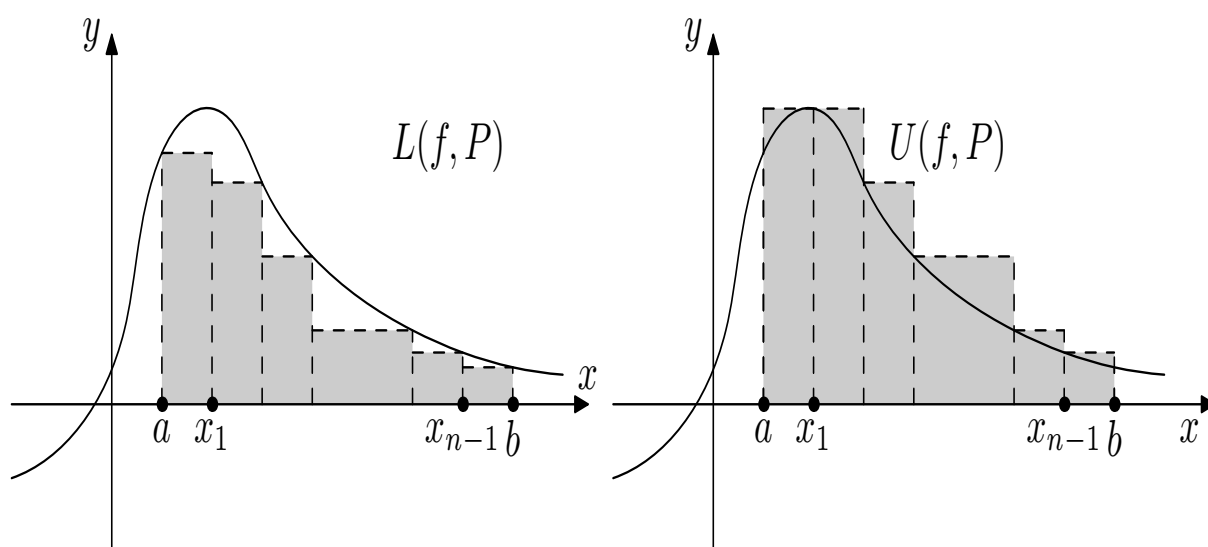


The University of Sussex – Department of Mathematics

G5095 – Further Analysis

Lecture Notes
Autumn Term 2010

Kerstin Hesse



Contents

Introduction	iii
1 Power Series, Taylor Series, and Taylor's Formula	1
1.1 Power Series	2
1.2 Taylor series	14
1.3 Taylor's Formula	17
2 Introduction of the Riemann Integral	27
2.1 Lower and Upper Sum of a Function With Respect to a Partition . .	27
2.2 Lower and Upper Riemann Integral	37
2.3 Proof of Lemma 2.16	40
3 Darboux's Theorem, Criteria for Integrability, Properties of the Integral	43
3.1 Darboux's Theorem and its Applications	43
3.2 Continuous and Monotone Functions are Riemann Integrable	58
3.3 Properties of the Integral	62
4 Techniques and Results of Integral Calculus	81
4.1 Locally Riemann Integrable Functions	82
4.2 The Primitive of a Function	84
4.3 The Indefinite Integral	86
4.4 Fundamental Theorem of Calculus	91
4.5 Integration by Parts	93
4.6 Integration by Substitution	100
4.7 Integral Test for the Convergence of a Series	105

5	Uniform Convergence	113
5.1	Uniform Convergence of Sequences of Functions	113
5.2	Results on Uniform Convergence	123
5.3	Interchange of Limit and Integral	129
5.4	Interchange of Limit and Differentiation	133
5.5	Uniform Convergence of Series of Functions and Weierstrass M -Test .	135
5.6	The Convergence of Power Series	139
6	Metric Spaces and Normed Linear Spaces	151
6.1	Metric Spaces and Normed Linear Spaces	153
6.1.1	Definitions and Basic Examples	154
6.1.2	Norms on \mathbb{R}^n	157
6.1.3	Spaces of Functions With Various Norms	162
6.1.4	Inner Product Spaces	169
6.2	Sequences in Metric Spaces and Normed Linear Spaces	176
6.2.1	Convergent Sequences and Cauchy Sequences	176
6.2.2	Complete Metric Spaces and Complete Normed Linear Spaces	181
6.2.3	Bounded Sets and the Bolzano-Weierstrass Theorem for \mathbb{R}^n . .	184
6.3	Open and Closed Subsets in Metric and Normed Linear Spaces	191
6.3.1	Interior Points and Open and Closed Sets	191
6.3.2	Accumulation Points and Characterizations of Closed Sets . .	200
A	Appendix: Handout ‘Derivatives and Integrals’	209

Introduction

The aim of this **introduction** and also of the very first lecture in this course is to give an **idea and overview what the course is about**: naturally you will not yet understand everything mentioned in this introduction, since you will encounter some of the topics for the first time in this course. Likewise the introduction should not be seen as a comprehensive summary of the course, because some important concepts will only be mentioned by name but not explained.

Let us start by taking a look at the **table of contents**: The course has **six chapters**, and the **diagram in Figure 1 below shows how the various chapters are interrelated**. You may notice that the last chapter is the longest one. In fact, Chapter 6 is rather important and has the purpose to bind together several individual concepts encountered in this course, but also in your first year courses, and give them a common framework. You will also find that the material in Chapter 6 is rather useful for analysis and calculus based courses that you might take in your third year!

I will now give a **brief overview** of the topics covered in this course.

We start the course in **Chapter 1** with a revision of material that you should have seen in some form in your first year classes: a **power series** centred at x_0 is a series of the form

$$\sum_{n=0}^{\infty} c_n (x - x_0)^n = c_0 + c_1 (x - x_0) + c_2 (x - x_0)^2 + \dots + c_n (x - x_0)^n + \dots,$$

and we want to know for which values of x the power series converges. For this purpose we determine the **radius of convergence** ρ : we then know that the power series converges for all x satisfying $|x - x_0| < \rho$ and diverges for all x with $|x - x_0| > \rho$. A special type of power series is the **Taylor series** centred at x_0 of an infinitely often differentiable function f , given by

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n.$$

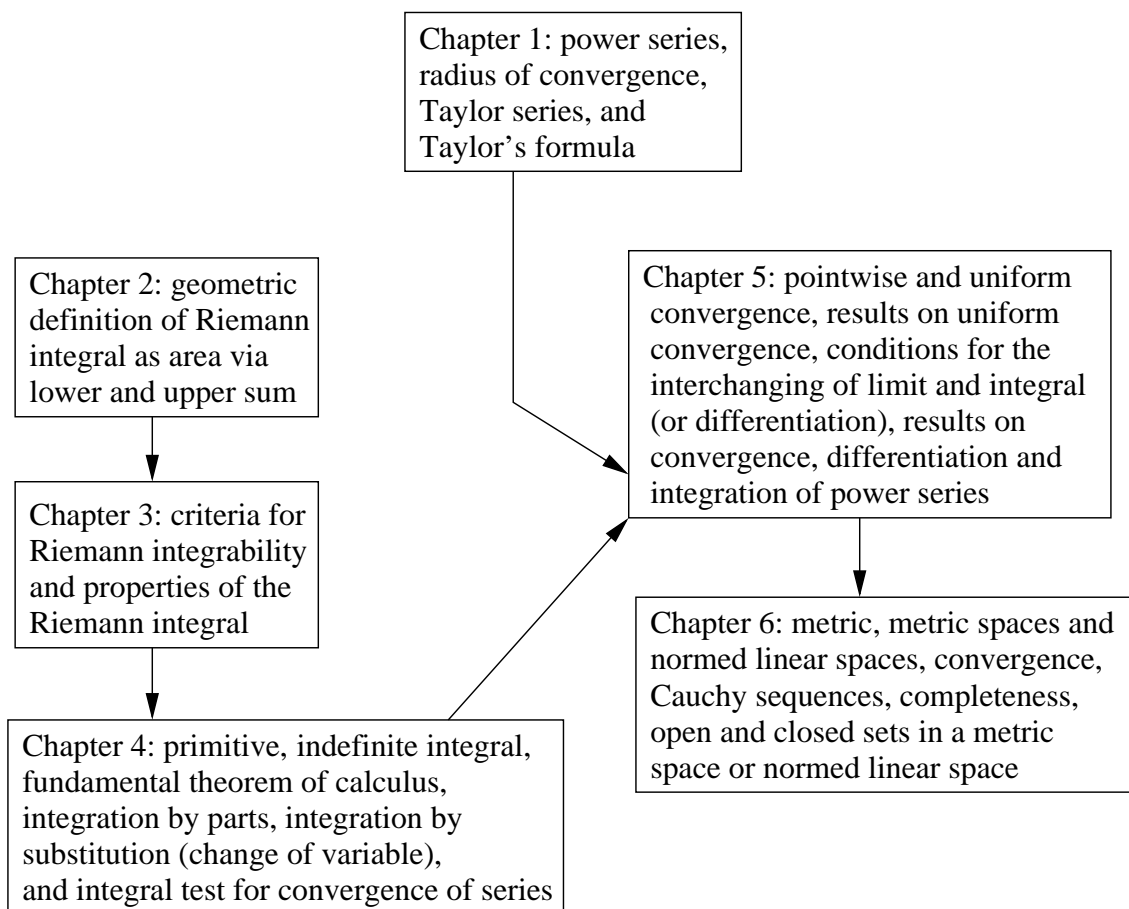


Figure 1: Interrelation of the various topics covered in this course.

We also encounter **Taylor's formula** which gives a polynomial approximation of a $(k+1)$ -times continuously differentiable function f by the terms of the Taylor series with $n \leq k$.

The **Chapters 2, 3, and 4** are all devoted to the **Riemann integral**: The Riemann integral is the usual integral that you will have encountered in school; the **Riemann integral** of a (continuous) function f over an interval $[a, b]$ is defined as the **area under the graph** of f from $x = a$ to $x = b$, as illustrated in Figure 2 below.

In order to compute the area under the graph of f from $x = a$ to $x = b$, we will start by filling out (or covering) the area under the graph by rectangles as illustrated in the left picture (and right picture, respectively) in Figure 3. The sum of the areas of the rectangles is clearly an approximation of the area under the graph, and we call the sum of the areas of the rectangles in the left picture a **lower sum** and the sum of the areas of the rectangles in the right picture an **upper sum**, respectively. The

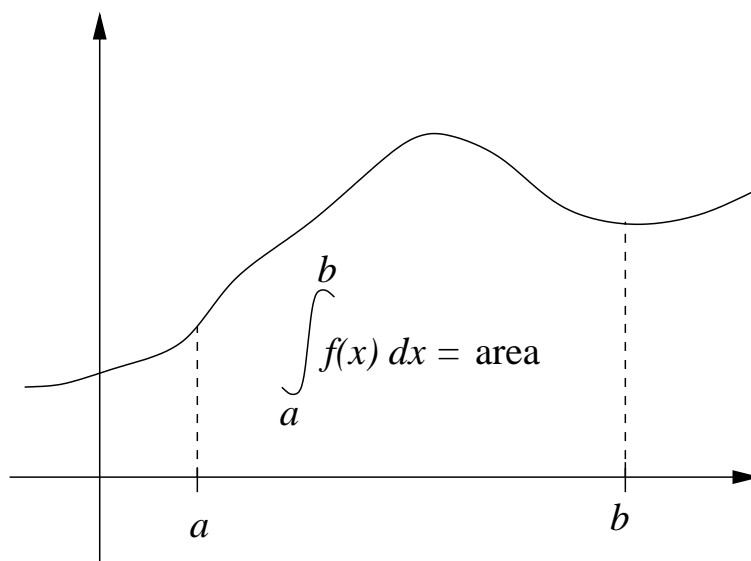


Figure 2: Riemann integral as area under the graph.

idea is then to shrink the base of each of these rectangles further and further and so to obtain in the limit the area under the graph. This is discussed in **Chapter 2**.

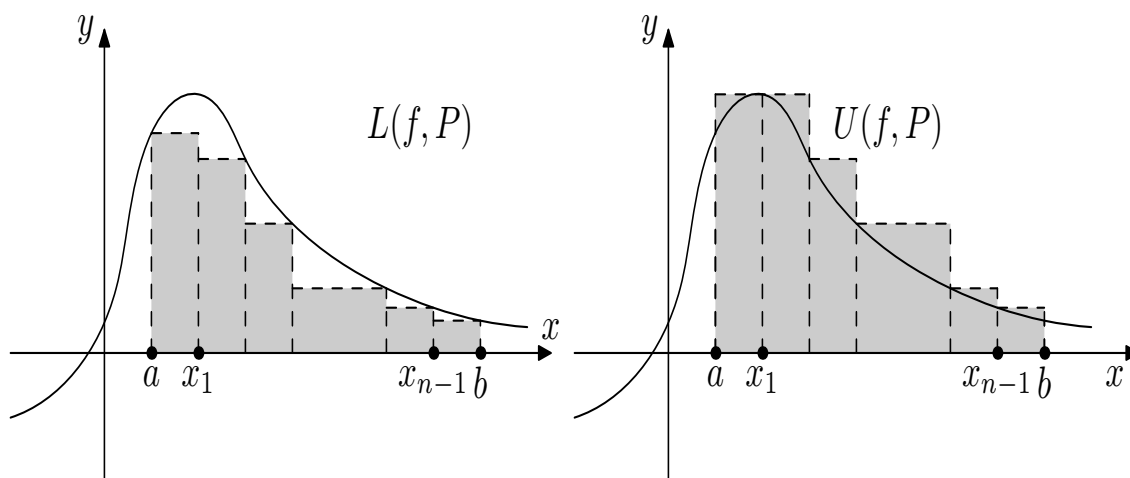


Figure 3: Approximation of the integral by the lower sum and the upper sum.

In **Chapter 3**, we derive some **criteria** to determine whether a function is Riemann integrable, and we prove important **properties of the Riemann integral**.

In **Chapter 4**, we introduce the notion of a primitive: a **primitive** (or **antiderivative**) of a given function f is a differentiable function F such that

$$F' = f.$$

For example, the function $F(x) = x^3/3$ is a primitive of $f(x) = x^2$, since $(x^3/3)' = (3x^2/3) = x^2$, and $G(x) = \sin x$ is a primitive of $g(x) = \cos x$ because $(\sin x)' = \cos x$. With the help of the so-called **indefinite integral** we can describe primitives, and we use the indefinite integral to prove the fundamental theorem of calculus which links differentiation and integration. The **fundamental theorem of calculus** says the following: *Let f be a continuous function and let F be a primitive of f , then*

$$\int_a^b f(x) dx = F(b) - F(a).$$

If we replace f by F' (since F is a primitive of f), then the **link between differentiation and integration** becomes even more obvious:

$$\int_a^b F'(x) dx = F(b) - F(a);$$

and we see why differentiation and integration are ‘inverse’ operations to each other.

In **Chapter 4**, we also prove techniques for evaluating integrals, namely, **integration by parts** and **integration by substitution** (or **change of variable**). We also learn the so-called **integral test** for checking the convergence of a series of real numbers.

In **Chapter 5**, we discuss the **convergence of sequences of functions**. Given a sequence $\{f_n\}$ of functions $f_n : [a, b] \rightarrow \mathbb{R}$, we can ask if there is some function $f : [a, b] \rightarrow \mathbb{R}$ such that

$$\lim_{n \rightarrow \infty} f_n(x) = f(x) \quad \text{for all } x \in [a, b]. \quad (0.1)$$

If this is the case we say that the sequence $\{f_n\}$ **converges pointwise** to the function f . (The notion ‘pointwise’ is motivated by the fact that at every fixed point $x \in [a, b]$ the sequence $\{f_n(x)\} \subset \mathbb{R}$ converges to the real number $f(x)$.) If (0.1) holds and if all the functions f_n and f are Riemann integrable, can we then **interchange the limit and the integral**, that is, does

$$\lim_{n \rightarrow \infty} \int_a^b f_n(x) dx = \int_a^b \left(\lim_{n \rightarrow \infty} f_n(x) \right) dx = \int_a^b f(x) dx \quad (0.2)$$

hold? The answer is in general **no**! In order to have (0.2), the sequence $\{f_n\}$ needs to converge in a stronger sense than pointwise. This new stronger type of convergence is called **uniform convergence on the interval** $[a, b]$; it is a non-trivial concept and we will therefore not try to explain it here. Once we have introduced uniform

convergence, we will discuss the interchanging of the limit and the integral (as in (0.2)) and also under which conditions we have

$$\lim_{n \rightarrow \infty} \frac{df_n(x)}{dx} = \frac{d}{dx} \left(\lim_{n \rightarrow \infty} f_n(x) \right),$$

if the sequence $\{f_n\}$ consists of differentiable functions f_n . Finally we will apply all these results to **series of functions**, with a particular attention to **power series**.

Chapter 6 finally ties various ideas together that we have discussed in this course and that you have learnt in previous mathematics courses. We start by discussing **distance functions**, usually called **metrics**, and so-called **norms** for linear spaces. The simplest example of a **distance function** is the usual distance on the real line \mathbb{R} , given by

$$\text{distance of } x \in \mathbb{R} \text{ and } y \in \mathbb{R} = d(x, y) = |x - y|.$$

An example of a norm is the **Euclidean norm** of \mathbb{R}^3 ,

$$\|(x, y, z)\|_2 = \sqrt{x^2 + y^2 + z^2}, \quad (x, y, z) \in \mathbb{R}^3.$$

In the Euclidean space \mathbb{R}^3 with $\|\cdot\|_2$, we measure distances by

$$d((x, y, z), (x', y', z')) = \|(x, y, z) - (x', y', z')\|_2 = \sqrt{(x - x')^2 + (y - y')^2 + (z - z')^2},$$

and we see, for the example of the Euclidean norm, that a norm induces a distance function. A set with a distance function is called a **metric space**, and a linear space with a norm is called a **normed linear space**.

In a metric space we can measure **distances** with the metric (or distance function), and this allows us to discuss **convergence** and **Cauchy sequences**. (Remember the ε -definition of convergence in \mathbb{R} : it really makes only use of having a distance function.) Definitions of convergence and Cauchy sequences are now given for general metric spaces (or normed linear spaces), but we find that **all notions of convergence and Cauchy sequences that we have learnt before** (for example, convergence in \mathbb{R} and uniform convergence) **are special cases of these general definitions**. We also discuss the concept of **completeness** of a metric space (or a normed linear space): A metric space is **complete** if every Cauchy sequence converges to some element in the space. For example, the real numbers \mathbb{R} with the distance function $d(x, y) = |x - y|$ are a complete metric space, since every Cauchy sequence $\{a_n\}$ in \mathbb{R} converges to some $a \in \mathbb{R}$.

Finally we define **open (sub)sets** and **closed (sub)sets** in a metric space (or a normed linear space), and we will derive some **criteria for checking whether a**

set is closed. These concepts are entirely new material, and it would be difficult to try to explain them here. As an elementary example, consider the real line \mathbb{R} with the usual distance function $d(x, y) = |x - y|$: here open intervals (a, b) are indeed open subsets of \mathbb{R} , and closed intervals $[a, b]$ are indeed closed subsets of \mathbb{R} .

Chapter 1

Power Series, Taylor Series, and Taylor's Formula

This chapter is mainly revision of material covered in first year. In Section 1.1, we will introduce **power series** and investigate their convergence. More precisely, we will learn how to determine the **radius of convergence** ρ of a given power series

$$\sum_{n=0}^{\infty} c_n (x - x_0)^n = c_0 + c_1 (x - x_0) + c_2 (x - x_0)^2 + \dots + c_n (x - x_0)^n + \dots,$$

and we will learn that the series converges absolutely for all $x \in \mathbb{R}$ with $|x - x_0| < \rho$ and diverges for all $x \in \mathbb{R}$ with $|x - x_0| > \rho$. The **Taylor series** of an infinitely often differentiable function f centred at x_0 , given by

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n = f(x_0) + f'(x_0) (x - x_0) + \dots + \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n + \dots,$$

is a special type of power series, and Taylor series are briefly discussed in Section 1.2. We raise the question whether, at those points x with $|x - x_0| < \rho$, the Taylor series converges to the function f . The answer is in general no, and we give an example to illustrate this. In Section 1.3, we introduce **Taylor's formula** which allows us to approximate a given $(n + 1)$ -times differentiable function by a so-called Taylor polynomial (the Taylor series up to a degree n) and gives an representation of the approximation error. Taylor's formula can be used to investigate whether the Taylor series of an infinitely often differentiable function f converges for x with $|x - x_0| < \rho$ to the function f .

1.1 Power Series

In this section we define **power series centred at** x_0 , and we learn that there exists an ρ with $0 \leq \rho \leq \infty$ such that the power series converges for all x with $|x - x_0| < \rho$ and diverges for all x with $|x - x_0| > \rho$. This number ρ is called the **radius of convergence**. We will learn two useful theorems that can often be used to find the radius of convergence, namely the **ratio test** and the **root test** for power series. These are proved with the help of the ratio test and root test for series of real numbers. We discuss several examples.

We start with a motivating example.

Example 1.1 (geometric series)

We know by the ratio test that the **geometric series**

$$\sum_{n=0}^{\infty} x^n = 1 + x + x^2 + x^3 + x^4 + \dots + x^n + \dots$$

is absolutely convergent if $|x| < 1$ and is divergent when $|x| \geq 1$. If $|x| < 1$, the limit is $1/(1 - x)$. If we treat x in the above example as a variable, we obtain a function $f : \mathbb{R} \rightarrow \mathbb{R}$ given by the series

$$f(x) := \sum_{n=0}^{\infty} x^n, \quad x \in \mathbb{R},$$

and we know that its value at $|x| < 1$ is the limit $1/(1 - x)$ of the series, whereas the series diverges for $|x| \geq 1$. This is a simple example of a power series. \square

Power series are to be viewed as a **special type of sequences of functions** as we will elaborate below.

Definition 1.2 (power series)

A **power series** centred at x_0 is an expression of the form

$$\sum_{n=0}^{\infty} c_n (x - x_0)^n = c_0 + c_1 (x - x_0) + c_2 (x - x_0)^2 + \dots + c_n (x - x_0)^n + \dots, \quad (1.1)$$

where $x \in \mathbb{R}$ is the variable, $x_0 \in \mathbb{R}$ a fixed point, and $c_n \in \mathbb{R}$, $n \in \mathbb{N}_0$, are the **coefficients**. The point x_0 is called the **centre** of the power series, and we will also say that the power series (1.1) is **centred at** x_0 .

Example 1.3 (Example 1.1 continued)

In Example 1.1 above, the centre of the power series is $x_0 = 0$ and the coefficients are $c_n = 1$ for all $n \in \mathbb{N}_0$. \square

Example 1.4 (power series)

The power series

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} (x+1)^n = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} (x - (-1))^n$$

has the centre $x_0 = -1$ and the coefficients $c_n = (-1)^n/(2n+1)!$, $n \in \mathbb{N}_0$. \square

A central question is the **convergence** of a power series (1.1). More precisely, we will want to know **at which points** $x \in \mathbb{R}$ a given power series converges. For a **fixed** $x \in \mathbb{R}$, the series (1.1) is just a **series of real numbers**, and we can make use of our knowledge about series in \mathbb{R} .

As a starting point, recall the definition of **convergence of a series in \mathbb{R}** and two useful criteria for determining whether the series converges.

Definition 1.5 (convergence and absolute convergence of a series in \mathbb{R})

Let $\{a_n\}$ be a sequence in \mathbb{R} . The series

$$\sum_{n=0}^{\infty} a_n \tag{1.2}$$

is **convergent** if the sequence $\{s_m\}$ of the **partial sums**

$$s_m := \sum_{n=0}^m a_n$$

converges, that is, if for all $\varepsilon > 0$ there exists $N = N(\varepsilon) \in \mathbb{N}$, such that

$$|s_m - s_k| = \left| \sum_{n=k+1}^m a_n \right| < \varepsilon \quad \text{for all } m > k \geq N.$$

The series (1.2) **converges absolutely** if the series

$$\sum_{n=0}^{\infty} |a_n|$$

converges.

Lemma 1.6 (necessary but not sufficient condition for convergence)

If a series (1.2) of real numbers converges then we have $\lim_{n \rightarrow \infty} |a_n| = 0$. The condition $\lim_{n \rightarrow \infty} |a_n| = 0$ is necessary but **not sufficient** for the convergence of (1.2). **If** $\lim_{n \rightarrow \infty} |a_n| \neq 0$, **then the series (1.2) diverges**.

We give some examples of convergent and divergent series of real numbers.

Example 1.7 (series of real numbers)

- (a) $\sum_{n=1}^{\infty} a^n$ converges if $|a| < 1$ and diverges if $|a| \geq 1$ (geometric series).
- (b) $\sum_{n=1}^{\infty} n$ diverges, because the sequence $\{n\}$ does not have the limit zero.
- (c) $\sum_{n=1}^{\infty} \frac{1}{n^\gamma}$ converges if $\gamma > 1$ and diverges if $\gamma \leq 1$ (not obvious).

We will learn later in this course how to prove statement (c) very neatly with the so-called integral test. \square

The above definition of the convergence of a series of real numbers is in practice not very useful. Useful criteria for testing the convergence of a series are given by the **ratio test** and the **root test**.

Lemma 1.8 (ratio test for series in \mathbb{R})

Let $\{a_n\}$ be a sequence of real numbers. Then the series

$$\sum_{n=0}^{\infty} a_n \tag{1.3}$$

converges absolutely if there exists a real number α with $0 < \alpha < 1$ and an integer $N \in \mathbb{N}$ such that

$$\left| \frac{a_{n+1}}{a_n} \right| \leq \alpha < 1 \quad \text{for all } n \geq N.$$

In particular, if

$$\lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| = \beta,$$

then the series **converges absolutely** if $\beta < 1$ and **diverges** if $\beta > 1$. If $\beta = 1$, then the ratio test is inconclusive: the series could either converge or diverge.

Lemma 1.9 (root test for series in \mathbb{R})

Let $\{a_n\}$ be a sequence of real numbers. Then the series

$$\sum_{n=0}^{\infty} a_n \quad (1.4)$$

converges absolutely if there exists a real number α with $0 < \alpha < 1$ and an integer $N \in \mathbb{N}$ such that

$$\sqrt[n]{|a_n|} \leq \alpha < 1 \quad \text{for all } n \geq N.$$

In particular, if

$$\lim_{n \rightarrow \infty} \sqrt[n]{|a_n|} = \beta,$$

then the series **converges absolutely** if $\beta < 1$ and **diverges** if $\beta > 1$. If $\beta = 1$, then the root test is inconclusive: the series could either converge or diverge.

It is important not to forget the **absolute values** in the definitions above.

We note that the first condition for the convergence in Lemmas 1.8 and 1.9 is more general, since the limits $\lim_{n \rightarrow \infty} |a_{n+1}|/|a_n|$ and $\lim_{n \rightarrow \infty} \sqrt[n]{|a_n|}$ need not exist.

Let us return to **power series**. In analogy to series in \mathbb{R} , think of a power series

$$\sum_{n=0}^{\infty} c_n (x - x_0)^n, \quad x \in \mathbb{R}, \quad (1.5)$$

as a **sequence $\{s_m\}$ of partial sums**, where the partial sum s_m is given by

$$s_m(x) := \sum_{n=0}^m c_n (x - x_0)^n, \quad x \in \mathbb{R}. \quad (1.6)$$

Note that s_m is a polynomial of degree m . In order to determine for which x (1.5) converges, we need to determine for which values of x the sequence $\{s_m(x)\} \subset \mathbb{R}$ converges. For each fixed x , (1.5) is a series in \mathbb{R} , and we can apply the **ratio test** or the **root test** to determine whether the series converges for this x .

Example 1.10 (convergence of geometric series)

Consider the geometric series

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

Then for a fixed $x \in \mathbb{R}$

$$\lim_{n \rightarrow \infty} \sqrt[n]{|x^n|} = \lim_{n \rightarrow \infty} \sqrt[n]{|x|^n} = \lim_{n \rightarrow \infty} |x| = |x|,$$

and the **root test** yields that the series converges for $|x| < 1$ and diverges for $|x| > 1$. For the case $|x| = 1$ the root test gives no information. Likewise, for a fixed $x \in \mathbb{R}$

$$\lim_{n \rightarrow \infty} \left| \frac{x^{n+1}}{x^n} \right| = \lim_{n \rightarrow \infty} |x| = |x|,$$

and the **ratio test** yields that the series converges for $|x| < 1$ and diverges for $|x| > 1$. For $|x| = 1$, the ratio test gives no information. \square

Definition 1.11 (radius of convergence)

Consider a power series

$$\sum_{n=0}^{\infty} c_n (x - x_0)^n. \quad (1.7)$$

Then the quantity

$$\rho := \sup \left\{ |x - x_0| : x \in \mathbb{R} \text{ for which } \sum_{n=0}^{\infty} c_n (x - x_0)^n \text{ converges} \right\}$$

is called the **radius of convergence** of the series (1.7).

Remark 1.12 (radius of convergence)

Definition 1.11 allows $0 \leq \rho \leq \infty$. (Remember that the supremum is the least upper bound.) We have to distinguish essentially **three cases** where the radius of convergence is concerned:

- (1) $\rho = 0$ means that the series converges only at $x = x_0$ (clear from the definition of ρ).
- (2) $\rho = \infty$ implies that the series converges for all $x \in \mathbb{R}$ (not obvious).
- (3) $0 < \rho < \infty$ means that the series converges absolutely for $x \in (x_0 - \rho, x_0 + \rho)$ and diverges for $x \in \mathbb{R} \setminus [x_0 - \rho, x_0 + \rho]$ (not obvious).

The next Theorem shows that statements (2) and (3) in Remark 1.12 are true.

Theorem 1.13 (radius of convergence)

Consider the power series

$$\sum_{n=0}^{\infty} c_n (x - x_0)^n,$$

and let ρ be its radius of convergence. Then the series **converges absolutely** for all $x \in (x_0 - \rho, x_0 + \rho)$. The series **diverges** for all $x \in \mathbb{R} \setminus [x_0 - \rho, x_0 + \rho]$.

Note: The statement $x \in (x_0 - \rho, x_0 + \rho)$ is equivalent to saying that x satisfies $|x - x_0| < \rho$. The statement that $x \in \mathbb{R} \setminus [x_0 - \rho, x_0 + \rho]$ is equivalent to saying that x satisfies $|x - x_0| > \rho$.

Note: The properties of ρ in Theorem 1.13, namely that the power series converges absolutely for all $x \in (x_0 - \rho, x_0 + \rho)$ and diverges for all $x \in \mathbb{R} \setminus [x_0 - \rho, x_0 + \rho]$, determine ρ uniquely.

Proof of Theorem 1.13: Consider a fixed $x \in (x_0 - \rho, x_0 + \rho)$. From the definition of the radius of convergence, there exists some $y \in (x_0 - \rho, x_0 + \rho)$ such that $|x - x_0| < |y - x_0| < \rho$ and such that

$$\sum_{n=0}^{\infty} c_n (y - x_0)^n$$

converges. Thus $\{|c_n (y - x_0)^n|\}$ is a sequence that tends to zero as $n \rightarrow \infty$, and is therefore bounded, that is, there exists some $K > 0$ such that $|c_n (y - x_0)^n| \leq K$ for all $n \in \mathbb{N}_0$.

We will now estimate $|c_n (x - x_0)^n|$. From $|c_n (y - x_0)^n| \leq K$ for all $n \in \mathbb{N}_0$, we have

$$\begin{aligned} |c_n (x - x_0)^n| &= \left| c_n (y - x_0)^n \frac{(x - x_0)^n}{(y - x_0)^n} \right| \\ &= |c_n (y - x_0)^n| \left| \frac{x - x_0}{y - x_0} \right|^n \\ &\leq K \left| \frac{x - x_0}{y - x_0} \right|^n. \end{aligned} \tag{1.8}$$

We see that from (1.8), $|x - x_0|/|y - x_0| < 1$ (since $|x - x_0| < |y - x_0|$), and the geometric series that

$$\sum_{n=0}^{\infty} |c_n (x - x_0)^n| \leq \sum_{n=0}^{\infty} K \left| \frac{x - x_0}{y - x_0} \right|^n = K \sum_{n=0}^{\infty} \left(\underbrace{\left| \frac{x - x_0}{y - x_0} \right|}_{<1} \right)^n < \infty.$$

Thus the power series converges absolutely at x .

That the series diverges for $x \in \mathbb{R} \setminus [x_0 - \rho, x_0 + \rho]$, that is, for x with $|x - x_0| > \rho$, follows directly from the definition of the radius of convergence. \square

With the help of the **ratio test** for sequences of real numbers (see Lemma 1.8) we can derive a convenient test for finding the radius of convergence of a power series.

Theorem 1.14 (ratio test for power series)

Consider the power series

$$\sum_{n=0}^{\infty} c_n (x - x_0)^n. \quad (1.9)$$

Suppose that the limit

$$\sigma := \lim_{n \rightarrow \infty} \frac{|c_{n+1}|}{|c_n|} \quad (1.10)$$

exists, where we also allow infinity as the value of the limit. Then $\rho := 1/\sigma$ is the **radius of convergence** of the power series, and the power series (1.9) **converges absolutely** for all $x \in (x_0 - \rho, x_0 + \rho)$, and **diverges** for all $x \in \mathbb{R} \setminus [x_0 - \rho, x_0 + \rho]$.

Note that the ratio test assumes that the limit (1.10) **exists** in $[0, \infty]$, which is not always the case.

Proof of Theorem 1.14: Let us for the moment consider an arbitrary fixed x . By the ratio test (see Lemma 1.8), the series

$$\sum_{n=0}^{\infty} c_n (x - x_0)^n \quad (1.11)$$

converges absolutely if the limit

$$\lim_{n \rightarrow \infty} \frac{|c_{n+1} (x - x_0)^{n+1}|}{|c_n (x - x_0)^n|} = \lim_{n \rightarrow \infty} \frac{|c_{n+1}|}{|c_n|} |x - x_0| = \sigma |x - x_0| \quad (1.12)$$

is less than 1. Thus the series (1.11) converges absolutely if

$$\sigma |x - x_0| < 1 \quad \Leftrightarrow \quad |x - x_0| < 1/\sigma.$$

If $\sigma |x - x_0| > 1$, we know from (1.12) and the ratio test for sequences of real numbers that the series (1.11) diverges.

We found that the power series (1.9) converges absolutely for x with $|x - x_0| < 1/\sigma$ and diverges for x with $|x - x_0| > 1/\sigma$. Thus we see from Theorem 1.13 that $\rho = 1/\sigma$ is the radius of convergence. \square

Example 1.15 (radius of convergence determined with ratio test)

Find the radius of convergence ρ for each of the following power series:

$$(a) \sum_{n=0}^{\infty} n! x^n, \quad (b) \sum_{n=0}^{\infty} \frac{x^n}{n!}, \quad (c) \sum_{n=0}^{\infty} r^n x^n, \quad (d) \sum_{n=0}^{\infty} \frac{(-1)^n}{(n+1)^2} (x-1)^n.$$

Solution: Using the ratio test yields the following:

- (a) $\sigma = \lim_{n \rightarrow \infty} \left| \frac{(n+1)!}{n!} \right| = \lim_{n \rightarrow \infty} (n+1) = \infty$, thus $\rho = 1/\sigma = 0$.
- (b) $\sigma = \lim_{n \rightarrow \infty} \left| \frac{1/(n+1)!}{1/n!} \right| = \lim_{n \rightarrow \infty} \frac{n!}{(n+1)!} = \lim_{n \rightarrow \infty} \frac{1}{n+1} = 0$, thus $\rho = 1/\sigma = \infty$.
- (c) $\sigma = \lim_{n \rightarrow \infty} \left| \frac{r^{n+1}}{r^n} \right| = \lim_{n \rightarrow \infty} |r| = |r|$, thus $\rho = 1/\sigma = 1/|r|$.
- (d) $\sigma = \lim_{n \rightarrow \infty} \frac{|(-1)^{n+1}/(n+2)^2|}{|(-1)^n/(n+1)^2|} = \lim_{n \rightarrow \infty} \frac{(n+1)^2}{(n+2)^2} = 1$, thus $\rho = 1/\sigma = 1$. \square

With the help of Lemma 1.9, we can also derive a **root test** for determining the radius of convergence of a power series.

Theorem 1.16 (root test for power series)

Consider the power series

$$\sum_{n=0}^{\infty} c_n (x - x_0)^n. \quad (1.13)$$

Suppose that the limit

$$\sigma = \lim_{n \rightarrow \infty} \sqrt[n]{|c_n|}$$

exists, where we also allow infinity as the value of the limit. Then $\rho := 1/\sigma$ is the **radius of convergence** of the power series (1.13), and the power series (1.13) **converges absolutely** for all $x \in (x_0 - \rho, x_0 + \rho)$, and **diverges** for all $x \in \mathbb{R} \setminus [x_0 - \rho, x_0 + \rho]$.

Proof of Theorem 1.16: Let us consider a fixed arbitrary $x \in \mathbb{R}$. From the root test for series in \mathbb{R} (see Lemma 1.9), we know that the series

$$\sum_{n=0}^{\infty} c_n (x - x_0)^n \quad (1.14)$$

converges absolutely if

$$\lim_{n \rightarrow \infty} \sqrt[n]{|c_n (x - x_0)^n|} = \lim_{n \rightarrow \infty} \sqrt[n]{|c_n|} |x - x_0| = \sigma |x - x_0| < 1.$$

This means that the power series (1.14) converges absolutely at x if $|x - x_0| < 1/\sigma$. From the root test for series of real numbers, we know that the series (1.14) diverges if

$$\lim_{n \rightarrow \infty} \sqrt[n]{|c_n (x - x_0)^n|} = \sigma |x - x_0| > 1.$$

Hence the power series (1.13) diverges for x with $|x - x_0| > 1/\sigma$.

Since we have shown that the power series (1.13) converges absolutely for x with $|x - x_0| < 1/\sigma$ and diverges for $|x - x_0| > 1/\sigma$, we see from Theorem 1.13 that $\rho = 1/\sigma$ is the radius of convergence. \square

Note that the root test assumes that the limit $\lim_{n \rightarrow \infty} \sqrt[n]{|c_n|}$ **exists** in $[0, \infty]$ which is not always the case. The root test is not so easy to apply as the ratio test, since it is not always straight-forward to determine $\lim_{n \rightarrow \infty} \sqrt[n]{|c_n|}$.

Example 1.17 (radius of convergence determined with root test)

Determine the radius of convergence of

$$(a) \quad \sum_{n=0}^{\infty} r^n x^n, \quad (b) \quad \sum_{n=0}^{\infty} \frac{(-1)^n}{(n+1)^2} (x-1)^n$$

with the root test.

Solution:

$$(a) \quad \sigma = \lim_{n \rightarrow \infty} \sqrt[n]{|r^n|} = \lim_{n \rightarrow \infty} \sqrt[n]{|r|^n} = \lim_{n \rightarrow \infty} |r| = |r|, \quad \text{thus} \quad \rho = 1/\sigma = 1/|r|.$$

$$(b) \quad \sigma = \lim_{n \rightarrow \infty} \sqrt[n]{|(-1)^n/(n+1)^2|} = \lim_{n \rightarrow \infty} (n+1)^{-2/n} = \lim_{n \rightarrow \infty} e^{-2 \ln(n+1)/n} = 1$$

$$(\text{since } \lim_{n \rightarrow \infty} \ln(n+1)/n = 0), \quad \text{thus} \quad \rho = 1/\sigma = 1. \quad \square$$

If all coefficients c_n with even indices or all coefficients c_n with odd indices vanish, then we cannot write down the ratio $|c_{n+1}/c_n|$ since it is undefined for even n and odd n respectively. Thus the ratio test for power series cannot be applied! However, we can just use the ratio test for sequences in \mathbb{R} (that is, Lemma 1.8) to find the radius of convergence as shown in the next two examples.

Example 1.18 (ratio test for power series cannot be applied)

In the power series

$$\sum_{n=0}^{\infty} x^{2n} = 1 + x^2 + x^4 + x^6 + \dots$$

all coefficients c_{2k+1} with odd indices $n = 2k+1$, $k \in \mathbb{N}_0$, are zero. Thus the quotient $|c_{n+1}/c_n|$ is undefined if n is odd. Hence we cannot apply Theorem 1.14. However, we may apply the ratio test for sequences in \mathbb{R} (see Lemma 1.8) to the series directly for each (fixed) x , ignoring the vanishing terms. We have

$$\frac{|x^{2(n+1)}|}{|x^{2n}|} = |x^2| = |x|^2 \rightarrow |x|^2 \quad \text{as } n \rightarrow \infty.$$

Therefore the series is convergent if $|x|^2 < 1$, that is, if $|x| < 1$, and diverges if $|x|^2 > 1$, that is, if $|x| > 1$. Thus we conclude that the radius of convergence is $\rho = 1$. \square

Note: If we want to only describe the coefficients c_n with **odd indices** n , we represent all odd non-negative integers by $n = 2k + 1$, $k \in \mathbb{N}_0$, and work with c_{2k+1} . Likewise, if we want to only describe the coefficients c_n with **even indices** n , we represent all even non-negative integers by $n = 2k$, $k \in \mathbb{N}_0$, and work with c_{2k} .

Example 1.19 (ratio test for power series cannot be applied)

Determine the radius of convergence of the power series

$$\sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \dots + (-1)^n \frac{x^{2n}}{(2n)!} + \dots$$

Solution: Since $c_{2k+1} = 0$, $k \in \mathbb{N}_0$, we cannot apply Theorem 1.14, but we may apply the ratio test for sequences in \mathbb{R} . Since

$$\lim_{n \rightarrow \infty} \frac{|(-1)^{n+1} x^{2n+2} / (2n+2)!|}{|(-1)^n x^{2n} / (2n)!|} = \lim_{n \rightarrow \infty} \frac{(2n)!}{(2n+2)!} |x|^2 = \lim_{n \rightarrow \infty} \frac{|x|^2}{(2n+1)(2n+2)} = 0 < 1$$

for all $x \in \mathbb{R}$, the series converges absolutely for all $x \in \mathbb{R}$. Thus the radius of convergence is $\rho = \infty$. \square

A **general formula** for determining the radius of convergence of a power series that can **always** be applied can be given with the help of the limit superior.

Definition 1.20 (limit superior and limit inferior)

The **limit superior** $\limsup_{n \rightarrow \infty} a_n$ of a sequence $\{a_n\} \subset \mathbb{R}$ is defined by

$$\limsup_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} \left(\sup \{a_m : m \geq n\} \right).$$

The **limit inferior** $\liminf_{n \rightarrow \infty} a_n$ of a sequence $\{a_n\} \subset \mathbb{R}$ is defined by

$$\liminf_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} \left(\inf \{a_m : m \geq n\} \right).$$

Example 1.21 (limit superior and limit inferior)

The alternating sequence $\{a_n\}$ defined by

$$a_{2k} = 1, \quad a_{2k+1} = -\frac{1}{k}, \quad k \in \mathbb{N}_0,$$

does not have a limit, but it has a limit superior and a limit inferior. Indeed,

$$\limsup_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} \left(\sup \{a_m : m \geq n\} \right) = \lim_{n \rightarrow \infty} 1 = 1,$$

and

$$\liminf_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} \left(\inf \{a_m : m \geq n\} \right) = \lim_{k \rightarrow \infty} -\frac{1}{k} = 0.$$

We see that the limit superior and the limit inferior are in general not the same. \square

A sequence $\{a_n\} \subset \mathbb{R}$ converges if and only if the limit inferior and the limit superior are the same.

Lemma 1.22 (convergence of a sequence of real numbers)

A sequence $\{a_n\} \subset \mathbb{R}$ converges **if and only if** both the limit superior and the limit inferior of $\{a_n\}$ exist and

$$\limsup_{n \rightarrow \infty} a_n = \liminf_{n \rightarrow \infty} a_n.$$

After this preparation, we can give the general formula for the radius of convergence.

Theorem 1.23 (general formula for the radius of convergence)

The **radius of convergence** ρ of the power series

$$\sum_{n=0}^{\infty} c_n (x - x_0)^n$$

is given by the formula

$$\rho := \frac{1}{\sigma}, \quad \text{where} \quad \sigma := \limsup_{n \rightarrow \infty} \sqrt[n]{|c_n|}. \quad (1.15)$$

The **geometric series formula**

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x} \quad \text{for } x \in (-1, 1)$$

(which we have also used in the proof of Theorem 1.13) is the basis for a number of useful power series expansions. For example, if $|(-x^2)| < 1$ then

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

Here we may think of $-x^2$ as a new variable, that is, by letting $y := -x^2$, we have $1/(1-y) = \sum_{n=0}^{\infty} y^n$ for $|y| < 1$. Replacing y by $-x^2$ we obtain the above power

series for $1/(1+x^2)$ which is absolutely convergent for $|-x^2| < 1$, that is, for $|x| < 1$. If $|x| \geq 1$, we see that the series diverges, since $\{|(-1)^n x^{2n}|\}$ does no longer tend to zero as $n \rightarrow \infty$. Thus we find that $\rho = 1$.

Example 1.24 (power series derived from the geometric series)

Expand each of the following functions into a power series at the given centre and find its radius of convergence:

$$(a) \frac{x^2}{1-x^3} \text{ with } x_0 = 0, \quad (b) \frac{x}{a^2+x^2} \text{ with } a > 0, x_0 = 0, \quad (c) \frac{1}{x} \text{ with } x_0 \neq 0.$$

Solution:

(a) We expand $1/(1-x^3)$ with the geometric series with the argument $y = x^3$ and obtain

$$\frac{x^2}{1-x^3} = x^2 \sum_{n=0}^{\infty} x^{3n} = \sum_{n=0}^{\infty} x^{3n+2}, \quad \text{if } |x^3| < 1.$$

The series converges absolutely if $|x^3| < 1$, that is, if $|x| < 1$, and it diverges if $|x| \geq 1$ (since for $|x| \geq 1$, $\{|x^{3n+2}|\}$ does not tend to zero as $n \rightarrow \infty$). Thus the radius of convergence is $\rho = 1$.

(b) We rewrite the function and then expand with the help of the geometric series:

$$\begin{aligned} \frac{x}{a^2+x^2} &= \frac{x}{a^2 \left[1 + \left(\frac{x}{a}\right)^2\right]} = \frac{x}{a^2 \left[1 - \left(-\left(\frac{x}{a}\right)^2\right)\right]} = \frac{x}{a^2} \sum_{n=0}^{\infty} \left(-\left(\frac{x}{a}\right)^2\right)^n \\ &= \frac{x}{a^2} \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{a^{2n}} = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{a^{2(n+1)}} \quad \text{if } \left|-\left(\frac{x}{a}\right)^2\right| < 1. \end{aligned}$$

By the ratio test for series in \mathbb{R} , we have

$$\frac{|(-1)^{n+1} x^{2(n+1)+1} / a^{2((n+1)+1)}|}{|(-1)^n x^{2n+1} / a^{2(n+1)}|} = \frac{|x|^2}{a^2} \rightarrow \frac{|x|^2}{a^2} \quad \text{as } n \rightarrow \infty.$$

Thus the series is convergent if $|x|^2/a^2 < 1$, that is, if $|x| < a$, and the series diverges if $|x|^2/a^2 > 1$, that is, if $|x| > a$. Thus the radius of convergence is $\rho = a$.

(c) We rewrite the function and then expand with the help of the geometric series:

$$\frac{1}{x} = \frac{1}{(x-x_0)+x_0} = \frac{1}{x_0 \left(1 - \frac{x_0-x}{x_0}\right)} = \frac{1}{x_0} \sum_{n=0}^{\infty} \left(\frac{x_0-x}{x_0}\right)^n = \sum_{n=0}^{\infty} (-1)^n \frac{(x-x_0)^n}{x_0^{n+1}},$$

if $|x_0 - x|/|x_0| < 1$. By the ratio test, we have

$$\sigma = \lim_{n \rightarrow \infty} \frac{|(-1)^{n+1} / x_0^{n+2}|}{|(-1)^n / x_0^{n+1}|} = \lim_{n \rightarrow \infty} |x_0|^{-1} = |x_0|^{-1}.$$

Thus the radius of convergence is $\rho = 1/\sigma = |x_0|$.

Alternatively, we could have determined the radius of convergence as follows: We know that the geometric series has the radius of convergence $\rho = 1$, and hence the power series is absolutely convergent if $|(x_0 - x)/x_0| < 1$, that is, if $|x - x_0| < |x_0|$. From the convergence of the geometric series, we also know that the power series diverges if $|(x_0 - x)/x_0| > 1$, that is, $|x - x_0| > |x_0|$. Thus we find that the radius of convergence is $\rho = |x_0|$ \square

Remark 1.25 (functions similar to $1/(1 - x)$)

*In the examples above, we try to **rewrite** the function in the form $1/(1 - y)$ with y of the form $y = c(x - x_0)^\gamma$, where c is a constant and $\gamma \in \mathbb{N}$ a fixed integer. The known expansion of the geometric series with y as argument gives then a power series expansion centred at x_0 .*

1.2 Taylor series

In this section we will briefly discuss a particular type of power series, the so-called **Taylor series** of an infinitely often differentiable function centred at x_0 .

Definition 1.26 (Taylor series of f centred at x_0)

*Let I be an open interval, and $f : I \rightarrow \mathbb{R}$ be a function that is infinitely often continuously differentiable on I . Let $x_0 \in I$. Then the **Taylor series of f centred at x_0** is defined by*

$$\begin{aligned} & \sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n \\ &= f(x_0) + \frac{f'(x_0)}{1!} (x - x_0) + \frac{f''(x_0)}{2!} (x - x_0)^2 + \dots + \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n + \dots, \end{aligned}$$

where by definition $0! := 1$.

As with every other power series we can determine the **radius of convergence** of the Taylor series. The fact that we have determined the Taylor series starting with an infinitely often continuously differentiable function f raises the following question.

Question: If ρ denotes the radius of convergence of the Taylor series of f centred at x_0 , does this Taylor series **converge** for all $x \in (x_0 - \rho, x_0 + \rho)$ to $f(x)$?

The answer is in general **no**!

Example 1.27 (Taylor series of e^x centred at $x_0 = 0$)

Compute the Taylor series of $f(x) = e^x$ centred at $x_0 = 0$ and inspect it for convergence.

Solution: The exponential function $f(x) = e^x$ is infinitely often continuously differentiable on \mathbb{R} . We know that $(e^x)' = e^x$. Thus $f^{(n)}(x) = e^x$ for all $n \in \mathbb{N}$, and

$$f(0) = f'(0) = \dots = f^{(n)}(0) = e^0 = 1,$$

and we obtain the Taylor series

$$\sum_{n=0}^{\infty} \frac{1}{n!} x^n = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \dots \quad (1.16)$$

Since

$$\left| \frac{1/(n+1)!}{1/n!} \right| = \frac{n!}{(n+1)!} = \frac{1}{n+1} \rightarrow 0 \quad \text{as } n \rightarrow \infty,$$

we see that $\sigma = 0$, and, from the ratio test, the radius of convergence is $\rho = \infty$. The Taylor series of e^x centred at $x_0 = 0$ converges absolutely for all $x \in \mathbb{R}$. \square

Example 1.28 (Taylor series of $\cos x$ centred at $x_0 = 0$)

Determine the Taylor series of $f(x) = \cos x$ centred at $x_0 = 0$, and find its radius of convergence.

Solution: The function $f(x) = \cos x$ is infinitely often continuously differentiable on \mathbb{R} . Thus we can compute the Taylor series centered at $x_0 = 0$. We find that

$$\begin{aligned} f(x) &= \cos x, & f'(x) &= -\sin x, & f''(x) &= -\cos x, \\ \dots, & & f^{(2k)}(x) &= (-1)^k \cos x, & f^{(2k+1)}(x) &= (-1)^{k+1} \sin x. \end{aligned}$$

Since $\sin 0 = 0$ and $\cos 0 = 1$, we see that

$$\begin{aligned} f^{(2k)}(0) &= (-1)^k, & k &\in \mathbb{N}_0; \\ f^{(2k+1)}(0) &= 0, & k &\in \mathbb{N}_0. \end{aligned}$$

Thus the Taylor series of $\cos x$ centred at $x_0 = 0$ is given by

$$\sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} x^{2k} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \frac{x^8}{8!} - \frac{x^{10}}{10!} + \dots \quad (1.17)$$

From the ratio test for sequences of real numbers

$$\lim_{k \rightarrow \infty} \frac{|(-1)^{k+1} x^{2k+2}/(2k+2)!|}{|(-1)^k x^{2k}/(2k)!|} = \lim_{k \rightarrow \infty} \frac{|x|^2}{(2k+1)(2k+2)} = 0 < 1 \quad \text{for all } x \in \mathbb{R},$$

and the series converges for all $x \in \mathbb{R}$. Thus the radius of convergence is $\rho = \infty$. \square

Example 1.29 (Taylor series of $\ln x$ centred at $x_0 = 1$)

Determine the Taylor series of $f(x) = \ln x$ centred at $x_0 = 1$.

Solution: Since the function $f(x) = \ln x$ is infinitely often differentiable on $(0, \infty)$, we can compute the Taylor series of f centred at $x_0 = 1$. We have

$$f'(x) = \frac{1}{x}, \quad f''(x) = \frac{(-1)}{x^2}, \quad f'''(x) = \frac{2!}{x^3}, \quad \dots, \quad f^{(n)}(x) = \frac{(-1)^{n-1}(n-1)!}{x^n}.$$

Thus,

$$\begin{aligned} f(1) &= \ln 1 = 0; \\ f^{(n)}(1) &= (-1)^{n-1}(n-1)! \quad \text{for all } n \in \mathbb{N}, \end{aligned}$$

and the Taylor series of $f(x) = \ln x$ centred at $x_0 = 1$ is given by

$$\sum_{n=1}^{\infty} \frac{(-1)^{n-1}(n-1)!}{n!} (x-1)^n = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} (x-1)^n. \quad (1.18)$$

From the ratio test, we have

$$\frac{|(-1)^n/(n+1)|}{|(-1)^{n-1}/n|} = \frac{n}{(n+1)} \rightarrow 1 \quad \text{as } n \rightarrow \infty.$$

Thus $\sigma = 1$, and the radius of convergence is $\rho = 1/\sigma = 1$. The Taylor series of $f(x) = \ln x$ centred at $x_0 = 1$ converges for $x \in (1-1, 1+1) = (0, 2)$. \square

Example 1.30 (Taylor series of a polynomial)

Let $p : \mathbb{R} \rightarrow \mathbb{R}$ be a polynomial of degree m

$$p(x) = a_m x^m + a_{m-1} x^{m-1} + \dots + a_1 x + a_0$$

with the coefficients $a_0, a_1, \dots, a_{m-1}, a_m \in \mathbb{R}$. Then the Taylor series of p about $x_0 = 0$ is just the polynomial p itself.

Proof: We compute the derivatives of p at $x_0 = 0$. Since p is a polynomial of degree m , we have that $p^{(n)} = 0$ for all $n > m$. For $k \leq m$, we have

$$f(0) = a_0, \quad f'(0) = a_1 = 1! a_1, \quad f^{(2)}(0) = 2! a_2, \quad \dots, \quad f^{(n)}(0) = n! a_n, \quad \dots \quad (n \leq m).$$

Thus we see that the Taylor series of p about $x_0 = 0$ is given by

$$\sum_{n=0}^m \frac{n! a_n}{n!} x^n = \sum_{n=0}^m a_n x^n = p(x)$$

as claimed. \square

Example 1.31 (function whose Taylor series vanishes)

Let the function $f : \mathbb{R} \rightarrow \mathbb{R}$ be defined by

$$f(x) := \begin{cases} e^{-1/x^2} & \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases}$$

It can be shown that this function is arbitrarily often continuously differentiable (not trivial!). Thus we can compute its Taylor series centred at $x_0 = 0$. Computation of the derivatives and making use of $\lim_{x \rightarrow 0} x^{-k} e^{-1/x^2} = 0$ for all $k \in \mathbb{N}$, yields that all derivatives of f vanish at $x_0 = 0$. (These computations are not trivial!) Since $f^{(n)}(0) = 0$ for all $n = 0, 1, 2, \dots$, the Taylor series is zero. Since the function f is only at the point $x_0 = 0$ zero, we see in this case that the Taylor series converges on \mathbb{R} , but only at the point $x_0 = 0$ does it converge to f . \square

Remark 1.32 (Taylor series of f does not need to converge to f)

*From the last example we see that the Taylor series of an infinitely often differentiable function f (centred at a point x_0), with radius of convergence $\rho > 0$, **need not converge to the function f in any point from $(x_0 - \rho, x_0 + \rho)$, other than x_0 itself.***

In the next section we will learn a criterion for checking whether the Taylor series of an infinitely often continuously differentiable function f (centred at x_0) converges at a given x to $f(x)$.

1.3 Taylor's Formula

Taylor's formula (centred at x_0) gives an expansion of a $(n+1)$ -times continuously differentiable function f as a **polynomial of degree n** , given by

$$s_n(x) = \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k$$

and a **remainder term**, which describes the ‘error’. Taylor’s formula is a generalization of the mean value theorem.

Theorem 1.33 (Taylor’s Formula)

Let $f : (a, b) \rightarrow \mathbb{R}$ be $(n+1)$ -times continuously differentiable, and let $x_0 \in (a, b)$. Then for any $x \in (a, b)$, $x \neq x_0$, there exists some ξ strictly between x_0 and x (that is, $\xi \in (x_0, x)$ if $x_0 < x$ and $\xi \in (x, x_0)$ if $x < x_0$) such that

$$\begin{aligned} f(x) &= \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k + R_n \\ &= f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2!} (x - x_0)^2 + \dots + \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n + R_n \end{aligned} \quad (1.19)$$

with the **remainder term**

$$R_n := \frac{f^{(n+1)}(\xi)}{(n+1)!} (x - x_0)^{n+1}, \quad (1.20)$$

where by definition $0! := 1$. (Note that ξ occurs only in the remainder term.) We call the sum on the right-hand side of (1.19), that is,

$$\begin{aligned} s_n(x) &:= \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k \\ &= f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2!} (x - x_0)^2 + \dots + \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n \end{aligned}$$

the **Taylor polynomial of f of degree n centred at x_0** . The formula (1.19) is called **Taylor’s formula** (up to degree n centred at x_0). Note that ξ depends on n , x_0 and x .

Remark 1.34 (some properties of Taylor’s formula)

(1) For $n = 0$ we get the **mean value theorem**:

$$f(x) = f(x_0) + f'(\xi)(x - x_0), \quad \text{with } \xi \in (x, x_0) \text{ if } x < x_0 \text{ and } \xi \in (x_0, x) \text{ if } x_0 < x.$$

(2) For polynomials p of degree m the last term disappears if we take $n = m$, so that Taylor’s formula reads

$$p(x) = p(x_0) + p'(x_0)(x - x_0) + \frac{p''(x_0)}{2!} (x - x_0)^2 + \dots + \frac{p^{(m)}(x_0)}{m!} (x - x_0)^m.$$

(3) We shall interpret Taylor's formula as a **polynomial approximation** of the function f with an **explicit remainder term** R_n .

Remark 1.35 (connection with Taylor series)

If the function f in Taylor's formula is infinitely often continuously differentiable on (a, b) , then we can also compute its Taylor series centred at x_0 . We see that the Taylor polynomial of f up degree n is the partial sum

$$s_n(x) := \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k$$

of the Taylor series centred at x_0 , and we can write Taylor's formula (1.19) as

$$f(x) - s_n(x) = f(x) - \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k = R_n = \frac{f^{(n+1)}(\xi)}{(n+1)!} (x - x_0)^{n+1}.$$

We see that the Taylor series of f centred at x_0 converges at the point x to $f(x)$ if and only if

$$\lim_{n \rightarrow \infty} |f(x) - s_n(x)| = \lim_{n \rightarrow \infty} |R_n| = \lim_{n \rightarrow \infty} \left| \frac{f^{(n+1)}(\xi)}{(n+1)!} (x - x_0)^{n+1} \right| = 0.$$

With the help of Remark 1.35, we can now check for which x the Taylor series of the function f (centred at $x_0 = 0$) in Examples 1.27 and 1.28 does converge towards $f(x)$. Whether the Taylor series of $f(x) = \ln x$ centred at $x_0 = 1$ (see Example 1.29) converges to $f(x) = \ln(x)$ for $x \in (0, 2)$ is not obvious from considering the remainder term, and therefore we will not discuss it here.

Example 1.36 (Examples 1.27 and 1.28 continued)

Before we discuss the individual examples, we observe that

$$\lim_{n \rightarrow \infty} \frac{\alpha^n}{n!} = 0 \quad \text{for all } \alpha \geq 0. \quad (1.21)$$

This can be seen as follows: Choose a fixed integer N such that $2\alpha \leq N$, or equivalently, $\alpha/N \leq 1/2$. Then for all $n \geq N$

$$0 \leq \frac{\alpha^n}{n!} = \frac{\alpha^N}{N!} \left(\frac{\alpha}{(N+1)} \frac{\alpha}{(N+2)} \cdots \frac{\alpha}{n} \right) \leq \frac{\alpha^N}{N!} \left(\frac{1}{2} \right)^{n-N} \rightarrow 0 \quad \text{as } n \rightarrow \infty,$$

where the estimate follows from $\alpha/m \leq \alpha/N \leq 1/2$ for all $m \geq N$. From the sandwich theorem we see that (1.21) holds true.

Now we can show that in Examples 1.27 and 1.28 the Taylor series of the function f centred at $x_0 = 0$ converges for all $x \in \mathbb{R}$ to $f(x)$.

(a) In Example 1.27, we considered the Taylor series of $f(x) = e^x$ centred at $x_0 = 0$, which converges absolutely for all $x \in \mathbb{R}$. For any $x \in \mathbb{R}$, we have

$$0 \leq |R_n| = \left| \frac{f^{(n+1)}(\xi)}{(n+1)!} x^{n+1} \right| = \left| \frac{e^\xi}{(n+1)!} x^{n+1} \right| \leq \frac{e^{|\xi|} |x|^{n+1}}{(n+1)!} \leq \frac{e^{|x|} |x|^{n+1}}{(n+1)!},$$

where we have used that $0 < |\xi| < |x|$ (since ξ lies strictly between 0 and x). Thus from the sandwich theorem and from (1.21),

$$0 \leq \lim_{n \rightarrow \infty} |R_n| \leq \lim_{n \rightarrow \infty} \frac{e^{|x|} |x|^{n+1}}{(n+1)!} = 0,$$

that is, $\lim_{n \rightarrow \infty} |R_n| = 0$. Thus the Taylor series (1.16) of $f(x) = e^x$ centred at $x_0 = 0$ converges for every $x \in \mathbb{R}$ to e^x .

(b) In Example 1.28, we computed the Taylor series of $f(x) = \cos x$ centred at $x_0 = 0$, which converges absolutely for all $x \in \mathbb{R}$. We have, for all $x \in \mathbb{R}$,

$$0 \leq |R_n| = \left| \frac{f^{(n+1)}(\xi)}{(n+1)!} x^{n+1} \right| = \begin{cases} \left| \frac{(-1)^k \cos \xi}{(2k)!} x^{2k} \right| \leq \frac{|x|^{2k}}{(2k)!} & \text{if } n+1 = 2k, \\ \left| \frac{(-1)^{k+1} \sin \xi}{(2k+1)!} x^{2k+1} \right| \leq \frac{|x|^{2k+1}}{(2k+1)!} & \text{if } n+1 = 2k+1, \end{cases}$$

where we have used that $|\sin \xi| \leq 1$ and $|\cos \xi| \leq 1$. From (1.21), we see that both upper bounds tend to zero as $n \rightarrow \infty$, and thus $k \rightarrow \infty$, and from the sandwich theorem $\lim_{n \rightarrow \infty} |R_n| = 0$. Thus the Taylor series of $f(x) = \cos x$ centred at $x_0 = 0$ converges for every $x \in \mathbb{R}$ to $\cos x$. \square

Now we will prove Theorem 1.33. The proof uses **Rolle's theorem** which says the following: Let $a < b$, and let $f : [a, b] \rightarrow \mathbb{R}$ be a continuous function satisfying $f(a) = f(b)$. If f is differentiable in (a, b) , then there exists a point $\xi \in (a, b)$ such that $f'(\xi) = 0$.

Proof of Theorem 1.33: For simplicity we will only consider the case $x > x_0$. (The case $x < x_0$ can be treated analogously.) Fix $x > x_0$ and consider the function

$$R_n(y) := f(x) - f(y) - f'(y)(x-y) - \frac{f''(y)}{2!}(x-y)^2 - \dots - \frac{f^{(n)}(y)}{n!}(x-y)^n,$$

where y is the independent variable. Let us find the derivative of R_n :

$$R'_n(y) = -f'(y) - (f''(y)(x-y) - f'(y)) - \left(\frac{f'''(y)}{2!}(x-y)^2 - \frac{f''(y)}{2!} 2(x-y) \right)$$

$$- \dots - \left(\frac{f^{(n+1)}(y)}{n!} (x-y)^n - \frac{f^{(n)}(y)}{n!} n(x-y)^{n-1} \right).$$

Carefully making all cancellations we get

$$R'_n(y) = - \frac{f^{(n+1)}(y)}{n!} (x-y)^n. \quad (1.22)$$

Now we consider the function

$$G_n(y) := R_n(y) - \left(\frac{x-y}{x-x_0} \right)^{n+1} R_n(x_0).$$

We observe that $G_n(x_0) = G_n(x) = 0$, where we have used that $R_n(x) = 0$. Since G_n is continuously differentiable on (x_0, x) , we now from Rolle's theorem that there exists a $\xi \in (x_0, x)$ such that $G'_n(\xi) = 0$. From (1.22) we obtain for G'_n

$$\begin{aligned} G'_n(y) &= R'_n(y) + \frac{(n+1)(x-y)^n}{(x-x_0)^{n+1}} R_n(x_0) \\ &= - \frac{f^{(n+1)}(y)}{n!} (x-y)^n + \frac{(n+1)(x-y)^n}{(x-x_0)^{n+1}} R_n(x_0), \end{aligned}$$

and Substituting $x = \xi$ yields

$$0 = G'_n(\xi) = - \frac{f^{(n+1)}(\xi)}{n!} (x-\xi)^n + \frac{(n+1)(x-\xi)^n}{(x-x_0)^{n+1}} R_n(x_0).$$

We bring the first term on the other side of the equation and multiply with the factor in front of $R_n(x_0)$. Thus

$$\frac{f^{(n+1)}(\xi)}{(n+1)!} (x-x_0)^{n+1} = R_n(x_0).$$

Substituting the definition of $R_n(x_0)$ into the formula above, now yields (1.19) and completes the proof. \square

Now we will use Taylor's formula to obtain a **polynomial approximation** of a function as explained in Remark 1.34 (3).

Example 1.37 (polynomial approximation with Taylor's formula for e^x)

We have $f^{(k)}(x) = e^x$ for all $k \in \mathbb{N}$, hence, $f(0) = f'(0) = \dots = f^{(n)}(0) = e^0 = 1$. Thus Taylor's formula centred at $x_0 = 0$ up to order n reads

$$e^x = 1 + x + \frac{x^2}{2!} + \dots + \frac{x^n}{n!} + \frac{e^\xi}{(n+1)!} x^{n+1}, \quad (1.23)$$

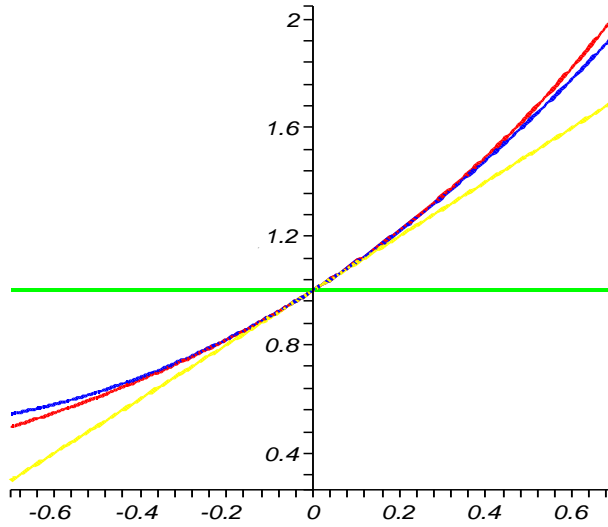


Figure 1.1: Approximation of $f(x) = e^x$ (red) by its Taylor polynomial centred at $x_0 = 0$ of degree $n = 0$ (green), $n = 1$ (yellow) and $n = 2$ (blue).

with some ξ strictly between 0 and x . We want to use Taylor's formula to approximate $e^{0.1}$ by a polynomial of degree $n = 2$ and to get an error estimate for the quality of the approximation. From Taylor's formula (1.23) with $n = 2$,

$$\left| e^x - \left(1 + x + \frac{x^2}{2} \right) \right| = \left| \frac{e^\xi}{6} x^3 \right| \leq \frac{e^{|x|}}{6} |x|^3, \quad (1.24)$$

where we have used in the last step that $0 < |\xi| < |x|$ (since ξ is strictly between 0 and x) and that the function $f(x) = e^x$ is monotonically increasing. The right-hand side gives an estimate for the approximation of e^x by its Taylor polynomial of degree $n = 2$. For $x = 0.1$, we find that the error has the upper bound

$$\left| e^{0.1} - \left(1 + 0.1 + \frac{(0.1)^2}{2} \right) \right| \leq \frac{e^{0.1}}{6} (0.1)^3 \approx 0.000184,$$

Since $x = 0.1$ is rather small, we obtain a fairly decent approximation of $e^{0.1}$ by the Taylor polynomial of degree $n = 2$. For large x , the approximation of e^x by the Taylor polynomial of degree n is only a good approximation for rather large n . \square

Example 1.38 (Taylor's formula for $\cos x$ centred at $x = \pi$)

Write down Taylor's formula up to the power $2n$ for $f(x) = \cos x$ centred at $x_0 = \pi$.

Solution: As we have seen in Example 1.28, we have for $f(x) = \cos x$

$$f^{(2k)}(x) = (-1)^k \cos x, \quad f^{(2k+1)}(x) = (-1)^{k+1} \sin x, \quad k \in \mathbb{N}_0.$$

Since $\cos \pi = -1$ and $\sin \pi = 0$, we have $f(\pi) = -1$, and

$$f^{(2k)}(\pi) = (-1)^{k+1} \quad \text{and} \quad f^{(2k+1)}(\pi) = 0, \quad k \in \mathbb{N}_0.$$

Thus Taylor's formula up to degree $2n$ for $f(x) = \cos x$ centred at $x_0 = \pi$ reads

$$\cos x = -1 + \frac{(x - \pi)^2}{2!} - \frac{(x - \pi)^4}{4!} + \dots + (-1)^{n+1} \frac{(x - \pi)^{2n}}{(2n)!} + \frac{(-1)^{n+1} \sin \xi}{(2n+1)!} (x - \pi)^{2n+1},$$

with some ξ strictly between x and π . \square

Example 1.39 (Taylor's formula as a polynomial approximation)

Consider $f(x) = \cos x$. We want to estimate the difference between $f(x)$ and its Taylor's polynomials of degree n centred at $x_0 = 0$ for $n = 0, 1, 2, 3$.

(a) Let $n = 0$. Then, from $f'(x) = -\sin x$,

$$f(x) = f(0) + f'(\xi)x \quad \Leftrightarrow \quad \cos x = 1 - (\sin \xi)x,$$

for some ξ strictly between 0 and x . Since $|\sin \xi| \leq 1$, we have

$$|\cos x - 1| = |\sin \xi| |x| \leq |x|. \quad (1.25)$$

(b) Let $n = 1$. Then from $f'(x) = -\sin x$ and $f''(x) = -\cos x$,

$$f(x) = f(0) + f'(0)x + \frac{f''(\xi)}{2}x^2 \quad \Leftrightarrow \quad \cos x = 1 - (\sin 0)x - \frac{\cos \xi}{2}x^2.$$

with some ξ strictly between 0 and x . Since $f'(0) = \sin 0 = 0$, the second term on the right-hand side vanishes, and we have, because $|\cos \xi| \leq 1$,

$$\cos x - 1 = -\frac{\cos \xi}{2}x^2 \quad \Rightarrow \quad |\cos x - 1| = \left| \frac{\cos \xi}{2}x^2 \right| \leq \frac{|x|^2}{2}. \quad (1.26)$$

While the approximation of $\cos x$ is in both (1.25) and (1.26) the constant function with value 1, we observe that (1.26) is a better estimate if $|x| < 1$.

(c) Let $n = 3$. Then $f'(x) = -\sin x$, $f''(x) = -\cos x$, and $f'''(x) = \sin x$, and Taylor's formula up to the order 2 centred at $x_0 = 0$ reads

$$\cos x = 1 - \frac{x^2}{2} + \frac{\sin \xi}{6}x^3,$$

with some ξ strictly between 0 and x . Thus

$$\left| \cos x - \left(1 - \frac{x^2}{2} \right) \right| = \frac{|\sin \xi|}{6} |x|^3 \leq \frac{|x|^3}{6}. \quad (1.27)$$

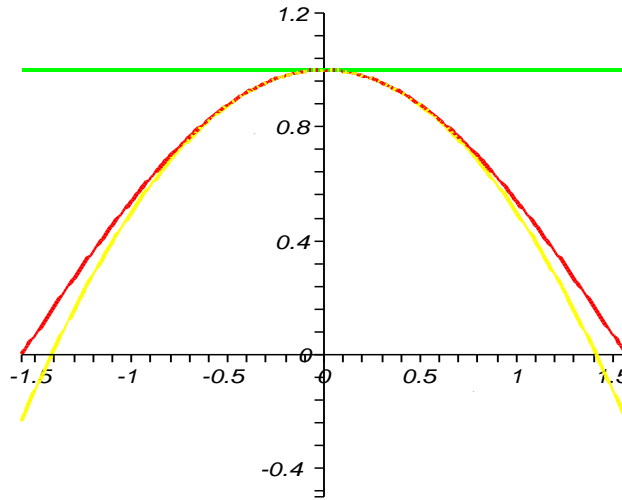


Figure 1.2: Approximation of $f(x) = \cos x$ (red) by its Taylor polynomial of degree $n = 0$ (green) and $n = 2$ (yellow)

(d) Repeat the same argument with $n = 3$. From $f'(x) = -\sin x$, $f''(x) = -\cos x$, $f'''(x) = \sin x$, and $f^{(4)}(x) = \cos x$, and $\sin 0 = 0$ and $\cos 0 = 1$, we find

$$\cos x = 1 - \frac{x^2}{2} + \frac{\cos \xi}{4!} x^4,$$

with some ξ strictly between x and 0. This implies that

$$\left| \cos x - \left(1 - \frac{x^2}{2} \right) \right| = \frac{|\cos \xi|}{4!} |x|^4 \leq \frac{|x|^4}{24}.$$

This formula gives a very good approximation for $\cos x$ with small x . Namely, if $x = 0.1$, we have

$$|\cos(0.1) - 0.995| \leq \frac{1}{240000} \approx 4.167 \times 10^{-6}.$$

This is why our interpretation of Taylor's formula as a polynomial approximation with an explicit remainder makes sense. \square

Remark 1.40 (other forms of Taylor's formula)

Under the same assumptions as in Theorem 1.33, we have the following other forms of Taylor's formula:

(1) For any $x \in (a, b)$, $x \neq x_0$, there exists some $\theta \in (0, 1)$ such that

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 + \dots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n + R_n$$

with the remainder term

$$R_n = \frac{f^{(n+1)}(x_0 + \theta(x - x_0))}{(n+1)!} (x - x_0)^{n+1}.$$

This follows directly from (1.19) by taking $\theta = (\xi - x_0)/(x - x_0)$.

(2) For any h with $x_0 + h \in (a, b)$ and $h \neq 0$ there exists some $\theta \in (0, 1)$ such that

$$f(x_0 + h) = f(x_0) + f'(x_0)h + \frac{f''(x_0)}{2!}h^2 + \dots + \dots + \frac{f^{(n)}(x_0)}{n!}h^n + R_n$$

with the remainder term

$$R_n = \frac{f^{(n+1)}(x_0 + \theta h)}{(n+1)!} h^{n+1}.$$

This version of the formula follows from the previous one with $h = x - x_0$.

(3) There is an integral form of Taylor's formula obtained by integration by parts which will be discussed later.

Chapter 2

Introduction of the Riemann Integral

In this chapter we will introduce the Riemann integral $\int_a^b f(x) dx$ as the area under the curve/graph of f from $x = a$ to $x = b$. The Riemann integral encompasses the usual notion of an integral which you have encountered in school, but we will see that the precise mathematical definition of the Riemann integral allows us to integrate functions that are not continuous and that you could not integrate with those methods that you learnt in school.

2.1 Lower and Upper Sum of a Function With Respect to a Partition

We want to define the **Riemann integral**

$$\int_a^b f(x) dx$$

as the **signed area under the graph/curve of f from $x = a$ to $x = b$** . See Figures 2.1, 2.2, and 2.3 for illustration.

For approximating the area under the curve we partition the interval $[a, b]$ and then approximate the area by columns as indicted in Figure 2.4. We will define this more rigorously once we have introduced some more notation.

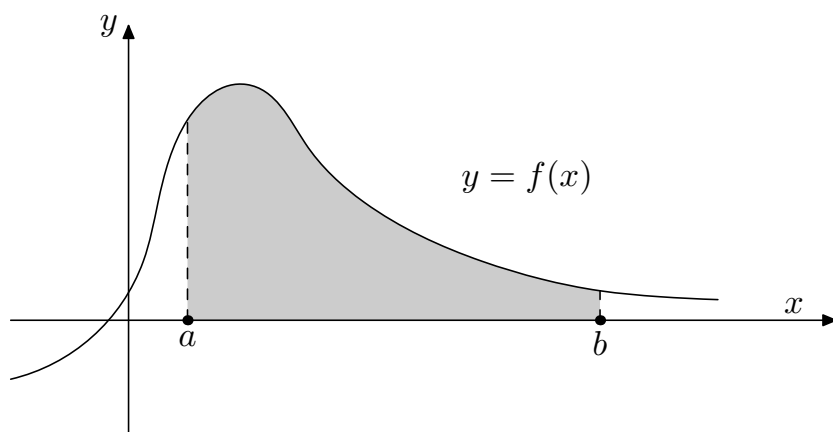


Figure 2.1: Positive area under the graph of f from $x = a$ to $x = b$.

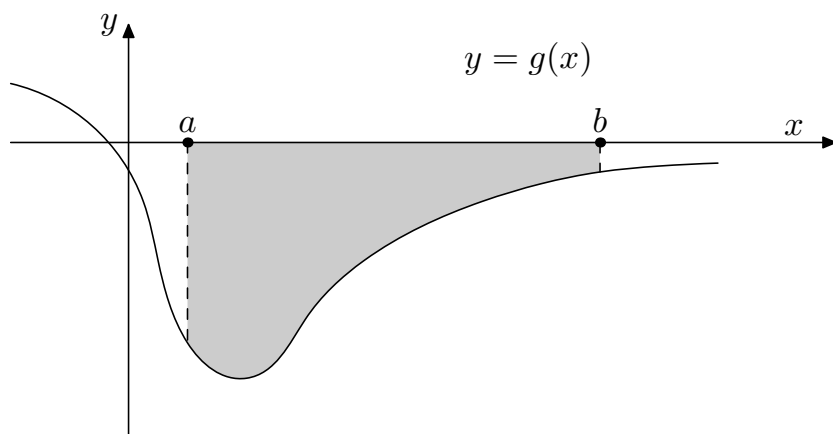


Figure 2.2: Negative area ‘under’ the graph of g from $x = a$ to $x = b$.

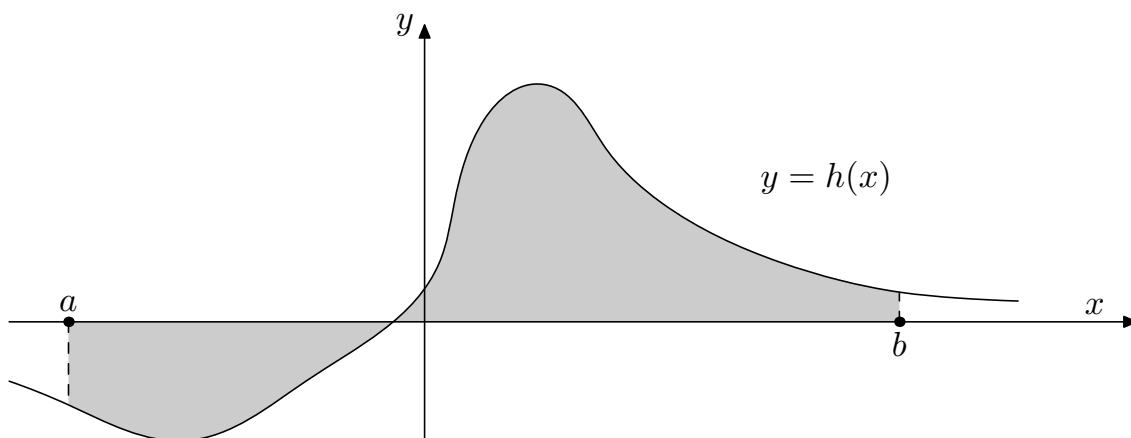


Figure 2.3: Area ‘under’ the graph of h from $x = a$ to $x = b$ with different signs.

We start by introducing the notion of a **bounded function**, that is, a function whose values do not get arbitrarily large and arbitrarily small. Then we introduce **partitions**: a partition of an interval $[a, b]$ is a collection of points $P := \{x_0, x_1, \dots, x_{n-1}, x_n\}$ from $[a, b]$, such that,

$$a = x_0 < x_1 < x_2 < \dots < x_{k-1} < x_k < \dots < x_{n-1} < x_n = b.$$

The name partition is motivated by the fact that the points $x_0, x_1, \dots, x_{n-1}, x_n$ give a **subdivision** of the interval $[a, b]$ into the intervals

$$[x_0, x_1], [x_1, x_2], \dots, [x_{k-1}, x_k], \dots, [x_{n-1}, x_n] = [x_{n-1}, b].$$

To **approximate the area under the graph of a function** f , we will erect over each subinterval $[x_{k-1}, x_k]$ a rectangular box that approximates the area under the graph over this subinterval. An approximation of the area under the graph is then given by the sum of the areas of the rectangular boxes for all subintervals. By shrinking the width of these subintervals we will get a better and better approximation of the area under the graph, and this will lead us to a definition of the **Riemann integral**.

Definition 2.1 (bounded function)

A function $f : [a, b] \rightarrow \mathbb{R}$ is called **bounded** if there exist $m, M \in \mathbb{R}$ such that

$$m \leq f(x) \leq M \quad \text{for all } x \in [a, b].$$

Equivalently, a function $f : [a, b] \rightarrow \mathbb{R}$ is called **bounded** if there exists $K \in \mathbb{R}$ such that

$$|f(x)| \leq K \quad \text{for all } x \in [a, b].$$

The **set of bounded functions** $f : [a, b] \rightarrow \mathbb{R}$ will be denoted by $\mathcal{B}([a, b])$.

Definition 2.2 (partition, width of a partition, and refinement)

- (i) A **partition** $P = \{x_0, \dots, x_n\}$ of an interval $[a, b]$ is a finite set of points satisfying $a = x_0 < x_1 < \dots < x_{n-1} < x_n = b$. The **set of all partitions of a given interval** $[a, b]$ will be denoted by $\mathcal{P}([a, b])$.
- (ii) The **width** of a partition P is the number $w(P) = \max_{k=1,2,\dots,n} (x_k - x_{k-1})$.
- (iii) Let P and Q be two partitions of the same interval $[a, b]$. We say that Q is a **refinement** of P if $P \subset Q$.

Example 2.3 (Examples of partitions on $[0, 1]$)

- (a) Let $P = \{0, \frac{1}{3}, 1\}$ and $Q = \{0, \frac{1}{3}, \frac{1}{2}, 1\}$. Here Q is a refinement of P , and we have $w(P) = 2/3$, $w(Q) = 1/2$.
- (b) Let $P = \{0, \frac{1}{3}, 1\}$ and $Q = \{0, \frac{1}{4}, \frac{1}{3}, 1\}$. Here Q is a refinement of P , and we have $w(P) = 2/3$, $w(Q) = 2/3$.
- (c) Example where neither partition is a refinement of the other: let $P = \{0, \frac{1}{3}, 1\}$ and $Q = \{0, \frac{1}{2}, 1\}$. Here we have $w(P) = 2/3$ and $w(Q) = 1/2$.
- (d) Take $P = \{0, \frac{1}{3}, 1\}$ and $Q = \{0, \frac{1}{2}, 1\}$. Then $P \cup Q = \{0, \frac{1}{3}, \frac{1}{2}, 1\}$ is a refinement of both P and Q . This is a general fact: **$P \cup Q$ is always a refinement of both P and Q .**
- (e) Partition of $[0, 1]$ into n equal subintervals (equally spaced partition),

$$P_n = \left\{0, \frac{1}{n}, \frac{2}{n}, \dots, \frac{n-1}{n}, 1\right\}.$$

For an arbitrary interval $[a, b]$, the **equally spaced partition** into n subintervals of equal length is given by

$$P_n := \left\{x_k := a + k \frac{b-a}{n} : k = 0, 1, 2, \dots, n\right\}.$$

Lemma 2.4 (partition width decreases under refinement)

If $P, Q \in \mathcal{P}([a, b])$ and $P \subset Q$, then $w(P) \geq w(Q)$.

Proof of Lemma 2.4: Since we have additional points in the partition Q , the maximal length of any subinterval in Q is smaller or equal to the maximal length of any subinterval in P . Thus $w(P) \geq w(Q)$. \square

In order to define the rectangular box over each of the subintervals $[x_{k-1}, x_k]$ that approximates the area of f over $[x_{k-1}, x_k]$, we need the notion of the **infimum** and the **supremum** which you have already encountered in your first year at university.

Definition 2.5 (infimum)

Let $X \subset \mathbb{R}$ be a non-empty subset of the real numbers. A number α is called the **infimum** of X (and denoted $\alpha = \inf X$) if

(i) $\alpha \leq x$ for all $x \in X$, and

(ii) for every $\varepsilon > 0$ there exists an $x \in X$ such that $x < \alpha + \varepsilon$.

Note: infimum = largest lower bound.

Definition 2.6 (supremum)

Let $X \subset \mathbb{R}$ be a non-empty subset of the real numbers. A number β is called the **supremum** of X (and denoted $\beta = \sup X$) if

- (i) $\beta \geq x$ for all $x \in X$, and
- (ii) for every $\varepsilon > 0$ there exists an $x \in X$ such that $x > \beta - \varepsilon$.

Note: supremum = least upper bound.

From the **completeness axiom** for the real numbers we know that any bounded set of real numbers has an infimum and supremum.

Example 2.7 (infimum and supremum of $(-1, 1]$)

Find the infimum and the supremum of $X = \{x \in \mathbb{R} : -1 < x \leq 1\} = (-1, 1]$.

Solution: We claim that the number $\alpha = -1$ is the infimum of X and that the number $\beta = 1$ is the supremum of X .

Proof that $\inf(-1, 1] = -1$:

- (i) From the definition of the interval $(-1, 1]$, $-1 \leq x$ for all $x \in (-1, 1]$.
- (ii) Let $\varepsilon > 0$ be arbitrary. We must find $x \in (-1, 1]$ such that $x < -1 + \varepsilon$. Try taking $x = -1 + \frac{\varepsilon}{2}$, but then for large ε we would get $x \notin (-1, 1]$. This difficulty can be overcome by taking $x = \min\{-1 + \frac{\varepsilon}{2}, 0\}$. Because

$$x = \min\left\{-1 + \frac{\varepsilon}{2}, 0\right\} \leq -1 + \frac{\varepsilon}{2} < -1 + \varepsilon,$$

we will still have $x < -1 + \varepsilon$ and it is guaranteed that $x \in (-1, 1]$.

Proof that $\sup(-1, 1] = 1$:

- (i) From the definition of the interval $(-1, 1]$, $x \leq 1$ for all $x \in (-1, 1]$
- (ii) Let ε be arbitrary. Take $x = \max\{0, 1 - \frac{\varepsilon}{2}\}$. Then $x \in (-1, 1]$, $x < 1$, and $x = \max\{0, 1 - \frac{\varepsilon}{2}\} \geq 1 - \frac{\varepsilon}{2} > 1 - \varepsilon$.

Thus $\inf(-1, 1] = -1$ and $\sup(-1, 1] = 1$. □

Remark 2.8 (infimum and supremum may belong to X or not)

With an analogous proof as in Example 2.7, we can show that the closed interval $[a, b]$, the open interval (a, b) , and the half-open intervals $(a, b]$ and $[a, b)$ have all the

infimum $\alpha = a$ and the supremum $\beta = b$. We see that the supremum $\sup X$ (and the infimum $\inf X$) of a set of real numbers $X \subset \mathbb{R}$ **may belong** to the set X or **not**.

Example 2.9 (infimum and supremum of $\{1/n : n \in \mathbb{N}\}$)

Determine the infimum and supremum of

$$X := \left\{ \frac{1}{n} : n \in \mathbb{N} \right\} = \left\{ 1, \frac{1}{2}, \frac{1}{3}, \dots, \frac{1}{n}, \frac{1}{n+1}, \dots \right\}.$$

Solution: Claim: $\inf X = 0$ and $\sup X = 1$

Proof that $\inf X = 0$: (i) Clearly $0 < \frac{1}{n}$ for all $n \in \mathbb{N}$. (ii) Let $\varepsilon > 0$ be arbitrary. We choose $x = \frac{1}{m}$ with some $m \in \mathbb{N}$ such that

$$\frac{1}{m} < 0 + \varepsilon = \varepsilon.$$

Then $x \in X$ and $x < 0 + \varepsilon$.

Proof that $\sup X = 1$: (i) Clearly $x \leq 1$ for all $x \in X$. (ii) For any $\varepsilon > 0$ we take $x = 1$. Then $x \in X$ and $x > 1 - \varepsilon$. Thus $\inf X = 0$ and $\sup X = 1$. \square

Now we can finally introduce the **approximation of the area under the graph of a bounded function f with rectangles**. As implied before, we choose a partition, and over each subinterval $[x_{k-1}, x_k]$, we take the **largest rectangle that can still be fitted under the graph**. The sum of the areas of all these rectangles yields the **lower sum**. If we take over each subinterval $[x_{k-1}, x_k]$ instead the **smallest rectangle that still contains the graph**, and then take the sum of the areas of all these rectangles, then we get the **upper sum**. See Figure 2.4 for illustration.

Definition 2.10 (lower sum and upper sum)

Let $f \in \mathcal{B}([a, b])$, and let $P \in \mathcal{P}([a, b])$ be given by $P = \{x_0, \dots, x_n\}$. Then the **lower sum** of f with respect to the partition P is defined by

$$L(f, P) := \sum_{k=1}^n \left(\inf_{x \in [x_{k-1}, x_k]} f(x) \right) (x_k - x_{k-1}),$$

and the **upper sum** of f with respect to the partition P is defined by

$$U(f, P) := \sum_{k=1}^n \left(\sup_{x \in [x_{k-1}, x_k]} f(x) \right) (x_k - x_{k-1}).$$

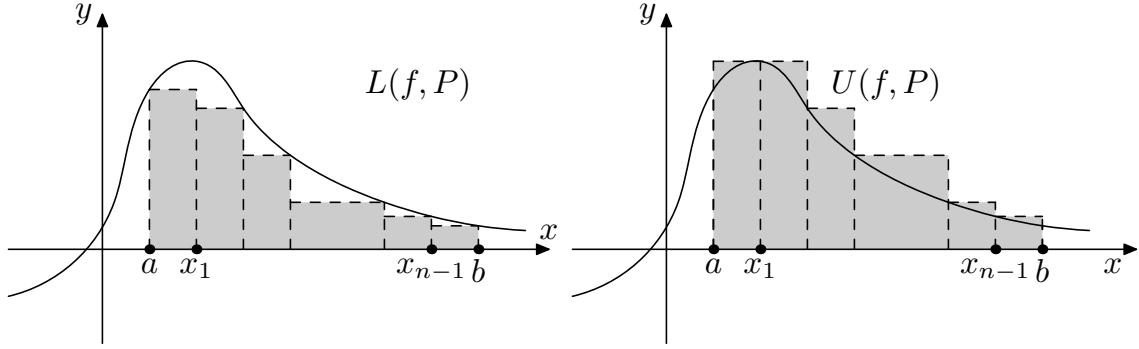


Figure 2.4: The lower sum $L(f, P)$ in the left picture and the upper sum $U(f, P)$ in the right picture.

Note: Since $f \in \mathcal{B}([a, b])$, we know that $\{f(x) : x \in [c, d]\}$ is bounded for any $a \leq c < d \leq b$ and hence

$$\inf_{x \in [c, d]} f(x) := \inf \{f(x) : x \in [c, d]\} \quad \text{and} \quad \sup_{x \in [c, d]} f(x) := \sup \{f(x) : x \in [c, d]\}$$

do exist and are finite. Thus

$$\begin{aligned} \inf_{x \in [x_{k-1}, x_k]} f(x) &:= \inf \{f(x) : x \in [x_{k-1}, x_k]\}, \\ \sup_{x \in [x_{k-1}, x_k]} f(x) &:= \sup \{f(x) : x \in [x_{k-1}, x_k]\} \end{aligned}$$

exist and are finite. With the abbreviated notation

$$m_k(f) := \inf_{x \in [x_{k-1}, x_k]} f(x), \quad M_k(f) := \sup_{x \in [x_{k-1}, x_k]} f(x),$$

we may write $L(f, P)$ and $U(f, P)$ as

$$L(f, P) = \sum_{k=1}^n m_k(f) (x_k - x_{k-1}), \quad U(f, P) = \sum_{k=1}^n M_k(f) (x_k - x_{k-1}).$$

Let us consider some examples.

Example 2.11 (lower and upper sum usually differ)

Let $f : [7, 11] \rightarrow \mathbb{R}$ be defined by

$$f(x) := \begin{cases} 2 & \text{if } x \text{ is rational,} \\ 5 & \text{if } x \text{ is irrational.} \end{cases}$$

Let P be an arbitrary partition of $[7, 11]$. Find $L(f, P)$ and $U(f, P)$.

Solution: Let $P = \{x_0, \dots, x_n\}$ be an arbitrary partition of $[7, 11]$. Any interval $[x_{k-1}, x_k]$ contains both rationals and irrationals, therefore

$$\inf_{x \in [x_{k-1}, x_k]} f(x) = 2 \quad \text{and} \quad \sup_{x \in [x_{k-1}, x_k]} f(x) = 5.$$

Hence the lower and upper sum of f with respect to P are given by

$$\begin{aligned} L(f, P) &= \sum_{k=1}^n \left(\inf_{x \in [x_{k-1}, x_k]} f(x) \right) (x_k - x_{k-1}) = 2 \sum_{k=1}^n (x_k - x_{k-1}) = 2(11 - 7) = 8, \\ U(f, P) &= \sum_{k=1}^n \left(\sup_{x \in [x_{k-1}, x_k]} f(x) \right) (x_k - x_{k-1}) = 5 \sum_{k=1}^n (x_k - x_{k-1}) = 5(11 - 7) = 20. \end{aligned}$$

We see that for this example the lower and the upper sum have each the same value for all partitions, and the value of the lower sum and the value of the upper sum differ. \square

Example 2.12 (bounded function with discontinuity)

Let $f : [1, 2] \rightarrow \mathbb{R}$ be defined by

$$f(x) := \begin{cases} 0 & \text{if } x \neq \sqrt{2}, \\ 1 & \text{if } x = \sqrt{2}. \end{cases}$$

Let P_n , $n \in \mathbb{N}$, be the partition of $[1, 2]$ into n intervals of equal length, that is,

$$P_n = \left\{ 1, 1 + \frac{1}{n}, \dots, 1 + \frac{n-1}{n}, 2 \right\} = \left\{ 1 + \frac{k}{n} : k = 0, 1, \dots, n \right\},$$

and $w(P_n) = 1/n$. Find $L(f, P_n)$ and $U(f, P_n)$.

Solution: Since $f(x) = 0$ for all $x \neq \sqrt{2}$, it is clear that

$$L(f, P_n) = \sum_{k=1}^n \left(\inf_{x \in [1 + \frac{k-1}{n}, 1 + \frac{k}{n}]} f(x) \right) \frac{1}{n} = \sum_{k=1}^n 0 \cdot \frac{1}{n} = 0.$$

Since $\sqrt{2}$ is irrational, there exists exactly one $j \in \{1, 2, \dots, n\}$ such that we have $\sqrt{2} \in [1 + \frac{j-1}{n}, 1 + \frac{j}{n}]$. Therefore

$$\begin{aligned} U(f, P_n) &= \sum_{k=1}^n \left(\sup_{x \in [1 + \frac{k-1}{n}, 1 + \frac{k}{n}]} f(x) \right) \frac{1}{n} \\ &= \left(\sup_{x \in [1 + \frac{j-1}{n}, 1 + \frac{j}{n}]} f(x) \right) \frac{1}{n} + \sum_{\substack{k=1, \\ k \neq j}}^n \left(\sup_{x \in [1 + \frac{k-1}{n}, 1 + \frac{k}{n}]} f(x) \right) \frac{1}{n} \end{aligned}$$

$$= \frac{1}{n} + 0 = \frac{1}{n}.$$

We observe that for $n \rightarrow \infty$, we find that $U(f, P_n) \rightarrow 0$ which is the value of any lower sum $L(f, P_n)$. \square

Example 2.13 (piecewise constant function)

Let $f : [0, 2] \rightarrow \mathbb{R}$ be defined by

$$f(x) := \begin{cases} 0 & \text{if } 0 \leq x \leq 1, \\ 7 & \text{if } 1 < x \leq 2. \end{cases}$$

For $n \in \mathbb{N}$, consider the partition P_{2n} of $[0, 2]$ into $2n$ subintervals of equal length, given by

$$P_{2n} := \left\{ 0, \frac{1}{n}, \dots, \frac{2n-1}{n}, 2 \right\} = \left\{ \frac{k}{n} : k = 0, 1, \dots, 2n \right\}.$$

Find $L(f, P_{2n})$ and $U(f, P_{2n})$.

Solution: It is relatively easy to see that

$$\inf_{x \in [\frac{k-1}{n}, \frac{k}{n}]} f(x) = \begin{cases} 0 & \text{if } 1 \leq k \leq n+1, \\ 7 & \text{if } n+2 \leq k \leq 2n, \end{cases}$$

and

$$\sup_{x \in [\frac{k-1}{n}, \frac{k}{n}]} f(x) = \begin{cases} 0 & \text{if } 1 \leq k \leq n, \\ 7 & \text{if } n+1 \leq k \leq 2n. \end{cases}$$

Hence

$$L(f, P_{2n}) = \sum_{k=1}^{2n} \left(\inf_{x \in [\frac{k-1}{n}, \frac{k}{n}]} f(x) \right) \frac{1}{n} = \sum_{k=n+2}^{2n} 7 \frac{1}{n} = 7 \frac{1}{n} \sum_{k=n+2}^{2n} 1 = 7 \frac{n-1}{n}$$

and

$$U(f, P_{2n}) = \sum_{k=1}^{2n} \left(\sup_{x \in [\frac{k-1}{n}, \frac{k}{n}]} f(x) \right) \frac{1}{n} = \sum_{k=n+1}^{2n} 7 \frac{1}{n} = 7 \frac{1}{n} \sum_{k=n+1}^{2n} 1 = 7 \frac{1}{n} n = 7.$$

We observe that for $n \rightarrow \infty$, we have $\lim_{n \rightarrow \infty} L(f, P_{2n}) = \lim_{n \rightarrow \infty} 7 \frac{n-1}{n} = 7 = \lim_{n \rightarrow \infty} U(f, P_{2n})$. \square

Next we derive some bounds for the lower sum and the upper sum.

Lemma 2.14 (bounds on lower and upper sum I)

Let $f \in \mathcal{B}([a, b])$ and let $P \in \mathcal{P}([a, b])$. Then

$$\left(\inf_{x \in [a, b]} f(x) \right) (b - a) \leq L(f, P) \leq U(f, P) \leq \left(\sup_{x \in [a, b]} f(x) \right) (b - a). \quad (2.1)$$

Proof of Lemma 2.14: Let $P = \{x_0, x_1, x_2, \dots, x_n\}$. For each interval $[x_{k-1}, x_k]$ of the partition P , we have

$$\inf_{x \in [a, b]} f(x) \leq \inf_{x \in [x_{k-1}, x_k]} f(x) \leq \sup_{x \in [x_{k-1}, x_k]} f(x) \leq \sup_{x \in [a, b]} f(x).$$

If we multiply the inequality above with $(x_k - x_{k-1})$ and sum over $k = 1, 2, \dots, n$, then we obtain (2.1). \square

Remark 2.15 (lower and upper sum are bounded)

Lemma 2.14 implies that the sets of real numbers $\{L(f, P) : P \in \mathcal{P}([a, b])\}$ and $\{U(f, P) : P \in \mathcal{P}([a, b])\}$ are bounded.

The next lemma is technical and is needed as an aid to prove Lemma 2.17 below. It is not worth memorizing Lemma 2.16 below, but you should know Lemma 2.14 and Lemma 2.17. We will prove Lemma 2.16 at the end of this chapter.

Lemma 2.16 (bounds on lower and upper sum II)

Let $f \in \mathcal{B}([a, b])$, and let $P, Q \in \mathcal{P}([a, b])$, $P \subset Q$, and let Q have j more points than P . Let K be an upper bound for $|f|$ on $[a, b]$. Then

$$L(f, Q) \geq L(f, P) \geq L(f, Q) - 2jK w(P), \quad (2.2)$$

$$U(f, Q) \leq U(f, P) \leq U(f, Q) + 2jK w(P). \quad (2.3)$$

Lemma 2.17 (lower sum \leq upper sum)

Let $f \in \mathcal{B}([a, b])$. For any partitions P and Q in $\mathcal{P}([a, b])$ we have

$$L(f, P) \leq U(f, Q). \quad (2.4)$$

Proof of Lemma 2.17: The result is intuitive from the pictures in Figure 4. The proof can be given by considering the partition $P \cup Q$. From Lemma 2.14, we have

$$L(f, P \cup Q) \leq U(f, P \cup Q). \quad (2.5)$$

From (2.2) and (2.3) in Lemma 2.16 we obtain, since $P \cup Q$ is a refinement of both P and Q ,

$$L(f, P \cup Q) \geq L(f, P), \quad U(f, P \cup Q) \leq U(f, Q). \quad (2.6)$$

The inequalities (2.5) and (2.6) now imply (2.4). \square

2.2 Lower and Upper Riemann Integral

After these preparations we can now introduce the **lower Riemann integral** and the **upper Riemann integral** of a bounded function. We have explained before that the idea and the picture to keep in mind is that we shrink the subintervals in the partition and make them smaller and smaller and then take the limit. This is not what happens formally in the definition below, but we will see in the next chapter that the more complicated definition below does indeed imply that the intuitive idea of shrinking the width of the partition to zero is correct.

Definition 2.18 (lower and upper Riemann integral)

Let $f \in \mathcal{B}([a, b])$. We define the **lower Riemann integral** of f over $[a, b]$ by

$$\int_a^b f(x) dx := \sup \{ L(f, P) : P \in \mathcal{P}([a, b]) \}. \quad (2.7)$$

We define the **upper Riemann integral** of f over $[a, b]$ by

$$\overline{\int_a^b} f(x) dx := \inf \{ U(f, P) : P \in \mathcal{P}([a, b]) \}. \quad (2.8)$$

We can elementary show that the upper and lower Riemann integral are bounded.

Remark 2.19 (lower and upper Riemann integral are finite)

Because from Lemmata 2.14 and 2.17 for any partitions $P, Q \in \mathcal{P}([a, b])$ and for any $f \in \mathcal{B}([a, b])$

$$\left(\inf_{x \in [a, b]} f(x) \right) (b - a) \leq L(f, P) \leq U(f, Q) \leq \left(\sup_{x \in [a, b]} f(x) \right) (b - a),$$

we know that the infimum and the supremum in (2.7) and (2.8), respectively, exist and are finite.

Now we finally define the Riemann integral and say what it means if a function is Riemann integrable. In words, a bounded function $f \in \mathcal{B}([a, b])$ is **Riemann integrable** if the lower Riemann integral and the upper Riemann integral have the same value, and this common value is then the value of the Riemann integral.

Definition 2.20 (Riemann integrable function and Riemann integral)

We say that $f \in \mathcal{B}([a, b])$ is **Riemann integrable** over $[a, b]$ if

$$\int_a^b f(x) dx = \overline{\int_a^b f(x) dx},$$

and in this case the common value of $\int_a^b f(x) dx$ and $\overline{\int_a^b f(x) dx}$ will be called the **Riemann integral** of f over $[a, b]$ and is denoted by

$$\int_a^b f(x) dx := \int_a^b f(x) dx = \overline{\int_a^b f(x) dx}.$$

The **set of all Riemann integrable functions** over the interval $[a, b]$ will be denoted by $\mathcal{R}([a, b])$.

In Lemma 2.17, we have seen that for any two partitions $P, Q \in \mathcal{P}([a, b])$, we have that the lower sum $L(f, P)$ is always less than or equal to the upper sum $U(f, Q)$. This implies that the lower Riemann integral is always less than or equal to the upper Riemann integral.

Lemma 2.21 (lower Riemann integral \leq upper Riemann integral)

Let $f \in \mathcal{B}([a, b])$. Then

$$\int_a^b f(x) dx \leq \overline{\int_a^b f(x) dx}.$$

Proof of Lemma 2.21: From Lemma 2.17 we know that for any two partitions $P, Q \in \mathcal{P}([a, b])$,

$$L(f, P) \leq U(f, Q).$$

This implies that

$$\sup \{L(f, P) : P \in \mathcal{P}([a, b])\} \leq \inf \{U(f, Q) : Q \in \mathcal{P}([a, b])\}$$

which proves the statement. \square

We consider some examples.

Example 2.22 (constant functions are Riemann integrable)

Any constant function $f : [a, b] \rightarrow \mathbb{R}$, $f(x) := C$, where $C \in \mathbb{R}$ is a fixed constant, is Riemann integrable and

$$\int_a^b f(x) dx = C(b - a). \quad (2.9)$$

Proof: Since

$$\inf_{x \in [c, d]} f(x) = \sup_{x \in [c, d]} f(x) = C$$

for any subinterval $[c, d] \subset [a, b]$, we can work out that

$$L(f, P) = U(f, P) = C(b - a) \quad \text{for any partition } P \in \mathcal{P}([a, b]).$$

Thus we obtain (2.9). \square

Example 2.23 (everywhere discontinuous function)

Show that the function $f : [7, 11] \rightarrow \mathbb{R}$, defined by

$$f(x) := \begin{cases} 2 & \text{if } x \text{ is rational,} \\ 5 & \text{if } x \text{ is irrational,} \end{cases}$$

is not Riemann integrable over $[7, 11]$.

Solution: In Example 2.11, we saw that for any partition $P \in \mathcal{P}([7, 11])$ we have $L(f, P) = 8$ and $U(f, P) = 20$. Thus

$$\begin{aligned} \int_7^{11} f(x) dx &= \sup \{L(f, P) : P \in \mathcal{P}([7, 11])\} = 8 \neq \\ \int_7^{11} f(x) dx &= \inf \{U(f, P) : P \in \mathcal{P}([7, 11])\} = 20. \end{aligned}$$

Thus f is not Riemann integrable over $[7, 11]$ because the lower and upper Riemann integral do not coincide. \square

Example 2.24 (Riemann integral of bounded function with discontinuity)

The function $f : [1, 2] \rightarrow \mathbb{R}$ from Example 2.24, defined by

$$f(x) := \begin{cases} 0 & \text{if } x \neq \sqrt{2}, \\ 1 & \text{if } x = \sqrt{2}, \end{cases}$$

is Riemann integrable and

$$\int_1^2 f(x) dx = 0.$$

We will show this in the next chapter. \square

2.3 Proof of Lemma 2.16

Finally we give the proof of the technical Lemma 2.16.

Proof of Lemma 2.16: The proof follows by induction over j . We will only explain the proof for the lower sums.

- (1) For $j = 0$ we have $P = Q$, and the statement is obvious: we have equalities.
- (2) For $j = 1$, the partition Q , has only one additional point compared to P . Let us denote $P = \{x_1, x_2, \dots, x_n\}$ and $Q = \{x_1, x_2, \dots, x_n\} \cup \{q\}$, and we assume that the point q lies in $[x_{k-1}, x_k]$. It is clear that the only contribution to the lower sums $L(f, P)$ and $L(f, Q)$ which may differ for P and Q is from the interval $[x_{k-1}, x_k]$. Let us denote the contribution to $L(f, P)$ from the interval $[x_{k-1}, x_k]$ by

$$S_1 = \left(\inf_{x \in [x_{k-1}, x_k]} f(x) \right) (x_k - x_{k-1}),$$

and let us denote the contributions to $L(f, Q)$ from the interval $[x_{k-1}, x_k]$ by

$$S_2 = \left(\inf_{x \in [x_{k-1}, q]} f(x) \right) (q - x_{k-1}), \quad S_3 = \left(\inf_{x \in [q, x_k]} f(x) \right) (x_k - q).$$

Since

$$\inf_{x \in [x_{k-1}, q]} f(x) \geq \inf_{x \in [x_{k-1}, x_k]} f(x), \quad \inf_{x \in [q, x_k]} f(x) \geq \inf_{x \in [x_{k-1}, x_k]} f(x),$$

we have $S_2 + S_3 \geq S_1$, and the first estimate in (2.2) is clear for $j = 1$.

To prove the second estimate we need to show that $S_1 \geq S_2 + S_3 - 2K w(P)$. We observe that either

$$\inf_{x \in [x_{k-1}, x_k]} f(x) = \inf_{x \in [x_{k-1}, q]} f(x)$$

or that

$$\inf_{x \in [x_{k-1}, x_k]} f(x) = \inf_{x \in [q, x_k]} f(x). \quad (2.10)$$

Without loss of generality we may assume that (2.10) holds true. Then we have

$$\begin{aligned} \inf_{x \in [x_{k-1}, x_k]} f(x) &= \inf_{x \in [x_{k-1}, q]} f(x) + \inf_{x \in [x_{k-1}, x_k]} f(x) - \inf_{x \in [x_{k-1}, q]} f(x) \\ &\geq \inf_{x \in [x_{k-1}, q]} f(x) - 2 \sup_{x \in [a, b]} |f(x)| \geq \inf_{x \in [x_{k-1}, q]} f(x) - 2K. \end{aligned} \quad (2.11)$$

Thus from (2.11) and (2.10)

$$\begin{aligned}
 S_1 &= \left(\inf_{x \in [x_{k-1}, x_k]} f(x) \right) (x_k - x_{k-1}) \\
 &= \left(\inf_{x \in [x_{k-1}, x_k]} f(x) \right) (q - x_{k-1}) + \left(\inf_{x \in [x_{k-1}, x_k]} f(x) \right) (x_k - q) \\
 &\geq \left[\left(\inf_{x \in [x_{k-1}, q]} f(x) \right) - 2K \right] (q - x_{k-1}) + \left(\inf_{x \in [q, x_k]} f(x) \right) (x_k - q) \\
 &\geq \left(\inf_{x \in [x_{k-1}, q]} f(x) \right) (q - x_{k-1}) - 2K w(P) + \left(\inf_{x \in [q, x_k]} f(x) \right) (x_k - q) \\
 &= S_2 - 2K w(P) + S_3,
 \end{aligned}$$

where we have used in the second last step the fact that $q - x_{k-1} \leq w(Q) \leq w(P)$ (since Q is a refinement of P).

(3) For $j > 1$ we use induction. Assume the statement has already been proved if Q has $j - 1$ points more than P . If Q is a refinement of P with j points more than P , then we remove an arbitrary point in $u \in Q \setminus P$ from Q and obtain a partition $\tilde{Q} := Q \setminus \{u\}$ that is a refinement of P and has $j - 1$ points more than P . For \tilde{Q} we apply (2.2). Then we consider the refinement Q of \tilde{Q} (which has one point more than \tilde{Q}) and repeat the argumentation for $j = 1$. The details of the induction step are left as an exercise. \square

Chapter 3

Darboux's Theorem, Criteria for Riemann Integrability, and Properties of the Integral

In Section 3.1, we will prove **Darboux's theorem** and will draw several conclusions from it that help to gain insight into the Riemann integral. One of those conclusions brings us back to our intuitive idea of taking a sequence of partitions whose widths tend to zero, and obtaining the lower (or upper) Riemann integral as the limit of the sequence of lower (or upper) sums as the width of the partitions tends to zero. In Section 3.2, we will prove that **continuous and monotone functions are Riemann integrable**, and in Section 3.3, we will discuss the **properties of the Riemann integral**. These properties of the Riemann integral are very useful and simplify the computation of Riemann integrals enormously. In Chapter 4, we will use these properties of the Riemann integral to derive all those statements and laws for the (Riemann) integral that you have learnt in school: integration by parts and integration by substitution, and the fundamental theorem of calculus which links integration and differentiation.

3.1 Darboux's Theorem and its Applications

We start by introducing Darboux's theorem. **Darboux's theorem** says that we can find a partition such that the lower sum (or the upper sum) get arbitrarily close to the value of the lower Riemann integral (or the upper Riemann integral).

Theorem 3.1 (Darboux's Theorem)

Let $f \in \mathcal{B}([a, b])$. Then given any $\varepsilon > 0$ there is a $\delta > 0$ such that $w(P) < \delta$ for $P \in \mathcal{P}([a, b])$ implies

$$\int_a^b f(x) dx - \varepsilon < L(f, P) \leq \int_a^b f(x) dx \quad (3.1)$$

and

$$\int_a^b f(x) dx \leq U(f, P) < \int_a^b f(x) dx + \varepsilon. \quad (3.2)$$

We can also write (3.1) and (3.2) in the following way.

Remark 3.2 (equivalent formulation of (3.1) and (3.2))

Note that (3.1) and (3.2) imply that for all $P \in \mathcal{P}([a, b])$ with $w(P) < \delta$ we have

$$0 \leq \int_a^b f(x) dx - L(f, P) < \varepsilon,$$

$$0 \leq U(f, P) - \int_a^b f(x) dx < \varepsilon.$$

Darboux's theorem implies that we can **compute the lower and upper Riemann integral by just considering the limit of the lower and upper sum, respectively, for a sequence of partitions whose width tends to zero.**

Remark 3.3 (lower and upper sum as limit)

Darboux's theorem says that $w(P) \rightarrow 0$ implies

$$L(f, P) \rightarrow \int_a^b f(x) dx \quad \text{and} \quad U(f, P) \rightarrow \int_a^b f(x) dx,$$

that is, the lower Riemann integral and the upper Riemann integral can be approached by shrinking $w(P)$.

Proof of Theorem 3.1: We prove the statement for the upper Riemann integral only, as the proof of the statement for the lower Riemann integral is similar. For any $\varepsilon > 0$, by definition of the upper integral

$$\int_a^b f(x) dx = \inf \{U(f, P) : P \in \mathcal{P}([a, b])\},$$

there is some $Q \in \mathcal{P}([a, b])$, such that

$$\overline{\int_a^b f(x) dx} \leq U(f, Q) \leq \overline{\int_a^b f(x) dx} + \frac{\varepsilon}{2}.$$

Suppose Q has j points, and let $K > 0$ be an upper bound of $|f|$ on $[a, b]$ (that is, $|f(x)| \leq K$ for all $x \in [a, b]$). Now we will find a way to choose $\delta > 0$. Suppose $P \in \mathcal{P}([a, b])$ with $w(P) < \delta$. Set $\tilde{Q} = P \cup Q$. Then \tilde{Q} has at most j points more than P . By Lemma 1.7, we have

$$U(f, P) \leq U(f, \tilde{Q}) + 2jK w(P), \quad U(f, \tilde{Q}) \leq U(f, Q), \quad (3.3)$$

and hence, using (3.3),

$$\begin{aligned} \overline{\int_a^b f(x) dx} &\leq U(f, P) \\ &\leq U(f, \tilde{Q}) + 2jK w(P) \\ &\leq U(f, Q) + 2jK w(P) \\ &< \frac{\varepsilon}{2} + 2jK \delta + \overline{\int_a^b f(x) dx} \end{aligned}$$

The above inequalities tell us how to choose $\delta > 0$ in order to prove (3.2): we need to have $2jK\delta \leq \varepsilon/2$. Thus we choose $\delta \leq \varepsilon/(4jK)$ and have for $P \in \mathcal{P}([a, b])$ with $w(P) < \delta$ that

$$\overline{\int_a^b f(x) dx} \leq U(f, P) < \frac{\varepsilon}{2} + 2jK \delta + \overline{\int_a^b f(x) dx} \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} + \overline{\int_a^b f(x) dx}$$

Thus the proof is finished. \square

Darboux's theorem allows several important conclusions that help investigating the Riemann integral. As a first application of Darboux's theorem we prove **Riemann's criterion for integrability** which says that we can find a partition such that the upper sum and the lower sum are as close together as we like.

Corollary 3.4 (Riemann's criterion for integrability)

Let $f \in \mathcal{B}([a, b])$. Then $f \in \mathcal{R}([a, b])$ **if and only if** for every $\varepsilon > 0$ there exists $P \in \mathcal{P}([a, b])$ such that

$$0 \leq U(f, P) - L(f, P) < \varepsilon.$$

Proof of Corollary 3.4: \Rightarrow : Suppose $f \in \mathcal{R}([a, b])$ and let $\varepsilon > 0$. By Darboux's Theorem, there is some $P \in \mathcal{P}([a, b])$ such that

$$\overline{\int_a^b} f(x) dx \leq U(f, P) < \overline{\int_a^b} f(x) dx + \frac{\varepsilon}{2}, \quad \underline{\int_a^b} f(x) dx \geq L(f, P) > \underline{\int_a^b} f(x) dx - \frac{\varepsilon}{2}. \quad (3.4)$$

Since $f \in \mathcal{R}([a, b])$, one has

$$\underline{\int_a^b} f(x) dx = \overline{\int_a^b} f(x) dx = \int_a^b f(x) dx,$$

and multiplying the second inequality in (3.4) by (-1) , (3.4) is therefore equivalent to

$$\int_a^b f(x) dx \leq U(f, P) < \int_a^b f(x) dx + \frac{\varepsilon}{2}, \quad (3.5)$$

$$-\int_a^b f(x) dx \leq -L(f, P) < -\int_a^b f(x) dx + \frac{\varepsilon}{2}. \quad (3.6)$$

Adding now (3.5) and (3.6), we have

$$0 \leq U(f, P) - L(f, P) < \varepsilon.$$

\Leftarrow : Now suppose that for every $\varepsilon > 0$, there is some $P \in \mathcal{P}([a, b])$ such that

$$0 \leq U(f, P) - L(f, P) < \varepsilon. \quad (3.7)$$

Since

$$L(f, P) \leq \underline{\int_a^b} f(x) dx \leq \overline{\int_a^b} f(x) dx \leq U(f, P), \quad (3.8)$$

we have from (3.7) and (3.8)

$$0 \leq \overline{\int_a^b} f(x) dx - \underline{\int_a^b} f(x) dx \leq U(f, P) - L(f, P) < \varepsilon.$$

Since ε was arbitrary, this implies that

$$\overline{\int_a^b} f(x) dx = \underline{\int_a^b} f(x) dx,$$

that is, $f \in \mathcal{R}([a, b])$. □

As already announced in Remark 3.3, Darboux's theorem allows us to take a sequence of partitions whose widths tend to zero and then compute the lower and upper Riemann integral of a bounded function as the limit of the corresponding sequence of lower and upper sums. The following lemma and theorem formalize this.

Definition 3.5 (limiting sequence of partitions (for $[a, b]$))

A sequence $\{P_n\}$ of partitions $P_n \in \mathcal{P}([a, b])$ will be called a **limiting sequence** for $[a, b]$ if $w(P_n) \rightarrow 0$ as $n \rightarrow \infty$.

We give the standard example of a limiting sequence of partitions.

Example 3.6 (equally spaced partition of $[a, b]$)

Partition of $[a, b]$ into n intervals of equal length, that is,

$$P_n := \left\{ a + k \frac{b-a}{n} : k = 0, 1, 2, \dots, n \right\}.$$

Then $w(P_n) = (b-a)/n \rightarrow 0$ as $n \rightarrow \infty$. □

Theorem 3.7 (upper and lower Riemann integral as limits)

Let $\{P_n\}$ be a limiting sequence of partitions for $[a, b]$, and let $f \in \mathcal{B}([a, b])$. Then

$$\int_a^b f(x) dx = \lim_{n \rightarrow \infty} L(f, P_n), \quad \overline{\int_a^b} f(x) dx = \lim_{n \rightarrow \infty} U(f, P_n), \quad (3.9)$$

and hence $f \in \mathcal{R}([a, b])$ **if and only if**

$$\lim_{n \rightarrow \infty} L(f, P_n) = \lim_{n \rightarrow \infty} U(f, P_n). \quad (3.10)$$

Proof of Theorem 3.7: Let us prove the first limit in (3.9): Let $\{P_n\} \subset \mathcal{P}([a, b])$ be a limiting sequence of partitions. Then we have to show that for every $\varepsilon > 0$ there exists $N = N(\varepsilon) \in \mathbb{N}$ such that

$$\left| \int_a^b f(x) dx - L(f, P_n) \right| < \varepsilon \quad \text{for all } n \geq N.$$

Let $\varepsilon > 0$ be arbitrary. From Darboux's theorem (see Theorem 3.1 and Remark 3.2) we know that there exists a $\delta > 0$ such that for all partitions $P \in \mathcal{P}([a, b])$ that satisfy $w(P) < \delta$ we have

$$\left| \int_a^b f(x) dx - L(f, P) \right| < \varepsilon. \quad (3.11)$$

By the definition of a limiting sequence, we have that for the $\delta > 0$ chosen above there exists $N \in \mathbb{N}$ such that for all $n \geq N$ we have $w(P_n) < \delta$. Thus for $n \geq N$, we can apply (3.11) for $P = P_n$ and find that

$$\left| \int_a^b f(x) dx - L(f, P_n) \right| < \varepsilon \quad \text{for all } n \geq N.$$

Since $\varepsilon > 0$ was arbitrary, the first limit in (3.9) has been proved.

The proof of the second limit in (3.9) is analogous.

That $f \in \mathcal{R}([a, b])$ if and only if (3.10) holds true follows immediately from (3.9) and the definition of the Riemann integral. \square

We give some examples where we use a sequence of equally spaced partitions and Theorem 3.7 to compute the lower and upper Riemann integral of a bounded function on an interval $[a, b]$.

Example 3.8 (bounded function with discontinuity)

As in Example 2.12, let $f : [1, 2] \rightarrow \mathbb{R}$ given by

$$f(x) := \begin{cases} 0 & \text{if } x \neq \sqrt{2}, \\ 1 & \text{if } x = \sqrt{2}, \end{cases}$$

and for every $n \in \mathbb{N}$ let

$$P_n := \left\{ 1, 1 + \frac{1}{n}, \dots, 1 + \frac{n-1}{n}, 2 \right\} = \left\{ 1 + k \frac{1}{n} : k = 0, 1, 2, \dots, n \right\}.$$

Show that the sequence $\{P_n\}$ is a limiting sequence. Use the result of Example 2.12 and Theorem 3.7 above to show that f is Riemann integrable in $[1, 2]$ and that

$$\int_1^2 f(x) dx = 0.$$

Solution: From Example 2.12, $L(f, P_n) = 0$ and $U(f, P_n) = 1/n$. As $w(P_n) = 1/n \rightarrow 0$ as $n \rightarrow \infty$, P_n is a limiting sequence, we have from Theorem 3.7

$$\begin{aligned} \int_1^2 f(x) dx &= \lim_{n \rightarrow \infty} L(f, P_n) = \lim_{n \rightarrow \infty} 0 = 0, \\ \int_1^2 f(x) dx &= \lim_{n \rightarrow \infty} U(f, P_n) = \lim_{n \rightarrow \infty} \frac{1}{n} = 0. \end{aligned}$$

Hence

$$\int_1^2 f(x) dx = \overline{\int_1^2 f(x) dx} = 0.$$

This implies, by the definition of the Riemann integral, that the function f is Riemann integrable over $[1, 2]$ and that

$$\int_1^2 f(x) dx = 0. \quad \square$$

Example 3.9 (piecewise constant function)

As in Example 2.13, let $f : [0, 2] \rightarrow \mathbb{R}$ be defined by

$$f(x) := \begin{cases} 0 & \text{if } 0 \leq x \leq 1, \\ 7 & \text{if } 1 < x \leq 2, \end{cases}$$

and let

$$P_{2n} := \left\{ 0, \frac{1}{n}, \dots, \frac{2n-1}{n}, 2 \right\}, - \left\{ \frac{k}{n} : k = 0, 1, 2, \dots, 2n \right\}, \quad n \in \mathbb{N}.$$

Find the lower and upper Riemann integral of f over $[0, 2]$. Is the function f Riemann integrable over $[0, 2]$, and if yes, what is the value of the Riemann integral?

Solution: In Example 2.13, we found that $L(f, P_{2n}) = 7(n-1)/n$ and $U(f, P_{2n}) = 7$. Since $w(P_{2n}) = 1/n \rightarrow 0$ as $n \rightarrow \infty$, $\{P_{2n}\}$ is a limiting sequence and we have from Theorem 3.7

$$\begin{aligned} \int_0^2 f(x) dx &= \lim_{n \rightarrow \infty} L(f, P_{2n}) = \lim_{n \rightarrow \infty} 7 \frac{n-1}{n} = \lim_{n \rightarrow \infty} 7 \left(1 - \frac{1}{n} \right) = 7, \\ \overline{\int_0^2 f(x) dx} &= \lim_{n \rightarrow \infty} U(f, P_{2n}) = \lim_{n \rightarrow \infty} 7 = 7. \end{aligned}$$

Since

$$\int_0^2 f(x) dx = \overline{\int_0^2 f(x) dx} = 7,$$

the function f is Riemann integrable over $[0, 2]$, and the Riemann integral of f over $[0, 2]$ has the value

$$\int_0^2 f(x) dx = 7. \quad \square$$

Example 3.10 (integral over x^2)

Let b be a positive number. Show that $f(x) = x^2$ is Riemann integrable over $[0, b]$, and that

$$\int_0^b x^2 dx = \frac{b^3}{3}.$$

Solution: Introduce the partition

$$P_n := \left\{ 0, \frac{b}{n}, \frac{2b}{n}, \dots, \frac{(n-1)b}{n}, b \right\} = \left\{ k \frac{b}{n} : k = 0, 1, 2, \dots, n \right\}.$$

We have $w(P_n) = b/n \rightarrow 0$ as $n \rightarrow \infty$, so $\{P_n\}$ is a limiting sequence of partitions for $[0, 1]$ and we can apply Theorem 3.7. We need to compute $U(f, P_n)$ and $L(f, P_n)$.

We have

$$U(f, P_n) = \sum_{k=1}^n \left(\sup_{x \in [\frac{(k-1)b}{n}, \frac{kb}{n}]} x^2 \right) \frac{b}{n} = \sum_{k=1}^n \left(\frac{kb}{n} \right)^2 \frac{b}{n} = \frac{b^3}{n^3} \sum_{k=1}^n k^2.$$

Using

$$\sum_{k=1}^n k^2 = \frac{n(n+1)(2n+1)}{6} \quad (3.12)$$

(which can be proved by induction), we obtain

$$U(f, P_n) = \frac{b^3}{n^3} \frac{n(n+1)(2n+1)}{6} = \frac{b^3}{6} \left(2 + \frac{3}{n} + \frac{1}{n^2} \right) \rightarrow \frac{b^3}{3}, \quad n \rightarrow \infty.$$

Due to Theorem 3.7,

$$\overline{\int_0^b} x^2 dx = \frac{b^3}{3}. \quad (3.13)$$

Similarly,

$$L(f, P_n) = \sum_{k=1}^n \left(\inf_{x \in [\frac{(k-1)b}{n}, \frac{kb}{n}]} x^2 \right) \frac{b}{n} = \sum_{k=1}^n \left(\frac{(k-1)b}{n} \right)^2 \frac{b}{n} = \frac{b^3}{n^3} \sum_{k=1}^n (k-1)^2.$$

Denote $j := k - 1$, then the above formula can be rewritten with (3.12) as

$$L(f, P_n) = \frac{b^3}{n^3} \sum_{j=0}^{n-1} j^2 = \frac{b^3}{n^3} \sum_{j=1}^{n-1} j^2 = \frac{b^3}{n^3} \frac{(n-1)n(2n-1)}{6} = \frac{b^3}{6} \left(2 - \frac{3}{n} + \frac{1}{n^2} \right) \rightarrow \frac{b^3}{3}$$

as $n \rightarrow \infty$. So, due to Theorem 3.7

$$\underline{\int_0^b} x^2 dx = \frac{b^3}{3}. \quad (3.14)$$

Comparing (3.13) and (3.14) we conclude that $f(x) = x^2$ is Riemann integrable over $[0, b]$ and that the Riemann integral of x^2 over $[0, b]$ is given by

$$\int_0^b x^2 dx = \frac{b^3}{3}. \quad \square$$

The following two applications of Darboux's Theorem, namely **Simpson's rule** and the **rectangular rule** are examples of relatively basic **numerical integration rules** (or **quadrature rules**). Numerical integration rules are (as the name implies) formulas for evaluating integrals numerically/computationally, that is, we approximate the integral

$$\int_a^b f(x) dx$$

by a finite sum

$$\sum_{k=0}^m w_k f(x_k),$$

where the real numbers w_k are the **weights** and the functions f is evaluated at some points x_k inside the interval $[a, b]$, the so-called **nodes** of the numerical integration rule. For example, we could approximate

$$\int_0^1 f(x) dx$$

by the sum

$$\frac{1}{3} \sum_{k=0}^2 f(k/2) = \frac{1}{3} f(0) + \frac{1}{3} f(1/2) + \frac{1}{3} f(1),$$

and we see that for constant functions this numerical integration rule is exact, since for $f(x) := c$

$$\frac{1}{3} \sum_{k=0}^2 f(k/2) = \frac{1}{3} c + \frac{1}{3} c + \frac{1}{3} c = c = c(1 - 0) = \int_0^1 c dx.$$

For any given partition the lower sum and the upper sum give each a numerical integration rule. The example above and the lower sum and upper sum are not 'very nice' numerical integration rules, and we will now derive better rules for numerical integration.

Corollary 3.11 (Simpson's rule)

Let $f : [a, b] \rightarrow \mathbb{R}$ be Riemann integrable, and let $Q_{2n} = \{y_0, \dots, y_{2n}\}$ be the partition of $[a, b]$ given by

$$y_k := a + k \frac{(b-a)}{2n}, \quad k = 0, 1, 2, \dots, 2n.$$

Define **Simpson's rule** by

$$\begin{aligned} S(f, Q_{2n}) &:= \frac{b-a}{6n} \left(f(y_0) + 4 \sum_{k=1}^n f(y_{2k-1}) + 2 \sum_{k=1}^{n-1} f(y_{2k}) + f(y_{2n}) \right) \\ &= \frac{b-a}{6n} [f(y_0) + 4f(y_1) + 2f(y_2) + 4f(y_3) + 2f(y_4) + \dots \\ &\quad \dots + 2f(y_{2n-2}) + 4f(y_{2n-1}) + f(y_{2n})] \\ &= \frac{b-a}{6n} \sum_{k=1}^n [f(y_{2k-2}) + 4f(y_{2k-1}) + f(y_{2k})]. \end{aligned} \quad (3.15)$$

Then

$$\lim_{n \rightarrow \infty} S(f, Q_{2n}) = \int_a^b f(x) dx, \quad (3.16)$$

that is, (3.15) defines an **approximation** of the Riemann integral of f over the interval $[a, b]$.

In order to remember and understand Simpson's rule, it is best to look at the last line in (3.15): each term including the factor in front of the sum is explicitly

$$\frac{b-a}{n} \left[\frac{f\left(a + (2k-2) \frac{(b-a)}{2n}\right)}{6} + \frac{4f\left(a + (2k-1) \frac{(b-a)}{2n}\right)}{6} + \frac{f\left(a + 2k \frac{(b-a)}{2n}\right)}{6} \right],$$

and we can interpret this as the contribution from the k th interval

$$\left[a + (2k-2) \frac{(b-a)}{2n}, a + 2k \frac{(b-a)}{2n} \right].$$

This interval has length $(b-a)/n$, and we take the values of f at both endpoints each with weight $1/6$ and the value of f at the middle point of the interval with weight $4/6$.

Proof of Corollary 3.11: To prove (3.16) we introduce the equally spaced partition $P_n = \{x_0, x_1, x_2, \dots, x_n\}$, given by

$$x_k := a + k \frac{(b-a)}{n}, \quad k = 0, \dots, n.$$

Then $x_k = y_{2k}$ for all $k = 0, 1, 2, \dots, n$. Consider our function f on the subinterval $[x_{k-1}, x_k] = [y_{2k-2}, y_{2k}]$. If we can show that

$$\inf_{x \in [x_{k-1}, x_k]} f(x) \leq \frac{f(y_{2k-2}) + 4f(y_{2k-1}) + f(y_{2k})}{6} \leq \sup_{x \in [x_{k-1}, x_k]} f(x), \quad (3.17)$$

then multiplication of (3.17) with $(x_k - x_{k-1}) = (b-a)/n$ and subsequent summation over $k = 1, 2, \dots, n$ gives us

$$L(f, P_n) \leq S(f, Q_{2n}) \leq U(f, P_n). \quad (3.18)$$

Taking in (3.18) the limit for $n \rightarrow \infty$ yields (since $w(P_n) \rightarrow 0$ as $n \rightarrow \infty$)

$$\int_a^b f(x) dx \leq \lim_{n \rightarrow \infty} S(f, Q_{2n}) \leq \int_a^b f(x) dx.$$

Since $f \in \mathcal{R}[a, b]$, the upper and lower Riemann integral coincide which proves (3.16).

It remains to verify that (3.17) is true. Since $y_{2k-2}, y_{2k-1}, y_{2k} \in [x_{k-1}, x_k]$, we have that

$$\begin{aligned} \inf_{x \in [x_{k-1}, x_k]} f(x) &\leq f(y_{2k-2}) \leq \sup_{x \in [x_{k-1}, x_k]} f(x), \\ \inf_{x \in [x_{k-1}, x_k]} f(x) &\leq f(y_{2k-1}) \leq \sup_{x \in [x_{k-1}, x_k]} f(x), \\ \inf_{x \in [x_{k-1}, x_k]} f(x) &\leq f(y_{2k}) \leq \sup_{x \in [x_{k-1}, x_k]} f(x). \end{aligned}$$

We apply these estimates to to get lower and upper bounds for the middle term in (3.17) and so obtain the estimate (3.17). This completes the proof. \square

We use Simpson's rule to integrate $f(x) = x^2$ over the interval $[0, b]$.

Example 3.12 (application of Simpson's rule)

Use Simpson's Rule to show that

$$\int_0^b x^2 dx = \frac{b^3}{3}.$$

Solution: From Example 3.10, we know that $x^2 \in \mathcal{R}([0, b])$. Therefore we can apply Simpson's rule. We have

$$S(f, Q_{2n}) = \frac{b}{6n} \sum_{k=1}^n \left[\left(\frac{(2k-2)b}{2n} \right)^2 + 4 \left(\frac{(2k-1)b}{2n} \right)^2 + \left(\frac{2kb}{2n} \right)^2 \right]$$

$$\begin{aligned}
&= \frac{b^3}{24n^3} \sum_{k=1}^n [(2k-2)^2 + 4(2k-1)^2 + (2k)^2] \\
&= \frac{b^3}{24n^3} \sum_{k=1}^n [24k^2 - 24k + 8] = \frac{b^3}{3n^3} \sum_{k=1}^n [3k^3 - 3k + 1] \\
&= \frac{b^3}{3n^3} \left[\frac{3n(n+1)(2n+1)}{6} - \frac{3n(n+1)}{2} + n \right] = \frac{b^3}{3n^3} n^3 = \frac{b^3}{3},
\end{aligned}$$

where we have used (3.12) and

$$\sum_{k=1}^n k = \frac{n(n+1)}{2}. \quad (3.19)$$

Thus we find

$$\int_0^b x^2 dx = \lim_{n \rightarrow \infty} S(f, Q_{2n}) = \lim_{n \rightarrow \infty} \frac{b^3}{3} = \frac{b^3}{3}$$

as claimed. \square

Another example of a numerical integration rule is the rectangular rule or midpoint rule.

Corollary 3.13 (rectangular rule or midpoint rule)

Let $f : [a, b] \rightarrow \mathbb{R}$ be Riemann integrable and let $P_n = \{x_0, x_1, \dots, x_n\}$ be the partition of $[a, b]$ given by

$$x_k := a + k \frac{(b-a)}{n}, \quad k = 0, 1, 2, \dots, n.$$

Define the **rectangular rule** (or **midpoint rule**) by

$$R(f, P_n) := \frac{b-a}{n} \sum_{k=1}^n f\left(\frac{x_{k-1} + x_k}{2}\right) = \frac{b-a}{n} \sum_{k=1}^n f\left(a + \frac{(2k-1)(b-a)}{2n}\right).$$

The rectangular rule satisfies

$$L(f, P_n) \leq R(f, P_n) \leq U(f, P_n),$$

and thus

$$\lim_{n \rightarrow \infty} R(f, P_n) = \int_a^b f(x) dx. \quad (3.20)$$

It is easy to see why the rule is called **midpoint rule** or **rectangular rule**: we

subdivide the interval $[a, b]$ into n subintervals

$$\left[a + (k-1) \frac{b-a}{n}, a + k \frac{b-a}{n} \right], \quad k = 1, 2, \dots, n,$$

of equal length $(b-a)/n$ and on each subinterval we evaluate the function f at the **midpoint** (that is, the point halfway between the endpoints of the interval)

$$\frac{1}{2} \left[\left(a + (k-1) \frac{b-a}{n} \right) + \left(a + k \frac{b-a}{n} \right) \right] = a + \frac{(2k-1)(b-a)}{2n}.$$

Thus in the sum the k th term contributes towards the integral the area of the rectangle over the k th interval with height determined by the value of f at the midpoint.

Proof of Corollary 3.13: Since $(x_{k-1} + x_k)/2$ is the point in the middle between x_{k-1} and x_k , we have $(x_{k-1} + x_k)/2 \in [x_{k-1}, x_k]$, and thus

$$\inf_{x \in [x_{k-1}, x_k]} f(x) \leq f\left(\frac{x_{k-1} + x_k}{2}\right) \leq \sup_{x \in [x_{k-1}, x_k]} f(x). \quad (3.21)$$

Multiplication of (3.21) with $(x_k - x_{k-1}) = (b-a)/n$ and subsequent summation over $k = 1, 2, \dots, n$ yields

$$L(f, P_n) \leq R(f, P_n) \leq U(f, P_n), \quad n \in \mathbb{N}. \quad (3.22)$$

Taking in (3.22) the limit for $n \rightarrow \infty$, yields from Theorem 3.7

$$\int_a^b f(x) dx \leq \lim_{n \rightarrow \infty} R(f, P_n) \leq \overline{\int_a^b f(x) dx}, \quad (3.23)$$

and since $f \in \mathcal{R}([a, b])$ the lower and upper Riemann integral coincide and we conclude from (3.23) that (3.20) holds true. \square

Now we use the rectangular rule to integrate x^2 over $[0, b]$.

Example 3.14 (application of rectangular rule)

Use the Rectangle Rule to show that for $b > 0$

$$\int_0^b x^2 dx = \frac{b^3}{3}.$$

Solution: From Example 3.10, we know that $f(x) = x^2$ is Riemann integrable over $[0, b]$, and we can apply the rectangular rule. With the help of (3.12) and (3.19),

$$R(x^2, P_n) = \frac{b}{n} \sum_{k=1}^n \left(\frac{1}{2} \left[\frac{(k-1)b}{n} + \frac{kb}{n} \right] \right)^2 = \frac{b}{n} \sum_{k=1}^n \left(\frac{(2k-1)b}{2n} \right)^2$$

$$\begin{aligned}
&= \frac{b^3}{n^3} \sum_{k=1}^n \left[k^2 - k + \frac{1}{4} \right] \\
&= \frac{b^3}{n^3} \left[\frac{n(n+1)(2n+1)}{6} - \frac{n(n+1)}{2} + \frac{n}{4} \right] = \frac{b^3}{3} \left[1 - \frac{1}{12n^2} \right].
\end{aligned}$$

Thus

$$\lim_{n \rightarrow \infty} R(x^2, P_n) = \lim_{n \rightarrow \infty} \frac{b^3}{3} \left[1 - \frac{1}{12n^2} \right] = \frac{b^3}{3}. \quad \square$$

Remark 3.15 (on Simpson's rule and the rectangular rule)

Both Simpson's rule and the rectangular rule are ***much better*** approximations of the integral than $U(f, P_n)$ and $L(f, P_n)$.

Yet another application of Darboux's Theorem tells us that bounded functions that differ only in a finite number of points have the same lower Riemann integral and the same upper Riemann integral, respectively.

Corollary 3.16 ('almost equal' functions)

Let $f, g \in \mathcal{B}([a, b])$, and suppose $f(x) \neq g(x)$ only at a finite number of points. Then

$$\int_a^b f(x) dx = \int_a^b g(x) dx, \quad \overline{\int_a^b} f(x) dx = \overline{\int_a^b} g(x) dx. \quad (3.24)$$

In particular, $f \in \mathcal{R}([a, b])$ if and only if $g \in \mathcal{R}([a, b])$.

Corollary 3.16 is very useful which is demonstrated by the next example.

Example 3.17 (integrals of 'almost equal' functions)

Consider the function $f : [1, 2] \rightarrow \mathbb{R}$, given by

$$f(x) := \begin{cases} 0 & \text{if } x \neq \sqrt{2}, \\ 1 & \text{if } x = \sqrt{2}, \end{cases}$$

from Examples 2.12 and 3.8. We have seen that in Examples 2.12 and 3.8 that $f \in \mathcal{R}([1, 2])$ and that

$$\int_1^2 f(x) dx = 0.$$

Corollary 3.16 gives a much simpler way to derive this: From Example 2.22 we know that the function $g(x) := 0$, $x \in [1, 2]$ is in $\mathcal{R}([1, 2])$, and that

$$\int_1^2 g(x) dx = 0(2 - 1) = 0.$$

We have that $f(x) = g(x)$ for all $x \in [1, 2] \setminus \{\sqrt{2}\}$. From Corollary 3.16, we know therefore that $f \in \mathcal{R}([1, 2])$ and that

$$\int_1^2 f(x) dx = \int_1^2 g(x) dx = 0. \quad \square$$

Proof of Corollary 3.16: We will only prove the equality of the lower Riemann integrals. The equality of the upper Riemann integrals follows analogously.

Let j be the number of points at which $f(x) \neq g(x)$, and let K be an upper bound for both $|f|$ and $|g|$ on $[a, b]$, that is, $\sup_{x \in [a, b]} |f(x)| \leq K$ and $\sup_{x \in [a, b]} |g(x)| \leq K$. Take a partition $P \in \mathcal{P}([a, b])$. We claim that

$$|L(f, P) - L(g, P)| \leq 4jK w(P). \quad (3.25)$$

This can be seen as follows: There are at most $2j$ subintervals on which

$$\sup_{x \in [x_{k-1}, x_k]} f(x) \neq \sup_{x \in [x_{k-1}, x_k]} g(x). \quad (3.26)$$

On any such subinterval

$$\left| \sup_{x \in [x_{k-1}, x_k]} f(x) - \sup_{x \in [x_{k-1}, x_k]} g(x) \right| \leq \left| \sup_{x \in [x_{k-1}, x_k]} f(x) \right| + \left| (-1) \sup_{x \in [x_{k-1}, x_k]} g(x) \right| \leq 2K. \quad (3.27)$$

From (3.26) and (3.27), we now obtain (3.25).

Now take a limiting sequence of partitions $\{P_n\} \subset \mathcal{P}([a, b])$. From the triangle inequality and (3.25) we see that

$$\begin{aligned} & \left| \int_a^b f(x) dx - \int_a^b g(x) dx \right| \\ & \leq \left| \int_a^b f(x) dx - L(f, P_n) \right| + |L(f, P_n) - L(g, P_n)| + \left| L(g, P_n) - \int_a^b g(x) dx \right| \\ & \leq \left| \int_a^b f(x) dx - L(f, P_n) \right| + 2jK w(P_n) + \left| L(g, P_n) - \int_a^b g(x) dx \right| \end{aligned}$$

From Corollary 3.7, we know that the first and third term tend to zero as $n \rightarrow \infty$, and the second term tends to zero because $w(P_n) \rightarrow 0$ as $n \rightarrow \infty$. Thus

$$\lim_{n \rightarrow \infty} \left| \int_a^b f(x) dx - \int_a^b g(x) dx \right| = 0$$

which implies the first equality in (3.24). \square

3.2 Continuous and Monotone Functions are Riemann Integrable

In this section we will show that continuous and monotone functions on an interval $[a, b]$ are Riemann integrable over $[a, b]$. Before we state these results more explicitly and prove them, we stop for a moment and consider why it is useful to know that a continuous or monotone function is Riemann integrable. So far, in order to verify that a function is Riemann integrable, we had to show that the lower and upper Riemann integral have the same value and this was a rather laborious and tedious process. We know from first year how to check that a function is continuous, and we know for example directly that the trigonometric functions sine and cosine, all polynomials, and the exponential function are continuous and hence are all Riemann integrable over any interval $[a, b]$. Once we will have learnt more techniques for computing Riemann integrals this knowledge will make it much easier to actually compute Riemann integrals. (**Note:** in order to apply any of the techniques for computing Riemann integrals you first have to know that the function is Riemann integrable!)

Notation: Let $\langle c, d \rangle$ denote any type of interval, that is, either (c, d) , $[c, d]$, $[c, d)$, or $(c, d]$, where $-\infty \leq c < d \leq +\infty$. Then $C(\langle c, d \rangle)$ denotes the **set of all continuous functions** on the interval $\langle c, d \rangle$.

Theorem 3.18 (continuous on $[a, b] \Rightarrow$ Riemann integrable on $[a, b]$)

Any continuous function on a closed interval $[a, b]$ is Riemann integrable over $[a, b]$, that is,

$$C([a, b]) \subset \mathcal{R}([a, b]).$$

Proof of Theorem 3.18: We prove the theorem by contradiction. Suppose that the claim is not true, then there exists some $f \in C([a, b])$ such that $f \notin \mathcal{R}([a, b])$. Riemann's criterion of integrability (see Corollary 3.4) implies that there exists some $\varepsilon_0 > 0$ such that for all $P = \{x_0, x_1, \dots, x_n\}$ in $\mathcal{P}([a, b])$

$$U(f, P) - L(f, P) \geq \varepsilon_0. \quad (3.28)$$

Hence with the notation

$$m_k(f) := \inf_{x \in [x_{k-1}, x_k]} f(x), \quad M_k(f) := \sup_{x \in [x_{k-1}, x_k]} f(x),$$

the estimate (3.28) reads

$$\sum_{k=1}^n [M_k(f) - m_k(f)](x_k - x_{k-1}) \geq \varepsilon_0. \quad (3.29)$$

There exists an index k_0 such that

$$M_{k_0}(f) - m_{k_0}(f) = \max_{1 \leq k \leq n} [M_k(f) - m_k(f)].$$

Then (3.29) can be estimated from above as follows

$$\begin{aligned} \varepsilon_0 &\leq \sum_{k=1}^n [M_k(f) - m_k(f)](x_k - x_{k-1}) \\ &\leq [M_{k_0}(f) - m_{k_0}(f)] \sum_{k=1}^n (x_k - x_{k-1}) \\ &= [M_{k_0}(f) - m_{k_0}(f)](b - a). \end{aligned}$$

Hence

$$M_{k_0}(f) - m_{k_0}(f) \geq \frac{\varepsilon_0}{b - a}. \quad (3.30)$$

Since

$$M_{k_0}(f) = \sup_{x \in [x_{k_0-1}, x_{k_0}]} f(x), \quad m_{k_0}(f) = \inf_{x \in [x_{k_0-1}, x_{k_0}]} f(x),$$

and since f is continuous on $[x_{k_0-1}, x_{k_0}]$, f reaches its infimum and supremum on the interval $[x_{k_0-1}, x_{k_0}]$. Hence there are $y, z \in [x_{k_0-1}, x_{k_0}]$ such that $f(y) = M_{k_0}(f)$, $f(z) = m_{k_0}(f)$ and (3.30) can be written as

$$|f(y) - f(z)| \geq \frac{\varepsilon_0}{b - a}. \quad (3.31)$$

Since $y, z \in [x_{k_0}, x_{k_0-1}]$, we also have

$$|y - z| \leq x_{k_0} - x_{k_0-1} \leq w(P). \quad (3.32)$$

Taking a limiting sequence $\{P_m\} \subset \mathcal{P}([a, b])$, there are $y_m, z_m \in [a, b]$ (given as explained above), such that from (3.32)

$$|y_m - z_m| \leq w(P_m) \quad (3.33)$$

and from (3.31)

$$|f(y_m) - f(z_m)| \geq \frac{\varepsilon_0}{b - a}. \quad (3.34)$$

Since every bounded sequence in \mathbb{R} has a convergent subsequence and since $\{y_m\} \subset [a, b]$ is bounded, there is a convergent subsequence $\{y_{m_j}\}$, that is, there exists some $x_0 \in [a, b]$ such that $y_{m_j} \rightarrow x_0$ as $j \rightarrow \infty$. Furthermore,

$$|z_{m_j} - x_0| \leq |z_{m_j} - y_{m_j}| + |y_{m_j} - x_0| \rightarrow 0 \quad \text{as } j \rightarrow \infty,$$

from (3.33) and $y_{m_j} \rightarrow x_0$ as $j \rightarrow \infty$. Thus $|z_{m_j} - x_0| \rightarrow 0$ as $j \rightarrow \infty$, that is, $z_{m_j} \rightarrow x_0$ as $j \rightarrow \infty$. Since f is continuous in $[a, b]$ and $y_{m_j} \rightarrow x_0$ and $z_{m_j} \rightarrow x_0$ as $j \rightarrow \infty$, we have $f(y_{m_j}) \rightarrow f(x_0)$ and $f(z_{m_j}) \rightarrow f(x_0)$ as $j \rightarrow \infty$. Passing to the limit in (3.34), we obtain

$$0 = \lim_{j \rightarrow \infty} |f(y_{m_j}) - f(z_{m_j})| \geq \frac{\varepsilon_0}{b - a}.$$

This leads to the contradiction $\varepsilon_0/(b - a) \leq 0$. \square

We give an example how Theorem 3.18 can be applied.

Example 3.19 (Elementary examples illustrating the use of Theorem 3.18)

(a) The functions e^x , $\cos x$, $\sin x$, and all polynomials

$$p(x) := a_m x^m + a_{m-1} x^{m-1} + \dots + a_1 x + a_0$$

are continuous on \mathbb{R} . Thus they are Riemann integrable over any interval $[a, b]$.

(b) The logarithm $\ln x$ is continuous on $\mathbb{R}^+ := \{x \in \mathbb{R} : x > 0\}$. Thus the logarithm is Riemann integrable over any interval $[a, b]$ with $a > 0$. \square

Next we introduce the notion of a monotone function.

Definition 3.20 (monotone functions)

- (i) A function $f : \langle c, d \rangle \rightarrow \mathbb{R}$ is called **increasing** (or **monotonically increasing**) if $x < y$, $x, y \in \langle c, d \rangle$ implies $f(x) \leq f(y)$.
- (ii) A function $f : \langle c, d \rangle \rightarrow \mathbb{R}$ is called **decreasing** (or **monotonically decreasing**) if $x < y$, $x, y \in \langle c, d \rangle$ implies $f(x) \geq f(y)$.
- (iii) A function $f : \langle c, d \rangle \rightarrow \mathbb{R}$ is called **monotone** if (i) or (ii) holds. The **set of all monotone functions** on the interval $\langle c, d \rangle$ will be denoted by $\mathcal{M}(\langle c, d \rangle)$.

Here are some examples of monotone functions.

Example 3.21 (monotone functions)

- (a) A constant function $f : \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = C$ for all $x \in \mathbb{R}$, where $C \in \mathbb{R}$ is some constant, is monotone (both increasing and decreasing).
- (b) The exponential function e^x is increasing on any interval $\langle c, d \rangle$.
- (c) The function $\sin x$ is not monotone on $[0, \frac{3\pi}{2}]$. However, $\sin x$ is increasing on $[0, \frac{\pi}{2}]$, and $\sin x$ is decreasing on $[\frac{\pi}{2}, \frac{3\pi}{2}]$.

(d) The logarithm $\ln x$ is increasing on $(0, \infty)$. \square

Now we get the second main result of this section: every monotone function is also Riemann integrable.

Theorem 3.22 (monotone on $[a, b] \Rightarrow$ Riemann integrable on $[a, b]$)

Any monotone function on $[a, b]$ is Riemann integrable on $[a, b]$, that is,

$$\mathcal{M}([a, b]) \subset \mathcal{R}([a, b]).$$

Remark 3.23 (monotone on $[a, b] \Rightarrow$ bounded on $[a, b]$)

If f is increasing then $f(a) \leq f(x) \leq f(b)$ for all $x \in [a, b]$, and if f is decreasing then $f(a) \geq f(x) \geq f(b)$, for all $x \in [a, b]$. Thus increasing functions and decreasing functions are bounded, that is,

$$\mathcal{M}([a, b]) \subset \mathcal{B}([a, b]).$$

Proof of Theorem 3.22: We will give the proof for the case of an increasing function, the case of decreasing functions is analogous.

Let f be an increasing function. If we can show for a limiting sequence $\{P_n\} \subset \mathcal{P}([a, b])$ that $U(f, P_n) - L(f, P_n) \rightarrow 0$ as $n \rightarrow \infty$, then by Theorem 3.7, $f \in \mathcal{R}([a, b])$.

Let us divide the interval $[a, b]$ into n equal parts, that is, let us consider the partition

$$P_n = \left\{ a, a + \frac{b-a}{n}, a + \frac{2(b-a)}{n}, \dots, b \right\} = \left\{ a + \frac{k(b-a)}{n} : k = 0, 1, 2, \dots, n \right\}.$$

As f is increasing

$$\sup_{x \in [x_{k-1}, x_k]} f(x) = f(x_k), \quad \inf_{x \in [x_{k-1}, x_k]} f(x) = f(x_{k-1}).$$

Thus, we have

$$U(f, P_n) = \sum_{k=1}^n \left(\sup_{x \in [x_{k-1}, x_k]} f(x) \right) (x_k - x_{k-1}) = \sum_{k=1}^n f(x_k) \frac{b-a}{n} = \frac{b-a}{n} \sum_{k=1}^n f(x_k)$$

and

$$L(f, P_n) = \sum_{k=1}^n \left(\inf_{x \in [x_{k-1}, x_k]} f(x) \right) (x_k - x_{k-1}) = \sum_{k=1}^n f(x_{k-1}) \frac{b-a}{n} = \frac{b-a}{n} \sum_{k=1}^n f(x_{k-1}).$$

Therefore

$$\begin{aligned} U(f, P_n) - L(f, P_n) &= \frac{b-a}{n} \sum_{k=1}^n f(x_k) - \frac{b-a}{n} \sum_{k=0}^n f(x_{k-1}) \\ &= \frac{b-a}{n} (f(x_n) - f(x_0)) = \frac{b-a}{n} (f(b) - f(a)) \rightarrow 0 \quad \text{as } n \rightarrow \infty. \end{aligned}$$

By Theorem 3.7, this implies that $f \in \mathcal{R}([a, b])$. \square

Note: Monotone functions are not necessarily continuous and continuous functions are not necessarily monotone, so Theorem 3.18 and Theorem 3.22 do not follow from one another.

Example 3.24 ($\text{sign}(x)$ is Riemann integrable)

The **sign function** is defined by

$$\text{sign}(x) := \begin{cases} -1 & \text{if } x < 0, \\ 0 & \text{if } x = 0, \\ +1 & \text{if } x > 0, \end{cases}$$

is monotone on \mathbb{R} and hence Riemann integrable over any interval $[a, b]$. \square

Example 3.25 (step function)

The **step function** $f : \mathbb{R} \rightarrow \mathbb{R}$ defined by $f(x) := n$ if $x \in [n, n+1)$, $n \in \mathbb{Z}$, is monotone on \mathbb{R} and hence Riemann integrable over any interval $[a, b]$. \square

Note that the functions in Examples 3.24 and 3.25 are both not continuous, but monotone. Note that monotone functions can even have an infinite number of discontinuities as in Example 3.25 above.

Note: If a function is neither monotone and nor continuous, then it can still be Riemann integrable! So for showing that a function is not Riemann integrable, it is not enough to show that it is neither continuous nor monotone.

3.3 Properties of the Integral

In this section we will discuss the **properties of the Riemann integral**. Before we start to discuss the Riemann integral we make an excursion into linear algebra and remember what we have learnt about **vector spaces** or **linear spaces**:

Definition 3.26 (real vector space/linear space)

A set V together with two associations, the **addition** $\oplus : V \times V \rightarrow V$, $(v, w) \mapsto v \oplus w$, and the **scalar multiplication** $\odot : \mathbb{R} \times V \rightarrow V$, $(\alpha, v) \mapsto \alpha \odot v$ is called a **linear space** (or **vector space**) over \mathbb{R} if (V, \oplus, \odot) satisfies the following conditions:

(i) The (additive) **associative law** holds true, namely

$$(u \oplus v) \oplus w = u \oplus (v \oplus w) \quad \text{for all } u, v, w \in V.$$

(ii) There exists an element $\mathcal{O} \in V$ (called the **neutral element**) such that

$$\mathcal{O} \oplus v = v \oplus \mathcal{O} = v \quad \text{for all } v \in V.$$

(iii) For every $v \in V$, there exists an element $w \in V$ (called the **inverse element** to v) such that

$$v \oplus w = w \oplus v = \mathcal{O}.$$

(iv) The addition is **commutative**, that is,

$$v \oplus w = w \oplus v \quad \text{for all } v, w \in V.$$

(v) We have $1 \odot v = v$ for all $v \in V$.

(vi) If $\alpha, \beta \in \mathbb{R}$ and $v \in V$ then $(\alpha\beta) \odot v = \alpha \odot (\beta \odot v)$.

(vii) The **distributive laws** hold, that is, for all $\alpha, \beta \in \mathbb{R}$ and $v, w \in V$

$$(\alpha + \beta) \odot v = (\alpha \odot v) \oplus (\beta \odot v), \quad \alpha \odot (v \oplus w) = (\alpha \odot v) \oplus (\alpha \odot w).$$

The most natural example of the vector space/linear space is the set \mathbb{R} of real numbers with the usual addition and multiplication of real numbers. A more fancy example is the space of all functions on the interval $\langle c, d \rangle$.

Example 3.27 (functions on $\langle c, d \rangle$ as a vector space)

We observe that the real-valued functions on an interval $\langle c, d \rangle$ form a **linear space** (or **vector space**) with the pointwise addition $+$ and the pointwise scalar multiplication. That is, if we denote the set of all functions on $\langle c, d \rangle$ by

$$\mathcal{F}(\langle c, d \rangle) := \{f : \langle c, d \rangle \rightarrow \mathbb{R} \mid f \text{ is a function}\}$$

then $\mathcal{F}(\langle c, d \rangle)$ with the **addition** $+$ defined via the **pointwise addition** of two

functions $f, g \in \mathcal{F}(\langle c, d \rangle)$,

$$(f + g)(x) := f(x) + g(x), \quad x \in \langle c, d \rangle, \quad (3.35)$$

and the **scalar multiplication** of any function $f \in \mathcal{F}(\langle c, d \rangle)$ with any scalar $\alpha \in \mathbb{R}$, defined as the **pointwise multiplication**

$$(\alpha f)(x) := \alpha f(x), \quad x \in \langle c, d \rangle, \quad (3.36)$$

is a **linear space** (or **vector space**). Indeed once, we have observed that the sum $f + g$ of two functions f and g in $\mathcal{F}(\langle c, d \rangle)$ is again a function in $\mathcal{F}(\langle c, d \rangle)$ and that αf for $f \in \mathcal{F}(\langle c, d \rangle)$ and $\alpha \in \mathbb{R}$ also defines a function in $\mathcal{F}(\langle c, d \rangle)$, then the vector space properties follow directly from the properties of the real numbers. \square

Example 3.28 (continuous functions on $\langle c, d \rangle$ as a vector space)

The set $\mathcal{C}(\langle c, d \rangle)$ of continuous functions on $\langle c, d \rangle$ with the usual pointwise addition (3.35) and the pointwise scalar multiplication (3.36) is a vector space/linear space.

To check this it is crucial to check the **closure**, that is, that the $f + g \in \mathcal{C}(\langle c, d \rangle)$ for any $f, g \in \mathcal{C}(\langle c, d \rangle)$ and that $\alpha f \in \mathcal{C}(\langle c, d \rangle)$ for any $f \in \mathcal{C}(\langle c, d \rangle)$ and any $\alpha \in \mathbb{R}$. We know that this is true. The vector space properties are now easily checked.

For example, the additive associative law clearly holds because, from the additive associative law for the real numbers,

$$(f(x) + g(x)) + h(x) = f(x) + (g(x) + h(x)) \quad \text{for all } x \in \langle c, d \rangle \text{ and all } f, g, h \in \mathcal{C}(\langle c, d \rangle).$$

Thus we see that

$$(f + g) + h = f + (g + h) \quad \text{for all } f, g, h \in \mathcal{C}(\langle c, d \rangle).$$

It is easily shown that the neutral element is the zero function $\mathcal{O}(x) := 0$, $x \in \langle c, d \rangle$, and that for $f \in \mathcal{C}(\langle c, d \rangle)$ the inverse element of f is given by $(-1)f \in \mathcal{C}(\langle c, d \rangle)$. We leave the details and the remaining properties as an exercise. \square

We would like to know whether the set $\mathcal{R}([a, b])$ of Riemann integrable function is also a vector space.

Notation: Let $f : \langle c, d \rangle \rightarrow \mathbb{R}$, and let $\langle c', d' \rangle \subset \langle c, d \rangle$ be a subinterval of $\langle c, d \rangle$, that is, $c' < d'$ and $c', d' \in \langle c, d \rangle$. Then $f|_{\langle c', d' \rangle}$ denotes the **restriction** of f to the set $\langle c', d' \rangle$.

The next two theorems state several important **properties** of the Riemann integral which you need to know! These properties will give us a means to establish that

functions that are 'built' from Riemann integrable functions (subject to certain rules) are Riemann integrable. These properties will also give us means to compute Riemann integrals easier and they will allow us to verify that $\mathcal{R}([a, b])$ is indeed a vector space.

Theorem 3.29 (properties of the Riemann integral I)

Linear properties:

(i) Let $f, g \in \mathcal{R}([a, b])$. Then $f + g \in \mathcal{R}([a, b])$ and

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

(ii) Let $f \in \mathcal{R}([a, b])$ and $\alpha \in \mathbb{R}$. Then $\alpha f \in \mathcal{R}([a, b])$ and

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

Domain splitting property: Let $f : [a, b] \rightarrow \mathbb{R}$ and let $c \in (a, b)$. Then $f \in \mathcal{R}([a, b])$ if and only if $f|_{[a, c]} \in \mathcal{R}([a, c])$ and $f|_{[c, b]} \in \mathcal{R}([c, b])$. Moreover, if $f \in \mathcal{R}([a, b])$ then

$$\int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx. \quad (3.37)$$

Order properties:

(i) Let $f \in \mathcal{R}([a, b])$ and $f(x) \geq 0$ for all $x \in [a, b]$. Then

$$\int_a^b f(x) dx \geq 0.$$

(ii) Let $f, g \in \mathcal{R}([a, b])$ and $f(x) \geq g(x)$ for all $x \in [a, b]$. Then

$$\int_a^b f(x) dx \geq \int_a^b g(x) dx. \quad (3.38)$$

(iii) Let $f \in \mathcal{C}([a, b])$ satisfy $f(x) \geq 0$ for all $x \in [a, b]$ and

$$\int_a^b f(x) dx = 0,$$

then $f(x) = 0$ for all $x \in [a, b]$.

Theorem 3.29 helps in determining whether a function is Riemann integrable over an interval $[a, b]$ and, if yes, it helps in computing its Riemann integral over $[a, b]$.

Example 3.30 (piecewise constant function)

The function $f : [0, 3] \rightarrow \mathbb{R}$ given by

$$f(x) := \begin{cases} 2 & \text{if } 0 \leq x \leq 1, \\ 5 & \text{if } 1 < x \leq 3, \end{cases}$$

is Riemann integrable because it is increasing. With the help of Theorem 3.29, we can now easily compute the Riemann integral of f over $[0, 3]$. From the domain splitting property

$$\int_0^3 f(x) dx = \int_0^1 f(x) dx + \int_1^3 f(x) dx.$$

Now we make use of the fact that Riemann integrable functions that differ only in a finite number of points have the same integral (see Corollary 3.16). Thus $g : [1, 3] \rightarrow \mathbb{R}$, $g(x) := 5$, has the same Riemann integral as $f|_{[1,3]}$, and we find

$$\int_0^3 f(x) dx = \int_0^1 2 dx + \int_1^3 5 dx = 2(1 - 0) + 5(3 - 1) = 12,$$

where we have used the fact that

$$\int_a^b C dx = C(b - a).$$

□

Theorem 3.31 (properties of the Riemann integral II)

Modulus property: Let $f \in \mathcal{R}([a, b])$. Then $|f| \in \mathcal{R}([a, b])$ and

$$\left| \int_a^b f(x) dx \right| \leq \int_a^b |f(x)| dx. \quad (3.39)$$

Product property: Let $f, g \in \mathcal{R}([a, b])$. Then $fg \in \mathcal{R}([a, b])$ and **Cauchy's inequality** holds:

$$\left| \int_a^b f(x) g(x) dx \right| \leq \sqrt{\int_a^b |f(x)|^2 dx} \sqrt{\int_a^b |g(x)|^2 dx}. \quad (3.40)$$

Quotient of functions: Let f, g be in $\mathcal{R}([a, b])$, and assume that

$$\inf_{x \in [a, b]} |g(x)| = K > 0.$$

Then the quotient function f/g is also in $\mathcal{R}([a, b])$.

We give an example how Theorem 3.31 can be used.

Example 3.32 (product of functions)

Show that the function $g : [0, 3] \rightarrow \mathbb{R}$, given by

$$g(x) := \begin{cases} 2x^2 & \text{if } 0 \leq x \leq 1, \\ 5 & \text{if } 1 < x \leq 3, \end{cases}$$

is Riemann integrable over $[0, 3]$.

Solution: We observe first that the function g is the product of the function f from the last example and the function $h : [0, 3] \rightarrow \mathbb{R}$, given by

$$h(x) := \begin{cases} x^2 & \text{if } 0 \leq x \leq 1, \\ 1 & \text{if } 1 < x \leq 3. \end{cases}$$

We know from Example 3.10 that $h|_{[0,1]}(x) = x^2 \in \mathcal{R}([0,1])$, and the function $\kappa(x) = 1$, $x \in [1, 3]$, is constant and thus in $\mathcal{R}([1, 3])$. Since $h|_{[1,3]}(x) = \kappa(x)$ for all $x \in [1, 3] \setminus \{1\}$, we know (see Corollary 3.16) that $h|_{[1,3]} \in \mathcal{R}([1, 3])$. From the domain splitting property we know therefore that $h \in \mathcal{R}([0, 3])$. Since $f \in \mathcal{R}([0, 3])$ (see Example 3.30) and since $h \in \mathcal{R}([0, 3])$, we know from Theorem 3.31 that $g = fh \in \mathcal{R}([0, 3])$. The integral of g over $[0, 3]$ can be computed with the domain splitting property, the linear properties, and Corollary 3.16.

$$\begin{aligned} \int_0^3 g(x) dx &= \int_0^1 g(x) dx + \int_1^3 g(x) dx = \int_0^1 2x^2 dx + \int_1^3 5 dx \\ &= 2 \int_0^1 x^2 dx + \int_1^3 5 dx = 2 \frac{1^3}{3} + 5(3 - 1) = \frac{2}{3} + 10 = \frac{32}{3}, \end{aligned}$$

where we have used that from Example 3.10 and Example 2.22

$$\int_0^b x^2 dx = \frac{b^3}{3} \quad \text{and} \quad \int_a^b C dx = C(b - a).$$

Clearly this is much simpler than working out the limits of the lower and upper sum for a limiting sequence of partitions. \square

Before we give the proof of all the properties in Theorems 3.29 and 3.31 (which is rather lengthy) we state some consequences of Theorems 3.29 and 3.31.

Remark 3.33 (lower limit \geq upper limit)

When the upper limit of integration is not necessarily strictly greater than the lower one, we define the integral as follows.

(1) We define

$$\int_a^a f(x) dx := 0$$

for any function which is defined at the point a (and takes a finite value at a).

(2) If $a > b$, we say that $\int_a^b f(x) dx$ exists if $\int_b^a f(x) dx$ exists and we define

$$\int_a^b f(x) dx := - \int_b^a f(x) dx.$$

Remark 3.34 (on the linear properties)

In terms of the addition of functions and the scalar multiplication with real numbers, **subtraction** is strictly speaking defined by

$$f - g := f + (-1)g, \quad f, g \in \mathcal{R}([a, b]).$$

We see that $f - g \in \mathcal{R}([a, b])$ if $f, g \in \mathcal{R}([a, b])$. The linear properties of the Riemann integral imply that for all $f, g \in \mathcal{R}([a, b])$ we then have

$$\int_a^b [f(x) - g(x)] dx = \int_a^b f(x) dx + (-1) \int_a^b g(x) dx = \int_a^b f(x) dx - \int_a^b g(x) dx, \quad (3.41)$$

as we did expect.

Remark 3.35 (on the domain splitting property)

If $f \in \mathcal{R}([a, b])$ and $c \in (a, b)$, then we have

$$\begin{aligned} \int_a^b f(x) dx - \int_a^c f(x) dx &= \int_c^b f(x) dx, \\ \int_a^b f(x) dx - \int_c^b f(x) dx &= \int_a^c f(x) dx. \end{aligned}$$

This follows from the domain splitting property:

$$\int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx,$$

and we obtain the claimed formulas by subtracting the first term and the second term on the right-hand side, respectively, from both sides of the equation.

Now we can answer the question that we have formulated at the beginning of this section, namely, whether the set $\mathcal{R}([a, b])$ is a linear space.

Lemma 3.36 ($\mathcal{R}([a, b])$ is a linear space)

The set $\mathcal{R}([a, b])$ with the usual pointwise addition (3.35) and the usual pointwise scalar multiplication (3.36) of functions is a linear space.

Now we will give the proofs of Theorems 3.29 and 3.31 and Lemma 3.36.

Proof of the first linear property: Let $f, g \in \mathcal{R}([a, b])$. Take an arbitrary partition $P = \{x_0, x_1, \dots, x_n\}$ and denote

$$m_k(f) := \inf_{x \in [x_{k-1}, x_k]} f(x), \quad M_k(f) := \sup_{x \in [x_{k-1}, x_k]} f(x), \quad (3.42)$$

$$m_k(g) := \inf_{x \in [x_{k-1}, x_k]} g(x), \quad M_k(g) := \sup_{x \in [x_{k-1}, x_k]} g(x), \quad (3.43)$$

and

$$m_k(f+g) := \inf_{x \in [x_{k-1}, x_k]} [f(x) + g(x)], \quad M_k(f+g) := \sup_{x \in [x_{k-1}, x_k]} [f(x) + g(x)]. \quad (3.44)$$

With the notation (3.42), (3.43), and (3.44), we have (from the definition of the infimum and the supremum)

$$m_k(f) + m_k(g) \leq m_k(f+g) \leq M_k(f+g) \leq M_k(f) + M_k(g). \quad (3.45)$$

Multiplying (3.45) by $(x_k - x_{k-1})$ and subsequently summing over $k = 1, 2, \dots, n$, gives

$$L(f, P) + L(g, P) \leq L(f+g, P) \leq U(f+g, P) \leq U(f, P) + U(g, P).$$

Take a limiting sequence of partitions $\{P_n\}$. Then

$$L(f, P_n) + L(g, P_n) \leq L(f+g, P_n) \leq U(f+g, P_n) \leq U(f, P_n) + U(g, P_n).$$

Letting $n \rightarrow \infty$ we conclude from Theorem 3.7 that (since $f, g \in \mathcal{R}([a, b])$)

$$\begin{aligned} \int_a^b f(x) dx + \int_a^b g(x) dx &\leq \int_a^b [f(x) + g(x)] dx \\ &\leq \overline{\int_a^b [f(x) + g(x)] dx} \\ &\leq \int_a^b f(x) dx + \int_a^b g(x) dx. \end{aligned}$$

Since the left-most expression and the right-most expression are the same, we conclude that all \leq have to be equalities. Thus

$$\int_a^b [f(x) + g(x)] dx = \overline{\int_a^b [f(x) + g(x)] dx},$$

that is, $f + g \in \mathcal{R}([a, b])$, and

$$\int_a^b f(x) dx + \int_a^b g(x) dx = \int_a^b [f(x) + g(x)] dx,$$

as claimed. \square

Proof of the second linear property: Let $f \in \mathcal{R}([a, b])$ and let $\alpha \in \mathbb{R}$. Take an arbitrary partition $P = \{x_0, x_1, \dots, x_n\}$. Then for $\alpha \geq 0$

$$\begin{aligned} m_k(\alpha f) &:= \inf_{x \in [x_{k-1}, x_k]} \alpha f(x) = \alpha \inf_{x \in [x_{k-1}, x_k]} f(x) = \alpha m_k(f), \\ M_k(\alpha f) &:= \sup_{x \in [x_{k-1}, x_k]} \alpha f(x) = \alpha \sup_{x \in [x_{k-1}, x_k]} f(x) = \alpha M_k(f), \end{aligned}$$

and for $\alpha < 0$

$$\begin{aligned} m_k(\alpha f) &:= \inf_{x \in [x_{k-1}, x_k]} \alpha f(x) = \alpha \sup_{x \in [x_{k-1}, x_k]} f(x) = \alpha M_k(f), \\ M_k(\alpha f) &:= \sup_{x \in [x_{k-1}, x_k]} \alpha f(x) = \alpha \inf_{x \in [x_{k-1}, x_k]} f(x) = \alpha m_k(f), \end{aligned}$$

where we have used the notation (3.42) and (3.43) with $g = \alpha f$. Thus we have for $\alpha \geq 0$

$$\alpha m_k(f) = m_k(\alpha f) \leq M_k(\alpha f) = \alpha M_k(f) \quad (3.46)$$

and for $\alpha < 0$

$$\alpha M_k(f) = m_k(\alpha f) \leq M_k(\alpha f) = \alpha m_k(f). \quad (3.47)$$

Multiplying (3.46) and (3.47) with $(x_k - x_{k-1})$ and subsequently summing over $k = 1, 2, \dots, n$ yields for $\alpha \geq 0$

$$\alpha L(f, P) = L(\alpha f, P) \leq U(\alpha f, P) = \alpha U(f, P) \quad (3.48)$$

and for $\alpha < 0$

$$\alpha U(f, P) = L(\alpha f, P) \leq U(\alpha f, P) = \alpha L(f, P). \quad (3.49)$$

Now we choose a limiting sequence of partitions $\{P_n\}$, substitute the P_n into (3.48) and (3.49), and take the limit for $n \rightarrow \infty$. Since $f \in \mathcal{R}([a, b])$, we have that

$$\lim_{n \rightarrow \infty} \alpha L(f, P_n) = \lim_{n \rightarrow \infty} \alpha U(f, P_n) = \alpha \int_a^b f(x) dx,$$

and therefore

$$\lim_{n \rightarrow \infty} L(\alpha f, P_n) = \lim_{n \rightarrow \infty} U(\alpha f, P_n).$$

Thus $\alpha f \in \mathcal{R}([a, b])$. Substituting a limiting sequence of partitions into (3.48) and (3.49) and taking the limit for $n \rightarrow \infty$ yields

$$\int_a^b \alpha f(x) dx = \alpha \int_a^b f(x) dx. \quad \square$$

Proof of the domain splitting property: First we will prove that $f \in \mathcal{R}([a, b])$ if and only if $f|_{[a, c]} \in \mathcal{R}([a, c])$ and $f|_{[c, b]} \in \mathcal{R}([c, b])$.

In this proof let $\{P_n\}$ denote a limiting sequence of partitions for $[a, b]$ with the property that each P_n contains the point c . Then $\{P_n \cap [a, c]\}$ and $\{P_n \cap [c, b]\}$ are limiting sequences of partitions for $[a, c]$ and $[c, b]$, respectively.

\Rightarrow : Assume that $f \in \mathcal{R}([a, b])$. From Theorem 3.7, we know that for every $\varepsilon > 0$ there exists an $N = N(\varepsilon) \in \mathbb{N}$ such that for all $n \geq N$

$$\begin{aligned} \varepsilon &> U(f, P_n) - L(f, P_n) \\ &= \sum_{k=0}^n \left[\left(\sup_{x \in [x_{k-1}, x_k]} f(x) \right) - \left(\inf_{x \in [x_{k-1}, x_k]} f(x) \right) \right] (x_k - x_{k-1}) \geq 0. \end{aligned}$$

Since all terms behind the sum are non-negative, and since

$$U(f, P_n) = U(f, P_n \cap [a, c]) + U(f, P_n \cap [c, b]), \quad (3.50)$$

$$L(f, P_n) = L(f, P_n \cap [a, c]) + L(f, P_n \cap [c, b]), \quad (3.51)$$

we see that also for all $n \geq N$

$$\begin{aligned} \varepsilon &> U(f, P_n \cap [a, c]) - L(f, P_n \cap [a, c]) \geq 0, \\ \varepsilon &> U(f, P_n \cap [c, b]) - L(f, P_n \cap [c, b]) \geq 0. \end{aligned}$$

Since $\varepsilon > 0$ we arbitrary, we see that

$$\begin{aligned} \lim_{n \rightarrow \infty} L(f, P_n \cap [a, c]) &= \lim_{n \rightarrow \infty} U(f, P_n \cap [a, c]), \\ \lim_{n \rightarrow \infty} L(f, P_n \cap [c, b]) &= \lim_{n \rightarrow \infty} U(f, P_n \cap [c, b]), \end{aligned}$$

and thus $f|_{[a, c]} \in \mathcal{R}([a, c])$ and $f|_{[c, b]} \in \mathcal{R}([c, b])$.

\Leftarrow : Now let us assume that $f|_{[a, c]} \in \mathcal{R}([a, c])$ and that $f|_{[c, b]} \in \mathcal{R}([c, b])$. Then we know from Theorem 3.7 that for every $\varepsilon > 0$, there exists an $N_1 = N_1(\varepsilon) \in \mathbb{N}$ such that for all $n \geq N_1$

$$0 \leq U(f, P_n \cap [a, c]) - L(f, P_n \cap [a, c]) < \frac{\varepsilon}{2}, \quad (3.52)$$

and there exists an $N_2 = N_2(\varepsilon) \in \mathbb{N}$ such that for all $n \geq N_2$

$$0 \leq U(f, P_n \cap [c, b]) - L(f, P_n \cap [c, b]) < \frac{\varepsilon}{2}. \quad (3.53)$$

Due to the special choice of P_n ,

$$\begin{aligned} U(f, P_n) &= U(f, P_n \cap [a, c]) + U(f, P_n \cap [c, b]), \\ L(f, P_n) &= L(f, P_n \cap [a, c]) + L(f, P_n \cap [c, b]). \end{aligned}$$

We see from adding (3.52) and (3.53) that for all $n \geq N := \max\{N_1, N_2\}$

$$0 \leq U(f, P_n) - L(f, P_n) \leq \varepsilon.$$

Thus from Theorem 3.7 (see also Corollary 3.16) $f \in \mathcal{R}([a, b])$.

To see that (3.37) is correct, we observe that if $f \in \mathcal{R}([a, b])$ then (3.50) and (3.51) are true. Taking the limit for $n \rightarrow \infty$ in either equation (and using that also $f|_{[a, c]} \in \mathcal{R}([a, c])$ and $f|_{[c, b]} \in \mathcal{R}([c, b])$) yields (3.37). \square

Proof of the first two order properties: If $f(x) \geq 0$ for all $x \in [a, b]$, then we have for any partition $P := \{x_0, x_1, \dots, x_n\}$ that

$$m_k(f) := \inf_{x \in [x_{k-1}, x_k]} f(x) \geq 0, \quad M_k(f) := \sup_{x \in [x_{k-1}, x_k]} f(x) \geq 0,$$

This implies that for any partition P , we have

$$L(f, P) \geq 0, \quad U(f, P) \geq 0.$$

Since the lower and upper sum are non-negative for any partition, the Riemann integral of $f \in \mathcal{R}([a, b])$ over $[a, b]$ is also non-negative. Thus

$$\int_a^b f(x) dx \geq 0,$$

which shows the first order property.

The second order property follows easily from the first: The assumption $f(x) \geq g(x)$ for all $x \in [a, b]$ can be rewritten as $f(x) - g(x) \geq 0$ for all $x \in [a, b]$. Now we apply the first order property to the function $f - g$ and obtain from the linear properties (see also Remark 3.34)

$$0 \leq \int_a^b [f(x) - g(x)] dx = \int_a^b f(x) dx - \int_a^b g(x) dx, \quad (3.54)$$

which implies (3.38). Note that in the second step we have used Remark 3.34. \square

Proof of the third order property: We prove this by contradiction. Suppose the claim is not true, then there is some $x_0 \in (a, b)$ such that $f(x_0) > 0$. (If such a point does not exist then $f(x) = 0$ for all $x \in (a, b)$, and by the continuity of f we see that $f(a) = 0$ and $f(b) = 0$). Now as f is continuous at x_0 , by definition, for $\varepsilon = f(x_0)/2 > 0$, there is a $\delta > 0$ such that $[x_0 - \delta, x_0 + \delta] \subset [a, b]$ and

$$|f(x) - f(x_0)| < \varepsilon = \frac{f(x_0)}{2}, \quad \text{for all } x \in [x_0 - \delta, x_0 + \delta].$$

This implies that

$$\begin{aligned} f(x) &= f(x_0) + (f(x) - f(x_0)) \geq f(x_0) - |f(x) - f(x_0)| \\ &> f(x_0) - \varepsilon = \frac{f(x_0)}{2} \quad \text{for all } x \in [x_0 - \delta, x_0 + \delta]. \end{aligned} \quad (3.55)$$

Now by our assumption that $f(x) \geq 0$ for all $x \in [a, b]$, from the domain splitting property and the first and second order property in Theorem 3.29, we have

$$\begin{aligned} 0 &= \int_a^b f(x) dx = \int_a^{x_0-\delta} f(x) dx + \int_{x_0-\delta}^{x_0+\delta} f(x) dx + \int_{x_0+\delta}^b f(x) dx \\ &\geq \int_{x_0-\delta}^{x_0+\delta} f(x) dx \geq \int_{x_0-\delta}^{x_0+\delta} \frac{f(x_0)}{2} dx = 2\delta \frac{f(x_0)}{2} = \delta f(x_0) > 0, \end{aligned}$$

where we have used (3.55) in the second step in the second line. This is a contradiction. \square

Proof of the modulus property: Let $f \in \mathcal{R}([a, b])$, and let $\{P_n\}$ be a limiting sequence of partitions $P_n := \{x_0, x_1, \dots, x_n\}$. If we can show that

$$\lim_{n \rightarrow \infty} [U(|f|, P_n) - L(|f|, P_n)] = 0 \quad (3.56)$$

then we know that $|f| \in \mathcal{R}([a, b])$. Then the estimate

$$\left| \sup_{x \in [x_{k-1}, x_k]} f(x) \right| \leq \sup_{x \in [x_{k-1}, x_k]} |f(x)|, \quad \text{for all } k = 1, 2, \dots, n,$$

implies that

$$\left| \sum_{k=1}^n \left(\sup_{x \in [x_{k-1}, x_k]} f(x) \right) (x_k - x_{k-1}) \right| \leq \sum_{k=1}^n \left| \sup_{x \in [x_{k-1}, x_k]} f(x) \right| (x_k - x_{k-1})$$

$$\leq \sum_{k=1}^n \left(\sup_{x \in [x_{k-1}, x_k]} |f(x)| \right) (x_k - x_{k-1}),$$

and thus

$$|U(f, P_n)| \leq U(|f|, P_n) \quad \text{for all } n \in \mathbb{N}. \quad (3.57)$$

Taking in (3.57) the limit for $n \rightarrow \infty$ yields the estimate (3.39).

It remains to show (3.56). We distinguish for each subinterval $[x_{k-1}, x_k]$ the following three cases:

- (1) $f(x) \geq 0$ for all $x \in [x_{k-1}, x_k]$,
- (2) $f(x) \leq 0$ for all $x \in [x_{k-1}, x_k]$, or
- (3) $f(x)$ assumes both positive and negative values on $[x_{k-1}, x_k]$.

In case (1), we have

$$0 \leq \sup_{x \in [x_{k-1}, x_k]} |f(x)| - \inf_{x \in [x_{k-1}, x_k]} |f(x)| = \sup_{x \in [x_{k-1}, x_k]} f(x) - \inf_{x \in [x_{k-1}, x_k]} f(x), \quad (3.58)$$

and in case (2) we have

$$\begin{aligned} 0 &\leq \sup_{x \in [x_{k-1}, x_k]} |f(x)| - \inf_{x \in [x_{k-1}, x_k]} |f(x)| \\ &= \left| \inf_{x \in [x_{k-1}, x_k]} f(x) \right| - \left| \sup_{x \in [x_{k-1}, x_k]} f(x) \right| \\ &= \sup_{x \in [x_{k-1}, x_k]} f(x) - \inf_{x \in [x_{k-1}, x_k]} f(x). \end{aligned} \quad (3.59)$$

In case (3), we have

$$\begin{aligned} 0 &\leq \sup_{x \in [x_{k-1}, x_k]} |f(x)| - \inf_{x \in [x_{k-1}, x_k]} |f(x)| \\ &= \max \left\{ \left| \inf_{x \in [x_{k-1}, x_k]} f(x) \right|, \left| \sup_{x \in [x_{k-1}, x_k]} f(x) \right| \right\} - \inf_{x \in [x_{k-1}, x_k]} |f(x)| \\ &= \max \left\{ (-1) \inf_{x \in [x_{k-1}, x_k]} f(x), \sup_{x \in [x_{k-1}, x_k]} f(x) \right\} - \inf_{x \in [x_{k-1}, x_k]} |f(x)| \\ &\leq \max \left\{ (-1) \inf_{x \in [x_{k-1}, x_k]} f(x), \sup_{x \in [x_{k-1}, x_k]} f(x) \right\} \\ &\leq \sup_{x \in [x_{k-1}, x_k]} f(x) - \inf_{x \in [x_{k-1}, x_k]} f(x), \end{aligned} \quad (3.60)$$

where we have used in the last step the fact that $\sup_{x \in [x_{k-1}, x_k]} f(x) \geq 0$ and $(-1) \inf_{x \in [x_{k-1}, x_k]} f(x) \geq 0$. From (3.58), (3.59), and (3.60), we see that in every case

$$0 \leq \sup_{x \in [x_{k-1}, x_k]} |f(x)| - \inf_{x \in [x_{k-1}, x_k]} |f(x)| \leq \sup_{x \in [x_{k-1}, x_k]} f(x) - \inf_{x \in [x_{k-1}, x_k]} f(x),$$

which implies that (multiply with $(x_k - x_{k-1})$ and sum over $k = 1, 2, \dots, n$)

$$0 \leq U(|f|, P_n) - L(|f|, P_n) \leq U(f, P_n) - L(f, P_n).$$

Since the right-hand side tends to zero for $n \rightarrow \infty$, we see that (3.56) holds true. This completes the proof. \square

Proof of the product property: We first show that for $f \in \mathcal{R}([a, b])$, we have that $f^2 = |f|^2$ is also in $\mathcal{R}([a, b])$. To do this let $\{P_n\}$ be a limiting sequence of partitions $P_n := \{x_0, x_1, \dots, x_n\}$. Then we have

$$\begin{aligned} 0 &\leq \sup_{x \in [x_{k-1}, x_k]} |f(x)|^2 - \inf_{x \in [x_{k-1}, x_k]} |f(x)|^2 \\ &= \left(\sup_{x \in [x_{k-1}, x_k]} |f(x)| \right)^2 - \left(\inf_{x \in [x_{k-1}, x_k]} |f(x)| \right)^2 \\ &= \left(\sup_{x \in [x_{k-1}, x_k]} |f(x)| + \inf_{x \in [x_{k-1}, x_k]} |f(x)| \right) \left(\sup_{x \in [x_{k-1}, x_k]} |f(x)| - \inf_{x \in [x_{k-1}, x_k]} |f(x)| \right) \\ &\leq \left(2 \sup_{x \in [a, b]} |f(x)| \right) \left(\sup_{x \in [x_{k-1}, x_k]} |f(x)| - \inf_{x \in [x_{k-1}, x_k]} |f(x)| \right). \end{aligned}$$

From this estimate we see that

$$0 \leq U(|f|^2, P_n) - L(|f|^2, P_n) \leq 2 \left(\sup_{x \in [a, b]} |f(x)| \right) [U(|f|, P_n) - L(|f|, P_n)].$$

Since $f \in \mathcal{R}([a, b])$, $|f|$ is also in $\mathcal{R}([a, b])$, and we know that the right-hand side tends to zero as $n \rightarrow \infty$ and thus $\lim_{n \rightarrow \infty} (U(|f|^2, P_n) - L(|f|^2, P_n)) = 0$. Thus $\lim_{n \rightarrow \infty} U(|f|^2, P_n) = \lim_{n \rightarrow \infty} L(|f|^2, P_n)$, and from Theorem 3.7 we know therefore that $f^2 = |f|^2$ is also in $\mathcal{R}([a, b])$.

Now let $f, g \in \mathcal{R}([a, b])$. Then $f + g \in \mathcal{R}([a, b])$ and $f - g \in \mathcal{R}([a, b])$. We have

$$\begin{aligned} &(f(x) + g(x))^2 - (f(x) - g(x))^2 \\ &= (|f(x)|^2 + 2f(x)g(x) + |g(x)|^2) - (|f(x)|^2 - 2f(x)g(x) + |g(x)|^2) \end{aligned}$$

$$= 4 f(x) g(x),$$

and thus

$$f(x) g(x) = \frac{1}{4} \left([f(x) + g(x)]^2 - [f(x) - g(x)]^2 \right). \quad (3.61)$$

From the first part of our proof we know that we know that $(f + g)^2 \in \mathcal{R}([a, b])$ and $(f - g)^2 \in \mathcal{R}([a, b])$ (since $f + g \in \mathcal{R}([a, b])$ and $f - g \in \mathcal{R}([a, b])$). Thus $(f + g)^2 - (f - g)^2$ is also in $\mathcal{R}([a, b])$, and the representation (3.61) implies therefore that $f g \in \mathcal{R}([a, b])$.

The Cauchy inequality follows now relatively easily. Let us first discuss the case where

$$\int_a^b |f(x)|^2 dx = 0 \quad \text{and} \quad \int_a^b |g(x)|^2 dx = 0. \quad (3.62)$$

Then the right-hand side of Cauchy's inequality is zero and we only have to show that the left-hand side is also zero. From

$$0 \leq (|f(x)| - |g(x)|)^2 = |f(x)|^2 - 2|f(x)| |g(x)| + |g(x)|^2$$

we have

$$|f(x) g(x)| = |f(x)| |g(x)| \leq \frac{1}{2} (|f(x)|^2 + |g(x)|^2)$$

Integrating over $[a, b]$ and using the linear properties and the first and second order property of the Riemann integral yields

$$\int_a^b |f(x) g(x)| dx \leq \frac{1}{2} \int_a^b |f(x)|^2 dx + \frac{1}{2} \int_a^b |g(x)|^2 dx = 0.$$

From the modulus property we then have

$$\left| \int_a^b f(x) g(x) dx \right| \leq \int_a^b |f(x) g(x)| dx \leq 0,$$

and thus

$$\left| \int_a^b f(x) g(x) dx \right| = 0.$$

Thus the Schwarz inequality holds true under the assumption (3.62).

We may from now on assume that

$$\int_a^b |f(x)|^2 dx > 0 \quad \text{or} \quad \int_a^b |g(x)|^2 dx > 0.$$

Let us assume that $\int_a^b |g(x)|^2 dx > 0$. For $f, g \in \mathcal{R}([a, b])$ and $\lambda \in \mathbb{R}$, we have that $(f - \lambda g)^2 \in \mathcal{R}([a, b])$ (from the previously shown properties of the integral), and from the first order property and the linear properties

$$\begin{aligned}
0 &\leq \int_a^b [f(x) - \lambda g(x)]^2 dx \\
&= \int_a^b [|f(x)|^2 - 2\lambda f(x)g(x) + \lambda^2 |g(x)|^2] dx \\
&= \lambda^2 \int_a^b |g(x)|^2 dx - 2\lambda \int_a^b f(x)g(x) dx + \int_a^b |f(x)|^2 dx \\
&= \int_a^b |g(x)|^2 dx \left(\lambda^2 - 2\lambda \frac{\int_a^b f(x)g(x) dx}{\int_a^b |g(x)|^2 dx} + \frac{\int_a^b |f(x)|^2 dx}{\int_a^b |g(x)|^2 dx} \right) \\
&= \int_a^b |g(x)|^2 dx \left(\lambda - \frac{\int_a^b f(x)g(x) dx}{\int_a^b |g(x)|^2 dx} \right)^2 \\
&\quad + \int_a^b |g(x)|^2 dx \left(\frac{\int_a^b |f(x)|^2 dx}{\int_a^b |g(x)|^2 dx} - \frac{\left(\int_a^b f(x)g(x) dx \right)^2}{\left(\int_a^b |g(x)|^2 dx \right)^2} \right) \\
&= \int_a^b |g(x)|^2 dx \left(\lambda - \frac{\int_a^b f(x)g(x) dx}{\int_a^b |g(x)|^2 dx} \right)^2 \\
&\quad + \frac{\int_a^b |f(x)|^2 dx \int_a^b |g(x)|^2 dx - \left(\int_a^b f(x)g(x) dx \right)^2}{\int_a^b |g(x)|^2 dx}. \tag{3.63}
\end{aligned}$$

Now we choose

$$\lambda = \frac{\int_a^b f(x)g(x) dx}{\int_a^b |g(x)|^2 dx}.$$

Then the first term on the right-hand side of (3.63) vanishes and we get, after multiplying with $\int_a^b |g(x)|^2 dx$,

$$0 \leq \int_a^b |f(x)|^2 dx \int_a^b |g(x)|^2 dx - \left(\int_a^b f(x)g(x) dx \right)^2,$$

which yields after rearranging (3.40). \square

Proof of the statement about the quotient of functions: Because we know that $fh \in \mathcal{R}([a, b])$ if $f, h \in \mathcal{R}([a, b])$, we see that it is enough to show that

$1/g \in \mathcal{R}([a, b])$. If $P := \{x_0, x_1, \dots, x_n\}$ is a partition then for any $x, y \in [x_{k-1}, x_k]$

$$\begin{aligned} \frac{1}{g(x)} - \frac{1}{g(y)} &= \frac{g(y) - g(x)}{g(x)g(y)} \leq \frac{|g(y) - g(x)|}{(\inf_{z \in [a, b]} |g(z)|)^2} \\ &\leq \frac{\sup_{u \in [x_{k-1}, x_k]} g(u) - \inf_{w \in [x_{k-1}, x_k]} g(w)}{(\inf_{z \in [a, b]} |g(z)|)^2}. \end{aligned} \quad (3.64)$$

Thus we have from (3.64)

$$\sup_{x \in [x_{k-1}, x_k]} \frac{1}{g(x)} - \inf_{y \in [x_{k-1}, x_k]} \frac{1}{g(y)} \leq \left(\inf_{z \in [a, b]} |g(z)| \right)^{-2} \left(\sup_{u \in [x_{k-1}, x_k]} g(u) - \inf_{w \in [x_{k-1}, x_k]} g(w) \right). \quad (3.65)$$

Multiplying (3.65) with $(x_k - x_{k-1})$ and subsequently summing over $k = 1, 2, \dots, n$, we see that

$$0 \leq U(1/g, P) - L(1/g, P) \leq \left(\inf_{z \in [a, b]} |g(z)| \right)^{-2} [U(g, P) - L(g, P)] \quad (3.66)$$

If we now consider a limiting sequence of partitions $\{P_n\}$, then we see that in (3.66) with $P = P_n$, the right-hand side tends to zero as $n \rightarrow \infty$ (since $g \in \mathcal{R}([a, b])$). Thus we have, from the sandwich theorem,

$$\lim_{n \rightarrow \infty} (U(1/g, P_n) - L(1/g, P_n)) = 0,$$

and due to Theorem 3.7, we know that $1/g \in \mathcal{R}([a, b])$. □

Finally we prove Lemma 3.36 and show that the set $\mathcal{R}([a, b])$ of Riemann integrable functions on $[a, b]$ with pointwise addition and pointwise scalar multiplication is indeed a vector space.

Proof of Lemma 3.36: From the linear properties, for any $f, g \in \mathcal{R}([a, b])$ and any $\alpha \in \mathbb{R}$, $f + g \in \mathcal{R}([a, b])$ and $\alpha f \in \mathcal{R}([a, b])$. Thus we have closure with respect to the addition (3.35) scalar multiplication (3.36).

(1) Because of the associative law for real numbers we have for all $f, g, h \in \mathcal{R}([a, b])$

$$(f(x) + g(x)) + h(x) = f(x) + (g(x) + h(x)) \quad \text{for all } x \in [a, b].$$

Thus

$$(f + g) + h = f + (g + h) \quad \text{for all } f, g, h \in \mathcal{R}([a, b]).$$

(2) Let $\mathcal{O}(x) := 0$, $x \in [a, b]$, be the zero function which is in $\mathcal{R}([a, b])$. Then for any $f \in \mathcal{R}([a, b])$

$$f(x) + \mathcal{O}(x) = \mathcal{O}(x) + f(x) = f(x) \quad \text{for all } x \in [a, b].$$

Thus the zero function \mathcal{O} is the neutral element, that is,

$$f + \mathcal{O} = \mathcal{O} + f = f \quad \text{for all } f \in \mathcal{R}([a, b]).$$

(3) Let $f \in \mathcal{R}([a, b])$. Define the function $g(x) := -f(x) = (-1)f(x)$, $x \in [a, b]$. Then $g \in \mathcal{R}([a, b])$ and

$$f(x) + g(x) = g(x) + f(x) = f(x) + (-f(x)) = 0 = \mathcal{O}(x) \quad \text{for all } x \in [a, b].$$

Thus $g := -f$ is the inverse element for f , that is,

$$f + g = g + f = \mathcal{O}.$$

(4) For any $f \in \mathcal{R}([a, b])$, we have $1 f(x) = f(x)$ for all $x \in [a, b]$. Thus $1 f = f$.

(5) For all $\alpha, \beta \in \mathbb{R}$ and $f \in \mathcal{R}([a, b])$, we have

$$(\alpha\beta)f(x) = \alpha\beta f(x) = \alpha(\beta f(x)) \quad \text{for all } x \in [a, b] \quad \Leftrightarrow \quad (\alpha\beta)f = \alpha(\beta f).$$

(6) Let $\alpha, \beta \in \mathbb{R}$ and $f, g \in \mathcal{R}([a, b])$. Then from the distributive law of real numbers

$$\begin{aligned} (\alpha + \beta)f(x) &= \alpha f(x) + \beta f(x) && \text{for all } x \in [a, b], \\ \alpha(f(x) + g(x)) &= \alpha f(x) + \alpha g(x) && \text{for all } x \in [a, b]. \end{aligned}$$

Thus both distributive laws hold true:

$$(\alpha + \beta)f = \alpha f + \beta f, \quad \alpha(f + g) = \alpha f + \alpha g$$

for all $\alpha, \beta \in \mathbb{R}$ and all $f, g \in \mathcal{R}([a, b])$. □

Chapter 4

Techniques and Results of Integral Calculus

This this chapter we encounter the main theorems for the Riemann integral which are of paramount importance for computing integrals. Most of these results you will have seen in some form in school.

In Section 4.1, we first extend the notion of Riemann integrability in some sense to arbitrary intervals by introducing **locally Riemann integrable functions**. (So far we have only talked about Riemann integrable functions on closed bounded intervals $[a, b]$ with $-\infty < a < b < \infty$.) In Section 4.2, we introduce the primitive of a function: a **primitive** (or **antiderivative**) of a function f is a differentiable function F such that $F' = f$. In Section 4.3, we then introduce the so-called **indefinite integral**. We will see that the indefinite integral of a continuous function defines a primitive (or antiderivative). In Section 4.4, we can then finally prove the **fundamental theorem of calculus** which links differentiation and integration and which is the ‘backbone’ of calculus. In Sections 4.5 and 4.6, we will learn **integration by parts** and **integration by substitution**, respectively. We will use integration by parts to derive another version of **Taylor’s formula** in which the remainder term is in integral form. In Section 4.7, we will learn the **integral test for the convergence of a series**.

4.1 Locally Riemann Integrable Functions

So far we have discussed whether functions $f : [a, b] \rightarrow \mathbb{R}$ are Riemann integrable on a **closed bounded interval** $[a, b]$ with $-\infty < a < b < \infty$. But what about Riemann integrability of functions defined on an open interval (c, d) or an half open interval $(c, d]$ or $[c, d)$ or even an unbounded interval, for example, $(-\infty, d]$ or (c, ∞) or $(-\infty, \infty) = \mathbb{R}$? This question leads to the notion of **locally Riemann integrable functions**.

As before $\langle c, d \rangle$ denotes any type of interval with $-\infty \leq c < d \leq \infty$. That is, $\langle c, d \rangle$ can be any of the following types of interval: (1) **bounded intervals** which are open intervals (c, d) , half-open intervals $(c, d]$, $[c, d)$, or closed intervals $[c, d]$, where $-\infty < c < d < \infty$ or (2) **unbounded intervals** $(-\infty, \infty) = \mathbb{R}$, $(-\infty, d)$, (c, ∞) , $(-\infty, d]$, or $[c, \infty)$, where c and d are now finite real numbers.

Definition 4.1 (locally Riemann integrable function)

A function $f : \langle c, d \rangle \rightarrow \mathbb{R}$ is called **locally Riemann integrable** over $\langle c, d \rangle$ if for any bounded closed subinterval $[a, b] \subset \langle c, d \rangle$ (that is, $a, b \in \langle c, d \rangle$ and $a < b$) we have $f|_{[a, b]} \in \mathcal{R}([a, b])$. The **set of all locally Riemann integrable functions** over the interval $\langle c, d \rangle$ will be denoted by $\mathcal{R}_{\text{loc}}(\langle c, d \rangle)$.

Remember that $f|_{[a, b]}$ denotes the restriction of $f : \langle c, d \rangle \rightarrow \mathbb{R}$ to the subinterval $[a, b] \subset \langle c, d \rangle$

We give some examples of locally Riemann integrable functions.

Example 4.2 (locally Riemann integrable functions)

- (a) $f : (-1, 1) \rightarrow \mathbb{R}$, $f(x) := (1 - x^2)^{-1}$, is locally Riemann integrable over $(-1, 1)$ (that is, $f \in \mathcal{R}_{\text{loc}}((-1, 1))$), because for any $a, b \in (-1, 1)$, $a < b$, the function $f|_{[a, b]}$ is in $\mathcal{C}([a, b])$, and thus, from Theorem 3.18, $f|_{[a, b]}$ is in $\mathcal{R}([a, b])$.
- (b) $f(x) := e^x$, is locally Riemann integrable over \mathbb{R} (that is, $f \in \mathcal{R}_{\text{loc}}(\mathbb{R})$), because for any $a, b \in \mathbb{R}$, $a < b$, the function $f|_{[a, b]}$ is in $\mathcal{C}([a, b])$, and thus, from Theorem 3.18, $f|_{[a, b]}$ is in $\mathcal{R}([a, b])$.
- (c) $f(x) := x^2$ is locally Riemann integrable over \mathbb{R} (that is, $f \in \mathcal{R}_{\text{loc}}(\mathbb{R})$), because for any $a, b \in \mathbb{R}$, $a < b$, the function $f|_{[a, b]}$ is in $\mathcal{C}([a, b])$, and thus, from Theorem 3.18, $f|_{[a, b]}$ is in $\mathcal{R}([a, b])$.
- (d) $\cos x$ and $\sin x$ are in $\mathcal{R}_{\text{loc}}(\mathbb{R})$. This can be seen as follows: From Theorem 3.18, they are continuous on \mathbb{R} , and hence on any bounded closed subinterval

$[a, b] \subset \mathbb{R}$. Thus $\cos x|_{[a,b]}$ and $\sin x|_{[a,b]}$ are in $\mathcal{R}([a, b])$ for all $[a, b] \subset \mathbb{R}$, and therefore $\cos x|_{[a,b]}, \sin x|_{[a,b]} \in \mathcal{R}_{\text{loc}}(\mathbb{R})$.

(e) $\ln x$, $x > 0$, is in $\mathcal{R}_{\text{loc}}((0, \infty))$. This follows from the fact that \ln is monotone on $(0, \infty)$, and from Theorem 3.22 $\ln x|_{[a,b]}$ is in $\mathcal{R}([a, b])$ for any bounded closed subinterval $[a, b] \subset (0, \infty)$. Thus $\ln x \in \mathcal{R}_{\text{loc}}((0, \infty))$.

(f) $\text{sign}(x)$ (see Example 3.24) and $f(x) := \max\{n \in \mathbb{Z} : n \leq x\}$ are monotone on \mathbb{R} . Restricted to any bounded closed subinterval $[a, b]$, they are in $\mathcal{M}([a, b])$ and hence in $\mathcal{R}_{\text{loc}}([a, b])$ (see Theorem 3.22). Thus they are locally Riemann integrable over \mathbb{R} .

The following two lemmas just formalize what we have used repeatedly in the examples above: Since continuity (or monotonicity) of a function f on an interval $\langle c, d \rangle$ implies continuity (or monotonicity) on any subinterval $[a, b] \subset \langle c, d \rangle$, we know from Theorem 3.18 (or Theorem 3.22, respectively) that all continuous (or monotone) functions on $\langle c, d \rangle$ are locally Riemann integrable on $\langle c, d \rangle$.

Lemma 4.3 (continuous on $\langle c, d \rangle \Rightarrow$ locally Riemann integrable on $\langle c, d \rangle$)

Any continuous function on $\langle c, d \rangle$ is locally Riemann integrable over $\langle c, d \rangle$, that is,

$$\mathcal{C}(\langle c, d \rangle) \subset \mathcal{R}_{\text{loc}}(\langle c, d \rangle).$$

Proof of Lemma 4.3: Let f be continuous on $\langle c, d \rangle$, and let $a, b \in \langle c, d \rangle$ with $a < b$ be arbitrary. Then $f|_{[a,b]}$ is in $\mathcal{C}([a, b])$, and thus, from Theorem 3.18, we know $f|_{[a,b]} \in \mathcal{R}([a, b])$. Thus $f \in \mathcal{R}_{\text{loc}}(\langle c, d \rangle)$. \square

Lemma 4.4 (monotone on $\langle c, d \rangle \Rightarrow$ locally Riemann integrable on $\langle c, d \rangle$)

Any monotone function on $\langle c, d \rangle$ is locally Riemann integrable over $\langle c, d \rangle$, that is,

$$\mathcal{M}(\langle c, d \rangle) \subset \mathcal{R}_{\text{loc}}(\langle c, d \rangle).$$

Proof of Lemma 4.4: Let f be monotone on $\langle c, d \rangle$, and let $a, b \in \langle c, d \rangle$ with $a < b$ be arbitrary. Then $f|_{[a,b]}$ is monotone on $[a, b]$, that is $f|_{[a,b]} \in \mathcal{M}([a, b])$, and thus, from Theorem 3.22, $f|_{[a,b]} \in \mathcal{R}([a, b])$. Thus $f \in \mathcal{R}_{\text{loc}}(\langle c, d \rangle)$. \square

What are the locally Riemann integrable functions on a **closed** interval?

Remark 4.5 ($\mathcal{R}_{\text{loc}}([c, d]) = \mathcal{R}([c, d])$)

If $\langle c, d \rangle = [c, d]$ is a closed bounded interval, then $\mathcal{R}_{\text{loc}}([c, d]) = \mathcal{R}([c, d])$.

Proof: Let $f \in \mathcal{R}_{\text{loc}}([c, d])$. Then by definition of $\mathcal{R}_{\text{loc}}([c, d])$, for any $a, b \in [c, d]$, $a < b$, we have $f|_{[a, b]} \in \mathcal{R}([a, b])$. We may choose $a = c$ and $b = d$, and thus $f \in \mathcal{R}([c, d])$. Thus we see that $\mathcal{R}_{\text{loc}}([c, d]) \subset \mathcal{R}([c, d])$.

Let $f \in \mathcal{R}([c, d])$. For any interval $[a, b]$ with $c \leq a < b \leq d$, we know from the domain splitting property that $f|_{[a, b]} \in \mathcal{R}([a, b])$. Thus $f \in \mathcal{R}_{\text{loc}}([c, d])$. \square

4.2 The Primitive of a Function

Differentiating a differentiable function g yields its derivative $g' = dg/dx$, and you have learnt rules for computing derivatives, such as the product rule, the quotient rule, and the chain rule. Given a function f we can also ask whether there is some differentiable function F such that $F' = f$. This leads us to the notion of the **primitive** or **antiderivative** of a function.

Definition 4.6 (primitive/antiderivative of a function)

Let $f : \langle c, d \rangle \rightarrow \mathbb{R}$. A **primitive** (or **antiderivative**) of f is a function $F : \langle c, d \rangle \rightarrow \mathbb{R}$ which is continuous on $\langle c, d \rangle$, differentiable on (c, d) , and satisfies $F'(x) = f(x)$ for all $x \in (c, d)$.

To find primitives we can use our knowledge about differentiation. We give some examples.

Example 4.7 (primitives)

(a) Let $f(x) := \sin x$. Then a primitive of f is $F(x) := -\cos x$, since

$$F'(x) = (-\cos x)' = \sin x = f(x).$$

(b) The function $F(x) := e^x$ is a primitive of $f(x) = e^x$, since

$$F'(x) = (e^x)' = e^x = f(x).$$

(c) The function $F(x) := e^x + 1$ is also a primitive of $f(x) := e^x$. Indeed

$$F'(x) = (e^x + 1)' = (e^x)' + (1)' = e^x + 0 = f(x).$$

(d) The function $f(x) = x^3 - 2x$ has a primitive $F(x) = x^4/4 - x^2$. Indeed, we have

$$F'(x) = \frac{d}{dx} \left(\frac{x^4}{4} - x^2 \right) = \frac{d}{dx} \left(\frac{x^4}{4} \right) - \frac{d(x^2)}{dx} = \frac{4x^3}{4} - 2x = x^3 - 2x,$$

which shows that F is a primitive of f . \square

From our knowledge about derivatives, we can easily conclude that, if a function has a primitive, then this **primitive is not unique**. This is explained in the next remark.

Remark 4.8 (comments on the primitive)

(1) We say **a** (and not the) primitive since a primitive is **not unique**. This is easily seen as follows: Let F be a primitive to f , and define $\tilde{F}(x) := F(x) + C$, $x \in \langle c, d \rangle$, where C is an arbitrary constant. From the properties of F , we find $\tilde{F} \in \mathcal{C}(\langle c, d \rangle)$, and that \tilde{F} is differentiable on (c, d) . Since the derivative of a constant functions is the zero function, we have

$$\tilde{F}'(x) = (F(x) + C)' = F'(x) + 0 = F'(x) = f(x) \quad \text{for all } x \in (c, d).$$

Thus \tilde{F} is clearly another primitive of f , and $\tilde{F} \neq F$ if $C \neq 0$.

(2) Remember that (c, d) is the interior of $\langle c, d \rangle$.

As you may expect, all primitives of a function can be obtained by taking one (arbitrarily chosen) primitive and adding constant functions.

Lemma 4.9 (description of all primitives)

If the function $f : \langle c, d \rangle \rightarrow \mathbb{R}$ has a primitive F on $\langle c, d \rangle$, then any primitive of f on $\langle c, d \rangle$ is of the form $F(x) + C$, $x \in \langle c, d \rangle$, with some constant $C \in \mathbb{R}$.

Proof of Lemma 4.9: We have to show that (1) $F(x) + C$ is a primitive, and (2) that every primitive \tilde{F} is of the form $F(x) + C$, with some constant $C \in \mathbb{R}$.

Proof of (1): From the properties of F , the function $F(x) + C$ is continuous on $\langle c, d \rangle$ and differentiable on (c, d) . Since the derivative of a constant function is zero and since $F'(x) = f(x)$ for all $x \in (c, d)$, we have

$$(F(x) + C)' = F'(x) + 0 = f(x) \quad \text{for all } x \in (c, d).$$

Thus $F(x) + C$ is a primitive of f on $\langle c, d \rangle$. This shows (1).

Proof of (2): Let \tilde{F} be any primitive of f . Then \tilde{F} is continuous on $\langle c, d \rangle$, and $\tilde{F}'(x) = f(x)$ for all $x \in (c, d)$. Thus

$$(F(x) - \tilde{F}(x))' = F'(x) - \tilde{F}'(x) = f(x) - f(x) = 0 \quad \text{for all } x \in (c, d). \quad (4.1)$$

Since the only functions whose derivatives vanish are the constant functions, we know from (4.1) that

$$F(x) - \tilde{F}(x) = C \quad \text{for all } x \in (c, d),$$

with some constant C . From the continuity of F and \tilde{F} , we conclude that

$$F(x) - \tilde{F}(x) = C \quad \text{for all } x \in \langle c, d \rangle.$$

This proves (2). \square

For more complicated functions than the ones in Example 4.7 we can not so easily guess a primitive from our knowledge of differentiation. The indefinite integral introduced in the next section allows us to formally write down a primitive of a continuous function. The link between differentiation and integration is formalized in the fundamental theorem of calculus which we will introduce in Section 4.4. Once we have learnt integration by parts and integration by substitution, we will be able to compute the indefinite integral and hence find primitives of fairly complicated functions.

4.3 The Indefinite Integral

If we treat one of the boundaries of the Riemann integral as a variable, then we obtain a new function. We call an integral with one boundary treated as a variable an **indefinite integral**.

Definition 4.10 (indefinite integral)

Let $f \in \mathcal{R}_{loc}(\langle c, d \rangle)$. An **indefinite integral** of f is a function $F : \langle c, d \rangle \rightarrow \mathbb{R}$ defined by

$$F(x) := \int_a^x f(t) dt, \quad x \in \langle c, d \rangle, \quad (4.2)$$

for some fixed $a \in \langle c, d \rangle$.

Example 4.11 (indefinite integral)

An indefinite integral of the constant function $f : \mathbb{R} \rightarrow \mathbb{R}$, $f(x) := 3$, is given by

$$\int_0^x f(t) dt = \int_0^x 3 dt = 3(x - 0) = 3x,$$

where we have made use of the knowledge that $\int_a^b C dt = C(b - a)$. \square

What happens if we change the fixed (lower) boundary a of the integral in (4.2)?

Lemma 4.12 (indefinite integrals of $f \in \mathcal{R}_{\text{loc}}(\langle c, d \rangle)$ differ by a constant)

Any two indefinite integrals of the same function $f \in \mathcal{R}_{\text{loc}}(\langle c, d \rangle)$ differ by a constant.

Proof of Lemma 4.12: Due to the domain splitting property (see also Remark 3.35), for $f \in \mathcal{R}_{\text{loc}}(\langle c, d \rangle)$ and $a, b \in \langle c, d \rangle$

$$\int_a^x f(t) dt - \int_b^x f(t) dt = \left(\int_a^b f(t) dt + \int_b^x f(t) dt \right) - \int_b^x f(t) dt = \int_a^b f(t) dt,$$

and the later integral has a fixed value, that is, is a constant. \square

The following example shows that taking an indefinite integral is a **smoothing operation**, that is, an indefinite integral of a function f is a smoother function than f itself.

Example 4.13 (integral of $\text{sign}(x)$)

Consider the function $\text{sign}(x)$, defined by

$$\text{sign}(x) := \begin{cases} -1 & \text{if } x < 0, \\ 0 & \text{if } x = 0, \\ +1 & \text{if } x > 0. \end{cases}$$

which is discontinuous. For simplicity let $a < 0$. From the domain splitting property and

$$\int_c^d C dt = C(d - c)$$

(see Example 2.22), we find that

$$\int_a^x \text{sign}(t) dt = \begin{cases} \int_a^x (-1) dt = (-1)(x - a) = a - x & \text{if } x \leq 0, \\ \int_a^0 (-1) dt + \int_0^x 1 dt = (-1)(0 - a) + (x - 0) = a + x & \text{if } x > 0. \end{cases}$$

We see that for $a < 0$

$$\int_a^x \text{sign}(t) dt = a + |x|$$

which is continuous. \square

The next theorem establishes important **properties of indefinite integrals**. In particular, the theorem states that if f is continuous on an interval $\langle c, d \rangle$ then an

indefinite integral $F(x) := \int_a^x f(t) dt$ of f , where $a \in \langle c, d \rangle$, defines a primitive of f on (c, d) . That is, we have for **continuous** $f : \langle c, d \rangle \rightarrow \mathbb{R}$ and any $a \in \langle c, d \rangle$,

$$F'(x) = \frac{d}{dx} \left(\int_a^x f(t) dt \right) = f(x), \quad x \in (c, d). \quad (4.3)$$

The formula (4.3) already links differentiation and integration, and from (4.3) it is only a small step to proving the fundamental theorem of calculus.

Theorem 4.14 (properties of the indefinite integral)

Let $f \in \mathcal{R}_{loc}(\langle c, d \rangle)$ and let $F : \langle c, d \rangle \rightarrow \mathbb{R}$ be an indefinite integral of f . Then the following statements hold true:

- (i) F is continuous on $\langle c, d \rangle$.
- (ii) F is differentiable at each point $x_0 \in (c, d)$ at which f is continuous, and at such a point $F'(x_0) = f(x_0)$.
- (iii) If f is continuous on $\langle c, d \rangle$ then F is a primitive of f .

We give an example of a the application of Theorem 4.14.

Example 4.15 (primitive of x^2)

We have seen in Example 3.10 that

$$\int_0^x t^2 dt = \frac{x^3}{3}. \quad (4.4)$$

Since $f(x) = x^2$ is continuous on \mathbb{R} , we know from (4.4) and Theorem 4.14 that $F(x) := x^3/3$ is a primitive of $f(x) = x^2$. Of course we knew this already from first year, since $(x^3/3)' = x^2$. \square

As mentioned before Theorem 4.14 will yield, with a little extra work, the fundamental theorem of calculus. As long as we ‘know’ a primitive of a function f from our knowledge about differentiation, Theorem 4.14 is not very helpful. But consider for example the function

$$f(x) := \frac{\cos x}{(\sin x)^3}, \quad x \in (0, \pi).$$

For this function it is not obvious (from our understanding of differentiation) how to find a primitive, but we do know from Theorem 4.14 that the indefinite integral

$$F(x) := \int_{\pi/2}^x \frac{\cos t}{(\sin t)^3} dt, \quad x \in (0, \pi),$$

defines a primitive of f . Once we have learnt techniques to compute complicated Riemann integrals, then Theorem 4.14 gives us a powerful tool for computing primitives.

Proof of Theorem 4.14 (i): Fix $x_0 \in \langle c, d \rangle$, we have

$$F(x) - F(x_0) = \int_a^x f(t) dt - \int_a^{x_0} f(t) dt = \int_{x_0}^x f(t) dt, \quad (4.5)$$

due to the domain splitting property (see Theorem 3.29 and Remark 3.35). We consider three cases due to the fact that the interval $\langle c, d \rangle$ is not necessarily open and if $\langle c, d \rangle$ is not open x_0 could be an endpoint.

Case 1: Suppose $x_0 \in (c, d)$, then there is a $\delta_0 > 0$ such that $I_0 := [x_0 - \delta_0, x_0 + \delta_0] \subset \langle c, d \rangle$. Now since $f \in \mathcal{R}_{loc}(\langle c, d \rangle)$, $f|_{I_0}$ is Riemann integrable on I_0 and hence $f|_{I_0}$ is bounded. Therefore, there is some $M > 0$ such that $|f(x)| \leq M$ for $x \in I_0$. From (4.5) above and the modulus property of the integral, we have for all $x \in I_0$

$$|F(x) - F(x_0)| = \left| \int_{x_0}^x f(t) dt \right| \leq \left| \int_{x_0}^x |f(t)| dt \right| \leq \left| \int_{x_0}^x M dt \right| = M|x - x_0|. \quad (4.6)$$

The continuity of F at x_0 follows from (4.6), as for any $\varepsilon > 0$, we can simply take $\delta := \min\{\delta_0, \varepsilon/(2M)\}$. Then $|x - x_0| < \delta$ implies $x \in I_0$ and from (4.6), we also have for $x \in \langle c, d \rangle$ with $|x - x_0| < \delta$

$$|F(x) - F(x_0)| \leq M|x - x_0| < M\delta \leq M \frac{\varepsilon}{2M} = \frac{\varepsilon}{2} < \varepsilon.$$

Thus F is continuous in x_0 and the proof for Case 1 is complete.

Case 2: Suppose $\langle c, d \rangle$ contains the left endpoint c , that is, $\langle c, d \rangle = [c, d)$ or $\langle c, d \rangle = [c, d]$, and assume $x_0 = c$. In this case we can find $\delta_0 > 0$ so that $I_1 := [x_0, x_0 + \delta_0] \subset \langle c, d \rangle$. Now since $f \in \mathcal{R}_{loc}(\langle c, d \rangle)$, $f|_{I_1}$ is Riemann integrable on I_1 and hence $f|_{I_1}$ is bounded, that is, there exists some $M > 0$ such that $|f(x)| \leq M$ for $x \in I_1$. From (4.5) above, we have, for $x \in I_1$,

$$|F(x) - F(x_0)| = \left| \int_{x_0}^x f(t) dt \right| \leq \left| \int_{x_0}^x |f(t)| dt \right| \leq \left| \int_{x_0}^x M dt \right| \leq M|x - x_0|. \quad (4.7)$$

The continuity of F at x_0 follows from (4.7), as for any $\varepsilon > 0$, we can simply take $\delta := \min\{\delta_0, \varepsilon/(2M)\}$. Then $|x - x_0| < \delta$ and $x \geq x_0$ implies $x \in I_1$, and, from (4.7), we have for all $x \in \langle c, d \rangle$ with $|x - x_0| < \delta$

$$|F(x) - F(x_0)| \leq M|x - x_0| < M\delta \leq M \frac{\varepsilon}{2M} = \frac{\varepsilon}{2} < \varepsilon.$$

Thus F is continuous at the end point $x_0 = c$ and the proof for Case 2 is complete.

Case 3: If $\langle c, d \rangle$ contains the right endpoint d , that is, $\langle c, d \rangle = (c, d]$ or $\langle c, d \rangle = [c, d]$, and if $x_0 = d$, then the proof proceeds analogously to the proof of Case 2. \square

Proof of Theorem 4.14 (ii): Let $x_0 \in (c, d)$ be an arbitrary fixed point at which f is continuous. By the definition of continuity, we have that for any $\varepsilon > 0$, there exists a $\delta > 0$, such that $I_\delta := (x_0 - \delta, x_0 + \delta) \subset \langle c, d \rangle$ and $|f(x) - f(x_0)| < \varepsilon/2$, whenever $x \in I_\delta$.

Now for $x \in I_\delta$, $x \neq x_0$, we have, from the domain splitting property, that

$$\begin{aligned} \left| \frac{F(x) - F(x_0)}{x - x_0} - f(x_0) \right| &= \left| \frac{\int_a^x f(t) dt - \int_a^{x_0} f(t) dt}{x - x_0} - f(x_0) \right| = \left| \frac{\int_{x_0}^x f(t) dt}{x - x_0} - f(x_0) \right| \\ &= \left| \frac{\int_{x_0}^x f(t) dt - f(x_0)(x - x_0)}{x - x_0} \right| = \left| \frac{\int_{x_0}^x f(t) dt - \int_{x_0}^x f(x_0) dt}{x - x_0} \right| \\ &\leq \frac{\left| \int_{x_0}^x |f(t) - f(x_0)| dt \right|}{|x - x_0|} \leq \frac{\left| \int_{x_0}^x \frac{\varepsilon}{2} dt \right|}{|x - x_0|} = \frac{\left| \frac{\varepsilon}{2}(x - x_0) \right|}{|x - x_0|} = \frac{\varepsilon}{2} < \varepsilon. \end{aligned}$$

Thus, from the definition of the derivative, $F'(x_0) = f(x_0)$. \square

Proof of Theorem 4.14 (iii): The proof of (iii) follows from (ii) and the definition of a primitive of f as a function F which is continuous on $\langle c, d \rangle$, differentiable on (c, d) , and satisfies $F'(x) = f(x)$ for all $x \in (c, d)$. \square

Fact: Every indefinite integral of a continuous function f is a primitive of f , but **not** every primitive of f is an indefinite integral. This is illustrated by the following example.

Example 4.16 (primitives of e^x)

Since the exponential function e^x is continuous on \mathbb{R} , according to Theorem 4.14, a primitive is given by

$$F(x) := \int_a^x e^t dt, \quad x \in \mathbb{R}, \quad (4.8)$$

with some constant $a \in \mathbb{R}$. From the order properties of the Riemann integral $F(x) := \int_a^x e^t dt \geq 0$ if $x \geq a$ and $F(x) := \int_a^x e^t dt < 0$ if $x < a$. Thus for any choice of a , the function F has positive and negative values.

Since we know that $(e^x)' = e^x$ for all $x \in \mathbb{R}$, the function $\tilde{F}(x) := e^x$, $x \in \mathbb{R}$, is another primitive of e^x . The primitive \tilde{F} has only positive values since $e^x > 0$ for all $x \in \mathbb{R}$. We see that the primitive $\tilde{F}(x) = e^x$ cannot be represented as an indefinite integral (4.8). \square

4.4 Fundamental Theorem of Calculus

Now we can finally prove the fundamental theorem of calculus.

Theorem 4.17 (fundamental theorem of calculus)

Let $f \in \mathcal{C}(\langle c, d \rangle)$ and let F be a primitive of f . Then for all $a, b \in \langle c, d \rangle$

$$\int_a^b f(x) dx = F(b) - F(a). \quad (4.9)$$

Notation: It is customary to use the abbreviated notation

$$F(x)|_a^b := F(b) - F(a).$$

Proof of Theorem 4.17: From Theorem 4.14, we know that $F_0 : \langle c, d \rangle \rightarrow \mathbb{R}$, defined by

$$F_0(x) := \int_a^x f(t) dt$$

is a primitive of f . We observe that we have

$$F_0(a) = \int_a^a f(t) dt = 0. \quad (4.10)$$

Now we consider an arbitrary primitive F . From Lemma 4.9 we know that this primitive F of f can be written in the form $F(x) = F_0(x) + C$, $x \in \langle c, d \rangle$, with some constant $C \in \mathbb{R}$. Thus we have from (4.10)

$$F(b) - F(a) = (F_0(b) + C) - (F_0(a) + C) = F_0(b) = \int_a^b f(t) dt$$

which proves (4.9). □

If we replace f by F' in (4.9), then (4.9) reads

$$\int_a^b F'(x) dx = F(b) - F(a),$$

and it becomes clear why integration is the **inverse operation** to differentiation. The fundamental theorem of calculus (see Theorem 4.17 above) forms the backbone of the calculation of integrals.

If we consider in Theorem 4.17 indefinite integrals then (4.9) reads

$$\int_a^x f(t) dt = \int_a^x F'(t) dt = F(x) - F(a),$$

for any primitive F of f .

We give some examples to illustrate the use of Theorem 4.17.

Example 4.18 (examples illustrating the use of Theorem 3.18)

(a) Since $\cos x$ is continuous on \mathbb{R} , it is in $\mathcal{R}_{\text{loc}}(\mathbb{R})$. We know that $(\sin x)' = \cos x$, for all $x \in \mathbb{R}$, that is, $\sin x$ is a primitive for $\cos x$. From the fundamental theorem of calculus (see Theorem 4.17) we have

$$\int_a^b \cos x \, dx = \sin b - \sin a.$$

(b) The integral

$$\int_0^1 e^x \cos(e^x + 1) \, dx$$

exists, because the function $e^x \cos(e^x + 1)$ is continuous on \mathbb{R} , and hence (from Theorem 3.18) we know that it is in $\mathcal{R}_{\text{loc}}(\mathbb{R})$. We observe that, from the chain rule, $(\sin(e^x + 1))' = e^x \cos(e^x + 1)$ for all $x \in \mathbb{R}$, that is, $\sin(e^x + 1)$ is a primitive of $e^x \cos(e^x + 1)$. Thus from Theorem 4.17,

$$\int_0^1 e^x \cos(e^x + 1) \, dx = \sin(e^x + 1)|_0^1 = \sin(e + 1) - \sin(2). \quad \square$$

Example 4.19 (integral representation of $\ln x$)

We know from calculus that $(\ln x)' = 1/x$, $x > 0$, that is, $\ln x$ is a primitive of $1/x$. Since $1/x$ is continuous for $x > 0$, we may define (according to Theorem 4.14) a primitive of $f(x) = 1/x$ via

$$F(x) = \int_a^x \frac{1}{t} \, dt, \quad x > 0, \quad (4.11)$$

where $a > 0$. From Lemma 4.9, we know that

$$\ln x = F(x) + C$$

with some constant C , since any two primitives differ by a constant. We would like to choose $a > 0$ such that $\ln x = F(x)$ (that is, we want to achieve $C = 0$). Thus we try to choose $a > 0$ in (4.11) suitably. Since $\ln(1) = 0$, we want in (4.11) $F(1) = 0$. If we choose $a = 1$, then

$$F(1) = \int_1^1 \frac{1}{t} \, dt = 0,$$

and thus $0 = \ln(1) = F(1) + C = 0 + C = C \Rightarrow C = 0$. We see that the logarithm has the integral representation

$$\ln x = \int_1^x \frac{1}{t} \, dt. \quad \square$$

4.5 Integration by Parts

A very useful technique for the computation of integrals is **integration by parts**. Integration by parts is the ‘inverse’ operation to the **product rule**

$$(G F)' = G' F + G F',$$

where F and G are continuously differentiable functions.

Theorem 4.20 (Integration by parts)

Consider $\langle c, d \rangle$, and let $a, b \in \langle c, d \rangle$ with $a < b$. Let $f, g \in \mathcal{C}(\langle c, d \rangle)$ and let F and G be primitives of f and g , respectively. Then

$$\int_a^b f(x) G(x) dx + \int_a^b F(x) g(x) dx = F(b) G(b) - F(a) G(a). \quad (4.12)$$

Proof of Theorem 4.20: Since F and G are primitives of f and g , respectively, we have $F' = f$ and $G' = g$, and $F G$ is continuously differentiable on $\langle c, d \rangle$. From the product rule we find that

$$(F G)' = F' G + F G' = f G + F g. \quad (4.13)$$

Since the function $(F G)'$ is continuous on $\langle c, d \rangle$ (and thus in $\mathcal{R}_{\text{loc}}(\langle c, d \rangle)$), we may take the Riemann integral over $[a, b]$ on both sides of (4.13) and obtain

$$\int_a^b (F G)'(x) dx = \int_a^b f(x) G(x) dx + \int_a^b F(x) g(x) dx. \quad (4.14)$$

Since $F G$ is the primitive of $(F G)'$, we may (due to the fundamental theorem of calculus (see Theorem 4.17)) rewrite the left-hand side in (4.14) as

$$\int_a^b (F G)'(x) dx = F(b) G(b) - F(a) G(a). \quad (4.15)$$

Combining (4.14) and (4.15) yields (4.12). □

Remark 4.21 (other variants of integration by parts)

Other notations of the integration by parts formula are

$$\int_a^b F'(x) G(x) dx + \int_a^b F(x) G'(x) dx = F(x) G(x) \Big|_a^b.$$

and the most common formulation

$$\int_a^b F'(x) G(x) dx = F(x) G(x) \Big|_a^b - \int_a^b F(x) G'(x) dx. \quad (4.16)$$

The **assumptions** are analogous to Theorem 4.17: $a, b \in \langle c, d \rangle$ with $a < b$, and F and G are continuously differentiable on $\langle c, d \rangle$.

We show how integration by parts is performed for some examples.

Example 4.22 (integral of $\ln x$)

Evaluate

$$\int_1^2 \ln x dx$$

with integration by parts.

Solution: We use the integration by parts formula (4.16) with $F(x) = x$, $G(x) = \ln x$. Then $F'(x) = 1$, $G'(x) = 1/x$, and from (4.16)

$$\int_1^2 \ln x dx = x \ln x \Big|_1^2 - \int_1^2 x \frac{1}{x} dx = 2 \ln(2) - 0 - \int_1^2 1 dx = 2 \ln(2) - x \Big|_1^2 = 2 \ln(2) - 1.$$

Thus we find that

$$\int_1^2 \ln x dx = 2 \ln(2) - 1. \quad \square$$

Example 4.23 (integral of $x \sin x$)

Evaluate the integral

$$\int_0^\pi x \sin x dx.$$

Solution: We use integration by parts (4.16) with $F(x) = -\cos x$, $G(x) = x$ and thus $F'(x) = \sin x$, $G'(x) = 1$. Then (with $\cos \pi = -1$ and $\sin 0 = \sin \pi = 0$)

$$\int_0^\pi x \sin x dx = -x \cos x \Big|_0^\pi + \int_0^\pi \cos x dx = -\pi \cos \pi + \sin x \Big|_0^\pi = \pi.$$

Thus we have

$$\int_0^\pi x \sin x dx = \pi. \quad \square$$

Example 4.24 (integral of $e^x \sin x$)

Evaluate the integral

$$I := \int_0^\pi e^x \sin x \, dx.$$

Solution: This is an example of an integral that cannot be directly evaluated. Instead, it can be ‘solved’ via a simple equation. We apply integration by parts (4.16) with $F(x) = e^x$, $G(x) = \sin x$, and thus $F'(x) = e^x$ and $G'(x) = \cos x$. Then

$$\int_0^\pi e^x \sin x \, dx = e^x \sin x \Big|_0^\pi - \int_0^\pi e^x \cos x \, dx = - \int_0^\pi e^x \cos x \, dx. \quad (4.17)$$

Now we apply integration by parts (4.16) a second time with $F(x) = e^x$, $G(x) = \cos x$, and thus $F'(x) = e^x$ and $G'(x) = -\sin x$. Then, using $\cos \pi = -1$ and $\cos 0 = 1$,

$$- \int_0^\pi e^x \cos x \, dx = -e^x \cos x \Big|_0^\pi - \int_0^\pi e^x \sin x \, dx = e^\pi + 1 - \int_0^\pi e^x \sin x \, dx. \quad (4.18)$$

Combining (4.17) and (4.18), we obtain

$$\int_0^\pi e^x \sin x \, dx = e^\pi + 1 - \int_0^\pi e^x \sin x \, dx. \quad (4.19)$$

The original integral I occurs in (4.19) again on the right-hand side with a negative sign. We add the original integral I on both sides and divide afterwards by 2 to obtain

$$I = \int_0^\pi e^x \sin x \, dx = \frac{e^\pi + 1}{2}. \quad \square$$

The trick used to evaluate the integral in the last example is common for a certain class of integrals. One performs integration by parts twice and obtains $I = C - I$, where I denotes the original integral and C is a real value obtained from the integration by parts. Now we can rearrange to determine I .

We can also use integration by parts to **find primitives**.

Example 4.25 (primitives of $\ln x$)

A primitive of $h(x) := \ln x$, $x > 0$, is given by

$$H_0(x) := \int_a^x \ln t \, dt \quad \text{with some } a > 0,$$

and we know (from Lemma 4.9) that every primitive of $h(x) = \ln x$ is of the form

$$H(x) = H_0(x) + C = \int_a^x \ln t \, dt + C, \quad x > 0,$$

where $C \in \mathbb{R}$ is a constant.

With $F(t) := t$ and $G(t) := \ln t$, and thus $F'(t) = 1$ and $G'(t) = 1/t$, we have from the integration by parts formula (4.16)

$$\begin{aligned} H_0(x) &= \int_a^x \ln t \, dt = t \ln t \Big|_a^x - \int_a^x t \frac{1}{t} \, dt = x \ln x - a \ln a - \int_a^x 1 \, dt \\ &= x \ln x - a \ln a - t \Big|_a^x = x \ln x - a \ln a - x + a = x \ln x - x + (a - a \ln a). \end{aligned}$$

Thus every primitive of $h(x) = \ln x$ is of the form

$$H(x) = x \ln x - x + (a - a \ln a) + C = x \ln x - x + \tilde{C},$$

with the new constant $\tilde{C} := a - a \ln a + C$. □

Example 4.26 (primitives of $(\sin x)^2 + 2x + 3$)

Find all primitives of the continuous function $h : \mathbb{R} \rightarrow \mathbb{R}$, given by

$$h(x) := (\sin x)^2 + 2x + 3.$$

Solution: A primitive of $h(x) = (\sin x)^2 + 2x + 3$ is given by the indefinite integral

$$H_0(x) := \int_0^x [(\sin t)^2 + 2t + 3] \, dt,$$

and any primitive of $h(x) = (\sin x)^2 + 2x + 3$ is of the form

$$H(x) = H_0(x) + C, \quad x \in \mathbb{R},$$

with some constant $C \in \mathbb{R}$.

From the linear properties of the integral, from $(x)' = 1$ and $(x^2)' = 2x$, and from the fundamental theorem of calculus (see Theorem 4.17), we find that

$$\begin{aligned} H_0(x) &= \int_0^x (\sin t)^2 \, dt + \int_0^x 2t \, dt + \int_0^x 3 \, dt \\ &= \int_0^x (\sin t)^2 \, dt + t^2 \Big|_0^x + 3t \Big|_0^x = \int_0^x (\sin t)^2 \, dt + x^2 + 3x. \end{aligned} \quad (4.20)$$

From the integration by parts formula (4.16) we have, with $F(t) := -\cos t$, $G(t) := \sin t$, and thus $F'(t) = \sin t$, $G'(t) = \cos t$,

$$\int_0^x (\sin t)^2 dt = -\cos t \sin t|_0^x + \int_0^x (\cos t)^2 dt = -\cos x \sin x + \int_0^x (\cos t)^2 dt. \quad (4.21)$$

Now we use that $(\sin t)^2 + (\cos t)^2 = 1$ and replace $(\cos t)^2 = 1 - (\sin t)^2$ in the remaining integral in (4.21). Thus

$$\begin{aligned} \int_0^x (\sin t)^2 dt &= -\cos x \sin x + \int_0^x [1 - (\sin t)^2] dt \\ &= -\cos x \sin x + \int_0^x 1 dt - \int_0^x (\sin t)^2 dt \\ &= -\cos x \sin x + t|_0^x - \int_0^x (\sin t)^2 dt \\ &= -\cos x \sin x + x - \int_0^x (\sin t)^2 dt, \end{aligned}$$

and adding the original integral on both sides yields

$$2 \int_0^x (\sin t)^2 dt = x - \cos x \sin x. \quad (4.22)$$

Thus we have from (4.20) and (4.22) that

$$H_0(x) = \frac{1}{2}(x - \cos x \sin x) + x^2 + 3x = x^2 + \frac{7}{2}x - \cos x \sin x,$$

and any primitive of $h(x) = (\sin x)^2 + 2x + 3$ is of the form

$$H(x) = x^2 + \frac{7}{2}x - \cos x \sin x + C,$$

with some constant $C \in \mathbb{R}$. □

Example 4.27 (primitives of $(\cos x)^{-2}$)

Find all primitives of the function $h : (-\frac{\pi}{2}, \frac{\pi}{2}) \rightarrow \mathbb{R}$, given by

$$h(x) := \frac{1}{(\cos x)^2}.$$

Solution: We know that all primitives of $h(x) = (\cos x)^{-2}$ are of the form

$$H(x) = \int_0^x \frac{1}{(\cos t)^2} dt + C, \quad x \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right), \quad \text{with some constant } C \in \mathbb{R}.$$

We use that $1 = (\cos x)^2 + (\sin x)^2$ to rewrite the indefinite integral as follows

$$\int_0^x \frac{1}{(\cos t)^2} dt = \int_0^x \frac{(\cos t)^2 + (\sin t)^2}{(\cos t)^2} dt = \int_0^x \left(1 + \frac{(\sin t)^2}{(\cos t)^2} \right) dt.$$

Thus from the linear properties of the integral

$$\int_0^x \frac{1}{(\cos t)^2} dt = \int_0^x 1 dt + \int_0^x \frac{(\sin t)^2}{(\cos t)^2} dt = t \Big|_0^x + \int_0^x \frac{(\sin t)^2}{(\cos t)^2} dt = x + \int_0^x \frac{(\sin t)^2}{(\cos t)^2} dt. \quad (4.23)$$

We apply the integration by parts formula (4.16) with $F(t) = 1/\cos t$, $G(t) = \sin t$, and thus $F'(t) = \sin t/(\cos t)^2$, $G'(t) = \cos t$ to the remaining integral, and obtain

$$\int_0^x \frac{(\sin t)^2}{(\cos t)^2} dt = \frac{\sin t}{\cos t} \Big|_0^x - \int_0^x \frac{1}{\cos t} \cos t dt = \frac{\sin x}{\cos x} - \int_0^x 1 dt = \frac{\sin x}{\cos x} - t \Big|_0^x = \tan x - x. \quad (4.24)$$

Substituting (4.24) into (4.23) yields

$$\int_0^x \frac{1}{(\cos t)^2} dt = x + \tan x - x = \tan x.$$

Thus all primitives of $h(x) = 1/(\cos x)^2$ are of the form

$$H(x) = \tan x + C, \quad x \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right), \quad \text{with some constant } C \in \mathbb{R}. \quad \square$$

As an application of integration by parts, we obtain another version of Taylor's formula (see Theorem 1.33) with a remainder term in form of an integral.

Theorem 4.28 (integral version of Taylor's formula)

Let $f : (a, b) \rightarrow \mathbb{R}$ be a function which is $(n+1)$ -times continuously differentiable on (a, b) , and let $x_0 \in (a, b)$. Then for any $x \in (a, b)$, $x \neq x_0$, we have

$$f(x) = f(x_0) + \sum_{k=1}^n \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k + \frac{1}{n!} \int_{x_0}^x (x - t)^n f^{(n+1)}(t) dt. \quad (4.25)$$

Note that the assumption that f is $(n+1)$ -times continuously differentiable implies that the $(n+1)$ th derivative $f^{(n+1)}$ is locally Riemann integrable over (a, b) (since it is continuous on (a, b)).

Remark 4.29 (k -times continuously differentiable)

If we say that a function is k -times **continuously differentiable** on an interval (a, b) , we mean that f and all its derivatives up to order k exist and are continuous on (a, b) . The set of k -times continuously differentiable functions on (a, b) is often denoted by $\mathcal{C}^k((a, b))$.

Proof of Theorem 4.28: Using the integration by parts formula (4.16) we first show that if $n \geq 1$ then

$$\frac{1}{(n-1)!} \int_{x_0}^x f^{(n)}(t) (x-t)^{n-1} dt = \frac{f^{(n)}(x_0)}{n!} (x-x_0)^n + \frac{1}{n!} \int_{x_0}^x f^{(n+1)}(t) (x-t)^n dt. \quad (4.26)$$

Proof of (4.26): Define

$$F(t) := -\frac{1}{n!} (x-t)^n, \quad G(t) := f^{(n)}(t),$$

and thus

$$F'(t) = \frac{1}{(n-1)!} (x-t)^{n-1}, \quad G'(t) = f^{(n+1)}(t).$$

Then by the integration by parts formula (4.16)

$$\begin{aligned} & \frac{1}{(n-1)!} \int_{x_0}^x f^{(n)}(t) (x-t)^{n-1} dt \\ &= -f^{(n)}(t) \frac{1}{n!} (x-t)^n \Big|_{x_0}^x + \frac{1}{n!} \int_{x_0}^x f^{(n+1)}(t) (x-t)^n dt \\ &= \frac{f^{(n)}(x_0)}{n!} (x-x_0)^n + \frac{1}{n!} \int_{x_0}^x f^{(n+1)}(t) (x-t)^n dt. \end{aligned}$$

This proves (4.26).

Now we can prove (4.25) by induction:

(1) In the case $n = 0$ we have by the fundamental theorem of calculus

$$\int_{x_0}^x f'(t) dt = f(x) - f(x_0) \quad \Leftrightarrow \quad f(x) = f(x_0) + \int_{x_0}^x f'(t) dt. \quad (4.27)$$

Thus the formula is true for $n = 0$.

(2) For $n = 1$, we use that the formula is true for $n = 0$, thus (4.27) holds and we can apply (4.26) with $n = 1$ to rewrite the integral in (4.27) as (note $0! = 1$)

$$\int_{x_0}^x f'(t) dt = \frac{f'(x_0)}{1!} (x-x_0)^1 + \frac{1}{1!} \int_{x_0}^x f^{(2)}(t) (x-t)^1 dt. \quad (4.28)$$

Substituting (4.28) into (4.27) yields (4.25) for $n = 1$.

(3) Let us now assume that we have already proved (4.25) for $n = 0, 1, \dots, m-1$. For $n = m$ we may use the formula (4.25) for $n = m-1$ and obtain

$$f(x) = f(x_0) + \sum_{k=1}^{m-1} \frac{f^{(k)}(x_0)}{k!} (x-x_0)^k + \frac{1}{(m-1)!} \int_{x_0}^x (x-t)^{m-1} f^{(m)}(t) dt. \quad (4.29)$$

Now we apply the formula (4.26) to rewrite the integral in (4.29) as

$$\frac{1}{(m-1)!} \int_{x_0}^x f^{(m)}(t) (x-t)^{m-1} dt = \frac{f^{(m)}(x_0)}{m!} (x-x_0)^m + \frac{1}{m!} \int_{x_0}^x f^{(m+1)}(t) (x-t)^m dt. \quad (4.30)$$

Substitution of (4.30) into (4.29) yields now (4.25) with $n = m$. \square

4.6 Integration by Substitution

Another important technique for evaluating integrals is **substitution** or **change of variable**. Integration by substitution is the ‘inverse operation’ to the **chain rule**

$$\frac{d}{dt} F(\varphi(t)) = F'(\varphi(t)) \varphi'(t)$$

for a composition $(F \circ \varphi)(t) = F(\varphi(t))$ of two continuously differentiable functions.

Theorem 4.30 (integration by substitution – change of variable)

Let $\varphi : (c, d) \rightarrow \mathbb{R}$ be continuously differentiable. Let (c', d') be an open interval such that $\varphi((c, d)) = \{\varphi(t) : t \in (c, d)\} \subset (c', d')$, and let $f : (c', d') \rightarrow \mathbb{R}$ be continuous. Then for any $a, b \in (c, d)$

$$\int_a^b f(\varphi(t)) \varphi'(t) dt = \int_{\varphi(a)}^{\varphi(b)} f(x) dx. \quad (4.31)$$

For getting a better understanding of integration by substitution, it is helpful to consider a primitive F of f . Then the composite function $(F \circ \varphi)(t) = F(\varphi(t))$ has the derivative (from the chain rule)

$$(F \circ \varphi)'(t) = \frac{d}{dt} F(\varphi(t)) = F'(\varphi(t)) \varphi'(t) = f(\varphi(t)) \varphi'(t)$$

which is the integrand on the left-hand side of (4.31). This is the main idea of the proof of Theorem 4.30.

The formula (4.31) is called integration by substitution or change because we **substitute** $x = \varphi(t)$. With this substitution we change from the variable t on the left-hand side to the variable x on the right-hand side.

Proof of Theorem 4.30: Let $\gamma \in (c', d')$ and introduce the indefinite integral

$$F(s) := \int_{\gamma}^s f(x) dx, \quad s \in (c', d'),$$

of f , which defines a primitive of f . Then consider the function $(F \circ \varphi)(t) := F(\varphi(t))$, that is,

$$(F \circ \varphi)(t) = F(\varphi(t)) = \int_{\gamma}^{\varphi(t)} f(x) dx, \quad t \in (c, d).$$

Now differentiate $F(\varphi(t))$ with the chain rule and obtain

$$(F \circ \varphi)'(t) = F'(\varphi(t)) \varphi'(t) = f(\varphi(t)) \varphi'(t), \quad t \in (c, d), \quad (4.32)$$

where we have used that $F' = f$, since F is a primitive for f . Now we integrate (4.32) over $[a, b]$ and obtain

$$\int_a^b (F \circ \varphi)'(t) dt = \int_a^b f(\varphi(t)) \varphi'(t) dt.$$

We apply the fundamental theorem of calculus to the left-hand side and obtain

$$F(\varphi(b)) - F(\varphi(a)) = \int_a^b f(\varphi(t)) \varphi'(t) dt. \quad (4.33)$$

Now we substitute the definition of F in the expression on the left-hand side and obtain from the domain splitting property (see also Remark 3.35)

$$F(\varphi(b)) - F(\varphi(a)) = \int_{\gamma}^{\varphi(b)} f(x) dx - \int_{\gamma}^{\varphi(a)} f(x) dx = \int_{\varphi(a)}^{\varphi(b)} f(x) dx. \quad (4.34)$$

The formulas (4.33) and (4.34) now yield (4.31). \square

We give an example of integration by substitution.

Example 4.31 (integral of $(\sin t)^5 \cos t$)

Show that

$$\int_0^{\pi/2} (\sin t)^5 \cos t dt = \frac{1}{6}.$$

Solution: We substitute $\varphi = \varphi(t) := \sin t$. Then

$$\varphi'(t) = \frac{d\varphi}{dt} = \cos t \quad \Rightarrow \quad d\varphi = \cos t dt,$$

and $\varphi(0) = \sin 0 = 0$, $\varphi(\pi/2) = \sin(\pi/2) = 1$. Performing the substitution, we obtain

$$\int_0^{\pi/2} (\sin t)^5 \cos t dt = \int_0^{\pi/2} \underbrace{(\sin t)^5}_{=\varphi^5} \underbrace{\cos t dt}_{=d\varphi} = \int_0^1 \varphi^5 d\varphi = \left. \frac{\varphi^6}{6} \right|_0^1 = \frac{1}{6}$$

as claimed. \square

Note: In many practical situations, we need to **combine integration by parts and change of variable**. This is illustrated in the next examples.

Example 4.32 (integral of $x^5 \sin(x^3)$)

Evaluate the integral

$$\int_0^{\sqrt[3]{2\pi}} x^5 \sin(x^3) dx.$$

Solution: Introducing a new variable

$$y = y(x) = x^3, \quad \frac{dy}{dx} = 3x^2 \quad \Leftrightarrow \quad dy = 3x^2 dx,$$

we obtain with $y(0) = 0$ and $y(\sqrt[3]{2\pi}) = 2\pi$ from the change of variable $y = y(x)$

$$\int_0^{\sqrt[3]{2\pi}} x^5 \sin(x^3) dx = \frac{1}{3} \int_0^{\sqrt[3]{2\pi}} \underbrace{x^3}_{=y} \underbrace{\sin(x^3)}_{=\sin y} \underbrace{3x^2 dx}_{=dy} = \frac{1}{3} \int_0^{2\pi} y \sin y dy.$$

Now we use the integration by parts formula (4.16) with $F(y) = -\cos y$, $G(y) = y$ and thus $F'(y) = \sin y$ and $G'(y) = 1$ and obtain

$$\frac{1}{3} \int_0^{2\pi} y \sin y dy = \frac{1}{3} \left(-y \cos y \Big|_0^{2\pi} + \int_0^{2\pi} \cos y dy \right) = -\frac{2\pi}{3} + \frac{1}{3} \sin y \Big|_0^{2\pi} = -\frac{2\pi}{3}.$$

Thus

$$\int_0^{\sqrt[3]{2\pi}} x^5 \sin(x^3) dx = -\frac{2\pi}{3}. \quad \square$$

Example 4.33 (integral of $e^{2x} \cos(e^x - 1)$)

Evaluate the integral

$$\int_0^1 e^{2x} \cos(e^x - 1) dx$$

by using change of variable and integration by parts.

Solution: Introducing a new variable

$$y = y(x) = e^x - 1 \quad \Leftrightarrow \quad e^x = y + 1, \quad \frac{dy}{dx} = e^x \quad \Leftrightarrow \quad dy = e^x dx$$

yields with $y(0) = 0$ and $y(1) = e - 1$

$$\int_0^1 e^{2x} \cos(e^x - 1) dx = \int_0^1 \underbrace{e^x}_{=y+1} \underbrace{\cos(e^x - 1)}_{=\cos y} \underbrace{e^x dx}_{=dy} = \int_0^{e-1} (y+1) \cos y dy. \quad (4.35)$$

Now we use the integration by parts formula (4.16) with $F(y) = \sin y$, $G(y) = y + 1$ and thus $F'(y) = \cos y$ and $G'(y) = 1$ to obtain

$$\begin{aligned} \int_0^{e-1} (y+1) \cos y \, dy &= (y+1) \sin y \Big|_0^{e-1} - \int_0^{e-1} \sin y \, dy \\ &= e \sin(e-1) + \cos x \Big|_0^{e-1} = e \sin(e-1) + \cos(e-1) - 1. \end{aligned} \quad (4.36)$$

From (4.35) and (4.36) we obtain

$$\int_0^1 e^{2x} \cos(e^x - 1) \, dx = e \sin(e-1) + \cos(e-1) - 1. \quad \square$$

We can also use integration by substitution to find all primitives of a continuous function.

Example 4.34 (primitives of $(\sin x)^2 \cos x$)

Find all primitives of the function $f : \mathbb{R} \rightarrow \mathbb{R}$, given by

$$f(x) := (\sin x)^2 \cos x.$$

Solution: A primitive of the function $f(x) = (\sin x)^2 \cos x$ is given by

$$F_0(x) := \int_0^x (\sin t)^2 \cos t \, dt, \quad x \in \mathbb{R},$$

and we know (see Lemma 4.9) that all primitives are of the form

$$F(x) = F_0(x) + C = \int_0^x (\sin t)^2 \cos t \, dt + C, \quad x \in \mathbb{R},$$

with some constant $C \in \mathbb{R}$.

We use change of variable with the substitution

$$\varphi = \varphi(t) := \sin t, \quad \frac{d\varphi}{dt} = \cos t \quad \Rightarrow \quad d\varphi = \cos t \, dt,$$

with $\varphi(0) = \sin 0 = 0$ and $\varphi(x) = \sin x$. Then

$$F_0(x) = \int_0^x (\sin t)^2 \cos t \, dt = \int_0^x \underbrace{(\sin t)^2}_{=\varphi^2} \underbrace{\cos t \, dt}_{=d\varphi} = \int_0^{\sin x} \varphi^2 \, d\varphi = \frac{\varphi^3}{3} \Big|_0^{\sin x} = \frac{1}{3} (\sin x)^3.$$

Thus all primitives of $f(x) = (\sin x)^2 \cos x$ are of the form

$$F(x) = \frac{1}{3} (\sin x)^3 + C, \quad x \in \mathbb{R},$$

with some constant $C \in \mathbb{R}$. \square

Example 4.35 (primitives of $f_a(x) = a^x$)

Find all primitives of the function $f_a : \mathbb{R} \rightarrow \mathbb{R}$, given by

$$f_a(x) := a^x, \quad \text{where } a > 0 \text{ is some fixed positive real number.}$$

Solution: All primitives of $f_a(x) := a^x$ are of the form

$$F_a(x) = \int_0^x a^t dt + C, \quad x \in \mathbb{R},$$

with some constant $C \in \mathbb{R}$. We have $a^t = (e^{\ln a})^t = e^{(\ln a)t}$, and thus

$$\int_0^x a^t dt = \int_0^x e^{(\ln a)t} dt.$$

Now we have to distinguish two cases: $a = 1$ and $a \neq 1$. If $a = 1$, then we have

$$\int_0^x 1^t dt = \int_0^x 1 dt = t|_0^x = x.$$

Thus for $a = 1$ all primitives of $f_1(x) = 1^x = 1$ are of the form

$$F_1(x) = x + C, \quad x \in \mathbb{R}, \quad \text{with some constant } C \in \mathbb{R}.$$

For $a > 0$ and $a \neq 1$, we use the substitution

$$y = y(t) := (\ln a) t, \quad \frac{dy}{dt} = \ln a \quad \Rightarrow \quad dy = (\ln a) dt,$$

with $y(0) = 0$ and $y(x) = (\ln a) x$ and obtain

$$\int_0^x a^t dt = \int_0^x e^{(\ln a)t} dt = \int_0^{(\ln a)x} \frac{e^y}{\ln a} dy = \frac{e^y}{\ln a} \Big|_0^{(\ln a)x} = \frac{(e^{(\ln a)x} - 1)}{\ln a} = \frac{a^x}{\ln a} - \frac{1}{\ln a}.$$

Thus for $a > 0$ and $a \neq 1$, all primitives of $f_a(x) = a^x$ are of the form

$$F_a(x) = \frac{a^x}{\ln a} - \frac{1}{\ln a} + C = \frac{a^x}{\ln a} + \tilde{C}, \quad x \in \mathbb{R},$$

with some constant $\tilde{C} := C - 1/\ln a$. □

Example 4.36 (primitives of $1/\sqrt{1-x^2}$)

Find all primitives of the function $f : (-1, 1) \rightarrow \mathbb{R}$, given by

$$f(x) := \frac{1}{\sqrt{1-x^2}}, \quad x \in (-1, 1).$$

Solution: All primitives of the function $f(x) = 1/\sqrt{1-x^2}$ are of the form

$$F(x) = \int_0^x \frac{1}{\sqrt{1-t^2}} dt + C, \quad x \in (-1, 1), \quad \text{with some constant } C \in \mathbb{R}.$$

In order to compute the indefinite integral we use the substitution

$$t = \sin u \quad \Leftrightarrow \quad u = u(t) := \arcsin t = \sin^{-1} t, \quad \frac{dt}{du} = \cos u, \quad \Rightarrow \quad dt = \cos u \, du,$$

and with $u(0) = \arcsin 0 = 0$ and $u(x) = \arcsin x$, we find

$$\int_0^x \frac{1}{\sqrt{1-t^2}} dt = \int_0^{\arcsin x} \frac{1}{\sqrt{1-(\sin u)^2}} \cos u \, du. \quad (4.37)$$

Since $1 - (\sin u)^2 = (\cos u)^2$, we obtain

$$\int_0^{\arcsin x} \frac{\cos u}{\sqrt{1-(\sin u)^2}} du = \int_0^{\arcsin x} \frac{\cos u}{\cos u} du = \int_0^{\arcsin x} 1 \, du = u \Big|_0^{\arcsin x} = \arcsin x. \quad (4.38)$$

From (4.37) and (4.38), we see that all primitives of the function $f(x) = 1/\sqrt{1-x^2}$ are given by

$$F(x) = \arcsin x + C = \sin^{-1} x + C$$

with some constant $C \in \mathbb{R}$. □

Remark 4.37 *To find the ‘right’ method for computing an integral is mostly experience. There are some general rules of thumb how to attempt to evaluate certain types of integrals (see Handout ‘Derivatives and Integrals’ which is also included in Appendix A), but mostly it needs a lot of practice.*

4.7 Integral Test for the Convergence of a Series

Integration offers a simple way, the so-called **integral test**, to find out whether a certain type of series of positive real numbers converges or not.

Theorem 4.38 (integral test for the convergence of series)

Let r be a natural number and let $\phi : [r, \infty) \rightarrow (0, \infty)$ be a **decreasing non-negative valued** function. Then the series

$$\sum_{k=r}^{\infty} \phi(k)$$

converges **if and only if** the limit

$$\lim_{N \rightarrow \infty} \int_r^N \phi(x) dx$$

exists and is finite.

Theorem 4.38 and its proof are illustrated in Figure 4.1 below.

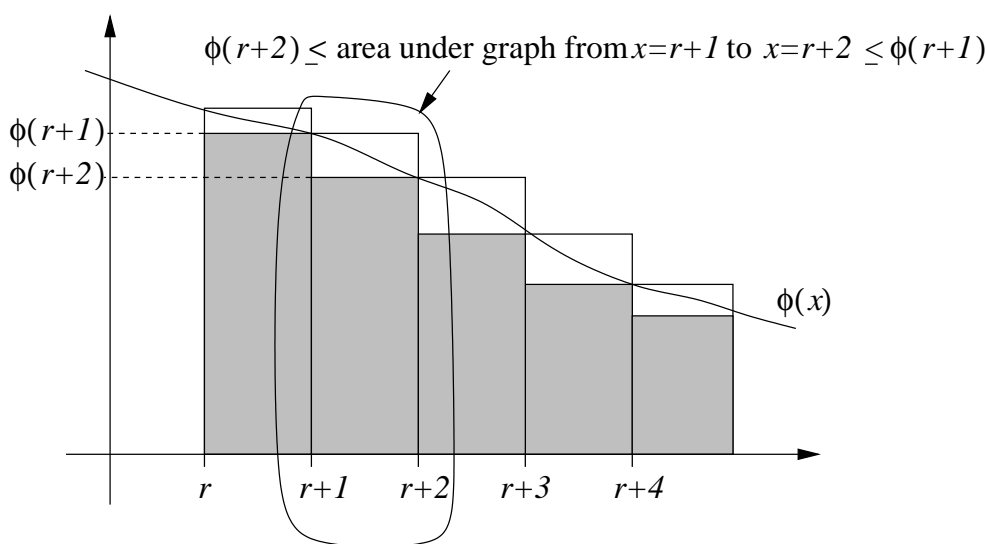


Figure 4.1: Illustration of the proof of Theorem 4.38.

Proof of Theorem 4.38 Since $\phi(x)$ is a decreasing function, we have

$$\phi(k+1) \leq \phi(x) \leq \phi(k) \quad \text{for all } x \in [k, k+1].$$

Now we integrate over $[k, k+1]$ and obtain (using the second order property of the integral)

$$\phi(k+1) = \int_k^{k+1} \phi(k+1) dx \leq \int_k^{k+1} \phi(x) dx \leq \int_k^{k+1} \phi(k) dx = \phi(k).$$

Taking the sum from $k = r$ to N , we have from the domain splitting property

$$\sum_{k=r}^N \phi(k+1) \leq \sum_{k=r}^N \int_k^{k+1} \phi(x) dx = \int_r^{N+1} \phi(x) dx \leq \sum_{k=r}^N \phi(k). \quad (4.39)$$

After these preparation we can give the proof:

\Leftarrow : Assume that the limit

$$\lim_{N \rightarrow \infty} \int_r^N \phi(x) dx$$

exists and has the value $M < \infty$. Then (4.39) implies (since ϕ is non-negative) that

$$\sum_{k=r}^N \phi(k+1) \leq \int_r^{N+1} \phi(x) dx \leq \lim_{N \rightarrow \infty} \int_r^{N+1} \phi(x) dx = M < \infty \quad \text{for all } N \geq r.$$

Hence

$$\sum_{k=r}^{\infty} \phi(k) \leq M$$

and the series $\sum_{k=0}^N \phi(k)$ converges.

\Rightarrow : The statement is equivalent to the following statement: If the limit

$$\lim_{N \rightarrow \infty} \int_r^N \phi(x) dx$$

of the integrals diverges (that is, has the value ∞) then the series $\sum_{k=r}^{\infty} \phi(k)$ diverges.

We will show this equivalent statement. Assume that

$$\lim_{N \rightarrow \infty} \int_r^N \phi(x) dx = \infty.$$

Then (4.39) implies that

$$\lim_{N \rightarrow \infty} \sum_{k=r}^N \phi(k) \geq \lim_{N \rightarrow \infty} \int_r^N \phi(x) dx = \infty,$$

hence the series is divergent. □

We give an example to show how the integral test is applied.

Example 4.39 (integral test for $\sum_{k=1}^{\infty} k^{\alpha}$)

Investigate the convergence of the series

$$\sum_{k=1}^{\infty} k^{\alpha}$$

in terms of $\alpha \in \mathbb{R}$.

Solution: Case 1: $\alpha > 0$. Here the integral test cannot be used as the function $\phi_\alpha(x) = x^\alpha$ is not decreasing, but the divergence of the series is obvious as $k^\alpha \rightarrow \infty$ as $k \rightarrow \infty$.

Case 2: $\alpha \leq 0$. Here the integral test can be used as the function $\phi_\alpha(x) = x^\alpha$ is positive and decreasing.

$$\int_1^N x^\alpha dx = \begin{cases} \left. \frac{x^{\alpha+1}}{\alpha+1} \right|_1^N = \frac{N^{\alpha+1} - 1}{\alpha+1} & \text{if } \alpha \neq -1, \\ \ln x \Big|_1^N = \ln N & \text{if } \alpha = -1. \end{cases}$$

We have $\ln N \rightarrow \infty$ as $N \rightarrow \infty$, and for $\alpha \neq -1$ we see that

$$\lim_{N \rightarrow \infty} \int_1^N x^\alpha dx = \lim_{N \rightarrow \infty} \left. \frac{x^{\alpha+1}}{\alpha+1} \right|_1^N = \lim_{N \rightarrow \infty} \left(\frac{N^{\alpha+1}}{\alpha+1} - \frac{1}{\alpha+1} \right) < \infty$$

if and only if $\alpha < -1$. **Thus $\sum_{k=1}^{\infty} k^\alpha$ converges if and only if $\alpha < -1$.** \square

Remark 4.40 (estimate from Theorem 4.38)

Under the assumption of Theorem 4.38 and if in addition ϕ is also defined and decreasing on $[r-1, \infty)$, then we have the following estimate if the series converges:

$$\lim_{N \rightarrow \infty} \int_r^N \phi(x) dx \leq \sum_{k=r}^{\infty} \phi(k) \leq \lim_{N \rightarrow \infty} \int_{r-1}^N \phi(x) dx \quad (4.40)$$

Proof: If the series converges, then we obtain from the second estimate in (4.39) for $N \rightarrow \infty$ that

$$\lim_{N \rightarrow \infty} \int_r^N \phi(x) dx = \lim_{N \rightarrow \infty} \int_r^{N+1} \phi(x) dx \leq \sum_{k=r}^{\infty} \phi(k),$$

which proves the first estimate in (4.40). To prove the second estimate in (4.40), we use the first estimate in (4.39) with r replaced by $r-1$.

$$\sum_{k=r}^{N+1} \phi(k) = \sum_{k=r-1}^N \phi(k+1) \leq \int_{r-1}^{N+1} \phi(x) dx,$$

and for $N \rightarrow \infty$ we obtain

$$\sum_{k=r}^{\infty} \phi(k) \leq \lim_{N \rightarrow \infty} \int_{r-1}^{N+1} \phi(x) dx = \lim_{N \rightarrow \infty} \int_{r-1}^N \phi(x) dx,$$

which proves the second estimate in (4.40). \square

Remark 4.41 (Comments on Theorem 4.38)

- (1) It is important to check that ϕ **is non-negative and decreasing!**
- (2) Since by the assumptions in Theorem 4.38, the function is decreasing on $[r, \infty)$, it is **Riemann integrable** over any interval $[r, N]$, with $N > r$.
- (3) Theorem 4.38 will often be stated saying that under the conditions in the theorem the series $\sum_{k=r}^{\infty} \phi(k)$ converges if and only if $\int_r^{\infty} \phi(x) dx$ exists. The so-called **improper integral** over the infinite interval $[r, \infty)$ is defined as

$$\int_r^{\infty} \phi(x) dx := \lim_{N \rightarrow \infty} \int_r^N \phi(x) dx,$$

and likewise

$$\int_{-\infty}^r f(x) dx := \lim_{N \rightarrow \infty} \int_{-N}^r f(x) dx, \quad \int_{-\infty}^{\infty} f(x) dx := \lim_{N \rightarrow \infty} \int_N^{-N} f(x) dx.$$

(Note that we have not defined improper integrals and therefore will not use this notation.)

We give some more examples of the application of the integral test.

Example 4.42 (integral test)

Prove that the series

$$\sum_{k=2}^{\infty} \frac{1}{(k+1) \ln(k+1)}$$

diverges.

Solution: Since

$$\phi(x) = (x+1)^{-1} (\ln(x+1))^{-1}$$

is positive and decreasing on $[2, \infty)$, the series converges if and only if

$$\lim_{N \rightarrow \infty} \int_2^N \frac{1}{(x+1) \ln(x+1)} dx$$

exists and is finite. We have with the substitution

$$y = y(x) := \ln(x+1), \quad \frac{dy}{dx} = \frac{1}{x+1} \quad \Leftrightarrow \quad dy = \frac{1}{x+1} dx$$

and $y(2) = \ln(3)$, $y(N) = \ln(N+1)$ that

$$\begin{aligned} \int_2^N \frac{1}{(x+1) \ln(x+1)} dx &= \int_{\ln(3)}^{\ln(N+1)} \frac{1}{y} dy \\ &= \ln y \Big|_{\ln(3)}^{\ln(N+1)} = \ln(\ln(N+1)) - \ln(\ln 3) \rightarrow \infty \quad \text{as } N \rightarrow \infty. \end{aligned}$$

So the integral does not exist, and consequently the series diverges. \square

Example 4.43 (integral test and estimate of value of the series)

Prove that the series

$$\sum_{k=1}^{\infty} \frac{1}{(2k+3)(\ln(2k+3))^2}$$

is convergent, and estimate the value of the series.

Solution: Since

$$\phi(x) = (2x+3)^{-1}(\ln(2x+3))^{-2} \quad (4.41)$$

is positive and decreasing on $[1, \infty)$ we may apply the integral test. The series converges if and only if

$$\lim_{N \rightarrow \infty} \int_1^N \frac{1}{(2x+3)(\ln(2x+3))^2} dx$$

exists and is finite. With the substitution

$$y = y(x) := \ln(2x+3), \quad \frac{dy}{dx} = \frac{2}{2x+3} \quad \Leftrightarrow \quad dy = \frac{2}{2x+3} dx$$

and $y(1) = \ln(5)$, $y(N) = \ln(2N+3)$, we have

$$\begin{aligned} \int_1^N \frac{1}{(2x+3)(\ln(2x+3))^2} dx &= \frac{1}{2} \int_{\ln 5}^{\ln(2N+3)} y^{-2} dy = -\frac{1}{2} y^{-1} \Big|_{\ln 5}^{\ln(2N+3)} \\ &= \frac{1}{2 \ln 5} - \frac{1}{2 \ln(2N+3)} \rightarrow \frac{1}{2 \ln 5} \quad \text{as } N \rightarrow \infty. \end{aligned} \quad (4.42)$$

So the integral exists, consequently the series converges and has a finite value.

Now we use Remark 4.40 to obtain an estimate for the value of the series. We have that ϕ given by (4.41) is also defined and decreasing on $[0, \infty)$. Thus we can apply Remark 4.40, and have that

$$\begin{aligned} \lim_{N \rightarrow \infty} \int_1^N \frac{1}{(2x+3)(\ln(2x+3))^2} dx \\ \leq \sum_{k=1}^{\infty} \frac{1}{(2k+3)(\ln(2k+3))^2} \leq \lim_{N \rightarrow \infty} \int_0^N \frac{1}{(2x+3)(\ln(2x+3))^2} dx, \end{aligned}$$

and from (4.42) we see that

$$\begin{aligned} \lim_{N \rightarrow \infty} \int_1^N \frac{1}{(2x+3)(\ln(2x+3))^2} dx &= \frac{1}{2 \ln 5}, \\ \lim_{N \rightarrow \infty} \int_0^N \frac{1}{(2x+3)(\ln(2x+3))^2} dx &= \frac{1}{2 \ln 3}, \end{aligned}$$

where we obtained the second integral by simply replacing $y(1) = \ln(5)$ by $y(0) = \ln(3)$ in (4.42). Thus

$$0.311 \approx \frac{1}{2 \ln 5} \leq \sum_{k=1}^{\infty} \frac{1}{(2k+3)(\ln(2k+3))^2} \leq \frac{1}{2 \ln 3} \approx 0.455. \quad \square$$

Chapter 5

Uniform Convergence

In Section 5.1, we will discuss the concepts of **pointwise convergence** and **uniform convergence** of sequences of functions. In Section 5.2, we will derive several **results related to pointwise and uniform convergence**. In Sections 5.3 and 5.4, we will discuss whether for a sequence of functions we may **interchange the limit and the integral**, and **interchange the limit and the differentiation**, respectively. Whether we may do this depends essentially on whether the series converges uniformly or not. Series of functions can be seen as sequences of functions, with the functions in the sequence being the partial sums. In Section 5.5 we will learn the **Weistrass M -test**, which gives a sufficient condition for the uniform convergence of a series of functions. In Section 5.6, we will apply the results derived in this chapter to **power series**.

5.1 Uniform Convergence of Sequences of Functions

We start the introduction of uniform convergence with the following question.

Query: If for every $x \in [a, b]$, $f_n(x) \rightarrow f(x)$ as $n \rightarrow \infty$, that is,

$$\lim_{n \rightarrow \infty} f_n(x) = f(x) \quad \text{for all } x \in [a, b], \quad (5.1)$$

does it imply that

$$\int_a^b f_n(x) dx \rightarrow \int_a^b f(x) dx \quad \text{as } n \rightarrow \infty?$$

In other words, **are we allowed to interchange the limit and the integral**, that is, is the statement

$$\lim_{n \rightarrow \infty} \int_a^b f_n(x) dx = \int_a^b \left(\lim_{n \rightarrow \infty} f_n(x) \right) dx$$

true? The answer is **in general no**, as we will learn later in this chapter! However, it is possible to interchange the limit and the integral under certain assumptions.

We give an example where we can see that it is **not** sufficient for interchanging limit and integral if a series of functions converges in the sense of (5.1) to a function f . We mention here already that (5.1) means that $\{f_n\}$ **converges pointwise** on $[a, b]$ to f , since at every fixed point $x \in [a, b]$ the sequence $\{f_n(x)\}$ of real numbers converges to the real number $f(x)$.

Example 5.1 (limits cannot be interchanged)

Consider the sequence $\{f_n\}$ of functions $f_n : [0, 1] \rightarrow \mathbb{R}$, given by

$$f_n(x) := \begin{cases} 0 & \text{if } x = 0, \\ n & \text{if } 0 < x \leq 1/n, \\ 0 & \text{if } 1/n < x \leq 1. \end{cases}$$

We have from the domain splitting property

$$\int_0^1 f_n(x) dx = \int_0^{1/n} n dx + \int_{1/n}^1 0 dx = \int_0^{1/n} n dx + 0 = n x \Big|_0^{1/n} = n \frac{1}{n} = 1.$$

For $x = 0$, we have $f_n(0) = 0$ for all n and thus $\lim_{n \rightarrow \infty} f_n(0) = 0$. For every fixed $x \in (0, 1]$, there exists some $N = N(x)$ such that $1/n < x$ for all $n \geq N$. Thus $f_n(x) = 0$ for all $n \geq N$, and thus $\lim_{n \rightarrow \infty} f_n(x) = 0$ for every fixed $x \in (0, 1]$. Thus we find that for every $x \in [0, 1]$, the sequence $\{f_n(x)\}$ of real numbers converges to 0 as $n \rightarrow \infty$. Thus $f(x) := 0$, $x \in [0, 1]$ is the so-called pointwise limit, where limit is here meant in the sense of (5.1). We find

$$\int_0^1 f(x) dx = \int_0^1 0 dx = 0$$

and thus clearly

$$\lim_{n \rightarrow \infty} \int_0^1 f_n(x) dx = 1 \neq \int_0^1 f(x) dx = 0,$$

Here we cannot interchange the limits. □

The question under which conditions the limit and the integral can be interchanged leads us to introduce the notions of **pointwise convergence** and **uniform convergence**.

Definition 5.2 (pointwise and uniform convergence I)

Let $\{f_n\}$ be a sequence of functions $f_n : \langle c, d \rangle \rightarrow \mathbb{R}$, and let $f : \langle c, d \rangle \rightarrow \mathbb{R}$ be another function.

- (i) The sequence $\{f_n\}$ **converges pointwise** on $\langle c, d \rangle$ to f if the following holds: For all $x \in \langle c, d \rangle$ and for all $\varepsilon > 0$ there exists $N = N(x, \varepsilon) \in \mathbb{N}$ such that $n \geq N$ implies $|f_n(x) - f(x)| < \varepsilon$.

Instead of ' $\{f_n\}$ converges pointwise on $\langle c, d \rangle$ to f ', we may also write more briefly ' $f_n \rightarrow f$ **pointwise** on $\langle c, d \rangle$ '.

- (ii) The sequence $\{f_n\}$ **converges uniformly on** $\langle c, d \rangle$ to f , if for all $\varepsilon > 0$ there exists $N = N(\varepsilon) \in \mathbb{N}$ such that $n \geq N$ implies

$$|f_n(x) - f(x)| < \varepsilon \quad \text{for all } x \in \langle c, d \rangle. \quad (5.2)$$

Instead of ' $\{f_n\}$ converges uniformly on $\langle c, d \rangle$ to f ' we may also write more briefly ' $f_n \rightarrow f$ **uniformly on** $\langle c, d \rangle$ '.

Remark 5.3 (difference between pointwise and uniform convergence)

(1) The difference between pointwise and uniform convergence is that in the pointwise case N depends on ε **and on** x whereas in the uniform case N depends **on** ε **only**. What does this mean?

In uniform convergence, given $\varepsilon > 0$, once the $N = N(\varepsilon)$ has been chosen, the estimate (5.2) is valid **for all** $x \in [a, b]$ and for all $n \geq N$.

In pointwise convergence, given $x \in \langle c, d \rangle$ and $\varepsilon > 0$, then there exists $N = N(x, \varepsilon)$ such that

$$|f_n(x) - f(x)| < \varepsilon \quad \text{for all } n \geq N(x, \varepsilon).$$

If we choose another point $y \in \langle c, d \rangle$ and the same $\varepsilon > 0$, then $N = N(y, \varepsilon)$ may be different from $N(x, \varepsilon)$. Since the interval $\langle c, d \rangle$ contains infinitely many points it is not clear whether, given $\varepsilon > 0$, we can find a common N such that

$$|f_n(x) - f(x)| < \varepsilon \quad \text{for all } n \geq N \quad \text{and all } x \in \langle c, d \rangle. \quad (5.3)$$

If and only if for every $\varepsilon > 0$ such a common N exists for which (5.3) holds, then the sequence $\{f_n\}$ converges uniformly on $\langle c, d \rangle$ to f .

(2) We observe that (5.2) is equivalent to the estimate

$$f(x) - \varepsilon < f_n(x) < f(x) + \varepsilon \quad \text{for all } x \in \langle c, d \rangle \text{ and for all } n \geq N(\varepsilon),$$

which is illustrated in Figure 5.1.

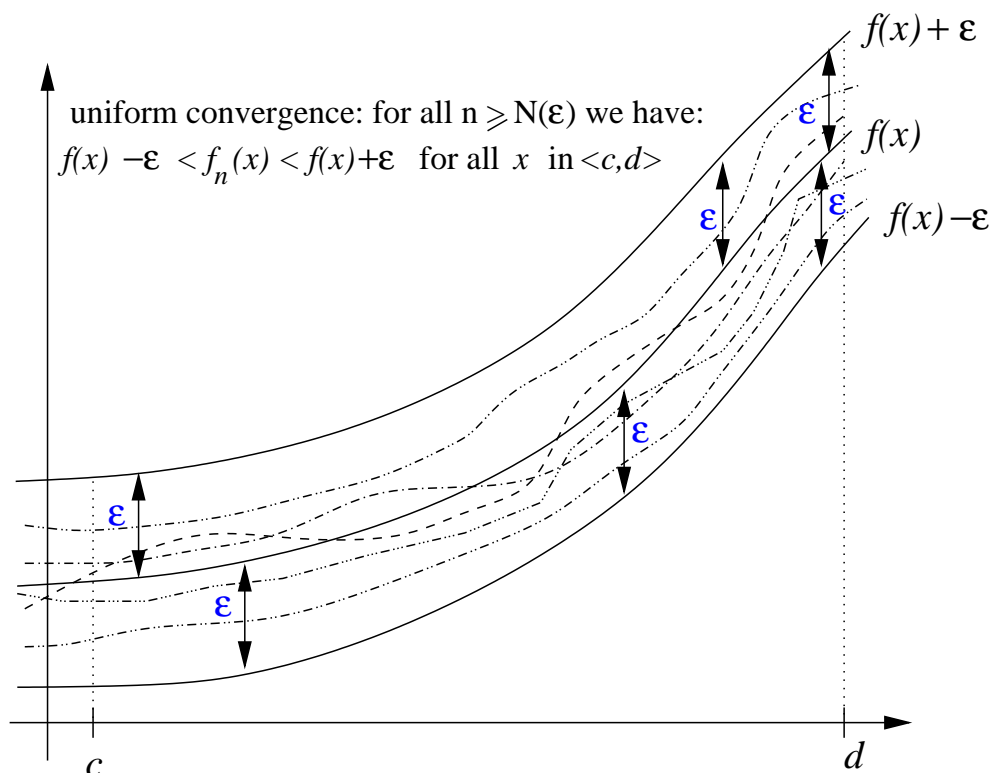


Figure 5.1: If $\{f_n\}$ converges uniformly on $\langle c, d \rangle$ to f , then we have that there exists an $N = N(\varepsilon)$ such that for all $n \geq N$, we have $f(x) - \varepsilon < f_n(x) < f(x) + \varepsilon$ for all $x \in \langle c, d \rangle$. The dash-dotted curves indicate some functions f_n with $n \geq N(\varepsilon)$.

We will show later that we may interchange the limit and the integral over a bounded closed interval $[a, b]$ if the sequence of functions $\{f_n\}$ converges uniformly on $[a, b]$ to some function $f : [a, b] \rightarrow \mathbb{R}$.

With the help of the so-called **supremum norm**, defined below, we can give an equivalent definition of uniform convergence (see Definition 5.5 below)

Definition 5.4 (supremum norm)

The **supremum norm** for bounded functions on $\langle c, d \rangle$ is defined by

$$\|f\|_\infty := \sup_{x \in \langle c, d \rangle} |f(x)|, \quad f \in \mathcal{B}(\langle c, d \rangle).$$

We now give the following definition of pointwise and uniform convergence which is equivalent to Definition 5.2.

Definition 5.5 (pointwise and uniform convergence II)

Let $\{f_n\}$ be a sequence of functions $f_n : \langle c, d \rangle \rightarrow \mathbb{R}$, and let $f : \langle c, d \rangle \rightarrow \mathbb{R}$ be another function.

(i) The sequence $\{f_n\}$ **converges pointwise** on $\langle c, d \rangle$ to f , if for all $x \in \langle c, d \rangle$ we have

$$\lim_{n \rightarrow \infty} f_n(x) = f(x).$$

(ii) The sequence $\{f_n\}$ **converges uniformly** on $\langle c, d \rangle$ to f , if

$$\lim_{n \rightarrow \infty} \left(\sup_{x \in \langle c, d \rangle} |f_n(x) - f(x)| \right) = \lim_{n \rightarrow \infty} \|f_n - f\|_\infty = 0.$$

Example 5.6 (uniformly convergent sequence)

Let $\{f_n\}$ be the sequence of functions $f_n : [0, \frac{2}{3}] \rightarrow \mathbb{R}$, given by

$$f_n(x) := x^n, \quad x \in \left[0, \frac{2}{3}\right].$$

Then $\lim_{x \rightarrow \infty} x^n = 0$ for all $x \in [0, \frac{2}{3}]$, and thus $\{f_n\}$ converges pointwise on $[0, \frac{2}{3}]$ to $f(x) := 0$, $x \in [0, \frac{2}{3}]$. We have that

$$\sup_{x \in [0, \frac{2}{3}]} |f_n(x) - f(x)| = \sup_{x \in [0, \frac{2}{3}]} |x^n| = \left(\frac{2}{3}\right)^n \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

Thus the sequence $\{f_n\}$ converges uniformly on $[0, \frac{2}{3}]$ to f . □

We come back to our introductory Example 5.1 in which we saw that pointwise convergence alone is not enough to guarantee that we can interchange limit and integral. Since we could not interchange limit and integral, we suspect that the sequence $\{f_n\}$ of functions in Example 5.1 does not converge uniformly, but how can we show this?

Remark 5.7 (Verifying that a sequence does not converge uniformly)

Let $\{f_n\}$ be a sequence of functions $f_n : \langle c, d \rangle \rightarrow \mathbb{R}$ that converges pointwise on $\langle c, d \rangle$ to $f : \langle c, d \rangle \rightarrow \mathbb{R}$. If we want to show that the sequence $\{f_n\}$ does **not converge uniformly on** $\langle c, d \rangle$ to f , we have to verify the following:

There exists $\varepsilon > 0$ such that for all $N \in \mathbb{N}$, there exists an $n \geq N$ with

$$|f_n(x) - f(x)| \geq \varepsilon \quad \text{for some } x \in \langle c, d \rangle.$$

Or equivalently: There exists $\varepsilon > 0$ such that for all $N \in \mathbb{N}$, there exists an $n \geq N$ with

$$\sup_{x \in \langle c, d \rangle} |f_n(x) - f(x)| \geq \varepsilon.$$

Also if we find that

$$\lim_{n \rightarrow \infty} \sup_{x \in \langle c, d \rangle} |f_n(x) - f(x)| \geq L > 0, \quad (5.4)$$

then we know that $\{f_n\}$ does not converge uniformly on $\langle c, d \rangle$ to f . However, the limit in (5.4) might not exist.

Example 5.8 (Example 5.1 continued)

Consider the sequence $\{f_n\}$ of functions $f_n : [0, 1] \rightarrow \mathbb{R}$, given by

$$f_n(x) := \begin{cases} 0 & \text{if } x = 0, \\ n & \text{if } 0 < x \leq 1/n, \\ 0 & \text{if } 1/n < x \leq 1. \end{cases}$$

In Example 5.1, we have seen that $\{f_n\}$ converges pointwise on $[0, 1]$ to the function $f(x) := 0$, $x \in [0, 1]$. Now we will show that the sequence $\{f_n\}$ does not converge uniformly on $[0, 1]$ to f . We have for all $n \in \mathbb{N}$ that

$$\sup_{x \in [0, 1]} |f_n(x) - f(x)| \geq |f_n(1/n) - 0| = |1 - 0| = 1.$$

Thus we have for $\varepsilon = 1$ that for all $N \in \mathbb{N}$

$$\sup_{x \in [0, 1]} |f_n(x) - f(x)| \geq 1 \quad \text{for all } n \geq N,$$

and we have proved that the sequence $\{f_n\}$ does not converge uniformly on $[0, 1]$ to the function f . \square

Notation: If we say that a sequence $\{f_n\}$ of functions $f_n : \langle c, d \rangle \rightarrow \mathbb{R}$ converges pointwise on $\langle c, d \rangle$ (or is pointwise convergent on $\langle c, d \rangle$), then we mean that is

converges pointwise on $\langle c, d \rangle$ to some function $f : \langle c, d \rangle \rightarrow \mathbb{R}$. If we say that a sequence $\{f_n\}$ of functions $f_n : \langle c, d \rangle \rightarrow \mathbb{R}$ converges uniformly on $\langle c, d \rangle$ (or is uniformly convergent on $\langle c, d \rangle$), then we mean that it converges uniformly on $\langle c, d \rangle$ to some function $f : \langle c, d \rangle \rightarrow \mathbb{R}$.

The next remark deals with the relation between uniform convergence and pointwise convergence. As will be explained, uniform convergence is the stronger type of convergence: **uniform convergence implies pointwise convergence**, but the reverse statement is **not** true!

Remark 5.9 (uniform convergence \Rightarrow pointwise convergence)

(1) Since for any $x \in \langle c, d \rangle$

$$|f_n(x) - f(x)| \leq \sup_{x \in \langle c, d \rangle} |f_n(x) - f(x)|,$$

we see that if $\{f_n\}$ converges uniformly on $\langle c, d \rangle$ to f , then $\{f_n\}$ converges pointwise on $\langle c, d \rangle$ to f . The converse is **not** true: pointwise convergence does **not** imply uniform convergence.

(2) From (1) we see that a sequence $\{f_n\}$ that does not converge pointwise on $\langle c, d \rangle$ cannot converge uniformly on $\langle c, d \rangle$. Pointwise convergence is a **necessary** condition for uniform convergence, but is it **not a sufficient** condition.

Here is how you should proceed when you want to investigate whether a sequence $\{f_n\}$ of functions $f_n : \langle c, d \rangle \rightarrow \mathbb{R}$ converges uniformly on $\langle c, d \rangle$.

Recipe for checking pointwise and uniform convergence: If we inspect a sequence $\{f_n\}$ of function on $\langle c, d \rangle$ for pointwise and uniform convergence, our first step is to check whether it converges pointwise by trying to find the pointwise limit. Only if it is pointwise convergent, we need to check whether it also converges uniformly towards the pointwise limit.

Example 5.10 (not uniformly convergent sequence)

Let $\{f_n\}$ be the sequence of functions $f_n : [0, 1] \rightarrow \mathbb{R}$, given by

$$f_n(x) := x^n, \quad x \in [0, 1].$$

Then $\lim_{x \rightarrow \infty} x^n = 0$ for all $x \in [0, 1)$ and $\lim_{n \rightarrow \infty} f_n(1) = \lim_{n \rightarrow \infty} 1^n = 1$. Thus $\{f_n\}$ converges pointwise on $[0, 1]$ to

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

Now we have that

$$\sup_{x \in [0,1]} |f_n(x) - f(x)| \geq \sup_{x \in [0,1)} |x^n - 0| = \sup_{x \in [0,1)} |x^n| = 1^n = 1.$$

Thus for $\varepsilon = 1$ and for all $N \in \mathbb{N}$, we have that

$$\sup_{x \in [0,1]} |f_n(x) - f(x)| \geq 1 \quad \text{for all } n \geq N,$$

and we see that the sequence $\{f_n\}$ does not converge uniformly on $[0, 1]$ to f .

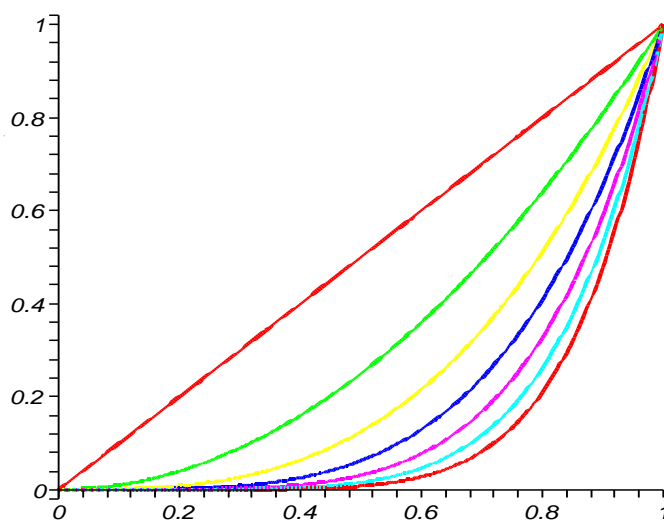


Figure 5.2: The functions $f_1(x) = x$, $f_2(x) = x^2$, $f_3(x) = x^3$, \dots , $f_7(x) = x^7$ on the interval $[0, 1]$.

In Figure 5.2, we have plotted $f_1(x) = x$, $f_2(x) = x^2$, $f_3(x) = x^3$, \dots , $f_7(x) = x^7$, and we see how the functions approach the pointwise limit. Keeping the picture in Figure 5.1 in mind, we can also see from Figure 5.2 why the convergence cannot be uniform: If we lay an ε -tube around the pointwise limit, then this ε -tube gives

$$f(x) - \varepsilon = -\varepsilon < f_n(x) < f(x) + \varepsilon = \varepsilon \quad \text{for all } x \in [0, 1),$$

where we have ignored the point $x = 1$ for the moment. Since the $f_n(x) = x^n$ are continuous and $f_n(1) = 1^n$, we have

$$\lim_{x < 1, x \rightarrow 1} f_n(x) = \lim_{x < 1, x \rightarrow 1} x^n = 1,$$

and therefore for any n , the value $f_n(x)$ will lie outside $(-\varepsilon, \varepsilon)$ for any $\varepsilon < 1$ if x is close enough to 1. \square

We discuss some more examples.

Example 5.11 (pointwise but not uniformly convergent sequence)

Consider $f_n(x) = nx^n(1-x^n)$ on $[0, 1]$. Show that $\lim_{n \rightarrow \infty} f_n(x) = 0$ for all $x \in [0, 1]$. This means that the pointwise limit is $f(x) = 0$, $x \in [0, 1]$. Then investigate whether the series converges uniformly on $[0, 1]$ to f .

Solution: Since $f_n(0) = 0$ and $f_n(1) = 0$ for all n , clearly $\lim_{n \rightarrow \infty} f_n(0) = 0$ and $\lim_{n \rightarrow \infty} f_n(1) = 0$. For $x \in (0, 1)$, we use that

$$\lim_{n \rightarrow \infty} nx^n = 0, \quad 0 < x < 1. \quad (5.5)$$

With $\lim_{n \rightarrow \infty} x^n = 0$ for $x \in (0, 1)$ and $x^{-n} = e^{-n \ln x}$, this follows from de l'Hopital's rule (with n as variable): for $x \in (0, 1)$

$$\lim_{n \rightarrow \infty} nx^n = \lim_{n \rightarrow \infty} \frac{n}{x^{-n}} = \lim_{n \rightarrow \infty} \frac{n}{e^{-n \ln x}} = \lim_{n \rightarrow \infty} \frac{1}{-(\ln x)e^{-n \ln x}} = \lim_{n \rightarrow \infty} -\frac{x^n}{(\ln x)} = 0.$$

From (5.5) and from $|1 - x^n| \leq 1$ for $x \in [0, 1]$, we get

$$|nx^n(1 - x^n)| = |nx^n| |1 - x^n| \leq |nx^n| \quad \text{for all } x \in (0, 1).$$

Thus, from the sandwich theorem,

$$0 \leq \lim_{n \rightarrow \infty} |nx^n(1 - x^n)| \leq \lim_{n \rightarrow \infty} |nx^n| = 0 \quad \Rightarrow \quad \lim_{n \rightarrow \infty} nx^n(1 - x^n) = 0, \quad x \in (0, 1).$$

This shows that $\lim_{n \rightarrow \infty} f_n(x) = 0$ for all $x \in [0, 1]$. Thus $f(x) := 0$, $x \in [0, 1]$, is the pointwise limit of $\{f_n\}$.

Now we consider

$$\sup_{x \in [0, 1]} |f_n(x) - f(x)| = \sup_{x \in [0, 1]} |nx^n(1 - x^n)|.$$

The function $f_n(x) - f(x) = f_n(x)$ is zero in $x = 0$ and $x = 1$ and we work out the maximum of $|nx^n(1 - x^n)| = |nx^n - nx^{2n}|$ on $[0, 1]$. To do this we compute the zeros of the derivative of $f_n(x) = nx^n(1 - x^n) = nx^n - nx^{2n}$

$$(nx^n(1 - x^n))' = n^2x^{n-1} - 2n^2x^{2n-1} = n^2x^{n-1}(1 - 2x^n) = 0 \quad \Rightarrow \quad x = 0 \text{ or } x^n = 1/2.$$

Since $f_n(0) = f_n(1) = 0$ and since $x = \sqrt[n]{1/2}$ is the only critical point in $(0, 1)$, $f_n(\sqrt[n]{1/2})$ must be the minimum or maximum of f_n , and thus the maximum of $|f_n|$. Thus

$$\sup_{x \in [0,1]} |f_n(x) - f(x)| = \sup_{x \in [0,1]} |n x^n (1 - x^n)| = \left| n \frac{1}{2} \left(1 - \frac{1}{2} \right) \right| = \frac{n}{4} \rightarrow \infty \quad \text{as } n \rightarrow \infty.$$

Thus the $\{f_n\}$ does not converge uniformly on $[0, 1]$ to $f(x) = 0$. \square

Example 5.12 (uniformly convergent sequence)

Define $f_n : [0, \pi] \rightarrow \mathbb{R}$ by

$$f_n(x) := \frac{(\cos x)^n}{2n + (\cos x)^n}, \quad x \in [0, \pi].$$

Find the pointwise limit, and determine whether the convergence is uniform on $[0, \pi]$.

Solution: We have $|(\cos x)^n| \leq 1$, and thus

$$|2n + (\cos x)^n| \geq |2n| - |(\cos x)^n| \geq 2n - 1 \quad \text{for all } x \in [0, \pi].$$

Hence we can estimate $|f_n(x)|$ as follows

$$0 \leq |f_n(x)| = \left| \frac{(\cos x)^n}{2n + (\cos x)^n} \right| = \frac{|(\cos x)^n|}{|2n + (\cos x)^n|} \leq \frac{|\cos x|^n}{2n - |\cos x|^n} \leq \frac{1}{2n - 1}. \quad (5.6)$$

From (5.6) and the sandwich theorem

$$0 \leq \lim_{n \rightarrow \infty} |f_n(x) - 0| = \lim_{n \rightarrow \infty} \left| \frac{(\cos x)^n}{2n + (\cos x)^n} \right| \leq \lim_{n \rightarrow \infty} \frac{1}{2n - 1} = 0,$$

and we see that the pointwise limit of $\{f_n\}$ is $f(x) := 0$, $x \in [0, \pi]$. Similarly,

$$0 \leq \lim_{n \rightarrow \infty} \sup_{x \in [0, \pi]} |f_n(x) - f(x)| = \lim_{n \rightarrow \infty} \sup_{x \in [0, \pi]} \left| \frac{(\cos x)^n}{2n + (\cos x)^n} - 0 \right| \leq \lim_{n \rightarrow \infty} \frac{1}{2n - 1} = 0.$$

Hence, from the sandwich theorem, $\{f_n\}$ converges uniformly on $[0, \pi]$ to the function $f(x) = 0$, $x \in [0, \pi]$. \square

Example 5.13 (uniformly convergent sequence)

Suppose $\{f_n\}$, $f_n : [-1, 1] \rightarrow \mathbb{R}$, converges uniformly on $[-1, 1]$ to $f : [-1, 1] \rightarrow \mathbb{R}$. Let $F_n(x) := |f_n(x)|$, $x \in [-1, 1]$, and let $\varphi_n(t) := f_n(\cos t)$, $t \in \mathbb{R}$. For each of the sequences $\{F_n\}$ on $[-1, 1]$ and $\{\varphi_n\}$ on \mathbb{R} find out whether or not it is pointwise and uniformly convergent.

Solution: Claim: $\{F_n\}$ converges uniformly on $[-1, 1]$ to $F(x) := |f(x)|$, $x \in [-1, 1]$, and $\{\varphi_n\}$ converges uniformly on \mathbb{R} to $\varphi(t) := f(\cos t)$, $t \in \mathbb{R}$.

Indeed, let $\varepsilon > 0$ be arbitrary. Since $\{f_n\}$ converges uniformly on $[-1, 1]$ to f , there exists $N = N(\varepsilon)$ such that $n \geq N$ implies

$$|f_n(x) - f(x)| < \varepsilon, \quad \text{for all } x \in [-1, 1]. \quad (5.7)$$

Hence for any $n \geq N$ we have from the lower triangle inequality

$$|F_n(x) - F(x)| = ||f_n(x)| - |f(x)|| \leq |f_n(x) - f(x)| < \varepsilon \quad \text{for all } x \in [-1, 1].$$

Since $\cos t \in [-1, 1]$ for all $t \in \mathbb{R}$, we have from (5.7) that for all $n \geq N$

$$|\varphi_n(t) - \varphi(t)| = |f_n(\cos t) - f(\cos t)| < \varepsilon \quad \text{for all } t \in \mathbb{R}.$$

Thus $\{F_n\}$ converges uniformly on $[-1, 1]$ to $F(x) = |f(x)|$, and $\{\varphi_n\}$ converges uniformly on \mathbb{R} to $\varphi(t) = f(\cos t)$. \square

5.2 Results on Uniform Convergence

In analogy to Cauchy sequences in \mathbb{R} , we introduce **uniform Cauchy sequences**, and in analogy to the real numbers \mathbb{R} we can prove a **uniform Cauchy principle**. Then we will prove the important statement that the **(uniform) limit of a uniformly convergent sequence of continuous functions is also continuous**. From this statement we draw a useful conclusion.

We start by remembering the definition of a Cauchy sequence of real numbers and the uniform Cauchy principle (for sequences in \mathbb{R}).

Definition 5.14 (Cauchy sequence)

A sequence of real numbers $\{a_n\} \subset \mathbb{R}$ is called a **Cauchy sequence** if for all $\varepsilon > 0$ there exists $N = N(\varepsilon) \in \mathbb{N}$ such that $m, n \geq N$ implies $|a_n - a_m| < \varepsilon$.

Lemma 5.15 (Cauchy Principle)

A sequence of real numbers $\{a_n\} \subset \mathbb{R}$ is a Cauchy sequence **if and only if** it is convergent.

The Cauchy Principle has a deep meaning and it will reappear in various forms in many other mathematical courses.

In analogy to Definition 5.14, we define **uniform Cauchy sequences**.

Definition 5.16 (uniform Cauchy sequence)

A sequence $\{f_n\}$ of functions $f_n : \langle c, d \rangle \rightarrow \mathbb{R}$ is called a **uniform Cauchy sequence** if for all $\varepsilon > 0$ there exists $N = N(\varepsilon) \in \mathbb{N}$ such that $m, n \geq N$ implies

$$|f_n(x) - f_m(x)| < \varepsilon \quad \text{for all } x \in \langle c, d \rangle.$$

Or equivalently: A sequence $\{f_n\}$ of functions $f_n : \langle c, d \rangle \rightarrow \mathbb{R}$ is called a **uniform Cauchy sequence** if for all $\varepsilon > 0$ there exists $N = N(\varepsilon) \in \mathbb{N}$ such that $m, n \geq N$ implies

$$\|f_n - f_m\|_\infty = \sup_{x \in \langle c, d \rangle} |f_n(x) - f_m(x)| < \varepsilon.$$

(Note: Here N depends only on ε .)

We give an example of a uniform Cauchy sequence.

Example 5.17 (uniform Cauchy sequence)

In Example 5.6 we have seen that the sequence $\{f_n\}$ of functions

$$f_n(x) := x^n, \quad x \in \left[0, \frac{2}{3}\right].$$

converges uniformly on $[0, 2/3]$ to the function $f(x) := 0$. We will now show that $\{f_n\}$ is a Cauchy sequence.

Indeed given $\varepsilon > 0$, choose $N = N(\varepsilon)$ such that $2(2/3)^N < \varepsilon$. Then we find $|x^n - x^m| < |x^n| + |x^m|$ and thus for all $n, m \geq N$

$$\begin{aligned} \sup_{x \in [0, 2/3]} |f_n(x) - f_m(x)| &= \sup_{x \in [0, 2/3]} |x^n - x^m| \leq \sup_{x \in [0, 2/3]} (|x^n| + |x^m|) \\ &= \left(\frac{2}{3}\right)^n + \left(\frac{2}{3}\right)^m \leq 2 \left(\frac{2}{3}\right)^N < \varepsilon. \end{aligned}$$

Thus $\{f_n\}$ is indeed a uniform Cauchy sequence. □

In analogy to the Cauchy principle for sequences of real numbers (see Lemma 5.15 above), we can prove a **uniform Cauchy principle**.

Theorem 5.18 (uniform Cauchy principle)

A sequence $\{f_n\}$ of functions $f_n : \langle c, d \rangle \rightarrow \mathbb{R}$, is a uniform Cauchy sequence **if and only if** it is uniformly convergent on $\langle c, d \rangle$.

Proof of Theorem 5.18: \Leftarrow : We first prove that uniform convergence implies that the sequence is a uniform Cauchy sequence. We will use an $\varepsilon/2$ argument and the triangle inequality.

Suppose $\{f_n\}$ converges uniformly on $\langle c, d \rangle$ to $f : \langle c, d \rangle \rightarrow \mathbb{R}$. Then for every $\varepsilon > 0$ there is some $N = N(\varepsilon) \in \mathbb{N}$ such that for all $n, m \geq N$

$$\begin{aligned} |f_n(x) - f(x)| &< \frac{\varepsilon}{2} && \text{for all } x \in \langle c, d \rangle, \\ |f_m(x) - f(x)| &< \frac{\varepsilon}{2} && \text{for all } x \in \langle c, d \rangle. \end{aligned}$$

Thus for all $n, m \geq N$ and for every $x \in \langle c, d \rangle$, we have from the triangle inequality that

$$\begin{aligned} |f_n(x) - f_m(x)| &= |(f_n(x) - f(x)) + (f(x) - f_m(x))| \\ &\leq |f_n(x) - f(x)| + |f(x) - f_m(x)| \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

Since $\varepsilon > 0$ was arbitrary, we see that $\{f_n\}$ is a uniform Cauchy sequence.

\Rightarrow : Now we prove that if $\{f_n\}$ is a uniform Cauchy sequence, then it is uniformly convergent on $\langle c, d \rangle$.

If $\{f_n\}$ is a uniform Cauchy sequence, then for every $\varepsilon > 0$ there exists an integer $N = N(\varepsilon) \in \mathbb{N}$ such that

$$|f_n(x) - f_m(x)| < \varepsilon \quad \text{for all } x \in \langle c, d \rangle \text{ and for all } n, m \geq N.$$

Thus for each fixed $x \in \langle c, d \rangle$, $\{f_n(x)\}$ is a Cauchy sequence of real numbers. By the Cauchy principle for sequences of real numbers, for each $x \in \langle c, d \rangle$, $\{f_n(x)\}$ is convergent, and we denote its limit by $f(x) := \lim_{n \rightarrow \infty} f_n(x)$. Now we need to show that $\{f_n\}$ converges uniformly on $\langle c, d \rangle$ to the function $f : \langle c, d \rangle \rightarrow \mathbb{R}$, $x \mapsto f(x)$, defined by the pointwise limits. Since $\{f_n\}$ is a uniform Cauchy sequence, for every $\varepsilon > 0$, there is some $M = M(\varepsilon) \in \mathbb{N}$ such that

$$|f_n(x) - f_m(x)| < \frac{\varepsilon}{2} \quad \text{for all } x \in \langle c, d \rangle \text{ and all } m, n \geq M.$$

We now fix n and let, for each $x \in \langle c, d \rangle$, $m \rightarrow \infty$ in the above inequality. We obtain

$$|f_n(x) - f(x)| \leq \frac{\varepsilon}{2} < \varepsilon \quad \text{for all } x \in \langle c, d \rangle \text{ and all } n \geq M.$$

Thus $\{f_n\}$ converges uniformly on $\langle c, d \rangle$ to f . □

The next theorem is of great importance. Later-on, we will come back to it and get an even deeper understanding of its relevance in a broader mathematical context.

Theorem 5.19 (uniform limit is continuous)

Let $\{f_n\}$ be a sequence of continuous functions $f_n : \langle c, d \rangle \rightarrow \mathbb{R}$ that converges uniformly on $\langle c, d \rangle$ to the function $f : \langle c, d \rangle \rightarrow \mathbb{R}$. Then f is continuous on $\langle c, d \rangle$.

Remark 5.20 (necessary condition for uniform convergence)

Theorem 5.19 provides a **necessary condition** for the uniform convergence of sequences of **continuous** functions: If the convergence is uniform, then the uniform limit is **continuous**. This implies the following (by contraposition):

If the pointwise limit of a sequence $\{f_n\}$ of continuous functions $f_n : \langle c, d \rangle \rightarrow \mathbb{R}$ is **not** continuous, then the sequence does **not** converge uniformly on $\langle c, d \rangle$.

Note that continuity of the limit of a sequence of continuous functions is only a **necessary** condition for uniform convergence, but **not** a **sufficient** one. If the pointwise limit of a pointwise convergent sequence of continuous functions is continuous, we **cannot** conclude that the sequence converges uniformly.

We give two examples to illustrate Theorem 5.19 and Remark 5.20.

Example 5.21 (sequence with discontinuous pointwise limit)

Let $\{f_n\}$ be given by $f_n(x) := x^n$, $x \in [0, 1]$. Since $\lim_{n \rightarrow \infty} x^n = 0$ for all $|x| < 1$ and $\lim_{n \rightarrow \infty} f_n(1) = \lim_{n \rightarrow \infty} 1 = 1$, the pointwise limit of $\{f_n\}$ is given by

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

The functions $f_n(x) = x^n$ are continuous on $[0, 1]$. Since the function f is not continuous at $x = 1$, we know from Theorem 5.19 that $\{f_n\}$ does not converge uniformly on $[0, 1]$ to f . \square

Example 5.22 (only pointwise convergent sequence with continuous limit)

If we define the sequence $\{g_n\}$ by $g_n(x) := x^n$, $x \in [0, 1)$, we see that $\{g_n\}$ converges pointwise on $[0, 1]$ to the continuous function $g(x) = 0$, $x \in [0, 1)$. (Note that in contrast to the previous example we consider the half-open interval $[0, 1)$.) Since $g(x) = 0$, $x \in [0, 1)$, is continuous, we cannot tell by Theorem 5.19 whether the

convergence is uniform or not. By using the definition of uniform convergence, we know can show that $\{g_n\}$ does not converge uniformly on $[0, 1)$ to g as follows: Since

$$\|g_n - g\|_\infty = \sup_{x \in [0, 1)} |x^n| = 1 \quad \text{for all } n \in \mathbb{N}.$$

we see that $\lim_{n \rightarrow \infty} \|g - g_n\|_\infty = 1 \neq 0$. Thus the sequence $\{g_n\}$ does not converge uniformly on $[0, 1)$ to g . \square

Proof of Theorem 5.19: Taking an arbitrary $x_0 \in \langle c, d \rangle$. Then we need to show that f is continuous at x_0 . This is done with an $\varepsilon/3$ argument as follows.

Let $x \in \langle c, d \rangle$ be close to x_0 . Then we have the following estimate by the triangle inequality:

$$|f(x) - f(x_0)| \leq |f(x) - f_n(x)| + |f_n(x) - f_n(x_0)| + |f_n(x_0) - f(x_0)|. \quad (5.8)$$

From the definition of the uniform convergence, for every $\varepsilon > 0$, there exists an $N = N(\varepsilon) \in \mathbb{N}$ such that for all $n \geq N$

$$|f_n(y) - f(y)| < \frac{\varepsilon}{3} \quad \text{for all } y \in \langle c, d \rangle.$$

Thus for $n \geq N$, the first and the third term in (5.8) are less than $\varepsilon/3$.

We need to use the continuity of f_n at x_0 to make the second term small. Let us now choose n in (5.8) to be $n = N$. (For this $n = N$, the first and the third term are less than $\varepsilon/3$ as explained above.) Since f_N is continuous on $\langle c, d \rangle$, it is, in particular, continuous at x_0 . Therefore by the definition of continuity, there is a $\delta > 0$ such that

$$|f_N(x) - f_N(x_0)| < \frac{\varepsilon}{3} \quad \text{for all } x \in \langle c, d \rangle \text{ with } |x - x_0| < \delta.$$

Thus we have for all $x \in \langle c, d \rangle$ with $|x - x_0| < \delta$,

$$\begin{aligned} |f(x) - f(x_0)| &\leq |f(x) - f_N(x)| + |f_N(x) - f_N(x_0)| + |f_N(x_0) - f(x_0)| \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon. \end{aligned}$$

Thus f is continuous at x_0 . This completes the proof. \square

We discuss some more examples.

Example 5.23 (sequence with discontinuous pointwise limit)

Define the continuous functions $f_n : [0, \frac{\pi}{2}] \rightarrow \mathbb{R}$ by

$$f_n(x) := \frac{(\sin x)^n}{2 + (\sin x)^n}, \quad x \in \left[0, \frac{\pi}{2}\right].$$

Find the pointwise limit of $\{f_n\}$ and determine whether the convergence is uniform by using Theorem 5.19.

Solution: If $x \in [0, \frac{\pi}{2})$, then $0 \leq \sin x < 1$ and $(\sin x)^n \rightarrow 0$ as $n \rightarrow \infty$. Hence $\lim_{n \rightarrow \infty} f_n(x) = 0$ for $x \in [0, \frac{\pi}{2})$. If $x = \pi/2$, then $\sin(\pi/2) = 1$ and $f_n(\pi/2) = 1/3$ for all n , and $\lim_{n \rightarrow \infty} f_n(\pi/2) = 1/3$. Thus $\{f_n\}$ converges pointwise on $[0, \frac{\pi}{2}]$ to

$$f(x) := \begin{cases} 0 & \text{if } x \in [0, \pi/2), \\ 1/3 & \text{if } x = \pi/2. \end{cases}$$

Since the functions f_n are all continuous but f is discontinuous, by Theorem 5.19, the convergence cannot be uniform. \square

Example 5.24 (Example 5.12 continued)

Define the sequence $\{f_n\}$ of functions $f_n : [0, \pi] \rightarrow \mathbb{R}$ by

$$f_n(x) := \frac{(\cos x)^n}{2n + (\cos x)^n}, \quad x \in [0, \pi].$$

In Example 5.12, we have seen that the sequence $\{f_n\}$ converges pointwise and uniformly on $[0, \pi]$, to the function $f(x) = 0$, $x \in [0, \pi]$. We see that the functions f_n are continuous and that the limit f is continuous. \square

Example 5.25 (uniformly convergent sequence)

Let $\{f_n\}$ be the sequence of continuous functions $f_n : [0, 1] \rightarrow \mathbb{R}$ defined by

$$f_n(x) := e^{x \frac{n+1}{n}}, \quad x \in [0, 1].$$

Show that the sequence $\{f_n\}$ converges pointwise and uniformly on $[0, 1]$, and find the pointwise and uniform limit.

Solution: We have (from $\lim_{n \rightarrow \infty} \frac{n+1}{n} = \lim_{n \rightarrow \infty} (1 + \frac{1}{n}) = 1$)

$$\lim_{n \rightarrow \infty} f_n(x) = \lim_{n \rightarrow \infty} e^{x \frac{n+1}{n}} = \lim_{n \rightarrow \infty} e^{x(1 + \frac{1}{n})} = e^x \quad \text{for all } x \in [0, 1].$$

Thus the pointwise limit is given by $f(x) := e^x$, $x \in [0, 1]$. Now we will show that the function converges also uniformly on $[0, 1]$ to $f(x) = e^x$.

$$\sup_{x \in [0, 1]} |f_n(x) - f(x)| = \sup_{x \in [0, 1]} \left| e^{x \frac{n+1}{n}} - e^x \right| = \sup_{x \in [0, 1]} |e^x (e^{x/n} - 1)| = |e^1 (e^{1/n} - 1)| \rightarrow 0$$

as $n \rightarrow \infty$. Thus we see that the sequence $\{f_n\}$ converges indeed uniformly on $[0, 1]$ to $f(x) = e^x$. Since the f_n are continuous, by Theorem 5.19, the uniform limit has also to be continuous. Obviously $f(x) = e^x$ is continuous. \square

Example 5.26 (pointwise convergent sequence with discontinuous limit)

Let the sequence $\{f_n\}$ of functions $f_n : [0, 1] \rightarrow \mathbb{R}$ be defined by

$$f_n(x) := e^{-nx}, \quad x \in [0, 1].$$

Investigate whether the sequence converges pointwise and uniformly on the interval $[0, 1]$, and find the pointwise limit if $\{f_n\}$ converges pointwise.

Solution: We have $\lim_{n \rightarrow \infty} f_n(x) = \lim_{n \rightarrow \infty} e^{-nx} = 0$ for $x \in (0, 1]$ and $\lim_{n \rightarrow \infty} f_n(0) = \lim_{n \rightarrow \infty} 1 = 1$. Thus the sequence $\{f_n\}$ converges pointwise to the function

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

Since the functions $f_n(x) = e^{-nx}$ are continuous on $[0, 1]$ but the pointwise limit f is not continuous at $x = 0$, we know from Theorem 5.19 that the sequence $\{f_n\}$ does not converge uniformly to f . \square

5.3 Interchange of Limit and Integral

Now we come back to our original question under which conditions on a sequence of functions we can **interchange limit and the integral**.

Theorem 5.27 (interchange of integral and limit)

Let $\{f_n\}$ be a sequence of functions $f_n : [a, b] \rightarrow \mathbb{R}$ which are Riemann integrable over $[a, b]$ and which converge uniformly on the bounded closed interval $[a, b]$ to the function $f : [a, b] \rightarrow \mathbb{R}$. Then the following statements hold:

(i) The uniform limit f is Riemann integrable over $[a, b]$.

(ii) The sequence $\{F_n\}$ of indefinite integrals

$$F_n(x) := \int_a^x f_n(t) dt$$

converges uniformly on $[a, b]$ to

$$F(x) := \int_a^x f(t) dt.$$

(iii) In particular,

$$\lim_{n \rightarrow \infty} \int_a^b f_n(t) dt = \int_a^b f(t) dt.$$

Before we prove Theorem 5.27, we discuss an example where we apply Theorem 5.27.

Example 5.28 (interchange of limit and integral)

Let $f_n : [1, 2] \rightarrow \mathbb{R}$ be defined by

$$f_n(x) := \frac{(\cos(x^3))^n}{x^4 + n}, \quad x \in [1, 2].$$

Find

$$\lim_{n \rightarrow \infty} \int_1^2 f_n(x) dx.$$

Solution: The functions $f_n : [1, 2] \rightarrow \mathbb{R}$ are continuous and hence (by Theorem 3.18) Riemann integrable over $[1, 2]$. Since $|\cos(x^3)| \leq 1$ for all $n \in \mathbb{N}$,

$$0 \leq |f_n(x)| = \left| \frac{(\cos(x^3))^n}{x^4 + n} \right| = \frac{|(\cos(x^3))^n|}{|x^4 + n|} = \frac{|\cos(x^3)|^n}{|x|^4 + n} \leq \frac{1}{1 + n} \quad \text{for all } x \in [1, 2]. \quad (5.9)$$

Therefore, from the sandwich theorem, $\lim_{n \rightarrow \infty} f_n(x) = 0$ for all $x \in [1, 2]$, and $f(x) := 0$, $x \in [1, 2]$, is the pointwise limit of $\{f_n\}$. Since $f(x) = 0$ for all $x \in [1, 2]$, we have from (5.9) that

$$0 \leq |f_n(x) - f(x)| = \left| \frac{(\cos(x^3))^n}{x^4 + n} - 0 \right| = \left| \frac{(\cos(x^3))^n}{x^4 + n} \right| \leq \frac{1}{1 + n} \quad \text{for all } x \in [1, 2].$$

Thus from the sandwich theorem

$$0 \leq \lim_{n \rightarrow \infty} \|f_n - f\|_\infty = \lim_{n \rightarrow \infty} \sup_{x \in [1, 2]} |f_n(x) - f(x)| \leq \lim_{n \rightarrow \infty} \frac{1}{1 + n} = 0,$$

and we see that $\{f_n\}$ converges uniformly on $[1, 2]$ to $f(x) = 0$, $x \in [1, 2]$, and Theorem 5.27 implies that

$$\lim_{n \rightarrow \infty} \int_1^2 f_n(x) dx = \int_1^2 \left(\lim_{n \rightarrow \infty} f_n(x) \right) dx = \int_1^2 f(x) dx = \int_1^2 0 dx = (2 - 1) 0 = 0. \quad \square$$

Proof of Theorem 5.27 (i): We prove (i) by using Riemann's criterion for integrability (see Corollary 3.4). From the uniform convergence of $\{f_n\}$ we know that for every $\varepsilon > 0$, there is some $N = N(\varepsilon) \in \mathbb{N}$ such that

$$|f_n(x) - f(x)| < \frac{\varepsilon}{3(b-a)} \quad \text{for all } x \in [a, b] \text{ and all } n \geq N$$

or equivalently

$$f_n(x) - \frac{\varepsilon}{3(b-a)} < f(x) < f_n(x) + \frac{\varepsilon}{3(b-a)} \quad \text{for all } x \in [a, b] \text{ and all } n \geq N. \quad (5.10)$$

Now we fix $n = N$. Since $f_N \in \mathcal{R}([a, b])$, by Riemann's criterion of integrability, there is a partition $P = \{x_0, x_1, \dots, x_m\}$, such that,

$$0 \leq U(f_N, P) - L(f_N, P) < \frac{\varepsilon}{3}. \quad (5.11)$$

From (5.10), we have for $n = N$

$$f_N(x) - \frac{\varepsilon}{3(b-a)} < f(x) < f_N(x) + \frac{\varepsilon}{3(b-a)} \quad \text{for all } x \in [a, b],$$

and this implies

$$m_k(f_N) - \frac{\varepsilon}{3(b-a)} \leq m_k(f) \leq M_k(f) \leq M_k(f_N) + \frac{\varepsilon}{3(b-a)}, \quad k = 1, 2, \dots, m, \quad (5.12)$$

where we have used the notation

$$m_k(g) := \inf_{x \in [x_{k-1}, x_k]} g(x), \quad M_k(g) := \sup_{x \in [x_{k-1}, x_k]} g(x).$$

Multiplying (5.12) by $(x_k - x_{k-1})$ and subsequently summing over k , we obtain

$$L(f_N, P) - \frac{\varepsilon}{3} \leq L(f, P) \leq U(f, P) \leq U(f_N, P) + \frac{\varepsilon}{3}.$$

This implies that

$$U(f, P) - L(f, P) \leq \left(U(f_N, P) + \frac{\varepsilon}{3} \right) - \left(L(f_N, P) - \frac{\varepsilon}{3} \right) < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon,$$

where we have used (5.11) in the second last step. By Riemann's criterion of integrability (see Corollary 3.4), we know that $f \in \mathcal{R}([a, b])$. \square

Proof of Theorem 5.27 (ii): Since $\{f_n\}$ converges uniformly on $[a, b]$ to f , for every $\varepsilon > 0$, there is some $N = N(\varepsilon) \in \mathbb{N}$ such that

$$|f_n(t) - f(t)| < \frac{\varepsilon}{2(b-a)} \quad \text{for all } t \in [a, b] \text{ and all } n \geq N. \quad (5.13)$$

Thus for $n \geq N$ and $x \in [a, b]$, we have from (5.13)

$$\begin{aligned} |F_n(x) - F(x)| &= \left| \int_a^x f_n(t) dt - \int_a^x f(t) dt \right| = \left| \int_a^x (f_n(t) - f(t)) dt \right| \\ &\leq \int_a^x |f_n(t) - f(t)| dt \leq \int_a^x \frac{\varepsilon}{2(b-a)} dt = \frac{\varepsilon(x-a)}{2(b-a)} \leq \frac{\varepsilon(b-a)}{2(b-a)} = \frac{\varepsilon}{2} < \varepsilon. \end{aligned}$$

Since $\varepsilon > 0$ was arbitrary, we see that $\{F_n\}$ converges uniformly on $[a, b]$ to F . \square

Proof of Theorem 5.27 (iii): Statement (iii) follows from (ii) by taking $x = b$. \square

We discuss some more examples.

Example 5.29 (interchange of limit and integral)

Let $f_n : [0, 1] \rightarrow \mathbb{R}$, where $n \in \mathbb{N}$, be defined by

$$f_n(x) := \frac{n + (\sin(e^x))^n}{2n + x^3}, \quad x \in [0, 1].$$

Find

$$\lim_{n \rightarrow \infty} \int_0^1 f_n(x) dx.$$

Solution: The functions f_n are continuous and hence Riemann integrable over $[0, 1]$. Since $|\sin(e^x)| \leq 1$ for all $x \in [0, 1]$ and $0 \leq x^3/2 \leq 1/2$ for all $x \in [0, 1]$, we have for all x in $[0, 1]$

$$0 \leq \left| f_n(x) - \frac{1}{2} \right| = \left| \frac{(\sin(e^x))^n - \frac{1}{2}x^3}{2n + x^3} \right| \leq \frac{|\sin(e^x)|^n + \frac{1}{2}|x|^3}{2n + x^3} \leq \frac{3/2}{2n + x^3} \leq \frac{3}{4n}. \quad (5.14)$$

Therefore from the sandwich theorem $\lim_{n \rightarrow \infty} |f_n(x) - 1/2| = 0$ for all $x \in [0, 1]$, and the pointwise limit of $\{f_n\}$ on $[0, 1]$ is $f(x) := 1/2$, $x \in [0, 1]$. From (5.14),

$$0 \leq \sup_{x \in [0, 1]} \left| f_n(x) - \frac{1}{2} \right| \leq \frac{3}{4n} \rightarrow 0 \quad \text{as } n \rightarrow \infty,$$

and, for $n \rightarrow \infty$, we see from the sandwich theorem that the series $\{f_n\}$ converges uniformly on $[0, 1]$ to $f(x) := 1/2$, $x \in [0, 1]$. Thus Theorem 5.27 implies that

$$\lim_{n \rightarrow \infty} \int_0^1 f_n(x) dx = \int_0^1 f(x) dx = \int_0^1 \frac{1}{2} dx = \frac{x}{2} \Big|_0^1 = \frac{1}{2}. \quad \square$$

Example 5.30 (Theorem 5.27 cannot be applied)

Consider $\{f_n\}$ defined by $f_n(x) := x^n$ on $[0, 1]$. From Example 5.21, we know that $\{f_n\}$ converges pointwise but not uniformly on $[0, 1]$, and that the pointwise limit is given by

$$f(x) := \lim_{n \rightarrow \infty} f_n(x) = \lim_{n \rightarrow \infty} x^n = \begin{cases} 0 & \text{if } 0 \leq x < 1, \\ 1 & \text{if } x = 1. \end{cases}$$

Since the convergence is not uniform, we cannot apply Theorem 5.27.

However we have $f_n \in \mathcal{R}([0, 1])$ for all $n \in \mathbb{N}$, $f \in \mathcal{R}([0, 1])$,

$$\int_0^1 f(x) dx = 0,$$

and

$$\int_0^1 f_n(x) dx = \int_0^1 x^n dx = \frac{x^{n+1}}{n+1} \Big|_0^1 = \frac{1}{n+1}.$$

Thus

$$\lim_{n \rightarrow \infty} \int_0^1 f_n(x) dx = \lim_{n \rightarrow \infty} \frac{1}{n+1} = 0 = \int_0^1 f(x) dx. \quad (5.15)$$

In this example the limit of the integrals converges to the integral of the limit, despite the fact the convergence is not uniformly! \square

Remark 5.31 (interchange of limit and integral)

Example tells us that it is sometimes possible to interchange the limit and the integral under weaker conditions! Indeed, the so-called Lebesgue integral, a generalization of the Riemann integral, which will be not discussed in this class, explains why (5.15) has to be true.

5.4 Interchange of Limit and Differentiation

Next we want to investigate under which conditions on the convergence of a sequence of differentiable functions, we can **interchange limit and differentiation**.

Theorem 5.32 (interchange of differentiation and limit.)

Let $\{f_n\}$ be a sequence of continuously differentiable functions $f_n : (a, b) \rightarrow \mathbb{R}$, and let $f : (a, b) \rightarrow \mathbb{R}$ and $g : (a, b) \rightarrow \mathbb{R}$ be two other functions. Suppose that $\{f_n\}$ converges pointwise on (a, b) to f , and that $\{f'_n\}$ converges uniformly on (a, b) to g . Then f is continuously differentiable and

$$f'(x) = g(x) \quad \text{for all } x \in (a, b).$$

Substituting $f(x)$ and $g(x)$ by $\lim_{n \rightarrow \infty} f_n(x)$ and $\lim_{n \rightarrow \infty} f'_n(x)$, respectively, yields

$$\left(\lim_{n \rightarrow \infty} f_n(x) \right)' = \lim_{n \rightarrow \infty} f'_n(x) \quad \text{for all } x \in (a, b),$$

*that is, we are allowed to **interchange the limit and the differentiation**.*

Proof of Theorem 5.32: Fix $c \in (a, b)$. By the fundamental theorem of calculus, we may write

$$\int_c^x f'_n(t) dt = f_n(x) - f_n(c), \quad x \in (a, b). \quad (5.16)$$

Now let $n \rightarrow \infty$ in (5.16). Since $\{f'_n\}$ converges uniformly to g , Theorem 5.27 implies that the left-hand side in (5.16) converges to

$$\lim_{n \rightarrow \infty} \int_c^x f'_n(t) dt = \int_c^x \left(\lim_{n \rightarrow \infty} f'_n(t) \right) dt = \int_c^x g(t) dt. \quad (5.17)$$

Since $\{f_n\}$ converges pointwise to f , the right-hand side in (5.16) converges to

$$\lim_{n \rightarrow \infty} (f_n(x) - f_n(c)) = f(x) - f(c). \quad (5.18)$$

Taking the limit for $n \rightarrow \infty$ on both sides of (5.16) and using (5.17) and (5.18) yields for all $x \in (a, b)$

$$\int_c^x g(t) dt = f(x) - f(c) \quad \Leftrightarrow \quad f(c) + \int_c^x g(t) dt = f(x). \quad (5.19)$$

We observe that the function g is continuous because it is the limit of the uniformly convergent series $\{f'_n\}$ of continuous functions f'_n (due to Theorem 5.19). Since g is continuous on (a, b) , the left hand side in (5.19) is differentiable and its derivative is $g(x)$ (see Theorem 4.14). Thus the right hand side is differentiable. Differentiating both sides in (5.19) yields $g(x) = f'(x)$. \square

We give two examples of the application of Theorem 5.32.

Example 5.33 (interchange of limit and differentiation)

Let the sequence $\{f_n\}$ of functions $f_n : (e+1, e+2) \rightarrow \mathbb{R}$ be given by

$$f_n(x) := \frac{1}{x^n}, \quad e+1 < x < e+2.$$

Since $\lim_{n \rightarrow \infty} x^n = \infty$ for all $x \in (e+1, e+2)$, the series $\{f_n\}$ converges pointwise on $(e+1, e+2)$ to $f(x) := 0$. The functions f_n are all continuously differentiable on $(e+1, e+2)$, and the series of the derivatives $\{f'_n\}$ is given by

$$f'_n(x) = -\frac{n}{x^{n+1}}, \quad e+1 < x < e+2.$$

Now we show that the series $\{f'_n\}$ converges uniformly on $(e+1, e+2)$ to $g(x) := 0$.

$$\sup_{x \in (e+1, e+2)} |f'_n(x) - g(x)| = \sup_{x \in (e+1, e+2)} \frac{n}{x^{n+1}} = \frac{n}{(e+1)^{n+1}} \rightarrow 0 \quad \text{as } n \rightarrow \infty. \quad (5.20)$$

That $\lim_{n \rightarrow \infty} n/(e+1)^{n+1} = 0$ can be concluded with de l'Hospital's rule as follows:

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{n}{(e+1)^{n+1}} &= \lim_{n \rightarrow \infty} \frac{n}{e^{(n+1)\ln(e+1)}} = \lim_{n \rightarrow \infty} \frac{(n)'}{(e^{(n+1)\ln(e+1)})'} \\ &= \lim_{n \rightarrow \infty} \frac{1}{\ln(e+1) e^{(n+1)\ln(e+1)}} = \lim_{n \rightarrow \infty} \frac{1}{\ln(e+1) (e+1)^{n+1}} = 0. \end{aligned}$$

From (5.20), we see that $\{f'_n\}$ converges uniformly on $(e+1, e+2)$ to $g(x) = 0$.

From Theorem 5.32 we can now conclude that $f(x) = 0$, $x \in (e+1, e+2)$, (the pointwise and uniform limit of $\{f_n\}$) is continuously differentiable on $(e+1, e+2)$ and that $f'(x) = g(x) = 0$ for all $x \in (e+1, e+2)$. \square

Example 5.34 (interchange of limit and differentiation)

The sequence $\{f_n\}$ of functions $f_n : \mathbb{R} \rightarrow \mathbb{R}$ (where $n \in \mathbb{N}$), given by

$$f_n(x) := e^x + \frac{(\cos x)^n}{2n^2 + \sin x}, \quad x \in \mathbb{R},$$

satisfies (from $|\sin x| \leq 1$, $|\cos x| \leq 1$ for all $x \in \mathbb{R}$)

$$0 \leq \sup_{x \in \mathbb{R}} |f_n(x) - e^x| = \sup_{x \in \mathbb{R}} \left| \frac{(\cos x)^n}{2n^2 + \sin x} \right| \leq \sup_{x \in \mathbb{R}} \frac{|\cos x|^n}{n^2 + (n^2 - |\sin x|)} \leq \frac{1}{n^2}.$$

Since the upper bound tends to zero as $n \rightarrow \infty$, we see from the sandwich theorem that the series converges pointwise and uniformly on \mathbb{R} to $f(x) := e^x$, $x \in \mathbb{R}$. The functions f_n are differentiable, and the derivative of f_n is given by

$$f'_n(x) = e^x + \frac{(-1)n(\cos x)^{n-1} \sin x (2n^2 + \sin x) - (\cos x)^n \cos x}{(2n^2 + \sin x)^2}.$$

Now we show that the sequence of derivatives $\{f'_n\}$ converges also uniformly on \mathbb{R} to $g(x) = e^x$, $x \in \mathbb{R}$. To verify this, we estimate $|f'_n(x) - e^x|$ as follows

$$\begin{aligned} 0 \leq |f'_n(x) - e^x| &= \left| \frac{(-1)n(\cos x)^{n-1} \sin x (2n^2 + \sin x) - (\cos x)^n \cos x}{(2n^2 + \sin x)^2} \right| \\ &\leq \frac{n|\cos x|^{n-1} |\sin x| (2n^2 + |\sin x|) + |\cos x|^{n+1}}{(2n^2 + |\sin x|)^2} \\ &\leq \frac{n(2n^2 + 1) + 1}{n^4} = \frac{2}{n} + \frac{1}{n^3} + \frac{1}{n^4} \rightarrow 0 \quad \text{as } n \rightarrow \infty, \end{aligned}$$

where we have used $|\sin x| \leq 1$ and $|\cos x| \leq 1$. Thus from the sandwich theorem

$$0 \leq \limsup_{n \rightarrow \infty} \sup_{x \in \mathbb{R}} |f'_n(x) - e^x| \leq \lim_{n \rightarrow \infty} \left(\frac{2}{n} + \frac{1}{n^3} + \frac{1}{n^4} \right) = 0,$$

and we see that $\{f'_n\}$ converges uniformly on \mathbb{R} to $g(x) := e^x$, $x \in \mathbb{R}$. From Theorem 5.32, we know that f is differentiable on \mathbb{R} and that

$$(e^x)' = f'(x) = \frac{d}{dx} \left(\lim_{n \rightarrow \infty} f_n(x) \right) = \lim_{n \rightarrow \infty} f'_n(x) = g(x) = e^x, \quad x \in \mathbb{R},$$

thus $(e^x)' = e^x$ for all $x \in \mathbb{R}$. □

5.5 Uniform Convergence of Series of Functions and Weierstrass M -Test

If we deal with **series of functions** (for example, **power series**), we can apply all the definitions and theorems that we learned in this chapter to the sequence of partial sums. For example:

Let $\{f_n\}$, $f_n : \langle c, d \rangle \rightarrow \mathbb{R}$, be a sequence of functions. Then we say that the series

$$s(x) := \sum_{n=1}^{\infty} f_n(x), \quad x \in \langle c, d \rangle, \quad (5.21)$$

converges uniformly on $\langle c, d \rangle$ if the sequence $\{s_m\}$ of partial sums

$$s_m(x) := \sum_{n=1}^m f_n(x), \quad x \in \langle c, d \rangle,$$

converges uniformly on $\langle c, d \rangle$. The uniform limit is then the function $s : \langle c, d \rangle \rightarrow \mathbb{R}$ defined by the series (5.21).

The so-called **Weierstrass M -test** gives us a criterion for the uniform convergence of a series of functions.

Theorem 5.35 (Weierstrass M -test)

Let $\{f_n\}$ be a sequence of bounded functions $f_n : \langle c, d \rangle \rightarrow \mathbb{R}$, and suppose that for each n there is a positive real number M_n such that

$$\|f_n\|_{\infty} = \sup_{x \in \langle c, d \rangle} |f_n(x)| \leq M_n$$

and that

$$\sum_{n=1}^{\infty} M_n < \infty.$$

Then the series

$$\sum_{n=1}^{\infty} f_n(x) \quad (5.22)$$

converges uniformly on $\langle c, d \rangle$.

It should be noted that the Weierstrass M -test gives a **sufficient** condition for the uniform convergence of a series of functions, but this condition is not necessary. In other words, if the assumptions in Theorem 5.35 are **not** satisfied then we **cannot** conclude anything: the series of functions could converge uniformly or not.

We demonstrate the application of the Weierstrass M -test for an example.

Example 5.36 (Weierstrass M -test)

Show that the series

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

converges uniformly on $[-\frac{1}{2}, \frac{1}{2}]$.

Solution We want to apply the Weierstrass M -test. Let $\{f_n\}$ be given by

$$f_n(x) := \frac{x^n}{1+x^n}, \quad \left[-\frac{1}{2}, \frac{1}{2}\right].$$

Since $|1+x^n| \geq 1/2$ for all $x \in [-\frac{1}{2}, \frac{1}{2}]$ and $|x^n| \leq (1/2)^n$ for all $x \in [-\frac{1}{2}, \frac{1}{2}]$, we have

$$\sup_{x \in [-\frac{1}{2}, \frac{1}{2}]} |f_n(x)| = \sup_{x \in [-\frac{1}{2}, \frac{1}{2}]} \left| \frac{x^n}{1+x^n} \right| \leq 2 \left(\frac{1}{2}\right)^n = \left(\frac{1}{2}\right)^{n-1} =: M_n.$$

From the geometric series

$$\sum_{n=1}^{\infty} M_n = \sum_{n=1}^{\infty} \left(\frac{1}{2}\right)^{n-1} = \sum_{\ell=0}^{\infty} \left(\frac{1}{2}\right)^{\ell} = \frac{1}{1-1/2} = 2 < \infty.$$

Thus from the Weierstrass M -test we know that

$$\sum_{n=1}^{\infty} f_n(x) = \sum_{n=1}^{\infty} \frac{x^n}{1+x^n}$$

converges uniformly on $[-\frac{1}{2}, \frac{1}{2}]$. □

Now we prove the Weierstrass M -test.

Proof of Theorem 5.35: Since $\sum_{n=1}^{\infty} M_n < \infty$, for any given $\varepsilon > 0$, there is some $N = N(\varepsilon) \in \mathbb{N}$, such that

$$\sum_{n=k}^{\infty} M_n < \varepsilon \quad \text{for all } k \geq N. \quad (5.23)$$

Now we use this to show that the sequence $\{s_m\}$ of partial sums

$$s_m(x) = \sum_{n=1}^m f_n(x), \quad x \in \langle c, d \rangle,$$

is a uniform Cauchy sequence. Then we know from the uniform Cauchy principle (see Theorem 5.18) that the sequence of partial sums $\{s_m\}$ converges uniformly on $\langle c, d \rangle$ and thus the series (5.22) converges uniformly on $\langle c, d \rangle$.

Let N be the integer N from (5.23), then for any $m > k \geq N$ and all $x \in \langle c, d \rangle$

$$0 \leq |s_m(x) - s_k(x)| = \left| \sum_{n=1}^m f_n(x) - \sum_{n=1}^k f_n(x) \right| = \left| \sum_{n=k+1}^m f_n(x) \right|$$

$$\leq \sum_{n=k+1}^m |f_n(x)| \leq \sum_{n=k+1}^m M_n \leq \sum_{n=k+1}^{\infty} M_n < \varepsilon,$$

where we have used (5.23) in the last step. Hence

$$0 \leq \sup_{x \in \langle c, d \rangle} |s_m(x) - s_k(x)| \leq \sum_{n=k+1}^{\infty} M_n < \varepsilon,$$

and we see that the sequence of partial sums $\{s_m\}$ is a uniform Cauchy sequence. Consequently, the series (5.22) is uniformly convergent. \square

We discuss some more examples.

Example 5.37 (Weierstrass M -test)

Show that the series

$$\sum_{n=1}^{\infty} \frac{(\sin x)^n}{n!}$$

converges uniformly on \mathbb{R} .

Solution: We want to apply the Weierstrass M -test. Let $\{f_n\}$ be given by

$$f_n(x) := \frac{(\sin x)^n}{n!}, \quad x \in \mathbb{R}.$$

Then, since $|\sin x| \leq 1$ for all $x \in \mathbb{R}$,

$$\sup_{x \in \mathbb{R}} |f_n(x)| = \sup_{x \in \mathbb{R}} \left| \frac{(\sin x)^n}{n!} \right| \leq \frac{1}{n!} =: M_n,$$

and

$$\sum_{n=1}^{\infty} M_n = \sum_{n=1}^{\infty} \frac{1}{n!} < \infty,$$

which can be shown with the ratio test for series of real numbers. Thus from the Weierstrass M -test we know that

$$\sum_{n=1}^{\infty} f_n(x) = \sum_{n=1}^{\infty} \frac{(\sin x)^n}{n!}$$

converges uniformly on \mathbb{R} . \square

Example 5.38 (Weierstrass M -test)

Prove that the series

$$\sum_{n=1}^{\infty} \frac{1 + n \sin(nx)}{n^{4-\cos(nx)}}$$

is uniformly convergent on $[0, 2\pi]$.

Solution: We want to apply the Weierstrass M -test. Let $\{f_n\}$ be defined by

$$f_n(x) := \frac{1 + n \sin(nx)}{n^{4-\cos(nx)}}, \quad x \in [0, 2\pi].$$

Since $|\sin(nx)| \leq 1$ and $|\cos(nx)| \leq 1$ for all $x \in [0, 2\pi]$, we have

$$0 \leq |f_n(x)| = \left| \frac{1 + n \sin(nx)}{n^{4-\cos(nx)}} \right| \leq \frac{1 + n |\sin(nx)|}{n^{4-|\cos(nx)|}} \leq \frac{1 + n}{n^3} \leq \frac{2n}{n^3} = \frac{2}{n^2}, \quad x \in [0, 2\pi].$$

Thus

$$0 \leq \sup_{x \in [0, 2\pi]} |f_n(x)| = \sup_{x \in [0, 2\pi]} \left| \frac{1 + n \sin(nx)}{n^{4-\cos(nx)}} \right| \leq \frac{2}{n^2} =: M_n,$$

and we have

$$\sum_{n=1}^{\infty} M_n = \sum_{n=1}^{\infty} \frac{2}{n^2} = 2 \sum_{n=1}^{\infty} \frac{1}{n^2} < \infty,$$

which can be seen from the integral test, since we have that $\phi(x) = 1/x^2$ is non-negative and decreasing and

$$\lim_{N \rightarrow \infty} \int_1^N x^{-2} dx = \lim_{N \rightarrow \infty} \left(-\frac{1}{x} \Big|_1^N \right) = \lim_{N \rightarrow \infty} \left(1 - \frac{1}{N} \right) = 1 < \infty.$$

(see Theorem 4.38). Thus, from the Weierstrass M -test, the series

$$\sum_{n=1}^{\infty} f_n(x) = \sum_{n=1}^{\infty} \frac{1 + n \sin(nx)}{n^{4-\cos(nx)}}$$

is uniformly convergent on $[0, 2\pi]$. □

5.6 The Convergence of Power Series

Now we will consider **power series** (see Chapter 1) which are a special class of series of functions. We will apply everything that we have learnt in this chapter to discuss the convergence of power series, and we will also investigate whether we may integrate and differentiate a power series term by term.

Remember that a **power series** centred at x_0 is of the form

$$\sum_{n=0}^{\infty} c_n (x - x_0)^n = c_0 + c_1 (x - x_0) + c_2 (x - x_0)^2 + \dots + c_n (x - x_0)^n + \dots,$$

and that there exists a real number $0 \leq \rho \leq \infty$, called the **radius of convergence**, such that:

- (i) if $\rho = 0$ the series converges absolutely only at $x = x_0$,
- (ii) if $\rho = \infty$ the series converges absolutely for all $x \in \mathbb{R}$, and
- (iii) otherwise the series converges absolutely for all $x \in \mathbb{R}$ with $|x - x_0| < \rho$ and diverges for all $x \in \mathbb{R}$ with $|x - x_0| > \rho$.

On the interval of convergence $(x_0 - \rho, x_0 + \rho)$, a power series defines a function $f : (x_0 - \rho, x_0 + \rho) \rightarrow \mathbb{R}$, via the pointwise limit

$$f(x) := \sum_{n=0}^{\infty} c_n (x - x_0)^n, \quad x \in (x_0 - \rho, x_0 + \rho).$$

In practice, the radius of convergence is usually found using either the ratio test (see Lemmas 1.14 and 1.8) or the root test (see Lemmas 1.16 and 1.9). The most general formula of the radius of convergence is given in Theorem 1.23.

The first result is that a power series centred about x_0 with radius of convergence ρ converges uniformly on every closed bounded interval $[a, b] \subset (x_0 - \rho, x_0 + \rho)$.

Theorem 5.39 (uniform convergence of a power series)

Consider a power series

$$\sum_{n=0}^{\infty} c_n (x - x_0)^n$$

*with radius of convergence ρ , and let $[a, b] \subset (x_0 - \rho, x_0 + \rho)$ be a **bounded closed** interval. Then the power series **converges uniformly** on $[a, b]$.*

Note: It is important that $[a, b] \subset (x_0 - \rho, x_0 + \rho)$ is a bounded closed interval!

Example 5.40 (uniform convergence of power series)

In Example 1.27 and Example 1.36, we have derived that the exponential function $f(x) := e^x$ has the Taylor series centred at $x_0 = 0$, given by

$$\sum_{n=0}^{\infty} \frac{1}{n!} x^n,$$

and we have verified that the series converges for all $x \in \mathbb{R}$ to the exponential function $f(x) = e^x$. In other words the radius of convergence is $\rho = \infty$ and the pointwise limit is $f(x) = e^x$. Thus we know from Lemma 5.39 that the power series converges on every closed bounded interval $[a, b] \subset (-\infty, \infty) = \mathbb{R}$ to the exponential function $f(x) = e^x$. \square

Proof of Lemma 5.39: Let $\delta := \max\{|b - x_0|, |x_0 - a|\}$. Then $\delta < \rho$, and since the power series converges on $(x_0 - \rho, x_0 + \rho)$, there exists $y \in (x_0 - \rho, x_0 + \rho)$ with

$$|x - x_0| \leq \delta < |y - x_0| < \rho \quad \text{for all } x \in [a, b], \quad (5.24)$$

and the power series converges absolutely at y , that is,

$$\sum_{n=0}^{\infty} |c_n (y - x_0)^n| < \infty.$$

In particular, (5.25) implies that $\{|c_n (y - x_0)^n|\}$ tends to zero for $n \rightarrow \infty$, and thus the sequence $\{|c_n (y - x_0)^n|\}$ is bounded, that is, there exists $M \in \mathbb{R}$ such that

$$|c_n (y - x_0)^n| \leq M \quad \text{for all } n \in \mathbb{N}_0. \quad (5.25)$$

We want to show that

$$\sum_{n=0}^{\infty} c_n (x - x_0)^n$$

converges uniformly on $[a, b]$ with the Weierstrass M -test. From (5.24) and (5.25)

$$\sup_{x \in [a, b]} |c_n (x - x_0)^n| = \sup_{x \in [a, b]} |c_n (y - x_0)^n| \left(\frac{|x - x_0|}{|y - x_0|} \right)^n \leq M \left(\frac{\delta}{|y - x_0|} \right)^n,$$

and from the geometric series

$$\sum_{n=0}^{\infty} M \left(\frac{\delta}{|y - x_0|} \right)^n = M \sum_{n=0}^{\infty} \left(\frac{\delta}{|y - x_0|} \right)^n < \infty$$

since $\delta/|y - x_0| < 1$ (from (5.24)). Thus, due to the Weierstrass M -test, the power series converges uniformly on $[a, b]$. \square

The next lemma is technical and is needed to prove Theorem 5.42 below, which is the main result of this section. We will prove Lemma 5.41 at the end of this section.

Lemma 5.41 (integration and differentiation term by term)

If the power series

$$\sum_{n=0}^{\infty} c_n (x - x_0)^n \quad (5.26)$$

has the radius of convergence ρ , then the two power series

$$\sum_{n=1}^{\infty} n c_n (x - x_0)^{n-1} \quad \text{and} \quad \sum_{n=0}^{\infty} \frac{c_n}{n+1} (x - x_0)^{n+1} \quad (5.27)$$

also have the radius of convergence ρ .

We observe that the two series in (5.27) are just the two power series that we obtain if we differentiate and integrate the original power series (5.26) term by term, respectively. Now we can state the main result of this section.

Theorem 5.42 (basic theorem on power series)

Consider a power series

$$\sum_{n=0}^{\infty} c_n (x - x_0)^n \quad (5.28)$$

with radius of convergence $\rho > 0$, and let $f : (x_0 - \rho, x_0 + \rho) \rightarrow \mathbb{R}$ be the function defined by this series. Then the following statements are true:

- (i) The power series (5.28) **converges uniformly** on any bounded closed interval $[a, b] \subset (x_0 - \rho, x_0 + \rho)$.
- (ii) The function f is **continuous** on the interval $(x_0 - \rho, x_0 + \rho)$.
- (iii) The function f is **Riemann integrable** over any bounded closed interval $[a, b] \subset (x_0 - \rho, x_0 + \rho)$, and the integral $\int_a^b f(x) dx$ can be computed by **integrating** the series (5.28) **‘term-by-term’**:

$$\begin{aligned} \int_a^b f(x) dx &= \int_a^b \left(\sum_{n=0}^{\infty} c_n (x - x_0)^n \right) dx = \sum_{n=0}^{\infty} \left(\int_a^b c_n (x - x_0)^n dx \right) \\ &= \sum_{n=0}^{\infty} \frac{c_n (x - x_0)^{n+1}}{n+1} \Big|_a^b = \sum_{n=0}^{\infty} \frac{c_n}{n+1} [(b - x_0)^{n+1} - (a - x_0)^{n+1}]. \end{aligned} \quad (5.29)$$

- (iv) The function f has a **primitive** $F : (x_0 - \rho, x_0 + \rho) \rightarrow \mathbb{R}$, defined by

$$F(x) = \sum_{n=0}^{\infty} \int_{x_0}^x c_n (y - x_0)^n dy = \sum_{n=0}^{\infty} \frac{c_n}{n+1} (x - x_0)^{n+1}. \quad (5.30)$$

The power series on the right-hand side of (5.30) has the same radius of convergence ρ as the power series (5.28).

- (v) The function f is **differentiable** on $(x_0 - \rho, x_0 + \rho)$, and its derivative f' can be computed by **differentiating** the series (5.28) **‘term-by-term’**:

$$f'(x) = \frac{d}{dx} \left(\sum_{n=0}^{\infty} c_n (x - x_0)^n \right) = \sum_{n=0}^{\infty} c_n \frac{d}{dx} (x - x_0)^n = \sum_{n=1}^{\infty} n c_n (x - x_0)^{n-1}, \quad (5.31)$$

where $x \in (x_0 - \rho, x_0 + \rho)$. The power series on the right-hand side of (5.31) has the same radius of convergence ρ as the power series (5.28).

Remark 5.43 (repeated differentiation/integration term by term)

According to Theorem 5.42, the power series and its derivative (and its primitive), obtained by differentiating (and integrating) term by term, have the same radius of convergence. Naturally, we can repeat the procedure and differentiate (and integrate, respectively) again. Repeating this procedure, we see that a **power series is infinitely often continuously differentiable on** $(x_0 - \rho, x_0 + \rho)$ and that the k th derivative can be obtained by **differentiating the series k -times term by term**. The series representing the k th derivative will have the same radius of convergence. Likewise we can **integrate the series term by term as often as we like** and the resulting series will also have the same radius of convergence.

We give some examples to show how Theorem 5.42 can be applied.

Example 5.44 (derivative of e^x)

We have seen in Example 1.27 that the exponential function e^x has the Taylor series centred at $x_0 = 0$, given by

$$\sum_{n=0}^{\infty} \frac{1}{n!} x^n,$$

and that this Taylor has the radius of convergence $\rho = \infty$ and converges to e^x for all $x \in \mathbb{R}$ (see also Example 1.36). According to Theorem 5.27 we may obtain a representation of the derivative of e^x by differentiating the Taylor series term by term:

$$\frac{d}{dx} \left(\sum_{n=0}^{\infty} \frac{1}{n!} x^n \right) = \sum_{n=1}^{\infty} \frac{1}{n!} \frac{d(x^n)}{dx} = \sum_{n=1}^{\infty} \frac{1}{n!} n x^{n-1} = \sum_{n=1}^{\infty} \frac{1}{(n-1)!} x^{n-1} = \sum_{m=0}^{\infty} \frac{1}{m!} x^m, \quad (5.32)$$

where we have shifted the index of summation in the last step by setting $m = n - 1$. We recover from (5.32) the well known formula $(e^x)' = e^x$. \square

Example 5.45 (primitive of $\cos x$)

In Example 1.28, we have seen that the Taylor series of $\cos x$ centred at $x_0 = 0$ is given by

$$\sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} x^{2k} \quad (5.33)$$

and that the series has the radius of convergence $\rho = \infty$ and converges to $\cos x$ for all $x \in \mathbb{R}$ (see also Example 1.36). According to Theorem 5.42, we may determine a primitive of $\cos x$ by taking the indefinite integral of (5.33) term by term.

$$\int_0^x \cos y \, dy = \int_0^x \left(\sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} y^{2k} \right) dy = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} \int_0^x y^{2k} \, dy$$

$$= \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} \frac{y^{2k+1}}{2k+1} \Big|_0^x = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} x^{2k+1} \quad (5.34)$$

which is the Taylor series of $\sin x$ centred at $x_0 = 0$. We know from Theorem 5.42 that the power series (5.34) of $\int_{x_0}^x \cos(y) dy$ has also the radius of convergence $\rho = \infty$ and thus converges for all $x \in \mathbb{R}$. Since we have recognized this series as the power series $\sin x$ which converges on \mathbb{R} to $\sin x$, we can deduce from (5.34) that $\sin x$ is a primitive of $\cos x$. \square

Example 5.46 (power series of $\ln(1+x)$ centred at $x_0 = 0$)

We want to derive a series representation for $\ln(1+x)$ with the help of Theorem 5.27. From the geometric series we know that for $|x| < 1$

$$\frac{1}{1+x} = \frac{1}{1-(-x)} = \sum_{n=0}^{\infty} (-x)^n = \sum_{n=0}^{\infty} (-1)^n x^n,$$

and the power series has the radius of convergence $\rho = 1$. According to Theorem 5.27, a primitive of $1/(1+x)$ is given by

$$\begin{aligned} \int_0^x \frac{1}{1+y} dy &= \int_0^x \left(\sum_{n=0}^{\infty} (-1)^n y^n \right) dy = \sum_{n=0}^{\infty} (-1)^n \int_0^x y^n dy \\ &= \sum_{n=0}^{\infty} (-1)^n \frac{y^{n+1}}{n+1} \Big|_0^x = \sum_{n=0}^{\infty} \frac{(-1)^n}{n+1} x^{n+1} = \sum_{m=1}^{\infty} \frac{(-1)^{m-1}}{m} x^m, \end{aligned} \quad (5.35)$$

where $x \in (-1, 1)$. Since $(\ln(1+x))' = 1/(1+x)$, $x \in (-1, 1)$, the function $\ln(1+x)$ is another primitive of $1/(1+x)$, and the primitive given by (5.35) equals $\ln(1+x) + C$, with some constant $C \in \mathbb{R}$, that is,

$$\sum_{m=1}^{\infty} \frac{(-1)^{m-1}}{m} x^m = \ln(1+x) + C, \quad (5.36)$$

where C is a real constant. We want to show that $C = 0$. If we evaluate (5.36) in $x = 0$, we obtain

$$0 = \ln(1) + C = C \quad \Rightarrow \quad C = 0.$$

Thus we obtain that

$$\sum_{m=1}^{\infty} \frac{(-1)^{m-1}}{m} x^m = \ln(1+x), \quad x \in (-1, 1). \quad \square$$

Example 5.47 (Taylor series of $\ln x$ centred at $x_0 = 1$)

In Example 1.29, we have seen that the power series of $\ln x$ centred at $x_0 = 1$ is given by

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

and we verified that its radius of convergence is given by $\rho = 1$, that is, the power series converges on $(0, 2)$. From Theorem 5.39, we know that this power series converges uniformly on any subinterval $[a, b] \subset (0, 2)$. With the help of Theorem 5.42, we can now show that this series converges for all $x \in (0, 2)$ pointwise to $\ln x$.

We start with the geometric series

$$\frac{1}{x} = \frac{1}{1 - (1-x)} = \sum_{n=0}^{\infty} (1-x)^n = \sum_{n=0}^{\infty} (-1)^n (x-1)^n, \quad |x-1| < 1,$$

with the radius of convergence $\rho = 1$. From Theorem 5.42, we know that a primitive of $1/x$, $x \in (0, 2)$, is given by

$$\begin{aligned} \int_1^x \frac{1}{y} dy &= \int_1^x \left(\sum_{n=0}^{\infty} (-1)^n (y-1)^n \right) dy = \sum_{n=0}^{\infty} \int_1^x (-1)^n (y-1)^n dy \\ &= \sum_{n=0}^{\infty} (-1)^n \frac{(y-1)^{n+1}}{n+1} \Big|_1^x = \sum_{n=0}^{\infty} \frac{(-1)^n}{n+1} (x-1)^{n+1} = \sum_{m=1}^{\infty} \frac{(-1)^{m-1}}{m} (x-1)^m, \end{aligned}$$

where $x \in (0, 2)$. Since we have $(\ln x)' = 1/x$ for all $x \in (0, 2)$, we see that

$$\sum_{m=1}^{\infty} \frac{(-1)^{m-1}}{m} (x-1)^m = \ln x + C, \quad x \in (0, 2), \quad (5.37)$$

with some constant $C > 0$. Substituting $x = 1$ in (5.37), we obtain

$$0 = \ln 1 + C = 0 + C = C.$$

Thus the constant C is zero and we obtain

$$\sum_{m=1}^{\infty} \frac{(-1)^{m-1}}{m} (x-1)^m = \ln x, \quad x \in (0, 2),$$

and we have verified that the Taylor series of $\ln x$ centred at $x_0 = 1$ converges for all $x \in (0, 2)$ to $\ln x$. \square

Finally we prove Lemma 5.41 and Theorem 5.42.

Proof of Lemma 5.41 The proof of this lemma is somewhat similar to that of the previous one. First, we observe that for $0 < \alpha < 1$

$$\sum_{n=1}^{\infty} n \alpha^n < \infty. \quad (5.38)$$

This can be seen as follows: From Theorem 4.38, we see that (5.38) holds true if and only if

$$\lim_{N \rightarrow \infty} \int_1^N x \alpha^x dx < \infty.$$

We work out the integral with integration by parts. With $F(x) := e^{x \ln \alpha} / \ln \alpha$, $G(x) := x$, and thus $F'(x) = e^{x \ln \alpha} = \alpha^x$ and $G'(x) = 1$, we have

$$\begin{aligned} \int_1^N x \alpha^x dx &= \int_1^N x e^{x \ln \alpha} dx = \frac{x e^{x \ln \alpha}}{\ln \alpha} \Big|_1^N - \int_1^N \frac{e^{x \ln \alpha}}{\ln \alpha} dx \\ &= \frac{N e^{N \ln \alpha}}{\ln \alpha} - \frac{e^{\ln \alpha}}{\ln \alpha} - \frac{e^{x \ln \alpha}}{(\ln \alpha)^2} \Big|_1^N = \frac{N e^{N \ln \alpha}}{\ln \alpha} - \frac{e^{\ln \alpha}}{\ln \alpha} - \frac{e^{N \ln \alpha}}{(\ln \alpha)^2} + \frac{e^{\ln \alpha}}{(\ln \alpha)^2} \\ &= \frac{N \alpha^N}{\ln \alpha} - \frac{\alpha}{\ln \alpha} - \frac{\alpha^N}{(\ln \alpha)^2} + \frac{\alpha}{(\ln \alpha)^2} \rightarrow -\frac{\alpha}{\ln \alpha} + \frac{\alpha}{(\ln \alpha)^2} \quad \text{as } N \rightarrow \infty, \end{aligned}$$

since $0 < \alpha < 1$. Thus we see from Theorem 4.38 that (5.38) holds true.

After this preparation we can give the proof that both series in the lemma have the same radius of convergence ρ as the series $\sum_{n=0}^{\infty} c_n (x - x_0)^n$. Let ρ_1 and ρ_2 denote the radius of convergence of the series

$$\sum_{n=1}^{\infty} c_n n (x - x_0)^{n-1} \quad \text{and} \quad \sum_{n=0}^{\infty} \frac{c_n}{n+1} (x - x_0)^{n+1}, \quad (5.39)$$

respectively. Then we need to show that $\rho_1 = \rho_2 = \rho$.

Let $x \in (x_0 - \rho, x_0 + \rho)$, that is, $|x - x_0| < \rho$, and we may assume that $x \neq x_0$. Then there exists $y \in (x_0 - \rho, x_0 + \rho)$ with $|x - x_0| < |y - x_0| < \rho$. Since the power series converges absolutely at y , we know that $\{|c_n (y - x_0)^n|\}$ tends to zero and, in particular, $|c_n (y - x_0)^n| \leq M$ for all $n \in \mathbb{N}_0$ with a constant $M > 0$. Thus for $x \neq x_0$

$$|c_n n (x - x_0)^{n-1}| = \frac{|c_n (y - x_0)^n|}{|x - x_0|} n \left(\frac{|x - x_0|}{|y - x_0|} \right)^n \leq \frac{M}{|x - x_0|} n \left(\frac{|x - x_0|}{|y - x_0|} \right)^n$$

and

$$\left| \frac{c_n (x - x_0)^{n+1}}{n+1} \right| \leq |c_n (x - x_0)^{n+1}|$$

$$\begin{aligned}
&= |c_n (y - x_0)^n| |x - x_0| \left(\frac{|x - x_0|}{|y - x_0|} \right)^n \\
&\leq M |x - x_0| \left(\frac{|x - x_0|}{|y - x_0|} \right)^n.
\end{aligned}$$

From (5.38), the geometric series and $|x - x_0|/|y - x_0| < 1$ we see now that for $x \neq x_0$

$$\begin{aligned}
\sum_{n=0}^{\infty} |c_n n (x - x_0)^{n-1}| &\leq \frac{M}{|x - x_0|} \sum_{n=0}^{\infty} n \left(\frac{|x - x_0|}{|y - x_0|} \right)^n < \infty, \\
\sum_{n=0}^{\infty} \left| \frac{c_n}{n+1} (x - x_0)^n \right| |x - x_0| &\leq M |x - x_0| \sum_{n=0}^{\infty} \left(\frac{|x - x_0|}{|y - x_0|} \right)^n < \infty.
\end{aligned}$$

Thus the series in (5.39) converge absolutely at the point x . Since $x \in (x_0 - \rho, x_0 + \rho)$ was arbitrary, we have shown that the series in (5.39) both converge absolutely for all x in $(x_0 - \rho, x_0 + \rho)$, and thus $\rho_1 \geq \rho$ and $\rho_2 \geq \rho$.

It remains to show that the series in (5.39) diverge for x with $|x - x_0| > \rho$, that is, that $\rho_1 \leq \rho$ and $\rho_2 \leq \rho$. We give the proof by contradiction: Assume that $\rho_1 > \rho$ or $\rho_2 > \rho$. Then there exist x_1 and y_1 with

$$\rho < |x_1 - x_0| < |y_1 - x_0| < \rho_1 \quad \Rightarrow \quad \frac{|x_1 - x_0|}{|y_1 - x_0|} < 1, \quad (5.40)$$

or there exist x_2 and y_2 with

$$\rho < |x_2 - x_0| < |y_2 - x_0| < \rho_2 \quad \Rightarrow \quad \frac{|x_2 - x_0|}{|y_2 - x_0|} < 1, \quad (5.41)$$

respectively. By the assumption, the first power series in (5.39) converges absolutely at x_1 and y_1 , or the second power series in (5.39) converges absolutely at x_2 and y_2 , respectively. In particular, this means that the sequence $\{|c_n n (y_1 - x_0)^{n-1}|\}$ or $\{c_n (y_2 - x_0)^{n+1}/(n+1)\}$, respectively, converges to zero and thus there exists a positive real constant M_1 or M_2 , respectively, such that

$$|c_n n (y_1 - x_0)^{n-1}| \leq M_1 \quad \text{for all } n \in \mathbb{N}_0, \quad \text{or} \quad (5.42)$$

$$\left| \frac{c_n}{n+1} (y_2 - x_0)^{n+1} \right| \leq M_2 \quad \text{for all } n \in \mathbb{N}_0, \quad (5.43)$$

respectively. Thus, using (5.42) and (5.43)

$$|c_n (x_1 - x_0)^n| = |c_n n (y_1 - x_0)^{n-1}| \frac{|y_1 - x_0|}{n} \left(\frac{|x_1 - x_0|}{|y_1 - x_0|} \right)^n$$

$$\leq M_1 |y_1 - x_0| \left(\frac{|x_1 - x_0|}{|y_1 - x_0|} \right)^n \quad \text{for all } n \in \mathbb{N}, \quad (5.44)$$

or

$$\begin{aligned} |c_n (x_2 - x_0)^n| &= \left| \frac{c_n}{n+1} (y_2 - x_0)^{n+1} \right| \frac{n+1}{|x_2 - x_0|} \left(\frac{|x_2 - x_0|}{|y_2 - x_0|} \right)^{n+1} \\ &\leq \frac{M_2}{|x_2 - x_0|} (n+1) \left(\frac{|x_2 - x_0|}{|y_2 - x_0|} \right)^{n+1} \quad \text{for all } n \in \mathbb{N}_0, \end{aligned} \quad (5.45)$$

respectively. Thus we see from (5.38), that (5.42) and (5.44), or (5.43) and (5.45), respectively, imply that the original series converges absolutely at x_1 or x_2 , respectively. Indeed, from (5.42) and (5.44),

$$\sum_{n=0}^{\infty} |c_n (x_1 - x_0)^n| \leq c_0 + M_1 |y_1 - x_0| \sum_{n=1}^{\infty} \underbrace{\left(\frac{|x_1 - x_0|}{|y_1 - x_0|} \right)^n}_{<1} < \infty,$$

where the convergence follows from the geometric series. Likewise, from (5.43) and (5.45),

$$\sum_{n=0}^{\infty} |c_n (x_2 - x_0)^n| \leq \frac{M_2}{|x_2 - x_0|} \sum_{n=0}^{\infty} (n+1) \underbrace{\left(\frac{|x_2 - x_0|}{|y_2 - x_0|} \right)^{n+1}}_{<1} < \infty,$$

where the convergence follows from (5.38).

The convergence of the original series $\sum_{n=0}^{\infty} c_n (x - x_0)^n$ at x_1 and x_2 , respectively, is a contradiction, since, from (5.40) and (5.41), $|x_1 - x_0| > \rho$ and $|x_2 - x_0| > \rho$, which means that the original series diverges at x_1 and x_2 . Thus we know that our assumption was wrong, and hence $\rho_1 \leq \rho$ and $\rho_2 \leq \rho$. \square

Proof of Theorem 5.42:

(i) This is the statement of Theorem 5.39.

(ii) Take any $\tilde{x} \in (x_0 - \rho, x_0 + \rho)$. Then there exists a bounded closed interval $[a, b] \subset (x_0 - \rho, x_0 + \rho)$ with $\tilde{x} \in [a, b]$. From (i), we know that the power series converges uniformly on $[a, b]$. The partial sums

$$s_m(x) := \sum_{n=0}^m c_n (x - x_0)^n \quad (5.46)$$

of the power series are polynomials and thus continuous. Thus the limit f of the power series is on $[a, b]$ the uniform limit of a sequence of continuous functions. From Theorem 5.19, we know that the uniform limit on $[a, b]$ of a sequence of continuous functions is continuous on $[a, b]$. Thus f is continuous on $[a, b]$ and in particular at the point \tilde{x} .

(iii) From Theorem 5.39, we know that the power series converges uniformly on $[a, b] \subset (x_0 - \rho, x_0 + \rho)$. Since the partial sums (5.46) are polynomials and hence Riemann integrable, we know from Theorem 5.27 (i) and (iii) that $f \in \mathcal{R}([a, b])$, and that we may interchange the sum and the integration. More precisely (since Theorem 5.27 (iii) was formulated for sequences of functions), we know that

$$\begin{aligned} \int_a^b \left(\sum_{n=0}^{\infty} c_n (x - x_0)^n \right) dx &= \int_a^b \left(\lim_{m \rightarrow \infty} \sum_{n=0}^m c_n (x - x_0)^n \right) dx \\ &= \lim_{m \rightarrow \infty} \int_a^b \left(\sum_{n=0}^m c_n (x - x_0)^n \right) dx. \end{aligned} \quad (5.47)$$

We evaluate the integrals on the right-hand side and obtain

$$\begin{aligned} \lim_{m \rightarrow \infty} \int_a^b \left(\sum_{n=0}^m c_n (x - x_0)^n \right) dx &= \lim_{m \rightarrow \infty} \sum_{n=0}^m c_n \int_a^b (x - x_0)^n dx \\ &= \lim_{m \rightarrow \infty} \sum_{n=0}^m c_n \left. \frac{(x - x_0)^{n+1}}{n+1} \right|_a^b = \sum_{n=0}^{\infty} \frac{c_n}{n+1} [(b - x_0)^{n+1} - (a - x_0)^{n+1}]. \end{aligned} \quad (5.48)$$

The formulas (5.47) and (5.48) now yield (5.29).

(iv) Since f is continuous on $(x_0 - \rho, x_0 + \rho)$ (due to (ii)), we know that

$$F(x) := \int_{x_0}^x f(y) dy = \int_{x_0}^x \left(\sum_{n=0}^{\infty} c_n (y - x_0)^n \right) dy, \quad x \in (x_0 - \rho, x_0 + \rho),$$

defines a primitive of f . We may assume without loss of generality that $x > x_0$. Then the power series converges uniformly on $[x_0, x] \subset (x_0 - \rho, x_0 + \rho)$, and from (iii), we know that

$$F(x) = \int_{x_0}^x \left(\sum_{n=0}^{\infty} c_n (y - x_0)^n \right) dy = \sum_{n=0}^{\infty} c_n \int_{x_0}^x (y - x_0)^n dy = \sum_{n=0}^{\infty} \frac{c_n}{n+1} (x - x_0)^{n+1},$$

which implies (5.30).

(v) Take any $x \in (x_0 - \rho, x_0 + \rho)$. Then there exists a closed interval $[a, b] \subset (x_0 - \rho, x_0 + \rho)$ with $x \in (a, b)$. From Theorem 5.39, we know that the power series converges uniformly on $[a, b]$. We know that all the partial sums

$$s_m(x) = \sum_{n=0}^m c_n (x - x_0)^n$$

are continuously differentiable (since they are polynomials), and from Lemma 5.41 we know the sequence $\{s'_m\}$ of the derivatives

$$s'_m(x) = \sum_{n=0}^m c_n n (x - x_0)^{n-1} = \sum_{n=1}^m c_n n (x - x_0)^{n-1}$$

of the partial sums s_m converges uniformly on $[a, b]$ to

$$\sum_{n=1}^{\infty} c_n n (x - x_0)^{n-1}.$$

Thus from Theorem 5.32, we know that

$$\sum_{n=1}^{\infty} c_n n (x - x_0)^{n-1} = \frac{d}{dx} \left(\sum_{n=0}^{\infty} c_n (x - x_0)^n \right) = \frac{df(x)}{dx}$$

which verifies (iv). □

Chapter 6

Metric Spaces and Normed Linear Spaces

In Section 6.1, we introduce **metric spaces** and **normed linear spaces**. An elementary example of a normed linear space are the real numbers \mathbb{R} with the absolute value $|x|$ as norm. The most important example of a metric and normed linear space is the Euclidean space \mathbb{R}^n with the Euclidean norm

$$\|\mathbf{x}\|_2 = \|(x_1, x_2, \dots, x_n)\|_2 := \sqrt{\sum_{j=1}^n x_j^2} = (x_1^2 + x_2^2 + \dots + x_n^2)^{1/2},$$

where we use the notation $\mathbf{x} := (x_1, x_2, \dots, x_n)$. The Euclidean norm induces the Euclidean **metric** (or **distance function**)

$$d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|_2 = \left(\sum_{j=1}^n (x_j - y_j)^2 \right)^{1/2}.$$

In Subsection 6.1.4, we will discuss a special class of normed linear spaces, namely those whose norm is generated by an **inner product**. Our most prominent example of such a normed linear space is again \mathbb{R}^n and the inner product is the usual Euclidean scalar product

$$(\mathbf{x}, \mathbf{y}) := \mathbf{x}^T \cdot \mathbf{y} = \sum_{j=1}^n x_j y_j = x_1 y_1 + x_2 y_2 + \dots + x_n y_n.$$

The Euclidean scalar product of \mathbb{R}^n generates the norm $\|\cdot\|_2$ of \mathbb{R}^n via the relation

$$\|\mathbf{x}\|_2 = \sqrt{(\mathbf{x}, \mathbf{x})}.$$

We will also call a space whose norm is generated by an inner product an **inner product space**. Not all normed linear spaces have a norm which is generated by an inner product. In an inner product space the so-called **Schwarz inequality** holds. For \mathbb{R}^n with the Euclidean inner product, you will most likely have seen the Schwarz inequality in first year: it reads

$$|(\mathbf{x}, \mathbf{y})| = \left| \sum_{j=1}^n x_j y_j \right| \leq \sqrt{\sum_{j=1}^n x_j^2} \sqrt{\sum_{j=1}^n y_j^2} = \|\mathbf{x}\|_2 \|\mathbf{y}\|_2.$$

In Section 6.2, we discuss **sequences** and **Cauchy sequences** in metric spaces and normed linear spaces. For the elementary example of the real numbers \mathbb{R} with the absolute value, these new general definitions of convergence and Cauchy sequences will just lead to the notions of convergence and Cauchy sequences that you have learnt in first year. For the real numbers \mathbb{R} , you have learnt in first year that every Cauchy sequence converges, and you may remember that this was described by saying that the real numbers are complete. As for the case of the real numbers, we will discuss for various examples of normed linear spaces and metric spaces whether all Cauchy sequences converge (to a limit in that normed linear space). If this is the case then we call the normed linear space or metric space **complete**. We also define **bounded sets** in metric spaces normed linear spaces, and we will prove the **Bolzan-Weierstrass theorem** for \mathbb{R}^n with the Euclidean norm.

In Section 6.3, we finally introduce the notions of **open sets** and **closed sets** in metric spaces and normed linear spaces. Although these definitions are rather abstract, we find that they are in correspondence with our intuitive use of the terms open and closed for real numbers: open intervals (a, b) are open and closed intervals $[a, b]$ are closed.

All definitions that we encounter in this chapter will **extend terminology** that you may have already seen for the real numbers \mathbb{R} and possibly for \mathbb{R}^n **to a more general setting: These more general definitions will include the examples that you have already encountered as special cases.**

You may wonder now whether this chapter is not topic-wise somewhat separate from the rest of the course material. However, this is not the case, since much of the material from Chapter 5 contributes to this chapter: for example, in Chapter 5, we have introduced the supremum norm

$$\|f\|_\infty := \sup_{x \in \langle c, d \rangle} |f(x)|, \quad f \in \mathcal{B}(\langle c, d \rangle),$$

for bounded functions on a interval $\langle c, d \rangle$. We will see in this chapter that the supremum norm is indeed a norm. We will show that uniform convergence is just

convergence (in the sense to be defined in this chapter) for the space of continuous functions $\mathcal{C}([a, b])$ equipped with the supremum norm $\|\cdot\|_\infty$. In Chapter 5, we have also defined uniform Cauchy sequences, and we will see that they are just Cauchy sequences (in the sense to be defined in this chapter) for $\mathcal{C}([a, b])$ equipped with the supremum norm $\|\cdot\|_\infty$. In Chapter 5, we learnt that every uniformly convergent sequence $\{f_n\}$ of continuous functions $f_n \in \mathcal{C}([a, b])$ has a continuous limit. This implies that $\mathcal{C}([a, b])$ with the supremum norm is a complete normed linear space in the sense to be defined in this chapter.

As illustrated for the example of uniform convergence, the topics discussed in this chapter will help us to get a much better understanding of many concepts discussed in this course and also in your first year and other second year mathematics courses: they will allow us to **see several individual concepts as examples and variants of more general concepts**; these general concepts are **metric spaces, normed linear spaces, and inner product spaces, convergence, Cauchy sequences, completeness, and open sets, closed sets**, as well as related notions. These concepts are also of fundamental importance for many courses that you might choose in third year, and it is very important to understand these concepts thoroughly!

6.1 Metric Spaces and Normed Linear Spaces

In Subsection 6.1.1, we introduce the notation a **metric** and a **metric space**. Roughly speaking a metric is a function that measures distances and has some specific properties. Then we introduce the notion of a **norm** and a **normed linear space**. **A norm induces a metric**, but the reverse is **not** true. We discuss various elementary examples of normed linear spaces and metric spaces. In Subsection 6.1.2, we discuss the **Euclidean space** \mathbb{R}^n with the Euclidean norm. However, we will also equip \mathbb{R}^n with various other norms and verify the norm properties. A given linear space can have several different norms, and, likewise on a given set, we can define several different metrics (distance functions). In Subsection 6.1.3, we discuss more complicated examples of normed linear spaces, namely **function spaces** such as $\mathcal{B}(\langle c, d \rangle)$ and $\mathcal{C}([a, b])$ with different norms. Here we will encounter the supremum norm again and we will also discuss the important L_1 -norm and L_2 -norm. In Subsection 6.1.4, we finally introduce the notion of an **inner product** and an **inner product space**. We will see that some of the normed linear spaces that we have encountered so far are examples of inner product spaces, and we will show that an **inner product induces a norm**. We will also prove the **Schwarz inequality** for an arbitrary inner product space.

6.1.1 Definitions and Basic Examples

On the real line \mathbb{R} we measure the distance $d(x, y)$ of two points x and y in \mathbb{R} by

$$d(x, y) := |x - y|.$$

This is the simplest example of a **metric** or **distance function**.

Definition 6.1 (metric space)

A metric space (X, d) is a non-empty set X with a **metric** (or **distance function**) $d : X \times X \rightarrow \mathbb{R}$ satisfying

- (i) $d(x, y) \geq 0$ for all $x, y \in X$; and $d(x, y) = 0$ if and only if $x = y$.
- (ii) $d(x, y) = d(y, x)$ for all $x, y \in X$ (**symmetry**).
- (iii) $d(x, y) \leq d(x, z) + d(z, y)$ for all $x, y, z \in X$ (**triangle inequality**).

Example 6.2 (metric space \mathbb{R} with absolute value metric $d(x, y) := |x - y|$)

The real line \mathbb{R} with $d : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$, defined by $d(x, y) := |x - y|$, is a metric space (where $|\cdot|$ is the usual absolute value).

Proof: We have to verify the conditions (i) to (iii). Clearly $d(x, y) = |x - y| \geq 0$ for all $x, y \in \mathbb{R}$; and $d(x, y) = |x - y| = 0$ if and only if $x = y$. Also, d is symmetric, since $d(x, y) = |x - y| = |y - x| = d(y, x)$ for all $x, y \in \mathbb{R}$. Finally, by using the triangle inequality $|a + b| \leq |a| + |b|$ for real numbers $a, b \in \mathbb{R}$, we have

$$d(x, y) = |x - y| = |(x - z) + (z - y)| \leq |x - z| + |z - y| = d(x, z) + d(z, y)$$

for all $x, y, z \in \mathbb{R}$, which proves the triangle inequality for \mathbb{R} with the absolute value metric $d(x, y) = |x - y|$. \square

Instead of the real line we can also consider an interval with the distance function $d(x, y) = |x - y|$.

Example 6.3 (metric space $(0, 1]$ with $d(x, y) := |x - y|$)

The interval $(0, 1] \subset \mathbb{R}$ with the distance function $d : (0, 1] \times (0, 1] \rightarrow \mathbb{R}$, defined by $d(x, y) = |x - y|$, is a metric space.

Proof: Let us verify the conditions (i) to (iii). Clearly $d(x, y) = |x - y| \geq 0$ for all $x, y \in (0, 1]$; and $d(x, y) = |x - y| = 0$ if and only if $x = y$. Also, $d(x, y) = |x - y| = |y - x| = d(y, x)$ for all $x, y \in (0, 1]$. Finally, by using the triangle inequality $|a + b| \leq |a| + |b|$ for real numbers $a, b \in \mathbb{R}$, we have

$$d(x, y) = |x - y| = |(x - z) + (z - y)| \leq |x - z| + |z - y| = d(x, z) + d(z, y)$$

for all $x, y, z \in (0, 1]$, which proves the triangle inequality for $(0, 1]$ with the metric $d(x, y) = |x - y|$. \square

The so-called **discrete metric** from the next example is a metric which is contrary to our everyday experience of distances: Either two points have distance $d(x, y) = 0$, and then $x = y$, or alternatively two points $x \neq y$ have distance $d(x, y) = 1$. The discrete metric assumes no other values apart from the **discrete values** 0 and 1.

Example 6.4 (\mathbb{R} with the discrete metric)

The real line \mathbb{R} with the **discrete metric**

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

is a metric space.

Proof: By the definition of the discrete metric, $d(x, y) \geq 0$ for all $x, y \in \mathbb{R}$; and $d(x, y) = 0$ if and only if $x = y$. Let $x, y \in \mathbb{R}$ and $x \neq y$. Then $d(x, y) = 1$ and $d(y, x) = 1$. Thus we have symmetry $d(x, y) = d(y, x)$ for all $x, y \in \mathbb{R}$.

Finally, we need to verify the triangle inequality, that is, we have to show

$$d(x, y) \leq d(x, z) + d(z, y) \quad \text{for all } x, y, z \in \mathbb{R}. \quad (6.1)$$

If $x = y$, then the left-hand side is zero, and since the metric has only non-negative values, the right-hand side is non-negative. Thus the inequality (6.1) is true if $x = y$. Now assume that $x \neq y$. Then the left-hand side has the value $d(x, y) = 1$, and the point z has to differ from either x or y (or from both). Thus we find that $d(x, z) = 1$ or $d(z, y) = 1$. Since the metric assumes only non-negative values, the value of the right-hand side is ≥ 1 . Thus the inequality (6.1) is true if $x \neq y$. \square

A very important class of metric spaces are the so-called **normed linear spaces**. While it is not obvious from the definition of a normed linear space below, we will show later-on that **every normed linear space is also a metric space with a metric that is defined in terms of the norm**.

Definition 6.5 (normed linear space)

A (real) **normed linear space** $(X, \|\cdot\|)$ is a real linear space (or real vector space) X with a **norm** $\|\cdot\| : X \rightarrow \mathbb{R}$ satisfying:

- (i) $\|x\| \geq 0$ for all $x \in X$; and $\|x\| = 0$ if and only if $x = \mathcal{O}$.
- (ii) $\|\alpha x\| = |\alpha| \|x\|$ for all $x \in X$ and all $\alpha \in \mathbb{R}$.
- (iii) $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in X$ (**triangle inequality**).

Remember that \mathcal{O} denotes the **neutral element** (or **zero vector**) in the linear space X .

We show that \mathbb{R} with the absolute value is a normed linear space.

Example 6.6 (\mathbb{R} with absolute value norm)

The space \mathbb{R} with the absolute value norm $|x|$ is a normed linear space.

Proof: From the definition of the absolute value, we have $|x| \geq 0$ for all $x \in \mathbb{R}$; and $|x| = 0$ if and only if $x = 0$. Likewise,

$$|\alpha x| = |\alpha| |x| \quad \text{for all } x \in \mathbb{R} \text{ and for all } \alpha \in \mathbb{R}.$$

The triangle inequality is just the usual triangle inequality for real numbers:

$$|x + y| \leq |x| + |y| \quad \text{for all } x, y \in \mathbb{R}. \quad \square$$

Here is an example of a set with the absolute value norm which is **not** a normed linear space.

Example 6.7 ($(0, 1]$ with absolute value is not a normed linear space)

The interval $(0, 1]$ with the absolute value norm $|x|$ is not a normed linear space.

Proof: The interval $(0, 1] \subset \mathbb{R}$ is not a linear space. For example, the closure is violated: $1 \in (0, 1]$ and $1 + 1 = 2$, but $2 \notin (0, 1]$. \square

Note that, unlike in the definition of metric spaces, we do assume in the definition of a normed linear space $(X, \|\cdot\|)$ that X is a **linear space** or **vector space**! If this is not the case than $(X, \|\cdot\|)$ cannot be a normed linear space and you need not check whether the norm properties are satisfied.

Now we show that any norm induces a metric, and thus any normed linear space is in particular also a metric space with this induced metric.

Lemma 6.8 (normed linear space \Rightarrow metric space)

A normed linear space $(X, \|\cdot\|)$ with the norm $\|\cdot\| : X \rightarrow \mathbb{R}$ is a **metric space** with the **distance function** $d : X \times X \rightarrow \mathbb{R}$, given by

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

Proof of Lemma 6.8: We have to verify that $d(x, y) = \|x - y\|$ satisfies the three properties of a metric.

(i) From the norm property (i), clearly $d(x, y) = \|x - y\| \geq 0$; and we have $d(x, y) = \|x - y\| = 0$ if and only if $x - y = \mathcal{O}$, or equivalently, if $x = y$.

(ii) From the norm property (ii), we see that for all $x, y \in X$

$$d(x, y) = \|x - y\| = \|(-1)(y - x)\| = |-1| \|y - x\| = \|y - x\| = d(y, x).$$

Thus d is symmetric.

(iii) The triangle inequality (iii) of the norm implies that

$$d(x, y) = \|x - y\| = \|(x - z) + (z - y)\| \leq \|x - z\| + \|z - y\| = d(x, z) + d(z, y)$$

for all $x, y, z \in X$.

Thus $d(x, y) = \|x - y\|$ is indeed a metric for X . □

Example 6.9 (normed linear space is also metric space)

We have seen that \mathbb{R} with the absolute value norm $|x|$ is a normed linear space. According to Lemma 6.8 we have that

$$d(x, y) := |x - y|$$

defines a metric for \mathbb{R} . This has already been verified in Example 6.2 at the beginning of this subsection. □

6.1.2 Norms on \mathbb{R}^n

In this subsection we will consider the linear space \mathbb{R}^n of n -tuples $\mathbf{x} = (x_1, x_2, \dots, x_n)$ or real numbers. We will equip \mathbb{R}^n with various norms. The most common example that you will already have encountered in first year is the **Euclidean norm**.

Notation: A vector in \mathbb{R}^n is denoted by $\mathbf{x} = (x_1, x_2, \dots, x_n)$, where the components x_1, x_2, \dots, x_n are real numbers. We have

$$\mathbb{R}^n := \{\mathbf{x} = (x_1, x_2, \dots, x_n) : x_1, x_2, \dots, x_n \in \mathbb{R}\}.$$

If $n = 1$, we get the real line \mathbb{R} , and the superscript 1 is usually dropped, that is, we write \mathbb{R} instead of \mathbb{R}^1 .

Definition 6.10 (Euclidean norm and Euclidean metric)

The **Euclidean norm** on \mathbb{R}^n is defined by

$$\|\mathbf{x}\|_2 = \left(\sum_{k=1}^n x_k^2 \right)^{1/2} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}, \quad \mathbf{x} \in \mathbb{R}^n.$$

The **Euclidean distance** (or **Euclidean metric**) of any two vectors $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{y} \in \mathbb{R}^n$ is then defined by

$$\begin{aligned} d_2(\mathbf{x}, \mathbf{y}) &:= \|\mathbf{x} - \mathbf{y}\|_2 = \left(\sum_{k=1}^n (x_k - y_k)^2 \right)^{1/2} \\ &= \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_n - y_n)^2}. \end{aligned}$$

The next lemma states that \mathbb{R}^n with the Euclidean norm $\|\cdot\|_2$ is a normed linear space.

Lemma 6.11 ($(\mathbb{R}^n, \|\cdot\|_2)$ is a normed linear space)

The vector space \mathbb{R}^n with the Euclidean norm $\|\mathbf{x}\|_2 := (\sum_{k=1}^n x_k^2)^{1/2}$ is a normed linear space.

To prove the triangle inequality for $(\mathbb{R}^n, \|\cdot\|_2)$, we need the following lemma.

Lemma 6.12 (Schwarz inequality for \mathbb{R}^n)

For any vectors $\mathbf{x} = (x_1, x_2, \dots, x_n)$ and $\mathbf{y} = (y_1, y_2, \dots, y_n)$ in \mathbb{R}^n ,

$$\left| \sum_{j=1}^n x_j y_j \right| \leq \sqrt{\sum_{j=1}^n x_j^2} \sqrt{\sum_{j=1}^n y_j^2} = \|\mathbf{x}\|_2 \|\mathbf{y}\|_2 \quad (6.2)$$

Proof of Lemma 6.12: For $\mathbf{x} = \mathbf{0}$, the inequality is obviously true, since both sides of (6.2) are zero. Now we assume that $\mathbf{x} \neq \mathbf{0}$ and consider the quadratic function $\lambda \mapsto f(\lambda)$, defined by

$$f(\lambda) := \|\mathbf{y} - \lambda \mathbf{x}\|_2^2, \quad \lambda \in \mathbb{R},$$

and expand it in terms of λ :

$$f(\lambda) = \|\mathbf{y} - \lambda \mathbf{x}\|_2^2 = \sum_{j=1}^n (y_j - \lambda x_j)^2 = \sum_{j=1}^n (y_j^2 - 2\lambda x_j y_j + \lambda^2 x_j^2)$$

$$= \lambda^2 \sum_{j=1}^n x_j^2 - 2\lambda \sum_{j=1}^n x_j y_j + \sum_{j=1}^n y_j^2 = A\lambda^2 - 2B\lambda + C \geq 0, \quad (6.3)$$

with

$$A = \sum_{j=1}^n x_j^2 = \|\mathbf{x}\|_2^2, \quad C = \sum_{j=1}^n y_j^2 = \|\mathbf{y}\|_2^2, \quad B = \sum_{j=1}^n x_j y_j.$$

We may write the right-hand side of (6.3) as

$$f(\lambda) = A\lambda^2 - 2B\lambda + C = A \left[\lambda^2 - 2\lambda \frac{B}{A} + \frac{C}{A} \right] = A \left[\left(\lambda - \frac{B}{A} \right)^2 + \left(\frac{C}{A} - \frac{B^2}{A^2} \right) \right].$$

Since the quadratic function $f(\lambda) = \|\mathbf{y} - \lambda \mathbf{x}\|_2^2$ is non-negative, and, since $A > 0$ (from $\mathbf{x} \neq 0$), we conclude that

$$0 \leq f(\lambda) = A \left[\left(\lambda - \frac{B}{A} \right)^2 + \left(\frac{C}{A} - \frac{B^2}{A^2} \right) \right] \quad \text{for all } \lambda \in \mathbb{R}.$$

Now we choose $\lambda := B/A$ and have (using $A = \|\mathbf{x}\|_2^2 > 0$)

$$\begin{aligned} 0 \leq f\left(\frac{B}{A}\right) &= A \left[\left(\frac{B}{A} - \frac{B}{A} \right)^2 + \left(\frac{C}{A} - \frac{B^2}{A^2} \right) \right] = A \left(\frac{C}{A} - \frac{B^2}{A^2} \right) \\ \Rightarrow \quad 0 \leq \left(\frac{C}{A} - \frac{B^2}{A^2} \right) &= \frac{1}{A^2} (AC - B^2) \quad \Leftrightarrow \quad B^2 \leq AC. \end{aligned}$$

Substituting A, B, C by their definitions and taking the square root yields the Schwarz inequality. \square

Now we can prove the triangle inequality for $(\mathbb{R}^n, \|\cdot\|_2)$

Lemma 6.13 (triangle inequality for the Euclidean norm)

Consider \mathbb{R}^n with the Euclidean norm $\|\mathbf{x}\|_2 = (\sum_{k=1}^n x_k^2)^{1/2}$. For $(\mathbb{R}^n, \|\cdot\|_2)$ the **triangle inequality** holds true, that is,

$$\|\mathbf{x} + \mathbf{y}\|_2 \leq \|\mathbf{x}\|_2 + \|\mathbf{y}\|_2 \quad \text{for all } \mathbf{x}, \mathbf{y} \in \mathbb{R}^n.$$

Proof of Lemma 6.13: We have from the Schwarz inequality that

$$\begin{aligned} \|\mathbf{x} + \mathbf{y}\|_2^2 &= \sum_{j=1}^n (x_j + y_j)^2 = \sum_{j=1}^n (x_j^2 + 2x_j y_j + y_j^2) = \sum_{j=1}^n x_j^2 + 2 \sum_{j=1}^n x_j y_j + \sum_{j=1}^n y_j^2 \\ &\leq \|\mathbf{x}\|_2^2 + 2 \left| \sum_{j=1}^n x_j y_j \right| + \|\mathbf{y}\|_2^2 \leq \|\mathbf{x}\|_2^2 + 2 \|\mathbf{x}\|_2 \|\mathbf{y}\|_2 + \|\mathbf{y}\|_2^2 = (\|\mathbf{x}\|_2 + \|\mathbf{y}\|_2)^2. \end{aligned}$$

Taking the square-root, we see that the triangle inequality is satisfied. \square

After these preparations we can finally prove that \mathbb{R}^n equipped with the Euclidean norm is a normed linear space.

Proof of Lemma 6.11: We have to verify the three properties of a norm

(i) For $\mathbf{x} \in \mathbb{R}^n$, we have

$$\|\mathbf{x}\|_2 = \sqrt{\sum_{k=1}^n x_k^2} \geq 0$$

(since the square-root of a non-negative value is non-negative). We also see that $\|\mathbf{x}\|_2 = 0$ if and only if $x_k^2 = 0$ for all $k = 1, 2, \dots, n$, that is, if and only if $x_k = 0$ for all $k = 1, 2, \dots, n$. Thus $\|\mathbf{x}\|_2 = 0$ if and only if $\mathbf{x} = 0$.

(ii) Let $\alpha \in \mathbb{R}$ and $\mathbf{x} \in \mathbb{R}^n$. Then, using $\alpha^2 = |\alpha|^2$,

$$\|\alpha \mathbf{x}\|_2 = \sqrt{\sum_{k=1}^n (\alpha x_k)^2} = \sqrt{\sum_{k=1}^n \alpha^2 x_k^2} = \sqrt{|\alpha|^2 \sum_{k=1}^n x_k^2} = |\alpha| \sqrt{\sum_{k=1}^n x_k^2} = |\alpha| \|\mathbf{x}\|_2.$$

(iii) The triangle inequality was proved in Lemma 6.13.

Thus \mathbb{R}^n with the Euclidean norm $\|\cdot\|_2$ is indeed a normed vector space. \square

We have seen in Examples 6.2 and 6.4 before that it is possible to define different metrics on the same set. We had \mathbb{R} equipped with the absolute value metric (see Example 6.2) and the discrete metric (see Example 6.4). It is also possible to **define different norms for the same linear space**.

Example 6.14 (different norms for \mathbb{R}^n)

The following functions define all norms on \mathbb{R}^n :

(a) the **1-norm**

$$\|\mathbf{x}\|_1 := \sum_{k=1}^n |x_k| \quad \text{for } \mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n,$$

(b) for any fixed $p \in \mathbb{N}$, the **p -norm**

$$\|\mathbf{x}\|_p := \left(\sum_{k=1}^n |x_k|^p \right)^{1/p} \quad \text{for } \mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n,$$

(c) the **infinity norm** (∞ -norm)

$$\|\mathbf{x}\|_{\infty} := \max_{1 \leq k \leq n} |x_k| \quad \text{for } \mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n.$$

Note that the p -norm includes both the case of the Euclidean norm $\|\cdot\|_2$ for $p = 2$ and the 1-norm $\|\cdot\|_1$ for $p = 1$.

Proof: In each case we have to verify the three properties of a norm.

(a) (i) Clearly $\|\mathbf{x}\|_1 \geq 0$ for all $\mathbf{x} \in \mathbb{R}^n$. If $\mathbf{x} = \mathbf{0}$ then $\|\mathbf{x}\|_1 = 0$. If $\|\mathbf{x}\|_1 = \sum_{k=1}^n |x_k| = 0$ then $|x_k| = 0$ for all $k = 1, 2, \dots, n$, and thus $x_k = 0$ for all $k = 1, 2, \dots, n$, that is, $\mathbf{x} = \mathbf{0}$. Thus $\|\mathbf{x}\|_1 = 0$ if and only if $\mathbf{x} = \mathbf{0}$.

(ii) For $\alpha \in \mathbb{R}$ and $\mathbf{x} \in \mathbb{R}^n$, we have

$$\|\alpha \mathbf{x}\|_1 = \sum_{k=1}^n |\alpha x_k| = \sum_{k=1}^n |\alpha| |x_k| = |\alpha| \sum_{k=1}^n |x_k| = |\alpha| \|\mathbf{x}\|_1.$$

(iii) The triangle inequality follows from the triangle inequality for the absolute value. For $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$,

$$\|\mathbf{x} + \mathbf{y}\|_1 = \sum_{k=1}^n |x_k + y_k| \leq \sum_{k=1}^n (|x_k| + |y_k|) = \sum_{k=1}^n |x_k| + \sum_{k=1}^n |y_k| = \|\mathbf{x}\|_1 + \|\mathbf{y}\|_1.$$

Thus $\|\cdot\|_1$ is a norm for \mathbb{R}^n .

(b) That $\|\cdot\|_p$ is a norm for \mathbb{R}^n for $p \neq 1, 2$ is not so easy to prove and we will not discuss this case here.

(c) (i) By its definition, we see that $\|\mathbf{x}\|_{\infty} \geq 0$. If $\mathbf{x} = \mathbf{0}$ then $\|\mathbf{x}\|_{\infty} = 0$; and if $\|\mathbf{x}\|_{\infty} = \max_{1 \leq k \leq n} |x_k| = 0$ then $|x_k| = 0$ for all $k = 1, 2, \dots, n$, and thus $x_k = 0$ for all $k = 1, 2, \dots, n$, that is, $\mathbf{x} = \mathbf{0}$. Hence $\|\mathbf{x}\|_{\infty} = 0$ if and only if $\mathbf{x} = \mathbf{0}$.

(ii) For $\alpha \in \mathbb{R}$ and $\mathbf{x} \in \mathbb{R}^n$,

$$\|\alpha \mathbf{x}\|_{\infty} = \max_{1 \leq k \leq n} |\alpha x_k| = \max_{1 \leq k \leq n} |\alpha| |x_k| = |\alpha| \max_{1 \leq k \leq n} |x_k| = |\alpha| \|\mathbf{x}\|_{\infty}.$$

(iii) We prove that the triangle inequality holds. Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ be arbitrary. By definition we have $\|\mathbf{x} + \mathbf{y}\|_{\infty} = \max_{1 \leq k \leq n} |x_k + y_k|$. Suppose the maximum is reached for $k = k_0$; we then have

$$\begin{aligned} \|\mathbf{x} + \mathbf{y}\|_{\infty} &= \max_{1 \leq k \leq n} |x_k + y_k| = |x_{k_0} + y_{k_0}| \\ &\leq |x_{k_0}| + |y_{k_0}| \leq \max_{1 \leq k \leq n} |x_k| + \max_{1 \leq k \leq n} |y_k| = \|\mathbf{x}\|_{\infty} + \|\mathbf{y}\|_{\infty}, \end{aligned}$$

where we have used the triangle inequality for real numbers.

Thus $\|\cdot\|_{\infty}$ is a norm for the vector space \mathbb{R}^n . □

6.1.3 Spaces of Functions With Various Norms

A nontrivial example of a normed linear space is the linear space $\mathcal{C}([a, b])$ of continuous functions on $[a, b] \subset \mathbb{R}$ equipped with the supremum norm.

Lemma 6.15 (space of continuous functions with various norms)

Let $[a, b]$ be a closed and bounded interval. The linear space $\mathcal{C}([a, b])$ of all continuous functions $f : [a, b] \rightarrow \mathbb{R}$ with the **pointwise addition**

$$(f + g)(x) := f(x) + g(x), \quad x \in [a, b],$$

and the **pointwise scalar multiplication**

$$(\alpha f)(x) := \alpha f(x), \quad x \in [a, b],$$

forms a linear space $\mathcal{C}([a, b])$. The following are norms of $\mathcal{C}([a, b])$:

(a) the **supremum norm**

$$\|f\|_{\infty} := \sup_{x \in [a, b]} |f(x)|,$$

(b) the **L_1 -norm**

$$\|f\|_1 := \int_a^b |f(x)| dx,$$

(c) the **L_2 -norm**

$$\|f\|_2 := \left(\int_a^b |f(x)|^2 dx \right)^{1/2}.$$

Note that we can define the norms $\|\cdot\|_1$ and $\|\cdot\|_2$ with the Riemann integral over $[a, b]$ since all continuous functions on $[a, b]$ are **Riemann integrable**.

Proof of Lemma 6.15: We have to show that in each case the three properties of a norm are satisfied.

(a) (i) Since $|f(x)| \geq 0$ for all $x \in [a, b]$ (by the definition of the absolute value), we have $\|f\|_{\infty} \geq 0$ for all $f \in \mathcal{C}([a, b])$. The condition

$$0 = \|f\|_{\infty} = \sup_{x \in [a, b]} |f(x)|$$

implies that $|f(x)| = 0$ for all $x \in [a, b]$, thus $f(x) = 0$ for all $x \in [a, b]$. On the other hand, if $f(x) = 0$ for all $x \in [a, b]$, then $\|f\|_{\infty} = 0$. Thus $\|f\|_{\infty} = 0$ if and only if $f(x) = 0$ for all $x \in [a, b]$.

(ii) Let $\alpha \in \mathbb{R}$ and $f \in \mathcal{C}([a, b])$. Then

$$\|\alpha f\|_\infty = \sup_{x \in [a, b]} |\alpha f(x)| = \sup_{x \in [a, b]} |\alpha| |f(x)| = |\alpha| \sup_{x \in [a, b]} |f(x)| = |\alpha| \|f\|_\infty.$$

(iii) Let $f, g \in \mathcal{C}([a, b])$. For any fixed $x \in [a, b]$, we have, from the triangle inequality for real numbers,

$$|f(x) + g(x)| \leq |f(x)| + |g(x)| \leq \sup_{y \in [a, b]} |f(y)| + \sup_{z \in [a, b]} |g(z)| = \|f\|_\infty + \|g\|_\infty.$$

Thus the right hand side is an upper bound for $|f(x) + g(x)|$ for all $x \in [a, b]$. Taking the supremum over $x \in [a, b]$ yields

$$\|f + g\|_\infty = \sup_{x \in [a, b]} |f(x) + g(x)| \leq \|f\|_\infty + \|g\|_\infty.$$

Thus the supremum norm is a norm for $\mathcal{C}([a, b])$.

(b) (i) From the first order property of the Riemann integral (see Theorem 3.29) we know that

$$\|f\|_1 = \int_0^1 |f(x)| dx \geq 0,$$

since $|f(x)| \geq 0$ for all $x \in [a, b]$. If $f(x) = 0$ for all $x \in [a, b]$ then $\|f\|_1 = 0$. From the third order property of the Riemann integral, we see that, for any $f \in \mathcal{C}([a, b])$, $\|f\|_1 = 0$ implies that $|f(x)| = 0$ for all $x \in [a, b]$, and hence $f(x) = 0$ for all $x \in [a, b]$. Thus $\|f\|_1 = 0$ if and only if $f(x) = 0$ for all $x \in [a, b]$.

(ii) Let $\alpha \in \mathbb{R}$ and $f \in \mathcal{C}([a, b])$. Then from the linear properties of the Riemann integral (see Theorem 3.29)

$$\|\alpha f\|_1 = \int_a^b |(\alpha f)(x)| dx = \int_a^b |\alpha| |f(x)| dx = |\alpha| \int_a^b |f(x)| dx = |\alpha| \|f\|_1.$$

(iii) Let $f, g \in \mathcal{C}([a, b])$. The triangle inequality follows from

$$|f(x) + g(x)| \leq |f(x)| + |g(x)| \quad \text{for all } x \in [a, b],$$

where we have used the triangle inequality for real numbers, and from the first linear property of the Riemann integral (see Theorem 3.29). Indeed, we have for $f, g \in \mathcal{C}([a, b])$

$$\begin{aligned} \|f + g\|_1 &= \int_a^b |f(x) + g(x)| dx \\ &\leq \int_a^b (|f(x)| + |g(x)|) dx = \int_a^b |f(x)| dx + \int_a^b |g(x)| dx = \|f\|_1 + \|g\|_1. \end{aligned}$$

Thus the L_1 -norm is a norm for the linear space $\mathcal{C}([a, b])$.

(c) (i) From the first order property of the Riemann integral (see Theorem 3.29) we know that

$$\|f\|_2 = \left(\int_a^b |f(x)|^2 dx \right)^{1/2} \geq 0,$$

since $|f(x)|^2 \geq 0$ for all $x \in [a, b]$. If $f(x) = 0$ for all $x \in [a, b]$ then $\|f\|_2 = 0$. From the third order property we see that, for any $f \in \mathcal{C}([a, b])$, $\|f\|_2 = 0$ implies that $|f(x)|^2 = 0$ for all $x \in [a, b]$, and hence $f(x) = 0$ for all $x \in [a, b]$. Thus $\|f\|_2 = 0$ if and only if $f(x) = 0$ for all $x \in [a, b]$.

(ii) Let $\alpha \in \mathbb{R}$ and $f \in \mathcal{C}([a, b])$. Then

$$\|\alpha f\|_2 = \left(\int_a^b |\alpha|^2 |f(x)|^2 dx \right)^{1/2} = \left(|\alpha|^2 \int_a^b |f(x)|^2 dx \right)^{1/2} = |\alpha| \|f\|_2.$$

(iii) Let $f, g \in \mathcal{C}([a, b])$. The triangle inequality follows from the Schwarz inequality for the L_2 -norm (which can be proved analogously to the Schwarz inequality for \mathbb{R}^n (see Lemma 6.12 and its proof)),

$$\int_a^b |f(x)g(x)| dx \leq \left(\int_a^b |f(x)|^2 dx \right)^{1/2} \left(\int_a^b |g(x)|^2 dx \right)^{1/2} = \|f\|_2 \|g\|_2 \quad (6.4)$$

for all $f, g \in \mathcal{C}([a, b])$, with the following argumentation: from the linear properties of the Riemann integral, the triangle inequality for $(\mathbb{R}^n, |\cdot|)$ and (6.4),

$$\begin{aligned} \|f + g\|_2^2 &= \int_a^b |f(x) + g(x)|^2 dx \leq \int_a^b (|f(x)| + |g(x)|)^2 dx \\ &= \int_a^b (|f(x)|^2 + 2|f(x)||g(x)| + |g(x)|^2) dx \\ &= \int_a^b |f(x)|^2 dx + 2 \int_a^b |f(x)g(x)| dx + \int_a^b |g(x)|^2 dx \\ &= \|f\|_2^2 + 2 \int_a^b |f(x)g(x)| dx + \|g\|_2^2 \\ &\leq \|f\|_2^2 + 2\|f\|_2 \|g\|_2 + \|g\|_2^2 = (\|f\|_2 + \|g\|_2)^2. \end{aligned}$$

Thus the L_2 -norm is a norm for the linear space $\mathcal{C}([a, b])$. □

We give some more examples of norms on function spaces.

Example 6.16 (norms on $\mathcal{C}([0, 1])$)

Determine which of the following are norms for $\mathcal{C}([0, 1])$.

(a) $\|f\| := f(1) + \sup_{x \in [0, 1/2]} |f(x)|,$

$$(b) \quad \|f\| := \left| f\left(\frac{1}{2}\right) \right| + \int_0^1 |f(x)| dx.$$

Solution: We check whether the norm properties are satisfied.

(a) Claim: $\|f\| = f(1) + \sup_{x \in [0, 1/2]} |f(x)|$ does not define a norm of $C([0, 1])$ because $\|f\| = 0$ does not imply $f(x) = 0$ for all $x \in [0, 1]$.

For example, if we choose a piecewise affine linear function $f(x)$ defined by

$$F(x) := \begin{cases} 0 & \text{if } 0 \leq x < \frac{1}{2}, \\ x - \frac{1}{2} & \text{if } \frac{1}{2} \leq x \leq \frac{3}{4}, \\ 1 - x & \text{if } \frac{3}{4} < x \leq 1, \end{cases}$$

then $\sup_{x \in [0, 1/2]} |f(x)| = 0$, and $f(1) = 0$. Hence $\|f\| = 0$, but f is not identically zero in $[0, 1]$. Thus the property (i) of a norm is violated.

(b) Claim:

$$\|f\| = \left| f\left(\frac{1}{2}\right) \right| + \int_0^1 |f(x)| dx$$

is a norm for $\mathcal{C}([0, 1])$.

(i) From the first order property of the Riemann integral and from the definition of the absolute value, we have

$$\|f\| = \left| f\left(\frac{1}{2}\right) \right| + \int_0^1 |f(x)| dx \geq 0 \quad \text{for all } f \in \mathcal{C}([0, 1]).$$

If $\|f\| = 0$ then

$$\|f\| = \left| f\left(\frac{1}{2}\right) \right| + \int_0^1 |f(x)| dx = 0 \quad \Rightarrow \quad \left| f\left(\frac{1}{2}\right) \right| = 0 \quad \text{and} \quad \int_0^1 |f(x)| dx = 0. \quad (6.5)$$

From Example 6.15 we know that

$$\|f\|_1 := \int_0^1 |f(x)| dx$$

is a norm on $\mathcal{C}([0, 1])$. Thus (6.5) implies $f(x) = 0$ for all $x \in [0, 1]$. On the other hand, if $f(x) = 0$ for all $x \in [0, 1]$, then $\|f\| = 0$. Thus $\|f\| = 0$ if and only if $f(x) = 0$ for all $x \in [0, 1]$.

(ii) For all $\alpha \in \mathbb{R}$ and $f \in \mathcal{C}([0, 1])$, we have from the linear properties of the integral

$$\begin{aligned}\|\alpha f\| &= \left| \alpha f\left(\frac{1}{2}\right) \right| + \int_0^1 |\alpha f(x)| dx = |\alpha| \left| f\left(\frac{1}{2}\right) \right| + |\alpha| \int_0^1 |f(x)| dx \\ &= |\alpha| \left(\left| f\left(\frac{1}{2}\right) \right| + \int_0^1 |f(x)| dx \right) = |\alpha| \|f\|.\end{aligned}$$

(iii) The triangle inequality follows from the triangle inequality for real numbers and from the properties of the integral

$$\begin{aligned}\|f + g\| &= \left| f\left(\frac{1}{2}\right) + g\left(\frac{1}{2}\right) \right| + \int_0^1 |f(x) + g(x)| dx \\ &\leq \left| f\left(\frac{1}{2}\right) \right| + \left| g\left(\frac{1}{2}\right) \right| + \int_0^1 (|f(x)| + |g(x)|) dx \\ &\leq \left| f\left(\frac{1}{2}\right) \right| + \left| g\left(\frac{1}{2}\right) \right| + \int_0^1 |f(x)| dx + \int_0^1 |g(x)| dx \\ &\leq \left(\left| f\left(\frac{1}{2}\right) \right| + \int_0^1 |f(x)| dx \right) + \left(\left| g\left(\frac{1}{2}\right) \right| + \int_0^1 |g(x)| dx \right) = \|f\| + \|g\|\end{aligned}$$

for all $f, g \in \mathcal{C}([0, 1])$.

Thus $\|\cdot\|$ is a norm for $\mathcal{C}([0, 1])$. □

Example 6.17 (subset of continuous functions)

For the linear space $\mathcal{C}([0, \infty))$ of continuous real-valued functions on $[0, \infty)$, define

$$\|f\| := \sup_{x \in [0, \infty)} e^{-x} |f(x)|, \quad f \in \mathcal{C}([0, \infty)).$$

Let X be the subset of $\mathcal{C}([0, \infty))$ defined by

$$X := \{f \in \mathcal{C}([0, \infty)) : \|f\| < \infty\}.$$

Then $(X, \|\cdot\|)$ is a normed linear space. The rapidly declining function e^{-x} is called a **weight function**.

Proof: We check the vector space properties for X and the three norm properties for $\|\cdot\|$.

First we check the vector space properties for X . First we have to check closure for the pointwise addition of functions and the pointwise scalar multiplication. If $f, g \in X$, then from the triangle inequality of real numbers

$$\|f + g\| = \sup_{x \in [0, \infty)} e^{-x} |f(x) + g(x)|$$

$$\begin{aligned}
&\leq \sup_{x \in [0, \infty)} e^{-x} (|f(x)| + |g(x)|) = \sup_{x \in [0, \infty)} (e^{-x}|f(x)| + e^{-x}|g(x)|) \\
&\leq \sup_{x \in [0, \infty)} e^{-x}|f(x)| + \sup_{x \in [0, \infty)} e^{-x}|g(x)| = \|f\| + \|g\| < \infty.
\end{aligned}$$

If $\alpha \in \mathbb{R}$ and $f \in X$, then

$$\|\alpha f\| = \sup_{x \in [0, \infty)} e^{-x} |\alpha f(x)| = |\alpha| \sup_{x \in [0, \infty)} e^{-x} |f(x)| = |\alpha| \|f\| < \infty.$$

Thus we have closure, that is, $f + g \in X$ for all $f, g \in X$, and $\alpha f \in X$ for all $\alpha \in \mathbb{R}$ and all $f \in X$.

(i) For all $f, g, h \in X$, we have $(f + g) + h = f + (g + h)$ because, from the associative law for the real numbers

$$(f(x) + g(x)) + h(x) = f(x) + (g(x) + h(x)) \quad \text{for all } x \geq 0 \text{ and for all } f, g, h \in X.$$

(ii) The neutral element is given by the zero function $\mathcal{O}(x) := 0$ for all $x \geq 0$. This follows from

$$\|\mathcal{O}\| := \sup_{x \in [0, \infty)} e^{-x} |\mathcal{O}(x)| = \sup_{x \in [0, \infty)} e^{-x} |0| = \sup_{x \in [0, \infty)} 0 = 0 < \infty,$$

that is, $\mathcal{O} \in X$, and from

$$\begin{aligned}
(f + \mathcal{O})(x) &= f(x) + \mathcal{O}(x) = f(x) + 0 = f(x) \quad \text{for all } x \geq 0 \text{ and for all } f \in X, \\
(\mathcal{O} + f)(x) &= \mathcal{O}(x) + f(x) = 0 + f(x) = f(x) \quad \text{for all } x \geq 0 \text{ and for all } f \in X.
\end{aligned}$$

(iii) For $f \in X$, the inverse element is given by $-f \in X$ because

$$\begin{aligned}
(f + (-f))(x) &= f(x) - f(x) = 0 = \mathcal{O}(x) \quad \text{for all } x \geq 0, \\
((-f) + f)(x) &= -f(x) + f(x) = 0 = \mathcal{O}(x) \quad \text{for all } x \geq 0.
\end{aligned}$$

(iv) Since the real numbers are commutative, we have

$$(f + g)(x) = f(x) + g(x) = g(x) + f(x) = (g + f)(x) \quad \text{for all } x \geq 0 \text{ and for all } f, g \in X,$$

that is, $f + g = g + f$ for all $f, g \in X$.

(v) We have $(1f)(x) = 1f(x) = f(x)$ for all $x \geq 0$ and for all $f \in X$.

(vi) We have $((\alpha\beta)f)(x) = \alpha\beta f(x) = \alpha(\beta f(x))$ for all $x \geq 0$ and for all $f \in X$ and for all $\alpha, \beta \in \mathbb{R}$. Thus $(\alpha\beta)f = \alpha(\beta f)$ for all $f \in X$ and all $\alpha, \beta \in \mathbb{R}$.

(vii) For all $f, g \in X$ and for all $\alpha, \beta \in \mathbb{R}$, we have from the distributive laws for the real numbers

$$((\alpha + \beta)f)(x) = (\alpha + \beta)f(x) = \alpha f(x) + \beta f(x) = (\alpha f)(x) + (\beta f)(x), \quad x \geq 0,$$

that is, $(\alpha + \beta)f = \alpha f + \beta f$ for all $f \in X$ and all $\alpha, \beta \in \mathbb{R}$, and

$$(\alpha(f+g))(x) = \alpha(f+g)(x) = \alpha(f(x)+g(x)) = \alpha f(x) + \alpha g(x) = (\alpha f)(x) = (\alpha g)(x)$$

for all $x \geq 0$, that is, $\alpha(f+g) = \alpha f + \alpha g$ for all $f, g \in X$ and all $\alpha \in \mathbb{R}$

Now we check the norm properties.

(i) Since $e^{-x} > 0$ for all $x \in [0, \infty)$ and $|f(x)| \geq 0$ for all $x \in [0, \infty)$, we clearly have $\|f\| \geq 0$ for all $f \in X$. If $f(x) = 0$ for all $x \in [0, \infty)$, then $\|f\| = 0$. If

$$\|f\| = \sup_{x \in [0, \infty)} e^{-x}|f(x)| = 0$$

then we conclude from $e^{-x} > 0$, $x \in [0, \infty)$, that $e^{-x}|f(x)| = 0$ for all $x \in [0, \infty)$. Since $e^{-x} > 0$ for all $x \in [0, \infty)$, this implies $f(x) = 0$ for all $x \in [0, \infty)$. Thus $\|f\| = 0$ if and only if f is the zero function.

(ii) Let $\alpha \in \mathbb{R}$ and $f \in X$. Then

$$\|\alpha f\| = \sup_{x \in [0, \infty)} e^{-x}|\alpha f(x)| = \sup_{x \in [0, \infty)} e^{-x}|\alpha| |f(x)| = |\alpha| \sup_{x \in [0, \infty)} e^{-x}|f(x)| = |\alpha| \|f\|.$$

(iii) Let $f, g \in X$. The triangle inequality can be derived with the help of

$$|f(x) + g(x)| \leq |f(x)| + |g(x)| \quad \text{for all } x \in [0, \infty),$$

which follows from the triangle inequality for real numbers. Indeed, since $e^{-x} > 0$ for all $x \in [0, \infty)$,

$$\begin{aligned} \|f + g\| &= \sup_{x \in [0, \infty)} e^{-x}|f(x) + g(x)| \\ &\leq \sup_{x \in [0, \infty)} e^{-x}(|f(x)| + |g(x)|) = \sup_{x \in [0, \infty)} (e^{-x}|f(x)| + e^{-x}|g(x)|) \\ &\leq \sup_{x \in [0, \infty)} e^{-x}|f(x)| + \sup_{x \in [0, \infty)} e^{-x}|g(x)| = \|f\| + \|g\|. \end{aligned}$$

Thus $\|\cdot\|$ fulfills the three properties of a norm.

We have proved that $(X, \|\cdot\|)$ is a normed linear space. □

6.1.4 Inner Product Spaces

A special class of normed linear spaces are those whose norm is defined by a so-called **inner product** or **scalar product**.

Definition 6.18 (inner product/scalar product)

Let X be a (real) linear space. A (real) **inner product** or **scalar product** is a function $(\cdot, \cdot) : X \times X \rightarrow \mathbb{R}$ with the following properties:

- (i) $(x, x) > 0$ for all $x \in X$ with $x \neq \mathcal{O}$.
- (ii) $(x, y) = (y, x)$ for all $x, y \in X$ (symmetry).
- (iii) $(x, y + z) = (x, y) + (x, z)$ for all $x, y, z \in X$.
- (iv) $(\lambda x, y) = \lambda (x, y) = (x, \lambda y)$ for all $x, y \in X$ and all scalars $\lambda \in \mathbb{R}$.

The normed linear space X with the inner product (\cdot, \cdot) is called an **inner product space**.

Firstly we observe that property (ii) and (iii) in Definition 6.18 imply that also

$$(y + z, x) = (y, x) + (z, x) \quad \text{for all } x, y, z \in X. \quad (6.6)$$

We also observe that property (i) of the inner product implies that

$$(x, x) \geq 0 \quad \text{for all } x \in X, \quad \text{and} \quad (x, x) = 0 \quad \text{if and only if } x = \mathcal{O}. \quad (6.7)$$

Indeed, since from (i) $(x, x) > 0$ for all $x \neq \mathcal{O}$, it only remains to show that $(x, x) = 0$ if $x = \mathcal{O}$. If $x = \mathcal{O}$, then $x = \mathcal{O} = 0 \mathcal{O}$, and thus from property (iv) of the inner product

$$(\mathcal{O}, \mathcal{O}) = (0 \mathcal{O}, \mathcal{O}) = 0 (\mathcal{O}, \mathcal{O}) = 0.$$

Our standard example is the so-called **Euclidean inner product** on \mathbb{R}^n .

Lemma 6.19 (\mathbb{R}^n with Euclidean inner product)

Define on \mathbb{R}^n the **Euclidean inner product**

$$(\mathbf{x}, \mathbf{y}) := \mathbf{x}^T \cdot \mathbf{y} = \sum_{j=1}^n x_j y_j = x_1 y_1 + x_2 y_2 + \dots + x_n y_n.$$

Then \mathbb{R}^n with the Euclidean inner product is an inner product space.

Proof of Lemma 6.19. We observe that

$$(\mathbf{x}, \mathbf{x}) = \sum_{j=1}^n x_j^2 = \|\mathbf{x}\|_2^2,$$

and therefore we can use our knowledge about the Euclidean norm.

(i) From the property (i) of the Euclidean norm, we know that $\|\mathbf{x}\|_2^2 = (\mathbf{x}, \mathbf{x}) > 0$ if $\mathbf{x} \neq \mathbf{0}$.

(ii) The symmetry follows from the commutative law for the multiplication of real numbers: for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$,

$$(\mathbf{x}, \mathbf{y}) = \sum_{j=1}^n x_j y_j = \sum_{j=1}^n y_j x_j = (\mathbf{y}, \mathbf{x}).$$

(iii) For any $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$, we have from the distributive law for real numbers

$$\begin{aligned} (\mathbf{x}, \mathbf{y} + \mathbf{z}) &= \sum_{j=1}^n x_j (y_j + z_j) = \sum_{j=1}^n [x_j y_j + x_j z_j] \\ &= \sum_{j=1}^n x_j y_j + \sum_{j=1}^n x_j z_j = (\mathbf{x}, \mathbf{y}) + (\mathbf{x}, \mathbf{z}). \end{aligned}$$

(iv) For any real $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and any real number $\lambda \in \mathbb{R}$, we have

$$(\lambda \mathbf{x}, \mathbf{y}) = \sum_{j=1}^n (\lambda x_j) y_j = \lambda \sum_{j=1}^n x_j y_j = \lambda (\mathbf{x}, \mathbf{y}).$$

From this relation and the symmetry of the inner product (property (ii) which we have already verified), we have for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and any real number $\lambda \in \mathbb{R}$

$$(\lambda \mathbf{x}, \mathbf{y}) = \lambda (\mathbf{x}, \mathbf{y}) = \lambda (\mathbf{y}, \mathbf{x}) = (\lambda \mathbf{y}, \mathbf{x}) = (\mathbf{x}, \lambda \mathbf{y}).$$

Thus all the properties of an inner product are satisfied, and \mathbb{R}^n with the Euclidean inner product is indeed an inner product space. \square

In the previous sections we have already seen that both in \mathbb{R}^n with the Euclidean norm as well as in $\mathcal{C}([a, b])$ with the L_2 -norm

$$\|f\|_2 = \left(\int_a^b |f(x)|^2 dx \right)^{1/2}$$

we have a Schwarz inequality. We will now see that a **Schwarz inequality** is a property of any inner product space.

Lemma 6.20 (Schwarz inequality)

Let X with the inner product $(\cdot, \cdot) : X \times X \rightarrow \mathbb{R}$ be an inner product space. Then the **Schwarz inequality** holds

$$|(x, y)| \leq \sqrt{(x, x)} \sqrt{(y, y)} \quad \text{for all } x, y \in X,$$

where we have equality if and only if $x = \alpha y$ with some $\alpha \in \mathbb{R}$ or $y = \beta x$ with some $\beta \in \mathbb{R}$.

Before we prove Lemma 6.20, we take a look at the Schwarz inequality for our standard example \mathbb{R}^n with the Euclidean inner product.

In Lemma 6.19, we have seen that \mathbb{R}^n with the Euclidean inner product $(\mathbf{x}, \mathbf{y}) = \sum_{j=1}^n x_j y_j$ is an inner product space, and in Lemma 6.12 we have proved that

$$\left| \sum_{j=1}^n x_j y_j \right| \leq \sqrt{\sum_{j=1}^n x_j^2} \sqrt{\sum_{j=1}^n y_j^2},$$

which we can now write as

$$|(\mathbf{x}, \mathbf{y})| \leq \sqrt{(\mathbf{x}, \mathbf{x})} \sqrt{(\mathbf{y}, \mathbf{y})} \quad \text{for all } \mathbf{x}, \mathbf{y} \in \mathbb{R}^n.$$

We see that the Schwarz inequality for \mathbb{R}^n with the Euclidean inner product (\cdot, \cdot) is indeed a special case of Lemma 6.20.

Proof of Lemma 6.20. First we consider the case that $x = \mathcal{O}$ or $y = \mathcal{O}$; in this case the left-hand side and the right-hand side are both zero, since $(\mathcal{O}, \mathcal{O}) = 0$ and

$$(\mathcal{O}, x) = (0 \mathcal{O}, x) = 0 (\mathcal{O}, x) = 0 \quad \text{for all } x \in X,$$

and we have an equality. We note that if $x = \mathcal{O}$, then $x = \mathcal{O} = 0 \mathbf{y}$ and that if $y = \mathcal{O}$ then $y = \mathcal{O} = 0 x$.

Now let us assume that both x and y are different from the zero vector \mathcal{O} , and consider the quadratic function

$$f(\lambda) := (y - \lambda x, y - \lambda x) = \lambda^2 (x, x) - 2 \lambda (x, y) + (y, y),$$

where we have used the properties (ii), (iii), and (iv) of the inner product to obtain the second representation of $f(\lambda)$. This is a quadratic equation in λ , and, using $(x, x) > 0$ since $x \neq \mathcal{O}$, we factorize

$$\begin{aligned} f(\lambda) &= \lambda^2 (x, x) - 2\lambda (x, y) + (y, y) \\ &= (x, x) \left[\lambda^2 - 2\lambda \frac{(x, y)}{(x, x)} + \frac{(y, y)}{(x, x)} \right] \\ &= (x, x) \left[\left(\lambda - \frac{(x, y)}{(x, x)} \right)^2 + \left(\frac{(y, y)}{(x, x)} - \frac{(x, y)^2}{(x, x)^2} \right) \right] \\ &= (x, x) \left[\left(\lambda - \frac{(x, y)}{(x, x)} \right)^2 + \left(\frac{(x, x)(y, y) - (x, y)^2}{(x, x)^2} \right) \right]. \end{aligned}$$

Now we choose $\lambda = (x, y)/(x, x)$ so that the first term in the angular brackets vanishes, and we also use $f(\lambda) = (y - \lambda x, y - \lambda x) \geq 0$ from the property (6.7) of the inner product. Thus

$$0 \leq f\left(\frac{(x, y)}{(x, x)}\right) = \frac{(x, x)(y, y) - (x, y)^2}{(x, x)}.$$

Since $(x, x) > 0$ (from $x \neq \mathcal{O}$), we have from multiplying by (x, x)

$$0 \leq (x, x)(y, y) - (x, y)^2 \quad \Leftrightarrow \quad (x, y)^2 \leq (x, x)(y, y).$$

Taking roots and observing that $(x, x) > 0$ and $(y, y) > 0$, since $x \neq \mathcal{O}$ and $y \neq \mathcal{O}$, yields

$$|(x, y)| \leq \sqrt{(x, x)} \sqrt{(y, y)},$$

which proves the Schwarz inequality.

It remains to show that we have equality if and only if $x = \alpha y$ with some $\alpha \in \mathbb{R}$ or $y = \beta x$ with some $\beta \in \mathbb{R}$. – If $x = \mathcal{O}$ or $y = \mathcal{O}$, then we have seen that this is indeed true. – Now consider again the case that $x \neq \mathcal{O}$ and $y \neq \mathcal{O}$. If $x = \alpha y$ with some $\alpha \in \mathbb{R}$, then we find that $y = \alpha^{-1} x$ (since $\alpha \neq 0$ due to $x \neq \mathcal{O}$), and thus it is enough to consider the case $y = \beta x$. From taking the inner product with x ,

$$y = \beta x \quad \Rightarrow \quad (x, y) = \beta (x, x) \quad \Rightarrow \quad \beta = \frac{(x, y)}{(x, x)},$$

and thus

$$y = \beta x = \frac{(x, y)}{(x, x)} x.$$

Inspecting the proof we find that now

$$f\left(\frac{(x, y)}{(x, x)}\right) = \left(y - \frac{(x, y)}{(x, x)}x, y - \frac{(x, y)}{(x, x)}x\right) = (\mathcal{O}, \mathcal{O}) = 0,$$

and thus the \leq in the Schwarz estimate becomes an equality. On the other hand, if $y \neq \beta x$ for all $\beta \in \mathbb{R}$, then from property (i) of the inner product

$$f\left(\frac{(x, y)}{(x, x)}\right) = \left(y - \frac{(x, y)}{(x, x)}x, y - \frac{(x, y)}{(x, x)}x\right) > 0,$$

since $y - [(x, y)/(x, x)]x \neq \mathcal{O}$. Inspecting the proof shows that the Schwarz inequality becomes a strict inequality. \square

In the next lemma we learn that any inner product space is also a normed linear space with a suitable norm ‘induced’ by the inner product.

Lemma 6.21 (inner product space \Rightarrow normed linear space)

Let X with $(\cdot, \cdot) : X \times X \rightarrow \mathbb{R}$ be an inner product space. Then

$$\|\cdot\| : X \rightarrow \mathbb{R}, \quad \|x\| := \sqrt{(x, x)},$$

defines a norm on X , and $(X, \|\cdot\|)$ is a normed linear space.

The proof follows straight-forward by making use of the properties of the inner product.

Proof of Lemma 6.21. We have to verify that $\|x\| = \sqrt{(x, x)}$ satisfies the three norm properties.

(i) From property (i) of the inner product we have

$$\|x\|^2 = (x, x) > 0 \quad \text{for all } x \neq 0 \quad \Rightarrow \quad \|x\| = \sqrt{(x, x)} > 0 \quad \text{for all } x \neq 0.$$

If $x = \mathcal{O}$, then $x = \mathcal{O} = 0\mathcal{O}$, and thus from property (iv) of the inner product

$$\|\mathcal{O}\|^2 = (\mathcal{O}, \mathcal{O}) = (0\mathcal{O}, \mathcal{O}) = 0(\mathcal{O}, \mathcal{O}) = 0 \quad \Rightarrow \quad \|\mathcal{O}\| = 0.$$

Thus we have shown that $\|x\| \geq 0$ for all $x \in X$, and $\|x\| = 0$ if and only if $x = \mathcal{O}$.

(ii) Let $\alpha \in \mathbb{R}$ and $x \in X$ be arbitrary. Then

$$\|\alpha x\|^2 = (\alpha x, \alpha x) = \alpha^2 (x, x) = |\alpha|^2 \|x\|^2 \quad \Rightarrow \quad \|\alpha x\| = |\alpha| \|x\|.$$

(iii) Let x and y be two arbitrary elements in X . From property (iii) of the inner product, from (6.6), and from property (ii) of the inner product

$$\begin{aligned}
 \|x + y\|^2 &= (x + y, x + y) \\
 &= (x, x + y) + (y, x + y) \\
 &= (x, x) + (x, y) + (y, x) + (y, y) \\
 &= \|x\|^2 + 2(x, y) + \|y\|^2 \\
 &\leq \|x\|^2 + 2|(x, y)| + \|y\|^2.
 \end{aligned} \tag{6.8}$$

From Lemma 6.20, the inner product satisfies the Schwarz inequality

$$|(x, y)| \leq \sqrt{(x, x)} \sqrt{(y, y)} = \|x\| \|y\| \quad \text{for all } x, y \in X,$$

which we now apply to estimate the middle term on the right-hand side of (6.8).

$$\|x + y\|^2 \leq \|x\|^2 + 2|(x, y)| + \|y\|^2 \leq \|x\|^2 + 2\|x\| \|y\| + \|y\|^2 = (\|x\| + \|y\|)^2,$$

and thus taking the root yields the triangle inequality

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

We have verified that $\|x\| = \sqrt{(x, x)}$ satisfies the three norm properties and is thus a norm on X . Thus X with $\|x\| = \sqrt{(x, x)}$ is a normed linear space. \square

After proving Lemma 6.21, we can also state the **Schwarz inequality** in the following form: Let $(X, (\cdot, \cdot))$ be an inner product space with the inner product $(\cdot, \cdot) : X \times X \rightarrow \mathbb{R}$. Then

$$\boxed{|(x, y)| \leq \|x\| \|y\| \quad \text{for all } x, y \in X, \quad \text{where } \|x\| := \sqrt{(x, x)}, \quad \|y\| := \sqrt{(y, y)}. } \tag{6.9}$$

Equality in (6.9) holds if and only if $x = \alpha y$ with some $\alpha \in \mathbb{R}$ or $y = \beta x$ with some $\beta \in \mathbb{R}$.

We give another example of an inner product space that we have already encountered when we discussed normed linear spaces.

Example 6.22 (L_2 -inner product for continuous functions)

The space of continuous functions $\mathcal{C}([a, b])$ on the interval $[a, b]$ with the L_2 -inner product

$$(f, g) = \int_a^b f(x) g(x) dx, \quad f, g \in \mathcal{C}([a, b]), \tag{6.10}$$

is an inner product space.

Proof: We verify the properties of an inner product.

(i) Consider a continuous function $f(x)$ which is different from the zero function. Then $|f(x)|^2 \geq 0$ for all $x \in [a, b]$, and since f is not the zero function we have $|f(x)|^2 > 0$ for some $x \in [a, b]$. Since f is continuous and $|f(x)| > 0$ for some $x \in [a, b]$, we know, from the third order property of the Riemann integral (see Theorem 3.29), that

$$(f, f) = \int_a^b |f(x)|^2 dx > 0.$$

(Indeed, assume that $(f, f) = 0$, then the third order property would imply that $|f(x)|^2 = 0$ for all $x \in [a, b]$ which is not true.)

(ii) Since the multiplication of real numbers is commutative, we have

$$(f, g) = \int_a^b f(x) g(x) dx = \int_a^b g(x) f(x) dx = (g, f) \quad f, g \in \mathcal{C}([a, b]).$$

(iii) From the distributive law of the real numbers and the first linear property of the Riemann integral (see Theorem 3.29) we have, for all $f, g, h \in \mathcal{C}([a, b])$,

$$\begin{aligned} (f, g + h) &= \int_a^b f(x) [g(x) + h(x)] dx = \int_a^b [f(x) g(x) + f(x) h(x)] dx \\ &= \int_a^b f(x) g(x) dx + \int_a^b f(x) h(x) dx = (f, g) + (f, h). \end{aligned}$$

(iv) From the second linear property of the Riemann integral (see Theorem 3.29), we have for any $f, g \in \mathcal{C}([a, b])$ and any $\lambda \in \mathbb{R}$

$$(\lambda f, g) = \int_a^b \lambda f(x) g(x) dx = \lambda \int_a^b f(x) g(x) dx = \lambda (f, g). \quad (6.11)$$

From the symmetry property (ii) of the inner product and from (6.11) we have for any $f, g \in \mathcal{C}([a, b])$ and any $\lambda \in \mathbb{R}$

$$(f, \lambda g) = (\lambda g, f) = \lambda (g, f) = \lambda (f, g). \quad (6.12)$$

From (6.11) and (6.12) we obtain property (iv) of the inner product.

Thus we have verified all the properties of an inner product and know therefore that $\mathcal{C}([a, b])$ with the inner product (6.10) is an inner product space. \square

6.2 Sequences in Metric Spaces and Normed Linear Spaces

In Subsection 6.2.1, we define the notion of **(convergent) sequences** and **Cauchy sequences**. In Subsection 6.2.2, the notion of completeness is introduced: A metric space or normed linear space is **complete** if every Cauchy sequence in the metric space or normed linear space converges (to an element in the space). In Subsection 6.2.3, we introduce the notion of **bounded sets** in metric spaces and normed linear spaces. Then we will prove the **Bolzano-Weierstrass theorem** for \mathbb{R}^n which says that any bounded sequence in \mathbb{R}^n with the Euclidean norm has a convergent subsequence. For $\mathbb{R}^1 = \mathbb{R}$ you should have learnt the Bolzano-Weierstrass theorem in your first year courses. Throughout this section we will discuss various examples.

6.2.1 Convergent Sequences and Cauchy Sequences

With the help of the **metric** or **distance function** $d(x, y)$ in a metric space, we can give an ε -definition of **convergent sequences** and **Cauchy sequences** in metric spaces and normed linear spaces.

Definition 6.23 (sequences in a metric space)

Let X be a metric space with the distance function $d : X \times X \rightarrow \mathbb{R}$.

- (i) A sequence $\{x_k\}$ in (X, d) **converges to** $x \in X$ (that is, $\lim_{k \rightarrow \infty} x_k = x$) if for every $\varepsilon > 0$ there exists an $N = N(\varepsilon) \in \mathbb{N}$ such that

$$d(x_k, x) < \varepsilon \quad \text{for all } k \geq N.$$

Equivalently, a sequence $\{x_k\}$ in (X, d) **converges to** $x \in X$ if

$$\lim_{k \rightarrow \infty} d(x_k, x) = 0.$$

- (ii) A sequence $\{x_k\}$ in (X, d) is a **Cauchy sequence** if for each $\varepsilon > 0$ there exists an $N = N(\varepsilon) \in \mathbb{N}$ such that

$$d(x_k, x_m) < \varepsilon \quad \text{for all } m, k \geq N.$$

Example 6.24 (Cauchy sequences in \mathbb{R} with $d(x, y) := |x - y|$)

We have seen in Example 6.2 that the real line \mathbb{R} with the metric $d(x, y) := |x - y|$,

defined with the absolute value $|\cdot|$, is a metric space. According to Definition 6.23, a sequence $\{x_k\}$ in $(\mathbb{R}, |\cdot|)$ is convergent to $x \in \mathbb{R}$ if for every $\varepsilon > 0$ there exists an $N = N(\varepsilon) \in \mathbb{N}$ such that

$$d(x_k, x) = |x_k - x| < \varepsilon \quad \text{for all } k \geq N.$$

We see that this is just the usual definition of the convergence of a sequence in \mathbb{R} which we know from first year.

According to Definition 6.23, a sequence $\{x_k\}$ in $(\mathbb{R}, |\cdot|)$ is a Cauchy sequence if for every $\varepsilon > 0$ there exists an $N = N(\varepsilon) \in \mathbb{N}$ such that

$$d(x_k, x_m) = |x_k - x_m| < \varepsilon \quad \text{for all } k, m \geq N.$$

This is just the usual definition of a Cauchy sequence in \mathbb{R} which we know from first year. In first year we also learned the Cauchy principle which states that a sequence in \mathbb{R} is convergent if and only if it is a Cauchy sequence (see Lemme 5.15). \square

Definition 6.23 implies for a normed linear space $(X, \|\cdot\|)$ the following definition by using the distance function $d(x, y) = \|x - y\|$.

Definition 6.25 (sequences in a normed linear space)

Let X be a normed linear space with norm $\|\cdot\| : X \rightarrow \mathbb{R}$.

- (i) A sequence $\{x_k\}$ in $(X, \|\cdot\|)$ **converges to** $x \in X$ (that is, $\lim_{k \rightarrow \infty} x_k = x$) if for every $\varepsilon > 0$ there exists $N = N(\varepsilon) \in \mathbb{N}$ such that

$$\|x_k - x\| < \varepsilon \quad \text{for all } k \geq N.$$

Equivalently, a sequence $\{x_k\}$ in $(X, \|\cdot\|)$ **converges to** $x \in X$ if

$$\lim_{k \rightarrow \infty} \|x_k - x\| = 0.$$

- (ii) A sequence $\{x_k\}$ in $(X, \|\cdot\|)$ is a **Cauchy sequence** if for each $\varepsilon > 0$ there exists an $N = N(\varepsilon) \in \mathbb{N}$ such that

$$\|x_k - x_m\| < \varepsilon \quad \text{for all } m, k \geq N.$$

We will discuss another example.

Example 6.26 (sequences in \mathbb{R} with the discrete metric)

In Example 6.4, we saw that \mathbb{R} with the **discrete metric**

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

is a metric space. We will now show that a convergent sequence $\{x_n\}$ in \mathbb{R} with the discrete metric that converges to $x \in \mathbb{R}$ satisfies $x_n = x$ for all $n \geq N$ with some $N \in \mathbb{N}$, that is, from some $n = N$ onwards the sequence is constant.

Proof: Let $\{x_n\}$ be a sequence in (\mathbb{R}, d) that converges to x . Then for every $\varepsilon > 0$, there exists $N = N(\varepsilon) \in \mathbb{N}$ such that

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

Now we choose $\varepsilon = 1$. Then there exists $N = N(1) \in \mathbb{N}$ such that

$$d(x_n, x) < 1 \quad \text{for all } n \geq N. \quad (6.13)$$

From the definition of the discrete metric, we know that $d(x_n, x)$ has either the value 0 or 1. But from (6.13) we have $d(x_n, x) < 1$, and thus $d(x_n, x) = 0$ for all $n \geq N$. Thus we have $x_n = x$ for all $n \geq N$ as claimed. \square

A more complicated example is given by the space of continuous functions $\mathcal{C}([a, b])$ on $[a, b]$ equipped with the supremum norm $\|\cdot\|_\infty$.

Example 6.27 ($\mathcal{C}([a, b])$ with the supremum norm)

In Chapter 5, we defined the supremum norm

$$\|f\|_\infty := \sup_{x \in [a, b]} |f(x)|, \quad f \in \mathcal{B}([a, b]),$$

and in Example 6.15 we have seen that the space of continuous functions $\mathcal{C}([a, b])$ with the supremum norm $\|\cdot\|_\infty$ is a normed linear space. Convergence in the normed linear space $(\mathcal{C}([a, b]), \|\cdot\|_\infty)$ is just **uniform convergence**, and a Cauchy sequence in $(\mathcal{C}([a, b]), \|\cdot\|_\infty)$ is a **uniform Cauchy sequence** as encountered in Chapter 5.

Proof: According to the definitions in this section a sequence $\{f_n\}$ in $(\mathcal{C}([a, b]), \|\cdot\|_\infty)$ converges to $f \in \mathcal{C}([a, b])$ if for every $\varepsilon > 0$ there exists some $N \in \mathbb{N}$ such that

$$\|f_n - f\|_\infty = \sup_{x \in [a, b]} |f_n(x) - f(x)| < \varepsilon \quad \text{for all } x \in [a, b] \text{ and for all } n \geq N.$$

This means exactly that $\{f_n\}$ converges uniformly to f .

According to the definitions in this section a sequence $\{f_n\}$ in $(\mathcal{C}([a, b]), \|\cdot\|_\infty)$ is a Cauchy sequence if for every $\varepsilon > 0$ there exists some $N \in \mathbb{N}$ such that

$$\|f_n - f_m\|_\infty = \sup_{x \in [a, b]} |f_n(x) - f_m(x)| < \varepsilon \quad \text{for all } x \in [a, b] \text{ and for all } n, m \geq N.$$

This means exactly that the sequence $\{f_n\}$ is a uniform Cauchy sequence. \square

Remark 6.28 (on definition of convergence and Cauchy sequence)

From Examples 6.24 and 6.27 we see that the definitions of convergence and the definition of a Cauchy sequence that we learned previously for particular examples of metric spaces or normed linear spaces (for example, $(\mathbb{R}, d(x, y) = |x - y|)$ and $(C([a, b]), \|\cdot\|_\infty)$) are just **special cases** of the general Definitions 6.23 and 6.25.

Now we will investigate convergence in the normed linear space \mathbb{R}^n with the Euclidean norm. The next lemma establishes that in order to test whether a sequence $\{\mathbf{x}^{(k)}\} \subset \mathbb{R}^n$ or vectors $\mathbf{x}^{(k)} = (x_1^{(k)}, x_2^{(k)}, \dots, x_n^{(k)})$ converges it is enough to verify whether the component sequences $\{x_j^{(k)}\} \subset \mathbb{R}$ converge for all $j = 1, 2, \dots, n$. This leads back to the definition of convergence of sequences of real numbers which you learnt in your first year at university.

Lemma 6.29 (criterion for convergence in \mathbb{R}^n)

Consider the linear space \mathbb{R}^n with the Euclidean norm $\|\mathbf{x}\|_2 := \left(\sum_{j=1}^n x_j^2\right)^{1/2}$.

- (i) A sequence $\{\mathbf{x}^{(k)}\} = \{(x_1^{(k)}, x_2^{(k)}, \dots, x_n^{(k)})\}$ in $(\mathbb{R}^n, \|\cdot\|_2)$ converges if and only if for each $j = 1, 2, \dots, n$ the component sequence $\{x_j^{(k)}\}$ converges in $(\mathbb{R}, |\cdot|)$.
- (ii) A sequence $\{\mathbf{x}^{(k)}\} = \{(x_1^{(k)}, x_2^{(k)}, \dots, x_n^{(k)})\}$ in $(\mathbb{R}^n, \|\cdot\|_2)$ is a Cauchy sequence if and only if for each index $j = 1, 2, \dots, n$ the component sequence $\{x_j^{(k)}\}$ is a Cauchy sequence in $(\mathbb{R}, |\cdot|)$.

Proof of Lemma 6.29: The proof of Lemma 6.29 is based on the use of

$$0 \leq |x_j - y_j| \leq \|\mathbf{x} - \mathbf{y}\|_2 \leq \sqrt{n} \max_{1 \leq i \leq n} |x_i - y_i|, \quad (6.14)$$

which can be proved as follows: Fixing an index j with $1 \leq j \leq n$, we have

$$(x_j - y_j)^2 \leq \underbrace{(x_1 - y_1)^2 + \dots + (x_n - y_n)^2}_{=\|\mathbf{x} - \mathbf{y}\|_2^2} \leq n \max_{1 \leq i \leq n} (x_i - y_i)^2 = n \left(\max_{1 \leq i \leq n} |x_i - y_i| \right)^2,$$

and taking the square root yields (6.14).

Now we give the proofs with the help of (6.14).

(i) \Rightarrow : Assume that $\{\mathbf{x}^{(k)}\}$ converges to $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$, that is, we have $\lim_{k \rightarrow \infty} \|\mathbf{x}^{(k)} - \mathbf{x}\|_2 = 0$. Then (6.14) implies that

$$0 \leq \lim_{k \rightarrow \infty} |x_j^{(k)} - x_j| \leq \lim_{n \rightarrow \infty} \|\mathbf{x}^{(k)} - \mathbf{x}\|_2 = 0 \quad \text{for all } j = 1, 2, \dots, n,$$

and, from the sandwich theorem,

$$\lim_{k \rightarrow \infty} |x_j^{(k)} - x_j| = 0 \quad \text{for all } j = 1, 2, \dots, n.$$

Thus, for each $j = 1, 2, \dots, n$, the sequence $\{x_j^{(k)}\}$ converges to x_j .

\Leftarrow : Assume that for each $j = 1, 2, \dots, n$, the sequence $\{x_j^{(k)}\}$ converges to some number $x_j \in \mathbb{R}$. This means that for every $\varepsilon > 0$, there exists for each $j = 1, 2, \dots, n$ a number $N_j = N_j(\varepsilon) \in \mathbb{N}$ such that

$$|x_j^{(k)} - x_j| < \frac{\varepsilon}{\sqrt{n}} \quad \text{for all } k \geq N_j. \quad (6.15)$$

Define $N := \max\{N_1, N_2, \dots, N_n\}$. Then from (6.15), we have

$$|x_j^{(k)} - x_j| < \frac{\varepsilon}{\sqrt{n}} \quad \text{for all } k \geq N \text{ and for all } j = 1, 2, \dots, n.$$

Thus

$$\max_{1 \leq j \leq n} |x_j^{(k)} - x_j| < \frac{\varepsilon}{\sqrt{n}} \quad \text{for all } k \geq N. \quad (6.16)$$

From (6.16) and (6.14), we see that with $\mathbf{x} := (x_1, x_2, \dots, x_n)$

$$\|\mathbf{x}^{(k)} - \mathbf{x}\|_2 \leq \sqrt{n} \max_{1 \leq j \leq n} |x_j^{(k)} - x_j| < \sqrt{n} \frac{\varepsilon}{\sqrt{n}} = \varepsilon \quad \text{for all } k \geq N.$$

Since $\varepsilon > 0$ was arbitrary, we see that $\{\mathbf{x}^{(k)}\}$ converges to \mathbf{x} .

(ii) \Rightarrow : Assume that $\{\mathbf{x}_k\}$ is a Cauchy sequence, that is, for every $\varepsilon > 0$ there exists an $N = N(\varepsilon) \in \mathbb{N}$ such that

$$\|\mathbf{x}^{(k)} - \mathbf{x}^{(m)}\|_2 < \varepsilon \quad \text{for all } k, m \geq N.$$

Then from (6.14) for all $j = 1, 2, \dots, n$,

$$|x_j^{(k)} - x_j^{(m)}| \leq \|\mathbf{x}^{(k)} - \mathbf{x}^{(m)}\|_2 < \varepsilon \quad \text{for all } k, m \geq N.$$

Since $\varepsilon > 0$ was arbitrary, we see that, for each $j = 1, 2, \dots, n$, the sequence $\{x_j^{(k)}\}$ is a Cauchy sequence in \mathbb{R} .

\Leftarrow : Let $\varepsilon > 0$ be arbitrary, and assume that for every $j = 1, 2, \dots, n$, $\{x_j^{(k)}\}$ is a Cauchy sequence, that is, there exists $N_j = N_j(\varepsilon) \in \mathbb{N}$ such that

$$|x_j^{(k)} - x_j^{(m)}| < \frac{\varepsilon}{\sqrt{n}} \quad \text{for all } k, m \geq N_j. \quad (6.17)$$

Define $N := \max\{N_1, N_2, \dots, N_n\}$. Then from (6.17)

$$|x_j^{(k)} - x_j^{(m)}| < \frac{\varepsilon}{\sqrt{n}} \quad \text{for all } k, m \geq N \text{ and for all } j = 1, 2, \dots, n.$$

Thus

$$\max_{1 \leq j \leq n} |x_j^{(k)} - x_j^{(m)}| < \frac{\varepsilon}{\sqrt{n}} \quad \text{for all } k, m \geq N. \quad (6.18)$$

From (6.14) and (6.18), we see that with $\mathbf{x} := (x_1, x_2, \dots, x_n)$

$$\|\mathbf{x}^{(k)} - \mathbf{x}^{(m)}\|_2 \leq \sqrt{n} \max_{j=1,2,\dots,n} |x_j^{(k)} - x_j^{(m)}| < \sqrt{n} \frac{\varepsilon}{\sqrt{n}} = \varepsilon \quad \text{for all } k \geq N.$$

Since $\varepsilon > 0$ was arbitrary, we see that $\{\mathbf{x}^{(k)}\}$ is a Cauchy sequence in \mathbb{R}^n . \square

With the help of Lemma 6.29, we can now easily prove a Cauchy principle for $(\mathbb{R}^n, \|\cdot\|_2)$.

Theorem 6.30 (Cauchy principle for $(\mathbb{R}^n, \|\cdot\|_2)$)

Consider the linear space \mathbb{R}^n with the Euclidean norm $\|\mathbf{x}\|_2 = \left(\sum_{j=1}^n x_j^2\right)^{1/2}$. A sequence $\{\mathbf{x}^{(k)}\} \subset \mathbb{R}^n$ is convergent if and only if it is a Cauchy sequence.

Proof of Theorem 6.30. This follows almost immediately from Lemma 6.29 and the Cauchy principle for $(\mathbb{R}, |\cdot|)$ (see Lemma 5.15).

\Rightarrow : Let $\{\mathbf{x}^{(k)}\} \subset \mathbb{R}^n$ be convergent in $(\mathbb{R}^n, \|\cdot\|_2)$. From Lemma 6.29 (i), this means that all component sequences $\{x_j^{(k)}\}$, where $j = 1, 2, \dots, n$, are convergent. From the Cauchy principle for \mathbb{R} (see Lemma 5.15), this implies that for each $j \in \{1, 2, \dots, n\}$ the sequence $\{x_j^{(k)}\}$ is a Cauchy sequence. From Lemma 6.29 (ii), this implies that $\{\mathbf{x}^{(k)}\}$ is a Cauchy sequence in $(\mathbb{R}^n, \|\cdot\|_2)$.

\Leftarrow : Assume that $\{\mathbf{x}^{(k)}\}$ is a Cauchy sequence in $(\mathbb{R}^n, \|\cdot\|_2)$. Then from Lemma 6.29 (ii), for each $j = 1, 2, \dots, n$, the component sequence $\{x_j^{(k)}\}$ is a Cauchy sequence in $(\mathbb{R}, |\cdot|)$. Thus we know from the Cauchy principle for \mathbb{R} (see Lemma 5.15) that, for each $j = 1, 2, \dots, n$, the sequence $\{x_j^{(k)}\}$ converges in \mathbb{R} . From Lemma 6.29 (i), this implies that $\{\mathbf{x}^{(k)}\}$ converges in $(\mathbb{R}^n, \|\cdot\|_2)$. \square

6.2.2 Complete Metric Spaces and Complete Normed Linear Spaces

Now we define the notion of **completeness** for metric spaces and normed linear spaces: a metric space or normed linear space is **complete** if every Cauchy sequence converges to some element in the space. We will give several examples of complete normed linear spaces and complete metric spaces.

Definition 6.31 (complete metric space)

Let X be a metric space with the distance function $d : X \times X \rightarrow \mathbb{R}$. The metric space (X, d) is called **complete** if every Cauchy sequence in X converges to some element x **in** X .

For the special case of normed linear spaces Definition 6.31 yields the following definition.

Definition 6.32 (complete normed linear space)

Let X be a normed linear space with norm $\|\cdot\| : X \rightarrow \mathbb{R}$. The normed linear space $(X, \|\cdot\|)$ is called **complete** if every Cauchy sequence in X converges to some element x **in** X .

We will look at some examples. Our first example is \mathbb{R} with the absolute value norm which is a complete normed linear space.

Example 6.33 (\mathbb{R} with the absolute value norm)

The real line \mathbb{R} with the absolute value norm $|\cdot|$ is a complete normed linear space.

Proof: In Example 6.6, we have seen that \mathbb{R} with the absolute value norm $|\cdot|$ is a normed linear space. From the Cauchy principle (see Lemma 5.15) we know that a sequence $\{x_n\} \subset \mathbb{R}$ is a Cauchy sequence if and only if it is convergent to some $x \in \mathbb{R}$. Thus \mathbb{R} with the absolute value norm is complete. \square

Here is an example of a metric space that is **not** complete.

Example 6.34 (metric space $(0, 1]$ with the $d(x, y) := |x - y|$)

The metric space $(0, 1]$ with the metric $d(x, y) := |x - y|$ is not complete.

Proof: In Example 6.3 we proved that $(0, 1]$ with the metric $d(x, y) := |x - y|$ is a metric space. We want to show that it is not complete. Consider the sequence $\{1/n\} \subset (0, 1]$. Let $\varepsilon > 0$ be arbitrary, and let $N = N(\varepsilon) \in \mathbb{N}$ be given by $N := \min\{k \in \mathbb{N} : k > 1/\varepsilon\}$. Then $1/N < \varepsilon$, and for all $n, m \geq N$,

$$d\left(\frac{1}{n}, \frac{1}{m}\right) = \left|\frac{1}{n} - \frac{1}{m}\right| = \left|\frac{m-n}{nm}\right| \leq \frac{\max\{n, m\}}{nm} = \frac{1}{\min\{n, m\}} \leq \frac{1}{N} < \varepsilon.$$

Thus $\{1/n\}$ is a Cauchy sequence. We have

$$\lim_{n \rightarrow \infty} d\left(\frac{1}{n}, 0\right) = \lim_{n \rightarrow \infty} \left|\frac{1}{n} - 0\right| = \lim_{n \rightarrow \infty} \frac{1}{n} = 0,$$

but the limit 0 of $\{1/n\}$ is not in $(0, 1]$. Thus $((0, 1], d)$ is not complete. \square

We look at some more complicated examples.

Example 6.35 (\mathbb{R} with the discrete metric)

In Example 6.4, we saw that \mathbb{R} with the **discrete metric**

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

is a metric space. If $\{x_n\}$ is a Cauchy sequence, then for every $\varepsilon > 0$ there exists some $N = N(\varepsilon) \in \mathbb{N}$ such that

$$d(x_n, x_m) < \varepsilon \quad \text{for all } n, m \geq N, \quad (6.19)$$

and, in particular,

$$d(x_N, x_m) < \varepsilon \quad \text{for all } m \geq N,$$

Now let $\varepsilon = 1$. Then (6.19) with $\varepsilon = 1$ implies that $x_m = x_N$ for all $m \geq N$. Thus the series $\{x_n\}$ converges to $x_N \in \mathbb{R}$. We see that the metric space \mathbb{R} with the discrete metric is complete. \square

Example 6.36 (\mathbb{Q} with the absolute value norm)

Let \mathbb{Q} denote the set of rational numbers with the metric $d(x, y) := |x - y|$. Then $(\mathbb{Q}, |\cdot|)$ is a metric space, but $(\mathbb{Q}, |\cdot|)$ is not complete.

Proof: We have proved in Example 6.2 that the metric $d(x, y) := |x - y|$ is a metric for \mathbb{R} . Inspection of the norm properties shows that $d(x, y) = |x - y|$ is also a metric for any subset of \mathbb{R} . Thus $d(x, y) := |x - y|$ is a metric for \mathbb{Q} , and we see that (\mathbb{Q}, d) is a metric space.

It remains to show that \mathbb{Q} is not complete. Consider the number $\sqrt{2} \in \mathbb{R} \subset \mathbb{Q}$. Since \mathbb{Q} is dense in \mathbb{R} , we can construct a sequence $\{x_n\} \subset \mathbb{Q}$ that converges to $\sqrt{2}$. Since this sequence converges in \mathbb{R} we know from the Cauchy principle (see Lemma 5.15) that $\{x_n\}$ is a Cauchy sequence with respect to the absolute value norm and thus it is a Cauchy sequence in \mathbb{Q} . The limit of this Cauchy sequence is $\sqrt{2}$ which is not in \mathbb{Q} , and thus \mathbb{Q} is not complete. \square

Now we want to discuss the example of \mathbb{R}^n with the Euclidean norm. Lemma 6.29 immediately leads to the fact that \mathbb{R}^n with the Euclidean norm is complete.

Theorem 6.37 ($(\mathbb{R}^n, \|\cdot\|_2)$ is complete)

The linear space \mathbb{R}^n with the Euclidean norm $\|\mathbf{x}\|_2 := \left(\sum_{j=1}^n x_j^2\right)^{1/2}$ is complete.

Proof of Theorem 6.37: The statement follows from the Cauchy principle for $(\mathbb{R}^n, \|\cdot\|_2)$ (see Theorem 6.30). Indeed, let $\{\mathbf{x}^{(k)}\} \subset \mathbb{R}^n$ be a Cauchy sequence in $(\mathbb{R}^n, \|\cdot\|_2)$. Then we know from Theorem 6.30 that $\{\mathbf{x}^{(k)}\}$ is convergent to some $\mathbf{x} \in \mathbb{R}^n$. Thus $(\mathbb{R}^n, \|\cdot\|_2)$ is complete. \square

Now we come back to the normed linear space $\mathcal{C}([a, b])$ of continuous functions with the supremum norm $\|\cdot\|_\infty$. From the results in Chapter 5, we can easily deduce that $\mathcal{C}([a, b])$ with the supremum norm $\|\cdot\|_\infty$ is complete.

Theorem 6.38 ($\mathcal{C}([a, b])$ with supremum norm is complete)

Let $[a, b]$ be a bounded closed interval. The normed linear space $\mathcal{C}([a, b])$ of continuous functions on $[a, b]$ equipped with the supremum norm

$$\|f\|_\infty = \sup_{x \in [a, b]} |f(x)|, \quad f \in \mathcal{C}([a, b]),$$

is complete.

Proof of Theorem 6.38: The proof of this non-trivial result follows relatively easily with our knowledge from the previous subsections and from Chapter 5: In Lemma 6.15 (a) we have seen that $\mathcal{C}([a, b])$ with the supremum norm $\|\cdot\|_\infty$ is a normed linear space. From Example 6.27, we know that convergence in $(\mathcal{C}([a, b]), \|\cdot\|_\infty)$ is uniform convergence and that a Cauchy sequence in $(\mathcal{C}([a, b]), \|\cdot\|_\infty)$ is a uniform Cauchy sequence.

Now let $\{f_n\}$ be a Cauchy sequence in $\mathcal{C}([a, b])$. From the uniform Cauchy principle (see Theorem 5.18), we know that $\{f_n\}$ converges uniformly on $[a, b]$ to some function $f : [a, b] \rightarrow \mathbb{R}$. Since the f_n are continuous, from Theorem 5.19, the limit is also continuous, that is, $f \in \mathcal{C}([a, b])$. Thus $\mathcal{C}([a, b])$ equipped with the supremum norm is complete. \square

6.2.3 Bounded Sets and the Bolzano-Weierstrass Theorem for \mathbb{R}^n

In first year, you will have learned the **Bolzano-Weierstrass theorem** for the real numbers: every bounded sequence $\{x_k\}$ in \mathbb{R} (with the absolute value norm) has a convergent subsequence. Bounded means here that there exists some number $r \in \mathbb{R}$ such that $|x_k| \leq r$ for all $k \in \mathbb{N}$.

In this subsection we will introduce **bounded sets** in metric spaces and in normed linear spaces. Then we will prove the **Bolzano-Weierstrass theorem** for \mathbb{R}^n with the

Euclidean norm $\|\cdot\|_2$: **every bounded sequence in $(\mathbb{R}^n, \|\cdot\|_2)$ has a convergent subsequence.** It has to be noted that the Bolzano-Weierstrass theorem does **not** hold in arbitrary normed linear spaces (or arbitrary metric spaces), and we will give an example to illustrate this.

We start by defining bounded sets in a metric space.

Definition 6.39 (bounded set in a metric space)

Let (X, d) be a metric space. The following definitions of a bounded subset of X are equivalent:

- (i) A set $E \subset X$ is said to be **bounded** if there exist $x_0 \in X$ and $r > 0$ such that

$$d(x, x_0) \leq r \quad \text{for all } x \in E.$$

- (ii) A set $E \subset X$ is said to be **bounded** if **for every** $x_0 \in X$ there exists $r > 0$ such that we have

$$d(x, x_0) \leq r \quad \text{for all } x \in E.$$

For the special case of a normed linear space Definition 6.39 implies the following definition.

Definition 6.40 (bounded set in a normed linear space)

Let $(X, \|\cdot\|)$ be a normed linear space. The following definitions of a bounded subset of X are equivalent:

- (i) A set $E \subset X$ is said to be **bounded** if there exist $x_0 \in X$ and $r > 0$ such that

$$\|x - x_0\| \leq r \quad \text{for all } x \in E.$$

- (ii) A set $E \subset X$ is said to be **bounded** if **for every** $x_0 \in X$ there exists $r > 0$ such that

$$\|x - x_0\| \leq r \quad \text{for all } x \in E.$$

- (iii) A set $E \subset X$ is **bounded**, if there exists some $r > 0$ such that

$$\|x\| \leq r \quad \text{for all } x \in E.$$

Example 6.41 (bounded sets in \mathbb{R})

The intervals $[a, b]$, (a, b) , $(a, b]$ and $[a, b)$ with $-\infty < a < b < \infty$ are bounded in \mathbb{R} with the absolute value norm $|\cdot|$.

Proof: We want to use Definition 6.40 (iii). Let $r := \max\{|a|, |b|\}$. Then we have for all $x \in [a, b]$ that $a \leq x \leq b$ and thus

$$|x| \leq \max\{|a|, |b|\} = r. \quad (6.20)$$

Since $(a, b) \subset [a, b]$, $(a, b) \subset [a, b]$, and $[a, b) \subset [a, b]$, we see that the estimate (6.20) holds also for all x in these intervals as well. Thus $[a, b]$, (a, b) , $(a, b]$, and $[a, b)$ with $-\infty < a < b < \infty$ are bounded in $(\mathbb{R}, |\cdot|)$. \square

Remark 6.42 (on Definitions 6.39 and 6.40)

*We want to investigate why the statements in Definitions 6.39 and 6.40 are **equivalent**.*

In both Definitions 6.39 and 6.40 it is clear that (ii) implies (i), but the reverse direction needs explaining.

Proof of (i) \Rightarrow (ii): Assume that there exist $x_0 \in X$ and $r > 0$ such that

$$d(x, x_0) \leq r \quad \text{for all } x \in E. \quad (6.21)$$

Now consider any other point $y_0 \in \mathbb{R} \setminus \{x_0\}$. Then we have from the triangle inequality and (6.21)

$$d(x, y_0) \leq d(x, x_0) + d(x_0, y_0) \leq r + d(x_0, y_0) \quad \text{for all } x \in E.$$

Thus with $\tilde{r} := r + d(x_0, y_0)$, we have

$$d(x, y_0) \leq \tilde{r} \quad \text{for all } x \in E,$$

and thus (ii) holds true in Definition 6.39. Since Definition 6.40 is just a special case of Definition 6.39 where the metric is defined by $d(x, y) := \|x - y\|$, we have also proved that (i) implies (ii) in Definition 6.40. \square

It remains to show that (iii) in Definition 6.40 is equivalent to (i) and (ii).

Since (ii) implies that for $x_0 = \mathcal{O}$ there exists $r > 0$ such that $\|x - x_0\| = \|x - \mathcal{O}\| = \|x\| \leq r$ for all $x \in E$, (ii) clearly implies (iii). Assume now that (iii) is true, that is, $\|x\| \leq r$ for all $x \in E$, and consider an arbitrary $x_0 \in X$. Then

$$\|x - x_0\| \leq \|x\| + \|x_0\| \leq r + \|x_0\| \quad \text{for all } x \in E.$$

Since $r + \|x_0\|$ is a positive constant, this implies (ii). \square

Remark 6.43 (on the Definitions 6.39 and 6.40)

The fact that (i) implies (ii) tells us the following: if for one $x_0 \in X$ there exists an $r = r(x_0) > 0$ such that $d(x, x_0) \leq r$ for all $x \in E$, and $\|x - x_0\| \leq r$ for all $x \in E$, respectively, then **for every** $y_0 \in X$ there exists an $\tilde{r} = \tilde{r}(y_0) > 0$ such that $d(x, y_0) \leq \tilde{r}$ for all $x \in E$, and $\|x - y_0\| \leq \tilde{r}$ for all $x \in E$, respectively. Thus for investigating whether E is bounded, it is enough to check whether

$$d(x, x_0) \leq r \quad \text{for all } x \in E, \quad \text{and} \quad \|x - x_0\| \leq r \quad \text{for all } x \in E, \quad (6.22)$$

respectively, holds true for one particular x_0 . (Indeed if (6.22) was not true for the x_0 of our choice then Definition 6.39 (ii), and Definition 6.40 (ii), respectively, would be violated.)

From the observations in the remark above we can draw the following conclusions.

Corollary 6.44 (test for boundedness in a metric space)

Let (X, d) be a metric space, and let E be a subset of X . Let $x_0 \in X$ be an **arbitrary** point in X . If there exists some $r > 0$ such that

$$d(x, x_0) \leq r \quad \text{for all } x \in E, \quad (6.23)$$

then E is **bounded**, and if it is impossible to find an $r > 0$ such that (6.23) is true then E is **not bounded**.

For the special case of a normed linear space we have the following corollary.

Corollary 6.45 (test for boundedness in a normed linear space)

Let $(X, \|\cdot\|)$ be a normed linear space, and let E be a subset of X . Let $x_0 \in X$ be an **arbitrary** point in X . If there exists some $r > 0$ such that

$$\|x - x_0\| \leq r \quad \text{for all } x \in E, \quad (6.24)$$

then E is **bounded**, and if it is impossible to find an $r > 0$ such that (6.24) is true then E is **not bounded**.

Since Corollary 6.45 follows from Corollary 6.44 by replacing the metric d by its definition $d(x, y) := \|x - y\|$, it is enough to prove Corollary 6.44.

Proof of Corollary 6.44: The first statement is essentially the definition (i) of boundedness (see Definition 6.39).

The second statement can be easily proved by contradiction: Assume that for the given x_0 it is impossible to find an $r > 0$ such that (6.23) holds true, but E is still

bounded. Then we know from (ii) in Definition 6.39 that for every $x_0 \in X$ there exists an $r > 0$ such that

$$d(x, x_0) \leq r \quad \text{for all } x \in X.$$

This is a contradiction to our assumption, and therefore E is unbounded. \square

We give some examples of bounded sets in metric spaces and normed linear spaces.

Example 6.46 (bounded sets in \mathbb{R}^n)

For the Euclidean space \mathbb{R}^n with the Euclidean norm $\|\cdot\|_2$ we have: a set E in \mathbb{R}^n is bounded if there exists $r > 0$ such that

$$d(\mathbf{x}, \mathbf{0}) = \|\mathbf{x} - \mathbf{0}\|_2 = \|\mathbf{x}\|_2 = \left(\sum_{k=1}^n x_k^2 \right)^{1/2} \leq r \quad \text{for all } \mathbf{x} \in E. \quad \square$$

Example 6.47 (bounded set in $(\mathcal{C}([0, 2]), \|\cdot\|_\infty)$)

The subset

$$E := \{f_n : [0, 2] \rightarrow \mathbb{R}, f_n(x) := e^{-x/n} : n \in \mathbb{N}\}$$

of continuous functions is bounded in $(\mathcal{C}([0, 2]), \|\cdot\|_\infty)$.

Proof: We have for every $n \in \mathbb{N}$ that

$$\|f_n\|_\infty = \|e^{-x/n}\|_\infty = \sup_{x \in [0, 2]} |e^{-x/n}| = e^{-0/n} = e^0 = 1.$$

Since every function $f_n(x) = e^{-x/n}$ in E satisfies $\|f_n\|_\infty = \|e^{-x/n}\|_\infty = 1$, we see that

$$\|f_n\|_\infty = \|e^{-x/n}\|_\infty \leq 1 \quad \text{for all } n \in \mathbb{N},$$

and thus the set $E \subset \mathcal{C}([0, 2])$ is bounded. \square

Example 6.48 (unbounded set in $(\mathcal{C}([0, 1]), \|\cdot\|_\infty)$)

Investigate whether the set

$$E := \{f_n : [0, 1] \rightarrow \mathbb{R}, f_n(x) := n e^{-x/n} : n \in \mathbb{N}\}$$

is bounded as a subset of $\mathcal{C}([0, 1])$ with the supremum norm $\|\cdot\|_\infty$.

Solution: The set E is unbounded in $(\mathcal{C}([0, 1]), \|\cdot\|_\infty)$. This can be seen as follows: We have

$$\|f_n\|_\infty = \sup_{x \in [0, 1]} |n e^{-x/n}| = n e^{-0/n} = n e^0 = n, \quad n \in \mathbb{N}.$$

Thus for every $r > 0$, we can choose $N \in \mathbb{N}$ such that $r < N$, and thus

$$\|f_n\|_\infty = \sup_{x \in [0,1]} |n e^{-x/n}| = n > r, \quad \text{for all } n \geq N.$$

This proves that E is unbounded. \square

Example 6.49 (bounded set in $(\mathcal{C}([0, 1]), \|\cdot\|_1)$)

Let the set $E \subset \mathcal{C}([0, 1])$ be defined by

$$E := \{f_\alpha : [0, 1] \rightarrow \mathbb{R}, f_\alpha(x) := x^\alpha : \alpha \in \mathbb{R} \text{ with } \alpha \geq 0\}.$$

The set E is bounded in $(\mathcal{C}([0, 1]), \|\cdot\|_1)$, where $\|\cdot\|_1$ is the L_1 -norm

$$\|f\|_1 = \int_0^1 |f(x)| dx, \quad f \in \mathcal{C}([0, 1]).$$

Proof: We work out the norms of the functions in E . For any $\alpha \in \mathbb{R}$ with $\alpha \geq 0$,

$$\|f_\alpha\|_1 = \int_0^1 |f_\alpha(x)| dx = \int_0^1 |x^\alpha| dx = \int_0^1 x^\alpha dx = \left. \frac{x^{\alpha+1}}{\alpha+1} \right|_0^1 = \frac{1}{\alpha+1}.$$

We have

$$\sup_{\alpha \geq 0} \|f_\alpha\|_1 = \sup_{\alpha \geq 0} \frac{1}{\alpha+1} = 1,$$

and thus, for every $\alpha \geq 0$, the function $f_\alpha(x) = x^\alpha$ satisfies $\|f_\alpha\|_1 \leq 1$. Thus the set E is bounded in $(\mathcal{C}([0, 1]), \|\cdot\|_1)$. \square

To prove the Bolzano-Weierstrass theorem for \mathbb{R}^n , we need the Bolzano-Weierstrass theorem for \mathbb{R} (from first year) which is stated below.

Theorem 6.50 (Bolzano-Weierstrass theorem for \mathbb{R})

Every bounded sequence $\{x_k\}$ in \mathbb{R} with the absolute value norm $|\cdot|$ has a convergent subsequence.

Now we can prove the Bolzano-Weierstrass theorem for \mathbb{R}^n .

Theorem 6.51 (Bolzano-Weierstrass Theorem for \mathbb{R}^n)

Every bounded sequence $\{\mathbf{x}^{(k)}\}$ in \mathbb{R}^n with the Euclidean norm $\|\cdot\|_2$ has a convergent subsequence.

Proof of Theorem 6.51 Let $\{\mathbf{x}^{(k)}\}$ be a bounded sequence in \mathbb{R}^n , that is, there exists $r > 0$ such that $\|\mathbf{x}^{(k)}\|_2 \leq r$ for all $k \in \mathbb{N}$. This implies that every component

sequence $\{x_j^{(k)}\}$ is bounded: for each $j = 1, 2, \dots, n$ and for all $k \in \mathbb{N}$

$$|x_j^{(k)}| \leq \left(\sum_{i=1}^n (x_i^{(k)})^2 \right)^{1/2} = \|\mathbf{x}^{(k)}\| \leq r. \quad (6.25)$$

Thus we know from the Bolzano-Weierstrass theorem in \mathbb{R} (see Theorem 6.50 above) that, since $\{x_1^{(k)}\}$ is bounded, $\{x_1^{(k)}\}$ has a convergent subsequence which we will denote by $\{x_1^{(k_1)}\}$ (where $\{k_1\}$ is a subsequence of $\{k\} = \mathbb{N}$). Now we consider $\{x_2^{(k_1)}\}$ which is bounded according to (6.25). (Note that we do only consider the subsequence of $\{x_2^{(k)}\}$ corresponding to the selection $k = k_1$ from the first step.) From the Bolzano-Weierstrass theorem, we know that, since $\{x_1^{(k_1)}\}$ is bounded there exists a convergent subsequence $\{x_2^{(k_2)}\}$ of $\{x_2^{(k_1)}\}$ (where $\{k_2\} \subset \{k_1\}$). We continue this process and find finally that there exists a convergent subsequence $\{x_n^{(k_n)}\}$. Since the sets of indices are nested, that is $\{k_n\} \subset \{k_{n-1}\} \subset \dots \subset \{k_2\} \subset \{k_1\} \subset \mathbb{N}$, we see that $\{x_j^{(k_n)}\}$ is a convergent subsequence for all $j = 1, 2, \dots, n$. Thus from Lemma 6.29, we know that $\{\mathbf{x}^{(k_n)}\}$ is a convergent subsequence of $\{\mathbf{x}^{(k)}\}$. \square

The Bolzano-Weierstrass theorem does **not** hold in arbitrary normed linear spaces.

Remark 6.52 (no Bolzano-Weierstrass Theorem for $(\mathcal{C}([a, b]), \|\cdot\|_\infty)$)

*The Bolzano-Weierstrass Theorem does **not** hold for the space of continuous functions with the supremum norm: A bounded sequence $\{f_n\}$ in $\mathcal{C}([a, b])$ with the supremum norm (that is, $\{f_n\}$ satisfies $\|f_n\|_\infty \leq r$ for all $n \in \mathbb{N}$ with some fixed $r > 0$) does in general **not** have a convergent subsequence.*

For example, consider $\mathcal{C}([0, 1])$ and let $\{f_n\}$ be defined by $f_n(x) := x^n$, $x \in [0, 1]$. Then $\{f_n\}$ and any subsequence $\{f_{n_1}\}$ of $\{f_n\}$ converge pointwise to the function $f : [0, 1] \rightarrow \mathbb{R}$ defined by

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

Since the pointwise limit is not continuous, we know that the convergence is not uniform, and the convergence of any subsequence will also **not** be uniform. The sequence $\{f_n\}$ is bounded because $\|f_n\|_\infty = \|x^n\|_\infty = 1$ for all $n \in \mathbb{N}$, but it does not have any uniformly convergent subsequence.

6.3 Open and Closed Subsets in Metric and Normed Linear Spaces

In this section we discuss **open sets** and **closed sets** in metric spaces and normed linear spaces and related terminology.

In Subsection 6.3.1, we start by defining **open balls** and **closed balls** in metric spaces and normed linear spaces. Then we use the open balls to define **interior points** of a subset of a metric space and normed linear space. With the notion of interior points we can then finally define **open subsets** and **closed subsets**

In Subsection 6.3.2, we define the notion of an **accumulation point** of a subset of a metric space or normed linear space. The concept of an accumulation point can be also described with the help of sequences, and will derive various useful **characterizations of closed subsets** with the help of the concept of accumulation points.

6.3.1 Interior Points and Open and Closed Sets

We start by introducing **open balls** and **closed balls** in metric spaces.

Definition 6.53 (open and closed ball in a metric space)

Let (X, d) be a metric space. The **open ball** centred at $y \in X$ with radius $r > 0$ is defined by

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

and the **closed ball** centred at $y \in X$ with radius $r > 0$ is defined by

$$\overline{B}(y; r) := \{x \in X : d(x, y) \leq r\}.$$

For the special case of a normed linear space we obtain from Definition 6.53 the following definition.

Definition 6.54 (open and closed ball in a normed linear space)

Let $(X, \|\cdot\|)$ be a normed linear space. The **open ball** centred at $y \in X$ with radius $r > 0$ is defined by

$$B(y; r) := \{x \in X : \|x - y\| < r\},$$

and the **closed ball** centred at $y \in X$ with radius $r > 0$ is defined by

$$\overline{B}(y; r) := \{x \in X : \|x - y\| \leq r\}.$$

We give some examples of open and closed balls.

Example 6.55 (open and closed balls in $(\mathbb{R}^n, \|\cdot\|_2)$)

In the Euclidean space \mathbb{R}^n with the Euclidean norm $\|\cdot\|_2$. The **open ball** centred at $\mathbf{y} \in \mathbb{R}^n$ with radius $r > 0$ is given by

$$B(\mathbf{y}; r) := \left\{ \mathbf{x} \in \mathbb{R}^n : \|\mathbf{x} - \mathbf{y}\|_2 = \left(\sum_{k=1}^n (y_k - x_k)^2 \right)^{1/2} < r \right\},$$

and the **closed ball** centred at \mathbf{y} with radius $r > 0$ is given by

$$\overline{B}(\mathbf{y}; r) = \left\{ \mathbf{x} \in \mathbb{R}^n : \|\mathbf{x} - \mathbf{y}\|_2 = \left(\sum_{k=1}^n (y_k - x_k)^2 \right)^{1/2} \leq r \right\},$$

respectively. A little later we will see that the open ball $B(\mathbf{y}; r)$ is really open and that the closed ball $\overline{B}(\mathbf{y}; r)$ is really closed. \square

For the case of the plane \mathbb{R}^2 with a given norm we can easily plot open and closed balls, and we will do this in the next example.

Example 6.56 (open balls and closed balls \mathbb{R}^2 with various norms)

Define and plot the open ball $B(0; 1)$ and closed ball $\overline{B}(0; 1)$ of radius $r = 1$ in \mathbb{R}^2 endowed with each of the following norms:

- (a) $\|(x_1, x_2)\|_1 := |x_1| + |x_2|;$
- (b) $\|(x_1, x_2)\|_2 = \sqrt{x_1^2 + x_2^2};$
- (c) $\|(x_1, x_2)\|_\infty := \max \{|x_1|, |x_2|\}.$

Solution:

(a) In $(\mathbb{R}^2, \|\cdot\|_1)$, we have

$$B(0; 1) = \{(x_1, x_2) : |x_1| + |x_2| < 1\} \quad \text{and} \quad \overline{B}(0; 1) = \{(x_1, x_2) : |x_1| + |x_2| \leq 1\},$$

and we have plotted $B(0; 1)$ and $\overline{B}(0; 1)$ in the left picture of Figure 6.1. The open ball $B(0; 1)$ is the interior of the slanted square (without the boundary), whereas the closed ball $\overline{B}(0; 1)$ is the slanted square including the boundary.

(b) In $(\mathbb{R}^2, \|\cdot\|_2)$, we have

$$B(0; 1) = \left\{ (x_1, x_2) : \sqrt{x_1^2 + x_2^2} < 1 \right\} \quad \text{and} \quad \overline{B}(0; 1) = \left\{ (x_1, x_2) : \sqrt{x_1^2 + x_2^2} \leq 1 \right\},$$

and we have plotted $B(0; 1)$ and $\overline{B}(0; 1)$ in the middle picture of Figure 6.1. The open ball $B(0; 1)$ is the interior of the disc (without the boundary), whereas the closed ball $\overline{B}(0; 1)$ is the disc including the boundary.

(c) In $(\mathbb{R}^2, \|\cdot\|_\infty)$, we have

$$B(0; 1) = \{(x_1, x_2) : \max\{|x_1|, |x_2|\} < 1\} \quad \text{and} \\ \overline{B}(0; 1) = \{(x_1, x_2) : \max\{|x_1|, |x_2|\} \leq 1\},$$

and we have plotted $B(0; 1)$ and $\overline{B}(0; 1)$ in the right picture of Figure 6.1. The open ball $B(0; 1)$ is the interior of the square (without the boundary), whereas the closed ball $\overline{B}(0; 1)$ is the square including the boundary. \square

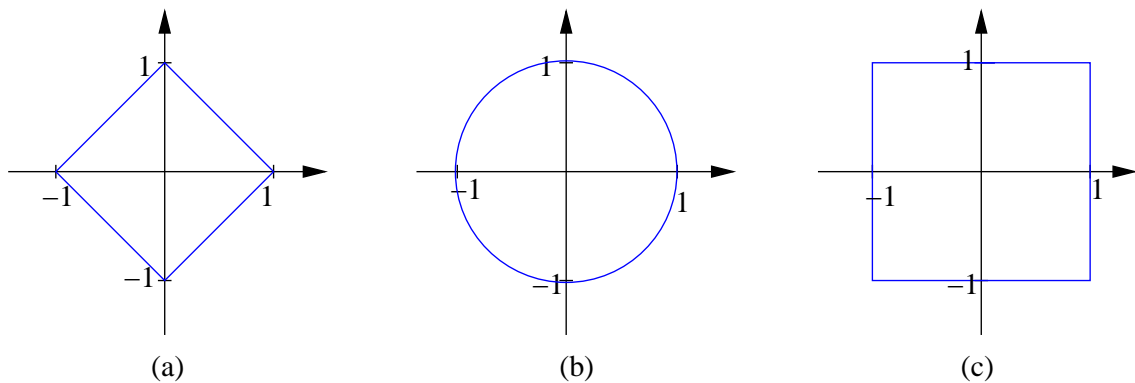


Figure 6.1: The open ball $B(0; 1)$ and closed ball $\overline{B}(0; 1)$ in the space (a) $(\mathbb{R}^2, \|\cdot\|_1)$, (b) $(\mathbb{R}^2, \|\cdot\|_2)$, and (c) $(\mathbb{R}^2, \|\cdot\|_\infty)$. In each plot $B(0; 1)$ is given by the area without the boundary, and $\overline{B}(0; 1)$ is given by the area including the boundary.

We have seen in the last example that balls are not always balls as we imagine them geometrically.

We consider another example.

Example 6.57 (open and closed balls in \mathbb{R} with the discrete metric)

In Example 6.4, we saw that \mathbb{R} with the **discrete metric**

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

is a metric space. We want to determine what the open and closed balls in (\mathbb{R}, d) are. Consider the open ball $B(y; r)$. From the definition of the discrete metric we see that

$$B(y; r) := \{x \in \mathbb{R} : d(x, y) < r\} = \begin{cases} \{y\} & \text{if } r \leq 1, \\ \mathbb{R} & \text{if } r > 1. \end{cases}$$

Likewise we see that the closed ball $\overline{B(y; r)}$ is given by

$$\overline{B(y; r)} := \{x \in \mathbb{R} : d(x, y) \leq r\} = \begin{cases} \{y\} & \text{if } r < 1, \\ \mathbb{R} & \text{if } r \geq 1. \end{cases}$$

So we see that the open ball $B(y; r)$ and the closed ball $\overline{B(y; r)}$ are either $\{y\}$ or \mathbb{R} , depending on the value of r . \square

Now we introduce, with the help of open balls, the notion of an **interior point** of a subset of a metric space. Then we can define **open (sub)sets** and **closed (sub)sets** in a metric space.

Definition 6.58 (interior points in a metric space)

Let (X, d) be a metric space, and let E be a subset of X . An element $y \in E$ is called an **interior point** of E if for some $r > 0$ we have $B(y; r) \subset E$. We write $\overset{\circ}{E}$ for the **set of all interior points** of E .

Definition 6.59 (open set in a metric space)

Let (X, d) be a metric space and let E be a subset of X . The set E is called **open** if every point in E is an interior point, that is, if $E = \overset{\circ}{E}$.

Definition 6.60 (closed set in a metric space)

Let (X, d) be a metric space, and let E be a subset of X . The set E is called **closed** if its **complement**

$$X \setminus E := \{x \in X : x \notin E\}$$

is open.

Note that the Definitions 6.58, 6.59, and 6.60 define automatically interiors points, open sets, and closed sets in a normed linear space, since every normed linear space $(X, \|\cdot\|)$ is, according to Lemma 6.8, a metric space with the metric

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

Thus we obtain the following definitions for the special case of a normed linear space.

Definition 6.61 (interior points in a normed linear space)

Let $(X, \|\cdot\|)$ be a normed linear space, and let E be a subset of X . An element $y \in E$ is called an **interior point** of E if for some $r > 0$ we have $B(y; r) \subset E$. We write $\overset{\circ}{E}$ for the **set of all interior points** of E .

Definition 6.62 (open set in a normed linear)

Let $(X, \|\cdot\|)$ be a normed linear space, and let E be a subset of X . The set E is called **open** if every point in E is an interior point, that is, if $E = \overset{\circ}{E}$.

Definition 6.63 (closed set in a normed linear)

Let $(X, \|\cdot\|)$ be a normed linear space, and let E be a subset of X . The set E is called **closed** if its **complement**

$$X \setminus E := \{x \in X : x \notin E\}$$

is open.

A direct consequence of Definitions 6.60 and 6.63 is the following lemma:

Lemma 6.64 (complement of an open set is closed)

Consider a metric space (X, d) or a normed linear space $(X, \|\cdot\|)$, and let E be a subset of X . The set E is open if and only if its complement $X \setminus E$ is closed.

Proof of Lemma 6.64 \Rightarrow : Suppose E is open. Then we see that $X \setminus E$ is closed because $X \setminus (X \setminus E) = E$ is open.

\Leftarrow : Now if $X \setminus E$ is closed, by definition its complement is open, that is, the set $X \setminus (X \setminus E) = E$ is open. \square

Now we will discuss some examples to become familiar with the notions of open and closed sets.

Example 6.65 (open and closed sets in \mathbb{R})

- (a) $(0, 1]$ is neither open nor closed in \mathbb{R} .
- (b) $(0, 1)$ is an open subset of \mathbb{R} .
- (c) $[0, 1]$ is a closed subset of \mathbb{R} .

Proof: We prove the three statements.

- (a) Consider the point $x = 1$ in $(0, 1]$. For every $\varepsilon > 0$, the open ball

$$B(1; \varepsilon) = \{x \in \mathbb{R} : |1 - x| < \varepsilon\}$$

contains the number $1 + \varepsilon/2$ (since $|1 - (1 + \varepsilon/2)| = |-\varepsilon/2| = \varepsilon/2 < \varepsilon$) which is not in $(0, 1]$. Thus 1 is not an interior point and $(0, 1]$ is not open. To show that $(0, 1]$ is not closed we consider the complement

$$\mathbb{R} \setminus (0, 1] = (-\infty, 0] \cup (1, \infty)$$

and show that it is not open. Consider the point $x = 0$ in $\mathbb{R} \setminus (0, 1]$. For every $\varepsilon > 0$, the open ball $B(0; \varepsilon)$ contains the number $\min\{\varepsilon/2, 1/2\}$, since

$$\left|0 - \min\left\{\frac{\varepsilon}{2}, \frac{1}{2}\right\}\right| = \min\left\{\frac{\varepsilon}{2}, \frac{1}{2}\right\} \leq \frac{\varepsilon}{2} < \varepsilon,$$

and from

$$0 < \min\left\{\frac{\varepsilon}{2}, \frac{1}{2}\right\} \leq \frac{1}{2},$$

we see that $\min\{\varepsilon/2, 1/2\}$ lies in $(0, 1]$. Thus $\min\{\varepsilon/2, 1/2\}$ lies not in $\mathbb{R} \setminus (0, 1]$, and consequently $B(0; \varepsilon)$ is not a subset of $\mathbb{R} \setminus (0, 1]$. Thus $0 \in \mathbb{R} \setminus (0, 1]$ is not an interior point of $\mathbb{R} \setminus (0, 1]$, and $\mathbb{R} \setminus (0, 1]$ is not open. Consequently $(0, 1]$ is not closed.

- (b) Consider an arbitrary point in $x \in (0, 1)$. Take

$$\delta := \min\left\{\frac{x - 0}{2}, \frac{1 - x}{2}\right\}$$

We claim that $B(x; \delta) = \{y \in \mathbb{R} : |y - x| < \delta\}$ is a subset of $(0, 1)$. If this is true then any $x \in (0, 1)$ is an interior point and hence $(0, 1)$ is open.

Now we show that $B(x; \delta) \subset (0, 1)$. Consider any $y \in B(x; \delta)$, then

$$\begin{aligned} y = x + (y - x) &\leq x + |y - x| \\ &< x + \delta = x + \min \left\{ \frac{x - 0}{2}, \frac{1 - x}{2} \right\} \\ &\leq x + \frac{1 - x}{2} = \frac{1 + x}{2} < 1 \quad \text{since } x < 1, \end{aligned}$$

and

$$\begin{aligned} y = x + (y - x) &\geq x - |y - x| \\ &> x - \delta = x - \min \left\{ \frac{x - 0}{2}, \frac{1 - x}{2} \right\} \\ &\geq x - \frac{x - 0}{2} = x - \frac{x}{2} = \frac{x}{2} > 0 \quad \text{since } x > 0. \end{aligned}$$

Thus $0 < y < 1$, that is, $y \in (0, 1)$ as claimed. Hence $B(x, \delta) \subset (0, 1)$, and we see that $(0, 1)$ is open.

(c) With a similar argumentation as in (b) we can show that the complement of $[0, 1]$, given by

$$\mathbb{R} \setminus [0, 1] = (-\infty, 0) \cup (1, \infty)$$

is open. Thus $[0, 1]$ is by definition closed. \square

Remark 6.66 (open and closed intervals in \mathbb{R})

In analogy to the previous example we can show the following for intervals in \mathbb{R} :

- (a) $[a, b] := \{x \in \mathbb{R} : a \leq x \leq b\}$ is closed.
- (b) $(a, b) := \{x \in \mathbb{R} : a < x < b\}$ is open.
- (c) $[a, b) := \{x \in \mathbb{R} : a \leq x < b\}$ and $(a, b] := \{x \in \mathbb{R} : a < x \leq b\}$ are neither open nor closed.
- (d) The space $\mathbb{R} = (-\infty, \infty)$ and the empty set \emptyset are both open and closed.

Next we show that open balls in \mathbb{R}^n equipped with the Euclidean norm are open subsets of \mathbb{R}^n .

Example 6.67 ($B(\mathbf{x}; r) \subset \mathbb{R}^n$ is open)

Consider \mathbb{R}^n with the Euclidean norm $\|\cdot\|_2$. For $\mathbf{x} \in \mathbb{R}^n$ and $r > 0$, show that the open ball $B(\mathbf{x}; r) \subset \mathbb{R}^n$ is an open set.

Proof: We need to show that every $\mathbf{y} \in B(\mathbf{x}; r)$ is an interior point of $B(\mathbf{x}; r)$, that is, for every $y \in B(\mathbf{x}; r)$ there exists some $\varepsilon > 0$, such that $B(\mathbf{y}; \varepsilon) \subset B(\mathbf{x}; r)$.

To prove this directly, we need some measurements. Since $\mathbf{y} \in B(\mathbf{x}; r)$, we have, by definition that $\|\mathbf{y} - \mathbf{x}\|_2 < r$. Let $\delta := \|\mathbf{y} - \mathbf{x}\|_2$, then $r - \delta > 0$, and we take $\varepsilon := (r - \delta)/2$. We want to show that $B(\mathbf{y}; \varepsilon) \subset B(\mathbf{x}; r)$, that is, if $\mathbf{z} \in B(\mathbf{y}; \varepsilon)$, then $\mathbf{z} \in B(\mathbf{x}; r)$. To see this, we use the triangle inequality as follows: For any $\mathbf{z} \in B(\mathbf{y}; \varepsilon)$,

$$\begin{aligned} \|\mathbf{z} - \mathbf{x}\|_2 &= \|(\mathbf{z} - \mathbf{y}) + (\mathbf{y} - \mathbf{x})\|_2 \\ &\leq \|\mathbf{z} - \mathbf{y}\|_2 + \|\mathbf{y} - \mathbf{x}\|_2 \\ &< \varepsilon + \delta = \frac{r - \delta}{2} + \delta = \frac{r + \delta}{2} < r, \end{aligned}$$

where we have used that $\delta < r$ in the last step. Since $\|\mathbf{z} - \mathbf{x}\|_2 < r$, we have $\mathbf{z} \in B(\mathbf{x}; r)$, and thus $B(\mathbf{y}; \varepsilon) \subset B(\mathbf{x}; r)$. \square

We discuss some more examples.

Example 6.68 (open sets in \mathbb{R} with the discrete metric)

In Example 6.4, we saw that \mathbb{R} with the **discrete metric**

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

is a metric space. Show that any subset E of \mathbb{R} is open.

Solution: Let E be an arbitrary subset $E \subset \mathbb{R}$, and consider an arbitrary $x \in E$. Then we have to show that x is an interior point, that is, there exists an $\varepsilon > 0$ such that $B(x; \varepsilon) \subset E$. For $\varepsilon = 1/2$, we have that (see Example 6.57)

$$B\left(x; \frac{1}{2}\right) = \left\{y \in \mathbb{R} : d(x, y) < \frac{1}{2}\right\} = \{x\} \subset E.$$

Thus x is an interior point, and the set E is open. \square

In the next example we want to determine the set of interior points of a given subset of a metric space and then use this information to determine whether the set is open or not.

Example 6.69 (interior points and open sets)

Find the set of interior points of each of the following subsets of the given normed linear space, and determine whether the subset is open or not.

$$(a) \quad B(\mathbf{0}; 1) := \left\{ \mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n : \|\mathbf{x}\|_\infty = \max_{1 \leq j \leq n} |x_j| < 1 \right\}$$

as a subset of \mathbb{R}^n with the norm $\|\mathbf{x}\|_\infty = \max_{1 \leq j \leq n} |x_j|$, $\mathbf{x} \in \mathbb{R}^n$.

(b) The set of constant functions

$$E := \{f \in \mathcal{C}(\mathbb{R}) : f(x) := C \text{ for all } x \in \mathbb{R} \text{ and for any } C \in \mathbb{R}\}$$

as a subset of $\mathcal{C}(\mathbb{R})$ with the supremum norm $\|f\|_\infty := \sup_{x \in \mathbb{R}} |f(x)|$, $f \in \mathcal{C}(\mathbb{R})$.

Solution:

(a) Claim: Every point x in $B(\mathbf{0}; 1)$ is an interior point, and thus $B(\mathbf{0}; 1)$ is open.

Proof: Consider $\mathbf{x} \in B(\mathbf{0}; 1)$. Then $\delta := \|\mathbf{x}\|_1$ satisfies $0 \leq \delta < 1$, and we will show that

$$B(\mathbf{x}; (1-\delta)/2) := \left\{ \mathbf{y} = (y_1, y_2, \dots, y_n) \in \mathbb{R}^n : \|\mathbf{y} - \mathbf{x}\|_\infty = \max_{1 \leq j \leq n} |y_j - x_j| < \frac{1-\delta}{2} \right\}$$

is a subset of $B(\mathbf{0}; 1)$. Consider an arbitrary point $\mathbf{y} \in B(\mathbf{x}; (1-\delta)/2)$; then from the triangle inequality and $\delta < 1$

$$\|\mathbf{y} - \mathbf{0}\|_1 = \|\mathbf{y}\|_1 \leq \|\mathbf{y} - \mathbf{x}\|_1 + \|\mathbf{x}\|_1 < \frac{1-\delta}{2} + \delta = \frac{1+\delta}{2} < 1.$$

Thus $\mathbf{y} \in B(\mathbf{0}; 1)$ and therefore $B(\mathbf{x}; (1-\delta)/2) \subset B(\mathbf{0}; 1)$, and we see that all $\mathbf{x} \in B(\mathbf{0}; 1)$ are interior points and that $B(\mathbf{0}; 1)$ is open. \square

(b) Claim no function f in E is an interior point, and thus E is not open.

Proof: Consider an arbitrary $f \in E$, that is, $f(x) := C$ for all $x \in \mathbb{R}$ with some constant $C \in \mathbb{R}$. For arbitrary $\varepsilon > 0$, the open ball

$$B(f; \varepsilon) := \left\{ g \in \mathcal{C}(\mathbb{R}) : \|g - f\|_\infty = \sup_{x \in \mathbb{R}} |g(x) - f(x)| < \varepsilon \right\}$$

contains the function $g : \mathbb{R} \rightarrow \mathbb{R}$, given by

$$g(x) := \frac{\varepsilon}{2} \sin x + C, \quad x \in \mathbb{R}.$$

Indeed, we have

$$\|g - f\|_\infty = \sup_{x \in \mathbb{R}} \left| \left(\frac{\varepsilon}{2} \sin x + C \right) - C \right| = \sup_{x \in \mathbb{R}} \left| \frac{\varepsilon}{2} \sin x \right| = \frac{\varepsilon}{2} \sin \left(\frac{\pi}{2} \right) = \frac{\varepsilon}{2} < \varepsilon,$$

and thus $g \in B(f; \varepsilon)$ but $g \notin E$. Thus the open ball $B(f; \varepsilon)$ is not contained in E , and since ε was arbitrary, we see that the constant function $f(x) = C$ is not an interior point. Since $f \in E$ was arbitrary, we see that E contains no interior points.

The set of interior points $\overset{\circ}{E}$ is empty, and the set E is not open. \square

6.3.2 Accumulation Points and Characterizations of Closed Sets

In order to find a simpler way for checking whether a set is closed, let us first introduce the notion of **accumulation points**. Then we will establish a simple tool, Lemma 6.77 below, for determining whether a set $E \subset \mathbb{R}^n$ is closed.

Definition 6.70 (accumulation point)

Let E be a non-empty subset of a metric space (X, d) or a non-empty subset of a normed linear space $(X, \|\cdot\|)$. A point $x \in X$ is said to be an **accumulation point** of E if for every $\varepsilon > 0$, there is $x' \neq x$ satisfying $x' \in B(x; \varepsilon) \cap E$. For a non-empty set $E \subset X$, we denote by E' the **set of all accumulation points** of E . Points in E who are not accumulation points are called **isolated points**.

Example 6.71 (normed linear space: interior pt. \Rightarrow accumulation pt.)

Let E be a non-empty subset of a normed linear space $(X, \|\cdot\|)$. Any interior point of E is an accumulation point. (Note this is **not** true in a metric space!)

Proof: We start by observing that, since X is a normed linear space, any open ball $B(\mathbf{y}; \gamma)$ contains elements other than \mathbf{y} itself. (This is not necessarily true in a metric space!) Now let x be an interior point of E . Then there exists $\delta > 0$ such that $B(x; \delta) \subset E$. Thus for any $\varepsilon > 0$ the set $B(x; \varepsilon) \cap B(x; \delta)$ is $B(x; \min\{\varepsilon; \delta\})$ and contains points other than x itself, and since $B(x; \min\{\varepsilon; \delta\}) \subset B(x; \delta) \subset E$, we have a point $\mathbf{y} \in B(x; \varepsilon) \cap E$. Thus \mathbf{x} is an accumulation point of E . \square

We find the accumulation points and isolated points for some subsets of the normed linear space \mathbb{R} with the absolute value norm $|\cdot|$.

Example 6.72 (accumulation points and isolated points)

Consider \mathbb{R} with the absolute value norm $|\cdot|$.

(a) The set of accumulation points of $(0, 1]$ is $[0, 1]$.

Proof: From Remark 6.71, we know that all $x \in (0, 1)$ are accumulation points since they are interior points. We claim that $x = 0$ and $x = 1$ are accumulation points. Indeed, let $\varepsilon > 0$ be arbitrary, and consider $B(0; \varepsilon)$ and $B(1; \varepsilon)$. Then $\varepsilon/2 \in B(0; \varepsilon) \cap (0, 1]$ and $1 - \varepsilon/2 \in B(1; \varepsilon) \cap (0, 1]$. Thus $x = 0$ and $x = 1$ are accumulation points.

For any point $x \in \mathbb{R} \setminus [0, 1]$, let $\delta := x - 1$ if $x > 1$ and $\delta := 0 - x = -x$ if $x < 0$. Then we can show that $B(x; \delta/2) \cap (0, 1] = \emptyset$. Thus $x \in \mathbb{R} \setminus [0, 1]$ is not an accumulation point.

We find that the set of all accumulation points is given by $E' = [0, 1]$. \square

(b) Similarly we can show that the set of all accumulation points of $[0, 1]$, of $[0, 1)$, and of $(0, 1)$ is also $[0, 1]$.

(c) Consider $E = [0, 1] \cup \{2\}$. Since $B(2; 1/2) \cap E = \{2\}$, we see that $x = 2$ is not an accumulation point. From (a), it follows that the set of accumulation points is $E' = [0, 1]$. The point $x = 2$ is an isolated point.

(d) The set $E := \{1/k : k \in \mathbb{N}\}$ has the set of accumulation points $E' = \{0\}$.

Proof: We have to show that for every $\varepsilon > 0$ the open ball $B(0; \varepsilon)$ contains an element from E . Let $K \in \mathbb{N}$ be such that

$$\frac{1}{K} < \varepsilon.$$

Then $1/K \in E \cap B(0; \varepsilon)$. Thus $x = 0$ is an accumulation point of E . Since for any point $1/k \in E$ we have $B(1/k; (2k(k+1))^{-1}) \cap E = \{1/k\}$, we see that all points in E are isolated points (since they are not accumulation points). Any point $x \notin E \cup \{0\}$ is not an accumulation point because $\delta := \inf\{|x - 1/k| : k \in \mathbb{N}\} > 0$ and thus $B(x; \delta/2) \cap E = \emptyset$. \square

Remark 6.73 (accumulation points may belong to E or not)

Note that an accumulation point may or may not belongs to $E \subset \mathbb{R}^n$. For example, let $E = (0, 1] \subset \mathbb{R}$. In Example 6.72 (a) we have seen that both $x = 0$ and $x = 1$ are accumulation points of $(0, 1]$. We observe that $1 \in E$ whereas $0 \notin E$. \square

We give another example.

Example 6.74 (subsets of \mathbb{R} with the discrete metric)

In Example 6.4, we saw that \mathbb{R} with the **discrete metric**

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

is a metric space. Any non-empty subset E of \mathbb{R} with the discrete metric contains only isolated points.

Proof: We have to show that no $x \in \mathbb{R}$ is an accumulation point of E . Consider an arbitrary $x \in \mathbb{R}$. From Example 6.57 we know that for any $\varepsilon \leq 1$, the open ball $B(x; \varepsilon)$ contains only x itself, that is,

$$B(x; \varepsilon) \cap \mathbb{R} = B(x; \varepsilon) = \{x\} \quad \text{for all } \varepsilon \leq 1.$$

Thus x is not an accumulation point of E . This implies that any $x \in E$ is not an accumulation point of E but an isolated point of E . Thus E contains only isolated points. \square

The next corollary is a consequence of the **Bolzano-Weistrass theorem** for \mathbb{R}^n with the Euclidean norm $\|\cdot\|_2$, and its proof shows that in $(\mathbb{R}^n, \|\cdot\|_2)$ there is a close link between accumulation points and convergent sequences.

Corollary 6.75 (infinite bounded set in \mathbb{R}^n has an accumulation point)

*Consider \mathbb{R}^n with the Euclidean norm $\|\mathbf{x}\| := \left(\sum_{j=1}^n x_j^2\right)^{1/2}$. Every non-empty **bounded infinite** set $E \subset \mathbb{R}^n$ has at least one accumulation point.*

Remark 6.76 (comment on Corollary 6.75)

Note that Corollary 6.75 does not hold for general metric spaces or general normed linear spaces. For example the closed bounded unit ball

$$\overline{B}(0; 1) := \{x \in \mathbb{R} : d(x, 0) \leq 1\}$$

in \mathbb{R} with the discrete metric

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

is bounded, and, in Examples 6.57 and 6.74, we have seen that $\overline{B}(0; 1) = \mathbb{R}$ and that \mathbb{R} contains only isolated points, thus $\overline{B}(0; 1) = \mathbb{R}$ can have no accumulation point.

Proof of Corollary 6.75: Since E is infinite and bounded, the set E contains a bounded infinite sequence $\{\mathbf{x}^{(k)}\} \subset E$ of distinct points $\mathbf{x}^{(k)}$. From the Bolzano-Weierstrass theorem (see Theorem 6.51), we know that $\{\mathbf{x}^{(k)}\}$ contains a convergent subsequence $\{\mathbf{x}^{(k_1)}\}$. Let $\mathbf{x} \in \mathbb{R}^n$ denote the limit of this subsequence. Then for every $\varepsilon > 0$ there exists $N = N(\varepsilon) \in \mathbb{N}$ such that

$$\|\mathbf{x}^{(k_1)} - \mathbf{x}\|_2 < \varepsilon \quad \text{for all } k_1 \geq N. \quad (6.26)$$

From (6.26) and because the elements $\mathbf{x}^{(k_1)}$ are distinct, we see that for every $\varepsilon > 0$, the open ball $B(\mathbf{x}; \varepsilon)$ contains an $\mathbf{x}^{(k_1)} \neq \mathbf{x}$. Thus \mathbf{x} is an accumulation point of the infinite bounded set E . \square

Now we learn two useful **criteria for checking whether a subset of a metric space or normed linear space is closed**.

Lemma 6.77 (closed = contains all accumulation points)

Consider a metric space (X, d) or a normed linear space $(X, \|\cdot\|)$. A set $E \subset X$ is closed if and only if E contains all its accumulation points, that is, $E' \subset E$.

Proof of Lemma 6.77: \Rightarrow : Let E be closed. Assume that E' is not contained in E , that is, there exists an accumulation point $x \in E' \setminus E$. Since E is closed, $X \setminus E$ is open. By the definition of an open set, there is some $\varepsilon > 0$, such that $B(x; \varepsilon) \subset X \setminus E$. On the other hand, since $x \in E'$, there is some $x' \neq x$ satisfying $x' \in B(x; \varepsilon) \cap E$. This contradicts the fact that $B(x; \varepsilon) \subset X \setminus E$. Thus we know that our assumption was wrong and that all accumulation points of E are in E .

\Leftarrow : To show that $E' \subset E$ implies that E is closed, we need to show that $X \setminus E$ is open. Given any point $x \in X \setminus E$, we have to show that x is an interior point of $X \setminus E$. Assume that $X \setminus E$ is not open. Then there exists some $x \in X \setminus E$ such that for every $\varepsilon > 0$, $B(x; \varepsilon)$ is not completely contained in $X \setminus E$. Then there is some $x' \in B(x; \varepsilon)$ that is not in $X \setminus E$, hence $x' \in E$. In other words, for every $\varepsilon > 0$ there exists $x' \in B(x; \varepsilon) \cap E$. Since $x \notin E$, we have that $x' \in B(x; \varepsilon) \cap E$ satisfies $x' \neq x$, and we see that x is an accumulation point of E . Since $x \notin E$, this contradicts the assumption that $E' \subset E$. Thus we know that our assumption, that $X \setminus E$ was not open, is wrong. Hence $X \setminus E$ is open and thus E is closed. \square

By adding all its accumulations points to a given set, we obtain the so-called closure of a set.

Definition 6.78 (Closure of E)

Let E be a non-empty subset of a metric space (X, d) or a non-empty subset of a normed linear space $(X, \|\cdot\|)$. The set $\overline{E} := E \cup E'$ is called the closure of E

We show that the closure is closed as the word seems to imply.

Lemma 6.79 (closure is closed)

Let E be a non-empty subset of a metric space (X, d) or a non-empty subset of a normed linear space $(X, \|\cdot\|)$. The closure \overline{E} of E is closed.

Proof of Lemma 6.79: The set $\overline{E} = E \cup E'$ contains E and all accumulation points of E . If we can show that this implies that \overline{E} contains also all its accumulation points, then we know that \overline{E} is closed (from Lemma 6.77). We will show that every accumulation point of \overline{E} is also an accumulation point of E and thus is contained in \overline{E} , since $\overline{E} = E \cup E'$.

Let x be an accumulation point of \overline{E} . Then for every $\varepsilon > 0$ there exists some $y \in \overline{E}$ with $y \neq x$ and $y \in B(x; \varepsilon/2)$. Either $y \in E$ or $y \in E'$. If $y \in E'$ but $y \notin E$, then there exists $z \in B(y; \varepsilon/2)$ with $z \in E$, since $y \in E'$ is an accumulation point of E .

From the triangle inequality

$$d(x, z) \leq d(x, y) + d(y, z) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

and thus $z \in B(x, \varepsilon)$. Thus for every $\varepsilon > 0$, there exists a point in $w \in B(x; \varepsilon) \cap E$ with $w \neq x$, and therefore x is also an accumulation point of E . Thus $x \in E'$ and consequently $x \in \overline{E} = E \cup E'$. Thus \overline{E} contains all its accumulation points and is closed, from Lemma 6.77. \square

The second criterion for a subset of a metric space being closed exploits the following fact: **Any accumulation point of a subset E of a metric space is the limit of a convergent sequence $\{x_k\}$ whose members are all distinct (that is, $x_k \neq x_m$ if $k \neq m$). The limit of any convergent sequence $\{x_k\}$ in a normed linear space whose members are all distinct is an accumulation point of the set $E := \{x_k : k \in \mathbb{N}\}$.** These two statements will be derived, explained, and used in the proof of Lemma 6.80 below.

Lemma 6.80 (criterion for being closed)

Consider a metric space (X, d) or a normed linear space $(X, \|\cdot\|)$. A set $E \subset X$ is closed if and only if for every sequence $\{x_k\} \subset E$ that is convergent in X to some $x \in X$, the limit x lies in E .

Proof of Lemma 6.80: \Leftarrow : Assume that for every sequence $\{x_k\}$ that is convergent to some $x \in X$ the limit x lies in E . We show that this condition implies that $E' \subset E$. Let $x \in E'$ be an arbitrary accumulation point. Then for every $k \in \mathbb{N}$, there is some $x_k \in B(x; 1/k) \cap E$ and $x_k \neq x$. Thus $d(x_k, x) < 1/k$ and $\{x_k\} \subset E$. Thus for every $\varepsilon > 0$, for $N := \min\{k \in \mathbb{N} : 1/k < \varepsilon\}$ we have $d(x_k, x) < \varepsilon$ for all $k \geq N$, and we see that $\{x_k\}$ converges in X to x . From the assumption, this implies $x \in E$. Thus all accumulation points lie in E and from Lemma 6.77 the set E is closed.

\Rightarrow : Suppose E is closed. We have to show that for every convergent sequence $\{x_k\}$ the limit lies in E . From Lemma 6.77, we know that all accumulation points lie in E . Now consider a sequence $\{x_k\} \subset E$ that converges in X to $x \in X$. We have to show two cases: (1) either $\{x_k\}$ contains infinitely many different points, or (2) $\{x_k\}$ contains only finitely many distinct points.

Let us first consider second case: If $\{x_k\}$ is convergent and contains only finitely many distinct points, then we have $x_k = x_N$ for all $k \geq N$ with some $N \in \mathbb{N}$. Since $x_N \in E$, we have that the limit $x = x_N$ is in E .

Now consider the first case: Let $\{x_k\}$ contain infinitely many distinct points. Because $\{x_k\}$ converges to x , for every $\varepsilon > 0$, there exists some N such that $d(x_k, x) < \varepsilon$ for all $k \geq N$. Thus there exists some $x_k \in B(x; \varepsilon) \cap E$ with $x_k \neq x$ and with $k \geq N$. This means that the limit x is an accumulation point of E and hence lies in E since E is closed. \square

Lemma 6.80 is a very convenient tool for studying open and closed sets.

Example 6.81 (\mathbb{R}^n and \emptyset are each open and closed)

In \mathbb{R}^n with the Euclidean norm $\|\mathbf{x}\|_2 := \left(\sum_{j=1}^n x_j^2\right)^{1/2}$ the set \mathbb{R}^n and the empty set \emptyset are each closed and open.

Proof: First we want to show that \mathbb{R}^n is open and \emptyset is closed. Consider an arbitrary point $\mathbf{x} \in \mathbb{R}^n$. Then $B(\mathbf{x}; \varepsilon) \subset \mathbb{R}^n$ for every $\varepsilon > 0$, and thus \mathbf{x} is an interior point. Since $\mathbf{x} \in \mathbb{R}^n$ was arbitrary, we know that all points in \mathbb{R}^n are interior points. Consequently, \mathbb{R}^n is open, and its complement $\emptyset = \mathbb{R}^n \setminus \mathbb{R}^n$ is closed.

Now we show that \mathbb{R}^n is closed and that \emptyset is open. Consider an arbitrary sequence in $\{\mathbf{x}^{(k)}\} \subset \mathbb{R}^n$ that converges in \mathbb{R}^n to some $\mathbf{x} \in \mathbb{R}^n$. Then the limit is in the subset \mathbb{R}^n , and we know from Lemma 6.80 that the subset \mathbb{R}^n is closed in \mathbb{R}^n . Thus \mathbb{R}^n is closed and its complement $\emptyset = \mathbb{R}^n \setminus \mathbb{R}^n$ is open. \square

Example 6.82 (closed balls in \mathbb{R}^n are closed)

Show that the closed ball $\overline{B}(\mathbf{y}; r) \subset \mathbb{R}^n$ is closed.

Proof: We want to apply Lemma 6.80. Let $\{\mathbf{x}^{(k)}\}$ be an arbitrary sequence in $\overline{B}(\mathbf{y}; r)$ with $\lim_{k \rightarrow \infty} \mathbf{x}^{(k)} = \mathbf{x}$. Then $\|\mathbf{x}^{(k)} - \mathbf{y}\|_2 \leq r$ for all $k \in \mathbb{N}$, and

$$\|\mathbf{y} - \mathbf{x}\|_2 = \|(\mathbf{y} - \mathbf{x}^{(k)}) + (\mathbf{x}^{(k)} - \mathbf{x})\|_2 \leq \|\mathbf{y} - \mathbf{x}^{(k)}\|_2 + \|\mathbf{x}^{(k)} - \mathbf{x}\|_2 \leq r + \|\mathbf{x} - \mathbf{x}^{(k)}\|_2.$$

For $k \rightarrow \infty$ we see that $\|\mathbf{y} - \mathbf{x}\|_2 \leq r$ because $\lim_{k \rightarrow \infty} \mathbf{x}^{(k)} = \mathbf{x}$ and consequently $\lim_{k \rightarrow \infty} \|\mathbf{x}^{(k)} - \mathbf{x}\|_2 = 0$. Thus $\mathbf{x} \in \overline{B}(\mathbf{y}; r)$. From Lemma 6.80 we conclude that $\overline{B}(\mathbf{x}, r)$ is closed. \square

Example 6.83 (closed set in \mathbb{R} with the absolute value norm)

Show that the set $E \subset \mathbb{R}$ (where \mathbb{R} is equipped with the absolute value norm $|\cdot|$), given by

$$E := \left\{ \frac{1}{k} : k \in \mathbb{N} \right\} \cup \{0\}$$

is closed.

Proof: We have seen in Example 6.72 that all points $1/k$, $k \in \mathbb{N}$ are isolated points, and that 0 is the only accumulation point. Since $0 \in E$, the set E contains all its accumulation points, and we know from Lemma 6.77 that E is closed in $(\mathbb{R}, |\cdot|)$. \square

The following example is a subset of \mathbb{R}^2 with the Euclidean norm which is neither open nor closed.

Example 6.84 (neither open nor closed set)

Let \mathbb{R}^2 be equipped with the Euclidean norm $\|\mathbf{x}\|_2 := \left(\sum_{j=1}^2 x_j^2\right)^{1/2} = \sqrt{x_1^2 + x_2^2}$. Show that the set

$$V := \{(x, y) \in \mathbb{R}^2 : x > 0, y \geq 0\}$$

is neither open or closed.

Proof: First we show that V is not closed. Consider the sequence $\{(1/k, 0)\} \subset V$. Then the limit in \mathbb{R}^2 of $\{(1/k, 0)\}$ is $\lim_{k \rightarrow \infty} (1/k, 0) = (0, 0)$, which is not in V . Thus according to Lemma 6.80, V is not closed.

Now we show that V is not open by showing that $\mathbb{R}^2 \setminus V$ is not closed. Let us consider the sequence $\{(1, -1/k)\} \in \mathbb{R}^2 \setminus V$. Clearly $\lim_{k \rightarrow \infty} (1, -1/k) = (1, 0)$, but, since $(1, 0) \in V$, we see that $(1, 0) \notin \mathbb{R}^2 \setminus V$. By Lemma 6.80, $\mathbb{R}^2 \setminus V$ is not closed. Thus V is not open. \square

We discuss some more complicated examples.

Example 6.85 (closed unit ball in $(\mathcal{C}([a, b]), \|\cdot\|_\infty)$)

The closed unit ball

$$\overline{B}(0; 1) := \{f \in \mathcal{C}([a, b]) : \|f\|_\infty \leq 1\}$$

in $\mathcal{C}([a, b])$ with the supremum norm

$$\|f\|_\infty = \sup_{x \in [a, b]} |f(x)|, \quad f \in \mathcal{C}([a, b])$$

is closed.

Proof: Consider an arbitrary sequence of functions $\{f_n\} \subset \overline{B}(0; 1)$ that converges uniformly to some $f \in \mathcal{C}([a, b])$. Then

$$\lim_{n \rightarrow \infty} \|f_n - f\|_\infty = 0$$

and $\|f_n\|_\infty \leq 1$ imply that

$$\|f\|_\infty = \|(f - f_n) + f_n\|_\infty \leq \|f - f_n\|_\infty + \|f_n\|_\infty \leq \|f - f_n\|_\infty + 1, \quad (6.27)$$

and thus, from $\lim_{n \rightarrow \infty} \|f - f_n\|_\infty = 0$, letting $n \rightarrow \infty$ in (6.27) yields

$$\|f\|_\infty \leq 1.$$

Since $\|f\|_\infty \leq 1$, we have $f \in \overline{B}(0; 1)$, and thus $\overline{B}(0; 1)$ is closed. \square

Example 6.86 (closed sets in \mathbb{R} with the discrete metric)

In Example 6.4, we saw that in \mathbb{R} with the **discrete metric**

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

a closed ball $\overline{B}(y; r)$ is $\{y\}$ if $r < 1$ and \mathbb{R} if $r \geq 1$. We want to show that any closed ball $\overline{B}(y; r)$ is indeed closed.

If $r \geq 1$ then $\overline{B}(y; r) = \mathbb{R}$, and every convergent sequence $\{x_k\}$ in $\overline{B}(y; r) = \mathbb{R}$ has its limit in \mathbb{R} . If $r < 1$ then $\overline{B}(y; r) = \{y\}$, every sequence in $\overline{B}(y; r) = \{y\}$ is constant, that is, $\{x_k\}$ satisfies $x_k = y$ for all $k \in \mathbb{N}$. Thus $\{x_k\}$ converges to $y \in \overline{B}(y; r)$. Thus we see that $\overline{B}(y; r)$ is closed. \square

Example 6.87 (closed set in $(\mathcal{C}(\mathbb{R}), \|\cdot\|_\infty)$)

The subset

$$E := \{f \in \mathcal{C}(\mathbb{R}) : f(x) := C \text{ for all } x \in \mathbb{R} \text{ and any } C \in \mathbb{R}\}$$

of $\mathcal{C}(\mathbb{R})$, endowed with the supremum norm $\|f\|_\infty := \sup_{x \in \mathbb{R}} |f(x)|$, is closed.

Proof: We use Lemma 6.80 to prove that E is closed. Consider any uniformly convergent sequence $\{f_n\} \subset E$. From the definition of E , we know that every function $f_n : \mathbb{R} \rightarrow \mathbb{R}$ is a constant function, that is,

$$f_n(x) := C_n \quad \text{with some constant } C_n \in \mathbb{R}.$$

Since the sequence $\{f_n\}$ converges uniformly to some function $f \in \mathcal{C}(\mathbb{R})$, we have in particular also that $\{f_n\}$ converges pointwise to f , that is

$$\lim_{n \rightarrow \infty} f_n(x) = \lim_{n \rightarrow \infty} C_n = f(x) \quad \text{for all } x \in \mathbb{R}.$$

Since for every $x \in \mathbb{R}$, the value $f(x)$ is the limit of the sequence of real numbers $\{C_n\}$, we see that

$$f(x) = \lim_{n \rightarrow \infty} C_n =: C \quad \text{for all } x \in \mathbb{R},$$

that is, $f : \mathbb{R} \rightarrow \mathbb{R}$ is constant and thus belongs to E . Since the limit of every uniformly convergent sequence $\{f_n\} \subset E$ lies in E , we know from Lemma 6.80 that E is closed in $(\mathcal{C}(\mathbb{R}), \|\cdot\|_\infty)$. \square

Appendix A

Appendix: Handout ‘Derivatives and Integrals’

You are expected to **know** the following derivatives and integrals, and the listed rules for dealing with derivatives and integrals.

1. Rules for dealing with derivatives

- **product rule:** Let f, g be differentiable on (a, b) . Then

$$\frac{d}{dx}(f(x)g(x)) = f'(x)g(x) + f(x)g'(x), \quad x \in (a, b).$$

- **quotient rule:** Let f, g be differentiable on (a, b) , and let $g(x) \neq 0$ for all $x \in (a, b)$. Then

$$\frac{d}{dx}\left(\frac{f(x)}{g(x)}\right) = \frac{f'(x)g(x) - f(x)g'(x)}{(g(x))^2}, \quad x \in (a, b).$$

- **chain rule:** Let $\varphi : (a, b) \rightarrow (c, d)$ be differentiable, and let $f : (c, d) \rightarrow \mathbb{R}$ be differentiable. Then

$$\frac{d}{dx}(f \circ \varphi)(x) = \frac{d}{dx}f(\varphi(x)) = f'(\varphi(x))\varphi'(x), \quad x \in (a, b).$$

2. Important derivatives

1. $\frac{d}{dx}(C) = 0$ for any constant C
2. $\frac{d}{dx}(x^n) = n x^{n-1}$
3. $\frac{d}{dx}(e^x) = e^x$
4. $\frac{d}{dx}(\ln x) = \frac{1}{x}$ for $x > 0$
5. $\frac{d}{dx}(\sin x) = \cos x$
6. $\frac{d}{dx}(\cos x) = -\sin x$
7. $\frac{d}{dx}(\sinh x) = \cosh x$
8. $\frac{d}{dx}(\cosh x) = \sinh x$
9. $\frac{d}{dx}(\tan x) = \frac{1}{(\cos x)^2}$ for $x \in (k\pi - \frac{\pi}{2}, k\pi + \frac{\pi}{2})$ with $k \in \mathbb{Z}$
10. $\frac{d}{dx}(\cot x) = -\frac{1}{(\sin x)^2}$ for $x \in (k\pi, (k+1)\pi)$ with $k \in \mathbb{Z}$

Note that you can easily work out the last two derivatives with the quotient rule.

3. Rules and techniques for manipulating integrals

- **fundamental theorem of calculus:** Let f be continuous on $\langle c, d \rangle$, and let F be a primitive of f . Then for all $a, b \in \langle c, d \rangle$

$$\int_a^b f(x) dx = F(b) - F(a).$$

- **integration by parts:** Integration by parts is based on the product rule.

Let F and G be continuously differentiable on $\langle c, d \rangle$ and let $a, b \in \langle c, d \rangle$. Then

$$\int_a^b F'(x) G(x) dx = F(x) G(x) \Big|_a^b - \int_a^b F(x) G'(x) dx.$$

- **substitution:** Substitution is based on the chain rule.

Let $f : (c', d') \rightarrow \mathbb{R}$ be continuous on (c', d') , and let $F : (c', d') \rightarrow \mathbb{R}$ be a primitive of f , that is, $F'(x) = f(x)$ for all $x \in (c', d')$. Let $\varphi : (c, d) \rightarrow (c', d')$ be continuously differentiable. Then for any $a, b \in (c, d)$

$$\int_a^b f(\varphi(x)) \varphi'(x) dx = \int_{\varphi(a)}^{\varphi(b)} f(u) du = F(u) \Big|_{\varphi(a)}^{\varphi(b)} = F(\varphi(b)) - F(\varphi(a)). \quad (*)$$

Formally we work this out by setting $u = \varphi(x)$. Then $du/dx = \varphi'(x)$, or equivalently $du = \varphi'(x) dx$, and replacing in the expression on the left of $(*)$ and using the assumptions yields $(*)$.

- **integrals of products and powers of trigonometric functions:** Here the following can for example be helpful:

- trigonometric identities
- (repeated) integration by parts
- substitution
- replace $(\sin x)^2 = 1 - (\cos x)^2$ or replace $(\cos x)^2 = 1 - (\sin x)^2$

- **trigonometric substitutions:**

- if the integrand involves $\sqrt{\alpha^2 - x^2}$, set $x = \alpha \sin \theta$
- if the integrand involves $\sqrt{\alpha^2 + x^2}$, set $x = \alpha \tan \theta$
- if the integrand involves $\sqrt{x^2 - \alpha^2}$, set $x = \alpha (\cos \theta)^{-1}$

- **partial fractions:** Partial fractions separation is used to evaluate integrals of the form

$$\int \frac{P(x)}{Q(x)} dx$$

where both P and Q are polynomials (that is, the integrand is a rational function) and where the polynomial degree of the numerator P is *strictly less* than the polynomial degree of the denominator Q .

If the polynomial degree of the numerator P is *not* strictly less than the polynomial degree of the denominator Q , then perform *long division*, in order to decompose into the sum of a polynomial and a rational function of the form $\tilde{P}(x)/\tilde{Q}(x)$, with \tilde{P} and \tilde{Q} polynomials and $\text{degree}(\tilde{P}) < \text{degree}(\tilde{Q})$. After this perform partial fraction separation for $\tilde{P}(x)/\tilde{Q}(x)$.

We can have the following three cases (as well as combinations of them):

1. *non-repeated linear factors*: The denominator Q is of the form

$$Q(x) = (x - r_1)(x - r_2) \cdots (x - r_n)$$

with all r_i distinct (that is, $r_i \neq r_j$ if $i \neq j$). Then there exist constants A_1, \dots, A_n such that

$$\frac{P(x)}{Q(x)} = \frac{A_1}{x - r_1} + \frac{A_2}{x - r_2} + \cdots + \frac{A_n}{x - r_n}.$$

2. *repeated linear factors*: The denominator Q is of the form

$$Q(x) = (x - r)^m \cdot \text{other factors},$$

where $m \geq 2$ is the largest possible power of $(x - r)$ in this factorization (This means the other factors do not contain any powers of $(x - r)$.) Then we obtain from the portion of the fraction corresponding to the factor $(x - r)^m$ a contribution

$$\frac{A_1}{x - r} + \frac{A_2}{(x - r)^2} + \cdots + \frac{A_m}{(x - r)^m}$$

to the partial fraction. Again the A_1, A_2, \dots, A_m are constants.

3. *quadratic factors*: The denominator Q is of the form

$$Q(x) = ((x - \alpha)^2 + \beta^2)^m \cdot \text{other factors},$$

where $m \geq 1$ is the largest possible power of $((x - \alpha)^2 + \beta^2)$ in this factorization. (That means the other factors contain no powers of $((x - \alpha)^2 + \beta^2)$.) Then we obtain from the portion of the fraction corresponding to the factor $((x - \alpha)^2 + \beta^2)^m$ a contribution

$$\frac{C_1 x + D_1}{(x - \alpha)^2 + \beta^2} + \frac{C_2 x + D_2}{((x - \alpha)^2 + \beta^2)^2} + \cdots + \frac{C_m x + D_m}{((x - \alpha)^2 + \beta^2)^m}$$

to the partial fraction. Here $C_1, D_1, C_2, D_2, \dots, C_m, D_m$ are constants.

4. Important integrals

$$1. \int_a^b C \, dx = C x|_a^b = C(b-a) \quad \text{for any constant } C \in \mathbb{R}$$

$$2. \int_a^b x^n \, dx = \frac{x^{n+1}}{n+1} \Big|_a^b = \frac{b^{n+1}}{n+1} - \frac{a^{n+1}}{n+1}$$

$$3. \int_a^b e^x \, dx = e^x|_a^b = e^b - e^a$$

$$4. \int_a^b \frac{1}{x} \, dx = (\ln|x|)|_a^b = \ln|b| - \ln|a| \quad \text{with either } a < b < 0 \text{ or } 0 < a < b$$

$$5. \int_a^b \sin x \, dx = -\cos x|_a^b = \cos a - \cos b$$

$$6. \int_a^b \cos x \, dx = \sin x|_a^b = \sin b - \sin a$$

$$7. \int_a^b \sinh x \, dx = \cosh x|_a^b = \cosh b - \cosh a$$

$$8. \int_a^b \cosh x \, dx = \sinh x|_a^b = \sinh b - \sinh a$$

$$9. \int_a^b \frac{1}{(\cos x)^2} \, dx = \tan x|_a^b = \tan b - \tan a$$

with $a < b$ and $a, b \in (k\pi - \frac{\pi}{2}, k\pi + \frac{\pi}{2})$, where $k \in \mathbb{Z}$

$$10. \int_a^b \frac{1}{(\sin x)^2} \, dx = -\cot x|_a^b = \cot a - \cot b$$

with $a < b$ and $a, b \in (k\pi, (k+1)\pi)$, where $k \in \mathbb{Z}$

$$11. \int_a^b \tan x \, dx = -(\ln|\cos x|)|_a^b = \ln|\cos a| - \ln|\cos b|$$

with $a < b$ and $a, b \in (k\pi - \frac{\pi}{2}, k\pi + \frac{\pi}{2})$, where $k \in \mathbb{Z}$

$$12. \int_a^b \cot x \, dx = (\ln|\sin x|)|_a^b = \ln|\sin b| - \ln|\sin a|$$

with $a < b$ and $a, b \in (k\pi, (k+1)\pi)$, where $k \in \mathbb{Z}$

Note that the last two integrals can be easily worked out with substitution.