Introduction to Analysis

Rafael Andrist

American University of Beirut, Fall 2019/20
9 August 2019

Contents

1 Notation and Elementary Notions **3**

2 Construction of Numbers **4**
2.1 Natural numbers 4
2.2 Integers 6
2.3 Rational numbers 8
2.4 Real Numbers 9
2.5 Complex Numbers 14

3 Sequences and Series **15**
3.1 Important Inequalities 15
3.2 Real sequences 16
3.3 Series 18
3.4 Tests for Convergence 21
3.5 Rearranging 24
3.6 Power Series 24

4 Continuity **26**
4.1 Open and Closed sets 26
4.2 Continuous Functions 26
4.3 Compactness 28
4.4 Connectedness 29
4.5 Uniform continuity 30

5 Differentiation and Integration **31**
5.1 Differentiation 31
5.2 Mean Value Theorem 40
5.3 Convexity 44
5.4 Taylor Series 45
5.5 Uniform convergence 49
5.6 Riemann Integration 54

1 Notation and Elementary Notions

$\exists$	there exists
$\forall$	for all
$\in$	in
$\notin$	not in
$\emptyset$	empty set
$\Longrightarrow$	implies
$\wedge$	logical and
$\vee$	logical or

A map f from a set X to a set Y assigns to every element $x \in X$ an element $f(x) \in Y$. We denote this by $f\colon X \to Y$, and call X the *domain* of f and Y the *co-domain* of f. The set of values assumed by f is denoted by $f(X) = \{f(x) : x \in X\}$ and is called its *range*.

Definition 1.0.1. A map $f\colon X \to Y$ is called

1. *injective* if
$$\forall a, b \in X \,:\, f(b) = f(a) \Longrightarrow a = b$$
2. *surjective* if
$$\forall y \in Y \; \exists x \in X \,:\, f(x) = y$$
3. *bijective* if it is injective and surjective.

Note that the notions of injectivity and surjectivity crucially depend on the choice of domain and co-domain.

2 Construction of Numbers

2.1 Natural numbers

Definition 2.1.1 (Peano axioms[1])**.** The set of *natural numbers*, denoted by $\mathbb{N}$, fulfills the following axioms where we denote by $S(n)$ the *successor* of n.

1. $0 \in \mathbb{N}$
2. $\forall n \in \mathbb{N}$ we have $S(n) \in \mathbb{N}$
3. $\forall n \in \mathbb{N}$ we have $S(n) \neq 0$
4. $\forall n, m \in \mathbb{N}$ we have that $S(m) = S(n) \Rightarrow m = n$
5. axiom of induction: $\forall X \subseteq \mathbb{N}$ with the property that $0 \in X \wedge (\forall n \in \mathbb{N}$ we have $(n \in X \Rightarrow S(n) \in X))$ we obtain $\mathbb{N} = X$

We denote the successor of 0 by 1, and in the decimal number system we continue by denoting the successor of 1 by 2, etc. and obtain the natural numbers

$$\mathbb{N} = \{0, 1, 2, 3, 4, 5, 6, 7, \dots\}$$

The last of the Peano axioms allows us to prove statements about the natural numbers using induction which is often stated as the following theorem.

Theorem 2.1.2 (Induction principle)**.** *A property concerning the natural numbers $\mathbb{N}$ that is true for $n = 0$, and that is true for all natural numbers up to $S(n)$ whenever it is true for all natural numbers up to n, must be true for all natural numbers.*

We would like to define *addition* of natural numbers. The *sum* of two natural numbers $m, n \in \mathbb{N}$ we denote by $m + n$ and define it as follows:

A1. $n + 0 = n$

A2. $n + S(m) = S(n + m)$

In particular we can now interpret $n + 1 = S(n)$.
We would like to define *multiplication* of natural numbers. The *product* of two natural numbers $m, n \in \mathbb{N}$ we denote by $m \cdot n$ and define it as follows:

M1. $n \cdot 0 = 0$

M2. $n \cdot S(m) = n + (n \cdot m)$

Exercise 2.1.3. Use the Peano axioms and the axioms for addition and multiplication of natural numbers to show the following properties.

[1] Giuseppe Peano, 1889

1. Show by induction in n that addition of natural numbers is associative.

$$\forall k, m, n \in \mathbb{N} \ : \ k + (m + n) = (k + m) + n)$$

2. Show by induction in n that addition of natural numbers is commutative.

$$\forall m, n \in \mathbb{N} \ : \ m + n = n + m$$

3. Show by induction in n the following distributivity law.

$$\forall k, m, n \in \mathbb{N} \ : \ k \cdot (m + n) = k \cdot m + k \cdot n$$

4. Show by induction in n that multiplication of natural numbers is commutative.

$$\forall m, n \in \mathbb{N} \ : \ m \cdot n = n \cdot m$$

5. Show by induction in n that multiplication of natural numbers is associative.

$$\forall k, m, n \in \mathbb{N} \ : \ k \cdot (m \cdot n) = (k \cdot m) \cdot n$$

Proof that addition is associative. We prove it by induction over n. For $n = 0$ we have for all $k, m \in \mathbb{N}$ that

$$k + (m + 0) = k + m = (k + m) + 0$$

Here is the induction step:

$$\begin{aligned} k + (m + (n + 1)) = k + (m + S(n)) = k + S(m + n) = S(k + (m + n)) \\ = S((k + m) + n) = (k + m) + S(n) = (k + m) + (n + 1) \end{aligned}$$

where we used the induction hypothesis for $k + (m + n) = (k + m) + n$. □

Definition 2.1.4. A *relation* $\sim$ between two sets X and Y is a subset R of the direct product $X \times Y = \{(x, y) \ : \ x \in X, y \in Y\}$. If we refer to only one set X, then we consider a subset of the direct product $X \times X$. We write $x \sim y$ if $(x, y) \in R$.

Definition 2.1.5. Let X be a set. An *order* is a relation $a \leq b$ between elements $a, b \in X$ with the following properties

1. reflexivity: $a \leq a$
2. anti-symmetry: $a \leq b \wedge b \leq a \Rightarrow a = b$
3. transitivity: $a \leq b \wedge b \leq c \Rightarrow a \leq c$

Remark 2.1.6. Note that this definition does not require that for all pairs $a, b \in X$ there is an order between between. If for all $a, b \in X$ we have $a \leq b$ or $b \leq a$ then we call the order a *total order*.
We may write $a < b$ if $a \leq b \wedge a \neq b$.

We define an order on the natural numbers by

$$a \leq b \iff \exists c : a + c = b$$

Lemma 2.1.7. *For all $a, b, c \in \mathbb{N}$ with $a \leq b$ we have:*

$$a + c \leq b + c$$
$$a \cdot c \leq b \cdot c$$

Moreover, the order on $\mathbb{N}$ is total.

Theorem 2.1.8 (Well-ordering principle)**.** *Every non-empty subset of $\mathbb{N}$ has a smallest element.*

Proof. We have that $\forall n \in \mathbb{N} : 0 \leq n$. Let $S \subseteq \mathbb{N}$. If $0 \in S$, then we are done. Assume now $0 \notin S$. Let $T := \mathbb{N} \setminus S$. We have $0 \in T$. Assume to get a contradiction that S has no smallest element. Consider any $n \in \mathbb{N}$ such that $1, 2, \ldots, n \in T$. If $n + 1 \notin T$, then $n + 1 \in S$ is the smallest element of S, hence $n + 1 \in T$. By induction, $T = \mathbb{N}$, a contradiction. □

2.2 Integers

Definition 2.2.1. Let X be a set. An *equivalence relation* is a relation $a \sim b$ between elements $a, b \in X$ with the following properties

1. reflexivity: $a \sim a$
2. symmetry: $a \sim b \Rightarrow b \sim a$
3. transitivity: $a \sim b \wedge b \sim c \Rightarrow a \sim c$

An *equivalence class* is the set that consists of all equivalent elements. To denote the equivalence class of a we often write $[a]$. The set of all equivalence classes is denoted by $X/\sim$.

Although we can add natural numbers, we can't solve the equation $x + n = 0$ within the natural numbers for any given $n \in \mathbb{N}, n \geq 1$. For this purpose we construct the *integers*, denoted by $\mathbb{Z}$, from the natural numbers. For each $n \in \mathbb{N}$, $n \neq 0$, we would like to introduce a negative number denoted by $-n$.
The more formal construction of the integers is done as follows:
Consider pairs of natural numbers $(a, b) \in \mathbb{N} \times \mathbb{N}$. We think of them as a being the positive and b being the negative part, representing "$a - b$". We introduce the following equivalence relation on $\mathbb{N} \times \mathbb{N}$:

$$(a, b) \sim (a', b') \iff a + b' = a' + b$$

Then we define the set of integers to be $\mathbb{Z} = (\mathbb{N} \times \mathbb{N})/\sim$. We define addition and multiplication by

$$[(a,b)] + [(c,d)] = [a+c, b+d]$$
$$[(a,b)] \cdot [(c,d)] = [ac+bd, bc+ad]$$

These operations are indeed well-defined, i.e. they do not depend on the choice of representatives. The natural numbers can be considered as a subset of the integers by $[(n,0)]$. We will write in short n for $[(n,0)]$ and $-n$ for $[(0,n)]$.
We can extend the order of the natural numbers to the integers by

$$a \leq b \iff \exists c \in \mathbb{N} \,:\, a + c = b$$

The integers satisfy the axioms of a *commutative ring*.

Definition 2.2.2. A *ring* is a set R together with an *addition* $+\colon R \times R \to R$, a multiplication $\cdot\colon R \times R \to R$ and an *additive neutral element* $0 \in R$ and a *multiplicative neutral element* $1 \in R$ with the following properties:

1. associativity of addition: $\forall a, b, c \,:\, (a+b)+c = a+(b+c)$
2. neutral element of addition: $\forall a \,:\, 0 + a = a = a + 0$
3. inverse element of addition: $\forall a \exists (-a) \,:\, a + (-a) = 0 = (-a) + a$
4. commutativity of addition: $\forall a, b \,:\, a + b = b + a$
5. associativity of multiplication: $\forall a, b, c \,:\, (a \cdot b) \cdot c = a \cdot (b \cdot c)$
6. neutral element of multiplication: $\forall a \,:\, 1 \cdot a = a = a \cdot 1$
7. distributivity: $\forall a, b, c$:

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

If we have commutativity of multiplication, i.e. $\forall a, b \,:\, a \cdot b = b \cdot a$, then we call the ring *commutative*.

Definition 2.2.3. The *absolute value* $|x|$ of an integer x is defined to be

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

Lemma 2.2.4. *For each integer $x \in \mathbb{Z}$ we have that $\pm x \leq |x|$.*

Theorem 2.2.5. *The absolute value satisfies $\forall x, y \in \mathbb{Z}$:*

1. $|x \cdot y| = |x| \cdot |y|$

2. triangle inequality: $|x+y| \leq |x| + |y|$

Proof of 2.

$$\begin{aligned} x + y \leq |x| + y &\leq |x| + |y| \\ -(x+y) = (-x) + (-y) \leq |x| + (-y) &\leq |x| + |y| \end{aligned}$$

where we used that $-a = (-1) \cdot a$ for $a \in \mathbb{Z}$ since $\mathbb{Z}$ is a ring. □

2.3 Rational numbers

We introduce an equivalence relation on $\mathbb{Z} \times \mathbb{Z}$ by declaring

$$(p, q) \sim (p', q') \iff p \cdot q' = q \cdot p'$$

We denote set of all the equivalence classes (p, q) with $q \neq 0$ by $\mathbb{Q}$ and call them *rational numbers*. We usually denote the equivalence class of (p, q) by $\frac{p}{q}$.
We define addition and multiplication the following way.

$$\begin{aligned} \frac{a}{b} + \frac{c}{d} &= \frac{a \cdot d + b \cdot c}{b \cdot d} \\ \frac{a}{b} \cdot \frac{c}{d} &= \frac{a \cdot c}{b \cdot d} \end{aligned}$$

The rational numbers satisfy the axioms of a *field*.

Definition 2.3.1. A *field* $(F, 0, 1, +, \cdot)$ is a ring where every element different from 0 has a multiplicative inverse:

$$\forall a \in F \setminus \{0\} \exists a^{-1} \in F \ : \ a \cdot a^{-1} = 1 = a^{-1} \cdot a$$

We can define an order for rational numbers the following way:

$$\frac{a}{b} \leq \frac{c}{d} \iff a \cdot d \leq b \cdot c$$

where we assume that $b > 0$ and $d > 0$.

Definition 2.3.2. A field F with a total order $\leq$ is called an *ordered field* if $\forall a, b, c \in F$

1. $a < b \implies a + c < b + c$
2. $0 < a, 0 < b \implies 0 < a \cdot b$

Lemma 2.3.3. *The rational numbers are an ordered field.*

Definition 2.3.4. The *absolute value* $|x|$ of a rational number x is defined to be

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

Theorem 2.3.5. *The absolute value satisfies* $\forall x, y \in \mathbb{Q}$*:*

- $|x \cdot y| = |x| \cdot |y|$
- *triangle inequality:* $|x + y| \leq |x| + |y|$

2.4 Real Numbers

Lemma 2.4.1. *"The number $\sqrt{2}$ is not rational", i.e. there exists no rational number whose square equals 2.*

Proof. Assume to get a contradiction that we find $p, q \in \mathbb{Z}$ with $\left(\frac{p}{q}\right)^2 = 2$. We can choose the equivalence class $\frac{p}{q}$ such that p and q have no common divisors. We obtain $p^2 = 2 \cdot q^2$. Hence, 2 divides p, but by assumption then can't divide q. This is a contradiction, since 2^2 now can't divide $2q^2$. □

Definition 2.4.2. A *sequence* of rational numbers $(x_n)_{n\in\mathbb{N}} \subseteq \mathbb{Q}$ is a map $\mathbb{N} \to \mathbb{Q}$, mapping n to x_n.

Definition 2.4.3. Let $(x_n)_{n\in\mathbb{N}} \subseteq \mathbb{Q}$ be a sequence of rational numbers. We call $\lim\limits_{n\to\infty} x_n \in \mathbb{Q}$ a *limit* of this sequence, if

$$\forall \varepsilon > 0 \exists N \in \mathbb{N} : m \geq N \Rightarrow |x_m - \lim_{n\to\infty} x_n| < \varepsilon.$$

If a limit exists, we call the sequence *convergent*. If no limit exists, we call the sequence *divergent*.

Example 2.4.4.

$$x_n = \frac{1}{n+1}$$

converges to 0. We need to show that given any $\varepsilon = \frac{p}{q} > 0$ with $p, q \in \mathbb{N}$ there exists an $N \in \mathbb{N}$ such that $x_n < \frac{p}{q}$ for all $n \geq N$. Choosing $N = q$ is sufficient.

Definition 2.4.5. A *Cauchy sequence* $(x_n)_{n\in\mathbb{N}} \subseteq \mathbb{Q}$ of rational numbers is a sequence such that $\forall \varepsilon > 0 \exists N \in \mathbb{N} : m, n \geq N \Rightarrow |x_n - x_m| < \varepsilon$.

Lemma 2.4.6. *Every convergent sequence of rational numbers is a Cauchy sequence.*

Proof. Choose N s.t. $n \geq N \Longrightarrow |x_n - a| < \varepsilon/2$. Then

$$|x_n - x_m| = |x_n - a + a - x_m| \leq |x_n - a| + |x_m - a| < \epsilon$$ □

Lemma 2.4.7. *Every Cauchy sequence of rational numbers is bounded.*

Proof. Assume that (x_n) is a Cauchy sequence, i.e. $\forall \varepsilon > 0 \exists N \in \mathbb{N} : m, n \geq N \Rightarrow |x_n - x_m| < \varepsilon$. Pick any $\varepsilon > 0$, e.g. $\varepsilon = 1$, and the corresponding $N \in \mathbb{N}$. Then all x_m with $m \geq N$ are bounded by $x_N + 1$. □

Corollary 2.4.8. *Convergent sequences are bounded.*

Definition 2.4.9. We call two Cauchy sequences (x_n), (y_n) equivalent, if $(x_n - y_n)$ converges to 0. The *real numbers*, denoted by $\mathbb{R}$ are the set of all such equivalence classes.

We view the rational numbers $\mathbb{Q}$ as a subset of the real numbers $\mathbb{R}$ in the following way: every rational number r corresponds to a Cauchy sequence $(r, r, r, \dots)$ which defines a real number. We define addition and multiplication the following way

$$(x_n) + (y_n) = (x_n + y_n)$$
$$(x_n) \cdot (y_n) = (x_n \cdot y_n)$$

These operations are indeed well-defined, i.e. they do not depend on the choice of the representative of the equivalence class. In order to prove the well-defined-ness, we need the following two results about limits:

Lemma 2.4.10. *Let $(x_n)_{n\in\mathbb{N}} \subseteq \mathbb{Q}$ and $(y_n)_{n\in\mathbb{N}} \subseteq \mathbb{Q}$ convergent sequences of rational numbers and let $r \in \mathbb{Q}$. Then the following holds:*

1. $\lim\limits_{n\to\infty}(r \cdot x_n) = r \cdot \lim\limits_{n\to\infty} x_n$
2. $\lim\limits_{n\to\infty}(x_n + y_n) = (\lim\limits_{n\to\infty} x_n) + (\lim\limits_{n\to\infty} y_n)$
3. $\lim\limits_{n\to\infty}(x_n \cdot y_n) = (\lim\limits_{n\to\infty} x_n) \cdot (\lim\limits_{n\to\infty} y_n)$

Proof. We denote $\lim_{n\to\infty} x_n = A$ and $\lim_{n\to\infty} y_n = B$.

1. If $r = 0$, then the statement is obvious. Assume $r \neq 0$. Given $\varepsilon > 0$, choose $N \in \mathbb{N}$ s.t. $n \geq N \implies |x_n - A| < \varepsilon/|r|$. Then $|rx_n - rA| = |r| \cdot |x_n - A| < \varepsilon$.
2. Given $\varepsilon > 0$, choose $N_1 \in \mathbb{N}$ s.t. $n \geq N_1 \implies |x_n - A| < \varepsilon/2$, and choose $N_2 \in \mathbb{N}$ s.t. $n \geq N_2 \implies |y_n - B| < \varepsilon/2$. Then $|(x_n + y_n) - (A + B)| \leq |x_n - A| + |y_n - B| < \varepsilon$ if $n \geq N := \max(N_1, N_2)$.
3. The statement is obvious for $A = 0$. Assume $A \neq 0$. Since x_n and y_n are convergent sequences, they are bounded. Let $M > 0$ be such that $\forall n \in \mathbb{N} : |y_n| < M$. Given $\varepsilon > 0$, choose $N_1 \in \mathbb{N}$ s.t. $n \geq N_1 \implies |x_n - A| < \varepsilon/(2M)$, and choose $N_2 \in \mathbb{N}$ s.t. $n \geq N_2 \implies |y_n - B| < \varepsilon/(2A)$. Then $|(x_n \cdot y_n) - (A \cdot B)| = |y_n \cdot (x_n - A) + A \cdot (y_n - B)| \leq M \cdot \varepsilon/(2M) + A \cdot \varepsilon/(2 \cdot |A|) < \varepsilon$ if $n \geq N := \max(N_1, N_2)$. □

Similarly we prove the following lemma.

Lemma 2.4.11. *Let $(x_n)_{n\in\mathbb{N}} \subseteq \mathbb{Q}$ and $(y_n)_{n\in\mathbb{N}} \subseteq \mathbb{Q}$ by Cauchy sequences and let $r \in \mathbb{Q}$. Then the following are Cauchy sequences:*

1. $(r \cdot x_n)$
2. $(x_n + y_n)$
3. $(x_n \cdot y_n)$

Remark 2.4.12. For proving the existence of a multiplicative inverse and for showing that the order relation that we are about to define, is indeed well-defined, it is useful to note that if $(w_n)_{n\in\mathbb{N}}$ is a sequence that does not converge to 0, then

$$\exists \varepsilon > 0 \forall N \exists n \geq N \, : \, |w_n| \geq \varepsilon$$

Lemma 2.4.13. *Let $(w_n)_{n\in\mathbb{N}}$ be a Cauchy sequence that is not equivalent to 0. Then there exists $\varepsilon > 0$ and an $N \in \mathbb{N}$ such that for all $n \geq N$ we have that $|w_n| \geq \varepsilon$.*

Proof. By the previous remark we have that

$$\exists \varepsilon > 0 \forall \widehat{N} \exists n \geq \widehat{N} \, : \, |w_n| \geq 2\varepsilon$$

By the definition of a Cauchy sequence, we find for this $\varepsilon > 0$ an $\widetilde{N}$ such that for all $n, m \geq \widetilde{N}$ we have $|w_n - w_m| < \varepsilon$. We choose an $m \geq \max(\widehat{N}, \widetilde{N})$ where the inequality $|w_m| \geq 2\varepsilon$ holds, and obtain $|w_n| \geq \varepsilon$ for $n \geq m$. Choose $N = m$. □

It is now easy to verify that:

Theorem 2.4.14. *The real numbers form a field.*

Keep in mind that the elements of $\mathbb{R}$ are equivalence classes of Cauchy sequences. We first need to define the notion of order and of "distance" for real numbers. Later we will also look at sequences of real numbers, i.e. sequences of Cauchy sequences.

Definition 2.4.15. We define an order for two real numbers

$$(x_n)_{n\in\mathbb{N}} \leq (y_n)_{n\in\mathbb{N}} \Longleftrightarrow ((x_n)_{n\in\mathbb{N}} \sim (y_n)_{n\in\mathbb{N}}) \vee (\exists N \, : \, n \geq N \Rightarrow x_n \leq y_n)$$

Again, this definition is indeed independent of the choice of the equivalence classes.

Theorem 2.4.16. *The real numbers are an ordered field.*

The basic properties can be verified as an exercise.

Proof that $\leq$ is well-defined. Let $(x_n) \sim (x'_n)$ and $(y_n) \sim (y'_n)$. If $(x_n) \sim (y_n)$ then $(x'_n) \sim (y'_n)$ by transitivity. Hence we assume $(x_n) \not\sim (y_n)$. Therefore, there exists an $\varepsilon > 0$ and an $\hat{N}$ such that $n \geq \hat{N} \Longleftrightarrow |x_n - y_n| \geq 3\varepsilon$. Moreover, by the definition of $(x_n) \leq (y_n)$ there exists an $\widetilde{N}$ such that $n \geq \widetilde{N} \Longrightarrow x_n \leq y_n$. For $n \geq \max(\hat{N}, \widetilde{N})$ we can conclude:

$$x_n + 3\varepsilon \leq y_n.$$

Now, since $\lim_{n\to\infty}(x_n - x'_n) = 0$ and $\lim_{n\to\infty}(y_n - y'_n) = 0$, for each $\varepsilon > 0$ there exist N_1 and N_2 such that $n \geq N_1 \Longrightarrow |x_n - x'_n| < \varepsilon$ and $n \geq N_2 \Longrightarrow || < \varepsilon$. We obtain

$$x'_n = x_n - x_n + x'_n + y_n - y'_n \leq x_n + \varepsilon + \varepsilon \leq y_n - \varepsilon \leq y'_n$$

for $n \geq N := \max(N_1, N_2, \hat{N}, \widetilde{N})$. □

Definition 2.4.17. The *absolute value* $|x|$ of a real number x is defined to be

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

This is equivalent to $|(x_n)_{n\in\mathbb{N}}| = (|x_n|)_{n\in\mathbb{N}}$. The triangle inequality

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

then follows for real numbers, because it holds for each element of the Cauchy sequence of rational numbers if the index is large enough.

Lemma 2.4.18 (Archimedean Axiom)**.** *For every real number $r \in \mathbb{R}$ there is natural number $n \in \mathbb{N}$ with $r \leq n$.*

Proof. We let $r = (r_0, r_1, r_2, \dots)$ be a Cauchy sequence of rationals. If there is no natural number $M \in \mathbb{N}$ such that $r \leq M$ according to the definition above, then the sequence $(r_0, r_1, r_2, \dots)$ is unbounded. (For rational numbers this obvious, since $p/q \leq p$ for $p, q \in \mathbb{N}, q > 0$.) □

Definition 2.4.19. Let $S \subseteq \mathbb{R}$.

1. An *upper bound* for S is any $r \in \mathbb{R}$ such that $\forall s \in S : s \leq r$.
2. The *supremum* of S, denoted by $\sup S$ is the smallest upper bound for S.
3. A *lower bound* for S is any $r \in \mathbb{R}$ such that $\forall s \in S : s \geq r$.
4. The *infimum* of S, denoted by $\inf S$ is the greatest lower bound for S.

Theorem 2.4.20 (Order-completeness)**.** *Any non-empty $S \subset \mathbb{R}$ that is bounded from above has a supremum $\sup S \in \mathbb{R}$.*

Proof. Let $S \subset \mathbb{R}$ be a subset that is bounded from above. Pick $b_0 \in \mathbb{Q}$ to be a upper bound for S and pick $a_0 \in \mathbb{Q}$ such that there exists at least one $s \in S$ with $a_0 < s$. We inductively construct two sequences of rational numbers (a_n) and (b_n). Let $m = \frac{1}{2}(a_{n-1} + b_{n-1})$. If m is an upper bound for S, then set $b_n = m$ and $a_n = a_{n-1}$. Otherwise, we set set $b_n = b_{n-1}$ and $a_n = m$. By construction $a_n < b_n$. Each a_n is an upper bound for S and no b_n is an upper bound for S. We claim that (b_n) is a Cauchy sequence[2], hence defines a real number and that this number is the supremum of S: A smaller upper bound L would need to satisfy $(a_n) \leq L \leq (b_n)$. But $(a_n) \sim (b_n)$. □

For a subset $S \subseteq \mathbb{R}$ that has no upper bound, we declare $\sup S = \infty$. For a subset $S \subseteq \mathbb{R}$ that has no lower bound, we declare $\inf S = -\infty$.
The existence of the supremum allows us to define roots:

[2]We use the fact that $1 + 1/2 + 1/4 + \cdots + 1/2^k = 2 - 1/2^k$ and that $2^k \geq k$ for $k \in \mathbb{N}$.

Definition 2.4.21. Let $k \in \mathbb{N}, k \geq 1$, and let $a \in \mathbb{R}, a \geq 0$. Then we define the *$k$-th root of a* to be the real number $x \geq 0$ such that $x^k = a$. We denote the x with this property by $\sqrt[k]{a}$.

Exercise 2.4.22 (Binomial formula)**.** Let R be a commutative ring. For $a, b \in R$ and $n \in \mathbb{N}$ we have

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} \cdot a^k \cdot b^{n-k}$$

Theorem 2.4.23. *For any $a \in \mathbb{R}$ with $a \geq 0$ and and any $k \in \mathbb{N}$, the root $\sqrt[k]{a}$ is a uniquely determined real number.*

Proof. Uniqueness: Assume $x^k = y^k = a$. If $x \neq y$, we may assume $x < y$. But this implies $x^k < y^k$, a contradiction. Existence: If $a = 0$, then $s = 0$. In the following we may assume $a > 0$. Let $X := \{x \in \mathbb{R} : x^k \leq a\}$. We have $0 \in X$, so it is non-empty. Moreover, it is bounded from above by $\max(1, a)$. Then its supremum exists, $s = \sup X \in \mathbb{R}$ and $s \geq 0$. We claim that $s^k = a$. Since s is the smallest upper bound, for every $1 > \delta > 0$ there exists $x \in X$ s.t. $s - \delta \leq x$ which implies $(s-\delta)^k \leq x^k \leq a$. By the binomial formula, this can be rewritten as

$$s^k \leq a + \sum_{m=1}^{k} \binom{k}{m} \cdot \delta \cdot (-\delta)^{m-1} \cdot s^{k-m} \leq a + \delta \cdot \sum_{m=1}^{k} \binom{k}{m} \cdot s^{k-m} \leq a + \varepsilon$$

if we choose $\delta = \varepsilon / (\sum_{m=1}^{k} \binom{k}{m} \cdot s^{k-m})$. Since this holds for any small $\varepsilon > 0$, we obtain $s^k \leq a$. Hence, $s \in X$. Assume to get a contradiction that $s^k < a$. In this case there would exist $\delta > 0$ such that $s^k + \delta = a$. We claim that $s + \varepsilon$ for a suitably chosen $\varepsilon > 0$ would be in X, contradicting that s is an upper bound for X:

$$(s+\varepsilon)^k = s^k + \varepsilon \cdot \sum_{m=1}^{k} \binom{k}{m} \cdot \varepsilon^{m-1} \cdot s^{k-m} \leq s^k + \varepsilon \cdot \sum_{m=1}^{k} \binom{k}{m} \cdot s^{k-m} \leq a - \delta/2$$

if we choose $\varepsilon < \delta/2/(\sum_{m=1}^{k} \binom{k}{m} \cdot s^{k-m})$. □

Exercise 2.4.24. For $x, y \in \mathbb{R}$ with $x \geq 0, y \geq 0$ we have that $\sqrt[k]{xy} = \sqrt[k]{x} \cdot \sqrt[k]{y}$.

We have a constructed the real numbers from Cauchy sequences of rational numbers. Now we would like to show that iterating this procedure does not lead any further, i.e. the reals are *Cauchy-complete*.
Let us first transfer the notions of sequences and limits from rational to real numbers:

Definition 2.4.25. A *sequence* of real numbers $(x_n)_{n \in \mathbb{N}} \subseteq \mathbb{R}$ is a map $\mathbb{N} \to \mathbb{R}$, mapping n to x_n.

Definition 2.4.26. Let $(x_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}$ be a sequence of real numbers. We call $A = \lim\limits_{n\to\infty} x_n \in \mathbb{R}$ a *limit* of this sequence, if

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

If a limit exists, we call the sequence *convergent*. If no limit exists, we call the sequence *divergent*.

Definition 2.4.27. A *Cauchy sequence* $(x_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}$ of real numbers is a sequence such that $\forall \varepsilon > 0 \exists N \in \mathbb{N} : m, n \geq N \Rightarrow |x_n - x_m| < \varepsilon$.

Theorem 2.4.28 (Cauchy-completeness)**.** *Let (a_n) be a Cauchy sequence of real numbers. Then (a_n) converges to a real number a.*

Proof. We deduce the Cauchy-completeness from the order-completeness:

1. A Cauchy sequence $(x_n)_{n\in\mathbb{N}} \subset \mathbb{R}$ has at most one accumulation point. Moreover, if such a sequence has an accumulation point $A \in \mathbb{R}$, then the sequence converges with limit A (as an exercise).

2. Let $(x_n)_{n\in\mathbb{N}} \subset \mathbb{R}$ be a Cauchy sequence. Let $S_n := \{x_k : k \in \mathbb{N}, k \geq n\}$. Then $A := \lim_{n\to\infty} S_n = \limsup_{n\to\infty} x_n$ exists, and that A is the limit of the Cauchy sequence $(x_n)_{n\in\mathbb{N}}$: This follows from the previous step, since A will be the only accumulation point. □

2.5 Complex Numbers

One problem of the real numbers is that we can't solve an equation like $x^2 = -1$. Real numbers are an ordered field and hence a square is always positive. By introducing complex numbers we will be able to find solutions for such equations. However, we in turn will need to give up something: the complex numbers can no longer be an ordered field.
As a set, we define the *complex numbers*, denoted by $\mathbb{C}$ to be $\mathbb{R} \times \mathbb{R}$ where we define the following addition and multiplication:

$$(x, y) + (u, v) = (x + u, y + v)$$
$$(x, y) \cdot (u, v) = (x \cdot u - y \cdot v, x \cdot v + y \cdot u)$$

Lemma 2.5.1. *The complex numbers form a field.*

Instead of writing pairs (x, y) for complex numbers, we usually write $x + iy$ where we call i the imaginary unit.
We define an *absolute value* for complex numbers by $|x + iy| = \sqrt{x^2 + y^2}$.

Theorem 2.5.2 (D'Alembert, Gauß)**.** *Every polynomial equation of the form*

$$x^n + c_{n-1}x^{n-1} + \cdots c_1 x + c_0 = 0$$

with $c_0, c_1, \ldots, c_{n-1} \in \mathbb{C}$ and $n \geq 1$ has at least one solution $x \in \mathbb{C}$

Proof. A proof will be given in an undergraduate course on Complex Analysis. □

3 Sequences and Series

3.1 Important Inequalities

Theorem 3.1.1 (Bernoulli's inequality)**.** *Let $x \in \mathbb{R}, x \geq -2$ and $k \in \mathbb{N}$. Then*

$$(1+x)^k \geq 1 + kx$$

The inequality is strict for $x > -2$ with $x \neq 0$ and $k \geq 2$.

Proof. For $k = 0$ and $k = 1$, equality holds for all $x \in \mathbb{R}$. Next we prove the case $k = 2$:

$$(1+x)^2 = 1 + 2x + x^2 \geq 1 + 2x$$

This holds for all $x \in \mathbb{R}$ and the inequality is strict precisely for $x \neq 0$. We proceed by induction in k, from k to $k+2$:

$$\begin{aligned}(1+x)^{k+2} &= (1 + 2x + x^2) \cdot (1+x)^k \\ &\geq (1 + 2x + x^2) \cdot (1 + kx) \\ &= 1 + (k+2) \cdot x + x^2 \cdot (2k + 1 + kx) \\ &\geq 1 + (k+2) \cdot x\end{aligned}$$

The summand $x^2 \cdot (k \cdot (2+x) + 1)$ is strictly positive for $x > -2$ and $k \geq 1$ unless $x = 0$ where it vanishes. □

Corollary 3.1.2.

$$0 < q < 1 \Longrightarrow \lim_{n\to\infty} q^n = 0$$

Proof. We need to show that given any $\varepsilon > 0$ we find an N such that $q^n < \varepsilon$ for $n \geq N$. This inequality is equivalent to $\frac{1}{\varepsilon} < \left(\frac{1}{q}\right)^n$. We write $1/q = 1 + x, x > 0$, and obtain $(1+x)^n \geq 1 + nx$ by Bernoulli's inequality.

$$\frac{1}{\varepsilon} < 1 + Nx \leq 1 + nx \leq (1+x)^n = \left(\frac{1}{q}\right)^n$$

We just need to choose $N > \frac{1}{x} \cdot (\frac{1}{\varepsilon} - 1)$ which is possible by the Archimedean axiom. □

Corollary 3.1.3.

$$a > 0 \Longrightarrow \lim_{n\to\infty} \sqrt[n]{a} = 1$$

Proof. We restrict ourselves to the case $a > 1$. For $0 < a < 1$, consider $\sqrt[n]{1/a} = 1/\sqrt[n]{a}$ instead. In this case we have $\sqrt[n]{a} > 1$, hence the inequality $|\sqrt[n]{a} - 1| < \varepsilon$ is equivalent to $a < (1+\varepsilon)^n$. By Bernoulli's inequality, it is enough to choose $N \geq n$ for an $N \in \mathbb{N}$ such that $a < 1 + n\varepsilon$, i.e. $N > \frac{a-1}{\varepsilon}$ which exists by the Archimedean axiom. □

Theorem 3.1.4 (Inequality of arithmetic and geometric mean (AGM inequality))**.** *Let $a_1, \ldots, a_n \in \mathbb{R}$ with $a_1 \geq 0, \ldots, a_n \geq 0$. Then the following inequality holds:*

$$\frac{a_1 + a_2 + \cdots + a_n}{n} \geq \sqrt[n]{a_1 \cdot a_2 \cdots a_n}$$

Equality holds if and only if $a_1 = a_2 = \cdots = a_n$.

Proof using Bernoulli's inequality. We prove inequality by induction in n. For $n = 1$ we have equality. Since the r.h.s. vanishes if any of the numbers is zero, we may assume $a_1 > 0, \ldots, a_n > 0$. For the induction step, assume that we have ordered the numbers s.t. $a_{n+1} = \max(a_1, \ldots, a_n, a_{n+1})$. Let $M = (a_1 + a_2 + \cdots + a_n)/n$ be the arithmetic mean of the first n numbers. We have $a_{n+1} \geq M$. We first observe that

$$\frac{a_1 + \cdots + a_n + a_{n+1}}{(n+1) \cdot M} = \frac{nM + M - M + a_{n+1}}{(n+1) \cdot M} = 1 + \frac{a_{n+1} - M}{(n+1) \cdot M}$$

Applying Bernoulli's inequality, we now obtain

$$\left(\frac{a_1 + \cdots + a_n + a_{n+1}}{(n+1) \cdot M}\right)^{n+1} = \left(1 + \frac{a_{n+1} - M}{(n+1) \cdot M}\right)^{n+1} \geq 1 + \frac{a_{n+1} - M}{M} = \frac{a_{n+1}}{M}$$

Multiplication with M^{n+1} and using the induction hypothesis yields:

$$\left(\frac{a_1 + \cdots + a_n + a_{n+1}}{(n+1)}\right)^{n+1} \geq a_{n+1} M^n \geq a_{n+1} \cdot a_1 \cdots a_n$$

In order to have equality, we would need $a_{n+1} = M$ for Bernoulli's inequality and $M = a_1 = \cdots = a_n$ from the induction hypothesis. □

3.2 Real sequences

Lemma 3.2.1. *Every convergent sequence of real numbers is a Cauchy sequence.*

Lemma 3.2.2. *Every Cauchy sequence of real numbers is bounded.*

Lemma 3.2.3. *Let $(x_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}$ and $(y_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}$ convergent sequences of real numbers and let $r \in \mathbb{R}$. Then the following holds:*

1. $\lim\limits_{n\to\infty} (r \cdot x_n) = r \cdot \lim\limits_{n\to\infty} x_n$
2. $\lim\limits_{n\to\infty} (x_n + y_n) = (\lim\limits_{n\to\infty} x_n) + (\lim\limits_{n\to\infty} y_n)$
3. $\lim\limits_{n\to\infty} (x_n \cdot y_n) = (\lim\limits_{n\to\infty} x_n) \cdot (\lim\limits_{n\to\infty} y_n)$
4. *If* $\lim_{n\to\infty} y_n \neq 0$, *then* $\lim\limits_{n\to\infty} (x_n / y_n) = (\lim\limits_{n\to\infty} x_n)/(\lim\limits_{n\to\infty} y_n)$

Proof of 4. Since $\lim_{n\to\infty} y_n \neq 0$, there exists $\delta > 0$ and and N such that $n \geq N \implies |y_n| \geq \delta$. Let $A = \lim_{n\to\infty} x_n$ and $B = \lim_{n\to\infty} y_n$. By the definition of limits, given $\varepsilon > 0$ there exist N_1 and N_2 such that $|x_n - A| < \delta\varepsilon/2$ for $n \geq N_1$ and $|y_n - B| < |B/A| \cdot \delta\varepsilon/2$ for $n \geq N_2$, assuming $A \neq 0$ (If $A = 0$, then the second estimate won't be needed at all.)

$$\left|\frac{x_n}{y_n} - \frac{A}{B}\right| = \left|\frac{Bx_n - Ay_n}{By_n}\right| \leq \frac{|B| \cdot |x_n - A| + |A| \cdot |B - y_n|}{|B| \cdot \delta} < \varepsilon$$

for $n \geq \max(N, N_1, N_2)$. □

Example 3.2.4.

$$\lim_{n\to\infty} \frac{4n^3 + n^2 + 3}{n^3 + 7n + 2} = \lim_{n\to\infty} \frac{4 + 1/n + 3/n^3}{1 + 7/n^2 + 2/n^3} = \frac{\lim\limits_{n\to\infty} 4 + \lim\limits_{n\to\infty} 1/n + \lim\limits_{n\to\infty} 3/n^3}{\lim\limits_{n\to\infty} 1 + \lim\limits_{n\to\infty} 7/n^2 + \lim\limits_{n\to\infty} 2/n^3} = 4$$

using the previously established $\lim_{n\to\infty} \frac{1}{n} = 0$ and all the properties 1, 2, 3, 4 from the preceding lemma.

Theorem 3.2.5. *A bounded and monotonously increasing (or decreasing) sequence of real numbers converges.*

Proof. Let (x_n) be the series under consideration. We have $\forall n \in \mathbb{N} : x_{n+1} \geq x_n$. Since the series is bounded, its supremum $s = \sup\{x_n : n \in \mathbb{N}\} \in \mathbb{R}$ exists. Indeed s is the limit: First, $\forall n \in \mathbb{N} : n \leq s$ and since there can't be be smaller upper bound, $\forall\varepsilon > 0 \exists m \in \mathbb{N} : 0 \geq s - x_m < \varepsilon$. Hence, for any given $\varepsilon > 0$, we choose m as above and obtain for all $m \geq m$ that $|s - x_n| < \varepsilon$. □

Lemma 3.2.6. *If the limit of a sequence exists, then it is unique.*

Definition 3.2.7. Let $(x_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}$ be a sequence. We call a an *accumulation point* of this sequence, if

$$\forall\varepsilon > 0 \forall N \in \mathbb{N} \exists m \geq N : |x_m - a| < \varepsilon.$$

Note that a limit is always an accumulation point.

Definition 3.2.8. Let $(x_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}$ be sequence of real numbers. We define

$$\liminf_{n\to\infty} x_n = \lim_{n\to\infty} \inf\{x_k : k \geq n\}$$
$$\limsup_{n\to\infty} x_n = \lim_{n\to\infty} \sup\{x_k : k \geq n\}$$

Remark 3.2.9. Note that the lim sup is the limit of a decreasing sequence. If the sequence of suprema is constantly equal to $-\infty$, we also set the lim sup to be $-\infty$. If the lim sup does not exist, then the sequence has no upper bound and we write $\liminf_{n\to\infty} x_n = +\infty$. Similarly, the lim inf is the limit of a increasing sequence. If the sequence of infima is constantly equal to $+\infty$, we also set the lim inf to be $+\infty$. If it does not exist, then the sequence has no lower bound and we write $\liminf_{n\to\infty} x_n = -\infty$.

Lemma 3.2.10. *For any sequence* $(x_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}$ *we have*

1. $\liminf_{n\to\infty} x_n \leq \limsup_{n\to\infty} x_n$
2. $\liminf_{n\to\infty} x_n = \limsup_{n\to\infty} x_n \iff (x_n)_{n\in\mathbb{N}}$ *is convergent, and* $\liminf_{n\to\infty} x_n = \limsup_{n\to\infty} x_n = \lim_{n\to\infty} x_n$.

3.3 Series

Definition 3.3.1. Let (x_n) be a sequence of real numbers. Then we define write (formal) *series* as:

$$\sum_{n=0}^{\infty} x_n$$

The *partial sums* $S_m = \sum_{n=1}^{m} x_n$ form a sequence of real numbers. If the sequence $(S_m)_{m\in\mathbb{N}}$ converges, we say that the series is *convergent*, otherwise we say that it is *divergent*. If it converges, we define *limit* to be the limit of the partial sums:

$$\sum_{n=0}^{\infty} x_n = \lim_{m\to\infty} \sum_{n=0}^{m} x_n$$

Theorem 3.3.2. *For* $q \in \mathbb{R}$ *with* $|q| < 1$ *the* geometric series $\sum_{n=0}^{\infty} q^n$ *converges towards* $\frac{1}{1-q}$. *The series diverges for* $|q| \geq 1$.

Proof. Let $S_m = \sum_{n=0}^{m} q^n$ be the m-th partial sum. Observe that $S_m - qS_m = 1 - q^{m+1}$, hence $S_m = \frac{1-q^{m+1}}{1-q}$ for $q \neq 1$. For $q = 1$ the series is obviously divergent. □

Lemma 3.3.3 (Cauchy criterion)**.** *A series* $\sum_{n=0}^{\infty} x_n$ *is convergent if and only if:*

$$\forall \varepsilon > 0 \exists N \in \mathbb{N} : n \geq m \geq N \Longrightarrow \left| \sum_{k=m}^{n} x_k \right| < \varepsilon$$

Corollary 3.3.4. *If a series* $\sum_{n=1}^{\infty} x_n$ *of real numbers converges, then* $\lim_{n\to\infty} x_n = 0$.

The converse does not hold, as the following important example shows.

Example 3.3.5. The *harmonic series*

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

is divergent. If it were convergent, then the partial sums would form a Cauchy sequence, i.e. for all $\varepsilon > 0$ there should be an $N \in \mathbb{N}$ such that for all $m \geq k \geq N$ we have

$$\sum_{n=k}^{m} \frac{1}{n} < \varepsilon$$

We choose $\varepsilon = \frac{1}{3}$. We choose $k = 2^L$ and $m = 2k$ for some $L \in \mathbb{N}$ such that $2^L \geq N$. Then $\frac{1}{2 \cdot 2^L}$ is a lower bound for each summand, and there are $2^L + 1$ summands. Hence, $\sum_{n=k}^{m} \frac{1}{n} \geq \frac{1}{2}$, a contradiction.

Lemma 3.3.6. *Let $A = \sum_{n=0}^{\infty} a_n$ and $B = \sum_{n=0}^{\infty} b_n$ be convergent series. Then $\sum_{n=0}^{\infty}(a_n + b_n)$ is a convergent series with limit $A + B$.*

Proof. Let A_n and B_n denote the resp. partial sums:

$$A_n = \sum_{k=0}^{n} a_k \quad \text{and} \quad B_n = \sum_{k=0}^{n} b_k$$

By associativity and commutativity of addition (for finite sums) have that

$$\sum_{k=0}^{n}(a_k + b_k) = A_n + B_n$$

The r.h.s. converges to $A + B$. □

Definition 3.3.7. We call a series $\sum_{n=0}^{\infty} x_n$ *absolutely convergent* if the series $\sum_{n=0}^{\infty} |x_n|$ converges.

Lemma 3.3.8. *An absolutely convergent series is convergent.*

A series that is convergent, but not absolutely convergent is called *conditionally convergent*. The following comparison test works only for positive sumands and is therefore often useful for showing/disproving absolute convergence.

Theorem 3.3.9 (Comparison test)**.**
Let $\sum_{n=0}^{\infty} a_n$ and $\sum_{n=0}^{\infty} b_n$ real series with $a_n \geq b_n \geq 0$ for every $n \in \mathbb{N}$.

1. *If $\sum_{n=0}^{\infty} a_n$ converges then $\sum_{n=0}^{\infty} b_n$ converges.*
2. *If $\sum_{n=0}^{\infty} b_n$ diverges then $\sum_{n=0}^{\infty} a_n$ diverges.*

Proof. Let A_n and B_n denote the resp. partial sums:

$$A_n = \sum_{k=0}^{n} a_k \quad \text{and} \quad B_n = \sum_{k=0}^{n} b_k$$

Since all of their summands are non-negative, A_n and B_n are monotonously increasing, and $a_n \geq b_n \geq 0$ implies $A_n \geq B_n$.

1. If A_n converges, monotonously increasing, to A, then we have $A \geq A_n \geq B_n \geq 0$. Hence, B_n is bounded and monotonously increasing and therefore convergent.
2. If B_n is not convergent, but monotonously increasing, then $\lim_{n\to\infty} B_n = +\infty$. Now $A_n \geq B_n$ implies that $\lim_{n\to\infty} A_n = +\infty$ as well. □

Let's look at the following motivation for the definition of the product of two series:

$$(a_0 + a_1x + a_2x^2) \cdot (b_0 + b_1x + b_2x^2) = a_0b_0 + (a_0b_1 + a_1b_0) \cdot x + (a_0b_2 + a_1b_1 + a_2b_0) \cdot x^2 + ...$$

So in front of the power x^k we have the coefficients whose indices sum up to k. For the definition of the product in the general case, we forget about the x (set $x = 1$).

Definition 3.3.10. *Cauchy product of two series* Let $\sum_{n=0}^{\infty} a_n$ and $\sum_{n=0}^{\infty} b_n$ be series. We define the *Cauchy product* as the following series:

$$\sum_{n=0}^{\infty} c_n, \quad \text{where } c_n = \sum_{k=0}^{n} a_k \cdot b_{n-k}$$

Theorem 3.3.11. *Let $A = \sum_{n=0}^{\infty} a_n$ and $B = \sum_{n=0}^{\infty} b_n$ be convergent series and assume that at least one of them is absolutely convergent. Then the Cauchy product of A and B is convergent with limit $A \cdot B$.*

Remark 3.3.12. The Cauchy product of two convergent, but not absolutely convergent series may diverge. The Cauchy product of two absolutely convergent series is absolutely convergent.

Lemma 3.3.13. *Let $\sum_{n=0}^{\infty} a_n$ and $\sum_{n=0}^{\infty} b_n$ be series and let A_n and B_n denote their respective partial sums, i.e. $A_n = \sum_{k=0}^{n} a_k$ and $B_n = \sum_{k=0}^{n} b_k$. Then the following holds:*

$$\sum_{k=0}^{n} \sum_{m=0}^{k} a_m b_{k-m} = \sum_{k=0}^{n} a_{n-k} B_k$$

Proof. Exercise. □

Proof of Theorem 3.3.11. Let $A_n = \sum_{k=0}^{n} a_k$, $B_n = \sum_{k=0}^{n} b_k$ and $C_n = \sum_{k=0}^{n} c_k$ denote the partial sums of these series. We need to show that for any given $\varepsilon > 0$ we can find an $N \in \mathbb{N}$ such that $|C_n - AB| < \varepsilon$ for $n \geq N$.
By the preceding lemma we obtain

$$C_n = \sum_{k=0}^{n} \sum_{m=0}^{k} a_m b_{k-m} = \sum_{k=0}^{n} a_{n-k} B_k = \sum_{k=0}^{n} a_{n-k}(B_k - B) + A_n \cdot B \tag{1}$$

Without loss of generality we can assume that $\sum_{k=0}^{\infty} a_k$ is absolutely convergent, hence $\sum_{k=0}^{\infty} |a_k| \in \mathbb{R}$. Since B_n converges to B, there exists a $N_1 \in \mathbb{N}$ such that

$$n \geq N_1 \Longrightarrow |B_n - B| < \frac{\varepsilon}{3(\sum_{k=0}^{\infty} |a_k| + 1)} \tag{2}$$

Since A_n converges to A we conclude that there exists a $N_2 \in \mathbb{N}$ such that

$$n \geq N_2 \Longrightarrow |A_n - A| < \frac{\varepsilon}{3(|B| + 1)} \tag{3}$$

Again since A_n converges to A we conclude, as a consequence of the Cauchy criterion, we have that $\lim_{n\to\infty} a_n = 0$, i.e. there exists a $N_3 \in \mathbb{N}$ such that

$$n \geq N_3 \Longrightarrow |a_n| < \frac{\varepsilon}{3N_1 \cdot (\sup\{|B_k - B| \,:\, 0 \leq k \leq N_1 - 1\} + 1)} \tag{4}$$

Note that the choice of N_3 depends on ε and on the previously chosen N_1. We can now estimate $|C_n - AB|$ if we choose $n \geq N := \max(N_1 + N_3, N_2)$

$$\begin{aligned} |C_n - AB| &\overset{(1)}{=} \left| \sum_{k=0}^{n} a_{n-k}(B_k - B) + (A_n - A) \cdot B \right| \\ &\leq \underbrace{\sum_{k=0}^{N_1-1} \left| a_{\underbrace{n-k}_{\geq N_3}}(B_k - B) \right|}_{<\varepsilon/3 \text{ by (4)}} + \underbrace{\sum_{k=N_1}^{n} |a_{n-k}(B_k - B)|}_{<\varepsilon/3 \text{ by (2)}} + \underbrace{|(A_n - A) \cdot B|}_{<\varepsilon/3 \text{ by (3)}} < \varepsilon \end{aligned}$$ □

3.4 Tests for Convergence

The geometric series is the basis for the following two convergence tests: ratio test and root test.

Theorem 3.4.1 (Ratio test)**.** *Let $\sum_{n=0}^{\infty} a_n$ be a series where $a_n \neq 0$ for all $n \in \mathbb{N}$. Let $L := \lim_{n\to\infty} \frac{|a_{n+1}|}{|a_n|}$. If $L < 1$, then this series converges absolutely. If $L > 1$, then this series diverges.*

Proof. Set $q = \frac{1+L}{2}$.

1. If $L < 1$, then $L < q < 1$. Since $\lim_{n\to\infty} \frac{|a_{n+1}|}{|a_n|} = L$, there exist $N \in \mathbb{N}$ s.t. $n \geq N \Longrightarrow 0 < \frac{|a_{n+1}|}{|a_n|} < q$ which leads to

   $$|a_{N+1}| < q \cdot |a_N|,\ |a_{N+2}| < |a_{N+1}| \cdot q < |a_N| \cdot q^2, \ldots, |a_{N+k}| < |a_N| \cdot q^k$$

   Hence we have that

   $$\sum_{n=0}^{\infty} |a_n| = \sum_{n=0}^{N-1} |a_n| + \sum_{n=N}^{\infty} |a_n| \leq \underbrace{\sum_{n=0}^{N-1} |a_n|}_{\text{finite sum}} + \underbrace{\sum_{n=N}^{\infty} |a_N| \cdot q^{n-N}}_{\text{geometric series with } |q|<1}$$

   and the series converges absolutely.

2. If $L > 1$, then $1 < q < L$. Since $\lim_{n\to\infty} \frac{|a_{n+1}|}{|a_n|} = L$, there exist $N \in \mathbb{N}$ s.t. $n \geq N \Longrightarrow \frac{|a_{n+1}|}{|a_n|} > q$ which similarly as above leads to $|a_{N+k}| > |a_N| \cdot q^k$. But since $q > 1$, this implies $|a_n| > |a_N| \neq 0$ for $n \geq N$, so in particular $\lim_{n\to\infty} a_n > |a_N| \neq 0$, hence the series diverges. □

Remark 3.4.2. Note that there is no statement for $L = 1$. In this case the ratio test does not make a statement, and absolute convergence, conditional convergence or divergence are possible.

Example 3.4.3. We define the *exponential function* $\exp\colon \mathbb{R} \to \mathbb{R}$ by

$$x \mapsto \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

We want to show that this series converges absolutely for every $x \in \mathbb{R}$. For $x = 0$, we are summing up zeros for $n \geq 1$, hence it is a finite sum, and $\exp(0) = 1$. For $x \neq 0$, we apply the ratio test:

$$L = \lim_{n\to\infty} \frac{\left|\frac{x^{n+1}}{(n+1)!}\right|}{\left|\frac{x^n}{n!}\right|} = |x| \cdot \lim_{n\to\infty} \frac{1}{n+1} = 0 < 1$$

Moreover, using the formula for the Cauchy product and the binomial formula, it is then easy to show that $\exp(x) \cdot \exp(y) = \exp(x + y)$. We also define the *Euler number* $e = \exp(1) = \sum_{n=0}^{\infty} \frac{1}{n!}$.

Theorem 3.4.4 (Root test)**.** *Let $\sum_{n=0}^{\infty} a_n$ be a series. Let $L := \limsup_{n\to\infty} \sqrt[n]{|a_n|}$. If $L < 1$, then this series converges absolutely. If $L > 1$, then this series diverges.*

Proof. Set $q = \frac{1+L}{2}$. We proceed similarly as in the proof of the quotient test.

1. If $L < 1$, then there exist $N \in \mathbb{N}$ s.t. $n \geq N \implies 0 \leq \sqrt[n]{|a_n|} < q < 1$ which leads to

   $$|a_n| < q^n$$

   Hence we have that

   $$\sum_{n=0}^{\infty} |a_n| \leq \underbrace{\sum_{n=0}^{N-1} |a_n|}_{\text{finite sum}} + \sum_{n=N}^{\infty} q^n$$

   and the series converges absolutely.

2. If $L > 1$, then $\limsup_{n\to\infty} \sqrt[n]{|a_n|} = L > q > 1$ implies that there is a subsequence with indices n_k such that $\sqrt[n_k]{|a_{n_k}|} > q$ which is equivalent to

   $$|a_{n_k}| > q^{n_k} > 1$$

   hence the series diverges. □

Example 3.4.5. Let $a_n = \begin{cases} 1/2^n, & n \text{ even} \\ 1/3^n, & n \text{ odd} \end{cases}$

Then the sequence $\sqrt[n]{|a_n|}$ has two accumulation points, namely $1/2$ and $1/3$, hence the limit of $\sqrt[n]{|a_n|}$ does not exist, but $\limsup \sqrt[n]{|a_n|} = 1/2 < 1$. This implies that $\sum_{n=0}^{\infty} a_n$ converges absolutely.

Theorem 3.4.6 (Cauchy condensation test)**.** *Let $a_n \geq a_{n+1} \geq 0$ be a sequence of non-increasing, non-negative real numbers. Then*

$$\sum_{n=0}^{\infty} a_n \text{ converges absolutely } \iff \sum_{n=0}^{\infty} 2^n \cdot a_{2^n} \text{ converges absolutely}$$

Proof. Since $a_n \geq 0$ we only need to show that the partial sums have an upper bound.
$\boxed{\Longrightarrow}$ We claim the following estimate:

$$2(a_0 + a_1 + a_2 + \cdots + a_{2^n-1}) \geq a_1 + 2a_2 + 2^2 a_{2^2} + \cdots + 2^n a_{2^n}$$

which then implies that the series on the right is bounded by twice the series on the left. Proof by induction in n.
$\boxed{\Longleftarrow}$ The estimate

$$2^n a_{2^n} \geq a_{2^n} + a_{2^n+1} + a_{2^n+2} + \cdots + a_{2^{n+1}-1}$$

shows that the series on the left is bounded from above by the series on the right. □

Example 3.4.7. We define the *natural logarithm* $\log\colon \mathbb{R}^+ \to \mathbb{R}$ to be the inverse function of the exponential function $\exp\colon \mathbb{R} \to \mathbb{R}^+$. Observe that this implies $\log(x \cdot y) = \log(x) + \log(y)$, in particular $\log(x^k) = k \cdot \log(x)$ for $k \in \mathbb{N}$. Moreover, monotonicity of exp implies monotonicity of log: $0 < x < y \Longrightarrow \log(x) < \log(y)$.
The series $\sum_{n=3}^{\infty} \frac{1}{n \cdot \log(n)}$ diverges: We apply the Cauchy condensation test to $a_n = \frac{1}{n \cdot \log(n)}$. Indeed, a_n is monotonously decreasing and positive. We have $2^n \cdot a_{2^n} = \frac{2^n}{2^n \log(2^n)} = \frac{1}{\log 2} \frac{1}{n}$ and the corresponding series diverges since the harmonic series diverges.

Theorem 3.4.8 (Dirichlet's test)**.** *Let $a_n \geq a_{n+1} \geq 0$ be a sequence of non-increasing, non-negative real numbers with $\lim_{n\to\infty} a_n = 0$. Let b_n be a sequence of real numbers such that $\exists M > 0\ \forall n \in \mathbb{N} : |\sum_{k=1}^{n} b_k| \leq M$. Then the following series converges:*

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

Proof. Let $S_n = \sum_{k=0}^{n} a_k b_k$ and $B_n = \sum_{k=0}^{n} b_k$. We need to show that the S_n form a convergent sequence. We make use of the following identity (proof by induction, as exercise):

$$S_n = a_n B_n + \sum_{k=0}^{n-1} B_k \cdot (a_k - a_{k+1})$$

Since $|B_n| \leq M$, we have $\lim_{n\to\infty} a_n B_n = 0$. Moreover, the remaining sum converges absolutely:

$$\sum_{k=0}^{n-1} |B_k \cdot (a_k - a_{k+1})| \leq \sum_{k=0}^{n-1} |M \cdot (a_k - a_{k+1})| = M \cdot \sum_{k=0}^{n-1} (a_k - a_{k+1}) \leq M \cdot a_0 \qquad \square$$

Example 3.4.9. The *alternating harmonic series* $\sum_{n=0}^{\infty} \frac{(-1)^n}{n+1}$ is conditionally convergent. We have already shown that it does not converge absolutely, since the harmonic series is divergent. We apply Dirichlet's test: choose $a_n = 1/n$ and $b_n = (-1)^n$, and then $M = 1$.

3.5 Rearranging

Definition 3.5.1. Let $\sum_{n=0}^{\infty} a_n$ be a series. Let $\varphi\colon \mathbb{N} \to \mathbb{N}$ be a bijective map. We call $\sum_{n=0}^{\infty} a_{\varphi(n)}$ a *rearranging* of $\sum_{n=0}^{\infty} a_n$.

Theorem 3.5.2. *Let $A = \sum_{n=0}^{\infty} a_n$ be an absolutely convergent series. Then every rearranging of $\sum_{n=0}^{\infty} a_{\varphi(n)}$ is also absolutely convergent with the same limit A.*

Proof. Consider the partial sums

$$A_n = \sum_{k=0}^{n} a_k, \quad B_n = \sum_{k=0}^{n} a_{\varphi(k)}$$

Given $n \in \mathbb{N}$ in there are $(n+1)$ values $\varphi(0), \varphi(1), \ldots, \varphi(n) \in \mathbb{N}$. Let $K(n)$ be their maximum. Then all the terms of B_n will also occur in $A_{K(n)}$ where $K(n) \geq n$.
Since $\sum_{k=0}^{\infty} a_k$ converges absolutely, the partial sums of $S_n = \sum_{k=0}^{n} |a_k|$ are monotonously increasing and bounded. But now $S_{K(n)} \geq T_n = \sum_{k=0}^{n} |a_{\varphi(k)}|$ which is also monotonously increasing. Hence $\sum_{k=0}^{\infty} a_{\varphi(k)}$ converges absolutely.
Let $M(n)$ be the maximum of $\varphi^{-1}(0), \varphi^{-1}(1), \ldots, \varphi^{-1}(n)$.
Using the Cauchy criterion for series, for any $\varepsilon > 0$ choose $L \in \mathbb{N}$ such that $n \geq m \geq L$ implies $\sum_{k=m}^{n} |a_k| < \varepsilon/2$ and $\sum_{k=m}^{n} |a_{\varphi(k)}| < \varepsilon/2$. Since we know that A_n and B_n are convergent, the equality $A = B$ follows if we can show that$\lim_{n\to\infty} |A_n - B_n| = 0$.
Let $n \geq N := M(K(L)) \geq K(L) \geq L$. Then

$$|A_n - B_n| = \left| \sum_{k=0}^{n} a_k - \sum_{k=0}^{n} a_{\varphi(k)} \right| \leq \sum_{k=K(L)+1}^{n} |a_k| + \sum_{k=L+1}^{n} |a_{\varphi(k)}| < \varepsilon$$

Since for $k \leq K(L)$, the terms $a_{\varphi(k)}$ of the second sum cancel with terms of the first sum, and for $k \leq M(K(L))$ the term a_k of the first sum cancel with terms of the second sum. □

Remark 3.5.3. A conditionally convergent series can always be rearranged to converge to any value.

3.6 Power Series

Definition 3.6.1. Let $p \in \mathbb{R}$ and let $(a_n)_{n\in\mathbb{N}} \subset \mathbb{R}$. We call

$$\sum_{n=0}^{\infty} a_n \cdot (x-p)^n$$

a *(formal) power series* around p.

Definition 3.6.2. Let

$$\sum_{n=0}^{\infty} a_n \cdot (x-p)^n$$

be a power series around $p \in \mathbb{R}$. We define its *radius of convergence* ρ by

$$\rho = \sup\{\delta \geq 0 \,:\, \exists x \in \mathbb{R} \text{ with } \sum_{n=0}^{\infty} a_n \cdot (x-p)^n \text{ convergent and } |x-p| = \delta\} \in [0, \infty]$$

Theorem 3.6.3. *Let*

$$\sum_{n=0}^{\infty} a_n \cdot (x-p)^n$$

be a power series around $p \in \mathbb{R}$ with radius of convergence ρ. Then for every $x \in \mathbb{R}$ with $|x-p| < \rho$ we have absolute convergence and for every and for every $x \in \mathbb{R}$ with $|x-p| > \rho$ we have divergence.

Proof. We are going to prove a slightly stronger statement: If the series converges for some $x_0 \in \mathbb{R}$, then it converges absolutely for all $x \in \mathbb{R}$ with $|x-p| < |x_0 - p|$. Choose $\delta \in \mathbb{R}$ with $|x-p| < \delta < |x_0 - p|$ and set

$$M := \sup\{|a_n| \cdot |x_0 - p|^n \,:\, n \in \mathbb{N}\}$$

This supremum is finite, since convergence implies that the summands tend to zero. We now obtain

$$|a_n \cdot (x-p)^n| < |a_n| \cdot \delta^n < M \cdot \left(\frac{\delta}{|x_0 - p|}\right)^n$$

where the series with summands as in the r.h.s. converges since it is a geometric series. Since the radius of convergence ρ is defined as the supremum of $|x_0 - p|$ for all such x_0 where we have convergence, the series necessarily diverges for points $x \in \mathbb{R}$ with $|x-p| > \rho$. □

Remark 3.6.4. The radius of convergence ρ does not make any statement what happens for points x with $|x-p| = \rho$. Convergence or divergence can occur in this case.

Example 3.6.5.

1. $\displaystyle\sum_{n=0}^{\infty} x^n$ has radius of convergence 1.
2. $\sin(x) := \displaystyle\sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}$ has radius of convergence ∞.
3. $\cos(x) := \displaystyle\sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}$ has radius of convergence ∞.

4 Continuity

4.1 Open and Closed sets

Definition 4.1.1. We call a set $U \subseteq \mathbb{R}$ *open* if for every $x \in U$ there exists an $\varepsilon > 0$ such that $(x - \varepsilon, x + \varepsilon) \subseteq U$.

Example 4.1.2. Any open interval $(a, b) := \{x \in \mathbb{R} : a < x < b\}$ is open.

Lemma 4.1.3. *We have the following basic properties:*

1. *The empty set $\emptyset$ and $\mathbb{R}$ are open.*
2. *Any union of open sets is open.*
3. *Any finite intersection of open sets is open.*

Note that an infinite intersection of open sets is in general not open.

Definition 4.1.4. A set $A \subseteq \mathbb{R}$ is called *closed* if its complement $\mathbb{R} \setminus A = \{y \in \mathbb{R} : x \notin A\}$ is open.

Theorem 4.1.5. *A set $A \subset \mathbb{R}$ is closed if and only if every sequence $(x_n)_{n\in\mathbb{N}} \subseteq A$ that converges in $\mathbb{R}$ has its limit in A.*

Definition 4.1.6. Let $X \subseteq \mathbb{R}$. A subset $A \subseteq X$ is called *dense* if for every point $x \in X$ there exists a sequence $(a_n)_{n\in\mathbb{N}} \subseteq A$ such that $\lim_{n\to\infty} a_n = x$.

Example 4.1.7. The rational numbers are dense in the real numbers.

4.2 Continuous Functions

Definition 4.2.1. Let $X \subset \mathbb{R}$ be a subset. A function $f\colon X \to \mathbb{R}$ is called *continuous* in $x \in X$ if

$$\forall \varepsilon > 0\, \exists \delta > 0\, \forall y \in X \ : \ |x - y| < \delta \Longrightarrow |f(x) - f(y)| < \varepsilon$$

If f is continuous for all $x \in X$, we say that f is continuous on X or is a continuous function from X to $\mathbb{R}$.

Theorem 4.2.2. *Let $X \subseteq \mathbb{R}$. Then the following hold:*

1. *A function $f\colon X \to \mathbb{R}$ is continuous in $x \in X$ if and only if for every sequence $(x_n)_{n\in\mathbb{N}} \subseteq X$ with $\lim_{n\to\infty} x_n = x$ it follows that $\lim_{n\to\infty} f(x_n) = f(x)$.*
2. *A function $f\colon X \to \mathbb{R}$ is continuous if and only if: $\forall\, U \subseteq \widetilde{X}$ open $\Longrightarrow f^{-1}(U) := \{x \in X : f(x) \in U\}$ open .*

Remark 4.2.3. We denote this characterizing property for continuous maps given above in Point 1 in short as $\lim\limits_{y\to x} f(y) = f(x)$.

Theorem 4.2.4. *Let $X \subseteq \mathbb{R}, Y \subseteq \mathbb{R}$ and let $f\colon X \to Y$ be a function that is continuous in $x \in X$ and $g\colon Y \to \mathbb{R}$ be a function that is continuous in $f(x) \in \widetilde{X}$. Then the composition $y \mapsto g \circ f(y) = g(f(y))$ is a function from $X \to \mathbb{R}$ that is continuous in x.*

Proof. We use the characterization of continuity by sequences: Let $\lim_{n\to\infty} a_n = x$ in X and define $b_n = f(a_n)$. Then by continuity of f in x, we have $\lim_{n\to\infty} b_n = \lim_{n\to\infty} f(a_n) = f(x)$. By continuity of g we have $\lim_{n\to\infty} g(b_n) = g(f(x))$ which is is equivalent to $\lim_{n\to\infty} g(f(a_n)) = g(f(x))$. □

Lemma 4.2.5. *Let $X \subseteq \mathbb{R}$. Let $f, g\colon X \to \mathbb{R}$ be continuous functions. Then the following hold:*

1. *$x \mapsto f(x) + g(x)$ is a continuous function from X to $\mathbb{R}$.*
2. *$x \mapsto f(x) \cdot g(x)$ is a continuous function from X to $\mathbb{R}$.*
3. *$x \mapsto f(x)/g(x)$ is a continuous function from $X \setminus \{p \in X : g(p) = 0\}$ to $\mathbb{R}$.*

Proof. For each case we use the characterization of continuity by sequences:
Let $\lim_{n\to\infty} a_n = x$ in X.

1. $f(x) + g(x) = \lim_{n\to\infty} f(a_n) + \lim_{n\to\infty} g(a_n) = \lim_{n\to\infty}(f(a_n) + g(a_n))$
2. $f(x) \cdot g(x) = \lim_{n\to\infty} f(a_n) \cdot \lim_{n\to\infty} g(a_n) = \lim_{n\to\infty}(f(a_n) \cdot g(a_n))$
3. $f(x)/g(x) = \lim_{n\to\infty} f(a_n) / \lim_{n\to\infty} g(a_n) = \lim_{n\to\infty}(f(a_n)/g(a_n))$ as long as we do not divide by zero □

Example 4.2.6.
1. The identity $\mathrm{id}(x) = x$ from $\mathbb{R}$ to $\mathbb{R}$ is continuous (choose $\delta = \varepsilon$).
2. Since products of continuous functions are continuous, also $x \mapsto x^n$ is continuous.
3. Since sums of continuous functions are continuous, polynomials are continuous.

Example 4.2.7. The exponential function is continuous: Let $x, y \in \mathbb{R}$. We want to show continuity in x. Assume w.l.o.g. $x < y$ which implies $\exp(x) < \exp(y)$. Given any $\varepsilon > 0$ we need to find $\delta > 0$ such that $y - x < \delta$ implies $\exp(y) - \exp(x) < \varepsilon$:

$$\exp(y) - \exp(x) = \exp(x) \cdot (\exp(y - x) - 1) \leq \exp(x) \cdot (\exp(\delta) - 1)$$

We choose $\delta = \log(\varepsilon \cdot \exp(-x) + 1)$.
Similarly, we can also show that $\log\colon \mathbb{R}^+ \to \mathbb{R}$ is a continuous function.

Example 4.2.8. For $\alpha \geq 0$ we can define the power function $x^\alpha := \exp(\alpha \cdot \log(x))$ for $x > 0$. For $\alpha \in \mathbb{N}$ it agrees with the usual power function. By the previous results it follows that x^α is a continuous function.

4.3 Compactness

Definition 4.3.1. Let $(x_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}$ a sequence. Let $(n_k)_{k\in\mathbb{N}} \subseteq \mathbb{N}$ be sequence of natural numbers with $n_{k+1} > n_k$ for all $k \geq 1$. The we call $(x_{n_k})_{k\in\mathbb{N}} \subseteq \mathbb{R}$ a *subsequence* of $(x_n)_{n\in\mathbb{N}}$.

Example 4.3.2. $x_n = (-1)^n$ is a sequence of real numbers. If we choose $n_k = 2k$, then obtain the subsequence $x_{n_k} = x_{2k} = 1$.

Definition 4.3.3. Let $X \subseteq \mathbb{R}$. We call a set $K \subseteq X$ *compact* if every sequence $(x_n)_{n\in\mathbb{N}} \subseteq K$ has a convergent subsequence with limit in K.

Lemma 4.3.4.

1. *The empty set is compact.*
2. *Intersections of compact sets are compact.*
3. *Finite unions of compact sets are compact.*

Remark 4.3.5. Infinite unions of compact sets are in general not compact.

Theorem 4.3.6. *A subset $X \subseteq \mathbb{R}$ is compact if and only if it is bounded and closed.*

Proof. We first show that a compact set has to be bounded and closed: If X would be unbounded, then there would exist a sequence $(x_n) \subseteq X$ with $x_{n+1} \geq x_n + 1$. such a sequence can't have a convergent subsequence. If X would not be closed, then there would be a subsequence $(x_n) \subseteq X$ convergent in $\mathbb{R}$ whose limit is not in X. By compactness, it has a convergent subsequence with limit in X. By uniqueness of limit, these two limits must agree. Next we show that a bounded and closed subset is compact: Since X is bounded, every sequence in X is bounded. We choose either an increasing or decreasing subsequence: Note that $\liminf_{n\to\infty} x_n$ and $\limsup_{n\to\infty} x_n$ are both finite. This bounded and monotonous subsequence has a limit in $\mathbb{R}$. By closedness, the limit lies in X. □

Example 4.3.7.

1. A bounded closed interval $[a, b] \subset \mathbb{R}$ is compact.
2. The set $\{0\} \cup \{\frac{1}{n+1} \ : \ n \in \mathbb{N}\}$ is compact.

Theorem 4.3.8. *For every $n \in \mathbb{N}$, let $K_n \subset \mathbb{R}$ be a compact subset such that $K_n \neq \emptyset$ and $K_{n+1} \subset K_n$ for every $n \in \mathbb{N}$. Then the intersection $\bigcap_{n\in\mathbb{N}} K_n$ contains at least one point $a \in \mathbb{R}$.*

Theorem 4.3.9. *Let $K \subseteq \mathbb{R}$ be a compact. Then every sequence $(x_n)_{n\in\mathbb{N}} \subseteq K$ has an accumulation point in K.*

Lemma 4.3.10. *Let $(x_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}$ be a sequence with an accumulation point $A \in \mathbb{R}$. Then there exists a subsequence $(x_{n_k})_{k\in\mathbb{N}}$ that converges to A.*

Theorem 4.3.11. *Every bounded sequence in $\mathbb{R}$ has an accumulation point.*

Theorem 4.3.12. *Let $K \subseteq X \subseteq \mathbb{R}$ be subsets. If K is compact and $f\colon X \to \mathbb{R}$ a continuous function, then $f(K)$ is also compact.*

Proof. Let $(b_n)_{n\in\mathbb{N}} \subseteq L = f(K)$ be any given sequence. To prove compactness of L we need to show that there exists a convergent subsequence. For every $b_n \in f(K)$ there exists a $a_n \in K$ with $f(a_n) = b_n$. Since K is compact, there exists a convergent subsequence $(a_{n_m})_{m\in\mathbb{N}} \subseteq K$ with limit $\lim_{m\to\infty} a_{n_m} = a \in K$. Continuity of f implies $\lim_{m\to\infty} f(a_{n_m}) = f(a) \in f(K) = L$. This means that $f(a_{n_m}) = b_{n_m}$ is the desired convergent subsequence. □

Corollary 4.3.13. *Let $K \subseteq X \subseteq \mathbb{R}$ be subsets. If K is compact and $f\colon X \to \mathbb{R}$ a continuous function, then there exists a point $a \in K$ such that $f(a) = \sup\{f(x) : x \in K\}$ and there exists a point $b \in K$ such that $f(b) = \inf\{f(x) : x \in K\}$.*

Remark 4.3.14. In this case, we call the supremum also a *maximum*, since the value is attained. Similarly, we call the infimum a *minimum*.

4.4 Connectedness

Definition 4.4.1. Let $X \subseteq \mathbb{R}$. We call X *connected* if the following holds: For any two open subsets $U, V \subseteq \mathbb{R}$ with $U \cap V = \emptyset$ and $(U \cup V) \cap X = X$ we have either $U = \emptyset, V = X$ or $V = \emptyset, U = X$.

Theorem 4.4.2. *An interval $(a, b) \subset \mathbb{R}$ with $a, b \in \mathbb{R}$ is connected.*

Proof. Assume that it is not connected. Then there exist two non-empty disjoint open subsets $X, Y \subset (a, b)$ with $X \cup Y = (a, b)$. Pick $x \in X$ and $y \in Y$. Assume w.l.o.g. that $x < y$. Let

$$z := \sup\{w \in (a, b) \ : \ [x, w) \subseteq X\}$$

Since we take the supremum of a bounded set, we have $z \in \mathbb{R}$ and $z \leq b$. Moreover, we know that X is an open set, hence $x < z$. We know have three possible cases, and each of them easily leads to a contradiction:

1. $z \in X$: Since X is open $\exists \varepsilon > 0 : (z - \varepsilon, z + \varepsilon) \subseteq X$. This contradicts the definition of the supremum.

2. $z \in Y$: Since Y is open $\exists \varepsilon > 0 : (z - \varepsilon, z + \varepsilon) \subseteq Y$. This contradicts the definition of the supremum.

3. $z = b$: Then $y \in X$, but this contradicts $X \cap Y = \emptyset$.

We reached a contraction, and conclude that (a, b) is connected. □

Remark 4.4.3. A rational interval $(a, b) \cap \mathbb{Q}$ is not connected in $\mathbb{Q}$. Choose a real, but not rational number $c \in (a, b)$. Then $(a, c) \cap \mathbb{Q}$ and $(c, b) \cap \mathbb{Q}$ are disjoint open subsets in $\mathbb{Q}$ whose union is $(a, b) \cap \mathbb{Q}$.

Theorem 4.4.4. *Let $X \subseteq \mathbb{R}$. Let $A \subseteq X$ be a connected set. Let $f\colon X \to \mathbb{R}$ be a continuous function. Then $f(A)$ is also connected.*

Proof. Assume to get a contradiction that $f(A)$ is not connected: there exist non-empty open subsets $U, V \subset \mathbb{R}$ such that

$$(U \cap f(A)) \cap (V \cap f(A)) = \emptyset \quad \text{and} \quad (U \cap f(A)) \cup (V \cap f(A)) = f(A)$$

By the characterization of continuity with open sets, we obtain that the pre-images $f^{-1}(U)$ and $f^{-1}(V)$ are non-empty open subsets of X. It is then easy to see that

$$(f^{-1}(U) \cap A) \cap (f^{-1}(V) \cap A) = \emptyset \quad \text{and } (f^{-1}(U) \cap A) \cup (f^{-1}(V) \cap A) = A$$

But this contradicts the connectedness of A. □

Theorem 4.4.5 (Intermediate Value Theorem)**.**
Let $[a, b] \subset \mathbb{R}$ be an interval and $f\colon [a, b] \to \mathbb{R}$ be a continuous function. Then f attains every real number between $f(a)$ and $f(b)$.

Proof. The interval $[a, b] \subset \mathbb{R}$ is connected. A continuous image of an interval is connected. Assume w.l.o.g. $f(a) < f(b)$. Assume to get a contradiction that there exists a point $c \in [f(a), f(b)]$ which is not attained by f. But then we can decompose $[f(a), f(b)] = [f(a), c) \cup (c, f(b)]$ and obtain a contradiction to connectedness. □

Corollary 4.4.6. *Every real, real-valued polynomial of odd degree has at least one real root.*

4.5 Uniform continuity

Definition 4.5.1. Let $X \subseteq \mathbb{R}$. A function $f\colon X \to \mathbb{R}$ is called *uniformly continuous* on X if

$$\forall \varepsilon > 0\, \exists \delta > 0\, \forall x \in X\, \forall y \in Y \ : \ |x - y| < \delta \Longrightarrow |f(x) - f(y)| < \varepsilon$$

The "uniformity" in this definition means that the choice of the δ depends only on the ε, but no longer on the point $x \in X$. It is immediate that a uniformly continuous map is also a continuous map.

Example 4.5.2.

1. $\sqrt{\cdot}\colon [0, \infty) \to [0, \infty)$ is uniformly continuous.
2. $x \mapsto x^2$ is not uniformly continuous on $\mathbb{R}$.

Theorem 4.5.3. *Let $K \subset \mathbb{R}$ be a compact. Then every continuous function $f\colon K \to \mathbb{R}$ is uniformly continuous.*

5 Differentiation and Integration

5.1 Differentiation

Definition 5.1.1. Let $g\colon X \to \mathbb{R}$ be a function and let $x_0 \in \mathbb{R}$. The *limit*

$$\lim_{h \to x_0} g(h) =: A \in \mathbb{R}$$

exists, if for every sequence $(h_n)_{n\in\mathbb{N}} \subseteq X$ it holds that

$$\lim_{n\to\infty} h_n = x_0 \Longrightarrow \lim_{n\to\infty} g(h_n) = A$$

Example 5.1.2. We can now say that a function g is continuous in x_0 if and only if

$$g(x_0) = \lim_{x \to x_0} g(x)$$

Definition 5.1.3. Let $f\colon (a,b) \to \mathbb{R}$ be a function. We call f *differentiable* in $x \in (a,b)$ if the limit

$$f'(x) := \frac{df}{dx}(x) := \lim_{h\to 0} \frac{f(x+h) - f(x)}{h} \in \mathbb{R}$$

exists. In this case we call $f'(x)$ the *derivative* of f in x. If f is differentiable in every $x \in (a,b)$ we call f differentiable on (a,b).

We can interpret the derivative as an "infinitesimal" rate of change and also as the slope of the graph of the function.

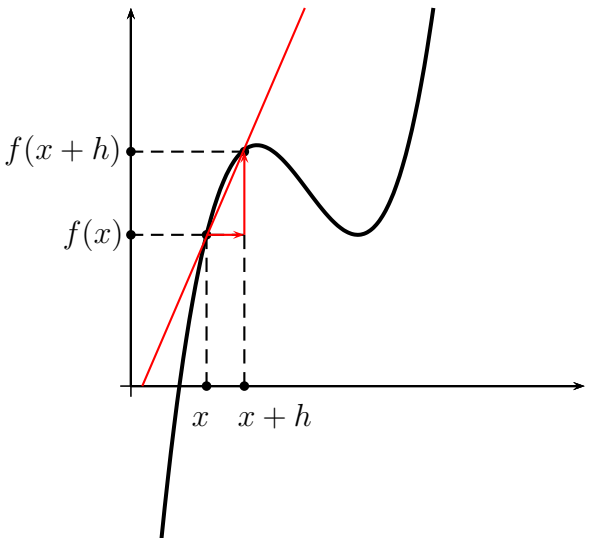


Figure 1: Geometric interpretation of the derivative

Example 5.1.4 (Affine-linear function)**.** Let $f\colon \mathbb{R} \to \mathbb{R}$ be given by $f(x) = ax + b$. We show that f is differentiable in every $x \in \mathbb{R}$ by calculating its derivative

$$f'(x) = \lim_{h\to 0} \frac{f(x+h) - f(x)}{h} = \lim_{h\to 0} \frac{a(x+h) + b - (ax+b))}{h} = \lim_{h\to 0} \frac{ah}{h} = a$$

Example 5.1.5 (A quadratic function)**.** Let $f\colon \mathbb{R} \to \mathbb{R}$ be given by $f(x) = x^2$. We show that f is differentiable in every $x \in \mathbb{R}$ by calculating its derivative

$$f'(x) = \lim_{h\to 0} \frac{f(x+h) - f(x)}{h} = \lim_{h\to 0} \frac{(x+h)^2 - x^2}{h} = \lim_{h\to 0} \frac{2xh + h^2}{h} = 2x$$

Remark 5.1.6. If $f\colon (a,b) \to \mathbb{R}$ is a differentiable function, then the map $x \mapsto f'(x)$ defines another function $f'\colon (a,b) \to \mathbb{R}$.

The derivative can be used to approximate a function by an affine-linear function.

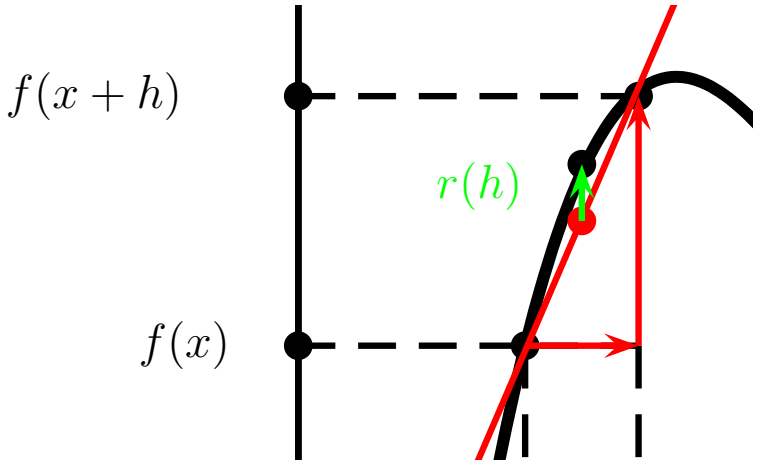


Figure 2: Approximation with an affine-linear function. The difference between the approximation and the original function is given by $r(h)$.

Lemma 5.1.7. *Let $f\colon (a,b) \to \mathbb{R}$ be a function. It is differentiable in $x \in (a,b)$ if and only if $\exists\, A \in \mathbb{R}$ and $r\colon (a,b) \to \mathbb{R}$ such that*

$$f(x+h) = f(x) + h \cdot A + r(h) \text{ and } \lim_{h\to 0} \frac{r(h)}{h} = 0$$

In case of differentiability, $A = f'(x)$.

Proof. $\boxed{\Longrightarrow}$ We know that $\lim_{h\to 0} \frac{f(x+h)-f(x)}{h} = f'(x) =: A$ exists. Define

$$r(h) := f(x+h) - f(x) - h \cdot A$$

Then the above equality holds trivially. We just need to check that

$$\lim_{h\to 0} \frac{r(h)}{h} = \lim_{h\to 0} \frac{f(x+h) - f(x) - h \cdot A}{h} = f'(x) - A = 0$$

$\boxed{\Longleftarrow}$ As before, we have that

$$\frac{f(x+h)-f(x)}{h} = A + \frac{r(h)}{h}$$

Knowing that $\lim_{h\to 0} \frac{r(h)}{h} = 0$, we obtain that $f'(x) = A$ exists. □

Definition 5.1.8. Let $f\colon (a,b) \to \mathbb{R}$ be a function. If f is differentiable, it is also called 1-*times differentiable*. We call f *n-times differentiable* if f' is $(n-1)$-times differentiable. We denote the derivative of f' by f'' etc. The *n-th derivative* of f is denoted by $f^{(n)}$ or $\frac{d^n f}{dx^n}$, and defined to be $\left(f^{(n-1)}\right)'$. If f is n-times differentiable and if and addition $f^{(n)}$ is continuous, we call f *n-times continuously differentiable.*

Remark 5.1.9. For a function $f\colon (a,b) \to \mathbb{R}$ we can't define n-times differentiable in a point $p \in (a,b)$ without assuming the the derivative $f^{(n-1)}$ exists in a neighborhood of p. If we say that f is n-times differentiable in a point $p \in (a,b)$ we implicitly assume that there exists $\varepsilon > 0$ such that f is $(n-1)$-times differentiable on $(p-\varepsilon, p+\varepsilon)$.

Example 5.1.10. $f(x) = x^2$ is differentiable for all $x \in \mathbb{R}$. Its derivative $f'(x) = 2x$ is also differentiable, $f''(x) = 2$, $f'''(x) = 0$, $f''''(x) = 0$ etc. It turns out that f is n-times differentiable for any $n \in \mathbb{N}$. In this case, we call the function also *infinitely many times differentiable* or *smooth.*

Theorem 5.1.11. *A function differentiable in x is continuous in x.*

Proof. We use the characterization

$$f(x+h) = f(x) + h \cdot A + r(h) \text{ and } \lim_{h\to 0} \frac{r(h)}{h} = 0$$

This implies $\lim_{h\to 0} r(h) = 0$, since $0 \le |r(h)| \le \left|\frac{r(h)}{h}\right|$ for $0 < |h| \le 1$. Taking the limit $h \to 0$ we obtain

$$\lim_{h\to 0} f(x+h) = \lim_{h\to 0} \left(f(x) + h \cdot A + r(h)\right) = f(x)$$

which is equivalent to continuity in x. □

Remark 5.1.12. There are continuous functions that are not differentiable, e.g. $x \mapsto |x|$ is not differentiable in $x = 0$. However, one can even construct continuous nowhere differentiable functions. In fact, one can show that "most" of the continuous function are nowhere differentiable.

Example 5.1.13. The function $f\colon \mathbb{R} \to \mathbb{R}$, $f(x) = |x|$, is not differentiable, but continuous in $x = 0$:

- $f(0) = 0$. And $\lim_{h\to 0} |h| = 0$, hence f is continuous in 0.

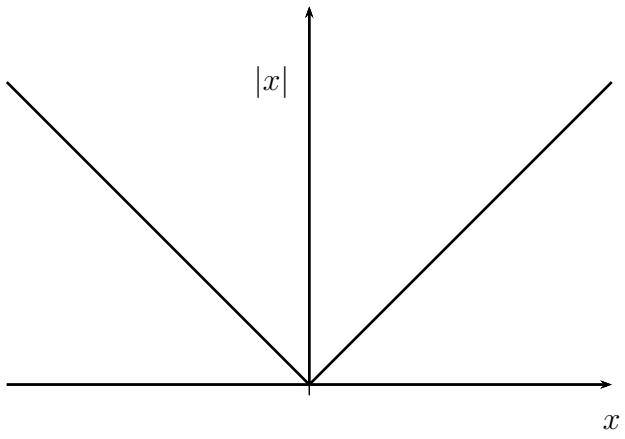


Figure 3: The absolute value is continuous in $x = 0$, but not differentiable in $x = 0$.

- f is not differentiable in 0. Consider the two limits:

$$\begin{aligned} \lim_{h \to 0, h>0} \frac{f(0+h) - f(0)}{h} &= \lim_{h \to 0, h>0} \frac{|h|}{h} &&= 1 \\ \lim_{h \to 0, h<0} \frac{f(0+h) - f(0)}{h} &= \lim_{h \to 0, h<0} \frac{|h|}{h} &&= -1 \end{aligned}$$

Lemma 5.1.14. *Let $f, g \colon (a, b) \to \mathbb{R}$ be functions that are differentiable in $x \in (a, b)$. Then the following hold:*

1. *$f + g$ is differentiable in x, and*

$$(f(x) + g(x))' = f'(x) + g'(x)$$

2. *$f \cdot g$ is differentiable in x, and*

$$(f(x) \cdot g(x))' = f'(x) \cdot g(x) + f(x) \cdot g'(x)$$

3. *f/g is a differentiable in x if $g(x) \neq 0$, and*

$$(f(x)/g(x))' = (f'(x) \cdot g(x) - f(x) \cdot g'(x))/(g^2(x))$$

Proof.

1. $$\begin{aligned}&\lim_{h\to 0}\frac{(f(x+h)+g(x+h))-(f(x)+g(x))}{h}\\ &=\lim_{h\to}\frac{f(x+h)-f(x)}{h}+\lim_{h\to}\frac{g(x+h)-g(x)}{h}\\ &=f'(x)+g'(x)\end{aligned}$$

2. $$\begin{aligned}&\lim_{h\to 0}\frac{f(x+h)\cdot g(x+h)-f(x)\cdot g(x)}{h}\\ &=\lim_{h\to 0}\frac{f(x+h)\cdot g(x+h)-f(x+h)\cdot g(x)+f(x+h)\cdot g(x)-f(x)\cdot g(x)}{h}\\ &=\lim_{h\to 0}\frac{f(x+h)\cdot(g(x+h)-g(x))+(f(x+h)-f(x))\cdot g(x)}{h}\\ &=\lim_{h\to 0}\frac{f(x+h)\cdot(g(x+h)-g(x))+(f(x+h)-f(x))\cdot g(x)}{h}\\ &=\lim_{h\to 0}f(x+h)\cdot\lim_{h\to 0}\frac{g(x+h)-g(x)}{h}+\lim_{h\to 0}\frac{f(x+h)-f(x)}{h}\cdot g(x)\\ &=f(x)\cdot g'(x)+f'(x)\cdot g(x)\end{aligned}$$

   where we used the continuity of f in x for the last equality.

3. Will follow from 2, if we can show it for $1/g$.

$$\begin{aligned}&\lim_{h\to 0}\frac{\frac{1}{g(x+h)}-\frac{1}{g(x)}}{h}=\lim_{h\to 0}\frac{\frac{1}{g(x+h)}-\frac{1}{g(x)}}{h}\cdot\frac{g(x)\cdot g(x+h)}{g(x)\cdot g(x+h)}\\ &=\lim_{h\to 0}\frac{g(x)-g(x+h)}{h}\cdot\lim_{h\to 0}\frac{1}{g(x)\cdot g(x+h)}\\ &=\frac{-g'(x)}{g^2(x)}\end{aligned}$$

   where we used the continuity of g in x for the last equality, as well as the fact that $g(x)\neq 0$ implies $g(x+h)\neq 0$ for h small enough. □

Corollary 5.1.15. *Polynomials are differentiable functions, and*

$$\left(c_0+c_1\cdot x+c_2\cdot x^2+c_3x^3+\cdots+c_n\cdot x^n\right)'=c_1+2c_2\cdot x+3c_3\cdot x^2+\cdots+nc_n\cdot x^{n-1}$$

Proof. By 1 of the preceding lemma, we need to consider the case $c\cdot x^k$ for $k\in\mathbb{N}$. For $k=0,1$ we already calculated the result:

$$\begin{aligned}(c)'&=0\\ (c\cdot x)'&=c\end{aligned}$$

We claim that $(c \cdot x^k)' = kc_k \cdot x^{k-1}$, and prove this by induction in $k \in \mathbb{N}$. Induction step from k to $k+1$:

$$\begin{aligned}\left(c \cdot x^{k+1}\right)' = \left(c \cdot x^k \cdot x\right)' &= \left(c \cdot x^k\right)' \cdot x + \left(c \cdot x^k\right) \cdot x' \\ &= c \cdot k \cdot x^{k-1} \cdot x + c \cdot x^k \cdot 1 \\ &= c \cdot (k+1) \cdot x^k\end{aligned}$$

□

Example 5.1.16. The exponential function is differentiable, and $\exp'(x) = \exp(x)$.

$$\begin{aligned}\exp'(0) &= \lim_{h\to 0} \frac{\exp(0+h) - \exp(0)}{h} = \lim_{h\to 0} \frac{\sum_{n=0}^{\infty} \frac{h^n}{n!} - 1}{h} \\ &= \lim_{h\to 0} \frac{\sum_{n=1}^{\infty} \frac{h^n}{n!}}{h} = \lim_{h\to 0} \sum_{n=1}^{\infty} \frac{h^{n-1}}{n!} = \lim_{h\to 0} \sum_{n=0}^{\infty} \frac{h^n}{(n+1)!} \\ &= \lim_{h\to 0} \left(1 + h \cdot \sum_{n=1}^{\infty} \frac{h^{n-1}}{(n+1)!}\right) = 1\end{aligned}$$

For the last equality, we observe that

$$\sum_{n=1}^{\infty} \left|\frac{h^{n-1}}{(n+1)!}\right| \overset{|h|\leq 1}{\leq} \sum_{n=1}^{\infty} \frac{1}{(n+1)!} = e - 2$$

Hence, the term $h \cdot \sum_{n=1}^{\infty} \frac{h^{n-1}}{(n+1)!}$ is bounded in absolute value by $|h| \cdot (e-2)$ and converges to 0 for $h \to 0$. Now,

$$\exp'(x) = \lim_{h\to 0} \frac{\exp(x+h) - \exp(x)}{h} = \exp(x) \cdot \lim_{h\to 0} \frac{\exp(0+h) - \exp(0)}{h} = \exp(x).$$

Theorem 5.1.17 (Chain rule)**.** *Let $f\colon (a_1, b_1) \to (a_2, b_2)$ and $g\colon (a_2, b_2) \to (a_3, b_3)$ be functions. Assume that f is differentiable in $x \in (a_1, b_1)$ and that g is differentiable in $f(x) \in (a_2, b_2)$. Then the the composition $z \mapsto g \circ f(z) = g(f(z))$ is a map from $(a_1, b_1) \to (a_3, b_3)$ that is differentiable in x and*

$$(g(f(x)))' = g'(f(x)) \cdot f'(x)$$

Proof. We use the characterization of differentiability by approximation with a linear function.

$$f(x+h) = f(x) + h \cdot A + r(h) \text{ and } \lim_{h\to 0} \frac{r(h)}{h} = 0$$

$$g(y+\widetilde{h}) = g(y) + \widetilde{h} \cdot B + s(\widetilde{h}) \text{ and } \lim_{\widetilde{h}\to 0} \frac{s(\widetilde{h})}{\widetilde{h}} = 0$$

where $y = f(x)$.

$$\begin{aligned} g(f(x+h) &= g(f(x) + \underbrace{h \cdot A + r(h)}_{=\widetilde{h}}) = g(f(x) + \widetilde{h}) \\ &= g(f(x)) + \widetilde{h} \cdot B + s(\widetilde{h}) \\ &= g(f(x)) + (h \cdot A + r(h)) \cdot B + s(h \cdot A + r(h)) \\ &= g(f(x)) + h \cdot (A \cdot B) + \underbrace{B \cdot r(h) + s(h \cdot A + r(h))}_{=R(h)} \end{aligned}$$

We know that $A = f'(x)$ and $B = g'(y) = g'(f(x))$. We need to check the behavior of R:

$$\lim_{h \to 0} \frac{R(h)}{h} = \lim_{h \to 0} \frac{B \cdot r(h) + s(h \cdot A + r(h))}{h} = \lim_{h \to 0} \frac{s(h \cdot A + r(h))}{h}$$

$$= \lim_{h \to 0} \frac{s(h \cdot A + r(h))}{h \cdot A + r(h)} \frac{h \cdot A + r(h)}{h} = \lim_{\widetilde{h} \to 0} \frac{s(\widetilde{h})}{\widetilde{h}} \cdot \lim_{h \to 0} \frac{h \cdot A + r(h)}{h} = 0 \cdot A = 0$$

Hence, $g \circ f$ is differentiable in x with derivative $A \cdot B = f'(x) \cdot g'(f(x))$. □

Example 5.1.18.

$$\left(\exp\left(\frac{1}{x^2-1}\right)\right)' = \exp'\left(\frac{1}{x^2-1}\right) \cdot \left(\frac{1}{x^2-1}\right)' = \exp\left(\frac{1}{x^2-1}\right) \cdot \frac{-2x}{(x^2-1)^2}$$

Definition 5.1.19. Let $X \subseteq \mathbb{R}$. A function $f\colon X \to \mathbb{R}$ is called

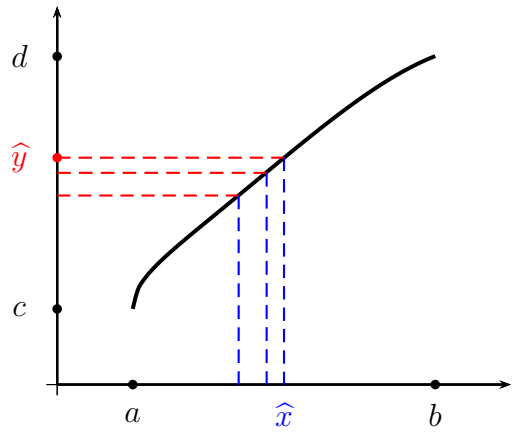
- *monotonously increasing* if $\forall x, y \in X,\ x < y \Longrightarrow f(x) \leq f(y)$
- *strictly monotonously increasing* if $\forall x, y \in X,\ x < y \Longrightarrow f(x) < f(y)$
- *monotonously decreasing* if $\forall x, y \in X,\ x < y \Longrightarrow f(x) \geq f(y)$
- *strictly monotonously decreasing* if $\forall x, y \in X,\ x < y \Longrightarrow f(x) > f(y)$

Lemma 5.1.20. *A bounded sequence $(x_n)_{n \in \mathbb{N}} \subseteq \mathbb{R}$ of real numbers with exactly one accumulation point $A \in \mathbb{R}$ converges to A.*

Corollary 5.1.21. *Let $(x_n)_{n \in \mathbb{N}} \subseteq \mathbb{R}$ be a bounded sequence of real numbers. If there exists $A \in \mathbb{R}$ such that every convergent subsequence of $(x_n)_{n \in \mathbb{N}}$ has limit A, then $(x_n)_{n \in \mathbb{N}}$ itself converges to A.*

Proof. Given any accumulation point A of $(x_n)_{n \in \mathbb{N}}$, we find a subsequence converging to it. But if all convergent subsequences have the same limit, there can be only one accumulation point. Be the preceding lemma, $(x_n)_{n \in \mathbb{N}}$ converges to A. □

Lemma 5.1.22. *A strictly monotonously increasing continuous function $f\colon (a, b) \to \mathbb{R}$ has a continuous inverse.*



Proof. First, note that a strictly monotonously increasing function is injective. Let $Y = f((a,b))$. Let $c = \lim_{x\to a, x>a} f(x)$ and let $d = \lim_{x\to b, x<b} f(x)$. By the intermediate value theorem, f will attain all values in (c,d). Moreover, since it is strictly monotonously increasing, it can't obtain any values outside this interval, hence $Y = (c,d)$, and thus $f\colon (a,b) \to (c,d)$ is bijective. Let $g := f^{-1}\colon (c,d) \to (a,b)$ be its inverse function.
Let $\widehat{y} \in (c,d)$ and let $(y_n)_{n\in\mathbb{N}} \subset (c,d)$ with $\lim_{n\to\infty} y_n = \widehat{y}$. For each y_n we have a unique $x_n = g(y_n) \in (a,b)$ with $f(x_n) = y_n$ and we have a unique $\widehat{x} = g(\widehat{y}) \in (a,b)$ with $f(\widehat{x}) = \widehat{y}$. Note that $x_n \in (a,b)$ is a bounded sequence. For any convergent subsequence, consider its limit $\widehat{z} \in [a,b]$.

1. If in fact $\widehat{z} \in (a,b)$, then, by continuity of f, we obtain $f(\widehat{z}) = \lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} y_n = \widehat{y} = f(\widehat{x})$, implying $\widehat{x} = \widehat{z}$ by injectivity of f. Hence, all convergent subsequences have the same limit $\widehat{x}$, and therefore the sequence $(x_n)_{n\in\mathbb{N}}$ itself is convergent with limit $\widehat{x}$.

2. If $\widehat{z} = a$ or $\widehat{z} = b$, we obtain a contradiction to f being strictly monotonously increasing, since $a < \widehat{x} < b$ with $f(\widehat{x}) = \widehat{y}$.

We now obtain $\lim_{n\to\infty} g(y_n) = \lim_{n\to\infty} x_n = \widehat{x} = g(\widehat{y})$. □

Theorem 5.1.23 (Derivative of the inverse function)**.** *Let $f\colon (a,b) \to (c,d)$ be a strictly increasing and differentiable function with $f'(x) > 0$ for all $x \in (a,b)$. Then the inverse function $f^{-1}\colon (c,d) \to (a,b)$ is differentiable and for all $y \in (a,b)$ we have that*

$$(f^{-1})'(y) = \frac{1}{f'(f^{-1}(y))}$$

Remark 5.1.24. If we would know that the inverse function is differentiable, we could easily

derive the formula above using the chain rule. For $y \in (c,d)$ we have that

$$\begin{aligned} f(f^{-1}(y)) &= y \\ \big(f(f^{-1}(y))\big)' &= f'(f^{-1}(y)) \cdot (f^{-1})'(y) = y' = 1 \end{aligned}$$

This would however be "cheating", since we do not know that f^{-1} is differentiable.

Remark 5.1.25. If the function is strictly decreasing and if $f'(x) < 0$, we can apply the theorem to $-f$ instead, and obtain the same statement. We will later improve this theorem, once we know the Mean Value Theorem.

Proof. Since f is differentiable in x, we have that

$$f(x+h) = f(x) + h \cdot A + r(h) \text{ and } \lim_{h \to 0} \frac{r(h)}{h} = 0$$

where $A = f'(x) \neq 0$. By the preceding lemma, $f\colon (a,b) \to (c,d)$ is a bijection.

$$\begin{aligned} f^{-1}(y + \widetilde{h}) &= f^{-1}(f(x+h)) = x + h = f^{-1}(y) + h \\ f^{-1}(y + \widetilde{h}) &= f^{-1}(f(x+h)) = f^{-1}(f(x) + h \cdot A + r(h)) \end{aligned}$$

Hence, $\widetilde{h} = h \cdot A + r(h)$, and

$$\begin{aligned} f^{-1}(y + \widetilde{h}) &= f^{-1}(y) + h = f^{-1}(y) + (\widetilde{h} - r(h))/A \\ &= f^{-1}(y) + \widetilde{h} \cdot \frac{1}{A} - \frac{r(h)}{A} \end{aligned}$$

If we can prove that $\lim\limits_{\widetilde{h} \to 0} \dfrac{r(h)}{\widetilde{h}} = 0$, then we obtain that $(f^{-1})'(y) = \dfrac{1}{A} = \dfrac{1}{f'(f^{-1}(y))}$. Since f^{-1} is continuous, he have that $\lim\limits_{\widetilde{h} \to 0} h = 0$. Now,

$$\lim_{\widetilde{h} \to 0} \frac{r(h)}{\widetilde{h}} = \lim_{\widetilde{h} \to 0} \frac{r(h)}{h \cdot A + r(h)} = \lim_{h \to 0} \frac{r(h)}{h \cdot A + r(h)} = \lim_{h \to 0} \frac{r(h)}{h \cdot (A + r(h)/h)} = \frac{0}{A + 0} = 0 \quad \square$$

Example 5.1.26.

1. $\big(\sqrt[k]{x}\big)' = \dfrac{1}{k \cdot (\sqrt[k]{x})^{k-1}}$
2. $(\log(x))' = \dfrac{1}{\exp'(\log(x))} = \dfrac{1}{x}$

Lemma 5.1.27. *Let $f\colon (a,b) \to \mathbb{R}$ be a function that a attains a local maximum or a local minimum in $x_0 \in (a,b)$. Assume that f is differentiable in x_0. Then $f'(x_0) = 0$.*

Proof. W.l.o.g. we can assume that f attains a maximum in x_0: if it is only a local maximum, then shrink (a, b). For a (local) minimum, replace f by $-f$. By the definition of a maximum we have that

$$\forall x \in (a, b)\ f(x) \leq f(x_0)$$

This implies

$$\frac{f(x_0 + h) - f(x_0)}{|h|} \leq \frac{0}{|h|} = 0$$

For $h > 0$ we obtain $\frac{f(x_0+h)-f(x_0)}{h} \leq 0$. And for $h < 0$ we obtain $\frac{f(x_0+h)-f(x_0)}{h} \geq 0$. Hence, since the limit $\lim_{h\to 0} \frac{f(x_0+h)-f(x_0)}{h} = f'(x)$ exists, it needs to be zero. □

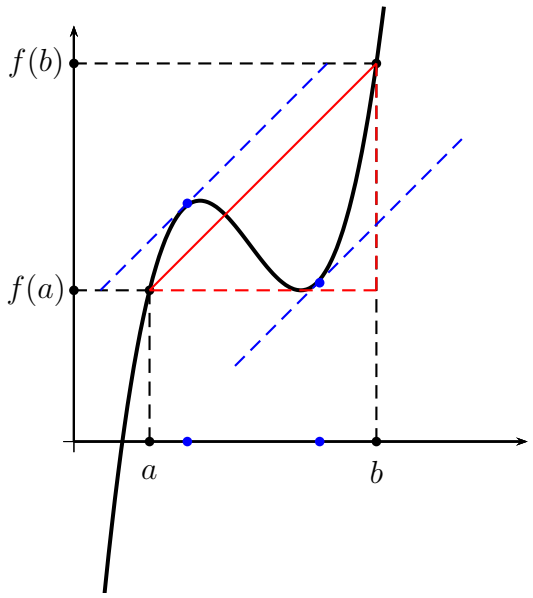
Definition 5.1.28. Let $f\colon (a, b) \to \mathbb{R}$ be a function. Assume that f is differentiable in $x_0 \in (a, b)$. If $f'(x_0) = 0$, then we call x_0 a *critical point* of f.

Remark 5.1.29. In other words, this lemma states that local minima/maxima are critical points. The converse is not true, e.g. $f\colon \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^3$ has a critical point in $x = 0$, but neither a local minimum or a local maximum in $x = 0$.

5.2 Mean Value Theorem

Theorem 5.2.1 (Mean Value Theorem (MVT))**.** *Let $f\colon [a, b] \to \mathbb{R}$ with $a < b$ be a continuous function that is differentiable on (a, b). Then there exists $c \in (a, b)$ such that*

$$f'(c) \cdot (b - a) = f(b) - f(a)$$



This Mean Value Theorem follows from a more general one which is due to Cauchy:

Theorem 5.2.2 (Cauchy's Mean Value Theorem)**.** *Let $f, g\colon [a,b] \to \mathbb{R}$ with $a < b$ be continuous functions that are differentiable on (a,b). Then there exists $c \in (a,b)$ such that*

$$f'(c) \cdot (g(b) - g(a)) = g'(c) \cdot (f(b) - f(a))$$

Proof. Define $h\colon [a,b] \to \mathbb{R}$ by

$$h(x) := (f(x) - f(a)) \cdot (g(b) - g(a)) - (f(b) - f(a)) \cdot (g(x) - g(a))$$

Observe that h is again continuous on $[a,b]$ and differentiable on (a,b). We have that $h(a) = 0 = h(b)$:

$$\begin{aligned} h(a) &= \overbrace{(f(a) - f(a))}^{=0} \cdot (g(b) - g(a)) - (f(b) - f(a)) \cdot \overbrace{(g(a) - g(a))}^{=0} = 0 \\ h(b) &= \underbrace{(f(b) - f(a)) \cdot (g(b) - g(a))}_{} - \underbrace{(f(b) - f(a)) \cdot (g(b) - g(a))}_{} = 0 \end{aligned}$$

For $x \in (a,b)$ we obtain

$$h'(x) = f'(x) \cdot (g(b) - g(a)) - (f(b) - f(a)) \cdot g'(x)$$

If h is constant, then we have $f'(x) \cdot (g(b) - g(a)) = (f(b) - f(a)) \cdot g'(x)$ for all $x \in (a,b)$ and $c \in (a,b)$ can be chosen arbitrarily. Since h is continuous, it will attain a minimum and a maximum on the interval $[a,b]$. If h is not constant, at least the minimum or the maximum will be attained in a point $c \in (a,b)$, since $h(a) = 0 = h(b)$. We then have $h'(c) = 0$ which implies the statement of the theorem. □

Remark 5.2.3. It is important to note that we do not require differentiability in the boundary points. The Mean Value Theorem (as an equality) does not hold for vector-valued functions. Generalizations in the form of an inequality are possible.

Corollary 5.2.4. *Let $f\colon [a,b] \to \mathbb{R}$ be a continuous function that is differentiable on (a,b).*

1. *If $\forall x \in (a,b)\ f'(x) > 0$ then f is strictly monotonously increasing.*
2. *If $\forall x \in (a,b)\ f'(x) \geq 0$ then f is monotonously increasing.*
3. *If $\forall x \in (a,b)\ f'(x) < 0$ then f is strictly monotonously decreasing.*
4. *If $\forall x \in (a,b)\ f'(x) \leq 0$ then f is monotonously decreasing.*

Proof. Assume that $\forall x \in (a,b)\ f'(x) > 0$. We want to show that f is strictly monotonously increasing. Let $z, w \in (a,b)$ with $z < w$. Then f is certainly continuous on $[z,w] \subset (a,b)$ and differentiable on (z,w). Hence we can apply the mean value theorem and obtain that $\exists x \in (z,w)$ such that

$$f'(x) \cdot (w - z) = f(w) - f(z)$$

Since $z < w$ and $f'(x) > 0$, the l.h.s. is strictly positive. This implies $f(w) > f(z)$.
The other cases are proved similarly. □

Remark 5.2.5. The converse does in general not hold, see again the example of $x \mapsto x^3$ for a strictly monotonously increasing function.

Corollary 5.2.6. *Let $f\colon [a,b] \to \mathbb{R}$ be a continuous function that is differentiable on (a,b). Then f is constant if and only if $f'(x) = 0$ for every $x \in (a,b)$.*

Proof. This follows from the previous corollary, since $f'(x) \leq 0$ and $f'(x) \geq 0$ imply that f is both monotonously increasing and monotonously decreasing. □

Corollary 5.2.7. *Let $f\colon (a,b) \to \mathbb{R}$ be a differentiable function. Let $x_0 \in (a,b)$.*

1. *If $f'(x) > 0$ for every $x \in (a, x_0)$ and if $f'(x) < 0$ for every $x \in (x_0, b)$, then f has a strict local maximum in x_0.*
2. *If $f'(x) \geq 0$ for every $x \in (a, x_0)$ and if $f'(x) \leq 0$ for every $x \in (x_0, b)$, then f has a local maximum in x_0.*
3. *If $f'(x) < 0$ for every $x \in (a, x_0)$ and if $f'(x) > 0$ for every $x \in (x_0, b)$, then f has a strict local minimum in x_0.*
4. *If $f'(x) \leq 0$ for every $x \in (a, x_0)$ and if $f'(x) \geq 0$ for every $x \in (x_0, b)$, then f has a local minimum in x_0.*

Proof. Assume that $f'(x) > 0$ for every $x \in (a, x_0)$, and assume that $f'(x) < 0$ for every $x \in (x_0, b)$. Then f is strictly monotonously increasing on (a, x_0) and strictly monotonously decreasing on (x_0, b). Since f is continuous in x_0, we have that $f(x) < f(x_0)$ for all $x \in (a, x_0)$, and $f(x) < f(x_0)$ for all $x \in (x_0, b)$. Thus, f has a strict local maximum in x_0. The other cases are proved similarly. □

Corollary 5.2.8. *Let $f\colon (a,b) \to \mathbb{R}$ be a differentiable function that is twice differentiable in $x_0 \in (a,b)$.*

1. *$f'(x_0) = 0$ and $f''(x_0) < 0 \implies f$ has a strict local maximum in x_0.*
2. *$f'(x_0) = 0$ and $f''(x_0) > 0 \implies f$ has a strict local minimum in x_0.*

Proof. Assume that $f'(x_0) = 0$ and $f''(x_0) < 0$. Recall the definition of the second derivative:

$$f''(x_0) = \lim_{h \to 0} \frac{f'(x_0 + h) - f'(x_0)}{h}$$

For $|h|$ small enough, we obtain that $\dfrac{f'(x_0 + h) - f'(x_0)}{h} < 0$. For $h > 0$ this implies $f'(x_0 + h) < f'(x_0) = 0$. For $h < 0$ this implies $f'(x_0 + h) > f'(x_0) = 0$. Hence, f is strictly monotonously increasing on some interval $(x_0 - \alpha, x_0)$, and strictly monotonously decreasing on some interval $(x_0, x_0 + \beta)$. Thus, f has a strict local maximum in x_0.
The other case is proved similarly. □

Remark 5.2.9. If $f''(x_0) = 0$ we can't make a statement. It is possible to refine the theorem in such a case if the function is three, four or more times differentiable (using Taylor approximation, see later).

The Mean Value Theorem can also be used for calculating limits:

Example 5.2.10.

$$\lim_{n\to\infty} \left(\sqrt[3]{n+1} - \sqrt[3]{n}\right) = 0$$

For each $n \in \mathbb{N}$, apply the MVT to $f\colon [n, n+1] \to \mathbb{R}$ given by $f(x) = \sqrt[3]{x}$. Then there exists $c \in [n, n+1]$ such that $0 \leq \sqrt[3]{n+1} - \sqrt[3]{n} = \dfrac{1}{3c^{2/3}} \leq \dfrac{1}{3n^{2/3}}$ which converges to 0 for $n \to \infty$.

Theorem 5.2.11 (Rule of Bernoulli–de l'Hôpital)**.** *Let $f, g\colon (a, b) \to \mathbb{R}$ be differentiable functions with $\lim_{x\to b} f(x) = 0$ and $\lim_{x\to b} g(x) = 0$. Assume that $g'(x) \neq 0$ for all $x \in (a, b)$ and that the limit*

$$\lim_{x\to b} \frac{f'(x)}{g'(x)}$$

exists. Then

$$\lim_{x\to b} \frac{f(x)}{g(x)} = \lim_{x\to b} \frac{f'(x)}{g'(x)}$$

Proof. Since f is continuous with $\lim_{x\to b} f(x) = 0$ we can set $f(b) = 0$ and obtain a continuous function f on $(a, b]$. Similarly, we can set $g(b) = 0$. We now apply the Cauchy mean value theorem. For each $x \in (a, b)$ we find $y \in (a, b)$ such that

$$\frac{f'(y)}{g'(y)} = \frac{f(b) - f(x)}{g(b) - g(x)} = \frac{0 - f(x)}{0 - g(x)} = \frac{f(x)}{g(x)}$$

We now take the limit $x \to a$. Since $y \in (x, a)$, this also forces $y \to a$, and we obtain

$$\lim_{y\to b} \frac{f'(y)}{g'(y)} = \lim_{x\to b} \frac{f(x)}{g(x)}$$ □

Remark 5.2.12. It easily follows that the conclusion of the theorem also holds

1. If $\lim_{x\to b} f(x) = \pm\infty$ and $\lim_{x\to b} g(x) = \pm\infty$.
2. If the upper boundary b is replaced by the lower boundary a.
3. If the boundaries a, b resp. are replaced by $-\infty$, $+\infty$ resp.

Example 5.2.13.

$$\lim_{x\downarrow 0} x \cdot \log(x) = -\lim_{x\downarrow 0} \frac{-\log(x)}{1/x} = -\lim_{x\downarrow 0} \frac{-1/x}{-1/x^2} = -\lim_{x\downarrow 0} x = 0$$

We first arrange the term as a fraction where both numerator and denominator converge to $+\infty$, and then we apply the rule of Bernoulli–de-l'Hôpital.

5.3 Convexity

Definition 5.3.1. A function $f\colon (a,b) \to \mathbb{R}$ is called *convex* if

$$\forall x, y \in (a,b) \text{ and } \forall t \in [0,1]\ f(tx + (1-t)y) \leq tf(x) + (1-t)f(y)$$

Remark 5.3.2. The region above the graph of a convex function is a convex region in $\mathbb{R}^2$. In fact, this can be used as an alternative characterization of the convexity of a function.

Theorem 5.3.3 (Jensen's inequality)**.** *Let $f\colon (a,b) \to \mathbb{R}$ be a convex function. Let $x_1, \ldots, x_m \in (a,b)$ and let $\lambda_1, \ldots, \lambda_m \in [0,1]$ s.t. $\sum_{k=1}^{m} \lambda_k = 1$. Then the following inequality holds:*

$$f\left(\sum_{k=1}^{m} \lambda_k \cdot c_k\right) \leq \sum_{k=1}^{m} f(\lambda_k \cdot c_k)$$

Theorem 5.3.4. *A convex function is continuous.*

Remark 5.3.5. In fact, stronger statements hold: the function is differentiable "almost everywhere" (which first has to be made precise).

Theorem 5.3.6. *Let $f\colon (a,b) \to \mathbb{R}$ a twice differentiable function. The function f is convex if and only if $f''(x) \geq 0$ for all $x \in (a,b)$.*

5.4 Taylor Series

Theorem 5.4.1 (Taylor's Theorem)**.** *Let $f\colon (a,b) \to \mathbb{R}$ be a n-times differentiable function in $p \in (a,b)$. Then there exists a function $R\colon (a,b) \to \mathbb{R}$ called* remainder *such that*

$$f(x) = f(p) + f'(p)\cdot(x-p) + \frac{f''(p)}{2}\cdot(x-p)^2 + \cdots + \frac{f^{(n)}(p)}{n!}\cdot(x-p)^n + R_n(x)$$
$$\lim_{p\to x}\frac{R(x)}{(x-p)^n} = 0$$

The polynomial $f(p) + f'(p)\cdot(x-p) + \frac{f''(p)}{2}\cdot(x-p)^2 + \cdots + \frac{f^{(n)}(p)}{n!}\cdot(x-p)^n$ is called the *n-th order Taylor polynomial of f around p.*

Remark 5.4.2. We denote the Taylor polynomial of degree n by t_n:

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

We notice that t_n and f have the same derivatives in the point p up to order n:

$$\begin{aligned}
\left(\frac{d}{dx}\right)^\ell t_n(x) &= \sum_{k=0}^{n}\frac{f^{(k)}(p)}{k!}\left(\frac{d}{dx}\right)^\ell (x-p)^k \\
&= \sum_{k=0}^{n}\frac{f^{(k)}(p)}{k!}\begin{cases} k\cdot(k-1)\cdots(k-\ell+1)\cdot(x-p)^{k-\ell}, & k \geq \ell \\ 0, & k < \ell \end{cases} \\
\left(\frac{d}{dx}\right)^\ell t_n(x)\bigg|_{x=p} &= \sum_{k=0}^{n}\frac{f^{(k)}(p)}{k!}\begin{cases} k\cdot(k-1)\cdots 1, & k = \ell \\ 0, & k \neq \ell \end{cases} \\
&= f^{(\ell)}(p)
\end{aligned}$$

Definition 5.4.3. If a function $f\colon (a,b) \to \mathbb{R}$ is infinitely many times differentiable, we can write down its *Taylor series* around $p \in (a,b)$.

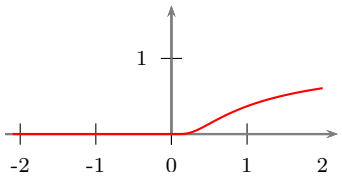
$$t_\infty(x) = \sum_{k=0}^{\infty}\frac{f^{(k)}(p)}{k!}(x-p)^k$$

Remark 5.4.4. Except for $x = p$ we do not know whether this Taylor series converges. And if it converges, we do not know whether it converges to the function f.

Example 5.4.5. The function $f\colon \mathbb{R} \to \mathbb{R}$ given by

$$f(x) = \begin{cases} 0, & x \leq 0 \\ \exp(\frac{-1}{x}), & x > 0 \end{cases}$$

is infinitely many times differentiable in 0 with $f^{(n)}(0) = 0$ for all $n \in \mathbb{N}$. Hence, its Taylor series will be zero which is obviously convergent, but does not converge to f.



Theorem 5.4.6 (Lagrange remainder theorem)**.** *Let $p, x \in (a, b) \subset \mathbb{R}$. If $f\colon (a, b) \to \mathbb{R}$ is $(n+1)$-times differentiable on (p, x) resp. (x, p) and $f^{(n)}$ continuous on $[x, p]$ resp. $[p, x]$, then following holds for the remainder R_n in Taylor's theorem: There exists a $\zeta \in (x, p)$ resp. (p, x) such that*

$$R_n(x) = \frac{f^{(n+1)}(\zeta)}{(n+1)!} \cdot (x-p)^{n+1}$$

Proof. For simplicity, assume that $p < x$. The other case can be treated in the same way.

1. We assume first that $f^{(k)}(p) = 0$ for $k = 0, 1, \ldots, n$, i.e. $t_n = 0$. Hence we need to show that there exists $\zeta \in (p, x)$ such that

   $$f(x) = \frac{(x-p)^{n+1}}{(n+1)!} f^{(n+1)}(\zeta)$$

   We prove the statement by induction in $n \in \mathbb{N}$. For $n = 0$ we obtain the existence of $\zeta \in (p, x)$ from the Mean Value Theorem:

   $$f(x) = (x-p) \cdot f'(\zeta)$$

   Induction step from $n-1$ to n: We define $g(y) := (y-p)^{n+1}$. This function is differentiable, and for every $y \in (p, x)$ we have $g'(y) = (n+1) \cdot (y-p)^n \neq 0$. We now apply Cauchy's Mean Value Theorem to f and g. There exists $\xi \in (p, x)$ such that

   $$\frac{f(x)}{g(x)} = \frac{f(x) - f(p)}{g(x) - g(p)} = \frac{f'(\xi)}{g'(\xi)} = \frac{f'(\xi)}{(n+1) \cdot (\xi - p)^n}$$

   We apply the induction hypothesis to f' on the interval (p, ξ). There exists $\zeta \in (p, \xi)$ such that

   $$f'(\xi) = \frac{(\xi-p)^n}{n!} (f')^{(n)}(\zeta) = \frac{(\xi-p)^n}{n!} f^{(n+1)}(\zeta)$$

   Combining these equations, we obtain

   $$\begin{aligned} f(x) &= g(x) \cdot \frac{f'(\xi)}{(n+1) \cdot (\xi-p)^n} \\ &= g(x) \cdot \frac{(\xi-p)^n}{n!} f^{(n+1)}(\zeta) \cdot \frac{1}{(n+1) \cdot (\xi-p)^n} \\ &= g(x) \cdot \frac{f^{(n+1)}(\zeta)}{(n+1)!} = (x-p)^{n+1} \cdot \frac{f^{(n+1)}(\zeta)}{(n+1)!} \end{aligned}$$

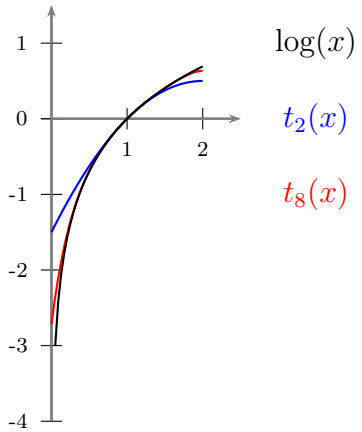
2. General case. Consider $h(y) = f(y) - t_n(y)$. Then h satisfies the assumptions of the previous case, and we obtain that there exists $\zeta \in (p, x)$ such that

$$f(x) - t_n(x) = h(x) = \frac{(x-p)^{n+1}}{(n+1)!} h^{(n+1)}(\zeta) = \frac{(x-p)^{n+1}}{(n+1)!} f^{(n+1)}(\zeta)$$

since $t_n^{(n+1)} = 0$. Adding t_n, we obtain the desired result. □

Example 5.4.7. Let $F(x) = \log(x)$. Its derivative is $F'(x) = f(x) = \frac{1}{x}$. The derivatives of $\frac{1}{x}$ can be calculated easily

$$\begin{aligned}
f(x) &= \frac{1}{x} \\
f'(x) &= \frac{-1}{x^2} \\
f''(x) &= \frac{(-1)\cdot(-2)}{x^3} \\
f'''(x) &= \frac{(-1)\cdot(-2)\cdot(-3)}{x^4} \\
f^{(n)}(x) &= \frac{(-1)^n \cdot n!}{x^{n+1}} \\
F^{(n)}(x) &= \frac{(-1)^{n-1}\cdot(n-1)!}{x^n}
\end{aligned}$$



and we obtain the following Taylor series around the point 1:

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

It is easy to see that the radius of convergence is 1. Hence we do not know whether the series converges for $x = 2$, and if it converges, it is not a priori clear whether it converges to $\log(2)$. We use the Lagrange remainder theorem:

$$\begin{aligned}
F(x) &= \sum_{k=0}^{n} \frac{F^{(k)}(p)}{k!} \cdot (x-p)^k && + \frac{F^{(n+1)}(\zeta)}{(n+1)!}(x-p)^{n+1} \\
\log(x) &= \sum_{k=1}^{n} \frac{(-1)^{k-1}}{k} \cdot (x-1)^k && + \frac{(-1)^{n+1} n!}{\zeta^{n+1}} \cdot \frac{(x-1)^{n+1}}{(n+1)!}
\end{aligned}$$

for some $\zeta \in (p, x)$.

For $x \in (0, 2]$ we can estimate the remainder by $\left|\frac{x-1}{\zeta}\right|^{n+1} \cdot \frac{1}{n+1}$. For $x \in [1, 2]$, we have $\left|\frac{x-1}{\zeta}\right| \leq |x-1| \leq 1$. For $x \in (0, 1)$, we have $0 < x < \zeta < 1$ and thus $\left|\frac{x-1}{\zeta}\right| < \left|\frac{x-1}{x}\right| \leq$

$1 \iff 1 - x \leq x$ which holds for $x \in [\frac{1}{2}, 1]$. We obtain for $x \in [\frac{1}{2}, 2]$ that

$$\left| \log(x) - \sum_{k=1}^{n} \frac{(-1)^{k-1}}{k} \cdot (x-1)^k \right| \leq \frac{1}{n+1} \to 0$$

and have thus also found $\log(2)$ as the value of the alternating harmonic series. In fact, this series also converges to $\log(x)$ for $x \in (0, \frac{1}{2})$, but this is not immediate from the form of the Lagrange remainder.

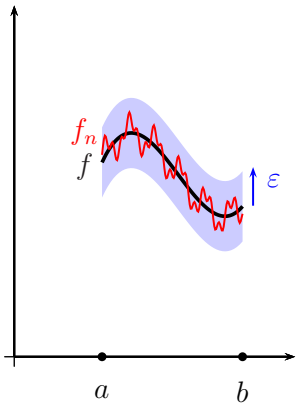
5.5 Uniform convergence

Definition 5.5.1. Let $X \subseteq \mathbb{R}$. For every $n \in \mathbb{N}$ let $f_n \colon X \to \mathbb{R}$ be a function. We say that the sequence of functions $(f_n)_{n\in\mathbb{N}}$ *converges uniformly* to a function $f \colon X \to \mathbb{R}$ if

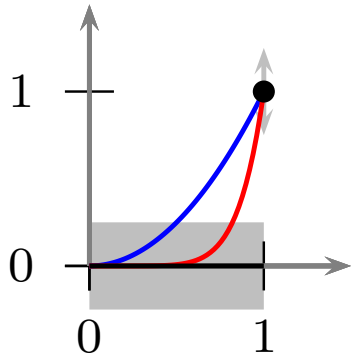
$$\forall \varepsilon > 0 \, \exists N \in \mathbb{N} \, \forall x \in X \; : \; n \geq N \Longrightarrow |f_n(x) - f(x)| < \varepsilon$$

Remark 5.5.2. In contrast, if $f_n(x)$ would converge to $f(x)$ for each point $x \in X$, which is called *point-wise convergence*, we would have:

$$\forall \varepsilon > 0 \, \forall x \in X \, \exists N \in \mathbb{N} \; : \; n \geq N \Longrightarrow |f_n(x) - f(x)| < \varepsilon$$



(a) Given $\varepsilon > 0$, then the graph of f_n (in red) for $n \geq N$ moves only in the shaded area around the limit function f.



(b) The sequence x^n does not converge uniformly on $[0, 1]$. The graph of x^2 is given in blue, the graph of x^7 is given in red.

Figure 4: Uniform convergence

Example 5.5.3. To illustrate what problems can occur, if we have point-wise convergence, i.e. $\forall x \in X \; : \; \lim_{n\to\infty} f_n(x) = f(x)$, but without uniform convergence, consider the following example:

$$f_n \colon [0, 1] \to \mathbb{R}, \; f_n(x) = x^n$$

We easily calculate the limit

$$f(x) = \begin{cases} 0, & x \in [0, 1) \\ 1, & x = 1 \end{cases}$$

and notice that f is not continuous, although all the f_n were continuous functions.

To formally prove that the convergence is not uniform, not even on $[0,1)$, assume to get a contradiction that it is uniform, i.e.

$$\forall \varepsilon > 0 \, \exists N \in \mathbb{N} \, \forall x \in [0,1) \, : \, n \geq N \Longrightarrow |x^n| < \varepsilon$$

Consider the point $x = 1 - \frac{1}{n}$. We obtain $\lim_{n\to\infty} \left(1 - \frac{1}{n}\right)^n = e^{-1}$. Thus, for any $0 < \varepsilon < e^{-1}$, we reach a contradiction.

Theorem 5.5.4. *The limit of a sequence of continuous functions that converge uniformly, is a continuous function.*

Proof. Let $f_n \colon X \to \mathbb{R}$ be a sequence of functions that converges uniformly to a limit function $f \colon X \to \mathbb{R}$, i.e.

$$\forall \epsilon > 0 \exists N \in \mathbb{N} \, \forall x \in X \, : \, n \geq N \Longrightarrow |f_n(x) - f(x)| < \varepsilon$$

Moreover, each of these functions is continuous, i.e.

$$\forall n \in \mathbb{N} \forall x \in X \forall \epsilon > 0 \exists \delta > 0 \forall y \in X \, : \, |x - y| < \delta \Longrightarrow |f_n(x) - f_n(y)| < \varepsilon$$

We need to estimate $|f(x) - f(y)|$. Give $\varepsilon > 0$, choose N such that $|f_n(y) - f(y)| < \varepsilon/3$ for all $y \in X$:

$$\begin{aligned} |f(x) - f(y)| &\leq |f(x) - f_n(x)| + |f_n(x) - f_n(y)| + |f_n(y) - f(y)| \\ &\leq \varepsilon/3 + |f_n(x) - f_n(y)| + \varepsilon/3 \end{aligned}$$

This holds for all $n \geq N$. We choose one such n and use the continuity of f_n in x to obtain there exists a $\delta > 0$ such that $d(x,y) < \delta$ implies $|f_n(x) - f_n(y)| < \varepsilon/3$, completing the estimate for $|f(x) - f(y)|$. □

Lemma 5.5.5. *Let $f_n \colon (a,b) \to \mathbb{R}$ be a uniform Cauchy sequence of functions. Then there exist a limit function $f \colon (a,b) \to \mathbb{R}$, and the functions f_n converge uniformly to f.*

Proof. Since $f_n \colon (a,b) \to \mathbb{R}$ is a uniform Cauchy sequence:

$$\forall \varepsilon > 0 \exists N \in \mathbb{N} \forall x \in (a,b) \, : \, m,n \geq N \Longrightarrow |f_n(x) - f_m(x)| < \varepsilon/2$$

In particular, for every $x \in (a,b)$, we have that $(f_n(x))_{n\in\mathbb{N}} \subset \mathbb{R}$ is a Cauchy sequence. Since $\mathbb{R}$ is complete, this sequence converges to a $f(x) \in \mathbb{R}$. This defines the limit function $f \colon (a,b) \to \mathbb{R}$. We rephrase the uniform Cauchy condition as follows:

$$\forall \varepsilon > 0 \exists N \in \mathbb{N} \forall x \in (a,b) \, : \, m,n \geq N \Longrightarrow f_n(x) - \varepsilon < f_m(x) < f_n(x) + \varepsilon/2$$

Taking the (point-wise) limit $m \to \infty$ yields

$$\forall \varepsilon > 0 \exists N \in \mathbb{N} \forall x \in (a,b) \, : \, n \geq N \Longrightarrow f_n(x) - \varepsilon \leq f(x) \leq f_n(x) + \varepsilon/2$$

which is equivalent to uniform convergence of f_n to f. □

Theorem 5.5.6. *Let $f_n\colon (a,b) \to \mathbb{R}$ be a sequence of differentiable functions. Assume that there exists at least one point $x_0 \in (a,b)$ where $\lim f_n(x_0) = A \in \mathbb{R}$. Assume that the sequence of derivatives f_n' converges uniformly on (a,b) to a function $g\colon (a,b) \to \mathbb{R}$. Then the sequence f_n converges uniformly to a function f that is differentiable with $f' = g$.*

Proof. Given $\varepsilon > 0$ choose $N \in \mathbb{N}$ such that $m, n \geq N$ implies

$$\forall x \in (a,b) \ : \ |(f_n'(x) - f_m'(x))| < \min\left(\frac{\varepsilon}{2(b-a)}, \frac{\varepsilon}{3}\right)$$

We use the Mean Value Theorem for each $f_n - f_m$ to find $\zeta \in (x, x_0)$ resp. $\zeta \in (x_0, x)$ such that

$$\begin{aligned} |(f_n(x) - f_m(x)) - (f_n(x_0) - f_m(x_0))| &= |f_n'(\zeta) - f_m'(\zeta)| \cdot |x - x_0| \\ &\leq \min\left(\frac{\varepsilon}{2(b-a)}, \frac{\varepsilon}{3}\right) \cdot |x - x_0| \end{aligned} \tag{$\star$}$$

Note that the right hand side does not depend on ζ anymore. Choose N even larger, if necessary, such that for $m, n \geq N$ we have

$$|f_n(x_0) - f_m(x_0))| < \frac{\varepsilon}{2}$$

We obtain that the f_n form a uniform Cauchy sequence:

$$\begin{aligned} |(f_n(x) - f_m(x))| &\leq |(f_n(x) - f_m(x)) - (f_n(x_0) - f_m(x_0))| + |(f_n(x_0) - f_m(x_0))| \\ &\leq \frac{\varepsilon}{2(b-a)} \cdot (b-a) + \varepsilon/2 = \varepsilon \end{aligned}$$

Hence, the f_n converge to some function $f\colon (a,b) \to \mathbb{R}$; in fact, by the preceding lemma, the convergence is uniform. We need to calculate the derivative of f and compare it to g. Let $x, x+h \in (a,b)$.

$$\left|\frac{f(x+h) - f(x)}{h} - g(x)\right| \leq$$
$$\underbrace{\left|\frac{f(x+h) - f(x)}{h} - \frac{f_n(x+h) - f_n(x)}{h}\right|}_{(1)} + \underbrace{\left|\frac{f_n(x+h) - f_n(x)}{h} - f_n'(x)\right|}_{(2)} + \underbrace{|f_n'(x) - g(x)|}_{(3)}$$

Choose N even larger, if necessary, such that for $n \geq N$ we have $|f_n'(x) - g(x)| < \varepsilon/3$. This takes care of (3).
We go back to the inequality $(\star)$ which now holds for all $x, x_0 \in (a,b)$, in particular for $x_0 = x + h$, i.e.

$$|(f_n(x) - f_m(x)) - (f_n(x+h) - f_m(x+h))| \leq \frac{\varepsilon}{3} \cdot |h|$$

which is equivalent to

$$\left| \frac{f_n(x) - f_n(x+h)}{h} - \frac{f_m(x) - f_m(x+h)}{h} \right| \leq \frac{\varepsilon}{3}$$

Taking the limit $m \to \infty$, we obtain the estimate for (1).
For fixed $n \in \mathbb{N}$, for small enough $|h| > 0$, the term (2) is smaller than $\varepsilon/3$ by the definition of derivative. □

Lemma 5.5.7. *Let*

$$f(x) = \sum_{n=0}^{\infty} c_n \cdot (x-p)^n$$

be a power series around point $p \in \mathbb{R}$ with strictly positive radius of convergence of $\rho > 0$. Then the partial sums

$$S_m(x) := \sum_{n=0}^{m} c_n \cdot (x-p)^n$$

and their derivatives

$$S'_m(x) := \sum_{n=1}^{m} c_n \cdot n \cdot (x-p)^{n-1}$$

converge uniformly on the the set $\{x \in \mathbb{N} : |x-p| < \sigma\}$ for every $0 < \sigma < \rho$.

Proof. Let $x_0 \in \mathbb{R}$ such that $|x_0 - p| < \rho$. We will revisit the proof about the convergence of the power series within its radius of convergence. Set

$$M := \sup\{|c_n| \cdot |x_0 - p|^n \, : \, n \in \mathbb{N}\}$$

This supremum is finite, since convergence implies that the summands tend to zero. For any $\sigma \in \mathbb{R}$ with $0 < \sigma < |x_0 - p|$ we now obtain

$$|c_n \cdot (x-p)^n| < |c_n| \cdot \sigma^n = \frac{|c_n| \cdot |x_0 - p|^n}{|x_0 - p|^n} \cdot \sigma^n \leq M \cdot \left(\frac{\sigma}{|x_0 - p|} \right)^n$$

where the series with summands as in the r.h.s. converges since it is a geometric series.
Since the radius of convergence ρ is defined as the supremum of $|x_0 - p|$ for all such x_0 where we have convergence, we obtain convergence for all $\sigma < \rho$.
Observe that $M \cdot \underbrace{\left(\frac{\sigma}{|x_0 - p|} \right)^n}_{=q} = Mq^n$ does not depend on x. We now obtain a uniform Cauchy sequence for $x \in [x_0 - \sigma, x_0 + \sigma]$:

$$\begin{aligned} |S_m(x) - S_n(x)| &= \left| \sum_{k=n+1}^{m} c_k \cdot (x-p)^k \right| \leq \sum_{k=n+1}^{m} Mq^k = M \frac{1-q^{m+1}}{1-q} - M \frac{1-q^{n+1}}{1-q} \\ &= M \frac{q^{n+1} - q^{m+1}}{1-q} = q^{n+1} \cdot \left(M \cdot \frac{1-q^{m-n}}{1-q} \right) \leq q^{n+1} \cdot \frac{M}{1-q} \end{aligned}$$

By the preceding lemma, a uniform Cauchy sequence is a uniformly convergent sequence. Now we consider the derivatives S'_m.
We derive the formula for a finite geometric sum w.r.t. to q:

$$\sum_{k=0}^{n} Mq^k = \frac{1-q^{n+1}}{1-q}$$
$$\sum_{k=0}^{n} Mkq^{k-1} = \frac{(1-(n+1)q^n)\cdot(1-q)-(1-q^{n+1})\cdot(-1)}{(1-q)^2}$$

So, we similarly obtain that

$$\begin{aligned}|S'_m(x)-S'_n(x)| &= \left|\sum_{k=n+1}^{m} c_k\cdot k\cdot(x-p)^{k-1}\right| \le \sum_{k=n+1}^{m} Mkq^{k-1}\\ &= \frac{(1-(m+1)q^m)\cdot(1-q)-(1-q^{m+1})\cdot(-1)}{(1-q)^2}\\ &\quad - \frac{(1-(n+1)q^n)\cdot(1-q)-(1-q^{n+1})\cdot(-1)}{(1-q)^2}\\ &= \frac{-n\,q^{n+1}+q^{n+1}+n\,q^n+m\,q^{m+1}-q^{m+1}-m\,q^m}{(1-q)^2 q}\end{aligned}$$

converges to 0, since $\lim_{n\to\infty} q^n = 0$ and $\lim_{n\to\infty} nq^n = 0$. We conclude that the derivatives form a uniform Cauchy sequence, and hence a uniformly convergent sequence as well. □

Theorem 5.5.8. *Let*

$$f(x) = \sum_{n=0}^{\infty} c_n\cdot(x-p)^n$$

be a power series around point $p\in\mathbb{R}$ with strictly positive radius of convergence of $\rho>0$. Inside its radius of convergence, a power series is infinitely many times differentiable, and we can calculate the derivative summand-wise

$$\left(\frac{d}{dx}\right)^{\ell} f(x) = \sum_{n=0}^{\infty} c_n\cdot\left(\frac{d}{dx}\right)^{\ell}(x-p)^n$$

Proof. By the preceding lemma, this statement is true for f and for f'. Since f' is a power series with the same radius of convergence, we can apply the preceding lemma inductively. □

Example 5.5.9. We know that the sine function is given by a power series

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

with an infinite radius of convergence. We can therefore derive the power series summandwise

$$\sin'(x) = \sum_{k=0}^{\infty}(-1)^k \frac{(2k+1)x^{2k}}{(2k+1)!} = \sum_{k=0}^{\infty}(-1)^k \frac{x^{2k}}{(2k)!} = \cos(x)$$

5.6 Riemann Integration

Definition 5.6.1. Let $[a,b] \subset \mathbb{R}$ be a bounded closed interval. We call the points $P = \{a_0, a_1, \ldots, a_n\}$ with the property

$$a = a_0 < a_1 < a_2 < \cdots < a_{n-1} < a_n = b$$

a *partition* P of $[a,b]$. Every $[a_j, a_{j+1}]$ where $j = 0, \ldots, n-1$ is called *subinterval* of this partition. The length of the longest subinterval is called the *mesh* σ of the partition:

$$\sigma = \max\{|a_{j+1} - a_j| \,:\, j = 0, 1, \ldots, n-1\}$$

Definition 5.6.2. We call a partition $a = c_0 < c_1 < c_2 < \cdots < c_{m-1} < c_m = b$ of $[a,b]$ a *refinement* of a partition $a = a_0 < a_1 < a_2 < \cdots < a_{n-1} < a_n = b$ of $[a,b]$, if for every $j = 0, 1, \ldots, n-1$ there exist $k, \ell \in \{0, 1, \ldots, m\}$ such that $a_j = c_k$ and $a_{j+1} = c_\ell$.

Lemma 5.6.3. *Given any two partitions of $[a,b]$ there exist a partition that is a refinement for each of those partition. This partition is called a* common refinement*.*

Definition 5.6.4. Let $f\colon [a,b] \to \mathbb{R}$ be a function. Let $a = a_0 < a_1 < a_2 < \cdots < a_{n-1} < a_n = b$ be a partition of an interval $[a,b]$. Let $\xi_j \in [a_j, a_{j+1}]$ for every $j = 0, 1, \ldots n-1$ be a point. These points are called *tag points*. We define the *Riemann sum* of f for this *tagged partition* to be

$$\sum_{j=0}^{n-1} f(\xi_j) \cdot (a_{j+1} - a_j)$$

Definition 5.6.5. We call a function $f\colon [a,b] \to \mathbb{R}$ *Riemann integrable* on $[a,b]$ and denote by

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

the *Riemann integral* of f on $[a,b]$ if there exists a real number $I \in \mathbb{R}$ such that for every $\epsilon > 0$ there exists a $\delta > 0$ with the following property: For every partition $a = a_0 < a_1 < a_2 < \cdots < a_{n-1} < a_n = b$ of $[a,b]$ with mesh $\sigma < \delta$ and any $\xi_j \in [a_j, a_{j+1}]$ the Riemann sum satisfies

$$\left| I - \sum_{j=0}^{n-1} f(\xi_j) \cdot (a_{j+1} - a_j) \right| < \varepsilon$$

Then we set

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

Definition 5.6.6. Let $P = \{a_0, a_1, \ldots, a_n\}$ be a partition of the interval $[a, b]$. Let $f\colon [a, b] \to \mathbb{R}$ be a function. We define the *lower sum* $\underline{S}_P(f)$ and the *upper sum* $\overline{S}_P(f)$ of f with respect to this partition.

$$\underline{S}_P(f) = \sum_{j=0}^{n-1} \inf\{f(x) \,:\, x \in [a_j, a_{j+1}]\} \cdot (a_{j+1} - a_j)$$
$$\overline{S}_P(f) = \sum_{j=0}^{n-1} \sup\{f(x) \,:\, x \in [a_j, a_{j+1}]\} \cdot (a_{j+1} - a_j)$$

Remark 5.6.7. Let $P = \{a_0, a_1, \ldots, a_n\}$ be a partition of the interval $[a, b]$. Let $f\colon [a, b] \to \mathbb{R}$ be a function. For any choice of tag points $\xi_j \in [a_j, a_{j+1}]$ the Riemann sum $S_P(f)$ will satisfy:

$$\underline{S}_P(f) \leq S_P(f) \leq \overline{S}_P(f)$$

In order to have well-defined upper and lower sums, the function f needs to be bounded.

Lemma 5.6.8. *Let $f\colon [a, b] \to \mathbb{R}$ be a bounded function. Let P and $\widetilde{P}$ be partitions of $[a, b]$. Then we have*

$$\underline{S}_{\widetilde{P}}(f) \leq \overline{S}_P(f)$$

Theorem 5.6.9 (Darboux Theorem)**.** *Let $f\colon [a, b] \to \mathbb{R}$. Then the function f is Riemann-integrable if and only if for every $\varepsilon > 0$ there exists partition P such that*

$$|\overline{S}_P(f) - \underline{S}_P(f)| < \varepsilon$$

Lemma 5.6.10. *Let $f, g\colon [a, b] \to \mathbb{R}$ be Riemann integrable functions and let $c \in \mathbb{R}$ and let $m \in (a, b)$. Then the following hold:*

1. *$f + g$ is Riemann integrable and $\int_a^b (f(x) + g(x))dx = \int_a^b f(x)dx + \int_a^b g(x)dx$*
2. *$c \cdot f$ is Riemann integrable and $\int_a^b c \cdot f(x)dx = c \cdot \int_a^b f(x)dx$*
3. *$\int_a^b f(x)dx = \int_a^m f(x)dx + \int_m^b f(x)dx$*
4. *If $f(x) \leq g(x)$ for every $x \in [a, b]$ then $\int_a^b f(x)dx \leq \int_a^b g(x)dx$.*
5. *$f \cdot g$ is Riemann integrable*

Lemma 5.6.11 (Triangle inequality)**.** *Let $f\colon [a, b] \to \mathbb{R}$ be a Riemann integrable function. Then $|f|$ is also Riemann integrable on $[a, b]$ and*

$$\left| \int_a^b f(x)dx \right| \leq \int_a^b |f(x)|dx$$

Definition 5.6.12. A function $f\colon [a,b] \to \mathbb{R}$ is called *piece-wise continuous*, if there exists a partition $P = \{a_0, a_1, \ldots, a_n\}$ of $[a,b]$ such that f is continuous on each open interval (a_j, a_{j+1}) for $j = 0, 1, \ldots, n-1$ and such that the limits $\lim\limits_{\substack{x \to a_j \\ x > a_j}} f(x)$ and $\lim\limits_{\substack{x \to a_{j+1} \\ x < a_{j+1}}} f(x)$ exist for each $j = 0, 1, \ldots, n-1$.

Theorem 5.6.13. *Every piece-wise continuous function $f\colon [a,b] \to \mathbb{R}$ is Riemann integrable on $[a,b]$*

Theorem 5.6.14 (Integral form of the MVT)**.** *Let $f\colon [a,b] \to \mathbb{R}$ be a continuous function. Then there exists a $c \in (a,b)$ such that*

$$\int_a^b f(x)dx = f(c) \cdot (b-a)$$

Theorem 5.6.15 (Fundamental Theorem of Calculus, first version)**.** *Let $f\colon [a,b] \to \mathbb{R}$ be a continuous function. For $x \in [a,b]$ define*

$$F(x) := \int_a^x f(t)dt$$

Then F is continuous on $[a,b]$, differentiable on (a,b) and $F'(x) = f(x)$ for every $x \in (a,b)$.

Theorem 5.6.16 (Fundamental Theorem of Calculus, second version)**.** *Let $f\colon [a,b] \to \mathbb{R}$ be a Riemann integrable function and let $F\colon [a,b] \to \mathbb{R}$ be a continuous function that is differentiable on (a,b) and such that $F'(x) = f(x)$ for every $x \in (a,b)$. Then*

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

Corollary 5.6.17 (Change of variable)**.** *Let $f\colon [a,b] \to \mathbb{R}$ be a continuous function and let $g\colon [c,d] \to [a,b]$ be a bijective, continuously differentiable function. Then*

$$\int_a^b f(y)dy = \int_c^d f(g(x)) \cdot g'(x)dx$$

Corollary 5.6.18 (Integration by parts)**.** *Let $f, g\colon [a,b] \to \mathbb{R}$ be continuously differentiable functions. Then the following holds:*

$$\int_a^b f'(x) \cdot g(x)dx = f(x) \cdot g(x)\big|_a^b - \int_a^b f(x) \cdot g'(x)dx$$

Definition 5.6.19. If $f\colon (a,b) \to \mathbb{R}$ is Riemann integrable, then we define the *improper Riemann integral* by

$$\int_a^b f(x)dx := \lim_{\substack{c \to a \\ c > a}} \lim_{\substack{d \to b \\ d < b}} \int_c^d f(x)dx$$

if the limits exist. We allow $a = -\infty$ and $b = +\infty$ in this definition.

Example 5.6.20.

$$\int_0^1 \frac{1}{\sqrt{x}}dx = \lim_{\substack{c \to 0 \\ c > 0}} \int_c^1 \frac{1}{\sqrt{x}}dx = \lim_{\substack{c \to 0 \\ c > 0}} \left(2\sqrt{1} - 2\sqrt{c}\right) = 2$$

References

[1] Walter Rudin, *Principles of mathematical analysis*, 3rd ed., McGraw-Hill Book Co., New York-Auckland-Düsseldorf, 1976. ISBN 978-1259064784.