) = \frac{e^x + e^{-x}}{2}, \quad \tanh(x) = \frac{\sinh(x)}{\cosh(x)} = \frac{e^x - e^{-x}}{e^x + e^{-x}}.$$

5.65. — It holds $\sinh'(x) = \cosh(x) > 0$ for all $x \in \mathbb{R}$. Thus, according to Proposition 5.44, the hyperbolic sine is strictly monotonically increasing. Since $\lim_{x\to\infty} \sinh(x) = \infty$ and $\lim_{x\to-\infty} \sinh(x) = -\infty$, by the Intermediate Value Theorem 3.29 we get that

$$\sinh : \mathbb{R} \to \mathbb{R}$$

is strictly monotonically increasing and bijective. The inverse function

$$\operatorname{arsinh} : \mathbb{R} \to \mathbb{R}$$

24 is called the **Inverse Hyperbolic Sine**. According to the theorem on differentiability of the inverse function, arsinh is differentiable and it holds, for $x \in \mathbb{R}$ and $s = \sinh(x)$,

$$\operatorname{arsinh}'(s) = \frac{1}{\cosh(x)} = \frac{1}{\sqrt{1+\sinh^2(x)}} = \frac{1}{\sqrt{1+s^2}}.$$

The inverse hyperbolic sine has a closed form, unlike the inverse functions arcsin, arccos, and arctan. In fact, starting from the relation $\sinh(x) = s$ we have

$$\frac{e^x - e^{-x}}{2} = s \quad \Longrightarrow \quad e^{2x} - 2se^x - 1 = 0.$$

Calling $y = e^x$, we see that $y^2 - 2sy - 1 = 0$, hence

$$y = s \pm \sqrt{1+s^2}.$$

Since $y = e^x > 0$, the only admissible root is $s + \sqrt{1+s^2}$, therefore

$$e^x = y = s + \sqrt{1+s^2} \quad \Longrightarrow \quad x = \log\left(s + \sqrt{1+s^2}\right).$$

5.66. — The hyperbolic cosine satisfies $\cosh'(x) = \sinh(x)$ and $\cosh''(x) = \cosh(x) > 0$ for all $x \in \mathbb{R}$. In particular, the cosine hyperbolic is strictly convex by Corollary 5.53 and has a global minimum at 0 by Corollary 5.28 (since $0 = \cosh'(x) = \sinh(x)$ implies $x = 0$). For $x > 0$ we have $\cosh'(x) > 0$, so cosh is strictly monotonically increasing on $\mathbb{R}_{\geq 0}$. Since $\cosh(0) = 1$ and $\lim_{x\to\infty} \cosh(x) = +\infty$, it follows that

$$\cosh : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 1}$$

is strictly monotonically increasing and bijective. The inverse

$$\operatorname{arcosh} : \mathbb{R}_{\geq 1} \to \mathbb{R}_{\geq 0}$$

is called the **Inverse Hyperbolic Cosine**, is differentiable on $\mathbb{R}_{>1}$ and satisfies

$$\operatorname{arcosh}'(s) = \frac{1}{\sinh(x)} = \frac{1}{\sqrt{s^2 - 1}}$$

for $s > 1$ and $s = \cosh(x)$ with $x > 0$. Furthermore

24

$$\operatorname{arcosh}(s) = \log\left(s + \sqrt{s^2 - 1}\right)$$

for all $s > 1$. We leave the proof of the above properties to those interested.

5.67. — The **Inverse Hyperbolic Tangent** is the inverse function

$$\operatorname{artanh} : (-1, 1) \to \mathbb{R}, \qquad \operatorname{artanh}(x) = \tfrac{1}{2} \log\left(\frac{1 + x}{1 - x}\right)$$

of the strictly monotonically increasing bijection $\tanh : \mathbb{R} \to (-1, 1)$. According to Theorem 5.21, artanh is differentiable and the following holds:

$$\operatorname{artanh}'(s) = \frac{1}{1 - s^2}$$

for all $s \in (-1, 1)$.

Exercise 5.68. — Check all assertions made in Paragraphs 5.65, 5.66, and 5.67.

Chapter 6

The Riemann Integral

In this chapter we will take the idea of section 1.1 and extend it to the notion of the Riemann integral with the help of the supremum and the infimum, i.e. implicitly the completeness axiom.

6.1 Step Functions and their Integral

6.1.1 Decompositions and Step Functions

Interlude: Partitions

Two sets A, B are called **disjoint** if $A \cap B = \varnothing$. For a collection $\mathcal{A}$ of sets, we say that the sets in $\mathcal{A}$ are **pairwise disjoint** if for all $A_1, A_2 \in \mathcal{A}$ with $A_1 \neq A_2$ it holds $A_1 \cap A_2 = \varnothing$.
Let X be a set. A **partition** of X is a family $\mathcal{P}$ of non-empty pairwise disjoint subsets of X such that

$$X = \bigcup_{P \in \mathcal{P}} P.$$

24 In other words, sets $P \in \mathcal{P}$ are non-empty, and each element of X is an element of exactly one $P \in \mathcal{P}$.

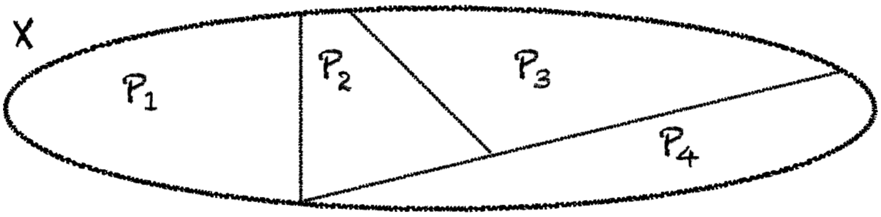


Figure 6.1: Schematic representation of a partition $\mathcal{P} = \{P_1, \ldots, P_4\}$ of a set X.

For the following discussion, we fix two real numbers $a < b$, and work with the compact interval $[a, b] \subset \mathbb{R}$.

Definition 6.1: Decomposition of an Interval

A **decomposition** of $[a, b]$ is a finite set of points

$$a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b$$

with $n \in \mathbb{N}$. The points $x_0, \ldots, x_n \in [a, b]$ are called the division points of the decomposition.

6.2. — Formally, a decomposition of $[a, b]$ is a finite subset of $[a, b]$ containing a and b, together with the listing of its elements in ascending order. A decomposition also induces a of partition of $[a, b]$, viz.

$$[a, b] = \{x_0\} \cup (x_0, x_1) \cup \{x_1\} \cup \cdots \cup (x_{n-1}, x_n) \cup \{x_n\}$$

which we will use implicitly from now on. A decomposition $a = y_0 < y_1 < \cdots < y_m = b$ is called a **refinement** of a decomposition $a = x_0 < x_1 < \cdots < x_n = b$ if

$$\{x_0, x_1, \ldots, x_n\} \subseteq \{y_0, y_1, \ldots, y_m\}.$$

The notion of refinement leads to an order relation on the set of all decompositions of $[a, b]$.
24 Note that any two decompositions of $[a, b]$ always have a common refinement (take the union of the points).

Definition 6.3: Step Functions

A function $f : [a, b] \to \mathbb{R}$ is called a **step function** if there exists a decomposition $a = x_0 < x_1 < \cdots < x_n = b$ of $[a, b]$ such that for $k = 1, 2, \ldots, n$ the restriction of f to the open interval (x_{k-1}, x_k) is constant. We also say that the function f is a step function with respect to the decomposition $a = x_0 < x_1 < \cdots < x_n = b$.

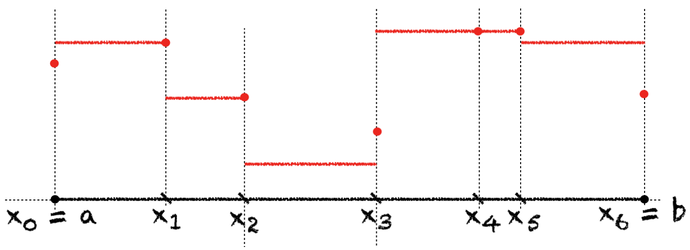


Figure 6.2: The graph of a step function on the interval $[a, b]$.

Proposition 6.4: Linearity of the Space of Step Functions

Let $f, g : [a, b] \to \mathbb{R}$ be step functions, and $\alpha, \beta \in \mathbb{R}$. Then $\alpha f + \beta g$ is also a step function.

Proof. There exist decompositions of $[a, b]$ with respect to which f and g are step functions. For these decompositions, there exists a common refinement $a = x_0 < x_1 < \cdots < x_n = b$ with respect to which f and g are step functions. Thus, the functions f and g are both constant on the open intervals (x_{k-1}, x_k), and consequently so is $\alpha f + \beta g$, which means that $\alpha f + \beta g$ is a step function with respect to $a = x_0 < x_1 < \cdots < x_n = b$. □

Example 6.5. — Constant functions are step functions.

Remark 6.6. — Just as in the proof of Proposition 6.4, one can show that the product of two step functions is again a step function. Also, we note that step functions are bounded, since they take finitely many values.

6.1.2 The Integral of a Step Function

Definition 6.7: Integral of Step Functions

Let $f : [a, b] \to \mathbb{R}$ be a step function with respect to a decomposition $a = x_0 < \cdots < x_n = b$ of $[a, b]$. We define the **integral** of f on $[a, b]$ as the real number

24

$$\int_a^b f(x)\,dx = \sum_{k=1}^{n} c_k(x_k - x_{k-1}) \tag{6.1}$$

where c_k denotes the value of f on the interval (x_{k-1}, x_k).

For the moment, in (6.1), the individual symbols $\int$ and dx have no meaning. Originally, the symbol $\int$ stands for an S for "sum", and the symbol dx indicates an "infinitesimal length", i.e. $x_k - x_{k-1}$ for an "infinitesimal fine" decomposition. The notation was introduced by Leibniz (1646-1716).

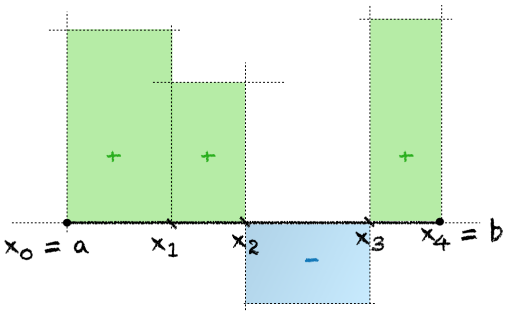


Figure 6.3: For a non-negative step function $f \geq 0$ we interpret (6.1) as the area of the, set $\{(x, y) \in \mathbb{R}^2 \mid a \leq x \leq b,\ 0 \leq y \leq f(x)\}$, and in general as the signed net area.

6.8. — The equation (6.1) defining the integral is not without problems. A priori, in fact, the right-hand side depends on the choice of a decomposition of the interval $[a, b]$. We must convince ourselves that this is only an apparent dependence. In other words, if $a = y_0 < \cdots < y_m = b$ is another decomposition of $[a, b]$ with respect to which f is a step function, then

$$\sum_{k=1}^{n} c_k(x_k - x_{k-1}) = \sum_{k=1}^{m} d_k(y_k - y_{k-1}) \tag{6.2}$$

where d_k denotes the constant value of f on the interval (y_{k-1}, y_k).

We consider this in three steps. In the first step, we assume that the decomposition $a = y_0 < \cdots < y_m = b$ is finer than $a = x_0 < \cdots < x_n = b$, and merely has an extra separation point $y_l \in (x_{\ell-1}, x_\ell) = (y_{l-1}, y_{l+1})$. But this means that the sums in (6.2) are the same, except that the term $c_l(x_l - x_{l-1})$ on the left becomes $d_l(y_l - y_{l-1}) + d_{l+1}(y_{l+1} - y_l)$ on the right, which does not change the value of the sum since $c_l = d_l = d_{l+1}$ holds. In a second step, using complete induction, we can conclude that (6.2) holds if $a = y_0 < \cdots < y_m = b$ is finer than $a = x_0 < \cdots < x_n = b$, with any number of additional separation points. Finally, we know that two decompositions of $[a, b]$ always have a common refinement. We can thus show (6.2) in full generality by comparing both sides with the corresponding sum for a common refinement of the given decompositions.

Proposition 6.9: Linearity of Integral of Step Functions

Let $f, g : [a, b] \to \mathbb{R}$ be step functions, and let $\alpha, \beta \in \mathbb{R}$. Then

24

$$\int_a^b (\alpha f + \beta g)(x)\,dx = \alpha \int_a^b f(x)\,dx + \beta \int_a^b g(x)\,dx$$

Proof. We have already shown in Proposition 6.4 that $\alpha f + \beta g$ is a step function. Let $a \leq x_0 < \cdots < x_n = b$ be a decomposition such that the functions f and g (and consequently $\alpha f + \beta g$) are constant on the intervals (x_{k-1}, x_k). If c_k is the value of f and d_k the value of g on (x_{k-1}, x_k), then $\alpha c_k + \beta d_k$ is the value of $\alpha f + \beta g$ on (x_{k-1}, x_k). Therefore

$$\begin{aligned}
\int_a^b (\alpha f + \beta g)(x)\,dx &= \sum_{k=1}^{n} (\alpha c_k + \beta d_k)(x_k - x_{k-1}) \\
&= \alpha \sum_{k=1}^{n} c_k(x_k - x_{k-1}) + \beta \sum_{k=1}^{n} d_k(x_k - x_{k-1}) \\
&= \alpha \int_a^b f(x)\,dx + \beta \int_a^b g(x)\,dx,
\end{aligned}$$

as desired. □

Proposition 6.10: Monotonicity for Integral of Step Functions

Let $f, g : [a, b] \to \mathbb{R}$ be step functions such that $f \leq g$. Then

$$\int_a^b f(x)\,dx \leq \int_a^b g(x)\,dx.$$

Proof. As in the proofs of Proposition 6.4 and Proposition 6.9, we can find a decomposition $a = x_0 < \cdots < x_n = b$ such that f and g are constant on the intervals (x_{k-1}, x_k). We again write c_k for the value of f and d_k for the value of g on (x_{k-1}, x_k). Now, because $f \leq g$ holds, i.e., $f(x) \leq g(x)$ for all $x \in [a,b]$, we get $c_k \leq d_k$ for all $k \in \{1, \ldots, n\}$. Therefore

$$\int_a^b f(x)\,dx = \sum_{k=1}^{n} c_k(x_k - x_{k-1}) \leq \sum_{k=1}^{n} d_k(x_k - x_{k-1}) = \int_a^b g(x)\,dx.$$

□

24 EXERCISE 6.11. — Let $[a,b], [b,c]$ be two bounded and closed intervals and let $f_1 : [a,b] \to \mathbb{R}$ and $f_2 : [b,c] \to \mathbb{R}$ be step functions. Show that the function

$$f : [a,c] \to \mathbb{R}, \qquad x \mapsto \begin{cases} f_1(x) & \text{if } x \in [a,b) \\ f_2(x) & \text{if } x \in [b,c] \end{cases}$$

is a step function on $[a,c]$. Then prove that the integral of f is given by

$$\int_a^c f(x)\,dx = \int_a^b f_1(x)\,dx + \int_b^c f_2(x)\,dx.$$

Finally, show that every step function on $[a,c]$ is of the form described above.

6.2 Definition and First Properties of the Riemann Integral

As in the last section, we consider functions on a compact interval $[a,b] \subset \mathbb{R}$. To alleviate the notation, we write $\mathcal{SF}$ for the set of step functions on $[a,b]$. Also, we often write $\int_a^b f\,dx$ in place of $\int_a^b f(x)\,dx$.

6.2.1 Integrability of Real-valued Functions

The following definition of integrability is a variant of Riemann's definition, which goes back to the French mathematician Jean-Gaston Darboux (1842–1917).

Definition 6.12: Lower and Upper Sums

Let $f : [a,b] \to \mathbb{R}$ be a function. Then we define the sets of **lower sums** $\mathcal{L}(f) \subset \mathbb{R}$ and **upper sums** $\mathcal{U}(f) \subset \mathbb{R}$ of f by

$$\mathcal{L}(f) = \left\{ \int_a^b \ell\,dx \;\middle|\; \ell \in \mathcal{SF} \text{ and } \ell \le f \right\} \qquad \mathcal{U}(f) = \left\{ \int_a^b u\,dx \;\middle|\; u \in \mathcal{SF} \text{ and } f \le u \right\}.$$

Note that, if f is bounded, then these sets are non-empty. Indeed, if $|f| \le M$, then
24 $\ell = -M \in \mathcal{L}(f)$ and $u = M \in \mathcal{U}(f)$.

For $\ell, u \in \mathcal{SF}$ with $\ell \le f \le u$, Proposition 6.10 implies that

$$\int_a^b \ell\,dx \le \int_a^b u\,dx,$$

therefore $s \le t$ for all $s \in \mathcal{L}(f)$ and $t \in \mathcal{U}(f)$. In particular, we have the inequality

$$\sup \mathcal{L}(f) \le \inf \mathcal{U}(f)$$

if f is bounded.

Definition 6.13: Riemann Integral

A bounded function $f : [a,b] \to \mathbb{R}$ is called **Riemann integrable** if $\sup \mathcal{L}(f) = \inf \mathcal{U}(f)$. In this case, this common value is called the **Riemann integral** of f, and is expressed as

$$\int_a^b f\,dx = \sup \mathcal{L}(f) = \inf \mathcal{U}(f).$$

6.14. — We call a the **lower (integration) limit** and b the **upper (integration) limit**, and the function f the **integrand** of the integral $\int_a^b f\,dx$. If $f \geq 0$ is Riemann integrable, then we interpret the number $\int_a^b f\,dx$ as the **area** of the set

$$\{(x,y) \in \mathbb{R}^2 \mid a \leq x \leq b, \text{ and } 0 \leq y \leq f(x)\}.$$

Remark 6.15. — Since, for the time being, we only know Riemann integrability and the Riemann integral, we will simply take the liberty of speaking of integrability and integral. Note however, that besides Riemann integration theory, there is another important such theory called **Lebesgue integral**.

Proposition 6.16: Riemann Integrability Condition

Let $f : [a,b] \to \mathbb{R}$ be bounded. The function f is Riemann integrable exactly if for every $\varepsilon > 0$ there exist step functions ℓ and u that satisfy

$$\ell \leq f \leq u \qquad \text{and} \qquad \int_a^b (u-\ell)\,dx < \varepsilon.$$

In such a case

25 $$\left|\int_a^b f\,dx - \int_a^b \ell\,dx\right| < \varepsilon, \qquad \left|\int_a^b u\,dx - \int_a^b f\,dx\right| < \varepsilon.$$

Proof. Let A and B be nonempty subsets of $\mathbb{R}$ with the property that $a \leq b$ holds for all $a \in A$ and all $b \in B$. Then $\sup A \leq \inf B$, and equality $\sup A = \inf B$ holds exactly if for every $\varepsilon > 0$ there is an $a \in A$ and an $b \in B$ with $b - a < \varepsilon$. This reasoning holds in particular for the sets $\mathcal{L}(f)$ and $\mathcal{U}(f)$. The implications

$$\begin{aligned} f \text{ is Riemann integrable} &\iff \sup \mathcal{L}(f) = \inf \mathcal{U}(f) \\ &\iff \forall \varepsilon > 0\, \exists s \in \mathcal{L}(f), t \in \mathcal{U}(f) \text{ with } t - s < \varepsilon \\ &\iff \forall \varepsilon > 0\, \exists \ell, u \in \mathcal{SF}, \text{ s.t. } \ell \leq f \leq u \text{ and } \int_a^b (u-\ell)\,dx < \varepsilon \end{aligned}$$

prove the proposition.

The last inequalities follow from

$$\int_a^b \ell\,dx \leq \int_a^b f\,dx \leq \int_a^b u\,dx \qquad \text{and} \qquad \int_a^b (u-\ell)\,dx < \varepsilon$$

□

6.17. — It is good to know that the Riemann integral is a generalisation of the integral of step functions, and in this sense we can simply speak of the Riemann integral of a step function. This is the subject of exercise 6.18.

Exercise 6.18. — Let $f : [a, b] \to \mathbb{R}$ be a step function. Show that f is Riemann integrable and that the Riemann integral of f is equal to the integral of f as a step function.

Exercise 6.19. — Repeat the proof of Proposition 1.1 and show, in the language of this section, that $f : [0, 1] \to \mathbb{R}$, $x \mapsto x^2$ is Riemann integrable with $\int_0^1 x^2 dx = \frac{1}{3}$. Also,

$$\mathcal{L}(f) = \big(-\infty, \tfrac{1}{3} \big) \qquad \text{and} \qquad \mathcal{U}(f) = \big(\tfrac{1}{3}, \infty \big).$$

Example 6.20. — Not all functions are Riemann integrable, as the following example shows. Consider $f : [0, 1] \to \mathbb{R}$ defined as

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

We claim that this function is not Riemann integrable.

To see this, let $u \in \mathcal{SF}$ with $f \le u$, and let $0 = x_0 < \cdots < x_n = 1$ be a decomposition such that u is constant on every interval (x_{k-1}, x_k), with value c_k. Since $\mathbb{Q}$ is dense in $\mathbb{R}$, there exists a $x \in (x_{k-1}, x_k)$ with $x \in \mathbb{Q}$. Because of $f \le u$, $1 = f(x) \le u(x) = c_k$ holds. Thus

25

$$\int_0^1 u(x)\, dx = \sum_{k=1}^{n} c_k (x_k - x_{k-1}) \ge \sum_{k=1}^{n} (x_k - x_{k-1}) = x_n - x_0 = 1$$

using telescopic sums. Thus, the upper integral of f is given by 1, since the step function with constant value 1 has integral 1 and u was arbitrary.

Similarly, one shows that the lower integral of f is given by 0. Thus, f is not Riemann integrable.

6.2.2 Linearity and Monotonicity of the Riemann Integral

Theorem 6.21: Linearity of the Riemann Integral

If $f, g : [a, b] \to \mathbb{R}$ are integrable and $\alpha, \beta \in \mathbb{R}$. Then $\alpha f + \beta g$ is integrable with integral

$$\int_a^b (\alpha f + \beta g)\, dx = \alpha \int_a^b f\, dx + \beta \int_a^b g\, dx.$$

Proof. Given $\varepsilon > 0$, thanks to Proposition 6.16 we can find step functions ℓ_1, ℓ_2, u_1, u_2 such that

$$\ell_1 \le f \le u_1, \quad \ell_2 \le g \le u_2, \quad \int_a^b (u_1 - \ell_1)\, dx < \varepsilon, \quad \int_a^b (u_2 - \ell_2)\, dx < \varepsilon,$$

$$\left|\int_a^b f\,dx - \int_a^b \ell_1\,dx\right| < \varepsilon, \qquad \left|\int_a^b g\,dx - \int_a^b \ell_2\,dx\right| < \varepsilon$$

Assume first that $\alpha, \beta \geq 0$. Then

$$\alpha\ell_1 \leq \alpha f \leq \alpha u_1, \quad \beta\ell_2 \leq \beta g \leq \beta u_2 \implies \alpha\ell_1 + \beta\ell_2 \leq \alpha f + \beta g \leq \alpha u_1 + \beta u_2,$$

and

$$\int_a^b \big((\alpha u_1 + \beta u_2) - (\alpha\ell_1 + \beta\ell_2)\big)\,dx = \alpha\int_a^b (u_1 - \ell_1)\,dx + \beta\int_a^b (u_2 - \ell_2)\,dx < (\alpha+\beta)\varepsilon.$$

Since ε is arbitrary, this shows that $\alpha f + \beta g$ is integrable. Also, by the traingle inequality and Proposition 6.9 applied to ℓ_1 and ℓ_2, we get

$$\begin{aligned}\left|\int_a^b (\alpha f + \beta g)\,dx - \alpha\int_a^b f\,dx - \beta\int_a^b g\,dx\right| &\leq \left|\int_a^b (\alpha f + \beta g)\,dx - \int_a^b (\alpha\ell_1 + \beta\ell_2)\,dx\right| \\ &\quad + \underbrace{\left|\int_a^b (\alpha\ell_1 + \beta\ell_2)\,dx - \alpha\int_a^b \ell_1\,dx - \beta\int_a^b \ell_2\,dx\right|}_{=0} \\ &\quad + \alpha\left|\int_a^b \ell_1\,dx - \int_a^b f\,dx\right| + \beta\left|\int_a^b \ell_2\,dx - \int_a^b g\,dx\right| \\ &\leq (\alpha+\beta)\varepsilon + \alpha\varepsilon + \beta\varepsilon = 2(\alpha+\beta)\varepsilon.\end{aligned}$$

25

which implies, again from the arbitrariness of ε, that $\int_a^b (\alpha f + \beta g)\,dx = \alpha\int_a^b f\,dx + \beta\int_a^b g\,dx$.

The case when α or β is negative is analogous, but one needs to reverse some inequalities. For instance, if $\alpha \geq 0$ but $\beta < 0$ then

$$\alpha\ell_1 \leq \alpha f \leq \alpha u_1, \quad \beta\ell_2 \geq \beta g \geq \beta u_2 \implies \alpha\ell_1 + \beta u_2 \leq \alpha f + \beta g \leq \alpha u_1 + \beta\ell_2,$$

and

$$\int_a^b \big((\alpha u_1 + \beta\ell_2) - (\alpha\ell_1 + \beta u_2)\big)\,dx = \alpha\int_a^b (u_1 - \ell_1)\,dx + \beta\int_a^b (\ell_2 - u_2)\,dx < (\alpha + |\beta|)\varepsilon.$$

Since ε is arbitrary, this shows that $\alpha f + \beta g$ is integrable, and analogously one proves that $\int_a^b (\alpha f + \beta g)\,dx = \alpha\int_a^b f\,dx + \beta\int_a^b g\,dx$. □

Exercise 6.22. — Let $f : [a,b] \to \mathbb{R}$ be Riemann integrable. Let $f^* : [a,b] \to \mathbb{R}$ be a function obtained by changing the value of f only at finitely many points in $[a,b]$. Show that f^* is Riemann integrable and has the same Riemann integral as f.

Proposition 6.23: Monotonicity of the Riemann Integral

Let $f, g : [a,b] \to \mathbb{R}$ be integrable. If $f \leq g$, then $\int_a^b f\,dx \leq \int_a^b g\,dx$.

Proof. For any step function $u : [a,b] \to \mathbb{R}$, if $u \le f$ then $u \le g$. This implies that $\mathcal{L}(f) \subseteq \mathcal{L}(g)$, and therefore

$$\int_a^b f\,dx = \sup \mathcal{L}(f) \le \sup \mathcal{L}(g) = \int_a^b g\,dx,$$

as desired. □

6.24. — Let $f : [a,b] \to \mathbb{R}$ be a function. We define functions f^+, f^-, and $|f|$ on $[a,b]$ by

$$f^+(x) = \max\{0, f(x)\}, \qquad f^-(x) = -\min\{0, f(x)\}, \qquad |f|(x) = |f(x)| \qquad \forall\, x \in [a,b].$$

The function f^+ is the **positive part**, f^- is the **negative part**, and $|f|$ is the **absolute value** of the function f. One can check that

$$f = f^+ - f^-, \qquad |f| = f^+ + f^-, \qquad f^+ = \frac{|f| + f}{2}, \qquad f^- = \frac{|f| - f}{2}.$$

In addition,

$$f \le g \implies f^+ \le g^+ \qquad \text{and} \qquad f \le g \implies f^- \ge g^-.$$

25

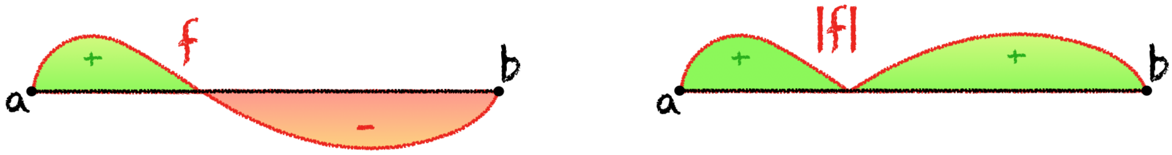


Figure 6.4: The graph of af function f is shown on the left, the graph of the corresponding function $|f|$ is shown on the right. The integral $\int_a^b f dx$ describes a net area and $\int_a^b |f| dx$ a gross area.

Theorem 6.25: Triangle Inequality for Riemann Integral

Let $f : [a,b] \to \mathbb{R}$ be an integrable function. Then f^+, f^-, and $|f|$ are also integrable, and

$$\left| \int_a^b f\,dx \right| \le \int_a^b |f|\,dx.$$

Proof. We start by showing that f^+ is Riemann integrable. Let $\varepsilon > 0$. Since f is integrable, there exist step functions ℓ and u with the property

$$\ell \le f \le u \qquad \text{and} \qquad \int_a^b (u - \ell)\,dx < \varepsilon.$$

The functions ℓ^+ and u^+ are also step functions, and $\ell^+ \le f^+ \le u^+$ holds. Since $u - \ell$ is non-negative, $u - \ell = (u - \ell)^+$ holds. Also, considering all possible cases (i.e., $u(x) \ge \ell(x) \ge 0$, $u(x) \ge 0 > \ell(x)$, or $0 > u(x) \ge \ell(x)$), one checks that

$$(u - \ell)^+(x) \ge u^+(x) - \ell^+(x) \qquad \forall\, x \in [a,b].$$

Therefore

$$\int_a^b (u^+ - \ell^+)\,dx \le \int_a^b (u-\ell)^+\,dx = \int_a^b (u-\ell)\,dx < \varepsilon$$

which shows that f^+ is integrable. Theorem 6.21 implies that $f^- = f^+ - f$ and $|f| = f^+ + f^-$ are also integrable. Finally

$$\left|\int_a^b f\,dx\right| = \left|\int_a^b f^+\,dx - \int_a^b f^-\,dx\right| \le \int_a^b f^+\,dx + \int_a^b f^-\,dx = \int_a^b |f|\,dx$$

where we used Theorem 6.21, as well as $\int_a^b f^+\,dx \ge 0$ and $\int_a^b f^-\,dx \ge 0$. □

Exercise 6.26. — Let $a < b < c$ be real numbers. Show that a function $f : [a,c] \to \mathbb{R}$ is integrable exactly when $f|_{[a,b]}$ and $f|_{[b,c]}$ are integrable, and that in this case

$$\int_a^c f\,dx = \int_a^b f|_{[a,b]}\,dx + \int_b^c f|_{[b,c]}\,dx.$$

Exercise 6.27. — Let $f : [a,b] \to \mathbb{R}$ be integrable, and $\lambda > 0$ be a real number. Let $g : [\lambda a, \lambda b] \to \mathbb{R}$ be the function given by $g(x) = f(\lambda^{-1}x)$. Show that g is integrable, and that

25

$$\lambda \int_a^b f\,dx = \int_{\lambda a}^{\lambda b} g\,dx.$$

Exercise 6.28. — Let $f : [a,b] \to \mathbb{R}$ be an integrable function. Show that the function $F : [a,b] \to \mathbb{R}$ given by

$$F(x) = \int_a^x f(t)\,dt$$

is continuous.

Exercise 6.29. — Let $\mathcal{C}$ be the space of continuous functions on $[a,b]$, and let $I : \mathcal{C} \to \mathbb{R}$ be the integration.

$$I(f) = \int_a^b f\,dx$$

Show that the function I is continuous, in the following sense: for all $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$|f(x) - g(x)| < \delta \quad \forall\, x \in D \implies |I(f) - I(g)| \le \varepsilon.$$

Exercise 6.30. — Let $f : [0,1] \to \mathbb{R}$ be an integrable function and $\varepsilon > 0$. Show that there exists a continuous function $g : [0,1] \to \mathbb{R}$ such that

$$\int_0^1 |f(x) - g(x)|\,dx < \varepsilon. \tag{6.3}$$

6.3 Integrability Theorems

6.3.1 Integrability of Monotone Functions

We consider as before a compact interval $[a, b] \subset \mathbb{R}$ for real numbers a, b with $a < b$. We note that monotone functions $f : [a, b] \to \mathbb{R}$ are bounded, as e.g. $f(a)$ is a lower bound and $f(b)$ is an upper bound if f is monotone increasing.

Theorem 6.31: Monotone Functions are Integrable

Every monotone function $f : [a, b] \to \mathbb{R}$ is Riemann integrable.

Proof. Without loss of generality, $f : [a, b] \to \mathbb{R}$ is increasing – if not, replace f with $-f$ and apply Proposition 6.21. We want to apply Proposition 6.16, that is, for a given $\varepsilon > 0$ we want to find two step functions $\ell, u \in \mathcal{SF}$ such that $\ell \leq f \leq u$ and $\int_a^b (u - \ell)\, dx < \varepsilon$.

We construct ℓ, u using a natural number $n \in \mathbb{N}$ which we will specify later, and the decomposition

$$a = x_0 < x_1 < \ldots < x_n = b$$

of $[a, b]$ given by $x_k = a + \frac{b-a}{n} k$ for $k \in \{0, \ldots, n\}$. Let ℓ, u be given, for $x \in [a, b]$, by

25

$$\ell(x) = \begin{cases} f(x_{k-1}) & \text{if } x \in [x_{k-1}, x_k) \text{ for } k \in \{1, \ldots, n\} \\ f(b) & \text{if } x = b \end{cases},$$

$$u(x) = \begin{cases} f(a) & \text{if } x = a \\ f(x_k) & \text{if } x \in (x_{k-1}, x_k] \text{ for } k \in \{1, \ldots, n\} \end{cases}.$$

Since f is increasing, $\ell \leq f \leq u$ holds. Indeed, for $x \in [a, b]$ either $x = b$, where $\ell(x) = f(x)$, or there is a $k \in \{1, \ldots, n\}$ with $x \in [x_{k-1}, x_k)$. In the latter case we get $\ell(x) = f(x_{k-1}) \leq f(x)$, and thus $\ell \leq f$ holds. An analogous argument yields $f \leq u$.

Recalling that $x_n = b$ and $x_0 = a$, this yields

$$\begin{aligned} \int_a^b (u - \ell)\, dx &= \sum_{k=1}^{n} \big(f(x_k) - f(x_{k-1})\big)(x_k - x_{k-1}) = \sum_{k=1}^{n} \big(f(x_k) - f(x_{k-1})\big) \frac{b-a}{n} \\ &= \big(f(x_n) - f(x_{n-1}) + f(x_{n-1}) - f(x_{n-1}) + \ldots + f(x_1) - f(x_0)\big) \frac{b-a}{n} \\ &= \big(f(b) - f(a)\big) \frac{b-a}{n}. \end{aligned}$$

Following Archimedes' principle, we can now choose $n \in \mathbb{N}$ such that $\int_a^b (u - \ell)\, dx < \varepsilon$. Thus, it follows from Proposition 6.16 that f is Riemann-integrable. □

Exercise 6.32. — Show that the function $x \in [0, 1] \mapsto \sqrt{1 - x^2} \in \mathbb{R}$ is Riemann integrable.

Using the addition property in Exercise 6.26, the statement of Theorem 6.31 can be extended to functions that are only piecewise monotonic.

Definition 6.33: Piecewise Monotone Functions

A function $f : [a, b] \to \mathbb{R}$ is called **piecewise monotone** if there is a decomposition

$$a = x_0 < x_1 < \ldots < x_n = b$$

of $[a, b]$ such that $f|_{(x_{k-1}, x_k)}$ is monotone for all $k \in \{1, \ldots, n\}$.

Corollary 6.34: Piecewise Monotone Functions are Integrable

Every piecewise monotone bounded function $f : [a, b] \to \mathbb{R}$ is Riemann integrable.

Proof. (Extra Material) This follows from Theorem 6.31, and Exercises 6.26 and 6.22. □

6.3.2 Integrability of Continuous Functions

In this section, using boundedness and uniform continuity of continuous functions on compact
25 intervals (Theorems 3.39 and 3.46), we show the following result.

Theorem 6.35: Continuous Functions are Integrable

Every continuous function $f : [a, b] \to \mathbb{R}$ is integrable.

Proof. Let $f : [a, b] \to \mathbb{R}$ be continuous, and $\varepsilon > 0$. By Theorem 3.46 f is uniformly continuous, so there exists $\delta > 0$ such that

$$|x - y| < \delta \implies |f(x) - f(y)| < \varepsilon$$

holds for all $x, y \in [a, b]$.

Now, let $a = x_0 < x_1 < \ldots < x_n = b$ be a decomposition of $[a, b]$ such that $x_k - x_{k-1} < \delta$ for all k, and for each $k \in \{1, \ldots, n\}$ define

$$c_k = \min\{f(x) \mid x_{k-1} \le x \le x_k\} \qquad \text{and} \qquad d_k = \max\{f(x) \mid x_{k-1} \le x \le x_k\}$$

where we used Theorem 3.42 for the existence of these extrema.

We note that, for all $k \in \{1, \ldots, n\}$, the inequality $d_k - c_k < \varepsilon$ holds. Indeed, if $y_k, z_k \in [x_{k-1}, x_k]$ as points of minimum and maximum, namely $f(y_k) = c_k$ and $f(z_k) = d_k$, then because $|y_k - z_k| \le x_k - x_{k-1} < \delta$, the uniform continuity implies that $d_k - c_k < \varepsilon$.

We now define step functions ℓ, u by

$$\ell(x) = \begin{cases} c_k & \text{if } x \in [x_{k-1}, x_k) \\ c_n & \text{if } x = b \end{cases} \qquad \text{and} \qquad u(x) = \begin{cases} d_k & \text{if } x \in [x_{k-1}, x_k) \\ d_n & \text{if } x = b \end{cases}$$

for $x \in [a, b]$. Since $c_k \le d_k$, we see that $\ell \le f \le u$. Also, because $d_k - c_k < \varepsilon$, it follows that

$$\int_a^b (u - \ell)\, dx = \sum_{k=1}^{n} (d_k - c_k)(x_k - x_{k-1}) < \varepsilon \sum_{k=1}^{n} (x_k - x_{k-1}) = \varepsilon(b - a).$$

Since $\varepsilon > 0$ is arbitrary, f is integrable. □

Again by the addition property in Exercise 6.26, the statement of Theorem 6.35 can be extended to functions that are only piecewise continuous.

Definition 6.36: Piecewise Continuous Functions

A function $f : [a, b] \to \mathbb{R}$ is called **piecewise continuous** if there is a decomposition

$$a = x_0 < x_1 < \ldots < x_n = b$$

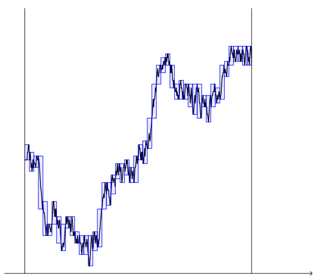
of $[a, b]$ such that $f|_{(x_{k-1}, x_k)}$ is continuous for all $k \in \{1, \ldots, n\}$, and both limits $\lim_{x \to x_{k-1}^+} f(x)$ and $\lim_{x \to x_k^-} f(x)$ exist. In other words, each function $f|_{(x_{k-1}, x_k)}$ can
25 be extended to a continuous function on $[x_{k-1}, x_k]$.

Corollary 6.37: Piecewise Continuous Functions are Integrable

Every piecewise continuous function $f : [a, b] \to \mathbb{R}$ is Riemann integrable.

Proof. (Extra Material) This follows from Theorem 6.35 applied to the continuous extension of $f|_{(x_{k-1}, x_k)}$ on $[x_{k-1}, x_k]$, and Exercises 6.26 and 6.22. □

6.38. — Most "common" functions are piecewise continuous or piecewise monotone, and in particular integrable according to Theorem 6.31 or according to Theorem 6.35. We note that there exist functions that are continuous on their domain of definition but are not monotone on any open subinterval.



Applet 6.39 (Integrability of a "shaky" function)**.** *We see that a continuous but shaky function as in the graph shown is also Riemann integrable. We can also note that the program GeoGebra sometimes has problems with the function used, and some of the displayed lower sums or upper sums are actually not displayed and calculated correctly. Regardless of this, we have proven the Riemann integrability, so we should not worry if there are some computational errors using*
25 *GeoGebra.*

EXERCISE 6.40. — Let $f : [a, b] \to \mathbb{R}$ be a continuous function. Prove that

$$f = 0 \iff \int_a^b |f(x)|\,dx = 0.$$

6.3.3 Integration and Sequences of Functions

Assume that a sequence of integrable functions $(f_n)_{n=0}^{\infty}$, $f_n : [a, b] \to \mathbb{R}$, converges pointwise or uniformly to a function $f : [a, b] \to \mathbb{R}$. Can we conclude that f is integrable and if so, does the equality

$$\lim_{n\to\infty} \int_a^b f_n \; dx = \int_a^b f \; dx$$

hold?

One can show that the pointwise limit of integrable functions is not integrable. More importantly, as the following example shows, if a sequence of integrable functions $(f_n)_{n\in\mathbb{N}}$ converges pointwise to an integrable function f, then the limit of the integrals does not necessarily coincide with the integral of the limit.

26 EXAMPLE 6.41. — Let $D = [0, 1]$ and let $f_n : D \to \mathbb{R}$ be given by.

$$f_n(x) = \begin{cases} n^2 x & \text{for } x \in \left[0, \frac{1}{2n}\right] \\ n^2(\frac{1}{n} - x) & \text{for } x \in \left[\frac{1}{2n}, \frac{1}{n}\right] \\ 0 & \text{for } x \in \left[\frac{1}{n}, 1\right] \end{cases}$$

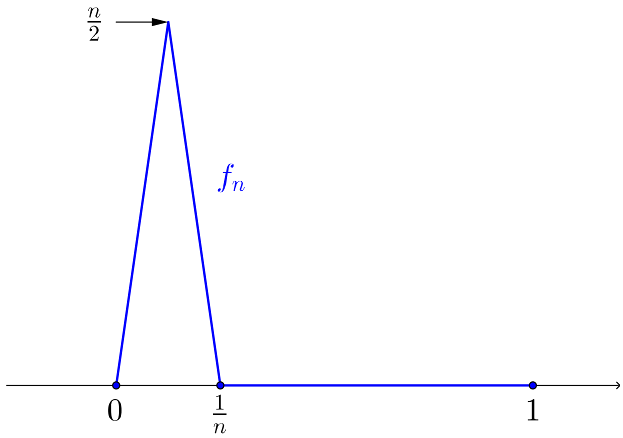
for $x \in [0, 1]$ and $n \in \mathbb{N}$. Then f_n is continuous and, in particular, integrable. Also, its integral is equal to the area of the triangle in the figure, which is $\frac{1}{2}\frac{1}{n}\frac{n}{2} = \frac{1}{4}$.

Note that the sequence $(f_n)_{n=0}^{\infty}$ converges pointwise to the constant function $f(x) = 0$. Indeed, $f_n(0)$ for every n, so $f_n(0) \to 0$. Also, for every $x > 0$ it follows that $f_n(x) = 0$ for every $n > \frac{1}{x}$ (since this is equivalent to $x > \frac{1}{n}$), so again $f_n(x) \to 0$.

However, for all $n \in \mathbb{N}$ the following is true:

$$\int_0^1 f_n(x)\,dx = \frac{1}{4} \neq 0 = \int_0^1 f(x)\,dx.$$

So, the limit of the integrals is not equal to the integral of the limit function, although all functions f_n and f are continuous.



On the other hand, if the convergence $f_n \to f$ is uniform, then both questions can be answered affirmatively:

Theorem 6.42: Uniform Convergence and Riemann Integral Commute

26 *Let $(f_n)_{n=0}^{\infty}$, with $f_n : [a,b] \to \mathbb{R}$, be a sequence of integrable functions converging uniformly to a function $f : [a,b] \to \mathbb{R}$. Then f is integrable and*

$$\int_a^b f \, dx = \lim_{n\to\infty} \int_a^b f_n \, dx \,. \tag{6.4}$$

Proof. Let $\varepsilon > 0$. By uniform convergence, there exists $N \in \mathbb{N}$ such that $f - \varepsilon \le f_n \le f + \varepsilon$ for all $n \ge N$. Since f_N is Riemann integrable by assumption, there exist step functions ℓ, u on $[a,b]$ with $\ell \le f_N \le u$ and $\int_a^b (u - \ell) \, dx < \varepsilon$. It follows that, for the step functions $\hat{\ell} = \ell - \varepsilon$ and $\hat{u} = u + \varepsilon$, we have

$$\hat{\ell} \le f \le \hat{u} \qquad \text{and} \qquad \int_a^b (\hat{u} - \hat{\ell}) \, dx < \varepsilon(2b - 2a + 1).$$

Since $\varepsilon > 0$ is arbitrary, the Riemann integrability of f follows from Proposition 6.16.

From the monotonicity and triangle inequality for the Riemann integral in Propositions 6.23 and 6.25, we also have

$$\left| \int_a^b f \, dx - \int_a^b f_n \, dx \right| = \left| \int_a^b (f - f_n) \, dx \right| \le \int_a^b |f - f_n| \, dx \le \varepsilon(b-a) \qquad \forall n \ge N.$$

This proves (6.4) and concludes the proof. □

Chapter 7

The Derivative and the Riemann Integral

In this chapter we will examine the connections between the Riemann integral from chapter 6 and the derivative from chapter 5. These connections are of fundamental importance for the further theory.

7.1 The Fundamental Theorem of Calculus

We specify for this section a compact interval $I \subseteq \mathbb{R}$ that is non-empty and does not consist of a single point. For brevity, we write integrable instead of Riemann-integrable.

7.1.1 The Fundamental Theorem

Definition 7.1: Primitive Function

Let $I \subset \mathbb{R}$ be an interval and $f : I \to \mathbb{R}$ a function. Any differentiable function $F : I \to \mathbb{R}$ such that $F' = f$ is called a **primitive** of f.

26

Remark 7.2. — In general, a primitive function needs neither exist (see the following exercise) nor be unique.

Exercise 7.3. — Show that there is no differentiable function $f : \mathbb{R} \to \mathbb{R}$ such that $f'(x) = \operatorname{sgn}(x)$ holds for all $x \in \mathbb{R}$.

Theorem 7.4: Fundamental Theorem of Calculus

Let $f : [a,b] \to \mathbb{R}$ be continuous. Then, for every $C \in \mathbb{R}$, the function $F : [a,b] \to \mathbb{R}$ defined as

$$F(x) = \int_a^x f(t)\,dt + C \tag{7.1}$$

is a primitive of f.
Moreover, any primitive $F : [a,b] \to \mathbb{R}$ has this form for some constant $C \in \mathbb{R}$.

Proof. By Theorem 6.35, f is integrable. We claim that the function

$$x \mapsto F(x) = \int_a^x f(t)\,dt + C$$

is differentiable on $[a,b]$ and a primitive function of f. Let $x_0 \in [a,b]$ and $\varepsilon > 0$. By the continuity of f, there exists $\delta > 0$ such that $|f(x) - f(x_0)| < \varepsilon$ for all $x \in [a,b]$ with $|x - x_0| < \delta$.

Now, for $x \in (x_0, x_0 + \delta) \cap [a,b]$, it follows from Exercise 6.26 that

$$\begin{aligned}\left|\frac{F(x)-F(x_0)}{x-x_0} - f(x_0)\right| &= \left|\frac{1}{x-x_0}\left(\int_a^x f(t)\,dt - \int_a^{x_0} f(t)\,dt\right) - f(x_0)\right| \\ &= \left|\frac{1}{x-x_0}\int_{x_0}^x f(t)\,dt - f(x_0)\right|.\end{aligned}$$

26

Noticing that

$$f(x_0) = f(x_0)\frac{1}{x-x_0}\int_{x_0}^x dt = \frac{1}{x-x_0}\int_{x_0}^x f(x_0)\,dt,$$

from the above formula and Theorem 6.25 we get

$$\begin{aligned}\left|\frac{F(x)-F(x_0)}{x-x_0} - f(x_0)\right| &= \left|\frac{1}{x-x_0}\int_{x_0}^x f(t)\,dt - \frac{1}{x-x_0}\int_{x_0}^x f(x_0)\,dt\right| \\ &= \left|\frac{1}{x-x_0}\int_{x_0}^x \big(f(t) - f(x_0)\big)\,dt\right| \\ &\le \frac{1}{x-x_0}\int_{x_0}^x |f(t) - f(x_0)|\,dt.\end{aligned}$$

Noticing that in the last integral $t \in [x_0, x] \subset [x_0, x_0+\delta) \cap [a,b]$, it follows from the continuity of f that $|f(t) - f(x_0)| < \varepsilon$, therefore

$$\left|\frac{F(x)-F(x_0)}{x-x_0} - f(x_0)\right| < \frac{1}{x-x_0}\int_{x_0}^x \varepsilon\,dt = \varepsilon\,\frac{1}{x-x_0}\int_{x_0}^x dt = \varepsilon.$$

Similarly, for $x \in (x_0 - \delta, x_0) \cap [a,b]$,

$$\left|\frac{F(x)-F(x_0)}{x-x_0} - f(x_0)\right| = \left|\frac{1}{x-x_0}\left(\int_a^x f(t)\,dt - \int_a^{x_0} f(t)\,dt\right) - f(x_0)\right|$$

$$
\begin{aligned}
&= \left| \frac{1}{x - x_0} \left(- \int_x^{x_0} f(t)\, dt \right) - f(x_0) \right| \\
&= \left| \frac{1}{x_0 - x} \int_x^{x_0} f(t)\, dt - f(x_0) \right| \\
&= \left| \frac{1}{x_0 - x} \int_x^{x_0} \big(f(t) - f(x_0) \big)\, dt \right| \\
&\leq \frac{1}{x_0 - x} \int_x^{x_0} |f(t) - f(x_0)|\, dt < \varepsilon.
\end{aligned}
$$

Since $\varepsilon > 0$ is arbitrary, this proves

$$
\lim_{x \to x_0} \frac{F(x) - F(x_0)}{x - x_0} = f(x_0),
$$

that is, $F'(x_0) = f(x_0)$. This shows that F is a primitive of f.

Uniqueness up to a constant follows by Corollary 5.46: if F is a primitive then

$$
\left(F(x) - \int_a^x f(t)\, dt \right)' = f(x) - f(x) = 0 \qquad \forall\, x \in [a, b],
$$

therefore $F(x) - \int_a^x f(t)\, dt$ is constant. □

26 7.5. — Illustration 7.1 shows the essential estimate in the proof of Theorem 7.4. The value $F(x) - F(x_0)$ can be written as $f(x_0)(x - x_0)$ plus the area in red that is smaller than $\varepsilon(x - x_0)$. Thus $\frac{F(x)-F(x_0)}{x-x_0}$, to an error less than ε, is given by $f(x_0)$.

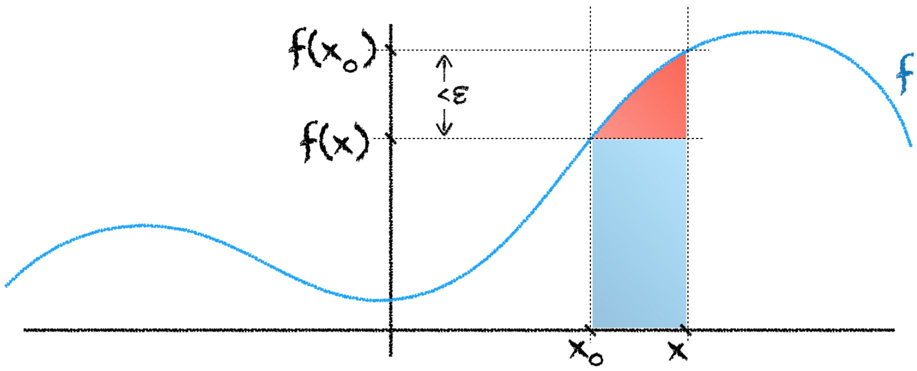


Figure 7.1

Theorem 7.4, as stated or in the form of one of the following corollaries, is known as **Fundamental Theorem of Integral and Differential Calculus** and goes back to the work of Leibniz, Newton and Barrow, which are largely the starting points of calculus. Isaac Barrow (1630–1677) was a theologian, but also a physics and mathematics professor at Cambridge. His most famous student was Isaac Newton.

Corollary 7.6: Integral vs Derivative

Let $F : [a,b] \to \mathbb{R}$ be continuously differentiable. Then, for all $x \in [a,b]$ we have

$$F(x) = F(a) + \int_a^x F'(t)\,dt.$$

Proof. By the definition of primitive, clearly F is a primitive of F'. Hence, according to Theorem 7.4, there exists a constant $C \in \mathbb{R}$ such that $F(x) = \int_a^x F'(t)\,dt + C$ for all $x \in [a,b]$. Choosing $x = a$ we obtain $F(a) = \int_a^a F'(t)\,dt + C = C$, thus

$$F(x) = \int_a^x F'(t)\,dt + F(a) \qquad \forall\, x \in [a,b].$$

□

Corollary 7.7: Riemann Integral and Primitives

Let $f : [a,b] \to \mathbb{R}$ be continuous and $F : [a,b] \to \mathbb{R}$ be a primitive of f. Then

$$\int_a^b f(t)\,dt = F(b) - F(a).$$

26 *Proof.* Apply Corollary 7.6 with $f = F'$ and $x = b$. □

Example 7.8. — For all $a < b$, $a, b \in \mathbb{R}$, it holds:

1. $\int_a^b e^x\, dx = \int_a^b (e^x)'\, dx = e^b - e^a$,
2. $\int_a^b \sin(x)\,dx = -\int_a^b \cos'(x)\,dx = -\cos(b) + \cos(a)$;
3. $\int_a^b \cos(x)\,dx = \int_a^b \sin'(x)\,dx = \sin(b) - \sin(a)$;
4. $\int_a^b \sinh(x)\,dx = \int_a^b \cosh'(x)\,dx = \cosh(b) - \cosh(a)$;
5. $\int_a^b \cosh(x)\,dx = \int_a^b \sinh'(x)\,dx = \sinh(b) - \sinh(a)$;
6. $\int_a^b x^\alpha\, dx = \frac{1}{1+\alpha}\int_a^b (x^{1+\alpha})'\, dx = \frac{1}{1+\alpha}\left(b^{1+\alpha} - a^{1+\alpha}\right)$ whenever $1+\alpha \neq 0$, $0 < a < b$;
7. $\int_a^b x^{-1}\, dx = \int_a^b (\log x)'\, dx = \log(b) - \log(a)$, whenever $0 < a < b$.

Exercise 7.9. — Let $f : [a,b]] \to \mathbb{R}$ be discontinuous at mostly finitely many points. Show that the function $F(x) = \int_a^x f(t)\,dt$ for is continuous in $[a,b]$, differentiable at all continuity points of f, and at such points it satisfies $F'(x) = f(x)$.

EXERCISE 7.10. — Let $f : I \to \mathbb{R}$ be continuous. Show that there exists $\xi \in (a, b)$ with

26

$$\int_a^b f(x)\,dx = f(\xi)(b-a).$$

Can you find two different proofs?

7.1.2 Integration by Parts and by Substitution

As a consequence of the fundamental theorem of calculus and the differentiation rules seen in the last section, we have the following fundamental results. We shall use the following convenient notation: given $h : [a, b] \to \mathbb{R}$,

$$\big[h(x)\big]_a^b = h(b) - h(a).$$

THEOREM 7.11: INTEGRATION BY PARTS

Let $f, g : [a, b] \to \mathbb{R}$ be continuously differentiable functions. Then

$$\int_a^b f(x)\,g'(x)\,dx = \big[f(x)g(x)\big]_a^b - \int_a^b f'(x)\,g(x)\,dx.$$

Proof. Thanks to Proposition 5.12 we have

$$fg' = (fg)' - f'g.$$

27

Integrating this identity on $[a, b]$ and using Corollary 7.7 yields

$$\int_a^b f(x)\,g'(x)\,dx = f(b)g(b) - f(a)g(a) - \int_a^b f'(x)\,g(x)\,dx,$$

as desired. □

From now on, as a convention, given $h : [a, b] \to \mathbb{R}$ we have

$$\int_b^a h(x)\,dx = -\int_a^b h(x)\,dx. \tag{7.2}$$

THEOREM 7.12: INTEGRATION BY SUBSTITUTION, 1ST FORM

Let $I, J \subset \mathbb{R}$ be intervals, let $f : I \to J$ be continuously differentiable and $g : J \to \mathbb{R}$ continuous. Then, for any $[a, b] \subset I$,

$$\int_a^b g(f(x))f'(x)\,dx = \int_{f(a)}^{f(b)} g(y)\,dy.$$

Proof. Fix $y_0 \in J$, and define $G(y) = \int_{y_0}^{y} g(t)\,dt$. By Theorem 7.4 we know that $G' = g$, so it follows from the Chain Rule (see Theorem 5.16) that $g \circ f\, f' = G' \circ f\, f' = (G \circ f)'$. Hence

$$\begin{aligned}\int_a^b g(f(x))f'(x)\,dx &= \int_a^b (G \circ f)'(x)\,dx = G(f(b)) - G(f(a)) \\ &= \int_{y_0}^{f(b)} g(y)\,dy - \int_{y_0}^{f(a)} g(y)\,dy = \int_{f(a)}^{f(b)} g(y)\,dy.\end{aligned}$$

□

Before stating the following result, we recall that if $h : [a,b] \to \mathbb{R}$ is continuously differentiable with $h' \neq 0$, then it follows from the Intermediate Value Theorem 3.29 that $h' > 0$ or $h' < 0$. So, Remark 5.45 implies that it is strictly monotone and therefore invertible. Thanks to Theorem 3.34 it follows that h^{-1} is continuous, and then Theorem 5.21 implies that h^{-1} is differentiable on $(h(a), h(b))$.

Theorem 7.13: Integration by Substitution, 2nd Form

27 *Let $I, J \subset \mathbb{R}$ be intervals, let $f : I \to J$ be continuously differentiable and $g : J \to \mathbb{R}$ continuous. Given $[a,b] \subset I$, assume that $f'(x) \neq 0$ for all $x \in [a,b]$, and let $f^{-1} : [f(a), f(b)] \to \mathbb{R}$ be the inverse of $f|_{[a,b]}$. Then*

$$\int_a^b g(f(x))\,dx = \int_{f(a)}^{f(b)} g(y)\,(f^{-1})'(y)\,dy.$$

Proof. By Theorem 7.13 applied with $\frac{g}{f'}$ in place of g we have

$$\int_a^b g(f(x))\,dx = \int_a^b \frac{g(f(x))}{f'(x)} f'(x)\,dx = \int_a^b \frac{g(f(x))}{f' \circ f^{-1}(f(x))} f'(x)\,dx = \int_{f(a)}^{f(b)} \frac{g(y)}{f'(f^{-1}(y))}\,dy.$$

Recalling that $\frac{1}{f' \circ f^{-1}} = (f^{-1})'$ (see Theorem 5.21), the result follows. □

7.1.3 Improper Integrals

Given a non-empty interval $I \subseteq \mathbb{R}$, we say that $f : I \to \mathbb{R}$ is locally integrable if $f|_{[a,b]}$ is integrable for every compact interval $[a,b] \subset I$.

DEFINITION 7.14: IMPROPER INTEGRALS

Let $I \subseteq \mathbb{R}$ be a nonempty interval, and $f : I \to \mathbb{R}$ be a locally integrable function. Set $c = \inf(I) \in \mathbb{R} \cup \{-\infty\}$ and $d = \sup(I) \in \mathbb{R} \cup \{\infty\}$, and choose $x_0 \in I$. We define the **improper integral** of f on I as

$$\int_c^d f(x)\,dx = \lim_{a \to c^+} \int_a^{x_0} f(x)\,dx + \lim_{b \to d^-} \int_{x_0}^b f(x)\,dx$$

whenever both limits exist and the sum makes sense (so, if the limits are infinite, we do not admit an expression of the form $\infty - \infty$).
If the limit is finite, we say that the improper integral converges. If the limit is ∞ or $-\infty$, we say that the improper integral is divergent. Otherwise, we call the improper integral not convergent.

EXAMPLE 7.15. — It holds

$$\int_0^\infty \frac{1}{1+x^2}\,dx = \lim_{b\to\infty} \int_0^b \frac{1}{1+x^2}\,dx = \lim_{b\to\infty} \big[\arctan(x)\big]_0^b = \lim_{b\to\infty} \arctan(b) = \frac{\pi}{2}.$$

EXAMPLE 7.16. — Given $\alpha \in \mathbb{R}$, it holds

27

$$\int_1^\infty x^{-\alpha}\,dx = \begin{cases} \frac{1}{\alpha-1} & \text{if } \alpha > 1 \\ \infty & \text{if } \alpha \leq 1. \end{cases}$$

In particular, the above improper integral is convergent exactly when $\alpha > 1$. In fact

$$\int_1^b x^{-\alpha}\,dx = \begin{cases} \left[\frac{1}{1-\alpha} x^{1-\alpha}\right]_1^b = \frac{1}{1-\alpha} b^{1-\alpha} - \frac{1}{1-\alpha} & \text{if } \alpha \neq 1 \\ [\log(x)]_1^b = \log(b) & \text{if } \alpha = 1 \end{cases}$$

and

$$\lim_{b\to\infty} \frac{1}{1-\alpha} b^{1-\alpha} = \begin{cases} \infty & \text{if } \alpha < 1 \\ 0 & \text{if } \alpha > 1 \end{cases}, \qquad \lim_{b\to\infty} \log(b) = \infty.$$

LEMMA 7.17: IMPROPER INTEGRAL OF NON-NEGATIVE FUNCTIONS

Let $a \in \mathbb{R}$, and $f : [a, \infty) \to \mathbb{R}_{\geq 0}$ be a non-negative locally integrable function. Then

$$\int_a^\infty f(x)\,dx = \sup\left\{ \int_a^b f(x)\,dx \mid b > a \right\}.$$

Proof. Since the function $b \in [a, \infty) \mapsto \int_a^b f(x)\,dx$ is monotonically increasing, it always has a limit as $b \to \infty$, which is equal to the supremum $\sup\left\{\int_a^b f(x)\,dx \mid b > a\right\}$. This supremum is either finite (in which case the improper integral converges) or it is infinite (in which case the improper integral diverges to ∞). □

EXAMPLE 7.18. — We want to study the improper integral

$$\int_{-\infty}^{\infty} \mathrm{e}^{-x^2}\, dx = \int_{-\infty}^{-1} \mathrm{e}^{-x^2}\, dx + \int_{-1}^{1} \mathrm{e}^{-x^2}\, dx + \int_{1}^{\infty} \mathrm{e}^{-x^2}\, dx.$$

The function $x \in \mathbb{R} \mapsto \mathrm{e}^{-x^2}$ is called the *Gaussian*. Due to Lemma 7.17, to prove that the integral above converges it suffices to find a "majorant function" which defines a convergent improper integral. Since $x^2 \geq x$ for $x \in [1, \infty)$, it follows that $\mathrm{e}^{-x^2} \leq \mathrm{e}^{-x}$, and therefore

$$\int_{1}^{\infty} \mathrm{e}^{-x^2}\, dx \leq \int_{1}^{\infty} \mathrm{e}^{-x}\, dx = \lim_{b\to\infty} \big[-e^{-x}\big]_1^b = \lim_{b\to\infty} \big(e^{-1} - e^{-b}\big) = e^{-1} < \infty.$$

This shows the convergence of the second improper integral. Therefore, due to the symmetry of the function, $\int_{-\infty}^{\infty} \mathrm{e}^{-x^2}\, dx = \int_{1}^{\infty} \mathrm{e}^{-x^2}\, dx$ is also convergent. Finally, since $e^{-x^2} \leq 1$, the integral on $[-1, 1]$ is bounded by $\int_{-1}^{1} 1\, dx = 2$.

This proves the convergence of the integral. However, we will not be able to calculate the exact value of this integral until the second semester.

THEOREM 7.19: INTEGRAL TEST FOR SERIES

Let $f : [0, \infty) \to \mathbb{R}_{\geq 0}$ be a monotonically decreasing function. Then

27

$$\sum_{n=1}^{N+1} f(n) \leq \int_0^{N+1} f(x)\, dx \leq \sum_{n=0}^{N} f(n).$$

In particular

$$\sum_{n=1}^{\infty} f(n) \leq \int_0^{\infty} f(x)\, dx \leq \sum_{n=0}^{\infty} f(n),$$

therefore, the series $\sum_{n=1}^{\infty} f(n)$ converges exactly when the improper integral $\int_0^{\infty} f(x)\, dx$ converges.

Proof. Due to the monotonicity of f, Theorem 6.31 implies that the function f is locally integrable. We consider the functions $\ell, u : [0, \infty) \to \mathbb{R}_{\geq 0}$ given by

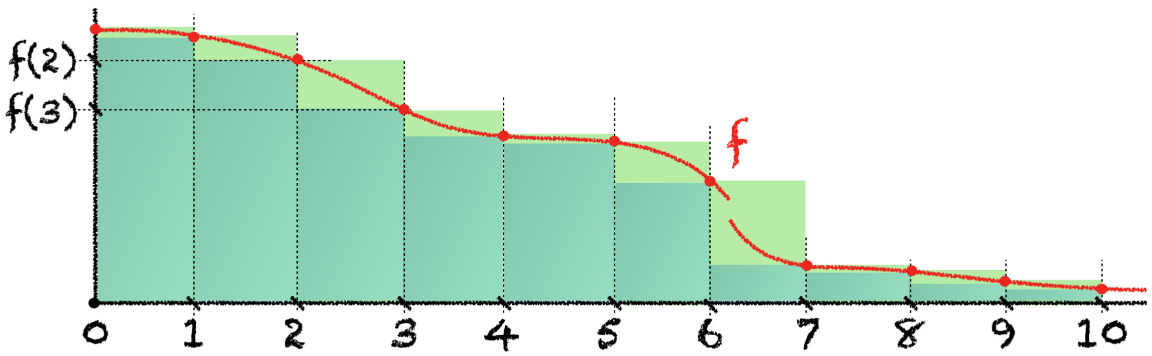
$$u(x) = f(\lfloor x \rfloor) \qquad \text{and} \qquad \ell(x) = f(\lceil x \rceil)$$

where $\lfloor \cdot \rfloor$ represents the rounding function (i.e., $\lfloor x \rfloor$ is the largest number $n \in \mathbb{N}$ smaller than x), and $\lceil \cdot \rceil$ represents the rounding up function (i.e., $\lceil x \rceil$ is the smallest number $n \in \mathbb{N}$ larger than x). With this choice, $\ell \leq f \leq u$. Therefore, for all $N \in \mathbb{N}$ with $N > 1$, we have

$$\sum_{n=1}^{N+1} f(n) = \int_0^{N+1} \ell(x)\, dx \leq \int_0^{N+1} f(x)\, dx \leq \int_0^{N+1} u(x)\, dx = \sum_{n=0}^{N} f(n),$$

which can also be seen in the following picture. The statement of the theorem follows by letting $N \to \infty$. □

27



EXAMPLE 7.20. — The harmonic series can be written as $\{f(n)\}_{n=0}^{\infty}$ with $f(x) = \frac{1}{1+x}$. Thus it diverges since $\int_0^{\infty} \frac{1}{1+x}\, dx = \infty$. On the other hand, the series $\sum_{n=1}^{\infty} \frac{1}{n^2}$ converges since $\int_0^{\infty} \frac{1}{(1+x)^2}\, dx < \infty$.

7.2 Integration and Differentiation of Power Series

From Example 7.8(6) we know that, for any $n \geq 0$, $\int_0^x t^n\,dt = \frac{1}{n+1}x^{n+1}$. In other words, $\frac{1}{n+1}x^{n+1}$ is a primitive of x^n. Also, thanks to Corollary 5.14 we know that $(x^n)' = nx^{n-1}$. These formulas allow us to compute integrals and derivatives of polynomials. We now want to understand when we can integrate/differentiate power series.

Before stating and proving the results, we recall that $\lim_{n\to\infty} \sqrt[n]{\frac{1}{n}} = 1$. This can be proved using that $\frac{\log(n)}{n} \to 0$ as $n \to \infty$. Indeed,

$$\sqrt[n]{\frac{1}{n}} = e^{\frac{1}{n}\log\left(\frac{1}{n}\right)} = e^{-\frac{1}{n}\log(n)} \to e^0 = 1 \qquad \text{as } n \to \infty. \tag{7.3}$$

Alternatively, one can use Exercise 3.37 and Proposition 2.96(3).

Theorem 7.21: Integration of Power Series

Let $f(x) = \sum_{n=0}^{\infty} a_n x^n$ be a power series with radius of convergence $R > 0$. Then the power series

$$F(x) = \sum_{n=0}^{\infty} \frac{a_n}{n+1} x^{n+1}$$

28 *has also radius of convergence R, and is a primitive of f on $(-R, R)$.*

Proof. We first check the assertion about the radius of convergence of F.

Let $\rho = \limsup_{n\to\infty} \sqrt[n]{|a_n|}$, so that $R = \rho^{-1}$, and define $c_0 = 0$ and $c_n = \frac{a_{n-1}}{n}$ for $n \geq 1$. In this way, it follows that

$$F(x) = \sum_{n=0}^{\infty} \frac{a_n}{n+1} x^{n+1} = \sum_{n=1}^{\infty} \frac{a_{n-1}}{n} x^n = \sum_{n=1}^{\infty} c_n x^n = \sum_{n=0}^{\infty} c_n x^n.$$

Recalling (7.3), given $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that $1 - \varepsilon \leq \sqrt[n]{\frac{1}{n}} \leq 1 + \varepsilon$ for all $n \geq N$. This implies that

$$\begin{aligned}\limsup_{n\to\infty} \sqrt[n]{|c_n|} &= \limsup_{n\to\infty} \sqrt[n]{\frac{|a_{n-1}|}{n}} \leq (1+\varepsilon) \limsup_{n\to\infty} \sqrt[n]{|a_{n-1}|} \\ &= (1+\varepsilon) \limsup_{n\to\infty} \left(\sqrt[n-1]{|a_{n-1}|}\right)^{\frac{n-1}{n}} = (1+\varepsilon)\rho,\end{aligned}$$

and analogously

$$\limsup_{n\to\infty} \sqrt[n]{|c_n|} \geq (1-\varepsilon) \limsup_{n\to\infty} \left(\sqrt[n-1]{|a_{n-1}|}\right)^{\frac{n-1}{n}} = (1-\varepsilon)\rho.$$

Thus, if we set $\bar{\rho} = \limsup_{n\to\infty} \sqrt[n]{|c_n|}$, this proves that

$$(1-\varepsilon)\rho \le \bar{\rho} \le (1+\varepsilon)\rho.$$

Since $\varepsilon > 0$ is arbitrary, this implies that $\bar{\rho} = \rho$, so the power series $F(x) = \sum_{n=0}^{\infty} c_n x^n$ has radius of convergence R.

We now want to prove that $F' = f$ on $(-R, R)$. To prove that, fix an interval $[a, b] \subset (-R, R)$, and consider the polynomial functions $f_n(t) = \sum_{k=0}^{n} a_k t^k$. We note that

$$\int_a^x f_n(t)\,dt = \sum_{k=0}^{n} a_k \int_a^x t^k\,dt = \sum_{k=0}^{n} \frac{a_k}{k+1} x^{k+1} - \sum_{k=0}^{n} \frac{a_k}{k+1} a^{k+1} \qquad \forall\, x \in [a,b].$$

By Theorem 4.42, the sequence of functions $(f_n)_{n=0}^{\infty}$ converges uniformly to f in $[a, x] \subset [a, b]$, so it follows from Theorem 6.42 that

$$\int_a^x f(t)\,dt = \lim_{n\to\infty} \int_a^x f_n(t)\,dt = F(x) - F(a) \qquad \forall\, x \in [a,b].$$

According to Theorem 7.4, this implies that $F'(x) = f(x)$ for all $x \in [a, b]$. Since $[a, b] \subset (-R, R)$ is an arbitrary interval, this implies that $F' = f$ on $(-R, R)$, as desired. □

Corollary 7.22: Differentiation of Power Series

28 *Let* $f(x) = \sum_{n=0}^{\infty} a_n x^n$ *be a power series with radius of convergence* $R > 0$. *Then* $f : (-R, R) \to \mathbb{R}$ *is differentiable and it holds*

$$f'(x) = \sum_{n=1}^{\infty} n a_n x^{n-1} \qquad \forall\, x \in (-R, R),$$

where the power series on the right has also radius of convergence R.

Proof. Let $c_n = (n+1)a_n$ for $n \ge 0$, and let $g(x) = \sum_{n=1}^{\infty} n a_n x^{n-1} = \sum_{n=0}^{\infty} c_n x^n$. Denote by $\bar{R}$ the radius of convergence of g. Then, according to Theorem 7.21, the series

$$G(x) = \sum_{n=0}^{\infty} \frac{c_n}{n+1} x^{n+1}$$

has radius of convergence $\bar{R}$ and is a primitive of g. We now note that

$$G(x) = \sum_{n=0}^{\infty} \frac{c_n}{n+1} x^{n+1} = \sum_{n=0}^{\infty} \frac{(n+1)a_n}{n+1} x^{n+1} = \sum_{n=1}^{\infty} a_n x^n = f(x) - a_0.$$

This implies that G and f have the same radius of convergence and $g = G' = (f - a_0)' = f'$. Therefore, we conclude that $\bar{R} = R$ and $f' = g$. □

EXERCISE 7.23. — Let $f(x) = \sum_{n=0}^{\infty} a_n x^n$ be a power series with radius of convergence $R > 0$. Show that $f : (-R, R) \to \mathbb{R}$ is smooth, and for each $n \in \mathbb{N}$ find a representation of the n-the derivative $f^{(n)}$ by a power series.

EXERCISE 7.24. — Let $f(x) = \sum_{n=0}^{\infty} a_n x^n$ and $g(x) = \sum_{n=0}^{\infty} b_n x^n$ be power series with radii of convergence $R_f, R_g > 0$. Let $R = \min\{R_f, R_g\}$ and suppose that $f(x) = g(x)$ for all $x \in (-R, R)$. Show that $c_n = d_n$ for all $n \in \mathbb{N}$, and therefore $R_f = R_g$.

EXERCISE 7.25. — Let $\alpha \in \mathbb{R}$. The goal of this exercise is to show that

$$(1+x)^{\alpha} = \sum_{n=0}^{\infty} \binom{\alpha}{n} x^n \qquad \forall\, x \in (-1, 1), \tag{7.4}$$

where the generalised binomial coefficients are defined as

$$\binom{\alpha}{n} = \frac{\prod_{j=0}^{n-1}(\alpha - j)}{n!} = \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!}.$$

(a) Show that, for $\alpha \notin \mathbb{N}$, the power series

28

$$f(x) = \sum_{n=0}^{\infty} \binom{\alpha}{n} x^n$$

has radius of convergence 1.

(b) Calculate the derivative of f and show that

$$f'(x) = \alpha \frac{f(x)}{1+x} \qquad \forall\, x \in (-1, 1). \tag{7.5}$$

(c) Define $g(x) = (1+x)^{\alpha}$ and use (8.20) to show that $\left(\frac{f}{g}\right)' = 0$ on $(-1, 1)$. Conclude the validity of (7.4) by noticing that $f(0) = 1 = g(0)$.

EXAMPLE 7.26. — We have already seen in Example 4.23 that, as a consequence of the Leibniz criterion in Proposition 4.22, the alternating harmonic series converges. However, with the results of Chapter 4 we could not determine the value of the series. Now, with the help of the fundamental theorem of integral and differential calculus, we can show that

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} = \log(2).$$

To prove this, using the formula for the geometric series (which has radius of convergence 1), we see that

$$(\log(1+x))' = \frac{1}{1+x} = \frac{1}{1-(-x)} = \sum_{n=0}^{\infty}(-1)^n x^n \qquad \forall\, x \in (-1,1).$$

According to Corollary 7.6 and Theorem 7.21, this implies that

$$\log(1+x) = \log(1) + \int_0^x \frac{1}{1+t}\,dt = \sum_{n=0}^{\infty}\frac{(-1)^n}{n+1}x^{n+1} = \sum_{k=1}^{\infty}\frac{(-1)^{k+1}}{k}x^k \qquad \forall\, x \in (-1,1).$$

Note now that, given $x \in [0,1]$, the sequence $a_k = \frac{x^k}{k+1}$ is non-negative, decreasing, and converging to zero. Hence it follows from Proposition 4.22 that, for $x \in [0,1]$,

$$\sum_{k=1}^{2n}\frac{(-1)^{k+1}}{k}x^k \le \sum_{k=1}^{\infty}\frac{(-1)^{k+1}}{k}x^k = \log(1+x) \le \sum_{k=1}^{2n+1}\frac{(-1)^{k+1}}{k}x^k \qquad \forall\, n \in \mathbb{N}.$$

Letting $x \to 1^-$, this implies that

$$\sum_{k=1}^{2n}\frac{(-1)^{k+1}}{k} \le \log(2) \le \sum_{k=1}^{2n+1}\frac{(-1)^{k+1}}{k} \qquad \forall\, n \in \mathbb{N}.$$

28 Finally, letting $n \to \infty$ in the above inequalities proves the result.

Example 7.27. — We can use the above method to show that

$$\sum_{n=0}^{\infty}\frac{(-1)^n}{2n+1} = \frac{\pi}{4}. \tag{7.6}$$

Using again the formula for the geometric series, we see that

$$\arctan'(x) = \frac{1}{1+x^2} = \sum_{k=0}^{\infty}(-1)^k x^{2k} \qquad \forall\, x \in (-1,1).$$

According to Corollary 7.6 and Theorem 7.21, this implies that

$$\arctan(x) = \arctan(0) + \int_0^x \frac{1}{1+t^2}\,dt = \sum_{k=0}^{\infty}\frac{(-1)^k}{2k+1}x^{2k+1}.$$

Again from Proposition 4.22, as in the previous example it follows that

$$\sum_{k=0}^{2n+1}\frac{(-1)^k}{2k+1}x^{2k+1} \le \arctan(x) \le \sum_{k=0}^{2n}\frac{(-1)^k}{2k+1}x^{2k+1} \qquad \forall\, n \in \mathbb{N}.$$

Letting first $x \to 1^-$ we get

$$\sum_{k=0}^{2n+1} \frac{(-1)^k}{2k+1} \leq \arctan(1) = \frac{\pi}{4} \leq \sum_{k=0}^{2n} \frac{(-1)^k}{2k+1} \qquad \forall\, n \in \mathbb{N},$$

so the result follows by letting $n \to \infty$ (note that the series converges, thanks to Leibniz criterion in Proposition 4.22).

Sometimes, the above methods for determining an indefinite integral of a function do not produce a result. This may be because the primitive function we are looking for cannot be expressed in terms of "known" functions.

28 EXAMPLE 7.28 (Integral Sine). — The **integral sine** is the primitive function $\mathrm{Si} : \mathbb{R} \to \mathbb{R}$ of the continuous function

$$x \in \mathbb{R} \mapsto \begin{cases} \frac{\sin(x)}{x} & \text{if } x \neq 0 \\ 1 & \text{if } x = 0 \end{cases}$$

with normalisation $\mathrm{Si}(0) = 0$, that is $\mathrm{Si}(x) = \int_0^x \frac{\sin(t)}{t}\, dt$. Thanks to Theorem 7.21, the function Si can be written as a power series:

$$\mathrm{Si}(x) = \int_0^x \frac{\sin(t)}{t}\, dt = \int_0^x \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} t^{2n}\, dt = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!(2n+1)} x^{2n+1}$$

for all $x \in \mathbb{R}$.

7.3 Integration Methods

Let $I \subseteq \mathbb{R}$ be an interval, and $f : I \to \mathbb{R}$ a function. The notation

$$\int f(x)\,dx = F(x) + C$$

means that F is a primitive function of f. In the expression $F(x) + C$, C is read as an indefinite constant - called **integration constant**. Since the domain I of f is an interval, two primitive functions of f differ by a constant, which makes the notation meaningful. One calls $F(x) + C$ the **indefinite integral** of f. Indefinite integrals of special functions can be found in tables or by means of computer algebra systems. In this section, we show general methods to determine indefinite integrals.

7.29. — We this whole section, $I \subseteq \mathbb{R}$ denotes a non-empty interval that does not consist only of one point. Also, all functions in this section are real-valued functions with domain I that are integrable on any compact interval $[a, b] \subseteq I$.

7.3.1 Integration by Parts and by Substitution in Leibniz Notation

In the computation of indefinite integrals, it is convenient to use Leibniz notation. This notation allows us to reformulate, in a natural formalism, both integration by parts and by
28 substitution (see Section 7.1.2). We recall that the derivative of a function h is denoted by h' or by $\frac{dh}{dx}$. In this section, the second notation will be useful.

7.30. — Let f and g be functions with primitives F, and G, respectively. Recall that, from the product rule for the derivative in Proposition 5.12, it follows that $(FG)' = fG + Fg$. This implies the **integration by parts** formula

$$\int F(x)g(x)\,dx = F(x)G(x) - \int f(x)G(x)\,dx + C. \tag{7.7}$$

In Leibniz notation, $f = \frac{dF}{dx}$ and $g = \frac{dG}{dx}$. This leads to the notation $f\,dx = dF$ and $g\,dx = dG$, and integration by parts is sometimes written as

$$\int F\,dG = FG - \int G\,dF + C,$$

which should be understood as a short form of formula (7.7).

7.31. — Let J be an interval and let $f : I \to J$ be a differentiable function. If $G : J \to \mathbb{R}$ is a primitive of g then, by the chain rule in Theorem 5.16, $[G(f(x))]' = g(f(x))f'(x)$ holds for all $x \in I$. From this it follows that

$$\int g(f(x))f'(x)\,dx = G(f(x)) + C.$$

Since $G(u) = \int g(u)\,du + C$ holds, we obtain

$$\int g(f(x))f'(x)\,dx = \int g(u)\,du + C \tag{7.8}$$

where we used the change of variables $u = f(x)$. The substitution rule is also called **change of variable**, as one has replaced the variable u in $\int g(u)\,du$ by $u = f(x)$. In Leibniz notation this is very natural: if $u = f(x)$ then $du = f'(x)dx$, and (7.8) follows.

We also recall the second form of the substitution rule: if $f' \neq 0$ we can set $x = f^{-1}(u)$ so that $\frac{dx}{du} = (f^{-1})'(u)$, and obtain

$$\int g(f(x))\,dx = \int g(u)\,\frac{dx}{du}\,du + C, \tag{7.9}$$

see Section 7.1.2.

28 7.3.2 Integration by Parts: Examples

Example 7.32. — We want to calculate the indefinite integral $\int xe^x\,dx$. Since $e^x = (e^x)'$, using (7.7) we get

$$\int xe^x\,dx = \int x(e^x)'\,dx = xe^x - \int x' \cdot e^x\,dx + C = xe^x - \int e^x + C.$$

Since $\int e^x = e^x + C$, we conclude that

$$\int xe^x\,dx = xe^x - e^x + C.$$

We note that it is sufficient to use only one integration constant C in such calculations, since several such constants can be combined into one.

Example 7.33. — We calculate the integral $\int \log(x)\,dx$:

$$\begin{aligned}\int \log(x)\,dx &= \int \log(x) \cdot 1\,dx = \int \log(x) \cdot x'\,dx = \log(x) \cdot x - \int \log'(x)x\,dx + C \\ &= \log(x) \cdot x - \int \frac{1}{x}x\,dx + C = \log(x) \cdot x - \int\,dx + C = x\log(x) - x + C.\end{aligned}$$

Suggestion: To ensure that the final result is correct, differentiate the result and check if you get the original function. For instance, in this case, one can easily check that

$$(x\log(x) - x + C)' = \log(x).$$

Exercise 7.34. — Give a recursive formula for calculating the indefinite integrals

$$\int x^n e^x \, dx \,, \qquad \int x^n \sin(x)\, dx \,, \qquad \int x^n \cos(x)\, dx$$

for $n \in \mathbb{N}$.

Exercise 7.35. — Calculate

$$\int x^s \log(x) dx \,, \qquad \int e^{ax} \sin(bx)\, dx$$

for all $s, a, b \in \mathbb{R}$. Note that the case $s = -1$ needs to be treated separately, in analogy with Example 7.8(6)-(7).

28

7.3.3 Integration by Substitution: Examples

Example 7.36. — We want to compute $\int \frac{x}{1+x^2}\, dx$. Let $u = f(x) = 1 + x^2$, so that $du = f'(x)\, dx = 2x\, dx$. Then we find

$$\int \frac{x}{1+x^2}\, dx = \frac{1}{2}\int \frac{1}{1+x^2} 2x\, dx = \frac{1}{2}\int \frac{1}{u} du = \frac{1}{2}\log|u| = \frac{1}{2}\log(1+x^2) + C.$$

Example 7.37. — Given $r > 0$, we want to compute the indefinite integral $\int \sqrt{r^2 - x^2}\, dx$. Due to the trigonometric identities $\sqrt{r^2 - r^2\sin(\theta)^2} = r\cos(\theta)$ it is convenient to use the change of variable $x = r\sin(\theta)$, $\theta \in (-\frac{\pi}{2}, \frac{\pi}{2})$. With this choice we have $dx = r\cos(\theta)\, d\theta$, therefore

$$\int \sqrt{r^2 - x^2}\, dx = r \int \sqrt{r^2 - r^2\sin(\theta)^2} \cos(\theta)\, d\theta = r^2 \int \cos^2(\theta) d\theta.$$

To compute $\int \cos^2(\theta) d\theta$ we use integration by parts as follows:

$$\begin{aligned} \int \cos^2(\theta) d\theta &= \int \cos(\theta)\sin'(\theta) d\theta = \cos(\theta)\sin(\theta) - \int \cos'(\theta)\sin(\theta) + C \\ &= \cos(\theta)\sin(\theta) + \int \sin^2(\theta) + C. \end{aligned}$$

Since $\sin^2(\theta) = 1 - \cos^2(\theta)$, we get

$$\int \cos^2(\theta) d\theta = \cos(\theta)\sin(\theta) + \int 1 - \int \cos^2(\theta) + C = \cos(\theta)\sin(\theta) + \theta - \int \cos^2(\theta) + C,$$

therefore

$$2\int \cos^2(\theta) d\theta = \cos(\theta)\sin(\theta) + \theta + C \qquad \Longrightarrow \qquad \int \cos^2(\theta) d\theta = \frac{1}{2}\big(\cos(\theta)\sin(\theta) + \theta\big) + C$$

28 (note that, since $C \in \mathbb{R}$ is arbitrary, in the last formula we still write C in place of $\frac{C}{2}$). This proves that

$$\int \sqrt{r^2 - x^2}\, dx = r^2 \int \cos^2(\theta)\, d\theta = \frac{r^2}{2}\big(\sin(\theta)\cos(\theta) + \theta\big) + C$$

(again, we write C instead of Cr^2). Recalling that $x = r\sin(\theta)$ with $\theta \in (-\frac{\pi}{2}, \frac{\pi}{2})$, it follows that $\theta = \arcsin\left(\frac{x}{r}\right)$ and $\cos(\theta) = \sqrt{1 - \frac{x^2}{r^2}}$, therefore

$$\int \sqrt{r^2 - x^2}\, dx = \frac{1}{2} x\sqrt{r^2 - x^2} + \frac{r^2}{2} \arcsin\left(\tfrac{x}{r}\right) + C.$$

7.38. — Substitution like in Example 7.41 is called **trigonometric substitutions**. We will not always argue carefully in these calculations and will rather trust the Leibniz notation, but recall that, to apply (7.9), there must be invertibility of the function when we express the old variable by the new variable. For the following list of trigonometric substitutions, let $n \in \mathbb{Z}$.

- In expressions of the form $(a^2 - x^2)^{\frac{n}{2}}$ for $a > 0$, as already seen in the example above, one considers the the substitution $x = a\sin(\theta)$ with $\theta \in (-\frac{\pi}{2}, \frac{\pi}{2})$, giving $dx = a\cos(\theta)\, d\theta$ and $(a^2 - x^2)^{\frac{1}{2}} = a\cos(\theta)$.
- 29 In expressions of the form $(a^2 + x^2)^{\frac{n}{2}}$ for $a > 0$, the substitution $x = a\tan(\theta)$ with $\theta \in (-\frac{\pi}{2}, \frac{\pi}{2})$ yields $dx = \frac{a}{\cos^2(\theta)}\, d\theta$ and $(a^2 + x^2)^{\frac{1}{2}} = \frac{a}{\cos(\theta)}$.
- Although this is not a trigonometric substitution, we still note the following: For the expression $x(a^2 - x^2)^{\frac{n}{2}}$ or the expression $x(a^2 + x^2)^{\frac{n}{2}}$, the substitutions $u = a^2 - x^2$ and $u = a^2 + x^2$, respectively, allow us to compute the indefinite integrals.

Example 7.39. — (i) Given $a > 0$, using the substitution $x = a\tan(\theta)$, recalling that $(a^2 + x^2)^{\frac{1}{2}} = \frac{a}{\cos(\theta)}$ and $dx = \frac{a}{\cos^2(\theta)}\, d\theta$, we get

$$\int \frac{1}{(a^2 + x^2)^{\frac{3}{2}}}\, dx = \int \frac{\cos^3(\theta)}{a^3} \frac{a}{\cos^2(\theta)}\, d\theta = \frac{1}{a^2}\int \cos(\theta) d\theta = \frac{1}{a^2}\sin(\theta) + C$$

$$= \frac{1}{a^2}\tan(\theta)\cos(\theta) + C = \frac{x}{a^2\sqrt{a^2+x^2}} + C.$$

(ii) Choosing $u = 1 - x^2$ (so that $du = -2x\,dx$), we have

$$\int x\sqrt{1-x^2}\,dx = -\frac{1}{2}\int u^{1/2}\,du = -\frac{1}{2}\frac{2}{3}u^{3/2} + C = -\frac{1}{3}(1-x^2)^{3/2} + C.$$

Certain indefinite integrals can be computed with hyperbolic substitutions. For instance, for expressions of the form $(x^2-a^2)^{\frac{n}{2}}$ with $a \in \mathbb{R}$, the substitution $x = a\cosh(u)$ yields $dx = a\sinh(u)\,du$ and $(x^2-a^2)^{\frac{1}{2}} = a\sinh(u)$.

Example 7.40. — Using the substitution $x = \cosh(u)$ (so $dx = \sinh(u)\,du$), we compute

$$\int \sqrt{x^2-1}\,dx = \int \sqrt{\cosh^2(u)-1}\,\sinh(u)\,du = \int \sinh^2(u)\,du.$$

In analogy to the argument used in Example 7.37, we compute $\int \sinh^2(u)\,du$ as follows:

29

$$\begin{aligned}\int \sinh^2(u)\,du &= \cosh(u)\sinh(u) - \int \cosh^2(u)\,du \\ &= \cosh(u)\sinh(u) - \int \left(1+\sinh^2(u)\right)du + C \\ &= \cosh(u)\sinh(u) - u - \int \sinh^2(u)\,du + C,\end{aligned}$$

This yields

$$2\int \sinh^2(u)\,du = \cosh(u)\sinh(u) - u + C \quad\implies\quad \int \sinh^2(u)\,du = \frac{\cosh(u)\sinh(u)-u}{2} + C,$$

hence

$$\int \sqrt{x^2-1}\,dx = \frac{\cosh(u)\sinh(u)-u}{2} + C = \frac{x\sqrt{x^2-1} - \operatorname{arcosh}(x)}{2} + C.$$

Another method that we would like to mention briefly here is the so-called **half-angle method** (or **Weierstrass substitution**). This is useful for the integral of expressions like $\frac{1}{\sin(x)}$ or $\frac{\cos^2(x)+\cos(x)+\sin(x)}{1+\sin(x)}$, see also Remark 7.46 below. We show this method in detail in the next example.

Example 7.41. — We want to compute $\int \frac{1}{\sin(x)}\,dx$, and we consider the change of variable $u = \tan\left(\frac{x}{2}\right)$. We can note that, by the doubling angle formulas for sine and cosine, it follows that

$$\sin(x) = \frac{2u}{1+u^2} \qquad \text{and} \qquad \cos(x) = \frac{1-u^2}{1+u^2}.$$

Indeed, using (4.15) we have

$$\frac{2u}{1+u^2} = \frac{2\tan\left(\frac{x}{2}\right)}{1+\tan^2\left(\frac{x}{2}\right)} = \frac{2\frac{\sin\left(\frac{x}{2}\right)}{\cos\left(\frac{x}{2}\right)}}{1+\frac{\sin^2\left(\frac{x}{2}\right)}{\cos^2\left(\frac{x}{2}\right)}} = 2\sin\left(\frac{x}{2}\right)\cos\left(\frac{x}{2}\right) = \sin(x),$$

and analogously for the second formula. Furthermore, the relation $u = \tan\left(\frac{x}{2}\right)$ implies that $x = 2\arctan(u)$, therefore $dx = \frac{2}{1+u^2}du$ (recall that $\arctan'(s) = \frac{1}{1+s^2}$).

Using these formulas, we get

$$\int \frac{1}{\sin(x)}\,dx = \int \frac{1+u^2}{2u}\frac{2}{1+u^2}\,du = \int \frac{1}{u}\,du = \log|u| + C = \log\left|\tan\left(\tfrac{x}{2}\right)\right| + C.$$

7.3.4 Integration of Rational Functions

A function of the form $f(x) = \frac{p(x)}{q(x)}$ for polynomials p and $q \neq 0$ is called a **rational function**. In this section, we show a procedure for computing the indefinite integral of a rational function $f = \frac{p}{q}$ on an interval I on which q has no zeros. By polynomial division with reminder, one can always write $f = \frac{p}{q}$ in the form $f = g + \frac{r}{q}$, where g and r are polynomials with $\deg r < \deg q$. The polynomial function g is easy to integrate. Therefore, we always assume that the degree
29 of p is smaller than the degree of q.

7.42. — We start by integrating some elementary rational functions. Let $a \in \mathbb{R}$ and $n \in \mathbb{N}$ with $n \geq 2$. Then:

$$\int \frac{1}{x-a}\,dx = \log|x-a| + C \tag{7.10}$$

$$\int \frac{1}{(x-a)^n}\,dx = \frac{1}{1-n}(x-a)^{1-n} + C \tag{7.11}$$

$$\int \frac{1}{a^2+x^2}\,dx = \frac{1}{a}\arctan\left(\tfrac{x}{a}\right) + C \tag{7.12}$$

$$\int \frac{x}{a^2+x^2}\,dx = \frac{1}{2}\log(a^2+x^2) + C \tag{7.13}$$

$$\int \frac{x}{(a^2+x^2)^n}\,dx = \frac{1}{2(1-n)}(a^2+x^2)^{1-n} + C \tag{7.14}$$

The integrals (7.10) and (7.11) are calculated with substitution $u = x-a$, for (7.12) substitute $u = \frac{x}{a}$, for (7.13) and (7.14) substitute $u = a^2 + x^2$.

7.43. — To integrate a general rational function, we use what is called the **partial fraction decomposition** of rational functions. Let p, q be polynomials without nontrivial common divisors such that $q \neq 0$ and $\deg p < \deg q$.

First, factorize the polynomial q into linear and quadratic factors

$$q(x) = (x-a_1)^{k_1} \cdots (x-a_n)^{k_n} (x^2+b_1x+c_1)^{\ell_1} \cdots (x^2+b_mx+c_m)^{\ell_m}$$

Then, the rational function $\frac{p(x)}{q(x)}$ can be rewritten as a linear combination of rational functions of the form

$$\frac{1}{(x-a_i)^k}, \qquad \frac{1}{(x^2+b_jx+c_j)^\ell}, \qquad \frac{x}{(x^2+b_jx+c_j)^\ell},$$

for some $k \le k_i$ and $\ell \le \ell_j$, and then one needs to integrate each of these individual terms.

Example 7.44. — We calculate the indefinite integral $\int \frac{x^4+1}{x^2(x+1)}\,dx$. First, we perform division with remainder:

$$\frac{x^4+1}{x^3+x^2} = x - 1 + \frac{x^2+1}{x^2(x+1)}$$

To obtain the partial fraction decomposition of $\frac{x^2+1}{x^2(x+1)}$, we set

$$\frac{x^2+1}{x^2(x+1)} = \frac{ax+b}{x^2} + \frac{c}{x+1}$$

for some real numbers a, b, c to be determined. To determine a, b, c we multiply both sides by $x^2(x+1)$, which gives

29 $$x^2+1 = ax(x+1) + b(x+1) + cx^2 = ax^2 + ax + bx + b + cx^2.$$

Comparing the coefficients, we find the linear system

$$a+c=1, \qquad a+b=0, \qquad b=1,$$

therefore $a=-1$, $b=1$, and $c=2$. So in summary

$$\frac{x^4+1}{x^2(x+1)} = x - 1 - \frac{1}{x} + \frac{1}{x^2} + \frac{2}{x+1}.$$

Therefore

$$\begin{aligned}\int \frac{x^4+1}{x^2(x+1)}\,dx &= \int x\,dx - \int 1\,dx - \int \frac{1}{x}\,dx + \int \frac{1}{x^2}\,dx + 2\int \frac{1}{x+1}\,dx \\ &= \frac{x^2}{2} - x - \log|x| - \frac{1}{x} + 2\log|x+1| + C.\end{aligned}$$

Example 7.45. — We calculate the indefinite integral $\int \frac{1}{x(x^2+2x+2)}\,dx$. Note that the polynomial x^2+2x+2 has no real zeros. For the partial fraction decomposition we make the approach

$$\frac{1}{x(x^2+2x+2)} = \frac{a}{x} + \frac{bx+c}{x^2+2x+2}.$$

Now we multiply both sides by $x(x^2+2x+2)$ and get

$$1 = a(x^2+2x+2) + (bx+c)x = ax^2 + 2ax + 2a + bx^2 + cx,$$

thus

$$a+b=0, \qquad 2a+c=0, \qquad 2a=1,$$

which gives $a=\frac{1}{2}$, $b=-\frac{1}{2}$, and $c=-1$. It follows that

$$\begin{aligned}
\int \frac{1}{x(x^2+2x+2)}\,dx &= \frac{1}{2}\int \frac{1}{x}\,dx - \frac{1}{2}\int \frac{x+2}{x^2+2x+2}\,dx \\
&= \frac{1}{2}\log|x| - \frac{1}{2}\int \frac{x+2}{(x+1)^2+1}\,dx \\
&= \frac{1}{2}\log|x| - \frac{1}{2}\int \frac{u+1}{u^2+1}\,dx \\
&= \frac{1}{2}\log|x| - \frac{1}{4}\log|u^2+1| - \frac{1}{2}\arctan(u) + C \\
&= \frac{1}{2}\log|x| - \frac{1}{4}\log((x+1)^2+1) - \frac{1}{2}\arctan(x+1) + C,
\end{aligned}$$

where we have set $u=x+1$ and used (7.13) and (7.14).

In some cases, the above procedure may also lead to the integral $\int \frac{1}{(a^2+x^2)^n}\,dx$ for an $a\in\mathbb{R}$
29 and $n\geq 2$, which (as explained previously) we can handle with the trigonometric substitution $\tan(u)=\frac{x}{a}$.

Remark 7.46. — Now that we know how to integrate rational functions, we can rediscuss the **half-angle method** introduced before. This allows one to compute the integral of rational expressions in sine and cosine. In fact, with the substitution $u=\tan\left(\frac{x}{2}\right)$, using that

$$\sin(x) = \frac{2u}{1+u^2}, \qquad \cos(x) = \frac{1-u^2}{1+u^2}, \qquad dx = \frac{2}{1+u^2}du,$$

(see Exercise 7.41), one ends up with the integral of a rational function in u.

Exercise 7.47. — Calculate the indefinite integral $\int \frac{\cos(x)}{2+\sin(x)}\,dx$ using the substitution $u=\tan\left(\frac{x}{2}\right)$.

Remark 7.48. — Sometimes one or the other substitution is carried out because there is a nested function in the function to be integrated and one simply has no other method available. For example, in the integral $\int \sin(\sqrt{x})\,dx$ none of the mentioned methods is available, but one is tempted to set $u=\sqrt{x}$, and this indeed leads to an integral that one can solve. Similarly, in an integral of the form $\int \frac{1}{1+e^x}\,dx$, one sets $u=e^x$.

7.3.5 Definite Integrals with Improper Integration Limits

We have so far only integrated functions on Compact Intervals $[a, b]$. In this section, we extend the notion of integral to also think about the Riemann integral of functions on unconstrained and not necessarily closed intervals.

EXAMPLE 7.49. — We compute the indefinite integral $\int_0^1 \log(x)\,dx$ using

$$\int_0^1 \log(x)\,dx = \lim_{a\to 0}\int_a^1 \log(x)\,dx = \lim_{a\to 0}\left[x\log(x) - x\right]_a^1$$
$$= \lim_{a\to 0}(\log(1) - 1 - a\log(a) + a) = -1$$

after Example 3.79.

29

EXAMPLE 7.50. — We consider the improper integral $\int_0^1 \frac{1}{x}\,dx$. This is improper since $x \mapsto \frac{1}{x}$ is unbounded as $x \to 0^+$. In this case, we get

$$\int_0^1 \frac{1}{x}\,dx = \lim_{a\to 0^+}\int_a^1 \frac{1}{x}\,dx = \lim_{a\to 0}\big(\log(1) - \log(a)\big) = \infty.$$

Thus the improper integral $\int_0^1 \frac{1}{x}\,dx$ diverges, and we can assign it the value ∞.

EXERCISE 7.51. — Calculate $\displaystyle\int_0^1 \frac{1}{\sqrt{x}}\,dx$ and $\displaystyle\int_0^{\frac{\pi}{2}} \tan(x)\,dx$.

EXERCISE 7.52. — Decide for which $p \in \mathbb{R}_{\geq 0}$ the improper integral $\int_0^\infty x\sin(x^p)\,dx$ converges.

7.3.6 The Gamma Function

7.53. — The **Gamma-function** Γ is defined, for $s \in (0, \infty)$, by the convergent improper integral

30

$$\Gamma(s) = \int_0^\infty x^{s-1}e^{-x}\,dx. \tag{7.15}$$

To verify that this improper integral indeed converges, we examine the integration limits 0 and ∞ separately. For $0 < a < b$ we find, using integration by parts,

$$\int_a^b x^{s-1}e^{-x}\,dx = \tfrac{1}{s}\left[x^s e^{-x}\right]_a^b + \tfrac{1}{s}\int_a^b x^s e^{-x}\,dx. \tag{7.16}$$

We obtain

$$\begin{aligned}\int_0^b x^{s-1}e^{-x}\,dx &= \lim_{a\to 0}\left(\tfrac{1}{s}\left[x^s e^{-x}\right]_a^b + \tfrac{1}{s}\int_a^b x^s e^{-x}\,dx\right)\\ &= \tfrac{1}{s}b^s\exp(-b) + \frac{1}{s}\int_0^b x^s e^{-x}\,dx,\end{aligned}$$

where the integral on the right is an actual Riemann integral since the function $x^s e^{-x}$ is continuous on $[0, b]$. To investigate the upper limit of integration, we note that there exists $R > 0$ such that $e^x > x^{s+2}$ holds for all $x > R$. Thus

$$\int_0^\infty x^s e^{-x}\,dx \le \int_0^R x^s e^{-x}\,dx + \int_R^\infty x^{-2}\,dx < \infty$$

which shows that (7.16) converges as $b \to \infty$. Specifically, we obtain

$$\int_0^\infty x^{s-1}e^{-x}\,dx = \lim_{b\to\infty}\left(\tfrac{1}{s}b^s\exp(-b) + \frac{1}{s}\int_0^b x^s e^{-x}\,dx\right) = \frac{1}{s}\int_0^\infty x^s e^{-x}\,dx.$$

This shows that the gamma function satisfies the functional equation

$$\Gamma(s+1) = s\,\Gamma(s) \tag{7.17}$$

30 for all $s \in (0,\infty)$.

7.54. — The Gamma function extends the factorial function from $\mathbb{N}$ to $(0,\infty)$. In fact

$$\Gamma(1) = \int_0^\infty x^0 e^{-x}\,dx = e^0 - \lim_{x\to\infty} e^{-x} = 1$$

therefore (7.17) implies that

$$\Gamma(n+1) = n\Gamma(n) = n(n-1)\Gamma(n-1) = \ldots = n!\,\Gamma(1) = n! \qquad \forall n \in \mathbb{N}.$$

At the moment, it is not clear whether the Gamma function is continuous. Eventually, it will turn out that Γ is smooth. Also, for example, we cannot calculate the value $\Gamma(\frac{1}{2})$ with the integration methods we know so far, but we will see later by means of a two-dimensional integral that it is $\sqrt{\pi}$.

7.55. — David Hilbert (1862–1943), in his 1893 article [Hil1893], used improper integrals in the style of the Gamma function to prove that e (as first proved by Hermite in 1873) and π (as first proved by Lindemann in 1882) are transcendental. We note here that the irrationality of these numbers is much easier to prove. Transcendence proofs are generally much more difficult. The difficulty in making such statements is perhaps illustrated by the fact that it is still not known whether $e + \pi$ is a transcendental number or not.

7.4 Taylor Series

7.4.1 Taylor Approximation

The derivative $f'(x_0)$ of a real-valued differentiable function f on an interval gives the slope of the tangent of the graph of f at x_0. The corresponding affine function

$$x \mapsto f(x_0) + f'(x_0)(x - x_0)$$

approximates the function f to within an error $f(x) - y(x) = o(x - x_0)$ as $x \to x_0$. The "quality" of the approximation can be increased by considering higher polynomial approximations instead of affine approximations.

In this section, it will be convenient to use the following abuse of notation: Given $a, b \in \mathbb{R}$, irrespective of the order between a and b, $[a, b]$ denotes the interval between them. In other words, for all $a, b \in \mathbb{R}$, $[a, b]$ and $[b, a]$ denote the same interval.

We also recall that, if $a < b$, then

$$\left| \int_a^b f(x)\, dx \right| \le \int_a^b |f(x)|\, dx,$$

30 see Theorem 6.25. If instead $b < a$, then a minus sign appears (recall (7.2)) and we get

$$\left| \int_a^b f(x)\, dx \right| = \left| \int_b^a f(x)\, dx \right| \le \int_b^a |f(x)|\, dx = -\int_a^b |f(x)|\, dx = \left| \int_a^b |f(x)|\, dx \right|.$$

In conclusion, independent of the order of a with respect to b, we always have

$$\left| \int_a^b f(x)\, dx \right| \le \left| \int_a^b |f(x)|\, dx \right|.$$

We shall use this inequality many times in this section.

7.56. — Let $D \subseteq \mathbb{R}$ be an open interval, and $f : D \to \mathbb{C}$ be an n times differentiable function. The n-th **Taylor approximation** of f around a point $x_0 \in D$ is the polynomial function

$$P_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k. \tag{7.18}$$

The coefficients are chosen so that $P^{(k)}(x_0) = f^{(k)}(x_0)$ for $k \in \{0, \ldots, n\}$.

We will state and prove several versions of Taylor's Theorem. We begin with this first version:

Theorem 7.57: Taylor Expansion to Order n with Integral Remainder

Let $n \geq 1$, $f : [a, b] \to \mathbb{R}$ a n-times continuously differentiable function, and fix $x_0 \in [a, b]$. Then, for all $x \in [a, b]$,

$$f(x) = P_{n-1}(x) + \int_{x_0}^{x} f^{(n)}(t) \frac{(x-t)^{n-1}}{(n-1)!}\, dt, \tag{7.19}$$

where P_{n-1} is the $(n-1)$-th Taylor approximation of f defined in (7.18).

Remark 7.58. — In the above theorem, the assumption that f is a n-times continuously differentiable function guarantees that the integral of the continuous function $t \mapsto f^{(n)}(t) \frac{(x-t)^{n-1}}{(n-1)!}$ exists.

Proof. The proof follows by induction on n and integration by parts. If $n = 1$ then $f : [a, b] \to \mathbb{R}$ is continuously differentiable and, by Corollary 7.6, we get

$$f(x) = f(x_0) + \int_{x_0}^{x} f'(t) dt.$$

If f is twice continuously differentiable, we can apply integration by parts to the above integral with $u(t) = f'(t)$ and $v(t) = t - x$. Indeed, since $v' = 1$ and $v(x) = 0$, we get

30

$$\begin{aligned} f(x) &= f(x_0) + \int_{x_0}^{x} f'(t) v'(t) dt \\ &= f(x_0) + \big[f'(t)v(t)\big]_{x_0}^{x} - \int_{x_0}^{x} f''(t) v(t)\, dt \\ &= f(x_0) + f'(x_0)(x - x_0) + \int_{x_0}^{x} f''(t)(x-t)\, dt \\ &= P_1(x) + \int_{x_0}^{x} f^{(2)}(t) \frac{(x-t)^1}{1!}\, dt. \end{aligned}$$

This proves the case $n = 2$.

More generally, assume that the statement of the theorem is true for $n \geq 1$ and that $f : [a, b]] \to \mathbb{R}$ is a $(n+1)$ times continuously differentiable function. Then, by the induction hypothesis,

$$f(x) = \sum_{k=0}^{n-1} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k + \int_{x_0}^{x} f^{(n)}(t) \frac{(x-t)^{n-1}}{(n-1)!}\, dt$$

for all $x \in D$. Now, if we set $u(t) = f^{(n)}(t)$ and $v(t) = -\frac{(x-t)^n}{n!}$, since $v'(t) = \frac{(x-t)^{n-1}}{(n-1)!}$ it follows from integration by parts that

$$f(x) = \sum_{k=0}^{n-1} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k - \left[f^{(n)}(t) \frac{(x-t)^n}{n!}\right]_{x_0}^{x} + \int_{x_0}^{x} f^{(n+1)}(t) \frac{(x-t)^n}{n!} dt$$

$$= \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k + \int_{x_0}^{x} f^{(n+1)}(t)\frac{(x-t)^n}{n!}\,dt.$$

This proves the induction step, and hence the result. □

We can now state our two versions of Taylor's Approximation, using first the big-O notation, and then a refined version with the little-o notation.

Corollary 7.59: Taylor Approximation with Big-O

Let $n \geq 1$, $f : [a, b] \to \mathbb{R}$ a n-times continuously differentiable function, and fix $x_0 \in [a, b]$. Then, for all $x \in [a, b]$,

$$f(x) = P_{n-1}(x) + O(|x - x_0|^n) \qquad \text{as } x \to x_0. \tag{7.20}$$

Proof. Recalling (7.19), we have

$$f(x) = P_{n-1}(x) + \int_{x_0}^{x} f^{(n)}(t)\frac{(x-t)^{n-1}}{(n-1)!}\,dt.$$

Also, since $f^{(n)}$ is continuous on $[a, b]$, it is bounded (recall Theorem 8.25). Hence, there exists
30 a constant M such that $|f^{(n)}| \leq M$ on $[a, b]$. This implies that

$$|f(x) - P_{n-1}(x)| \leq \left|\int_{x_0}^{x} f^{(n)}(t)\frac{(x-t)^{n-1}}{(n-1)!}\,dt\right| \leq M\left|\int_{x_0}^{x} \frac{|x-t|^{n-1}}{(n-1)!}\,dt\right| \qquad \forall\, x \in [a, b].$$

Observe now that the sign of the integrand $(x - t)$ is constant on the interval $[x_0, x]$, so

$$\left|\int_{x_0}^{x} \frac{|x-t|^{n-1}}{(n-1)!}\,dt\right| = \left|\int_{x_0}^{x} \frac{(x-t)^{n-1}}{(n-1)!}\,dt\right|$$

and the last integral can be computed with a change of variable: setting $s = x - t$ we get

$$\left|\int_{x_0}^{x} \frac{(x-t)^{n-1}}{(n-1)!}\,dt\right| = \left|\int_{0}^{x-x_0} \frac{s^{n-1}}{(n-1)!}\,ds\right| = \left|\frac{(x-x_0)^n}{n!}\right| = \frac{|x-x_0|^n}{n!}$$

This shows that

$$|f(x) - P_{n-1}(x)| \leq M\frac{|x-x_0|^n}{n!} \qquad \forall\, x \in [a, b],$$

therefore $f(x) - P_{n-1}(x) = O(|x - x_0|^n)$ as $x \to x_0$, concluding the proof. □

We now show that by replacing P_{n-1} with P_n, we can improve the previous result using the little-o notation.

Corollary 7.60: Taylor Approximation with Little-o

Let $n \geq 1$, $f : [a,b] \to \mathbb{R}$ a n-times continuously differentiable function, and fix $x_0 \in [a,b]$. Then, for all $x \in [a,b]$,

$$f(x) = P_n(x) + o(|x - x_0|^n) \qquad \text{as } x \to x_0. \tag{7.21}$$

Proof. Thanks to Theorem 7.57, we can write

$$f(x) = P_{n-1}(x) + \int_{x_0}^{x} f^{(n)}(t) \frac{(x-t)^{n-1}}{(n-1)!}\, dt.$$

Also, we note that

$$\int_{x_0}^{x} f^{(n)}(x_0) \frac{(x-t)^{n-1}}{(n-1)!}\, dt = f^{(n)}(x_0) \int_{x_0}^{x} \frac{(x-t)^{n-1}}{(n-1)!}\, dt = f^{(n)}(x_0) \frac{(x-x_0)^n}{n!}.$$

Therefore

30

$$\begin{aligned} f(x) &= P_{n-1}(x) + f^{(n)}(x_0)\frac{(x-x_0)^n}{n!} + \int_{x_0}^{x} \big(f^{(n)}(t) - f^{(n)}(x_0)\big) \frac{(x-t)^{n-1}}{(n-1)!}\, dt \\ &= P_n(x) + \int_{x_0}^{x} \big(f^{(n)}(t) - f^{(n)}(x_0)\big) \frac{(x-t)^{n-1}}{(n-1)!}\, dt. \end{aligned} \tag{7.22}$$

Now, given $\varepsilon > 0$, it follows from the continuity of $f^{(n)}$ at x_0 that there exists $\delta > 0$ such that $|f^{(n)}(x) - f^{(n)}(x_0)| < \varepsilon$ for all $x \in (x_0 - \delta, x_0 + \delta) \cap [a,b]$. Hence, if $x \in (x_0 - \delta, x_0 + \delta) \cap [a,b]$, we can bound the integrand in the last integral by

$$\left| \big(f^{(n)}(t) - f^{(n)}(x_0)\big) \frac{(x-t)^{n-1}}{(n-1)!} \right| \leq |f^{(n)}(t) - f^{(n)}(x_0)| \frac{|x-t|^{n-1}}{(n-1)!} < \varepsilon \frac{|x-t|^{n-1}}{(n-1)!} \qquad \forall\, t \in [x_0, x].$$

This implies

$$\begin{aligned} |f(x) - P_n(x)| &\leq \left| \int_{x_0}^{x} \big(f^{(n)}(t) - f^{(n)}(x_0)\big) \frac{(x-t)^{n-1}}{(n-1)!}\, dt \right| \\ &< \varepsilon \left| \int_{x_0}^{x} \frac{|x-t|^{n-1}}{(n-1)!}\, dt \right| = \varepsilon \frac{|x-x_0|^n}{n!} \leq \varepsilon |x - x_0|^n, \end{aligned}$$

where the last integral has been computed as in the proof of Corollary 7.59. This proves that

$$\frac{|f(x) - P_n(x)|}{|x - x_0|^n} < \varepsilon \qquad \forall\, x \in (x_0 - \delta, x_0 + \delta) \cap [a,b],$$

which shows that $f(x) - P_n(x) = o(|x - x_0|^n)$ as $x \to x_0$. □

EXAMPLE 7.61. — 1. If $f \in C^1$, then Corollary 7.59 implies that

$$f(x) = f(x_0) + O(|x - x_0|) \qquad \text{as } x \to x_0,$$

while Corollary 7.60 yields

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + o(|x - x_0|) \qquad \text{as } x \to x_0.$$

2. If $f \in C^2$, then Corollary 7.59 gives

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + O(|x - x_0|^2) \qquad \text{as } x \to x_0,$$

while Corollary 7.60 yields

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2} f''(x_0)(x - x_0)^2 + o(|x - x_0|^2) \qquad \text{as } x \to x_0.$$

3. If f is smooth, then Corollary 7.60 yields

$$f(x) = P_n(x) + o(|x - x_0|^n) \qquad \text{as } x \to x_0,$$

31 while Corollary 7.59 applied with $n+1$ in place of n gives

$$f(x) = P_n(x) + O(|x - x_0|^{n+1}) \qquad \text{as } x \to x_0.$$

Hence, while in the case where f has a finite number of derivatives Corollary 7.60 provides a stronger result, in the case when f is smooth, the bound on $f - P_n$ provided by Corollary 7.59 is more convenient.

While for proving Corollary 7.60 the continuity of $f^{(n)}$ plays a crucial role, in the proof of Corollary 7.59 we mainly used that $f^{(n)}$ is bounded (the continuity of $f^{(n)}$ is needed only to guarantee that $f^{(n)}$ is integrable). In fact, it is possible to prove Corollary 7.59 under the weaker assumption that the n-th derivative exists and is bounded (but is not necessarily continuous). For this, we first prove the following alternative version of Taylor Theorem. Note that in the case $n = 1$, this result corresponds to the Mean Value Theorem 5.31.

THEOREM 7.62: TAYLOR EXPANSION TO ORDER n WITH LAGRANGE REMAINDER

*Let $n \geq 1$, $f : [a, b] \to \mathbb{R}$ a n-times continuously differentiable function, and fix $x_0 \in [a, b]$. Then, for all $x \in [a, b]$ there exists $\xi_L \in (x_0, x)$ such that the **Lagrange remainder** formula holds:*

$$f(x) - P_{n-1}(x) = \frac{1}{n!} f^{(n)}(\xi_L)(x - x_0)^n. \tag{7.23}$$

Proof. (Extra Material) Fix $x \in (a,b)$ and consider the function $F : (a,b) \to \mathbb{R}$ defined as

$$F(t) = f(t) + f^{(1)}(t)(x-t) + \ldots + \frac{f^{(n-1)}(t)}{(n-1)!}(x-t)^{n-1} = \sum_{k=0}^{n-1} \frac{f^{(k)}(t)}{k!}(x-t)^k. \tag{7.24}$$

Then $F(x) = f(x)$ and $F(x_0) = P_{n-1}(x)$. Also, its derivative is given by

$$\begin{aligned} F'(t) &= \sum_{k=0}^{n-1} \frac{f^{(k+1)}(t)}{k!}(x-t)^k - \sum_{k=0}^{n-1} \frac{f^{(k)}(t)}{k!}k(x-t)^{k-1} \\ &= \sum_{k=0}^{n-1} \frac{f^{(k+1)}(t)}{k!}(x-t)^k - \sum_{k=1}^{n-1} \frac{f^{(k)}(t)}{(k-1)!}(x-t)^{k-1} \\ &= \sum_{k=0}^{n-1} \frac{f^{(k+1)}(t)}{k!}(x-t)^k - \sum_{k=0}^{n-2} \frac{f^{(k+1)}(t)}{k!}(x-t)^k = \frac{f^{(n)}(t)}{(n-1)!}(x-t)^{n-1}. \end{aligned}$$

Hence, applying Cauchy Mean Value Theorem 5.35 in the interval $[x_0, x]$ to the functions $f(t) = F(t)$ and $g(t) = -(x-t)^n$ we deduce the existence of a point $\xi_L \in (x_0, x)$ such that

$$\frac{f(x) - P_{n-1}(x)}{(x-x_0)^n} = \frac{F(x) - F(x_0)}{g(x) - g(x_0)} = \frac{f'(\xi_L)}{g'(\xi_L)} = \frac{\frac{f^{(n)}(\xi_L)}{(n-1)!}(x-\xi_L)^{n-1}}{n(x-\xi_L)^{n-1}} = \frac{f^{(n)}(\xi_L)}{n!}.$$

31 This implies (7.23) and concludes the proof. □

Corollary 7.63: Taylor Approximation with Big-O, 2nd version

Let $n \geq 1$, $f : [a,b] \to \mathbb{R}$ a n-times continuously differentiable function, and fix $x_0 \in [a,b]$. Assume that there exists $M > 0$ and such that $|f^{(n)}(x)| \leq M$ for all $x \in [a,b]$. Then

$$f(x) = P_{n-1}(x) + O(|x-x_0|^n) \qquad \text{as } x \to x_0. \tag{7.25}$$

Proof. (Extra Material) Given $x \in [a,b]$, we apply (7.23) to find a point $\xi_L \in (x_0, x)$ such that

$$f(x) - P_{n-1}(x) = \frac{1}{n!} f^{(n)}(\xi_L)(x-x_0)^n.$$

Since $|f^{(n)}(\xi_L)| \leq M$, it follows that

$$|f(x) - P_{n-1}(x)| \leq \frac{M}{n!}|x-x_0|^n \qquad \forall\, x \in [a,b],$$

therefore $f(x) - P_{n-1}(x) = O(|x-x_0|^n)$ as $x \to x_0$, as desired. □

Another version of Taylor formula is the one with the so-called Cauchy remainder. We discuss it in the following exercise.

EXERCISE 7.64. — Let $f : [a, b] \to \mathbb{R}$ be a n-times differentiable function, and fix $x, x_0 \in [a, b]$. Prove that there exists $\xi_C \in [x_0, x]$ such that the **Cauchy remainder** formula holds:

$$f(x) - P_{n-1}(x) = \frac{1}{(n-1)!} f^n(\xi_C)(x - \xi_C)^{n-1}(x - x_0). \tag{7.26}$$

Hint: Consider the function F defined in (7.24) and apply to it the Mean Value Theorem 5.31 in the interval $[x_0, x]$.

EXAMPLE 7.65. — We can use the Taylor approximation to refine the discussion in Section 5.2.1. Let $f : (a, b) \to \mathbb{R}$ be a n-times continuously differentiable function. Suppose $x_0 \in (a, b)$ satisfies

$$f'(x_0) = \ldots = f^{(n-1)}(x_0) = 0.$$

Then the following implications hold:

- If $f^{(n)}(x_0) < 0$ and n is even, then f has an isolated local maximum in x_0.
- If $f^{(n)}(x_0) > 0$ and n is even, then f has an isolated local minimum in x_0.
- If $f^{(n)}(x_0) \neq 0$ and n is odd, then x_0 is not a local extremum of f.

All three statements follow from (7.23), which, in this case, takes the form

31

$$f(x) = f(x_0) + \frac{1}{n!} f^{(n)}(\xi_L)(x - x_0)^n, \qquad \xi_L \in (x_0, x).$$

Indeed, if $f^{(n)}(x_0) > 0$, by continuity there exists $\delta > 0$ such that $f^{(n)}(\xi_L) > 0$ for $\xi_L \in (x_0, x) \subset (x_0 - \delta, x_0 + \delta)$. If n is even, then $(x - x_0)^n > 0$ for $x \neq 0$ and we deduce that $f(x) > f(x_0)$ for $x \in (x_0 - \delta, x_0 + \delta)$ with $x \neq 0$. If n is odd, then $(x - x_0)^n$ changes sign when considering $x > x_0$ and $x < x_0$, so x_0 is not a local extremum of f.

On the other hand, if $f^{(n)}(x_0) < 0$ and n is even, the same argument as above shows that $f(x) < f(x_0)$ for $x \in (x_0 - \delta, x_0 + \delta)$ with $x \neq 0$, while in the case n odd x_0 is not a local extremum of f.

Applet 7.66 (Taylor approximations)**.** *We present some Taylor approximations at shiftable footpoints to known functions.*

7.4.2 Analytic Functions

Motivated by Taylor's Theorem, one may be tempted to say that if one replaces the polynomial P_n with the **Taylor series**

$$\sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k,$$

then one should recover $f(x)$. Unfortunately, this is false, and functions that satisfy such a property are rather special.

Note that the Taylor series is centered at x_0 instead of 0 (i.e., x^n is replaced with $(x-x_0)^n$). Hence, all theorems about power series from Section 4.4 still hold, but taking into account that now x_0 plays the role of the center. In particular, if the series has radius of convergence $R > 0$, then it converges for all $x \in (x_0 - R, x_0 + R)$, while it diverges for $|x - x_0| > R$.

DEFINITION 7.67: ANALYTIC FUNCTIONS

Let $I \subseteq \mathbb{R}$ be an interval and $x_0 \in I$. A smooth function $f : I \to \mathbb{R}$ is called *analytic* at x_0 if there exists $\delta > 0$ such that the Taylor series of f around x_0 has radius of convergence $R > \delta$ and

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n \qquad \forall\, x \in (x_0 - \delta, x_0 + \delta) \cap I.$$

We say f is analytic in I if f is analytic at all points in I.

In other words, analytic functions $f : I \to \mathbb{R}$ are characterized by the fact that, for every point $x_0 \in I$, there exists a power series that converges to f in a neighborhood of x_0.

31 As the next example shows, there are smooth functions f whose Taylor series converges to a function different from f.

EXAMPLE 7.68. — Consider the function $\psi : \mathbb{R} \to \mathbb{R}$ defined by

$$\psi : x \in \mathbb{R} \mapsto \begin{cases} \exp\left(-\frac{1}{x}\right) & \text{if } x > 0 \\ 0 & \text{if } x \leq 0 \end{cases}.$$

As shown in Exercise 5.25, ψ is smooth on $\mathbb{R}$ and satisfies $\psi^{(n)}(0) = 0$ for all $n \in \mathbb{N}$. Hence, the Taylor series of the function ψ at the point $x_0 = 0$ is the zero series:

$$\sum_{n=0}^{\infty} \frac{\psi^{(n)}(0)}{n!} x^n = \sum_{n=0}^{\infty} \frac{0}{n!} x^n = 0.$$

This series has an infinite radius of convergence and converges to the function 0. Since $\psi(x) > 0$ holds for all $x > 0$, the Taylor series does not converge to ψ, and so ψ is not analytic at the point $x_0 = 0$.

The next result provides a criterion that guarantees that the Taylor series of f converges to f in a neighborhood of x_0.

Theorem 7.69: A Criterion for Analyticity at x_0

Let $I \subseteq \mathbb{R}$ be an interval and $f : I \to \mathbb{R}$ a smooth function. Given $x_0 \in I$, assume that there exist constants $r, c, A > 0$ such that

$$|f^{(n)}(x)| \le cA^n n! \qquad \text{for all } x \in (x_0 - r, x_0 + r) \cap I, \quad n \in \mathbb{N}.$$

Then f is analytic at x_0.

Proof. We first estimate the radius of convergence R of the Taylor series. If we define $a_n = \frac{f^{(n)}(x_0)}{n!}$, then the Taylor series is equal to $\sum_{n=0}^{\infty} a_n (x - x_0)^n$. Also, thanks to our assumption on the size of $|f^{(n)}|$ it follows that

$$|a_n| \le \frac{cA^n n!}{n!} = cA^n.$$

This implies that

$$\limsup_{n\to\infty} \sqrt[n]{|a_n|} \le \limsup_{n\to\infty} \sqrt[n]{cA^n} = \limsup_{n\to\infty} \sqrt[n]{c}A = A,$$

thus, by the definition of radius of convergence (see Definition 4.39), $R \ge \frac{1}{A}$.

Now, fix $\delta < \min\left\{r, \frac{1}{A}\right\}$. Given $x \in (x_0 - \delta, x_0 + \delta) \cap I$, we apply (7.23) and our assumption
31 on the size of $|f^{(n)}|$ to deduce that

$$|f(x) - P_{n-1}(x)| \le \frac{1}{n!} |f^{(n)}(\xi_L)|\, |x - x_0|^n \le cA^n |x - x_0|^n \le c(A\delta)^n.$$

Since $A\delta < 1$ by assumption (so, in particular, $\delta < \frac{1}{A} \le R$), letting $n \to \infty$ we conclude that

$$f(x) = \lim_{n\to\infty} P_{n-1}(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n \qquad \forall\, x \in (x_0 - \delta, x_0 + \delta) \cap I,$$

as desired. □

As a direct consequence of Theorem 7.69, we immediately deduce the following:

Corollary 7.70: A Criterion for Analyticity

Let $f : [a, b] \to \mathbb{R}$ be a smooth function, and assume there exist constants $c, A > 0$ such that

$$|f^{(n)}(x)| \le cA^n n! \qquad \text{for all } x \in [a, b], \quad n \in \mathbb{N}. \tag{7.27}$$

Then f is analytic on $[a, b]$.

Exercise 7.71. — 1. Show that the functions $\exp, \sin, \sinh$ satisfy the property (7.27) on any interval $[a,b] \subset \mathbb{R}$.

31 2. Show that the function log satisfies (7.27) on any interval $[a,b] \subset (0,\infty)$.

3. Let $f, g : [a,b] \to \mathbb{R}$ be functions satisfying (7.27). Show that $f+g$ and $f \cdot g$ also satisfy this property (possibly with different constants c and A).

Chapter 8

Ordinary Differential Equations

Setting up and solving differential equations stands as a primary practical use of calculus. These equations are instrumental in addressing a wide range of challenges in fields such as physics, chemistry, biology, and more. Moreover, disciplines like structural analysis, modern economics, and information technology heavily rely on differential equations, making them indispensable in these domains.

8.1 Ordinary Differential Equations (ODEs)

Ordinary differential equations (ODEs) are fundamental in mathematics and various applied sciences, offering a vital framework for modeling and understanding diverse phenomena. These equations involve functions of one variable and their derivatives, providing a deep insight into the behavior of many physical, biological, and economic systems. Although derivatives are usually denoted using $'$ (so f', f'', etc.), it is common to use a $\dot{}$ to denote derivatives with respect to time (so $\dot{f}$, $\ddot{f}$, etc.).

Definition 8.1: ODEs

32 An ODE is a relationship that involves a function $u : \mathbb{R} \to \mathbb{R}$ of a real variable $x \in \mathbb{R}$, and its derivatives. The general form of an n-th order ODE is

$$G\big(x, u(x), u'(x), u''(x), \ldots, u^{(n)}(x)\big) = 0, \tag{8.1}$$

where $G : \mathbb{R}^{n+1} \to \mathbb{R}$ is a given function.

ODEs can be categorized according to different criteria:

1. *Order:* An ODE is of order n if $u^{(n)}$ is the maximal derivative appearing in the ODE. For instance:

   (a) $u'' + u = 0 \rightsquigarrow$ second order.

   (b) $u^{(3)} = x^2u + x \rightsquigarrow$ third order.
   (c) $(u')^2 + u - x^3 = 0 \rightsquigarrow$ first order.

2. *Linearity:* An ODE is linear if it is linear in u and its derivatives. Otherwise, it is nonlinear.
   (a) $u'' + u = 0 \rightsquigarrow$ linear.
   (b) $u'' + u^2 = 0 \rightsquigarrow$ nonlinear.
   (c) $u'' + u'u = 0 \rightsquigarrow$ nonlinear.
   (d) $u^{(3)} = x^2u + x \rightsquigarrow$ linear.
   (e) $(u')^2 + u - x^3 = 0 \rightsquigarrow$ nonlinear.

3. *Homogeneity (for linear ODEs):* A linear ODE is homogeneous if all terms involve the function or its derivatives (this is equivalent to asking that if u is a solution, then Au is a solution for all $A \in \mathbb{R}$). It is non-homogeneous if there is an additional term independent of the function.
   (a) $u'' + u = 0 \rightsquigarrow$ homogeneous.
   (b) $u^{(3)} = x^2u + x \rightsquigarrow$ non-homogeneous.
   (c) $u^{(3)} = x^2u \rightsquigarrow$ homogeneous.

32

ODEs are pivotal in fields such as physics, engineering, biology, and economics, as they model phenomena where the rate of change of a quantity is related to the quantity itself.

Example 8.2. — Here we present some classic examples of ODEs and their applications:

1. *Newton's Law of Cooling:* In the field of heat transfer, Newton's Law of Cooling plays a pivotal role in understanding the dynamics of temperature change. This law states that:

   > "The rate of heat loss of a body is directly proportional to the difference in the temperatures between the body and its surrounding environment."

   This principle leads to the formulation of a differential equation that governs the temperature dynamics of an object. The equation is expressed as:

$$\dot{T}(t) = -k\big(T(t) - T_{\text{env}}\big),$$

   where:

   - $T(t)$ represents the temperature of the object at time t;
   - T_{env} is the temperature of the surrounding environment;

- k is a positive constant, which represents the proportionality factor in the rate of heat transfer.

This ODE is linear, non-homogeneous, and of first order

The equation illustrates how the rate of temperature change in an object is contingent upon the temperature difference with its environment, a concept widely applied in engineering, meteorology, and various scientific studies related to heat transfer.

2. *Harmonic Oscillator:* In the realm of classical mechanics, the concept of a harmonic oscillator is central to understanding various physical systems. It refers to a system where, upon being displaced from its equilibrium position, there is a restoring force F that is directly proportional to the displacement x.

   The fundamental principle of the simple harmonic oscillator can be described as follows:

   > "A system displaced from an equilibrium position experiences a restoring force F, which is proportional to the displacement x."

   From Newton's second law of motion $\ddot{x}(t) = -F(x(t))$, this leads to the formulation of the following ODE for the simple harmonic oscillator:

   $$\ddot{x}(t) + \omega^2 x(t) = 0,$$

32 where ω denotes the angular frequency of the oscillations. This ODE is linear, homogeneous, and of second order.

   In real oscillators, friction slows the motion of the system. In many vibrating systems, the frictional force can be modeled as being proportional to the velocity $\dot{x}$ of the object. This leads to the formulation of the following ODE for the damped harmonic oscillator:

   $$\ddot{x}(t) + 2\zeta\omega\dot{x}(t) + \omega^2 x(t) = 0,$$

   where $\zeta \geq 0$ is called "damping ratio".

   This equation is fundamental in the study of oscillatory systems and finds applications across various fields, including physics, engineering, and even economics.

3. *Logistic Population Growth:* In population dynamics, the logistic growth model is a fundamental concept that elucidates the effects of resource limitations on population growth by introducing the concept of carrying capacity, a threshold beyond which resource scarcity hinders a further increase in population.

   The logistic growth model is mathematically articulated as

   $$\dot{P}(t) = rP(t)\left(1 - \frac{P(t)}{K}\right),$$

   where:

- $P(t)$ denotes the population size at time t;
- r represents the intrinsic growth rate, indicating the potential increase rate in the absence of resource constraints;
- K is the carrying capacity, defined as the maximum sustainable population size given the prevailing environmental conditions.

This ODE is nonlinear and of first order.

This model is used to understand real-world population dynamics, as it accounts for the practical constraints of resource availability and environmental capacity.

4. *Bessel Equation:* The Bessel equation is a significant differential equation in mathematics and physics, particularly in problems involving cylindrical coordinates. For a parameter $\alpha \in \mathbb{R}$, the Bessel equation is given by

$$x^2 u''(x) + x u'(x) + (x^2 - \alpha^2) u(x) = 0.$$

This ODE is linear, homogeneous, and of second order.

Solutions to the Bessel equation are known as Bessel functions. These functions are particularly significant in physics for values of $\alpha \in \mathbb{Z}$ and $\alpha + \frac{1}{2} \in \mathbb{Z}$. They describe various physical phenomena such as heat conduction or wave propagation in cylindrical
32 media and also appear in quantum mechanics.

5. *Airy Equation:* Another related equation is the Airy equation, which is expressed as

$$u''(x) - \alpha^2 x u(x) = 0.$$

The Airy function is a special solution to this differential equation. It is particularly relevant in quantum mechanics, where it is correlated with the Schrödinger equation for a particle in a triangular potential well.

8.3. — So far, we have only talked about single differential equations, but one may also study systems of differential equations. In addition, solutions of a differential equation are required to satisfy some "initial conditions" such as $u(0) = 0$ (for example, this can correspond to prescribing position at time 0) and/or $u'(0) = 1$ (this can correspond to prescribing velocity at time 0). These are called **boundary conditions**.

8.1.1 Linear First Order ODEs

In this section, we show how to solve linear differential equations of first order. We specify an interval $I \subseteq \mathbb{R}$ that is non-empty and does not consist of a single point.

We start with the homogeneous case.

Proposition 8.4: Homogeneous Linear 1st Order ODEs

Let $f : I \to \mathbb{R}$ be a continuous function and consider the homogeneous first-order linear ODE

$$u'(x) + f(x)u(x) = 0 \qquad \forall\, x \in I. \tag{8.2}$$

Let $F : I \to \mathbb{R}$ be a primitive of f. Then, all C^1 solutions $u : I \to \mathbb{R}$ of the above ODE are of the form

$$u(x) = Ae^{-F(x)}, \qquad \text{with } A \in \mathbb{R}.$$

In other words, the set of solutions of (8.2) forms a one-dimensional linear subspace of $C^1(I)$.

Proof. Given $A \in \mathbb{R}$, define $u(x) = Ae^{-F(x)}$. Then

$$u'(x) = -F'(x)Ae^{-F(x)} = -f(x)Ae^{-F(x)} = -f(x)u(x) \qquad \forall\, x \in I,$$

that is, u solves the ODE.

Vice versa, let $u \in C^1(I)$ solve (8.2) and define $v(x) = e^{F(x)}u(x)$. Then

$$v'(x) = \big(e^{F(x)}\big)'u(x) + e^{F(x)}u'(x) = e^{F(x)}\big(f(x)u(x) - f(x)u(x)\big) = 0 \qquad \forall\, x \in I.$$

32 By Corollary 5.46, we deduce that $v(x) = A$ for some $A \in \mathbb{R}$, or equivalently $u(x) = Ae^{-F(x)}$. □

Remark 8.5. — As we have seen, solutions of (8.2) are defined in terms of a primitive of f. Since primitives are defined up to a constant, one could wonder what happens if one replaces F by $F + C$ for some constant $C \in \mathbb{R}$. This would correspond to replacing $Ae^{-F(x)}$ with $Ae^{-C}e^{-F(x)}$, but since $A \in \mathbb{R}$ is arbitrary, this plays no essential role in the final statement.

8.6. — We can now investigate the solvability of non-homogeneous linear first order ODEs, namely

$$u'(x) + f(x)u(x) = g(x) \qquad \forall\, x \in I. \tag{8.3}$$

To motivate the next result, we look for a special solution by applying the method of **variation of constants**. The idea is that, instead of looking for solutions of the form $x \mapsto Ae^{-F(x)}$ (that we know solve the homogeneous equation), we look for solutions $u(x) = H(x)e^{-F(x)}$ for some C^1 function $H : I \to \mathbb{R}$. With this choice it follows that

$$u'(x) = \Big(H'(x)e^{-F(x)} - H(x)F'(x)e^{-F(x)}\Big) = H'(x)e^{-F(x)} - f(x)u(x).$$

Hence, if we want u to solve (8.3) we need to impose that $H'(x)e^{-F(x)} = g(x)$, or equivalently, H is a primitive of $g(x)e^{F(x)}$.

Motivated by this discussion, we can now prove the following:

Proposition 8.7: Non-Homogeneous Linear 1st Order ODEs

Let $f, g : I \to \mathbb{R}$ be continuous functions and consider the non-homogeneous first-order linear ODE (8.3)*. Let $F : I \to \mathbb{R}$ be a primitive of f, and $H : I \to \mathbb{R}$ a primitive of ge^F. Then, all C^1 solutions $u : I \to \mathbb{R}$ of the above ODE are of the form*

$$u(x) = H(x)e^{-F(x)} + Ae^{-F(x)}, \qquad \text{with } A \in \mathbb{R}.$$

In other words, the set of solutions of (8.3) *forms a one-dimensional affine subspace of $C^1(I)$.*

Proof. First, given $A \in \mathbb{R}$ and $u(x) = (H(x) + A)e^{-F(x)}$, it follows that

$$\begin{aligned} u'(x) &= H'(x)e^{-F(x)} - F'(x)(H(x) + A)e^{-F(x)} \\ &= g(x)e^{F(x)}e^{-F(x)} - f(x)u(x) = g(x) - f(x)u(x) \qquad \forall\, x \in I, \end{aligned}$$

so u solves (8.3).

Vice versa, if u solves (8.3) then $v(x) = u(x) - H(x)e^{-F(x)}$ solves

32

$$\begin{aligned} v'(x) &= u'(x) - H'(x)e^{-F(x)} - F'(x)H(x)e^{-F(x)} \\ &= u'(x) - g(x)e^{F(x)}e^{-F(x)} + f(x)H(x)e^{-F(x)} \\ &= u'(x) - g(x) + f(x)\big(u(x) - v(x)\big) = -f(x)v(x). \end{aligned}$$

In other words, $v(x)$ solves (8.2), so Proposition 8.7 implies that $v(x) = Ae^{-F(x)}$ for some $A \in \mathbb{R}$. Since $u(x) = v(x) + H(x)e^{-F(x)}$, this proves the result. □

The previous results give us formulas to solve every linear first order ODE. However, in a concrete case, the difficulty will be determining the primitive F of f and then the one of $g(x)e^{F(x)}$. As we have seen above, solutions are uniquely determined up to a free parameter $A \in \mathbb{R}$. This will be used to impose the boundary condition.

Example 8.8. — We want to solve the ODE

$$u'(x) - 2xu(x) = e^{x^2}, \qquad u(0) = 1, \tag{8.4}$$

on $\mathbb{R}$. Following Proposition 8.7, we set $f(x) = -2x$ and $g(x) = e^{x^2}$. Then a primitive of f is the function $F(x) = -x^2$, while a primitive of $g(x)e^{F(x)} = e^{x^2}e^{-x^2} = 1$ is given by x. So, u must be of the form $u(x) = (x + A)e^{x^2}$. Imposing the boundary condition $u(0) = 1$ we obtain $A = 1$, therefore the solution to the above ODE is given by

$$u(x) = (x + 1)e^{x^2}. \tag{8.5}$$

Remark 8.9. — If one forgets the formula from Proposition 8.7, one can try to remember the following procedure to solve (8.3).

Recalling that multiplying by a function of the form $e^{w(x)}$ for some function w is "useful" (based on what we have seen in previous pages), we multiply (8.3) by $e^{w(x)}$, so to get

$$u'(x)e^{w(x)} + f(x)u(x)e^{w(x)} = g(x)e^{w(x)}.$$

Then we look for a special choice of w so that the left-hand side above is equal to $\big(u(x)e^{w(x)}\big)'$. This is equivalent to asking

$$u'(x)e^{w(x)} + w'(x)u(x)e^{w(x)} = u'(x)e^{w(x)} + f(x)u(x)e^{w(x)} \qquad \Longrightarrow \qquad w'(x) = f(x).$$

Hence, if we choose $w = F$ a primitive of f (note that here we can choose any primitive of f without worrying about the additional constant C, since all that matters is that $F' = f$),
32 then

$$\Big(u(x)e^{F(x)}\Big)' = g(x)e^{F(x)},$$

therefore

$$u(x)e^{F(x)} = \int ge^F + A,$$

for some $A \in \mathbb{R}$. In other words, if H is a primitive of ge^F then

$$u(x)e^{F(x)} = H(x) + A \qquad \Longrightarrow \qquad u(x) = H(x)e^{-F(x)} + Ae^{-F(x)}.$$

Exercise 8.10. — Find a solution to the ODE

$$u' - \big(\tfrac{4}{x} + 1\big)u = x^4 \qquad\qquad u(1) = 1$$

in the interval $(0, \infty)$.

8.1.2 Autonomous First Order ODEs

Autonomous first-order ordinary differential equations (ODEs) are a class of differential equations where the rate of change of a variable is a function of the variable itself, independent of the independent variable (often time). Mathematically, they are expressed as:

33

$$u'(x) = f(u(x)) \tag{8.6}$$

for some continuous function $f : \mathbb{R} \to \mathbb{R}$. The general solution of an autonomous first-order ODE can be found using **separation of variables**. The process involves rearranging the equation to separate the functions of u and x, and then integrating both sides. More precisely, assuming that $f(u(x)) \neq 0$, we can divide both side and obtain

$$\frac{u'(x)}{f(u(x))} = 1.$$

Integrating both side and using the change of variable formula (7.8), we get

$$\int \frac{1}{f(u)}\,du = \int \frac{1}{f(u(x))}u'(x)\,dx = \int 1\,dx + C = x + C, \tag{8.7}$$

where C is the constant of integration. Hence, if H is a primitive of $\frac{1}{f}$, then we get

$$H(u(x)) = x + C \qquad \Longrightarrow \qquad u(x) = H^{-1}(x + C).$$

Note that, since by assumption $\frac{1}{f} \neq 0$ on the domain of integration (otherwise we could not divide by $f(u(x))$), it means that $H' = \frac{1}{f} \neq 0$, so H is invertible (since it is either strictly increasing or strictly decreasing).

EXAMPLE 8.11. — Consider the logistic growth model used in population dynamics:

$$u'(x) = ru(x)\left(1 - \frac{u(x)}{K}\right) \tag{8.8}$$

33 where $u(x) \in (0, K)$ represents the population at "time" x, r is the growth rate, and K is the carrying capacity (see Example 8.2(3)). To solve this, we rearrange and integrate:

$$\frac{Ku'(x)}{u(x)\,(K - u(x))} = r \qquad \Longrightarrow \qquad \int \frac{K}{u\,(K-u)}\,du = \int r\,dx = rx + C$$

The left-hand side is the integral of a rational function, that can be solved as discussed in Section 7.3.4: one can observe that

$$\frac{K}{u\,(K-u)} = \frac{1}{u} + \frac{1}{K-u},$$

hence (recall that, in this model, $0 < u < K$)

$$\int \frac{K}{u\,(K-u)}\,du = \int \frac{1}{u}\,du + \int \frac{1}{K-u}\,du = \log u - \log(K-u) = \log\left(\frac{u}{K-u}\right).$$

This implies

$$\begin{aligned} \log\left(\frac{u(x)}{K-u(x)}\right) = C + rx \qquad &\Longrightarrow \qquad \frac{u(x)}{K-u(x)} = e^{C+rx} \\ &\Longrightarrow \qquad u(x) = Ke^{C+rx} - u(x)e^{C+rx} \\ &\Longrightarrow \qquad \left(1 + e^{C+rx}\right)u(x) = Ke^{C+rx} \\ &\Longrightarrow \qquad u(x) = K\frac{e^{C+rx}}{1+e^{C+rx}}. \end{aligned}$$

If $u_0 \in (0, K)$ denotes the initial population, then setting $x = 0$ (here x plays the role of time) we get

$$u_0 = K\frac{e^C}{1+e^C} \qquad \Longrightarrow \qquad e^C = \frac{u_0}{K - u_0},$$

which gives

$$u(x) = \frac{Ku_0}{u_0 + (K - u_0)\,e^{-rx}}.$$

Example 8.12. — Exponential decay can be modeled by an autonomous first-order ODE:

$$u'(x) = -ku(x)$$

where $k > 0$ is the decay constant. This ODE could be solved using (8.4), but we show an alternative approach via separation of variables.

More precisely, if u is identically zero then there is nothing to prove. Otherwise, if u is non-zero in some interval, then we can divide by u and integrate to get

$$\int \frac{1}{u}\,du = -k\int dx \qquad \Longrightarrow \qquad \log|u| = -kx + C. \tag{8.9}$$

This proves that, in each interval I where u does not vanish, there exists a constant $C \in \mathbb{R}$ such
33 that $|u(x)| = e^C e^{-kx}$. Since u has to be continuous, this implies that either $u(x) = e^C e^{-kx}$ or $u(x) = -e^C e^{-kx}$ on the whole $\mathbb{R}$. Therefore, in conclusion, $u(x) = ae^{-kx}$ for some $a \in \mathbb{R}$. Imposing the condition that $u(0) = u_0$, this gives

$$u(x) = u_0 e^{-kx} \qquad \forall\, x \in \mathbb{R}. \tag{8.10}$$

Remark 8.13 (Method of Separation of Variables). — The method described above can be generalized to ODEs of the form

$$u'(x) = f(u(x))g(x) \tag{8.11}$$

for some continuous functions $f, g : \mathbb{R} \to \mathbb{R}$. More precisely, assuming as before that $f(u(x)) \neq 0$, we can divide both side and obtain

$$\frac{u'(x)}{f(u(x))} = g(x).$$

Integrating both side and using the change of variable formula (7.8), we get

$$\int \frac{1}{f(u)}\,du = \int \frac{1}{f(u(x))}u'(x)\,dx = \int g(x)\,dx + C, \tag{8.12}$$

where C is the constant of integration. Hence, if H is a primitive of $\frac{1}{f}$ and G is a primitive of G, then we get

$$H(u(x)) = G(x) + C \qquad \Longrightarrow \qquad u(x) = H^{-1}(G(x) + C).$$

Finally, the boundary condition uniquely identifies C.

8.1.3 Homogeneous Linear Second Order ODEs with Constant Coefficients

Second order linear differential equations are considerably more difficult to solve than first order equations. We start by considering the simplest type of second order differential equation, namely homogeneous linear ODEs with constant coefficients, namely,

$$u'' + a_1 u' + a_0 u = 0 \tag{8.13}$$

for given $a_0, a_1 \in \mathbb{R}$. Let us first consider some examples.

Example 8.14. — Solutions of $u'' = 0$ are affine functions, that is, $u(x) = Ax + B$ for constants $A, B \in \mathbb{R}$.

Solutions of $u'' - u = 0$ are functions of the form

33

$$u(x) = Ae^{x} + Be^{-x}$$

for constants $A, B \in \mathbb{R}$.

Solutions of $u'' + u = 0$ are functions of the form

$$u(x) = A\sin(x) + B\cos(x)$$

for constants $A, B \in \mathbb{R}$. Since sine and cosine can be written in terms of $e^{\pm ix}$, one can also look for solutions of the form

$$u(x) = Ce^{ix} + De^{-ix}$$

with $C, D \in \mathbb{C}$ and then re-express the solution in terms of sine and cosine (recall that we are interested in real-valued functions).

Exercise 8.15. — Check the assertions made in Example 8.14.

8.16. — The last two examples above suggest the approach of looking for solutions of (8.13) of the form $u(x) = e^{\alpha x}$ for some $\alpha \in \mathbb{C}$. Indeed, with this choice we see that

$$u''(x) + a_1 u'(x) + a_0 u(x) = \big(\alpha^2 + a_1\alpha + a_0\big)u(x) = 0.$$

In other words, α must be a zero of the so-called **characteristic polynomial**

$$p(t) = t^2 + a_1 t + a_0.$$

We distinguish three cases, depending on whether the discriminant $\Delta = a_1^2 - 4a_0$ is positive, negative, or zero.

• Case 1: $\Delta > 0$. In this case, the characteristic polynomial $p(t)$ has two distinct real roots

$$\alpha = \frac{-a_1 + \sqrt{\Delta}}{2}, \qquad \beta = \frac{-a_1 - \sqrt{\Delta}}{2}. \tag{8.14}$$

This implies that the real-valued functions $x \mapsto e^{\alpha x}$ and $x \mapsto e^{\beta x}$ are solutions, and therefore

$$u(x) = Ae^{\alpha x} + Be^{\beta x}$$

is a solution of (8.13) for every $A, B \in \mathbb{R}$.

• Case 2: $\Delta < 0$. In this case, the characteristic polynomial $p(t)$ has two distinct complex roots

$$\alpha + i\beta = -\frac{a_1}{2} + i\frac{\sqrt{-\Delta}}{2}, \qquad \alpha - i\beta = -\frac{a_1}{2} - i\frac{\sqrt{-\Delta}}{2}. \tag{8.15}$$

This means that the two complex-valued functions $x \mapsto e^{(\alpha \pm i\beta)x}$ solve (8.13), therefore their real and imaginary parts are real-valued solutions. This gives that

33

$$u(x) = Ae^{\alpha x}\sin(\beta x) + Be^{\alpha x}\cos(\beta x)$$

is a solution of (8.13) for every $A, B \in \mathbb{R}$.

• Case 3: $\Delta = 0$. In this case, the characteristic polynomial $p(t)$ has only one real zero

$$\alpha = -\frac{a_1}{2}, \tag{8.16}$$

thus $x \mapsto e^{\alpha x}$ is a solution of (8.13). To find another solution we recall the example $u'' = 0$. In this case two solutions are given by 1 and x, and these two solutions can be written as $e^{\gamma x}$ and $xe^{\gamma x}$ with $\gamma = 0$.

Motivated by this observation, one could wonder whether $x \mapsto xe^{\alpha x}$ is a solution. This is indeed the case:

$$\big(xe^{\alpha x}\big)'' + a_1\big(xe^{\alpha x}\big)' + a_0 xe^{\alpha x} = \big(\underbrace{\alpha^2 + a_1\alpha + a_0}_{=0}\big)xe^{\alpha x} + \big(\underbrace{2\alpha + a_1}_{=0}\big)e^{\alpha x} = 0$$

where the first term vanishes because α is a root of the characteristic polynomial, while the second term vanishes because of (8.16). This shows that

$$u(x) = Ae^{\alpha x} + Bxe^{\alpha x}$$

solves (8.13) for every $A, B \in \mathbb{R}$.

It is customary, for second order ODEs, to prescribe both the value of u and the value of its derivative at some point (for instance, $u(0) = 1$ and $u'(0) = 0$). The fact that we have two constants A, B guarantees that we can choose them so as to impose these two boundary conditions.

Now that we have found solutions to (8.13), we want to prove that they are the only ones. This is the content of the next:

Proposition 8.17: Existence and Uniqueness: the Homogeneous Case

Following the terminology from Paragraph 8.16 above, consider the following solutions to the homogeneous ODE (8.13)*:*

$$
\begin{array}{llll}
\underline{\Delta > 0:} & u_1(x) = e^{\alpha x}, & u_2(x) = e^{\beta x}, & \alpha, \beta \text{ as in } (8.14), \\
\underline{\Delta < 0:} & u_1(x) = e^{\alpha x}\sin(\beta x), & u_2(x) = e^{\alpha x}\cos(\beta x), & \alpha, \beta \text{ as in } (8.15), \\
\underline{\Delta = 0:} & u_1(x) = e^{\alpha x}, & u_2(x) = x e^{\alpha x}, & \alpha \text{ as in } (8.16).
\end{array}
$$

If $u \in C^2(I)$ solves (8.13)*, then there exist $A, B \in \mathbb{R}$ such that $u = Au_1 + Bu_2$.*
In other words, the set of solutions of (8.13) *forms a two-dimensional linear subspace of $C^2(I)$.*

Proof. (Extra material) We consider the case $\Delta > 0$ (the other cases can be treated analogously). Assume for simplicity that $0 \in I$ (the general case can be treated similarly, choosing
33 a point $x_0 \in I$ and arguing in a similar fashion with x_0 in place of 0). Since

$$u_1(0) = u_2(0) = 1, \qquad u_1'(0) = \alpha > \beta = u_2'(0).$$

if we define

$$v_1(x) = \frac{\alpha u_2(x) - \beta u_1(x)}{\alpha - \beta}, \qquad v_2(x) = \frac{u_1(x) - u_2(x)}{\alpha - \beta},$$

then v_1 and v_2 are two solutions of (8.13) satisfying

$$v_1(0) = 1, \qquad v_1'(0) = 0, \qquad v_2(0) = 0, \qquad v_2'(0) = 1.$$

Now, given $u \in C^2(I)$ solution of (8.13), consider $w(x) = u(x) - u(0)v_1(x)$. This is still a solution and $w(0) = u(0) - u(0)v_1(0) = 0$. We then consider the function

$$W(x) = w(x)v_2'(x) - w'(x)v_2(x)$$

(this function is called "Wronskian"). Using that both w and v_2 solve (8.13) one can check that $W' = 0$, thus W is constant. Since $W(0) = 0$ (because $w(0) = v_2(0) = 0$), we conclude that

$$w(x)v_2'(x) - w'(x)v_2(x) = 0.$$

Now, if w is identically zero, then there is nothing to prove (since this means that $u = u(0)v_1$). Otherwise, if w is not identically zero, by continuity we can find a small interval where both

w and v_2 are non-zero, and we get

$$\frac{w'(x)}{w(x)} - \frac{v_2'(x)}{v_2(x)} = 0 \quad \Longrightarrow \quad \log' |w(x)| - \log' |v_2(x)| = 0,$$

which implies that $\log |w(x)| - \log |v_2(x)| = C$ for some $C \in \mathbb{R}$. This shows that, in each interval I where w and v_2 do not vanish, there exists a constant $C \in \mathbb{R}$ such that $|w(x)| = e^C|v_2(x)|$. By continuity this implies that as long as v_2 does not vanish, then also w does not vanish and $w(x) = av_2(x)$ for some constant $a \in \mathbb{R}$. Since in our case v_2 vanished only at 0 (as one can easily check), then we deduce that there exist two constants $b, c \in \mathbb{R}$ such that

$$w = bv_2 \quad \text{on } (-\infty, 0), \qquad w = cv_2 \quad \text{on } (0, \infty)$$

This implies that

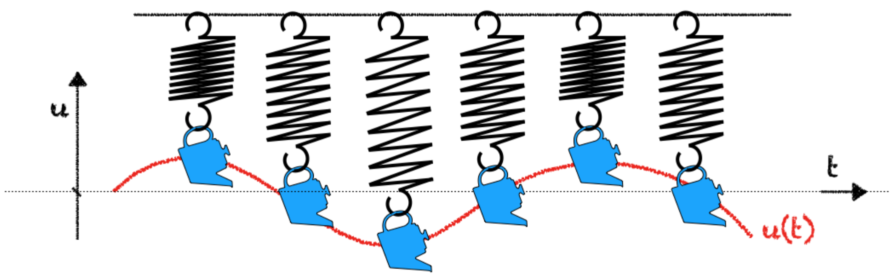
$$w_-'(0) = bv_2'(0) = b, \qquad w_+'(0) = cv_2'(0) = c.$$

So, since $w \in C^2(I)$ by assumption, the only option is $b = c$, which proves that $w(x) = bv_2(x)$ in $\mathbb{R}$. In conclusion

$$u(x) = u(0)v_1(x) + bv_2(x) = u(0)\frac{\alpha u_2(x) - \beta u_1(x)}{\alpha - \beta} + b\frac{u_1(x) - u_2(x)}{\alpha - \beta},$$

33 which implies the result. □

Example 8.18. — We attach a weight to a spring so that it is free to oscillate in the vertical direction, and want to determine the position $u(t)$ of the weight as a function of time t. We choose the coordinate system so that $u = 0$ corresponds to the state of equilibrium where the weight does not move. According to Newton's fundamental laws of motion, the second derivative u'' multiplied by the mass m of the weight is equal to the force acting on the weight.



A component of this force arises from the expansion of the spring and is oriented towards rest. According to Hooke's law, this force is given by $-ku$, where the real number $k > 0$ is called the spring constant. Furthermore, friction forces generally act movement. We assume that the corresponding force action is given by $-du'$, where $d \geq 0$ is the damping constant. The differential equation describing the motion $u(t)$ of the mass is thus $mu'' = -du' - ku$, or

$$u'' + \frac{d}{m}u' + \frac{k}{m}u = 0$$

which is, therefore, a linear homogeneous differential equation of second order. If we set $\omega = \sqrt{\frac{k}{m}}$ and $\zeta = \frac{d}{2m\omega}$, then the ODE becomes

$$u'' + 2\zeta\omega u' + \omega^2 u = 0$$

(recall Example 8.2(2)), whose characteristic polynomial is

$$p(t) = t^2 + 2\zeta\omega t + \omega^2$$

with discriminant $\Delta = 4(\zeta^2 - 1)\omega^2$.

If $\Delta < 0$ holds (equivalently $\zeta < 1$), friction is small compared to the spring strength, and we obtain solutions of the form

$$u(t) = e^{-\zeta\omega t}\big(A\sin(\gamma t) + B\cos(\gamma t)\big),$$

with $\gamma = \sqrt{1-\zeta^2}\,\omega$. The constants A and B depend on the initial position $u(0)$ and the initial velocity $u'(0)$ of the mass. In the case $\zeta = 0$, the oscillation is undamped and u is a
33 periodic function.

If friction is large compared to the strength of the spring so that $\Delta \geq 0$ (this happens when $\zeta \geq 1$), then the oscillating behavior disappears and the weight returns exponentially fast to its steady state: if $\zeta > 1$ then

$$u(t) = Ae^{-\lambda_1 t} + Be^{-\lambda_2 t}, \qquad \lambda_{1,2} = \left(\zeta \pm \sqrt{\zeta^2 - 1}\right)\omega,$$

while if $\zeta = 1$ then

$$u(t) = Ae^{-\omega t} + Bte^{-\omega t}.$$

One can note that $\zeta - \sqrt{\zeta^2 - 1} < 1$ for $\zeta > 1$, so the fastest exponential convergence is achieved when $\zeta = 1$. This behavior is desirable, for example, in a door-closing mechanism.

8.1.4 Non-Homogeneous Linear Second Order ODEs with Constant Coefficients

8.19. — After having studied the homogeneous case, we now want to solve the constant coefficients non-homogeneous linear second order ODE

$$u'' + a_1 u' + a_0 u = g \tag{8.17}$$

for given $a_0, a_1 \in \mathbb{R}$ and $g \in C^0(I)$. Following the terminology from Paragraph 8.16, consider the solutions to the homogeneous ODE:

$$\begin{array}{llll} \underline{\Delta > 0:} & u_1(x) = e^{\alpha x}, & u_2(x) = e^{\beta x}, & \alpha, \beta \text{ as in (8.14)}, \\ \underline{\Delta < 0:} & u_1(x) = e^{\alpha x}\sin(\beta x), & u_2(x) = e^{\alpha x}\cos(\beta x), & \alpha, \beta \text{ as in (8.15)}, \\ \underline{\Delta = 0:} & u_1(x) = e^{\alpha x}, & u_2(x) = x e^{\alpha x}, & \alpha \text{ as in (8.16)}. \end{array}$$

We want to solve (8.17) using the method of **variation of constants**, that is, we look for a solution u of the form $u(x) = H_1(x)u_1(x) + H_2(x)u_2(x)$ for some functions $H_1, H_2 \in C^2(I)$.

Calculating u' and u'' gives

$$u' = (H_1'u_1 + H_2'u_2) + (H_1u_1' + H_2u_2'),$$

$$u'' = (H_1'u_1 + H_2'u_2)' + (H_1'u_1' + H_2'u_2') + (H_1u_1'' + H_2u_2'').$$

Therefore

$$\begin{aligned} u'' + a_1u' + a_0u = {} & (H_1'u_1 + H_2'u_2)' + (H_1'u_1' + H_2'u_2') + (H_1u_1'' + H_2u_2'') \\ & + a_1(H_1'u_1 + H_2'u_2) + a_1(H_1u_1' + H_2u_2') + a_0(H_1u_1 + H_2u_2). \end{aligned}$$

Using that u_1 and u_2 solve the homogeneous equations, the expression above gives

33

$$u'' + a_1u' + a_0u = (H_1'u_1 + H_2'u_2)' + (H_1'u_1' + H_2'u_2') + a_1(H_1'u_1 + H_2'u_2).$$

Hence, we deduce that u is a solution of (8.17) provided the following holds:

$$H_1'(x)u_1(x) + H_2'(x)u_2(x) = 0, \qquad H_1'(x)u_1'(x) + H_2'(x)u_2'(x) = g(x).$$

Using the first equation, one can express H_2' in terms of H_1', that is,

$$H_2' = -\frac{u_1}{u_2}H_1'. \tag{8.18}$$

Then, inserting this relation into the second equation, we obtain

$$H_1'u_1' - \frac{u_1}{u_2}u_2'H_1' = g \qquad \Longrightarrow \qquad H_1' = \frac{u_2g}{u_1'u_2 - u_2'u_1}.$$

Substiting this relation back into (8.18), we conclude that

$$H_1 = \int \frac{u_2g}{u_1'u_2 - u_2'u_1}\,dx, \qquad H_2 = \int \frac{u_1g}{u_2'u_1 - u_1'u_2}\,dx.$$

In other words, if H_1 and H_2 are primitives of the functions appearing in the integral above, then $u = H_1u_1 + H_2u_2$ is a particular solution of (8.17). Finally, the set of all solutions can be found by adding to u solutions to the homogeneous equations, that is

$$u = Au_1 + Bu_2 + (H_1u_1 + H_2u_2), \qquad A, B \in \mathbb{R}.$$

We summarize this discussion in the following:

Proposition 8.20: Existence and Uniqueness: the Non-Homogeneous Case

Following the terminology from Paragraph 8.16 above, consider the following solutions to the homogeneous ODE (8.13)*:*

$$\begin{array}{llll} \underline{\Delta > 0:} & u_1(x) = e^{\alpha x}, & u_2(x) = e^{\beta x}, & \alpha, \beta \text{ as in } (8.14), \\ \underline{\Delta < 0:} & u_1(x) = e^{\alpha x}\sin(\beta x), & u_2(x) = e^{\alpha x}\cos(\beta x), & \alpha, \beta \text{ as in } (8.15), \\ \underline{\Delta = 0:} & u_1(x) = e^{\alpha x}, & u_2(x) = xe^{\alpha x}, & \alpha \text{ as in } (8.16). \end{array}$$

Also, let H_1 and H_2 denote two primitives of $\frac{u_2 g}{u_1' u_2 - u_2' u_1}$ and $\frac{u_1 g}{u_2' u_1 - u_1' u_2}$, respectively. If $u \in C^2(I)$ solves (8.17)*, then there exist $A, B \in \mathbb{R}$ such that*

$$u = Au_1 + Bu_2 + (H_1u_1 + H_2u_2).$$

In other words, the set of solutions of (8.13) *forms a two-dimensional affine subspace of $C^2(I)$.*

33

Although this method is very general, computationally it can be very involved. So, in some (very special) cases it may be easier to "guess" a particular solution to the homogeneous equation by looking at functions of the form $p(x)e^{\gamma x}$, $p(x)e^{\gamma x}\cos(\eta x)$, or $p(x)e^{\gamma x}\sin(\eta x)$, where $p(x)$ is a polynomial and $\gamma, \eta > 0$ (depending on the structure of g, one of these functions may work).

Example 8.21. — Solve the ODE

$$u''(x) + u(x) = 1, \qquad u(0) = 0, \quad u'(0) = 1.$$

In this case, all solutions to the homogeneous equation are $A\cos(x) + B\sin(x)$. Therefore, we look for a solution of the form $u(x) = H_1(x)\cos(x) + H_2(x)\sin(x)$.

This leads to the two equations

$$H_1'(x)\cos(x) + H_2'(x)\sin(x) = 0, \qquad -H_1'(x)\sin(x) + H_2'(x)\cos(x) = 1,$$

and then, solving the linear system as we did above, we get

$$H_1 = -\int \frac{\sin(x)}{\sin^2(x) + \cos^2(x)}\,dx = -\int \sin(x)\,dx$$

$$H_2 = \int \frac{\cos(x)}{\cos^2(x) + \sin^2(x)}\, dx = \int \cos(x)\, dx.$$

Hence, we can take $H_1 = \cos(x)$ and $H_2 = \sin(x)$, which leads to the particular solution $u(x) = \cos^2(x) + \sin^2(x) = 1$ (you see that, in this case, one may have tried to guess it!). So, the general solution is given by

$$u(x) = 1 + A\cos(x) + B\sin(x).$$

Imposing the boundary conditions $u(0) = 0$ and $u'(0) = 1$, we conclude that

$$u(x) = 1 - \cos(x) + \sin(x).$$

Example 8.22. — Solve the ODE

$$u''(x) + u(x) = \sin(x), \qquad u(0) = 0, \quad u'(0) = 1.$$

In this case, the method of constants gives

$$H_1 = -\int \frac{\sin(x)\sin(x)}{\sin^2(x) + \cos^2(x)}\, dx = -\int \sin(x)^2\, dx$$

33

$$H_2 = \int \frac{\cos(x)\sin(x)}{\cos^2(x) + \sin^2(x)}\, dx = \int \cos(x)\sin(x)\, dx.$$

In this case we can take $H_1 = \frac{1}{2}\big(\cos(x)\sin(x) - x\big)$ and $H_2 = -\frac{1}{2}\cos^2(x)$, which leads to the particular solution $u(x) = \frac{1}{2}\big(\cos^2(x)\sin(x) - x\cos(x) - \cos^2(x)\sin(x)\big) = -\frac{1}{2}x\cos(x)$. So, the general solution is given by

$$u(x) = -\frac{x\cos(x)}{2} + A\cos(x) + B\sin(x).$$

Imposing the boundary conditions $u(0) = 0$ and $u'(0) = 1$, we conclude that

$$u(x) = -\frac{1}{2}x\cos(x) + \frac{3}{2}\sin(x).$$

Remark 8.23. — In the solution of Example 8.22, one may note the presence of x in front of $\cos(x)$. This is due to the fact that $\sin(x)$ and $\cos(x)$ are solutions to the homogeneous equation, so the solution to the non-homogeneous problem cannot be just a linear combination of them. As a general strategy, in such situations, a special solution to the homogeneous equation is sought by multiplying the solutions of the homogeneous equation by x.

Exercise 8.24. — Solve the following ODEs:

1. $u''(x) + u(x) = \sin(2x), \qquad u(0) = 0, \quad u'(0) = 1.$
   Hint: Look for a special solution of the form $a\sin(2x) + b\cos(2x)$.

2. $u''(x) + 4u(x) = \cos(2x), \qquad u(0) = 1, \quad u'(0) = 0.$
   Hint: Look for a special solution of the form $ax\cos(2x) + bx\sin(2x)$.

33 3. $u''(x) + u'(x) - 2u(x) = x^2, \qquad u(0) = 2, \quad u'(0) = 1.$
   Hint: Look for a special solution of the form $ax^2 + bx + c$.

4. $u''(x) + 2u'(x) - 3u(x) = \cos(x) + x, \qquad u(0) = 1, \quad u'(0) = 1.$
   Hint: Look for a special solution of the form $a\sin(x) + b\cos(x) + cx + d$.

8.2 Existence and Uniqueness for ODEs

8.2.1 Existence and Uniqueness for First Order ODEs

Our goal now is to present the general theory of first-order ODEs for real-valued functions on the real line. In general, a first order ODE would be an equation of the form $G(x, u(x), u'(x)) = 0$. However, we shall assume that we can "isolate" u' so to express it as a function of x and u.

Definition 8.25: First Order ODEs

A first-order ordinary differential equation is an equation involving a function and its first derivative. The general form is

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

where $f : \mathbb{R}^2 \to \mathbb{R}$ is a given function.

The Cauchy-Lipschitz Theorem, also known as the Picard-Lindelöf Theorem, is a fundamental result in the theory of ODEs. It ensures the existence and uniqueness of solutions under suitable conditions on f.

34 In the next theorem we need to assume that f is continuous as a function of the two variables x and y. This means that, for any point (x_0, y_0) in the domain and for any $\varepsilon > 0$, there exists $\delta > 0$ such that

$$|x - x_0| < \delta \quad \text{and} \quad |y - y_0| < \delta \quad \Longrightarrow \quad |f(x,y) - f(x_0, y_0)| < \varepsilon.$$

Theorem 8.26: Cauchy-Lipschitz: Global Version

Let $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ satisfy the following conditions:

1. *f is continuous in $\mathbb{R} \times \mathbb{R}$;*
2. *f is Lipschitz with respect to the second variable; that is, there exists a constant $L > 0$ such that*

$$|f(x, y_1) - f(x, y_2)| \le L|y_1 - y_2| \qquad \forall\, x \in \mathbb{R},\ y_1, y_2 \in \mathbb{R}.$$

Then, for any point $(x_0, y_0) \in \mathbb{R} \times \mathbb{R}$ there exists a unique C^1 function $u : \mathbb{R} \to \mathbb{R}$ such that

$$\begin{cases} u'(x) = f(x, u(x)) & \text{for all } x \in \mathbb{R}, \\ u(x_0) = y_0. \end{cases} \tag{8.19}$$

As we shall see in Section 8.2.2 below, the proof is based on the method of successive approximations, also known as Picard iterations. It involves constructing a sequence of continuous functions that converge to the solution of the differential equation. Before diving into the proof of this important theorem, we first discuss some examples and generalizations.

Theorem 8.26 guarantees that solutions to the first order ODE (8.19) are unique when f is Lipschitz in the second variable. This assumption is crucial, as the next example shows.

EXAMPLE 8.27. — Consider the ODE

$$u'(x) = |u(x)|^\alpha, \qquad u(0) = 0, \tag{8.20}$$

with $\alpha \in (0, 1]$.

• For $\alpha = 1$ the function $f(y) = |y|$ is Lipschitz, since

$$|f(y_1) - f(y_2)| = ||y_1| - |y_2|| \leq |y_1 - y_2|.$$

Hence, Theorem 8.26 guarantees that the solution is unique. Since $u = 0$ is a solution, this is the unique solution.

• For $\alpha < 1$ the function $f(y) = |y|^\alpha$ is not Lipschitz. Indeed, choosing $y_2 = 0$ in the definition
34 of Lipschitz continuity, if this function were Lipschitz there would exist a constant $L > 0$ such that

$$|f(y_1) - f(0)| = |y_1|^\alpha \leq L|y_1| \qquad \forall y_1 \in \mathbb{R}.$$

This would imply

$$1 \leq L|y_1|^{1-\alpha} \qquad \forall y_1 \in \mathbb{R} \setminus \{0\},$$

but this is false since $\lim_{y_1 \to 0} |y_1|^{1-\alpha} = 0$ (recall that $\alpha < 1$).

Note now that, also in this case, the function $u = 0$ is a solution. We now try to use separation of variables (recall Section 8.1.2) to find a second solution that is not zero, say with $u(x) > 0$ somewhere:

$$u'(x) = u(x)^\alpha \quad \Longrightarrow \quad \frac{u'(x)}{u(x)^\alpha} = 1,$$

so, by integration, we obtain

$$\int \frac{du}{u^\alpha} = \int dx + C \quad \Longrightarrow \quad \frac{u(x)^{1-\alpha}}{1-\alpha} = x + C.$$

Choosing $C = 0$ (which is compatible with $u(0) = 0$) we get

$$u(x) = \big((1-\alpha)x\big)^{\frac{1}{1-\alpha}}$$

which is positive for $x > 0$. Hence, the function

$$u(x) = \begin{cases} 0 & \text{for } x \leq 0, \\ \big((1-\alpha)x\big)^{\frac{1}{1-\alpha}} & \text{for } x > 0, \end{cases}$$

is a second solution of (8.20). Actually, given $x_0 \geq 0$, all the functions

$$u_{x_0}(x) = \begin{cases} 0 & \text{for } x \leq x_0, \\ \big((1-\alpha)(x-x_0)\big)^{\frac{1}{1-\alpha}} & \text{for } x > x_0, \end{cases}$$

solve (8.20), so there are infinitely many solutions.

Motivated by the previous example, one may wonder if the solution of (8.20) is unique for $\alpha > 1$. We begin by this observation, that we state as an exercise.

Exercise 8.28. — Let $\alpha > 1$. Prove that the function $f : \mathbb{R} \to \mathbb{R}$, $y \mapsto |y|^\alpha$, is locally Lipschitz (i.e., it is Lipschitz in every compact interval $[a,b]$), but it is not Lipschitz on the whole $\mathbb{R}$.
Hint: To prove local Lipschitz continuity, use that $f(y_1) - f(y_2) = \int_{y_1}^{y_2} f'(x)\,dx$.

By the previous exercise, we see that Theorem 8.26 does not apply to (8.20) when $\alpha > 1$.
34 Still, since this function is locally Lipschitz, one may hope that some existence and uniqueness theorem still holds.

This is indeed the case, as implied by the local version of Cauchy-Lipschitz Theorem stated below. As we shall discuss later, since now the function f is only assumed to be locally Lipschitz, in general we cannot find a solution u defined in whole $\mathbb{R}$.

Theorem 8.29: Cauchy-Lipschitz: Local Version

Let $I \subset \mathbb{R}$ be an interval, and let $f : I \times \mathbb{R} \to \mathbb{R}$ satisfy the following conditions:

1. *f is continuous in $I \times \mathbb{R}$;*

2. *f is locally Lipschitz with respect to the second variable; that is, for every compact intervals $[a,b] \subset I$ and $[c,d] \subset \mathbb{R}$ there exists a constant $L > 0$ such that*

$$|f(x,y_1) - f(x,y_2)| \leq L|y_1 - y_2| \qquad \forall\, x \in [a,b],\ y_1, y_2 \in [c,d].$$

Then, for any point $(x_0, y_0) \in I \times \mathbb{R}$ there exist an interval $I' \subset I$ containing x_0, and a unique C^1 function $u : I' \to \mathbb{R}$, such that

$$\begin{cases} u'(x) = f(x, u(x)) & \text{for all } x \in I', \\ u(x_0) = y_0. \end{cases} \tag{8.21}$$

In other words, under a local Lipschitz assumption, one can only guarantee the existence and uniqueness of a solution for some interval around x_0. Also, as long as the solution $u(x)$ remains bounded within I', then one can continue applying Theorem 8.29 to extend the interval I' as much as possible.

To better understand why solutions are defined only in some interval $I' \subset I$, we consider the following example.

Example 8.30. — Consider the ODE

$$u'(x) = u(x)^2, \qquad u(0) = 1.$$

The function $f(y) = y^2$ is locally Lipschitz, so Theorem 8.29 applies.

To find the solution, we use separation of variables. More precisely, since $u(0) = 1 > 0$, u will be positive in a neighborhood of 0 and we get

$$u'(x) = u(x)^2 \quad \Longrightarrow \quad \frac{u'(x)}{u(x)^2} = 1,$$

so, by integration, we get

34

$$\int \frac{du}{u^2} = \int dx + C \quad \Longrightarrow \quad -\frac{1}{u(x)} = x + C.$$

Choosing $x = 0$ this implies $C = -1$ and we get

$$u(x) = \frac{1}{1-x}.$$

Note that this function solves the ODE in $(-\infty, 1)$, but $\lim_{x\to 1^-} u(x) = \infty$, so we cannot extend this solution beyond $x = 1$.

Exercise 8.31. — Consider the ODE

$$u'(x) = u(x)^\alpha, \qquad u(0) = 1,$$

with $\alpha > 1$. Show that Theorem 8.29 applies and find the unique solution.

Although Theorem 8.29 guarantees that most nonlinear ODEs have a unique solution (at least locally in x), nonlinear ODEs are very difficult to solve and there are no general techniques to tackle such problems, neither in practice nor in theory. Therefore, in applications, one often resorts to numerical methods.

8.2.2 Proof of Theorem 8.26 (Extra material)

REMARK 8.32. — In the proof, we shall use the following fact: If $v : \mathbb{R} \to \mathbb{R}$ is a continuous function, then also the function $s \mapsto f(s, v(s))$ is continuous. This is a consequence of the continuity of f and the fact that the composition of continuous functions is continuous. Although we did not prove this fact in this specific setting where f depends on two variables, this can be proved the same way as in Proposition 3.20.

Proof. We first prove existence in four steps. In the fifth step, we prove uniqueness.

• **Step 1: An equivalent integral equation.** We claim that $u : \mathbb{R} \to \mathbb{R}$ is a C^1 solution to (8.19) if and only if $u : \mathbb{R} \to \mathbb{R}$ is a continuous function satisfying

$$u(x) = y_0 + \int_{x_0}^{x} f(s, u(s))\, ds \qquad \forall\, x \in \mathbb{R}. \tag{8.22}$$

Indeed, if u solves the ODE, then by integration (see Corollary 7.6) we deduce the validity of (8.22).

Vice versa, if u is a continuous function satisfying (8.22), then Theorem 7.4 and Remark 8.32 imply that $u(x)$ is a primitive of the continuous function $f(x, u(x))$, therefore $u'(x) = f(x, u(x))$. In particular, u is C^1 (since its derivative is continuous). Finally, choosing $x = x_0$ in (8.22) we deduce that $u(x_0) = y_0$.

35 Therefore, to prove the existence, it suffices to construct a solution to (8.22). This will be accomplished by constructing what are known as Picard approximations, a sequence of functions that converge to a solution of (8.22).

• **Step 2: Construction of Picard approximations.** First, we define the continuous function $u_0 : \mathbb{R} \to \mathbb{R}$ as

$$u_0(x) = y_0 \qquad \forall\, x \in \mathbb{R}$$

Then, we define $u_1 : \mathbb{R} \to \mathbb{R}$ as

$$u_1(x) = y_0 + \int_{x_0}^{x} f\left(s, u_0(s)\right)\, ds.$$

Note that the integral is well defined since u_0 is continuous and therefore also the function $s \mapsto f(s, u_0(s))$ is continuous (see Remark 8.32). We also observe that u_1 is the primitive of a continuous function, so it is C^1 (and, in particular, continuous).

More generally, given $n \in \mathbb{N}$, once the continuous function $u_n : \mathbb{R} \to \mathbb{R}$ is constructed, then we define

$$u_{n+1}(x) = y_0 + \int_{x_0}^{x} f\left(s, u_n(s)\right)\, ds.$$

Again, since u_n is continuous, also $s \mapsto f(s, u_n(s))$ is continuous, and therefore u_{n+1} is C^1 (and in particular continuous).

• **Step 3: Convergence of Picard approximations.** Fix $\tau = \frac{1}{2M}$ (so that $\tau M = \frac{1}{2}$). We now prove the uniform convergence of the sequence of Picard approximations u_n on the interval $[x_0 - \tau, x_0 + \tau]$ by proving that this sequence corresponds to the partial sums of a uniformly convergent series of functions. More precisely, define $v_k = u_k - u_{k-1}$, so that

$$u_n = u_0 + \sum_{k=1}^{n} v_k.$$

We want to prove that the series

$$u_\infty(x) = y_0 + \sum_{k=1}^{\infty} v_k(x)$$

converges absolutely for every $x \in [x_0 - \tau, x_0 + \tau]$, so that the function $u_\infty : [x_0 - \tau, x_0 + \tau] \to \mathbb{R}$ is well defined, and that the sequence of function $\{u_n\}_{n=0}^{\infty}$ converges uniformly to u_∞ on $[x_0 - \tau, x_0 + \tau]$.

To prove this, we observe that

$$v_{n+1}(x) = u_{n+1}(x) - u_n(x) = \int_{x_0}^{x} \big(f(s, u_n(s)) - f(s, u_{n-1}(s))\big)\, ds \qquad \forall\, x \in \mathbb{R},$$

therefore, by the Lipschitz regularity of f in the second variable,

35

$$\begin{aligned} |v_{n+1}(x)| &\le \left| \int_{x_0}^{x} |f(s, u_n(s)) - f(s, u_{n-1}(s))|\, ds \right| \\ &\le M \left| \int_{x_0}^{x} |u_n(s) - u_{n-1}(s)|\, ds \right| = M \left| \int_{x_0}^{x} |v_n(s)|\, ds \right| \end{aligned} \tag{8.23}$$

for every $x \in I$ and $n \ge 1$. In particular, if we define $a_n = \max_{x \in [x_0 - \tau, x_0 + \tau]} |v_n(x)|$, given $x \in [x_0 - \tau, x_0 + \tau]$ it follows that $[x_0, x] \subset [x_0 - \tau, x_0 + \tau]$. Therefore

$$\left| \int_{x_0}^{x} |v_n(s)|\, ds \right| \le a_n \left| \int_{x_0}^{x} ds \right| = a_n |x - x_0| \le a_n \tau,$$

that combined with (8.23) yields

$$|v_{n+1}(x)| \le M a_n \tau = \frac{a_n}{2}.$$

Since x is an arbitrary point inside $[x_0 - \tau, x_0 + \tau]$, this proves that

$$a_{n+1} \le M a_n \tau = \frac{a_n}{2} \qquad \forall\, n \ge 1.$$

This implies (by induction) that $a_{n+1} \le 2^{-n} a_1$ which gives, for $x \in [x_0 - \tau, x_0 + \tau]$,

$$u_\infty(x) = y_0 + \sum_{k=1}^{\infty} v_k(x), \qquad |v_{k+1}(x)| \le 2^{-k} a_1.$$

By the majorant criterion, the series above converges absolutely. Also, for $x \in [x_0 - \tau, x_0 + \tau]$,

$$|u_\infty(x) - u_n(x)| \leq \sum_{k=n+1}^{\infty} |v_k(x)| \leq a_1 \sum_{k=n+1}^{\infty} 2^{k-1} = a_1 2^{n-1} \to 0 \qquad \text{as } n \to \infty,$$

which proves that the sequence of function $\{u_n\}_{n=0}^{\infty}$ converge uniformly to u_∞ on $[x_0 - \tau, x_0 + \tau]$.

• **Step 4: The limit function** $u_\infty(x)$ **solves** (8.22). We want to take the limit, as $n \to \infty$, in the formula

$$u_{n+1}(x) = y_0 + \int_{x_0}^{x} f\left(s, u_n(s)\right) ds, \qquad \forall\, x \in [x_0 - \tau, x_0 + \tau].$$

We first observe that the term on the left-hand side converges to $u_\infty(x)$ as $n \to \infty$.

To prove the convergence of the right-hand side, we estimate

$$\left|\int_{x_0}^{x} |f\left(s, u_\infty(s)\right) - f\left(s, u_n(s)\right)| \, ds\right| \leq M \left|\int_{x_0}^{x} |u_\infty(s) - u_n(s)| \, ds\right|,$$

and the last term converges to 0 as $n \to \infty$ thanks to Theorem 6.42 (since $|u_\infty - u_n| \to 0$ uniformly). This proves that

$$\int_{x_0}^{x} f\left(s, u_n(s)\right) ds \to \int_{x_0}^{x} f\left(s, u_\infty(s)\right) ds \qquad \text{as } n \to \infty,$$

35 so, we conclude that u_∞ solves (8.22) on $[x_0 - \tau, x_0 + \tau]$.

This proves the existence of a solution u_∞ on $[x_0 - \tau, x_0 + \tau]$. Now, to prove the existence of a solution on the whole interval I we need to iterate this argument. More precisely, we define $x_1 = x_0 + \tau$, $y_1 = u_\infty(x_1)$, and we consider the ODE

$$\begin{cases} u'(x) = f(x, u(x)) & \text{for all } x \in [x_1, x_1 + \tau], \\ u(x_1) = y_1. \end{cases}$$

Repeating the argument above on $[x_1, x_1 + \tau]$, we construct a solution $u_\infty : [x_1, x_1 + \tau] \to \mathbb{R}$. Then, we define $x_2 = x_1 + \tau = x_0 + 2\tau$ and $y_2 = u_\infty(x_2)$, we consider the ODE

$$\begin{cases} u'(x) = f(x, u(x)) & \text{for all } x \in [x_2, x_2 + \tau], \\ u(x_2) = y_2, \end{cases}$$

we find a solution $u_\infty : [x_2, x_2 + \tau] \to \mathbb{R}$, and so on. In this way, we construct a solution on $[x_0 - \tau, \infty)$.

Analogously, we define $x_{-1} = x_0 - \tau$ and $y_{-1} = u_\infty(x_{-1})$, we consider the ODE

$$\begin{cases} u'(x) = f(x, u(x)) & \text{for all } x \in [x_{-1} - \tau, x_1], \\ u(x_{-1}) = y_{-1}, \end{cases}$$

and we find a solution on $u_\infty : [x_{-1} - \tau, x_1] \to \mathbb{R}$. Then we define $x_{-2} = x_{-1} - \tau$ and $y_{-2} =$

$u_\infty(x_{-2})$, and we continue analogously. This allows us to construct a function $u_\infty : \mathbb{R} \to \mathbb{R}$ that solves (8.22) on the whole $\mathbb{R}$.

• **Step 5: Uniqueness**. Let $u_1, u_2 : \mathbb{R} \to \mathbb{R}$ be two solutions of (8.19), and therefore of (8.22). Then, similarly to Step 2, we note that

$$|u_1(x) - u_2(x)| = \left| \int_{x_0}^{x} |f(s, u_1(s)) - f(s, u_2(s))| \, ds \right| \leq M \left| \int_{x_0}^{x} |u_1(s) - u_2(s)| \, ds \right|.$$

Hence, if we define $a = \max_{x \in [x_0 - \tau, x_0 + \tau]} |u_1(x) - u_2(x)|$, then the inequality above yields

$$|u_1(x) - u_2(x)| \leq Ma \left| \int_{x_0}^{x} ds \right| = Ma|x - x_0| \leq M\tau a \qquad \forall\, x \in [x_0 - \tau, x_0 + \tau],$$

therefore

$$a \leq M\tau a = \frac{a}{2}.$$

This implies that $a = 0$, that is, $u_1 = u_2$ on $[x_0 - \tau, x_0 + \tau]$.

We can now repeat this argument in subsequent intervals as we did in Step 4. More precisely, we first repeat the argument in $[x_0 + \tau, x_0 + 2\tau]$ to show that $u_1 = u_2$ there, then in $[x_0 + 2\tau, x_0 + 3\tau]$, and so on. In this way, we deduce that $u_1 = u_2$ on $[x_0 - \tau, \infty)$. Then, we repeat the argument in $[x_0 - 2\tau, x_0 - \tau]$, then in $[x_0 - 3\tau, x_0 - 2\tau]$, etc., until we conclude that $u_1 = u_2$ on the whole $\mathbb{R}$. □

35 8.2.3 Higher Order ODEs (Extra material)

Suppose we are given an n-th order ODE $G\big(x, u(x), u'(x), u''(x), \ldots, u^{(n)}(x)\big) = 0$, where it is possible to isolate the highest order derivative $u^{(n)}(x)$ and express it as a function of x and lower order derivatives of u. Specifically, we assume that

$$u^{(n)}(x) = f\big(x, u(x), u'(x), \ldots, u^{(n-1)}(x)\big), \tag{8.24}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is a given function.

Our goal is to demonstrate that this ODE can be reformulated as a system of first-order ODEs. To achieve this, we introduce variables $U_1, U_2, \ldots, U_n$ to denote the derivatives of u up to the $(n-1)$-th order. These variables are defined as follows:

$$\begin{cases} U_1 = u, \\ U_2 = u' = U_1', \\ U_3 = u'' = U_2', \\ \vdots \\ U_n = u^{(n-1)} = U_{n-1}'. \end{cases}$$

Consequently, Equation (8.24) is transformed into

$$U_n'(x) = u^{(n)}(x) = f\big(x, u(x), u'(x), \ldots, u^{(n-1)}(x)\big) = f\big(x, U_1(x), U_2(x), \ldots, U_n(x)\big),$$

thus converting the original n-th order ODE into a system of first-order ODEs:

$$\begin{cases} U_1' &= U_2, \\ U_2' &= U_3, \\ &\vdots \\ U_{n-1}' &= U_n, \\ U_n' &= f\big(x, U_1(x), U_2(x), \ldots, U_n(x)\big). \end{cases}$$

The Cauchy-Lipschitz Theorem can be extended to systems of first-order ODEs, ensuring existence and uniqueness of solutions whenever f is locally Lipschitz continuous in each of the last $n-1$ variables. More precisely, once one prescribes the values of

$$U_1(x_0) = u(x_0), \quad U_2(x_0) = u'(x_0), \quad \ldots, \quad U_n(x_0) = u^{(n-1)}(x_0)$$

for some $x_0 \in I$, then there exists a unique solution (local or global, depending on the assumptions) satisfying these boundary conditions. As a consequence of this result, one can for instance recover (and generalize) Proposition 8.17

Proposition 8.33: Existence and Uniqueness for Linear 2nd Order ODEs

Given $a_0, a_1 : I \to \mathbb{R}$ continuous and bounded, consider the linear homogeneous second order ODE

35

$$u''(x) + a_1(x)u'(x) + a_0(x)u(x) = 0 \qquad \forall\, x \in I. \tag{8.25}$$

Then, the set of solutions to this equation forms a two-dimensional linear subspace of $C^2(I)$.

Proof. Fix $x_0 \in I$. According to the previous discussion, if we define $U_1 = u$ and $U_2 = u'$, then (8.25) can be rewritten as

$$\begin{cases} U_1' &= U_2, \\ U_2' &= -a_1(x)U_2 - a_0(x)U_1, \end{cases}$$

which has a unique solution once we prescribe the values of $U_1(x_0)$ and $U_2(x_0)$. Equivalently, (8.25) has a unique solution once we prescribe the values of $u(x_0)$ and $u'(x_0)$.

Now, let u_1 denote the unique solution satisfying $u_1(x_0) = 1$ and $u_1'(x_0) = 0$, and let u_2 denote the unique solution satisfying $u_2(x_0) = 0$ and $u_2'(x_0) = 1$. Then, the function $u = Au_1 + Bu_2$ is a solution of (8.25) for every $A, B \in \mathbb{R}$.

Vice versa, if $u \in C^2(I)$ solves (8.25) and we set $A = u(x_0)$ and $B = u'(x_0)$, then $v = u - Au_1 - Bu_2$ is a solution of (8.25) satisfying $v(x_0) = v'(x_0) = 0$. By uniqueness v must be identically zero (since the zero function is a solution), therefore $u = Au_1 + Bu_2$ as desired. □

The higher-dimensional version of Cauchy-Lipschitz will be explored in Analysis 2. Understanding the proof of Theorem 8.26 (in the simpler context of a single first-order ODE under a global Lipschitz condition) will provide a solid foundation for studying the generalized form of the Cauchy-Lipschitz Theorem for systems.

Index

Bibliography

[ACa2003] N. A'Campo, *A natural construction for the real numbers* arXiv preprint 0301015, (2003)

[Apo1983] T. Apostol, *A proof that Euler missed: Evaluating* $\zeta(2)$ *the easy way* The Mathematical Intelligencer **5** no.3, p. 59–60 (1983)

[Aig2014] M. Aigner and G. M. Ziegler, *Das BUCH der Beweise* Springer, (2014)

[Amm2006] H. Amann und J. Escher, *Analysis I*, 3. Auflage, Grundstudium Mathematik, Birkhäuser Basel, (2006)

[Bla2003] C. Blatter, *Analysis I* ETH Skript, https://people.math.ethz.ch/ blatter/dlp.html (2003)

[Bol1817] B. Bolzano, *Rein analytischer Beweis des Lehrsatzes, daß zwischen je zwei Werthen, die ein entgegengesetztes Resultat gewähren, wenigstens eine reelle Wurzel der Gleichung liege*, Haase Verl. Prag (1817)

[Boo1847] G. Boole, *The mathematical analysis of logic* Philosophical library, (1847)

[Can1895] G. Cantor, *Beiträge zur Begründung der transfiniten Mengenlehre* Mathematische Annalen **46** no.4, 481–512 (1895)

[Cau1821] A.L. Cauchy, *Cours d'analyse de l'école royale polytechnique* L'Imprimerie Royale, Debure frères, Libraires du Roi et de la Bibliothèque du Roi. Paris, (1821)

[Ded1872] R. Dedekind, *Stetigkeit und irrationale Zahlen* Friedrich Vieweg und Sohn, Braunschweig (1872)

[Die1990] J. Dieudonné, *Elements d'analyse* Editions Jacques Gabay (1990)

[Hat02] A. Hatcher, *Algebraic Topology* Cambridge University Press (2002)

[Hil1893] D. Hilbert, *Über die Transzendenz der Zahlen* e *und* π Mathematische Annalen **43**, 216-219 (1893)

[Hos1715] G.F.A. Marquis de l'Hôpital, *Analyse des Infiniment Petits pour l'Intelligence des Lignes Courbes* 2nde Edition, F. Montalant, Paris (1715)

[Lin1894] E. Lindelöf, *Sur l'application des méthodes d'approximations successives à l'étude des intégrales réelles des équations différentielles ordinaires* Journal de mathématiques pures et appliquées **10** no.4, 117–128 (1894)

[Rus1903] B. Russell, *The principles of mathematics* WW Norton & Company, (1903)

[Rot88] J. J. Rotman, *An introduction to Algebraic Topology* Graduate Texts in Mathematics **119** Springer 1988

[Smu1978] R. Smullyan, *What is the name of this book?* Prentice-Hall, (1978)

[Zag1990] D. Zagier, *A one-sentence proof that every prime* $p \equiv 1 \mod 4$ *is a sum of two squares.* Amer. Math. Monthly **97**, no.2, p. 144 (1990)