Mathematics for computer science

Frank-Olaf Schreyer

January 22, 2020

Wichtige Informationen

- Homepage der Arbeitsgruppe:

  https://www.math.uni-sb.de/ag/schreyer/
- Vorlesungshomepage:
  https://www.math.uni-sb.de/ag/schreyer/index.php/teaching/ws-19-20/133-mathematik-fuer-informatiker-1
- Kurs- bzw. Übungsgruppenanmeldung bis Freitag, 18.10.2019, 13 Uhr auf der Vorlesungshomepage.
- Übungen beginnen in der dritten Vorlesungswoche (ab Montag, 28.10.2019) mit einem Präsenzblatt.
- Sprachtest für die Teilnehmer der MfI unter
  https://www.szsb.uni-saarland.de/sprachkurse/kurs.asp?KursNr=SZENW1992
- Alle Informationen finden Sie auch auf der Vorlesungshomepage.

- The exponential function
- The Logarithm
- Trigometric functions

Asymptotic behavior and L' Hospital's rule

- Asymptotic behaviour of rational functions

Integration

- Antiderivative
- Improper Integrals

Taylor expansions

- Taylor polynomial
- Uniform convergence
- Application to power series

Fourier expansion

Sets, logic and proofs

Definition

A *set* is a collection of well-defined *elements*.

Examples:

- $A = \{a, b, c, \ldots, z\}$, the set of letters of the alphabet.
- $B = \{\text{students in this lecture}\}$.
- $C = \{2, 3, 5, 7\} = \{p \mid p \text{ is a prime number } \leq 10\}$.

Sets can be specified by listing its elements or by a characterising property of its elements. Some special sets:

- $\emptyset = \{\ \}$ the empty set,
- $\mathbb{N} = \{1, 2, \ldots\}$ the set natural numbers,
- $\mathbb{N}_0 = \{0, 1, 2, \ldots\}$,
- $\mathbb{Z} = \{0, \pm 1, \pm 2, \ldots\}$ integral numbers,
- $\mathbb{Q} = \{\frac{a}{b} \mid a, b \in \mathbb{Z}, b \neq 0\}$ rational numbers,
- $\mathbb{R} = \{\text{real numbers}\} =$
  $\{$ not necessarily periodic decimal numbers$\}$

Notation

- If x is an element of a set M, then we write

$$x \in M.$$

- If every element of a set N is also an element of a set M, then we write

$$N \subset M.$$

- $N \subseteq M$ is also in use. We use the notation $N \subsetneq M$ for $N \subset M$ and $N \neq M$.
- $N \not\subset M$ means N is not a subset of M.

Definition

We denote by $|M|$ the number of elements of a set M.

- $|\emptyset| = 0, \quad |\{0\}| = 1,$
- $|\{a, b, c, \ldots, z\}| = 26.$

If M has infinitely many elements, then we write

$$|M| = \infty.$$

Alternative notation: $\#M := |M|$.

- $\#\{\text{students in this lecture older than } 40\} = 0\,?$

Definition (Intersections, unions, complements)

Let A, B denote sets. Then

$$A \cup B = \{x \mid x \in A \text{ or } x \in B\}$$

denotes the *union*, and

$$A \cap B = \{x \mid x \in A \text{ and } x \in B\}$$

denotes the *intersection*.
A and B are called *disjoint* if $A \cap B = \emptyset$. If A and B are disjoint, then

$$|A \cup B| = |A| + |B|.$$

In general,

$$|A \cup B| + |A \cap B| = |A| + |B|$$

holds because in the sum $|A| + |B|$ the elements of $A \cap B$ are counted twice.

Distribution laws:
Let A, B, C denote sets. Then

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C),$$

and

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

holds.
If we consider only subset of a fixed set M, then

$$\overline{A} = \{x \in M \mid x \notin A\}$$

denotes the *complement* of A in M.

De Morgan's laws:
$A, B \subset M$. Then

- $\overline{A \cap B} = \overline{A} \cup \overline{B}$,
- $\overline{A \cup B} = \overline{A} \cap \overline{B}$,
- $\overline{\overline{A}} = A$

hold.
More general, $A \setminus B = \{a \in A \mid a \notin B\}$ denotes the *difference* of B in A.
Notice:

$$A \setminus B = A \setminus (B \cap A).$$

Definition (Cartesian product, power set)

If A, B are sets, then

$$A \times B = \{(a, b) \mid a \in A, b \in B\}$$

denotes the set of pairs (a, b), the *cartesian product* of A and B.

Examples:

$$\{a, \ldots, h\} \times \{1, \ldots, 8\} = \{(a, 1), \ldots, (h, 8)\}$$

is used in chess. We can describe with

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

the set of points in the plane.

Note

$$|A \times B| = |A| \cdot |B|.$$

Definition

For a fixed set M the set of all subsets

$$2^M = \{A \mid A \subset M\}$$

is called the *power set* of M.

Theorem

If M is a finite set, then

$$|2^M| = 2^{|M|}.$$

In other words: a set with n elements has 2^n subsets.

Examples:

- $2^{\emptyset} = \{\emptyset\}$, $2^{\{1\}} = \{\emptyset, \{1\}\}$
- $2^{\{1,2\}} = \{\emptyset, \{1\}, \{2\}, \{1,2\}\}$

How to prove such a theorem? In this section we will introduce two general methods for proofs:

1. Proof by contradiction
2. Proof by induction

The above theorem is proved by induction. We postpone this proof and start the discussion with a famous proof by contradiction.
Let A be a statement which can be either true or false. We want prove that A is true. For this we deduce from the assumption A *is false* a contradiction. Then the statement A is not true is false, hence A is true.

Recall a prime number p is a natural number $p \geq 2$ which has precisely two factors, namely 1 and p.

Theorem

There exist infinitely many prime numbers.

Notation (Propositional logic)

It is useful to have a short logical notation in complicated proofs. Let A and B be statements. Then

- $A \wedge B$ denotes the statement *A and B are true*,
- $A \vee B$ denotes *A or B (or both) are true*,
- $A \Rightarrow B$ denotes *if A is true, then B is true*,
  or in other words *A implies B*,
- $A \Leftrightarrow B$ denotes *A is true if and only if B is true*,
  or in other words *A is equivalent to B*,
- $\neg A$ denotes the statement *A is not true*.

If $A \Rightarrow B$ and $B \Rightarrow C$ hold, then also $A \Rightarrow C$ holds. Formulated differently, the statement

$$((A \Rightarrow B) \wedge (B \Rightarrow C)) \Rightarrow (A \Rightarrow C)$$

is always true. Clearly, this fact is used in many proofs to divide the argument into smaller pieces.
The statement $A \Rightarrow B$ should not be confused with $B \Rightarrow A$.
Example: We consider the statements
$A_1 =$ The gate is closed and $B_1 =$ Some train crosses.
Then
$A_1 \Rightarrow B_1$ is false because the gate is closed earlier,
and
$B_1 \Rightarrow A_1$ is (hopefully) true.

Theorem (Laws of de Morgan)

Let A and B be logical statements. Then

$$\neg(A \wedge B) = (\neg A) \vee (\neg B)$$
$$\neg(A \vee B) = (\neg A) \wedge (\neg B)$$

Proposition

Let A and B be logical statements. Then

$$A \Rightarrow B = (\neg A) \vee B.$$

Proposition

Let A, B, C be logical statements . Then

- *associative*

$$A \vee (B \vee C) = (A \vee B) \vee C,$$

$$A \wedge (B \wedge C) = (A \wedge B) \wedge C$$

- *commutative*

$$A \vee B = B \vee A \text{ and } A \wedge B = B \wedge A$$

- *distributative*

$$(A \vee B) \wedge C = (A \wedge C) \vee (B \wedge C),$$

$$(A \wedge B) \vee C = (A \vee C) \wedge (B \vee C)$$

Proposition

- *idempotent*

$$A \vee A = A \text{ and } A \wedge A = A$$

- *tertium non datur*

$$A \vee \neg A = 1,$$

$$A \wedge \neg A = 0$$

- *double negation*

$$\neg(\neg A) = A$$

- *units*

$$A \wedge 1 = A, \quad A \wedge 0 = 0,$$

$$A \vee 1 = 1, \quad A \vee 0 = A$$

Due to the associative law we can drop brackets and write $A \vee B \vee C$ for $(A \vee B) \vee C = (A \vee B) \vee C$ and similarly for more factors or $\wedge$. Similar to the convention $a + b \cdot c = a + (b \cdot c)$ for addition and multiplication of numbers we may drop some brackets by giving $\neg$ the strongest binding strength $\wedge$ and $\vee$ middle binding strength and $\Rightarrow$, $\Leftrightarrow$ the weakest strength.
That the statement

$$T = (A \Rightarrow B) \wedge (B \Rightarrow C) \Rightarrow (A \Rightarrow C)$$

is always true can be proved by the rules above.

Using the above proposition, one can remove $\Rightarrow$ and $\Leftrightarrow$ from any logical expression. Using de Morgan and the further rules, any logical expression S in finitely many boolean variables $A_1, A_2, \ldots, A_r$ can be brought into two normal forms:

- conjunctive normal form

$$(x_{11} \vee x_{12} \vee \ldots \vee x_{1n_1}) \wedge (x_{21} \vee x_{22} \vee \ldots \vee x_{2n_2}) \wedge \ldots \wedge (x_{m1} \vee x_{m2} \vee \ldots \vee x_{2n_m})$$

- disjunctive normal form

$$(y_{11} \wedge y_{12} \wedge \ldots \wedge y_{1n_1}) \vee (y_{21} \wedge y_{22} \wedge \ldots \wedge y_{2n_2}) \vee \ldots \vee (y_{m1} \wedge y_{m2} \wedge \ldots \wedge y_{2n_m})$$

where the $x_{ij}, y_{ij} \in \{A_1, \neg A_1, \ldots, A_r, \neg A_r\}$.

In the disjunctive normal form it is easy to check whether the formula S is satisfiable, meaning that there exists values 0 or 1 for each A_i which gives S the value 1. In programs logical statements which assure the correctness of the program most frequently are in the conjunctive normal form. No fast general algorithm is known which answers the question whether an expression in conjunctive normal form is satisfiable, and it is believed that a fast algorithm does not exit. This is at the heart of the famous $N \neq NP$ problem of complexity theory.

Principal of induction

Suppose that we have a statement $A(n)$ or each $n \in \mathbb{N}_{\geq 1}$. If

1. (base of the induction) $A(1)$ holds and
2. (induction step) $A(n) \Rightarrow A(n+1)$ holds for all $n \geq 1$,

then $A(n)$ holds for all $n \geq 1$.

Indeed $A(1) \Rightarrow A(2) \Rightarrow A(3) \Rightarrow \ldots \Rightarrow A(n) \Rightarrow \ldots$. The principal of induction is not a theorem but rather an axiom which specifies our intuition about natural numbers.

Remark

Sometimes one takes $A(0)$ as the base of the induction. In the case above also $A(0)$ makes sense and is true.

Induction is frequently used in the proof of sum or product formulas.

Definition

Suppose $a_1, \ldots, a_n \in \mathbb{R}$ are real numbers. Then

$$\sum_{k=1}^{n} a_k = a_1 + \ldots + a_n$$

denotes their sum, and

$$\prod_{k=1}^{n} a_k = a_1 \cdot \ldots \cdot a_n$$

denotes their product. By convention

$$\sum_{k=1}^{0} a_k = 0 \text{ and } \prod_{k=1}^{0} a_k = 1$$

We make this convention since then the recursive formulas

$$\sum_{k=1}^{n} a_k = \sum_{k=1}^{n-1} a_k + a_n$$

and

$$\prod_{k=1}^{n} a_k = (\prod_{k=1}^{n-1} a_k) \cdot a_n$$

make sense for all $n \geq 1$. In an implementation on a computer we would use the recursive formulas in a **for**-loop.

Example:

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

For a different proof:

$$\begin{array}{cccccc} & 1 & +\,2 & +\,\ldots & +\,n & \\ & +n & +\,(n\text{-}1) & +\,\ldots & +\,1 & \\ \hline = & (n{+}1) & +\,(n{+}1) & +\,\ldots & +\,(n{+}1) & = n(n{+}1). \end{array}$$

An anecdote says that Gauß discovered this proof as a 4-th grader.

Definition

The number of k-element subsets A of $\{1, \ldots, n\}$ is denoted by

$$\binom{n}{k} = |\{A \subset \{1, \ldots, n\} \mid |A| = k\}|.$$

For any $m \in \mathbb{N}$ the integer m-factorial is

$$m! = \prod_{k=1}^{m} k = 1 \cdot 2 \cdot \ldots \cdot m.$$

Thus $0! = 1$ by convention.

Theorem

Let k, n be integers with $0 \leq k \leq n$. Then

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

Corollary (of the proof)

$$\binom{n+1}{k} = \binom{n}{k} + \binom{n}{k-1}$$

□

Theorem

Let $a, b \in \mathbb{R}$ be real numbers, and let $n \geq 1$ be an integer. Then

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

A quick way to compute all binomial coefficients $\binom{n}{k}$ for small n is to use Pascal's triangle:

						1								
					1		1							
				1		2		1						
			1		3		3		1					
		1		4		6		4		1				
	1		5		10		10		5		1			
1		6		15		20		15		6		1		
	7		21		35		35		21		7		1	

Maps and counting

Definition

Let M and N be two sets. A *map* $f : M \to N$ is a rule which associates for each $x \in M$ an element $f(x) \in N$. We write $x \mapsto f(x)$.

Example:

$$f : M \to N \text{ with } M = \{a, b, c, d\} \text{ and } N = \{1, 2, 3, 4\}$$

$$a \mapsto 1, b \mapsto 1, c \mapsto 3, d \mapsto 4$$

If $A \subset M$ is a subset, then $f(A) = \{f(x) \mid x \in A\} \subset N$ is called the *image* of A. For $B \subset N$ the set $f^{-1}(B) = \{x \in M \mid f(x) \in B\}$ is called the *preimage* of B under f.

In the example we have

$$f(\{a, b\}) = \{1\} \text{ and } f^{-1}(\{2\}) = \emptyset.$$

For one-element subsets $B = \{y\}$ we abbreviate
$f^{-1}(y) = f^{-1}(\{y\})$.
In general, the preimage of an element can have more than one element.
In the example: $f^{-1}(1) = \{a, b\}$.
If $f : M \to N$ is a map and $A \subset M$ a subset, then the *restriction* of f to A is defined by

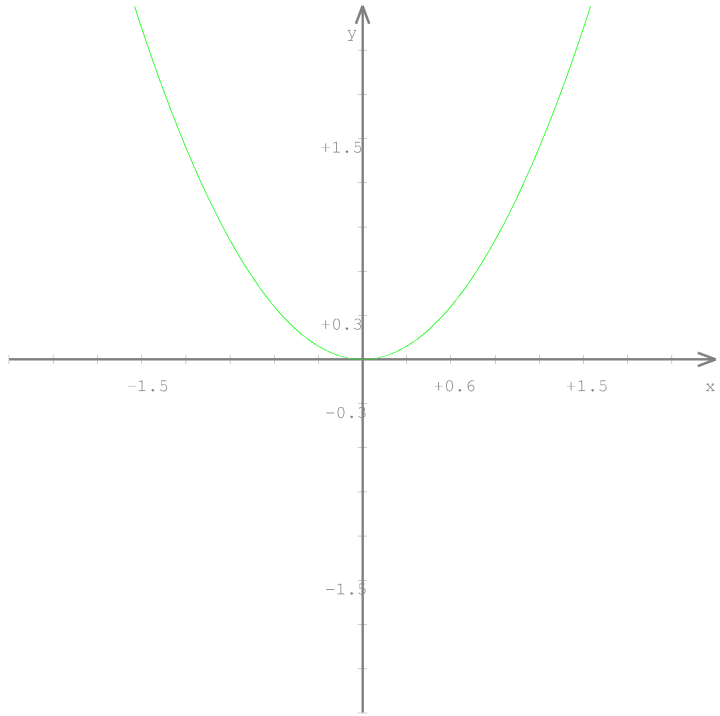
$$f|_A : A \to N, a \mapsto f(a).$$

The set

$$\Gamma_f = \{(x, y) \in M \times N \mid y = f(x)\}$$

is called the *graph* of f.

Example: $f : \mathbb{R} \to \mathbb{R}, x \mapsto f(x) = x^2$.



One can recover the map from its graph: If $x_0 \in M$ is mapped to $y_0 = f(x_0) \in N$, then

$$\Gamma_f \cap (\{x_0\} \times N) = \{(x_0, y_0)\} \subset M \times N.$$

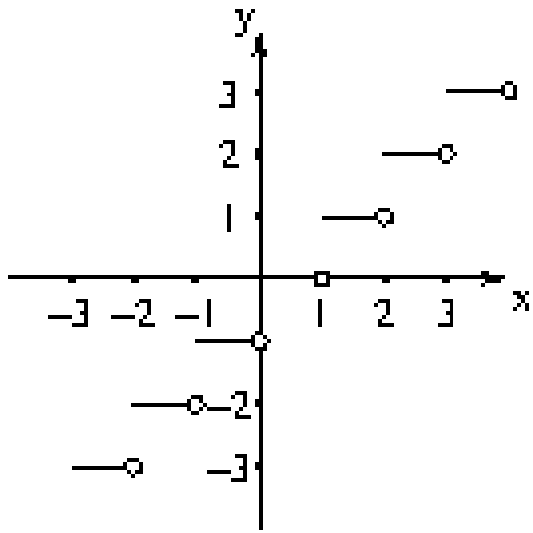
The most common maps are *real-valued functions*

$$f : D \to \mathbb{R}$$

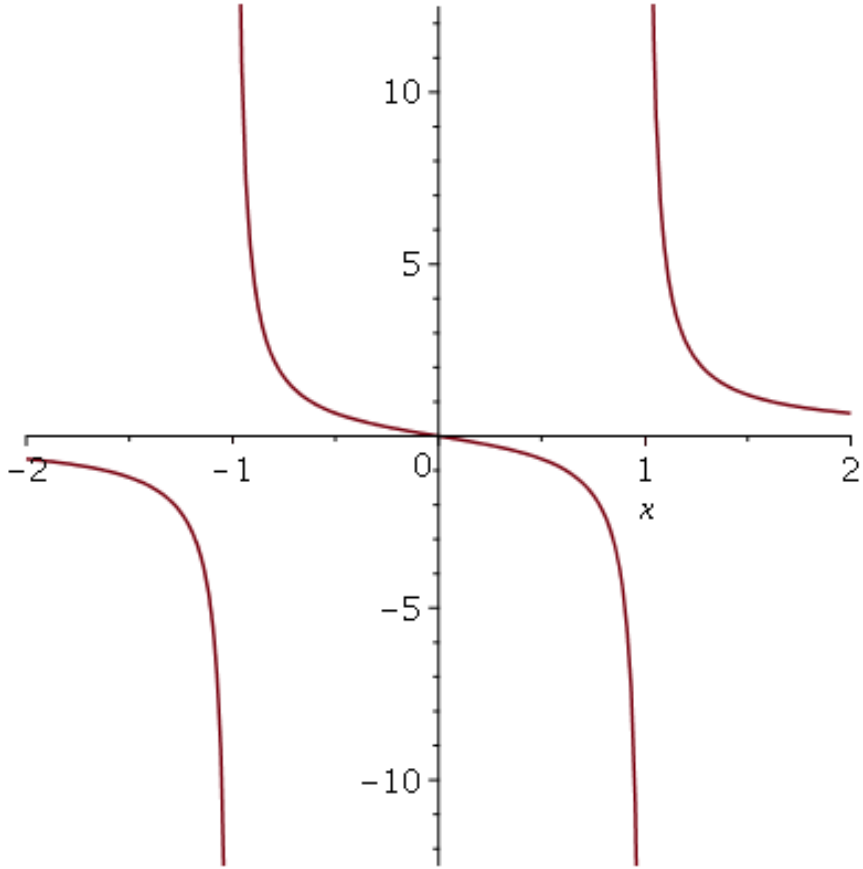
where $D \subset \mathbb{R}$. D is called the domain of definition of f.

Examples

1. $y = x$
2. $entier : \mathbb{R} \to \mathbb{R}, entier(x) := \lfloor x \rfloor = \max\{n \in \mathbb{Z} \mid n \leq x\}$



3. $y = x^2$
4. $y = \frac{x}{x^2-1}$.



In this case the natural domain of definition is $D = \mathbb{R} \setminus \{\pm 1\}$, since the formula does not define a value for $x = \pm 1$.

Two properties of maps deserve a special name.

Definition

A map $f : M \to N$ is called *injective* if

$$x_1, x_2 \in M, x_1 \neq x_2 \implies f(x_1) \neq f(x_2)$$

holds. $f : M \to N$ is called *surjective* if for all $y \in N$ there exists an $x \in M$ with $f(x) = y$.
A map $f : M \to N$ is called *bijective* if it is both injective and surjective. In this case there is the *inverse map*

$$f^{-1} : N \to M$$

defined by $f^{-1}(y) = x$ for the unique $x \in M$ with $f(x) = y$.

Note that the symbol f^{-1} is overloaded. For bijective f it usually refers to the inverse map, while for arbitrary maps $f : M \to N$ it refers to the map

$$f^{-1} : 2^N \to 2^M, B \mapsto f^{-1}(B)$$

defined by taking the preimage.

Counting

If $f : M \to N$ is bijective, then $|M| = |N|$. Conversely, we have

Theorem

Let $f : M \to N$ be a map between finite sets with $|M| = |N|$. The following are equivalent (TFAE)

a) *f is injective,*

b) *f is surjective,*

c) *f is bijective.*

Corollary (of the proof)

If $f : M \to N$ is injective, then $|M| \leq |N|$. If $f : M \to N$ is surjective, then $|M| \geq |N|$.

Corollary (Pigeon hole principal)

A map $f : M \to N$ with $|M| > |N|$ is not injective.

Example: For arbitrary $n^2 + 1$ points $p_i = (x_i, y_i)$ in a square

$$Q = \{(x, y) \in \mathbb{R}^2 \mid 0 \leq x < n, 0 \leq y \leq n\}$$

of area n^2 there exist two points $p_i, p_j, i \neq j$ with distance $dist(p_i, p_j) \leq \sqrt{2}$.

Example:

$$\sum_{i=1}^{n} i \binom{n}{i} = n2^{n-1}$$

This can be proved by induction.

Notation

Given the sets M and N we denote by

$$N^M = \{f : M \to N\}$$

the set of all maps from M to N. This is compatible with the notion 2^M of the power set: For each subset $A \subset M$ we define its characteristic function

$$\chi_A : M \to \{0, 1\}$$

by

$$\chi_A(m) = \begin{cases} 0 & \text{if } m \notin A \\ 1 & \text{if } m \in A \end{cases}.$$

Then

$$\chi : 2^M \to \{0, 1\}^M, A \mapsto \chi_A$$

is a bijection. Note

$$|N^M| = |N|^{|M|}.$$

Definition

If $f : M \to N$ and $g : N \to K$ are maps, then the composition

$$g \circ f : M \to K$$

of f and g is defined by $(g \circ f)(m) = g(f(m))$.
Composition of maps is associative:

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

this holds for any triple of maps
$f : M \to N, g : N \to K, h : K \to L$. Indeed,

$$\begin{aligned}((h \circ g) \circ f)(m) &= (h \circ g)(f(m)) \\ &= h(g(f(m))) \\ &= h((g \circ f)(m)) \\ &= (h \circ (g \circ f))(m)\end{aligned}$$

holds for all $m \in M$.

Existence and all quantifiers. The phrases *for all* and *there exists* are used very often in mathematics. One abbreviates

$$\forall = \text{ for all} \qquad \text{and} \qquad \exists = \text{ there exists.}$$

Example:

$f : M \to N$ is surjective

$\iff$ for all $n \in N$ there exists an $m \in M$ such that $f(m) = n$

$\iff \forall n \in N \, \exists m \in M : f(m) = n.$

Here we replace ':' by *such that*.

Under negation the quantifiers $\forall$ and $\exists$ are interchanges in a de Morgan style rule.

$f : M \to N$ is not surjective

$\iff \neg(\forall n \in N \, \exists m \in M : f(m) = n)$

$\iff \exists n \in N : \neg(\exists m \in M : f(m) = n)$

$\iff \exists n \in N : \forall m \in M \, \neg(f(m) = n)$

$\iff \exists n \in N : \forall m \in M \, f(m) \neq n$

$\iff$ there exists an $n \in N$ s.t. for all $m \in M$, $f(m)$ is not equal to n

translates correctly under de Morgan's rule.

Arbitrary unions and intersections. Given an arbitrary family $(A_i)_{i\in I}$ of subsets of a set M, i.e., a map

$$I \to 2^M,\ i \mapsto A_i$$

we define the union and the intersection by

$$\bigcup_{i\in I} A_i = \{x \in M \mid \exists i \in I : x \in A_i\}$$

and

$$\bigcap_{i\in I} A_i = \{x \in M \mid \forall i \in I\, x \in A_i\}.$$

De Morgan's rule still holds:

$$\overline{\bigcup_{i\in I} A_i} = \bigcap_{i\in I} \overline{A_i} \quad \text{and} \quad \overline{\bigcap_{i\in I} A_i} = \bigcup_{i\in I} \overline{A_i}.$$

Indeed,

$$\begin{aligned}
\overline{\bigcup_{i\in I} A_i} &= \{x \in M \mid \neg(\exists i \in I : x \in A_i)\} \\
&= \{x \in M \mid \nexists i \in I : x \in A_i\} \\
&= \{x \in M \mid \forall i \in I \neg(x \in A_i)\} \\
&= \{x \in M \mid \forall i \in I \; x \notin A_i\} \\
&= \{x \in M \mid \forall i \in I \; x \in \overline{A_i}\} = \bigcap_{i\in I} \overline{A_i}.
\end{aligned}$$

Equivalence relations and congruences

In mathematics and computer science one frequently considers relations.

Example: $\geq$ (greater or equal) is a relation on $\mathbb{R}$: For any two real numbers x, y the relation

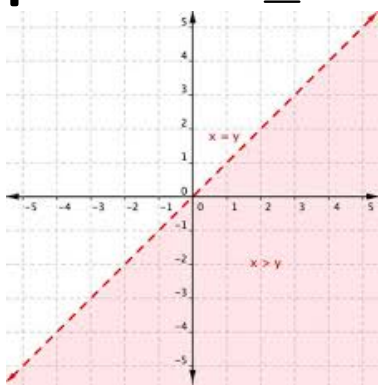
$$x \geq y \text{ is either true or false.}$$

Formally we define a relation as follows.

Definition

Let M be a set. A relation R is a subset $R \subset M \times M$. For $x, y \in M$ the relation R is satisfied if $(x, y) \in R$.

Examples: For $\geq$ we have



$$R_{\geq} = \{(x, y) \in \mathbb{R}^2 | x \geq y\}$$

$R_{=}$ is the diagonal in $\mathbb{R}^2$.

In this section we will study equivalence relations. Our goal is to weaken the notion of *equal* to a notion of *equivalent* or *similar*.
Example:

1) Let $f : M \to N$ be a map. We say $a, b \in M$ are equivalent, in symbols $a \sim b$, if $f(a) = f(b)$.
2) Let $M = \mathbb{Z}$ and let n be a positive integer. Two integers a, b are called congruent modulo n, in symbols

$$a \equiv b \mod n,$$

if $a - b$ is divisible by n.

What are the desired properties of an equivalence relation?

Definition

A subset $R \subset M \times M$ is called an *equivalence relation* (we write $a \sim b$, if $(a, b) \in R$) if the following properties hold:

i) (reflexive) $a \sim a \quad \forall a \in M$

ii) (symmetry) $a \sim b \Rightarrow b \sim a \quad \forall a, b \in M$

iii) (transitivity) $a \sim b$ and $b \sim c \Rightarrow a \sim c \quad \forall a, b, c \in M$

Examples:

0) Equality is an equivalence relation on any set M.
1) If $f : M \to N$ is a map, then $a \sim b \iff f(a) = f(b)$ is a equivalence relation because $=$ is a equivalence relation on N.
2) $a \equiv b \mod n$ is an equivalence relation:
   i) $a - a = 0 \cdot n$
   ii) $a - b = \alpha n \Rightarrow b - a = (-\alpha)n$
   iii) $a - b = \alpha n$ and $b - c = \beta n \Rightarrow a - c = (\alpha + \beta)n$
3) The relation $\geq$ on $\mathbb{R}$ is not an equivalence relation. i) and iii) are satisfied, since

$$x \geq x \; \forall x \in \mathbb{R}$$

and

$$x \geq y \text{ and } y \geq z \Rightarrow x \geq z \; \forall x, y, z \in \mathbb{R}$$

hold. But ii) does not hold: $x \geq y \not\Rightarrow y \geq x$.

Definition and Theorem

Let $\sim$ be an equivalence relation on M. For $a \in M$ we call

$$[a] = \{b \in M \mid b \sim a\}$$

the *equivalence class* of a. Any element $b \in [a]$ is called a *representative* of the equivalence class $[a]$.

Any two equivalence classes $[a]$ *and* $[b]$ *are either equal or disjoint.*

Example: The equivalence classes of $\equiv \mod 3$ are

$[0] = \{0, \pm 3, \pm 6, \ldots\}$

$[1] = \{\ldots, -5, -2, 1, 4, 7, \ldots\}$

$[2] = \{\ldots, -4, -1, 2, 5, 8, \ldots\}$.

Definition

Let $\sim$ be an equivalence relation on M. Then

$$M/\sim = \{[a] \mid a \in M\} \subset 2^M,$$

'M modulo $\sim$', denotes the set of equivalence classes.

$$\pi : M \to M/\sim,\ a \mapsto [a]$$

is called the canonical map or projection to the quotient $M/\sim$

Example: For $\equiv \mod 3$ the quotient map is

$$\begin{aligned} \pi : \mathbb{Z} &\to \{[0],[1],[2]\} \\ n &\mapsto [\text{remainder of } n \text{ divided by } 3]. \end{aligned}$$

Remark

Apparently we have

$$\pi(a) = \pi(b) \iff [a] = [b] \iff a \sim b$$

and

$$\pi^{-1}([a]) = [a].$$

Thus we can recover the equivalence relation from π. The main purpose of equivalence relations is to use suitable M and $\sim$ to construct new interesting sets $M/\sim$.

Example (The construction of $\mathbb{Q}$ from $\mathbb{Z}$):
We assume that $\mathbb{Z}$ together with the operations
$+ : \mathbb{Z} \times \mathbb{Z}, (a, b) \mapsto a + b$ and $\cdot : \mathbb{Z} \times \mathbb{Z}, (a, b) \mapsto a \cdot b$ are given.
We want to construct the field of rational numbers $\mathbb{Q}$.
For this we consider

$$M = \mathbb{Z} \times (\mathbb{Z} \setminus \{0\})$$

and the equivalence relation on M defined by

$$(p_1, q_1) \sim (p_2, q_2) \iff p_1 q_2 = p_2 q_1.$$

This is indeed a equivalence relation. i) and ii) are clear. For the transitivity, suppose

$$(p_1, q_1) \sim (p_2, q_2) \sim (p_3, q_3)$$

hence

$$p_1 q_2 = p_2 q_1 \text{ and } p_2 q_3 = p_3 q_2.$$

Then

$$p_1q_2q_3 = p_2q_1q_3 = p_2q_3q_1 = p_3q_2q_1$$

which implies

$$q_2(p_1q_3 - p_3q_1) = 0$$

Since $q_2 \neq 0$ this implies $p_1q_3 - p_3q_1 = 0 \in \mathbb{Z}$. So $(p_1, q_1) \sim (p_3, q_3)$.
As usual we denote the equivalence class $[(p, q)]$ by $\frac{p}{q}$. We now define $\mathbb{Q}$ as a set by

$$\mathbb{Q} = M/\sim = \mathbb{Z} \times (\mathbb{Z} \setminus \{0\})/\sim .$$

Addition and multiplication on $\mathbb{Q}$ are defined via representatives:

$$\frac{p}{q} + \frac{r}{s} := \frac{ps + qr}{qs}$$

and

$$\frac{p}{q}.\frac{r}{s} := \frac{pr}{qs}$$

Finally we can embed $\mathbb{Z}$ into $\mathbb{Q}$ by

$$\mathbb{Z} \hookrightarrow \mathbb{Q}, n \mapsto \frac{n}{1}.$$

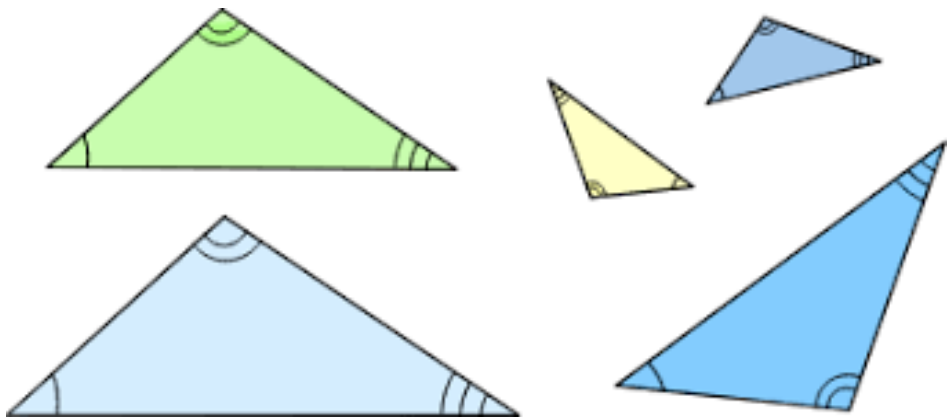
(The symbol $\hookrightarrow$ is used for injective maps which one would like to regard as an inclusion.)
Each element of $\mathbb{Q}$ has a distinguished representative, namely

$$(p, q) \in \frac{p}{q} \text{ with } p, q \text{ coprime and } q \geq 1.$$

Remark

In general there are no distinguished representatives for equivalence classes.

Example (Similarity of triangles):
Two triangles with angles (α, β, γ) and $(\alpha', \beta', \gamma')$ are similar, if the angles coincide up to their order. Similarity is an equivalence relation on the set of plane triangles.
How would you define a distinguished representative?



Congruences:

In the following we will study the equivalence relation $\equiv \mod n$ in detail. We abbreviate $\mathbb{Z}/(\equiv \mod n)$ by $\mathbb{Z}/n$. Every element of $\mathbb{Z}/n$ has a distinguished representative $i \in \{0, 1, \ldots, n-1\}$ given by the remainder of the division by n. The residue class of i is

$$[i] = \{i + kn | k \in \mathbb{Z}\} \subset \mathbb{Z}.$$

Frequently, the notation

$$\bar{i} = [i]$$

is used. The set

$$\mathbb{Z}/n = \{\bar{0}, \ldots, \overline{n-1}\}$$

has precisely n elements. The elements of $\mathbb{Z}/n$ can be added and multiplied:

$$\bar{i} + \bar{j} = \overline{i+j}, \bar{i} \cdot \bar{j} = \overline{i \cdot j}.$$

These $+$ and $\cdot$ on $\mathbb{Z}/n$ satisfy the usual laws of addition and multiplication:

- (assocative) $(\bar{i}+\bar{j})+\bar{k}=\bar{i}+(\bar{j}+\bar{k}), \quad (\bar{i}\cdot\bar{j})\cdot\bar{k}=\bar{i}\cdot(\bar{j}\cdot\bar{k}),$
- (distributive) $(\bar{i}+\bar{j})\cdot\bar{k}=\bar{i}\cdot\bar{k}+\bar{j}\cdot\bar{k},$
- (commutative) $\bar{i}+\bar{j}=\bar{j}+\bar{i}, \quad \bar{i}\cdot\bar{j}=\bar{j}\cdot\bar{i}.$

Example ($n = 6$): Then

$$(\overline{2} + \overline{2}) + \overline{5} = \overline{4} + \overline{5} = \overline{9} = \overline{3} \in \mathbb{Z}/6$$

coincides with

$$\overline{2} + (\overline{2} + \overline{5}) = \overline{2} + \overline{7} = \overline{2} + \overline{1} = \overline{3}$$

, and

$$(\overline{2} \cdot \overline{2}) \cdot \overline{5} = \overline{4} \cdot \overline{5} = \overline{20} = \overline{2} \in \mathbb{Z}/6$$

coincides with

$$\overline{2} \cdot (\overline{2} \cdot \overline{5}) = \overline{2} \cdot \overline{10} = \overline{2} \cdot \overline{4} = \overline{8} = \overline{2}.$$

To see these laws in general, we have to prove that in the definition

$$\bar{i} + \bar{j} = \overline{i + j}, \bar{i} \cdot \bar{j} = \overline{i \cdot j}$$

the result does not depend on the choice of the representative $i \in \bar{i}$ and $j \in \bar{j}$.

Remark

In $\mathbb{Z}/n$ it is possible that

$$\overline{a} \cdot \overline{b} = \overline{0} \text{ for some } \overline{a} \neq \overline{0} \text{ and } \overline{b} \neq \overline{0}.$$

Example: $\overline{2} \cdot \overline{3} = \overline{0} \in \mathbb{Z}/6$

Hence in general there is no sensible 'division' in $\mathbb{Z}/n$. An exception is the case $n = p$ is a prime number. (For $a, b \in \mathbb{Z}$ we write $a|b$ for the phrase a *divides* b).

$$\overline{a} \cdot \overline{b} \equiv \overline{0} \mod p \implies p|ab \implies p|a \text{ or } p|b.$$

Definition and Theorem

Let p be a prime number, and let $\overline{a} \in \mathbb{Z}/p$ be an element $\overline{a} \neq \overline{0}$. Then multiplication by $\overline{a}$ defines a bijective map

$$\mathbb{Z}/p \to \mathbb{Z}/p, \quad \overline{b} \mapsto \overline{b} \cdot \overline{a}.$$

The *inverse* of $\overline{a} \in \mathbb{Z}/p$ is the preimage of $\overline{1}$ which we denote

$$(\overline{a})^{-1}.$$

It is represented by a $u \in \mathbb{Z}$ such that $ua \equiv 1 \mod p$. u is called an *inverse* of $a \mod p$.

Simultaneous solutions of congruences:

Given two integers $m, n > 1$ we have a map

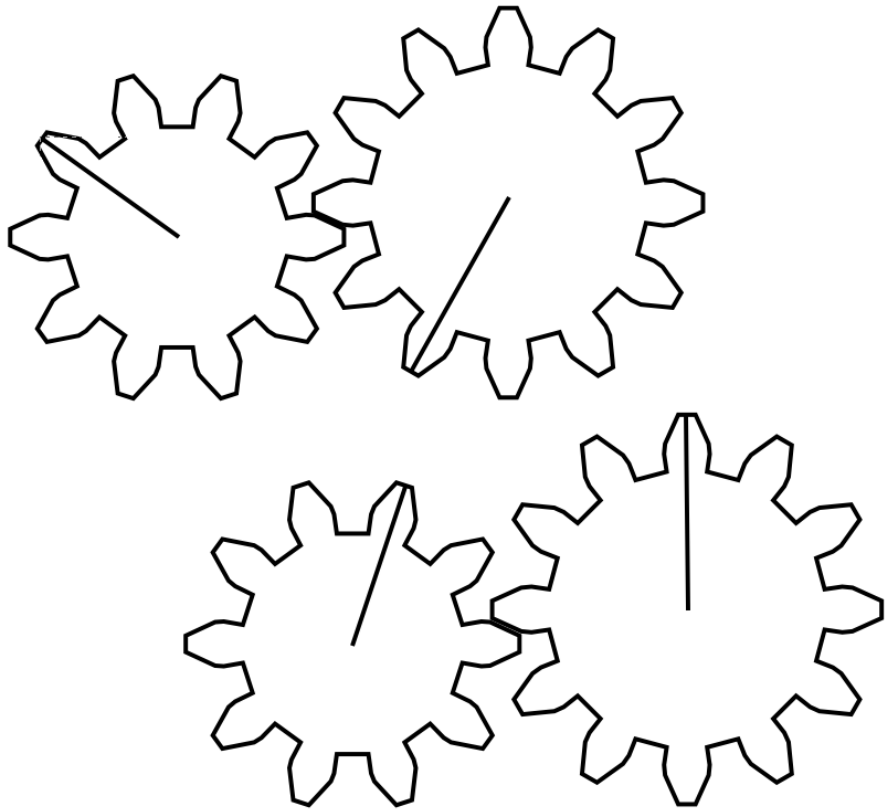
$$\mathbb{Z} \to \mathbb{Z}/m \times \mathbb{Z}/n, i \mapsto (i \bmod m, i \bmod n).$$

We ask whether a given pair $(\overline{a}, \overline{b}) \in \mathbb{Z}/m \times \mathbb{Z}/n$ lies in the image. In other words we ask whether the congruences

$$x \equiv a \mod m$$
$$x \equiv b \mod n$$

can be solved simultaneously by an $x \in \mathbb{Z}$.

Example (Cog wheels):



Is it possible to turn the wheels from first position to the second position?

How to compute the greatest common divisor? A method applied in high school uses factoring. Given $n, m \in \mathbb{N}$ we factor

$$n = p_1^{e_1} \cdots, p_r^{e_r} = \prod_{i=1}^{r} p_i^{e_i}$$

$$m = p_1^{f_1} \cdots, p_r^{f_r} = \prod_{i=1}^{r} p_i^{f_i}$$

into primes, where we allow also $e_i = 0$ or $f_i = 0$. Then

$$\gcd(m, n) = \prod_{i=1}^{r} p_i^{\min(e_i, f_i)}.$$

But factoring is hard. So hard that it is actually the basis of one of the first public-key cryptosystems.

RSA (Rivest-Shamir and Adleman, 1978):
Bob wants to send Alice a message through an open channel such that Eve who might listen to the channel cannot decipher the message. In classical cryptosystems Alice and Bob might share a common secret on which the encryption-decryption is build. For this however Bob and Alice should have met or should have interchanged the secret through a trust worthy third party. In the internet this is not feasible. The idea of public-key encryption overcomes this difficulty.
RSA relies on the difficulty of factoring and of Fermat's little Theorem.

Definition

The Euler φ-function

$$\varphi : \mathbb{N} \to \mathbb{N}$$

is defined by

$$\varphi(n) = |\{a \mid 1 \leq a < n \text{ with } \gcd(a, n) = 1\}|.$$

Clearly $\varphi(p) = p - 1$ for a prime number p. One of the remarkable properties of φ is $\varphi(mn) = \varphi(m)\varphi(n)$ if $\gcd(m, n) = 1$.

Theorem (Fermat)

Let $x, n \in \mathbb{Z}$ be integers with $\gcd(x, n) = 1$. Then

$$x^{\varphi(n)} \equiv 1 \mod n.$$

Now, Alice chooses two large prime numbers p_A, q_A and computes $n_A = p_A q_A$ and $\varphi(n_A) = (p_A - 1)(q_A - 1)$. In practise p_A and q_A will have at least 100 digits. In addition Alice chooses d_A, e_A with $d_A e_A \equiv 1 \mod \varphi(n_A)$. Then

$$n_A \text{ and } d_A$$

will be public, while

$$p_A, q_A, \varphi(n_A) \text{ and } e_A$$

remain secret.
Bob wants to send a message

$$x \in \{0, \ldots, n_A - 1\}$$

to Alice. With nearly 100% propabiltity we will have $\gcd(x, n_A) = 1$, since

$$\frac{\varphi(n_A)}{n_A} = \frac{p_A q_A - p_A - q_A + 1}{p_A q_A} \approx 1.$$

Bob computes

$$y \equiv x^{d_A} \mod n_A, y \in \{0, \ldots, n_A - 1\}$$

and sends y through the public channel to Alice. Alice computes

$$z \equiv y^{e_A} \mod n_A, z \in \{0, \ldots, n_A - 1\}$$

Since

$$\begin{aligned} z &\equiv \left(x^{d_A}\right)^{e_A} \equiv x^{1+k\varphi(n_A)} \mod n_A \\ &\equiv x \cdot \left(x^{\varphi(n_A)}\right)^k \equiv x \cdot 1 \mod n_A \end{aligned}$$

by Fermat's little theorem, Alice can recover x. Eve, who knows y, n_A and d_A, needs for the decryption the secret e_A. Knowing e_A allows with not too much effort to factor $n_A = p_A q_A$. However no fast factoring algorithm is known. So it is plausible that Eve cannot find e_A in a limited amount of time.

Algorithm

(Extended Euclidean Algorithm)
Input: Integers $a > b > 1$.
Output: $d = gcd(a, b)$ and integers $u, v \in \mathbb{Z}$ satisfying $d = ua + vb$.

1. (Initialize)
$$x_1 = a, \quad u_1 = 1, \ v_1 = 0$$
$$x_2 = b, \quad u_2 = 0, \ v_2 = 0$$
$$i = 2$$
2. **while** $x_i > 0$ **do** (
(Division with remainder) Compute $q_i \in \mathbb{Z}$ with $x_{i+1} = x_{i-1} - q_i x_i$ and $0 \leq x_{i+1} \leq x_i$;
Set $u_{i+1} = u_{i-1} - q_i u_i$, $v_{i+1} = v_{i-1} - v_i u_i$;
$i = i + 1$)
3. Set $n = i - 1$ and return $d = x_n$ and u_n, v_n.

Corollary

Let p be a prime number and $a, b \in \mathbb{Z}$. Then

$$p|ab \implies p|a \text{ or } p|b.$$

Theorem (Chinese remainder theorem)

Let $m, n \in \mathbb{Z}$ be positive integers with $\gcd(n, m) = d$ and let $a, b \in \mathbb{Z}$. The simultaneous congruence

$$x \equiv a \mod m$$
$$x \equiv b \mod n$$

has a solution if and only if $a \equiv b \mod d$. If existent, then the solution is unique modulo the least common multiple of m and n. If x is a solution, then

$$\{x + k \cdot \operatorname{lcm}(n, m) \mid k \in \mathbb{Z}\}$$

is the set of all solutions.

Theorem (Fundamental theorem of arithmetic)

Every integer $n \neq 0$ can factor

$$n = \epsilon \prod_{i=1}^{r} p_i$$

with $p_1, \ldots, p_r$ prime numbers and $\epsilon = \pm 1$. The factorisation is unique up to the order of the factors.

The real numbers

Starting in this section we will collect the essential properties of $\mathbb{R}$ in the form of axioms. All our results in real analysis will be deduced from these axioms. The axiomatic method was introduced in Euclid's elements. Hilbert in 20-th century reintroduced this approach into all parts of mathematics.
The first collection of axioms says that $\mathbb{R}$ together with the operations

$$+ : \mathbb{R} \times \mathbb{R} \to \mathbb{R}, (a, b) \mapsto a + b$$

and

$$\cdot : \mathbb{R} \times \mathbb{R} \to \mathbb{R}, (a, b) \mapsto a \cdot b$$

is a field in the sense of the following definition.

Definition

A triple $(K, +, \cdot)$ of a set K together with two maps $+ : K \times K \to K, (a, b) \mapsto a + b$ and $\cdot : K \times K \to K, (a, b) \mapsto a \cdot b$ is a *field* if the following axioms are satisfied.

K1 (Axioms for addition)

K1.1 (associativity of addition)

$$(a + b) + c = a + (b + c) \quad \forall\, a, b, c \in K$$

K1.2 (commutativity of addition)

$$a + b = b + a \quad \forall\, a, b \in K$$

K1.3 (existence of the zero element)

$$\exists\, 0 \in K \text{ such that } 0 + a = a \quad \forall\, a \in K$$

K1.4 (existence of negative elements)

$$\forall\, a \in K\ \exists\, a' \in K \text{ such that } a' + a = 0$$

We call a' the negative of a; denoted by $-a$.

K2 (Axioms for multiplication)

K2.1 (associativity of mutiplication)

$$(a \cdot b) \cdot c = a \cdot (b \cdot c) \quad \forall\, a, b, c \in K$$

K2.2 (commutativity of multiplication)

$$a \cdot b = b \cdot a \quad \forall\, a, b \in K$$

K2.3 (existence of the 1-element)

$$\exists\, 1 \in K \setminus \{0\} \text{ such that } 1 \cdot a = a \quad \forall\, a \in K$$

K2.4 (existence of inverse elements)

$$\forall\, a \in K \setminus \{0\}\ \exists\, a' \in K \text{ such that } a' \cdot a = 1$$

We call a' the inverse of a; denoted by a^{-1} or $\frac{1}{a}$.

K3 (distributivity)

$$(a + b) \cdot c = a \cdot c + b \cdot c \quad \forall\, a, b, c \in K$$

$$a \cdot (b + c) = a \cdot b + a \cdot c \quad \forall\, a, b, c \in K$$

Examples:

1. $\mathbb{Q}$ and $\mathbb{R}$ are fields.
2. $\mathbb{F}_p := (\mathbb{Z}/p, +, \cdot)$ for p a prime number is a field. In particular $\mathbb{F}_2$ is a field. Addition and multiplication can alternatively be defined by the Cayley tables:

$+$	0	1
0	0	1
1	1	0

$\cdot$	0	1
0	0	0
1	0	1

3. $(\mathbb{Z}, +, \cdot)$ is not a field, since $2 \in \mathbb{Z}$ has no inverse in $\mathbb{Z}$.

Remark

A triple $(R, +, \cdot)$ of a set together with with two operations, which satisfy the field axioms except $K2.4$ is called a commutative ring with 1. So $(\mathbb{Z}, +, \cdot)$ and $(\mathbb{Z}/n, +, \cdot)$ are examples of commutative rings with 1.

A (general) ring is a triple where one requires only the axioms K1, K2.1 and K3.

In a field we have $ab = 0 \implies a = 0$ or $b = 0$. Indeed $a \neq 0 \implies 0 = a^{-1}0 = a^{-1}(ab) = ((a^{-1}a)b = 1 \cdot b = b$ by the property 3 of the following proposition. In rings this is not necessarily true: $\overline{2} \cdot \overline{3} = 0 \in \mathbb{Z}/6$.

Proposition (Properties of fields)

Let $(K, +, \cdot)$ be a field. Then

1. 0 *and* 1 *are uniquely determined.*
2. *The negative of $a \in K$ and the inverse a^{-1} of $a \in K^* = K \setminus \{0\}$ are uniquely determined. We write $a - b$ for $a + (-b)$ and $\frac{a}{b}$ for $a \cdot b^{-1}$.*
3. $(-1) \cdot (-1) = 1$ *and* $0 \cdot a = 0 \; \forall a \in K$.
4. *In sums and products the result does not depend on how we set the brackets. The result does not depend on the order of the summands or factors respectively.*

Finite fields play an important role in computer science in particular for error-correcting codes and in cryptography. We give a further example.

Example: The Cayley tables

$+$	0	1	-1
0	0	1	-1
1	1	-1	0
-1	-1	0	1

$\cdot$	0	1	-1
0	0	0	0
1	0	1	-1
-1	0	-1	1

give $K = \{0, 1, -1\}$ the structure of a field. With the map $K \to \mathbb{Z}/3$ induced by $0 \mapsto \overline{0}, 1 \mapsto \overline{1}$ and $-1 \mapsto \overline{2}$, we see that K can be identified with $\mathbb{F}_3$.

In the exercises, we will answer the question whether there exist a field with precisely 4 elements.

$\mathbb{Q}$ and $\mathbb{R}$ are ordered fields in the sense of the following definition.

Definition

An *ordered field* is a field K together with a subset of positive elements

$$\{x \in K \mid x > 0\}$$

such that the following axioms hold:

A1: Each element $x \in K$ satisfies precisely one of the properties $x > 0$, $x = 0$ or $-x > 0$.

A2: If $x > 0$ and $y > 0$ then also $x + y > 0$.

A3: If $x > 0$ and $y > 0$ then also $x \cdot y > 0$.

In an ordered field we say $x > y$ if $x - y > 0$. We define $x \geq y :\Longleftrightarrow\ x > y$ or $x = y$. Note that $\geq$ is a transitive relation on K.

If K is ordered, then we define the *absolute value* of $x \in K$ by

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

Notice:

1. $|x| \geq 0$,
2. $|x| = 0$ if and only if $x = 0$,
3. $|x \cdot y| = |x| \cdot |y| \ \forall x, y \in K$,
4. (Δ-inequality) $|x + y| \leq |x| + |y| \ \forall x, y \in K$.

The last statement is proved by considering all possible cases $x > 0, x = 0, x < 0, y > 0, y = 0, y < 0$.

Remark.

If K is an ordered field than $\mathbb{Q}$ embeds into K as follows. We write 1_K to distinguish $1 \in \mathbb{Q}$ from the 1-element in K. Since $1_K \neq 0$ we have $-1_K > 0$ or $1_K > 0$ by A1. Since $1_K = 1_K \cdot 1_K = (-1_K) \cdot (-1_K)$ we have in fact $1_K > 0$ by A3. We define a map

$$\mathbb{N} \hookrightarrow K \text{ by } n \mapsto \sum_{i=0}^{n} 1_K = n \cdot 1_K.$$

By A2, the image consist of strictly positive elements of K. So the map is indeed injective because if $n \cdot 1_K = m \cdot 1_K$ for $n > m$ then $(n-m)1_K = 0_K$ would not be strictly positive in K. Next we extends this map to a map

$$\mathbb{Z} \hookrightarrow K \text{ by } 0 \mapsto 0_K \text{ and } -n \mapsto -(n \cdot 1_K).$$

Finally, we define

$$\iota : \mathbb{Q} \hookrightarrow K$$

by

$$\frac{n}{m} \mapsto (n1_K) \cdot (m1_K)^{-1}.$$

It is easy to see that ι respects the field structures: $\iota(a+b) = \iota(a) + \iota(b)$ and $\iota(a \cdot b) = \iota(a) \cdot \iota(b)$ holds for all $a, b \in \mathbb{Q}$. Thus we may regard $\mathbb{Q}$ as a subfield of K.
A finite field $\mathbb{F}$ cannot be ordered: There is no injective map $\mathbb{N} \to \mathbb{F}$ by the pigeon hole principal. If $\mathbb{N} \to K, n \mapsto n1_K$ is not injective then

$$p = \mathrm{char}(K) = \min\{n \in \mathbb{N} \mid n1_K = 0_K\}$$

is called the *characteristic* of K. The integer p is a prime number. Indeed, if $p = ab$ for $0 < a, b < p$ then $a1_K, b1_K \neq 0$ by the definition of p and $ab1_K = a1_K \cdot b1_K \neq 0$ gives a contradiction.
We say K has characteristic zero, $\mathrm{char}(K) = 0$, if $n1_K \neq 0 \, \forall n \in \mathbb{N}$. $\mathbb{Q}$ embeds in every field of characteristic zero by the argument above.

Definition

An ordered field K is called *archimedean* if the following axiom holds

A4: For every $x \in K$ there exits an $n \in \mathbb{N}$ such that $n1_K > x$.

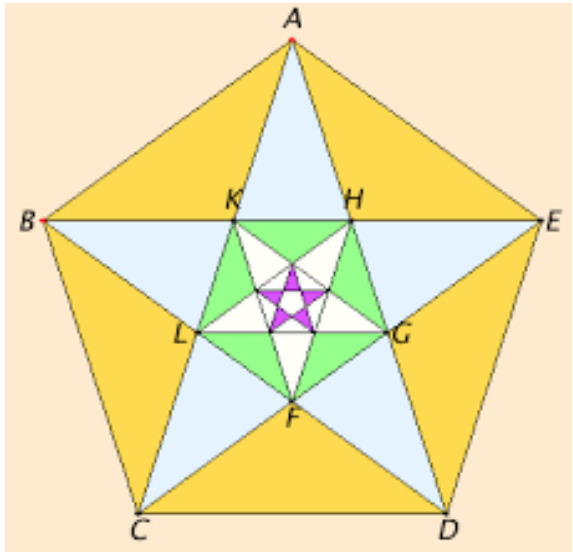
$\mathbb{Q}$ and $\mathbb{R}$ are archimedean ordered fields. There exists non-archimedean ordered fields, but right now we do not have techniques to describe an example. In a certain sense non-archimedian fields contain infinitely large elements: elements which are larger then any $n \in \mathbb{N} \subset \mathbb{Q} \subset K$.

Before we discuss the final axiom needed for $\mathbb{R}$, we recall why rational numbers are not sufficient. By Pythagoras the diagonal in a unit square has length c satisfying $c^2 = 1^2 + 1^2 = 2$. So $c = \sqrt{2}$. We claim, that

$$\sqrt{2} \notin \mathbb{Q}.$$

So rational numbers alone are not sufficient to compare the length of line segments.

Two line segments a and b are commensurable, if there exist a line segment d such that $a = md$ and $b = nd$.
In that case the ratio of the length $a : b = m : n$ is a rational number. If no such subsegment d exists, then a and b are called incommensurable. So the diagonal and the side of square are incommensurable.
If a and b are commensurable, then following Euclid's algorithm we can find the common line segment d by successively marking off the smaller segment from larger. After finitely many steps the remaining line sequences are equal and the algorithm stops with d.
The diagonal and the sides of a regular pentagon are visibly incommensurable:



We first observe that $\alpha = \measuredangle BAF$ and $\alpha' = \measuredangle BFA$ coincide. Indeed since $FBCD$ is a parallelogram, the angles $\alpha' = \measuredangle EFD = \measuredangle FBD$ coincide. By the rotational symmetry we have $\beta = \measuredangle ABE = \measuredangle EAD$ and $\alpha' + \beta = \alpha + \beta$, hence $\alpha' = \alpha$. Consequently ABF is an isosceles triangle and the length s of the side $\overline{AB}$ coincides with the length of $\overline{BF}$. Hence $\overline{EF}$ has the length $d - s$ where d denotes the length of the diagonal $\overline{BE}$. We now can draw a further regular pentagon side length $d - s$ and diagonal of length s within the triangle BZD. The process does not stop. Actually $x = \frac{d}{s}$ is the golden ratio: $\frac{d}{s} = \frac{s}{d-s}$ implies that $x = \frac{1}{x-1}$. Hence x is a root of the equation $x^2 - x - 1 = 0$ which has solutions $\frac{1 \pm \sqrt{5}}{2}$. Since $x > 0$ we get

$$\frac{d}{s} = \frac{1 + \sqrt{5}}{2}.$$

Convergence and the completeness axiom

Convergence is the central idea of analysis. We define when a sequence of real numbers converges. This idea goes back to the ancient Greek mathematicians Eudoxus (390 - 337 BC) and Archimedes (287 - 212 BC). Isaac Newton (1643-1726) introduced this concept into modern science.

Sequences

Definition

A *sequence* of real numbers $(a_n) = (a_n)_{n\in\mathbb{N}}$ is a map

$$\mathbb{N} \to \mathbb{R},\ n \mapsto a_n$$

where we usually do not give the map a name, but instead use the index notation a_n for the *n-th term* or *element* of the sequence.

Examples:

1. $(a_n) = (\frac{1}{n})$ i.e. $a_n = \frac{1}{n}$ has terms
$$1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots$$
2. $(b_n) = (n^2)$,
$$1, 4, 9, 16, \ldots$$
is the sequence of square numbers,
3. $(c_n) = (2^n)$,
$$2, 4, 16, 64, \ldots$$
is the sequence of powers of 2.

4. Sequences are used frequently in intelligence test, where the task is to get the next term from an initial part of a sequence. For example,

$$2, 4, 3, 6, 5, 10, 9, 18, \ldots$$

has the recursive rule

$$a_{n+1} = \begin{cases} 2a_n & \text{if } n \text{ is odd} \\ a_n - 1 & \text{if } n \text{ is odd} \end{cases}.$$

5. The sequence

$$(a_n) = (1, 2, 4, 6, 10, 12, 16, 18, 22, \ldots)$$

has the rule $a_n = n$-th prime number $- 1$.

6. Recursively defined sequence occur frequently.

$$(f_n) = (0, 1, 1, 2, 3, 5, 8, 13, \ldots)$$

defined by $f_0 = 0, f_1 = 1,$ and $f_{n+1} = f_n + f_{n-1}$ for $n \geq 2$ is called the sequence of Fibonacci numbers. In the exercise we will prove:

$$f_n = \frac{1}{\sqrt{5}}((\frac{1+\sqrt{5}}{2})^n - (\frac{1+\sqrt{5}}{2})^n)).$$

A general method to get closed formulas for sequences defined by linear recursions will be a topic of the lecture MfI 2.

7. In analysis, sequences are used to get arbitrary good approximations of a real number. For example,

$$(3, 3.1, 3.14, 3.141, 3.1415, \ldots)$$

are better and better approximations of Archimedes constant

$$\pi = 3.14592653589793238462643 3\ldots$$

The letter π for (half) the perimeter is in use since the 18-th century.

Convergence

We give the concept of arbitrary good approximations a precise meaning:

Definition

Let (a_n) be a sequence of real numbers and a a further real number. We say (a_n) converges to a, in sympols

$$\lim_{n\to\infty} a_n = a,$$

if $\forall\, \varepsilon > 0\, \exists n_0 \in \mathbb{N}$ such that $|a_n - a| < \varepsilon\ \forall n \geq n_0$. We call a the *limit* of the sequence (a_n).

Examples:

1. The limit of $(a_n) = (\frac{1}{n})$ is $\lim_{n\to\infty} a_n = 0$. Indeed given $\varepsilon > 0$, there exits an $n_0 \in \mathbb{N}$ with $n_0 > \frac{1}{\varepsilon}$ by the Archimedean axiom. Then
$$|\frac{1}{n} - 0| = \frac{1}{n} \leq \frac{1}{n_0} < \varepsilon\ \forall\, n \geq n_0.$$

2. We have
$$\lim_{n\to\infty} \frac{n+1}{n} = 1.$$
Indeed,
$$\left|\frac{n+1}{n} - 1\right| = \frac{1}{n} < \varepsilon \;\forall\, n \geq n_0$$
if we choose $n_0 = \lceil \frac{1}{\varepsilon} \rceil$.
3. The *constant sequence* (a_n) with $a_n = a$ for all n converges to a.
4. The sequence $(-1)^n)$ does not converges. Sequences which do not converge are called *divergent*.

Remark

If (a_n) converges to a, then for arbitrary small $\varepsilon > 0$ all but finitely many terms a_n lie in the interval

$$]-\varepsilon + a, a + \varepsilon[.$$

Sometimes $]-\varepsilon + a, a + \varepsilon[$ is called an ε-neighborhood of a.

Definition

Let $a < b$ be two real numbers. Then intervals with boundary point a, b are defined by

$$\begin{aligned}
[a,b] &:= \{x \in \mathbb{R} \mid a \leq x \leq b\} \quad (\textit{closed interval}),\\
]a,b[&:= \{x \in \mathbb{R} \mid a < x < b\} \quad (\textit{open interval}),\\
[a,b[&:= \{x \in \mathbb{R} \mid a \leq x < b\} \quad (\textit{half} - \textit{open interval}),\\
]a,b] &:= \{x \in \mathbb{R} \mid a < x \leq b\} \quad (\textit{half} - \textit{open interval}).
\end{aligned}$$

The notation $(a,b) =]a,b[$, $(a,b] =]a,b]$ etc. are also in use.

Remark

The limit $\lim a_n$ of a convergent sequence is uniquely determined.

Proposition (Calculation rules for limits)

Let (a_n) and (b_n) be convergent sequences with limits $a = \lim a_n$ and $b = \lim b_n$. Then the following holds:

1. *The sequence $(a_n + b_n)$ is also convergent with limit $a + b$. In other words:*

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

   if the right hand side exists.

2. *The sequence $(a_n \cdot b_n)$ is also convergent with limit $a \cdot b$. In other words:*

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

   if the right hand side exists.

3. *If $b_n \neq 0$ and $b \neq 0$ then the sequence $(\frac{a_n}{b_n})$ is also convergent and*

$$\lim_{n\to\infty} \frac{a_n}{b_n} = \frac{\lim a_n}{\lim b_n} = \frac{a}{b}.$$

Examples of sequences in Computer science

Let A be an algorithm which can have input of variable length n.
Example:[Addition of of integers] Let a_n be the maximal run time for the addition for two n-digit numbers.
Given two integers

$$d = \sum_{i=0}^{n-1} d_i 10^i, \quad e = \sum_{i=0}^{n-1} e_i 10^i$$

with digits d_i and e_i we add them with the scheme

$$\begin{array}{ccccccc} & d_{n-1} & \cdots & d_2 & d_1 & d_0 \\ & e_{n-1} & \cdots & e_2 & e_1 & e_0 \\ c_n & c_{n-1} & \cdots & c_2 & c_1 & \\ \hline f_n & f_{n-1} & \cdots & f_2 & f_1 & f_0. \end{array}$$

For the addition of two (respectively three) one digit numbers we can use a look-up table. If the look-up takes t machine clock cycles, which each take s seconds, then the run time of the algorithm is

$$a_n = t \cdot s \cdot n.$$

Definition (Landau symbols)

Let (a_n) be a sequence of positive numbers numbers and (b_n) a further sequence. We say

$$(b_n) \in O(a_n),\ (b_n) \text{ growth at most as } (a_n),$$

if there exist a constant $c \in \mathbb{R}_{>0}$ and an integer $n_0 \in \mathbb{N}$ such that

$$|b_n| \leq c a_n \ \forall n \geq n_0.$$

We write

$$(b_n) \in o(a_n),$$

if $\lim_{n\to\infty} \frac{b_n}{a_n} = 0$.

Hence

$$o(a_n) = \{(b_n) \in \mathbb{R}^{\mathbb{N}} \mid \lim_{n\to\infty} \frac{b_n}{a_n} = 0\}.$$

There are further Landau symbols, whose definitions we do not give here.

Example: $t \cdot s \cdot n \in O(n)$. This statement, that addition of two n digit numbers has run time in $O(n)$ is in many aspect better then the precise formula, because it does not depend on the hardware or implementation details, or the question whether we use binary or decimal expansion.

Example: The multiplication scheme for two n digits numbers requires n^2 memory task. Hence the naive multiplication scheme has run time $O(n^2)$. It can be done faster:

Algorithm (Karatsuba,1962)

Input: Two integers a, b with $n = 2^k$ binary digits.
Ouput: The product $a \cdot b$.

1. Write $a = a_0 + a_1 2^{k/2}$, $b = b_0 + b_1 2^{k/2}$ where a_0, a_1, b_0, b_1 have only $2^{k/2}$ binary digits.
2. Call the algorithm recursively to compute

$$a_0 b_0, (a_0 + a_1)(b_0 + b_1), a_1 b_1.$$

3. Return

$$a_0 b_0 + [(a_0 + a_1)(b_0 + b_1) - a_0 b_1 - a_1 b_1] \cdot 2^{k/2} + a_1 b_1 2^k$$

The third step involves the addition of several integers with $2^{k/2}$ binary digits, hence the additions altogether are in $O(n) = O(2^k)$ since $\sum_{i=0}^{k-1} 2^i = 2^k - 1$. The crucial point is that we use in the second step only three instead of four multiplication. Hence we obtain an algorithm for the multiplication of two n digits numbers of cost

$$O(n^{log_2 3}) \subset O(n^{1.59}).$$

That is much better than $O(n^2)$. Even faster algorithm exists:

Theorem (Schönhage-Strassen, 1981)

Two n digit numbers can be multiplied with cost

$$O(n \log n \log \log n).$$

The completeness axiom

We will formulate the completeness axiom for $\mathbb{R}$. Roughly speaking it says, every sequence which looks like an convergent sequence does converges.

Definition

A sequence (a_n) of real numbers is *bounded from above*, *bounded from below* or *bounded* if there exists a bound $M \in \mathbb{R}$ such that

$a_n \leq M \ \forall n$,

$a_n \geq M \ \forall n$, or

$|a_n| \leq M \ \forall n$ respectively.

The sequence (a_n) is called *increasing* or *decreasing*, if $a_{n+1} \geq a_n \ \forall n$ or $a_{n+1} \leq a_n \ \forall n$ respectively. (a_n) is called *monotone* if it is increasing or decreasing. We speak of *strictly increasing* or *strictly decreasing* sequences, if the inequalities are strict.
If (a_n) is a sequence and (n_k) a strictly increasing sequence of integers, then we call $(a_{n_k})_{k \in \mathbb{N}}$ a *subsequence* of (a_n).

Remark

Convergent sequences are bounded.

Our first version of the completeness axiom is

Theorem (Completness axiom, first version)

Every bounded monotone sequence (a_n) *of real numbers converges.*

Our second version uses the concept of Cauchy-sequences.

Definition

A sequence (a_n) of real numbers is a *Cauchy sequence*, if

$$\forall \varepsilon > 0 \; \exists n_0 \in \mathbb{N} \text{ such that } |a_n - a_m| < \varepsilon \; \forall n, m \geq n_0.$$

Remark

1. A convergent sequence is a Cauchy sequence.
2. Cauchy sequences are bounded.

Our second version of the completeness axiom says that the converse is true:

Theorem (Completeness axiom, second version; Cauchy criterion)

Every Cauchy sequence of real numbers converges.

For the proof that the Theorem on bounded monotone sequences implies the Cauchy criterion we need a lemma.

Lemma

Every sequence contains a monotone subsequence.

A variation of the idea above leads to a proof of the following very useful theorem.

Theorem (Bolzano-Weierstrass)

Every bounded sequence (a_n) *has a convergent subsequence* (a_{n_k}).

Definition

An *upper bound* of a subset $M \subset \mathbb{R}$ is a real number b such that $a \leq b \, \forall a \in M$. A subset $M \subset \mathbb{R}$ is bounded from above, if an upper bound of M exists. The supremum

$$\sup M$$

is a smallest upper bound b', i.e. an upper bound b' such that $b' \leq b$ for all other upper bounds of M.

Theorem (Existence of the supremum)

Every nonempty subset $M \subset \mathbb{R}$ which is bounded from above, has a supremum.

Remark

1. The above theorem is another property of $\mathbb{R}$ which is equivalent to the completeness axiom. In some courses this is used for the completeness axiom.
2. We some times write $\sup M = +\infty$ for subsets which have no upper bound and $\sup \emptyset = -\infty$ for the empty set.
3. The notion bounded from below, and the infimum $\inf M$ for the largest lower bound are defined similarly. Nearly the same argument as above proves the existence of an infimum for non-empty subset which are bounded from below. We set $\inf M = -\infty$, if M has no lower bound, and $\inf \emptyset = +\infty$.

Square roots

Further axioms are not needed to characterise $\mathbb{R}$. Let us prove the existence of square roots of real positive numbers using our axioms.

Theorem

Let $b \in \mathbb{R}_{>0}$. Let $a_0 \in \mathbb{R}_{>0}$ any starting value and consider the recursively defined sequence sequence

$$a_{n+1} = \frac{1}{2}(a_n + \frac{b}{a_n}).$$

Then (a_n) is well-defined and converges to limit $a \in \mathbb{R}_{>0}$ satisfying $a^2 = b$. We write $a = \sqrt{b}$.

Example: We apply the algorithm of the theorem to compute square roots approximately.

1. $b = 4$, $a_0 = 1$. The correct value is $\sqrt{b} = 2$. The algorithm gives

n	a_n	$\frac{b}{a_n}$
0	1	4
1	2.5	1.6
2	2.05	$1.\overline{95121}$
3	2.0006097 ...	

2. For $b = 2$ und $a_0 = 1$ gives $\sqrt{2}$:

n	a_n	$\frac{b}{a_n}$
0	1	2
1	1.5	1.3333 ...
2	1.41666 ...	1.411764 ...
3	1.414215 ...	1.414211 ...
4	1.41421356237 ...	

Already the fourth value has 12 correct digits!

Remark

Let $b > 0$. The sequence

$$a_{n+1} = \frac{1}{2}(a_n + \frac{b}{a_n})$$

converges to $a = \sqrt{b}$ remarkable fast. If we define the *relative error* f_n of a_n by the formula $a_n = a \cdot (1 + f_n)$. Then $f_n \geq 0$ for $n \geq 1$. Substituting into $a_{n+1} = \frac{1}{2}(a_n + \frac{b}{a_n})$ yields $a(1 + f_{n+1}) = \frac{1}{2}(a(1 + f_n) + \frac{a^2}{a(1+f_n)})$, hence $1 + f_{n+1} = \frac{1}{2}\big((1 + f_n) + \frac{1}{1+f_n}\big) = \frac{1}{2} \cdot \frac{2+2f_n+f_n^2}{1+f_n}$. We conclude

$$f_{n+1} = \frac{1}{2} \cdot \frac{f_n^2}{1 + f_n} \leq \frac{1}{2} \cdot \min(f_n, f_n^2).$$

If the relative error $f_n \geq 1$, then it will be halved in the next step. Once we reach $f_n < 1$, then $f_{n+1} = \frac{1}{2} \cdot f_n^2$. In this case, the number correct decimal digits will double with each iteration step. One speaks of *quadratic convergence*.

The existence of real numbers

After discussing the axioms it might (and should) become questionable whether a triple $(\mathbb{R}, +, \cdot)$ satisfying the axioms K1-K3,A1-A4 and the completeness axiom does exist. Certainly defining $\mathbb{R} = \{\text{decimal numbers}\}$ is not convincing since the definition of an infinite decimal numbers already uses the concepts of convergence. Certainly it will not explain properly, why

$$0.999\ldots = 0.\overline{9} = 1$$

holds. Also aliens might find it rather bizarre that we use ten digits. We will give two constructions starting from the rational numbers. But first let us establish that these axioms characterise $\mathbb{R}$ completely.

Theorem

Let R and R' be two Archimedean ordered fields satisfying the completeness axiom. Then there exists a unique bijection

$$\varphi : R \to R'$$

compatible with all structures, for example $\varphi(a+b) = \varphi(a) + \varphi(b), \varphi(a \cdot b) = \varphi(a) \cdot \varphi(b)$ *and* $a > b \Rightarrow \varphi(a) > \varphi(b)$.

By the theorem above, it does not matter how we construct $\mathbb{R}$.
We sketch two constructions as
$\{$Cauchy sequences in $\mathbb{Q}\}/\{$zero sequences$\}$ or as Dedekindt cuts.

Construction: ($\mathbb{R}$ as Cauchy sequences modulo zero sequences])
A *zero sequence* is a sequence which converges to 0. Consider

$$M = \{(a_n) \in \mathbb{Q}^{\mathbb{N}} \mid (a_n) \text{ is a Cauchy sequence}\}.$$

Element wise addition and multiplication gives M the structure of a commutative ring with 1, where 1 corresponds to the constant sequence (1). We define on M an equivalence relation by

$$(a_n) \sim (b_n) : \Longleftrightarrow \ (a_n - b_n) \text{ is a zero sequence,}$$

and define $\mathbb{R} := M/\sim$ as a set. Then

$$[(a_n)] + [(b_n)] := [(a_n + b_n)]$$

is well-defined since the sum of two zero-sequence is a zero sequence, and

$$[(a_n)] \cdot [(b_n)] := [(a_n \cdot b_n)]$$

is a well-defined, because the product of a bounded sequence with a zero sequence is a zero sequence.

The second approach has as its basic idea that a real number a is uniquely determined by the set

$$U = \{x \in \mathbb{Q} \mid x < a\}.$$

Definition

A *Dedekindt cut* (U, V) is a pair of non-empty subsets of $\mathbb{Q}$ satisfying

$$U \cup V = \mathbb{Q} \text{ and } u < v \; \forall u \in U \; \forall v \in V.$$

If $r \in \mathbb{Q}$ then $(U(r), V(r)) = (\{u \in \mathbb{Q} \mid u < r\}, \{v \in \mathbb{Q} \mid v \geq r\})$ is called a *well-chosen rational* Dedekindt cut, and $(U'(r), V'(r)) = (\{u \in \mathbb{Q} \mid u \leq r\}, \{v \in \mathbb{Q} \mid v > r\}$ is a badly chosen rational Dedekindt cut. All other Dedekindt cuts are called *irrational*. A Dedekindt cut is called *good* if it is a well-chosen rational or an irrational Dedekindt cut.

Construction: ($\mathbb{R}$ as a set of Dedekindt cuts)
We define $\mathbb{R}$ as a set as

$$\mathbb{R} := \{(U, V) \in 2^{\mathbb{Q}} \times 2^{\mathbb{Q}} \mid (U, V) \text{ is a good Dedekindt cut}\}.$$

To define the field structure we take as addition

$$(U_1, V_1) + (U_2, V_2) = (U_3, V_3)$$

by $U_3 = \{u_1 + u_2 \mid u_1 \in U_1, u_2 \in U_2\}$ and $V_3 = \mathbb{Q} \setminus U_3$. It is not difficult but a bit tedious to define all the structure. For example, the negative of (U, V) is not always
$(-V, -U) = (\{-v \mid v \in V\}, \{-u \mid u \in U\})$, since this might be a badly chosen rational cut. For details see E. Landau, xxx

The completeness axiom is best verified for bounded increasing sequences. If $((U_n, V_n))_{n \in N}$ is an increasing sequence bounded from above by (U_M, V_M), i.e.

$$U_n \subset U_{n+1} \subset U_M \; \forall n$$

then (U, V) with

$$U = \bigcup_{n \geq 1} U_n \subset U_M$$

and $V = \mathbb{Q} \setminus U$ is the limit. $\square$

Complex numbers

We complete the construction of number systems by constructing $\mathbb{C}$ from $\mathbb{R}$. This is by far the easiest step in the chain

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{R} \subset \mathbb{C}$$

of constructions.

Definition

As a set we define

$$\mathbb{C} = \mathbb{R} \times \mathbb{R}.$$

Addition and multiplication on $\mathbb{C}$ are defined by

$$(a_1, b_1) + (a_2, b_2) := (a_1 + a_2, b_1 + b_2)$$

and

$$(a_1, b_1) \cdot (a_2, b_2) = (a_1 a_2 - b_1 b_2, a_1 b_2 + b_1 a_2)$$

The associativity, commutativity and distributivity laws are verified in a straight forward way using these laws in $\mathbb{R}$.

The element $0 = (0, 0) \in \mathbb{C}$ is the zero element, and $1 = (1, 0) \in \mathbb{C}$ is the one element. The map

$$\mathbb{R} \hookrightarrow \mathbb{C}, a \mapsto (a, 0)$$

is an embedding, which respects addition and multiplication. The imaginary unit $i = (0, 1) \in \mathbb{C}$ is an element with $i^2 = -1 = (-1, 0) \in \mathbb{C}$. We usually write $c = a + ib$ instead of (a, b). To compute $(a_1 + ib_1)(a_2 + ib_2) = (a_1 a_2 - b_1 b_2) + i(a_1 b_2 + a_2 b_1)$ one only uses the usual laws and memorise $i^2 = -1$.
The inverse of $c = a - ib$ is given by the formula

$$c^{-1} = \frac{a - ib}{a^2 + b^2}.$$

Hence $(\mathbb{C}, +, \cdot)$ is a field. $\mathbb{C}$ cannot be ordered because $i \neq 0$ would imply $-1 = i^2 > 0$, a contradiction.

For $c = a + ib$ we call $\mathrm{Re}\, c = a, \mathrm{Im}\, c = b$ the real and imaginary part of c. The distance of c from 0

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

is called the absolute value of c. The number

$$\overline{c} = a - ib$$

is called the complex conjugate of c. The formula for the inverse can be also memorised by

$$c^{-1} = \frac{1}{c} = \frac{\overline{c}}{c\overline{c}} = \frac{\overline{c}}{|c|^2}.$$

Geometric interpretation of addition and multiplication.
Addition is simply vector addition.
Multiplication is best under stood in polar coordinates: If we write

$$c = r(\cos\alpha + i\sin\alpha)$$

then $r = |c|$ and α is called the argument of c

$$\begin{aligned} c_1 c_2 &= r_1(\cos\alpha_1 + i\sin\alpha_1)\cdot r_2(\cos\alpha_2 + i\sin\alpha_2) \\ &= r_1 r_2(\cos(\alpha_1+\alpha_2) + i\sin(\alpha_1+\alpha_2)) \end{aligned}$$

holds by the addition laws of sin and cos. Hence in multiplication the absolute values multiply and the arguments are added.

For $c \in \mathbb{C}^* = \mathbb{C} \setminus 0$ and $d \in \mathbb{C}$, the map

$$\mathbb{C} \to \mathbb{C}, z \mapsto cz + d$$

is a rotation by the argument α of c combined with a stretching by the factor $|c|$ followed by a translation by d.
The absolute value satisfies

1. $|z| \geq 0$, and $|z| = 0$ if and only if $z = 0$,
2. $|z \cdot w| = |z| \cdot |w| \; \forall z, w \in \mathbb{C}$,
3. (Δ-inequality) $|z + w| \leq |z| + |w| \; \forall z, w \in \mathbb{C}$.

The name triangle inequality is clear now: The third side of a triangle in $\mathbb{C}$ is at most as long as the sum of the length of the two other sides.

By definition, a sequence (z_n) of complex numbers converges to $z \in \mathbb{C}$, if

$$\forall \varepsilon > 0 \; \exists n_0 \text{ such that } |z_n - z| < \varepsilon \; \forall n \geq n_0,$$

equivalently, if the sequences of real and imaginary parts $(\mathrm{Re} z_n)$ and $(\mathrm{Im} z_n)$ converge to $\mathrm{Re} z$ and $\mathrm{Im} z$ respectively.
Complex numbers first showed up in the work of Cardano (1501–1576), who used them to give a formula for the three roots of a cubic polynomial with real coefficients. The name imaginary is used since first i was an imagined solution of the equation $x^2 + 1 = 0$. There is nothing mysterious about imaginary numbers, except that they do not lie on the real number line, but in the complex number plane.
Complex numbers play an important role in mathematics because of the following theorem.

Theorem (Fundamental theorem of algebra)

Let $p(z) = c_n z^n + c_{n-1} z^{n-1} + \ldots + c_1 z + c_0$ *be a polynomial of degree* n *with complex coefficients* c_k*. Then* p *has a complex root, equivalently, there exists* $z_1, \ldots, z_n \in \mathbb{C}$ *such that*

$$p = c_n \prod_{k=1}^{n} (z - z_k)$$

factors completely in linear factors.

Thus adding the imaginary root $i = \sqrt{-1}$ to our number system, all polynomials have roots. We do not give a proof of this theorem in this course.

For polynomials $p(x)$ with real coefficients, we have

$$p(w) = 0 \implies p(\overline{w}) = 0.$$

Thus the roots of real polynomials consist of a certain number n_1 of real roots (counted with multiplicity) and n_2 pairs of complex conjugate roots (counted with multiplicity), where

$$n_1 + 2n_2 = \deg p(x).$$

add up to the degree of $p(x)$. Since

$$q(x) = (x - w)(x - \overline{w}) = x^2 - 2\mathrm{Re}w + |w|^2$$

real polynomials factor in linear forms and quadric factors over $\mathbb{R}$.

In Cardano's formula, complex number are need to describe the roots of a cubic real polynomial with three real roots. In case there is only one real root, complex numbers can be avoided to give a formula for the single real root.
In physics complex numbers are used in quantum mechanics. In the course complex numbers will show up in the description of the range of convergence of power series. They are also crucial for the Fourier transform, which is used in sound and image compression in Computer science.

Countable sets

Definition

A set M is called countable if there exists a surjective map

$$\varphi : \mathbb{N} \to M.$$

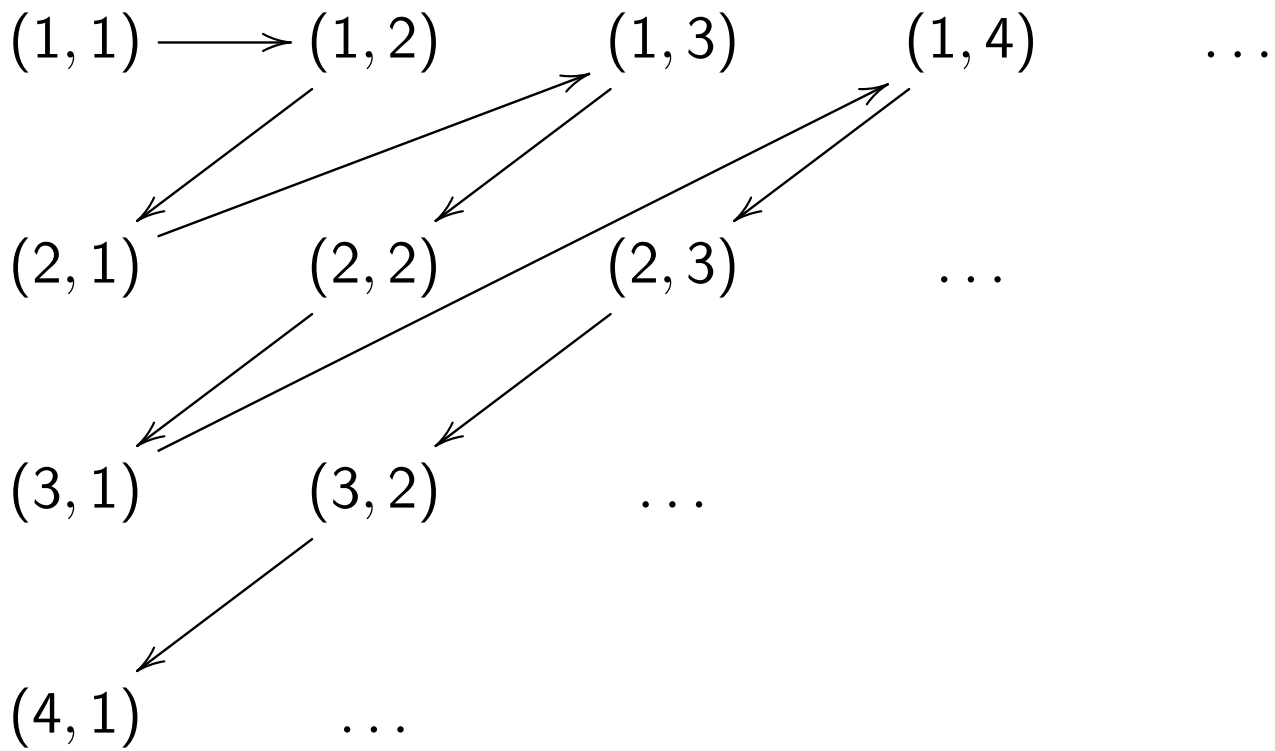
We call M uncountable if M is not countable.

Example:

1. Every finite set is countable.
2. $\mathbb{Z}$ is countable: $\mathbb{Z} = \{0, 1, -1, 2, -2, \dots\}$, more precisely

$$\varphi\colon \mathbb{N} \to \mathbb{Z},\ \varphi(n) = \begin{cases} 0, & n = 1, \\ \frac{1}{2}n, & n \text{ even}, \\ -\frac{1}{2}(n-1), & n \text{ odd } \geq 3. \end{cases}$$

3. $\mathbb{N} \times \mathbb{N}$ is countable. We define $\varphi \colon \mathbb{N} \to \mathbb{N} \times \mathbb{N}$ by



Remark

If M is countable infinite, then there exists also a bijective map $\psi : \mathbb{N} \to M$.

Theorem

A countable union $M = \bigcup_{k=1}^{\infty} M_k$ of countable sets M_k is countable.

Corollary

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

Theorem (Cantors second diagonal argument, 1877)

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

The set theory of Cantor extends counting from finite sets to infinite set.

Definition (Cantor)

Two sets M and N have the same cardinality, in notation

$$\mathrm{card}(M) = \mathrm{card}(N),$$

if there exists a bijection $M \to N$. We say that N has at least the cardinality of M, $\mathrm{card}(M) \leq \mathrm{card}(N)$, if there exists a injective map $M \to N$.

Using the axiom of choice of set theory one can prove that $\mathrm{card}(M) \leq \mathrm{card}(N)$ and $\mathrm{card}(N) \leq \mathrm{card}(M)$ implies $\mathrm{card}(M) = \mathrm{card}(N)$.

Axiom of choice. *Let $(M_i)_{i \in I}$ be a family of non-empty sets M_i. Then there exists a map*

$$a : I \to \bigcup_{i \in I} M_i$$

such that $a(i) \in M_i$ for all $i \in I$. In other words, it is possible to select elements $a(i) \in M_i$ simultaneously for all $i \in I$.

One can prove that for any set M, the power set 2^M has always strictly larger cardinality than M.

Infinite series

Definition

Let $(a_k)_{k\in\mathbb{N}}$ be a sequence of real (or complex) numbers. The sequence (s_n) of partial sums

$$s_n = \sum_{k=1}^{n} a_n$$

is called a series, which we denote by

$$\sum_{k=1}^{\infty} a_k.$$

If the sequence (s_n) converges, then

$$\sum_{k=1}^{\infty} a_k = \lim_{n\to\infty} s_n$$

also denotes the limit. Thus $\sum_{k=1}^{\infty} a_k$ is an overloaded notation, which can mean two things

1. the sequence of partial sums, or
2. the limit $\lim_{k\to\infty} s_n$.

Example:

1. $\sum_{k=1}^{\infty} \frac{1}{k}$
2. $\sum_{k=1}^{\infty} \frac{1}{2^k}$
3. (Decimal expansion): $d_k \in \{0, \ldots, 9\}$ digits and

$$\sum_{k=1}^{\infty} d_k \cdot 10^{-k}.$$

   We will see that series of this kind always converge.

4. Let (c_n) be a sequence. Then $a_1 = c_1$ and $a_k = c_k - c_{k-1}$ for $k \geq 2$ defines a series

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

   whose partial sums are $s_n = c_n$. Hence infinite sums is not really a new concept. Sometimes this can be used to compute the limit.

Example (Telescope series): We show

$$\sum_{k=1}^{\infty} \frac{1}{k(k+1)} = 1.$$

Theorem (Cauchy criterion for series)

A series $\sum_{k=1}^{\infty} a_n$ *converges if and only if*

$$\forall \varepsilon > 0 \; \exists n_0 \text{ such that } |\sum_{k=n}^{m} a_k| \leq \varepsilon \; \forall m, n \geq n_0.$$

In particular, the summands (a_k) *of a convergent series* $\sum_{k=1}^{\infty} a_k$ *form a zero sequence.*

Convergence criteria for infinite series

To prove the convergence of a series is often easy and possible without computing the limit.

Definition

An **alternating series** is a series of the form

$$\sum_{k=0}^{\infty}(-1)^k a_k$$

where all $a_k \geq 0$.

Theorem

If (a_k) is a monotone decreasing zero sequence, then

$$\sum_{k=0}^{\infty}(-1)^k a_k$$

converges.

Example: The series

$$\sum_{k=1}^{\infty}(-1)^{k+1}\frac{1}{k} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots$$

and

$$\sum_{k=0}^{\infty}(-1)^{k}\frac{1}{2k+1} = 1 - \frac{1}{3} + \frac{1}{5} - \cdots$$

converge. To compute the limits is much more difficult. In the end of the course we will be able to prove

$$\sum_{k=1}^{\infty}(-1)^{k+1}\frac{1}{k} = \ln 2$$

and

$$\sum_{k=0}^{\infty}(-1)^{k}\frac{1}{2k+1} = \frac{\pi}{4}.$$

Theorem

Let $q \in \mathbb{R}$. The **geometric series** $\sum_{k=0}^{\infty} q^n$ *converges if and only if* $|q| < 1$. *In this case*

$$\sum_{k=0}^{\infty} q^n = \frac{1}{1-q}.$$

Example:

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

Example: As known from high school

$$0.999\ldots = 0.\overline{9} = \sum_{n=1}^{\infty} 9 \cdot 10^{-n} = \frac{9}{10} \sum_{k=0}^{\infty} 10^{-k} = \frac{9}{10} \cdot \frac{1}{1 - \frac{1}{10}} = 1.$$

Example: The series

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

is called the harmonic series. It does not converges.

Remark

If (a_k) is a sequence of non-negative real numbers, then we write sometimes

$$\sum_{k=1}^{\infty} a_k < \infty$$

to indicate that this series converges. This is justified by our first version of the completeness axiom: Since $a_n \geq 0$, the sequence of partial sums (s_n) is monotonously increasing. Hence it is convergent if and only if it stays bounded.

We will work out the idea to compare a series with a simpler one.

Definition

Let $\sum_{n=1}^{\infty} b_n$, $\sum_{n=1}^{\infty} a_n$ be two series. Then $\sum a_n$ **majorizes** $\sum b_n$ if

$$|b_n| \leq a_n \quad \forall n.$$

Theorem (Comparison theorem)

Suppose $\sum a_n$ majorizes $\sum b_n$. Then the following holds

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

Example: The series $\sum_{n=1}^{\infty} \frac{1}{n^2}$ converges. To prove this, it suffices to prove that $\sum_{n=1}^{\infty} \frac{1}{(n+1)^2}$ converges. The telescope series $\sum_{n=1}^{\infty} \frac{1}{n(n+1)}$ is a convergent majorizing series. To compute the limit needs substantially more techniques. With the help of Fourier series one can prove

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}.$$

We deduce further convergence criteria from the comparison theorem.

Theorem (Quotient criterion)

Let $\sum_{n=0}^{\infty} a_n$ *be a series with* $a_n \neq 0 \; \forall n$. *Suppose there exists an* q *with* $0 < q < 1$ *such that*

$$\left|\frac{a_{n+1}}{a_n}\right| \leq q \; \forall n.$$

Then $\sum a_n$ *converges.*

It is enough to ask the bound $|\frac{a_{n+1}}{a_n}| \leq q$ for all but finitely many n. This yields:

Corollary (Quotient test)

Let $\sum_{n=0}^{\infty} a_n$ be a series with $a_n \neq 0 \, \forall n$. Suppose the limit

$$q = \lim_{n \to \infty} |\frac{a_{n+1}}{a_n}|$$

exists.

1. *If $q < 1$ then $\sum_{n=1}^{\infty} a_n$ converges.*
2. *if $q > 1$ then $\sum_{n=1}^{\infty} a_n$ diverges.*
3. *if $q = 1$ then this test yields no information.*

Example:

$$\lim_{n\to\infty} \frac{\frac{1}{n+1}}{\frac{1}{n}} = \lim_{n\to\infty} \frac{n}{n+1} = 1 = \lim_{n\to\infty} \frac{n^2}{(n+1)^2},$$

$\sum_n \frac{1}{n}$ diverges, while $\sum_n \frac{1}{n^2}$ converges.

Theorem (Root test)

Let (a_n) be a sequence and suppose the limit

$$r = \lim_{n\to\infty} \sqrt[n]{|a_n|}$$

exists.

1. *If $r < 1$ then $\sum_{n=1}^{\infty} a_n$ converges.*
2. *if $r > 1$ then $\sum_{n=1}^{\infty} a_n$ diverges.*
3. *if $r = 1$ then this test yields no information.*

Rearrangement of series

Consider the alternating harmonic series

$$\sum_{k=1}^{\infty}(-1)^k\frac{1}{k} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \cdots.$$

By Leibniz' criterium this sum has a limit s satisfying

$$1 = s_1 \geq s \geq s_2 = \frac{1}{2}.$$

Now consider the following rearrangements of the summands

$$1 - \frac{1}{2} - \frac{1}{4} + \frac{1}{3} - \frac{1}{6} - \frac{1}{8} + \frac{1}{5} - \frac{1}{10} - \frac{1}{12} + \cdots,$$

i.e., we consider

$$\sum_{k=1}^{\infty}(\frac{1}{2k-1} - \frac{1}{4k-2} - \frac{1}{4k}).$$

In this series each fraction $\frac{1}{n}$ occurs precisely once with the correct sign. Since $\frac{1}{2k-1} - \frac{1}{4k-2} = \frac{1}{4k-2}$, we obtain

$$\begin{aligned}\sum_{k=1}^{\infty}\left(\frac{1}{2k-1} - \frac{1}{4k-2} - \frac{1}{4k}\right) &= \sum_{k=1}^{\infty}\left(\frac{1}{4k-2} - \frac{1}{4k}\right) \\ &= \frac{1}{2}\sum_{k=1}^{\infty}\left(\frac{1}{2k-1} - \frac{1}{2k}\right) = \frac{1}{2}s \neq s,\end{aligned}$$

since $s \neq 0$. Thus in general an infinite sum has not the same limit if we rearrange its terms. Our next results say that under an additional assumption, arbitrary rearrangements of a series have the same limit.

Definition

A series $\sum_{n=1}^{\infty} a_n$ of real or complex numbers a_n is **absolutely convergent** if the series

$$\sum_{n=1}^{\infty} |a_n|$$

of the absolute values converges.

The triangle inequality and the Cauchy criterium give:

$$\text{absolute convergence} \quad \Longrightarrow \quad \text{convergence.}$$

Moreover in case of absolute convergence, we have a generalised triangle inequality

$$|\sum_{n=1}^{\infty} a_n| \leq \sum_{n=1}^{\infty} |a_n|$$

since for each finite sum
$|\sum_{n=1}^{N} a_n| \leq \sum_{n=1}^{N} |a_n| \leq \sum_{n=1}^{\infty} |a_n| < \infty$ holds.

Theorem

Let $\sum_{n=1}^{\infty} a_n$ be an absolutely convergent series. Then the following holds:

1. *(Little rearrangement theorem.) Let $\sigma : \mathbb{N} \to \mathbb{N}$ be a bijection. Then also $\sum_{n=1}^{\infty} a_{\sigma(n)}$ is absolutely convergent and*

$$\sum_{n=1}^{\infty} a_{\sigma(n)} = \sum_{n=1}^{\infty} a_n.$$

2. *(Big rearrangement theorem.) Let $(I_k)_{k\in\mathbb{N}}$ be a family of finite or infinite disjoint subsets $I_k \subset \mathbb{N}$ such that $\bigcup_{k=1}^{\infty} I_k = \mathbb{N}$. Then the sums $s_k = \sum_{n\in I_k} a_n$ and $\sum_{k=1}^{\infty} s_k$ are absolutely convergent, and*

$$\sum_{n=1}^{\infty} a_n = \sum_{k=1}^{\infty} s_k.$$

Theorem (Cauchy–Product of two series)

Let $\sum_{i=0}^{\infty} a_i$ and $\sum_{j=0}^{\infty} b_j$ be two absolutely convergent series and let (d_k) be defined by

$$d_k = \sum_{i=0}^{k} a_i b_{k-i}.$$

Then also the series $\sum_{k=0}^{\infty} d_k$ is absolutely convergent with limit

$$\sum_{k=0}^{\infty} d_k = \left(\sum_{i=0}^{\infty} a_i\right) \cdot \left(\sum_{j=0}^{\infty} b_j\right).$$

Definition

Let (q_n) be a sequence of numbers. The **infinite product** $\prod_{k=1}^{\infty} q_k$ is convergent, if the limit $\lim_{n\to\infty} \prod_{k=1}^{n} q_k$ of the partial products exists. In that case $q = \prod_{k=1}^{\infty} q_k$ denotes the limit as well.

Theorem (Euler product)

Let p_k denote the k-th prime number and let $s > 1$ denote an integer. Then

$$\prod_{k=1}^{\infty} \frac{1}{1 - p_k^{-s}} = \sum_{n=1}^{\infty} \frac{1}{n^s}.$$

For $s = 1$ the product is divergent. In particular there are infinitely many primes.

Corollary

Let N be a large integer and let ω_N be the probability that two randomly chosen integers $a, b \in \mathbb{N}$ with $0 < a \leq N$, $0 < b \leq N$ have no common factor. Then

$$\lim_{N \to \infty} \omega_N = \frac{6}{\pi^2} = 0.60792\cdots \approx 60\%.$$

Power series

Many important functions are defined via power series.

Definition

Let (a_n) be a sequence of numbers, and $x \in \mathbb{R}$. Then the series

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

is called a power series.

Examples:

1. $\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}$ for $|x| < 1$.
2. The exponential function

$$\exp : \mathbb{R} \to \mathbb{R}$$

is defined by

$$\exp(x) := \sum_{n=0}^{\infty} \frac{x^n}{n!}.$$

This power series converges for all $x \in \mathbb{R}$ by the quotient test:

$$\left| \frac{\frac{x^{n+1}}{(n+1)!}}{\frac{x^n}{n!}} \right| = \frac{|x|}{n+1} \underset{n\to\infty}{\longrightarrow} 0$$

The value $e = \exp(1) = 2.71828\cdots$ is called Euler's number. The notation

$$e^x := \exp(x)$$

is also in use.

3. Likewise we introduce sin and cos as power series

$$\sin(x) := \sum_{k=0}^{\infty}(-1)^k \frac{x^{2k+1}}{(2k+1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots$$

$$\cos(x) := \sum_{k=0}^{\infty}(-1)^k \frac{x^{2k}}{(2k)!} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots$$

The usual properties will be deduced from these formulas later.

The convergence behaviour of power series is best explained if we allow complex numbers.

Proposition

Let $\sum_{n=0}^{\infty} a_n z^n$ be a power series. If the power series converges for $z_0 \in \mathbb{C}$, then it converges absolutely for every element z of the disc

$$\{z \in \mathbb{C} \mid |z| < |z_0|\}.$$

Definition

Let $\sum_{n=0}^{\infty} a_n z^n$ be a power series. Then

$$R := \sup\left\{|z_0| \;\middle|\; \sum_{n=0}^{\infty} a_n z_0^n \text{ converges}\right\} \in [0, \infty]$$

is called the **radius of convergence** of the power series. The power series converges for all z in the disc $\{z \in \mathbb{C} \mid |z| < R\}$ and diverges for all z with $|z| > R$ by the above proposition.

Theorem

Let $(a_n)_{n\in\mathbb{N}_0}$ be a sequence of complex numbers with $a_n \neq 0\ \forall n$. If the limit $q = \lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right|$ exists, then the power series $\sum_{n=0}^{\infty} a_n z^n$ has the radius of convergence

$$R = \begin{cases} \frac{1}{q}, & \text{if } q > 0, \\ \infty, & \text{if } q = 0. \end{cases}$$

The formula does not always apply, e.g. for sin. A formula which always applies is the following:

Definition

Let (b_n) be a sequence real numbers. Then

$$\limsup_{n\to\infty}(b_n) := \lim_{n\to\infty} \sup\left\{ b_k \mid k \geq n \right\}$$

is called the **limit superior** of (b_n). If (b_n) is not bounded from above, then we set

$$\limsup b_n = +\infty.$$

Similarly,

$$\liminf_{n\to\infty} b_n := \lim_{n\to\infty} \inf\{b_k \mid k \geq n\},$$

is called the **limit inferior** of (b_n).

Theorem

Let (a_n) be a sequence of complex numbers and

$$q = \limsup_{n\to\infty}(\sqrt[n]{|a_n|}).$$

Then the power series

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

has radius of convergence

$$R = \begin{cases} 0, & \text{falls } q = \infty, \\ \frac{1}{q}, & \text{falls } 0 < q < \infty, \\ \infty, & \text{falls } q = 0. \end{cases}$$

The complex exponential function

Definition

The complex exponential function

$$\exp : \mathbb{C} \to \mathbb{C}$$

is defined by

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

Theorem (Functional equation of the exponential function)

$$\exp(z + w) = \exp(z) \cdot \exp(w)$$

holds for all $z, w \in \mathbb{C}$.

Corollary

The following holds

1. $\exp(0) = 1$,
2. $\exp(-z) = \frac{1}{\exp(z)}$. *In particular* $\exp(z) \in \mathbb{C}^* \ \forall z \in \mathbb{C}$, *where* $\mathbb{C}^* := \mathbb{C} \backslash \{0\}$.

Evaluating the exponential function at purely imaginary numbers $z = iy$ with $y \in \mathbb{R}$ gives

$$\begin{aligned}
\exp(iy) &= \sum_{n=0}^{\infty} \frac{(iy)^n}{n!} \\
&= \sum_{k=0}^{\infty} (-1)^k \frac{y^{2k}}{2k!} \; + \; i \cdot \sum_{k=0}^{\infty} (-1)^k \frac{y^{2k+1}}{(2k+1)!} \\
&= \cos(y) + i \sin(y),
\end{aligned}$$

i.e., a connection between the complex exponential function and sine and cosine. The functional equation of the complex exponential function implies the addition laws for sine and cosine.

Corollary (Addition laws for sine and cosine)

Let $\alpha, \beta \in \mathbb{R}$. *Then*

$$\begin{aligned} \cos(\alpha+\beta) &= \cos\alpha \cdot \cos\beta - \sin\alpha \cdot \sin\beta, \\ \sin(\alpha+\beta) &= \sin\alpha \cdot \cos\beta + \cos\alpha \cdot \sin\beta. \end{aligned}$$

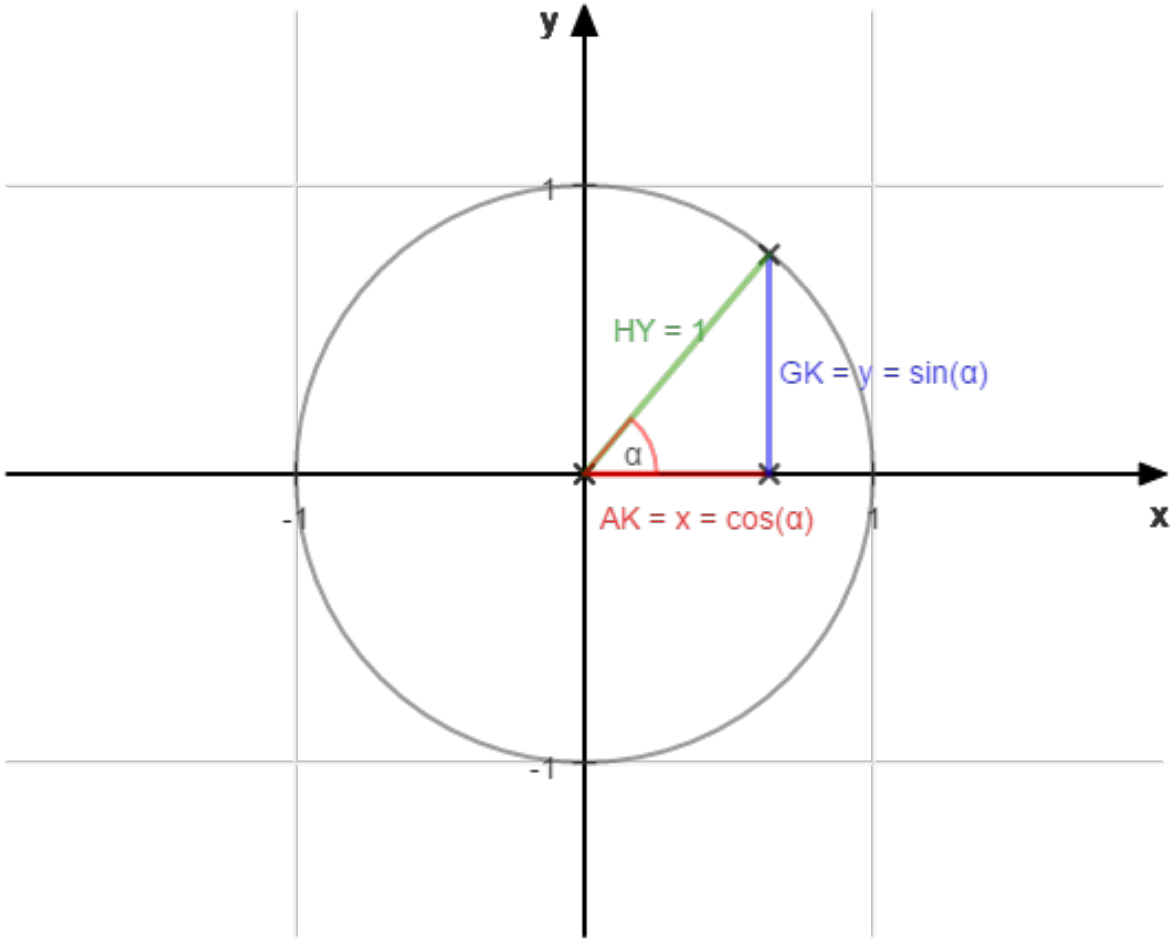
In particular (with the notation $\sin^k \alpha := (\sin\alpha)^k$ *and similarly for* cos*):*

$$1 = \sin^2\alpha + \cos^2\alpha.$$

The last formula says that

$$t \mapsto (\cos t, \sin t)$$

is a parametrization of the unit circle. Here t is measured in radian, as we will see later.



Continuity

Let $D \subset \mathbb{R}$ and $f : D \to \mathbb{R}$ be a real valued function. D is called the domain of definition of f, and typically D is a union of intervals.
The heaviside function $h : \mathbb{R} \to \mathbb{R}$ with

$$h(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } x \geq 0 \end{cases}$$

has a jump at $x = 0$. Roughly speaking, continuous functions are functions which do not have any jumps. We give a precise definition.

Definition

Let $f : D \to \mathbb{R}$ be function on a domain $D \subset \mathbb{R}$. f is called **continuous** at a point $x_0 \in D$ if

$$\forall \varepsilon > 0 \; \exists \, \delta > 0 : \; |f(x) - f(x_0)| < \varepsilon \; \forall x \in D \text{ with } |x - x_0| < \delta.$$

f is called continuous on D, if f is continuous at x_0 for all $x_0 \in D$.

Hence f is not continuous in x_0 if

$$\exists \varepsilon > 0 \text{ such that } \forall \delta > 0 \; \exists x \in D$$

$$\text{with } |x - x_0| < \delta \text{ and } |f(x) - f(x_0)| \geq \varepsilon.$$

Examples:

1. $f(x) = x$ is continuous in all points of $\mathbb{R}$. For $\varepsilon > 0$ given one can take $\delta = \varepsilon$.
2. $f(x) = x^2$ is continuous on $\mathbb{R}$: We have the estimate

$$\begin{aligned} |x^2 - x_0^2| &= |x + x_0| \cdot |x - x_0| \\ &\leq |2x_0 + 1| \cdot |x - x_0| \quad \forall x \text{ with } |x - x_0| < 1. \end{aligned}$$

   Hence $|x^2 - x_0^2| < \varepsilon \ \forall x$ with $|x - x_0| < \delta$, where we choose

$$\delta = \min\left\{1, \ \frac{\varepsilon}{2|x_0| + 1}\right\}.$$

   In this case, δ depends both on ε and x_0.
3. The heaviside function h from above is not continuous in $x_0 = 0$: For any $\varepsilon \leq 1$ and arbitrary small $\delta > 0$ there exists negative numbers $x \in]-\delta, \delta[$. For these x we have

$$|h(x) - h(x_0)| = |0 - 1| = 1 \not< \varepsilon.$$

Theorem (Limit criterion for continuity)

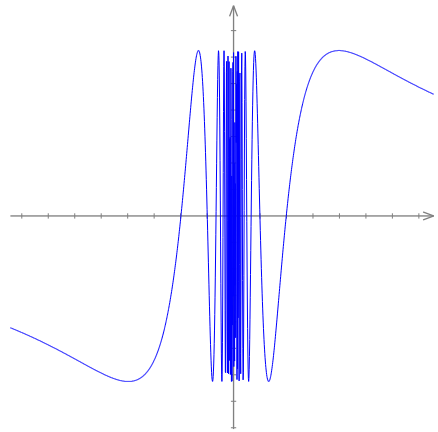
Let $f : D \to \mathbb{R}$ be a function and $x_0 \in D \subset \mathbb{R}$ a point. f is continuous in x_0 if and only if for all sequence $(x_n)_{n\in\mathbb{N}}$ with $x_n \in D$ and $\lim_{n\to\infty} x_n = x_0$ the identity

$$f(x_0) = \lim_{n\to\infty} f(x_n)$$

holds.

Example: Consider the function $f : \mathbb{R} \to \mathbb{R}$ with

$$f(x) = \begin{cases} 0 & \text{for } x = 0 \\ \sin(\frac{1}{x}) & \text{else} \end{cases}$$



As known from high school (we will reprove this later) one has

$$1 = \sin(\frac{1}{2}\pi) = \sin(\frac{1}{2}\pi + 2\pi n)$$

for $n \in \mathbb{Z}$. Thus $x_n = \frac{1}{\frac{1}{2}\pi + 2\pi n}$ is a sequence with $x_n \underset{n\to\infty}{\longrightarrow} 0$ and $f(x_n) = 1$. This shows that f is not continues in $x_0 = 0$. Similarly, $x'_n = \frac{1}{\frac{3}{2}\pi + 2\pi n}$ is a sequence with $x'_n \underset{n\to\infty}{\longrightarrow} 0$ and $f(x'_n) = -1$. Hence there is no way to give $f(0)$ a value, which would make the function continuous.

Theorem

Let $f, g\colon D \to \mathbb{R}$ be functions.

1. *If f, g are continuous in x_0, then $f + g$ and $f \cdot g$ are also continuous in x_0*
2. *If f, g are continuous in x_0 and $g(x_0) \neq 0$, then*

$$\frac{f}{g}\colon D' \to \mathbb{R}$$

with $D' = \{x \in D \mid g(x) \neq 0\} \subset D$ is also continuous in $x_0 \in D'$

Corollary

1. *Polynomials*

$$f(x) = a_n x^n + \ldots + a_1 x + a_0$$

with coefficients $a_k \in \mathbb{R}$ *define continuous functions* $f : \mathbb{R} \to \mathbb{R}$.

2. *A rational functions , i.e., maps* $\frac{f}{g} : D \to \mathbb{R}$ *with* f, g *defined by polynomials, are continues on the domain* $D = \{x \in \mathbb{R} \mid g(x) \neq 0\}$.

Intermediate value theorem and applications

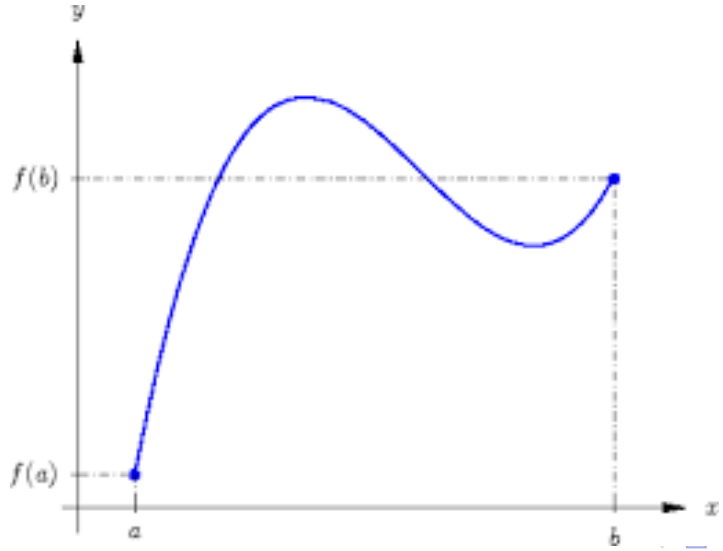
An important property of continuous function is the following:

Theorem (Intermediate value theorem)

Let $f : [a, b] \to \mathbb{R}$ be a continuous function, and let c be a value between $f(a)$ und $f(b)$, i.e., $f(a) \leq c \leq f(b)$ or $f(a) \geq c \geq f(b)$. Then there exists an $\xi \in [a, b]$ such that

$$f(\xi) = c.$$

In particular, every continuous function f with $f(a) < 0$ and $f(b) > 0$ has a zero in $[a, b]$.



Theorem (Existence of a maximum and a minimum)

Let $f : [a, b] \to \mathbb{R}$ be a continuous function on a closed bounded intervall. Then there exist $x_{\max}, x_{\min} \in [a, b]$ with

$$\begin{aligned} f(x_{\max}) &= \sup\left\{ f(x) \mid x \in [a, b] \right\}, \\ f(x_{\min}) &= \inf\left\{ f(x) \mid x \in [a, b] \right\}. \end{aligned}$$

In particular, f is bounded.

We say: f takes its **maximum** in $x_{\max}$ and write

$$\max_{x \in [a,b]} f(x) := f(x_{\max}).$$

Similarly for the **minimum**: $\min_{x \in [a,b]} f(x) := f(x_{\min})$.

Remark

The condition that $[a, b]$ is a bounded closed interval is essential. The function $f :]0, \infty[\to \mathbb{R}$ with $f(x) = \frac{1}{x}$ has neither a maximum nor a minimum.

Definition

A function : $I \to \mathbb{R}$ on an interval I **grows monotonously** (or is **monotone increasing**) if points $x_1 < x_2$ in I the values satisfy $f(x_1) \leq f(x_2)$. It is **strictly** monotone increasing, if $f(x_1) < f(x_2)$. **Monotone decreasing** and **strictly monotone decreasing** are defined similarly. f is (strictly) monotone if it is (strictly) monotone increasing or (strictly) monotone decreasing.

Proposition

1. *Let $f : I \to \mathbb{R}$ be a continuous function on an interval I. Then $J = f(I) \subset \mathbb{R}$ is a interval as well.*
2. *If in addition f is strictly monotone, then $f : I \to J$ is bijective, and the inverse map $f^{-1} : J \to I$ is also continuous.*

Definition

Let $f : D \to \mathbb{R}$ be a function and let $a \in \mathbb{R} \setminus D$ be a point such that there exists a sequence (x_n) in D with $\lim x_n = a$. If for each sequence (x_n) in D with $\lim x_n = a$ there exists the limit $b = \lim_{n\to\infty} f(x_n)$, and if these limits are equal for all possible sequences, then the common limit

$$\lim_{x\to a} f(x) = b$$

is called the limit of $f(x)$ when x approaches a. The notation

$$\lim_{x\nearrow a} f(x)$$

is used if we consider only sequence (x_n) with $\lim x_n = a$ and $x_n < a$. The limit from above $\lim_{x\searrow a} f(x)$ is defined similarly.

Example: The rational function $f(x) = \frac{x^2-1}{x-1}$ is only defined on $D = \mathbb{R} \setminus \{1\}$. However, since

$$\lim_{x \to 1} \frac{x^2 - 1}{x - 1} = \lim_{x \to 1} (x + 1) = 2,$$

we can extend this function to the continuous function $\tilde{f} : \mathbb{R} \to \mathbb{R}$ defined by $\tilde{f}(x) = x + 1$.

Differentiation

Definition

Let $f : I \to \mathbb{R}$ be a function on an interval I and let $x_0 \in I$ be a point. We call f **differentiable** at x_0 if the limit

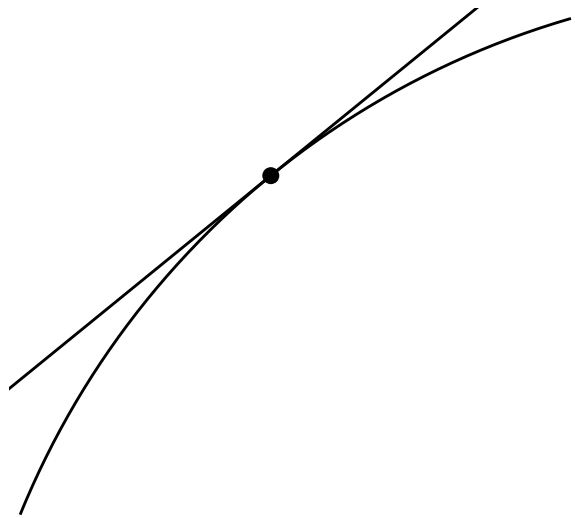
$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

exists.

Geometrically, we can interpret the difference quotient

$$\frac{f(x) - f(x_0)}{x - x_0}$$

as the slope of the secant line through the points $(x, f(x))$ and $(x_0, f(x_0))$ on the graph G_f. Hence the limit may be interpreted as the slope of the tangent line to G_f at $(x_0, f(x_0))$. Thus f is differentiable x_0, if we can associate in a sensible way a tangent line to G_f at $(x_0, f(x_0)$.



Definition

A function $f : I \to \mathbb{R}$ is called differentiable on I if f is differentiable at every point of I. In that case the function

$$f' : I \to \mathbb{R}$$

defined by

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

is called the **derivative** of f. Thus $f'(x_0)$ is the slope of the tangent to G_f at $(x_0, f(x_0))$. In case f' is continuous, we call f **continuously differentiable**.

Remark

1. In Physics the derivative is basic to even define the concept of speed. If $f : I \to \mathbb{R}, t \mapsto f(t)$ describes the motion of a point on a line, then $f'(t)$ is the speed at time t.
2. The notation $f'(x)$ (and $\dot{f}(t)$ for derivates after time) goes back to Newton. Leibniz used the notation $\frac{df}{dx}(x)$.

Proposition

Let $f : I \to \mathbb{R}$ be a function on an interval and let $x_0 \in I$. Then

$$f \text{ differentiable in } x_0 \implies f \text{ continuous in } x_0.$$

Examples:

1. $f(x) = x^2$ is differentiable, since

$$\lim_{x \to x_0} \frac{x^2 - x_0^2}{x - x_0} = \lim_{x \to x_0} (x + x_0) = 2x_0.$$

   Hence $f'(x) = 2x$.
2. Constant functions $f(x) = c \ \forall x$ have derivative 0.
3. Linear functions $f(x) = mx + c$ have derivative $f'(x) = m$, a constant function.

Theorem (Calculation rules for derivatives)

Let $f, g : I \to \mathbb{R}$ be functions on an interval which are differentiable at $x_0 \in I$. Then

1. *$f + g : I \to \mathbb{R}$ is differentiable in x_0 with*

$$(f + g)'(x_0) = f'(x_0) + g'(x_0).$$

2. **Product rule.** *$fg : I \to \mathbb{R}$ is differentiable in x_0 with*

$$(fg)'(x_0) = f'(x_0)g(x_0) + f(x_0)g'(x_0).$$

3. **Quotient rule.** *If $g(x_0) \neq 0$ then*

$$\frac{f}{g} : \{x \in I \mid g(x) \neq 0\} \to \mathbb{R}$$

is differentiable in x_0 with

$$(\frac{f}{g})'(x_0) = \frac{f'g - fg'}{g^2}(x_0).$$

Example: The function $f(x) = x^n$ is for arbitrary $n \in \mathbb{Z}$ on its domain of definition differentiable with the exponent rule

$$f'(x) = nx^{n-1}.$$

Rational functions $r = \frac{f}{g}$ are differentiable on their domain of definition. If f and g have no common factor, then a zero of g are called a **pole** of r.

Theorem (Chain rule)

Suppose $f : I \to \mathbb{R}$ *and* $g : J \to \mathbb{R}$ *are functions with* $f(I) \subset J$. *If* f *is differentiable at* x_0 *and* g *is differentiable at* $f(x_0)$, *then the composition* $g \circ f : I \to \mathbb{R}$ *is differentiable at* x_0 *and*

$$(g \circ f)(x_0) = g'(f(x_0))f'(x_0).$$

The factor $f'(x_0)$ *in this formula is called the* **inner derivative**.

Example We consider $f(x) = x^2 + 1$ and $g(x) = x^3$. The composition is $(g \circ f)(x) = (x^2 + 1)^3$. Using the chain rule, we obtain

$$(g \circ f)'(x) = 3(x^2 + 1)^2 2x.$$

First expanding and then taking the derivative,

$$(x^6 + 3x^4 + 3x^2 + 1)' = 6x^5 + 12x^3 + 6x,$$

gives the same result.

Theorem (Derivative of the inverse function)

Let $f : I \to \mathbb{R}$ be a strictly monotone function, and denote by $J = f(I)$. If f is differentiable in x_0 with $f'(x_0) \neq 0$, then the inverse function $f^{-1} : J \to I \subset \mathbb{R}$ is differentiable in $y_0 = f(x_0)$ and

$$(f^{-1})'(y_0) = \frac{1}{f'(f^{-1}(y_0))} = \frac{1}{f'(x_0)}.$$

Remark

A way to memorise the formula is to use the chain rule:
$f^{-1} \circ f = \mathrm{id}_I$ gives

$$1 = (f^{-1})'(f(x_0))f'(x_0) = (f^{-1})'(y_0)f'(x_0).$$

Example: The k-**th root** $g(x) = \sqrt[k]{x} = x^{1/k}$, $k \in \mathbb{N}$, is defined as the inverse function of the strictly monotonous function $f : \mathbb{R}_{\geq 0} \to \mathbb{R},\ f(x) = x^k$. We have $f'(x) = kx^{k-1} \neq 0$ for $x > 0$. Hence $g : \mathbb{R}_{\geq 0} \to \mathbb{R}$ is differentiable on $\mathbb{R}_{>0}$ with

$$g'(x) = \frac{1}{k(\sqrt[k]{x})^{k-1}} = \frac{1}{k}x^{\frac{1-k}{k}} = \frac{1}{k}x^{\frac{1}{k}-1}.$$

Again the exponent rule is valid.

Lokal extrema and the mean value theorem

The derivative is often used to find local maxima or minima.

Definition

Let $f : I \to \mathbb{R}$ be a function on an interval I, and let $x_0 \in I$. We say, f has in x_0 a **local maximum**, if $\exists\ h > 0$, such that $]x_0 - h, x_0 + h[\subset I$ and

$$f(x_0) \geq f(x) \quad \forall x \in]x_0 - h, x_0 + h[$$

holds. **Local minima** are defined similarly. A **local extremum** is a local maximum or local minimum.
If

$$f(x_0) > f(x)\ \forall x \in]x_0 - h, x_0 + h[,\ x \neq x_0,$$

holds, then we speak of an **isolated maximum**. Isolated minima and isolated extrema are defined accordingly.

An **absolute maximum** or **global maximum** is a point $x_0 \in I$, such that

$$f(x_0) \geq f(x) \ \forall x \in I$$

holds. **Absolute Minima** and **absolute extrema** are defined similarly.

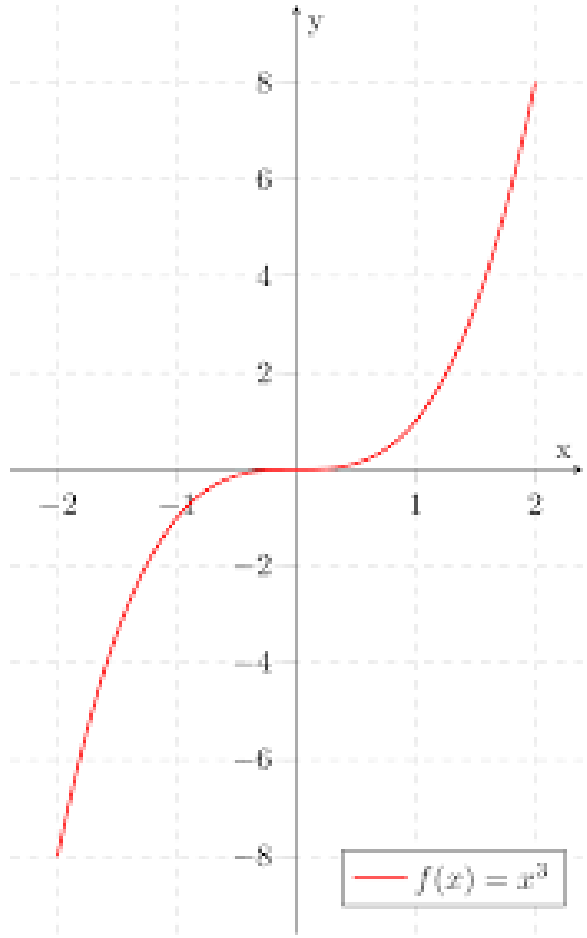
Theorem

Let $f :]a, b[\to \mathbb{R}$ be a differentiable function. If $x_0 \in]a, b[$ is a local extremum of f then

$$f'(x_0) = 0.$$

Remark

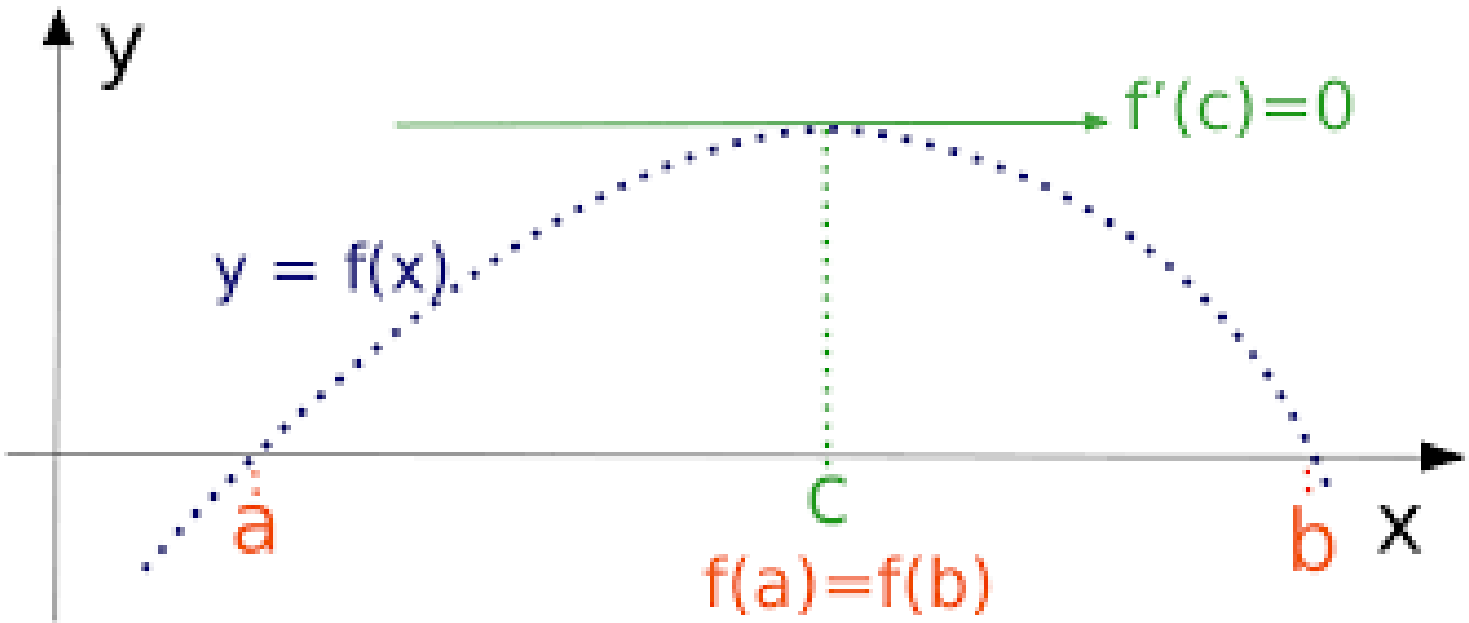
$f'(x) = 0$ is a necessary but not sufficient condition for a local extremum for a differentiable function. For example, the function $f(x) = x^3$ satisfies $f'(0) = 0$, but $x_0 = 0$ is not a local extremum.



Theorem (Rolle's theorem)

Let $f : [a, b] \to \mathbb{R}$ be a continuous function with $f(a) = f(b)$, which is differentiable on $]a, b[$. Then there exists an $\xi \in]a, b[$ with

$$f'(\xi) = 0.$$

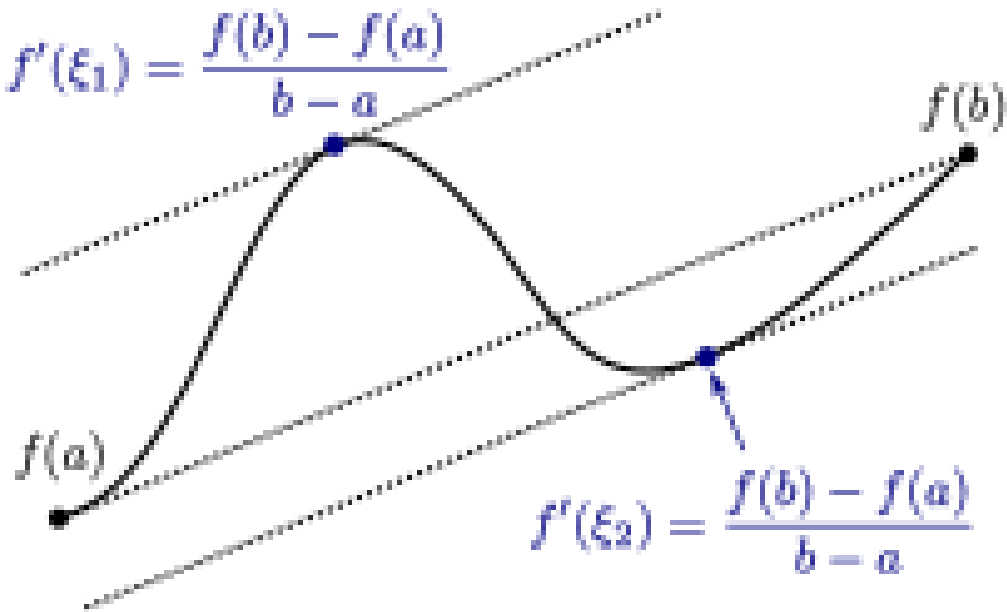


Theorem (Mean value theorem)

Let $f : [a, b] \to \mathbb{R}$ be a continues function, which is differentiable on $]a, b[$. Then there exists an $\xi \in]a, b[$ such that

$$\frac{f(b) - f(a)}{b - a} = f'(\xi)$$

.



Corollary

Let $f : [a, b] \to \mathbb{R}$ be continuous and in $]a, b[$ differentiable. Suppose we have constants $m, M \in \mathbb{R}$ such that $m \leq f'(x) \leq M$ for all $x \in]a, b[$. Then

$$m(x_2 - x_1) \leq f(x_2) - f(x_1) \leq M(x_2 - x_1)$$

holds for all $x_1 < x_2$ in $[a, b]$.

Corollary

Let $f : [a, b] \to \mathbb{R}$ be continuous and in $]a, b[$ differentiable. Suppose $\forall x \in]a, b[$ we have

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

Higher derivatives

Definition

Let $f : I \to \mathbb{R}$ be differentiable. f is twice differentiable, f' is differentiable again. In that case $f'' := f^{(2)} := (f')'$ denotes the second derivative.
Recursively, we define that f is n-times differentiable, if $f^{(n-1)}$ is differentiable, and then

$$f^{(n)} = (f^{(n-1)})'$$

denotes the n-th derivative.

Theorem (Sufficient criterion for local extrema)

Let $f :]a, b[\to \mathbb{R}$ be two-times differentiable, and let $x_0 \in]a, b[$. If $f'(x_0) = 0$ and $f''(x_0) \neq 0$, then f has an isolated local extremum at x_0.
It is a local maximum if $f''(x_0) < 0$ and a local minimum, if $f''(x_0) > 0$.

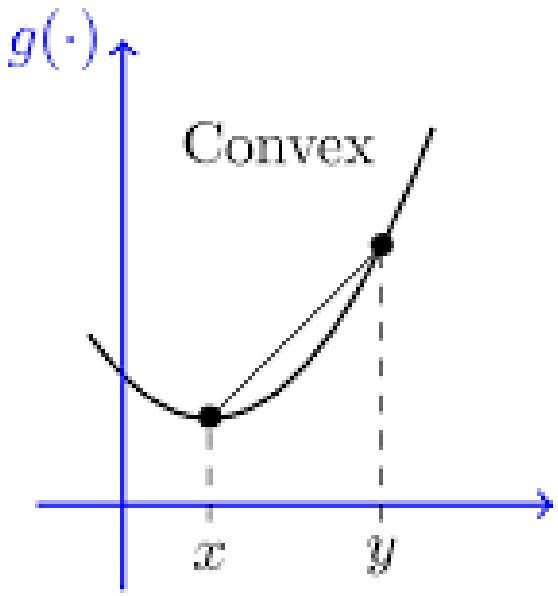
Example: Let $f(x) = x^2$. Then $f'(x) = 2x$ and $f''(x) = 2 > 0$. The function f has a local minimum at $x_0 = 0$.

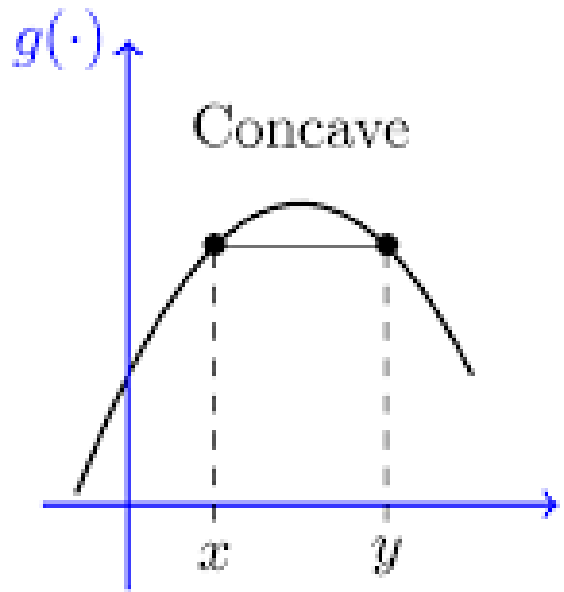
Definition

Let $I \subset \mathbb{R}$ be an interval. A function $f : I \to \mathbb{R}$ is called **convex** if $\forall x_1, x_2 \in I$ and all λ with $0 \leq \lambda \leq 1$ we have

$$f(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f(x_1) + (1 - \lambda)f(x_2).$$

f is called **concave**, if $-f$ is convex.



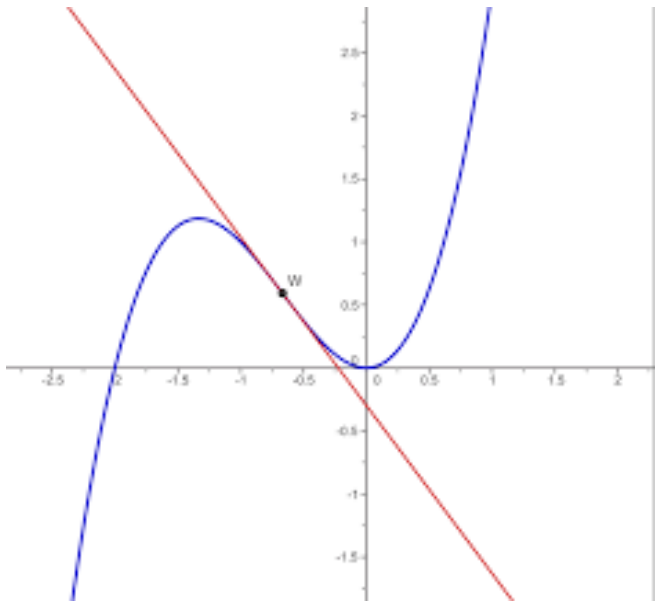


Theorem

Let $f : I \to \mathbb{R}$ be a twice differentiable function on an interval. f is convex if and only if $f''(x) \geq 0 \; \forall x \in I$.

Remark

If $f : I \to \mathbb{R}$ is three times differentiable, then points $x_0 \in I$ with $f'''(x_0) = 0$ are called a **flex** of f. Typically but not always, flexes are points where f'' changes sign. If this is the case, then the function turns form convex to concave or from concave to convex at x_0.



Newton method

The result above allows to make curve discussions for sufficiently often differentiable functions f. One issue which we have not discussed, is how to find zeroes of the function f and f'. The Newton method gives an answer.
Consider a point x_k with a non horizontal tangent. Then

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$$

is the intersection point of a tangent line with the x-axis.

Theorem (Newton method)

Let $f : [a, b] \to \mathbb{R}$ be a two times differentiable convex function with $f(a) < 0$ and $f(b) > 0$. Then the following holds.

1. *There is a unique zero $\xi \in [a, b]$ of f.*
2. *If $x_0 \in [a, b]$ is an arbitrary start value with $f(x_0) \geq 0$, then the sequence (x_n) with $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$ converges monotonously decreasing to ξ.*
3. *If $f'(x) \geq c > 0$ and $f''(x) < k \ \forall x \in [\xi, b]$, then we have the following estimate*

$$|x_{n+1} - x_n| \leq |\xi - x_n| \leq \frac{k}{2c} |x_n - x_{n+1}|^2.$$

Hence the Newton method converges **quadratically**.

Remark

For $f(x) = x^2 - a$ the Newton method recovers our algorithm to approximate the square root of a:

$$x_{k+1} = x_k - \frac{x_k^2 - a}{2x_k} = \frac{1}{2}(x_k + \frac{a}{x_k}).$$

The exponential function

We introduced the exponential function with a power series:

$$e^x = \exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

and established $e^{x_1+x_2} = e^{x_1}e^{x_2}$, $e^{-x} = \frac{1}{e^x}$, $e = \exp(1)$. What is the derivative of exp?

$$\frac{\exp(x) - 1}{x - 0} = \sum_{n=1}^{\infty} \frac{x^{n-1}}{n!} \quad \underset{x \to 0}{\to} \quad ?$$

Lemma (Error term of the exponential series)

For $N \in \mathbb{N}$ we define $r_{N+1}(x)$ by the identity

$$\exp(x) = \sum_{n=0}^{N} \frac{x^n}{n!} + r_{N+1}(x).$$

Then the estimate

$$|r_{N+1}(x)| \leq 2\frac{|x|^{n+1}}{(n+1)!} \quad \text{for } |x| \leq 1 + \frac{N}{2}$$

holds.

Theorem

The exponential function $\exp : \mathbb{R} \to \mathbb{R}$ *is differentiable with* $\exp'(x) = \exp(x)$.

Corollary

The exponential function exp *is strictly monotone increasing and convex.*

The reason for the frequent occurring of the exponential function in nature is that $y = e^{cx}$ is a solution of the differential equation $y' = cy$. More precisely:

Proposition

Let $f : I \to \mathbb{R}$ be a differentiable function on an interval satisfying $f' = cf$. Then

$$f(x) = f(x_0)e^{c(x-x_0)},$$

where $x_0 \in I$ is an arbitrary fixed point.

The map $\exp\colon \mathbb{R} \to \mathbb{R}_{>0}$ is bijektiv and is a so called *isomorphism of groups* $(\mathbb{R}, +) \to (\mathbb{R}_{>0}, \cdot)$ because $\exp(x + y) = \exp(x) \cdot \exp(y)$ for all $x, y \in \mathbb{R}$. Groups will be a topic in the next term.

The Logarithm

Definition

The inverse function

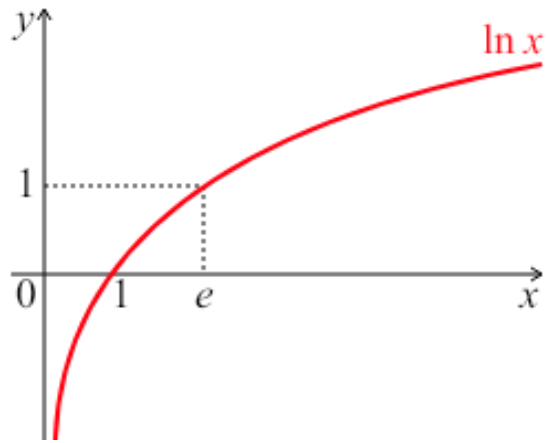
$$\ln : \mathbb{R}_{>0} \to \mathbb{R}$$

of exp is called the **natural logarithm**.

Proposition (Properties of the logarithm)

The following holds:

1. $\ln(x_1 \cdot x_2) = \ln x_1 + \ln x_2$.
2. ln *is differentiable with* $(\ln x)' = \frac{1}{x}$.
3. ln *is concave und monotone increasing*



Definition

Let $a \in \mathbb{R}_{>0}$ be a positive real number. We define the **exponentiation with base** a by

$$a^x := e^{x \cdot \ln a}.$$

Theorem

The following holds:

1. $a^{x_1+x_2} = a^{x_1} \cdot a^{x_2}$.
2. *If* $x = \frac{p}{q} \in \mathbb{Q}$ *is a rational number, then* $a^x = \sqrt[q]{a^p}$. *Hence* a^x *extends our previous notation* $a^{\frac{p}{q}} = \sqrt[q]{a^p}$.
3. *The function* $x \mapsto a^x$ *is differentiable with* $(a^x)' = \ln a \cdot a^x$.

Definition

For $a \in \mathbb{R}_{>0}$ we denote by

$$\log_a \colon \mathbb{R}_{>0} \to \mathbb{R}$$

the inverse function of $x \mapsto a^x$.

Remark

$\log_a x$ is differentiable with

$$(\log_a)'(x) = \frac{1}{\ln a \cdot a^{\log_a x}} = \frac{1}{x \ln a}.$$

In computer science $\log_2 n$ is important: $\lfloor \log_2 n \rfloor + 1$ is the number of **binary digits** of an integer $n > 0$, i.e., the number of bits needed to specify n.

Functions like $x \mapsto x^x = e^{x \ln x}$ show up in complexity theory: One of the best algorithmen known to factor an integer n with $x = \log_2 n$ binary digits has run time $O(e^{\frac{1}{2} x \ln x})$.

Trigometric functions

We have defined sine and cosine by power series.

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

Moreover, we already showed the addition laws and $\sin^2 x + \cos^2 x = 1 \ \forall x \in \mathbb{R}$.

Theorem

The functions $\sin, \cos \colon \mathbb{R} \to \mathbb{R}$ *are differentiable with*

$$\sin' = \cos, \quad \cos' = -\sin.$$

We next define Archimedian's constant π. To define π as the ratio of the diameter to the circumference of a circle or to prove that the unit disc has area π, one needs integration. We will come back to this later.

Proposition

The following holds

1. $\cos(0) = 1$, $\cos(2) < 0$.
2. $\sin(x) > 0$ *for* $x \in]0, 2[$.
3. *The function* cos *has a unique zero in* $[0, 2]$.

We now define π as follows:

Definition

We define **Archimedean's constant** π as the unique real number such that $\frac{\pi}{2}$ is the zero of cos in $[0, 2]$.

Note, that

$$\cos\left(\frac{\pi}{2}\right) = 0 \implies \sin\left(\frac{\pi}{2}\right) = 1.$$

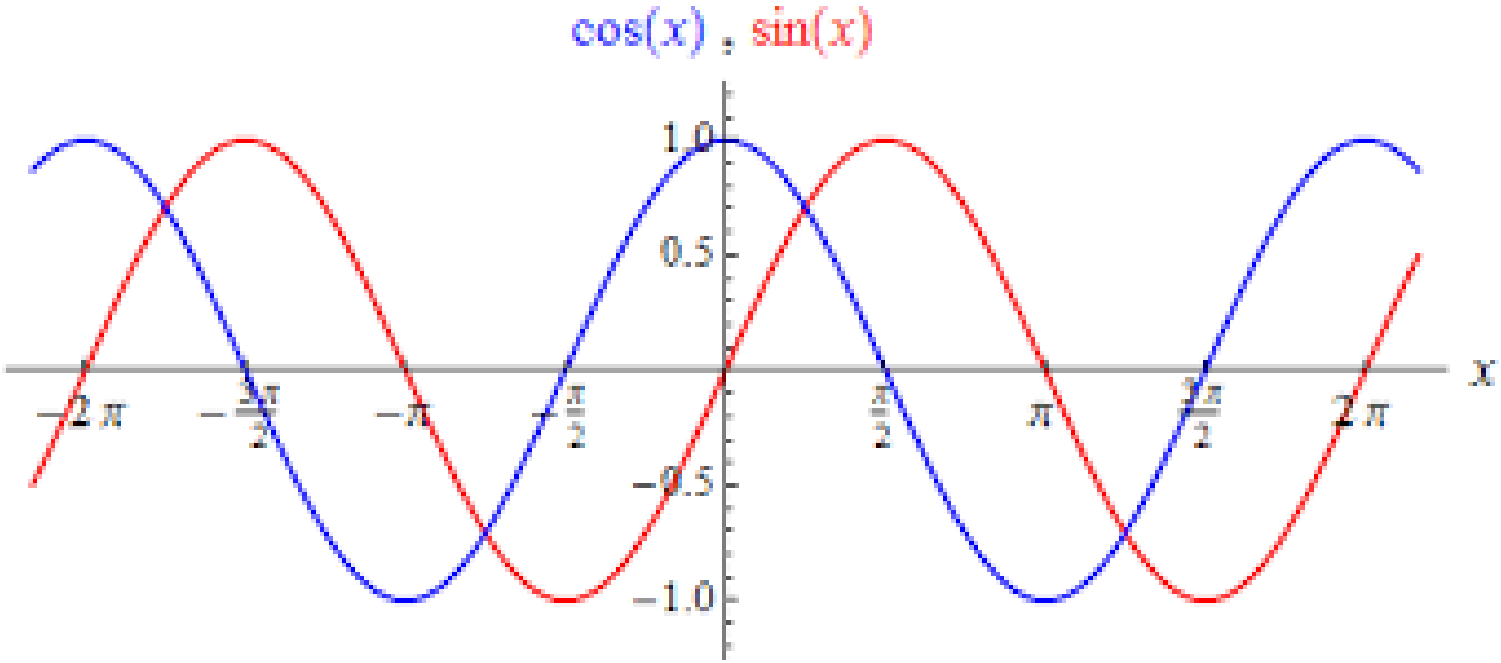
Hence the addition laws give the following identities:

Theorem (Phase shift of sine and cosine)

1. $\sin(x+\frac{\pi}{2}) = \cos x, \;\; \cos(x+\frac{\pi}{2}) = -\sin x.$
2. $\sin(x+\pi) = -\sin x, \;\; \cos(x+\pi) = -\cos x.$
3. $\sin(x+2\pi) = \sin x, \;\; \cos(x+2\pi) = \cos x.$

□

We say sin and cos are 2π-periodic functions. The value of π is $3.1415\ldots$.



Remark (Importance of sine und cosine)

1. $[0, 2\pi[\to \mathbb{R}^2, \quad t \mapsto (\cos t, \sin t)$ is a parametrization of the unit circle.
2. If f is a solution of the differential equation

$$y'' = -w^2 y,$$

then $f(x) = a\cos(wx) + b\sin(wx)$. Hence sine und cosine occur in oscillations.

Definition

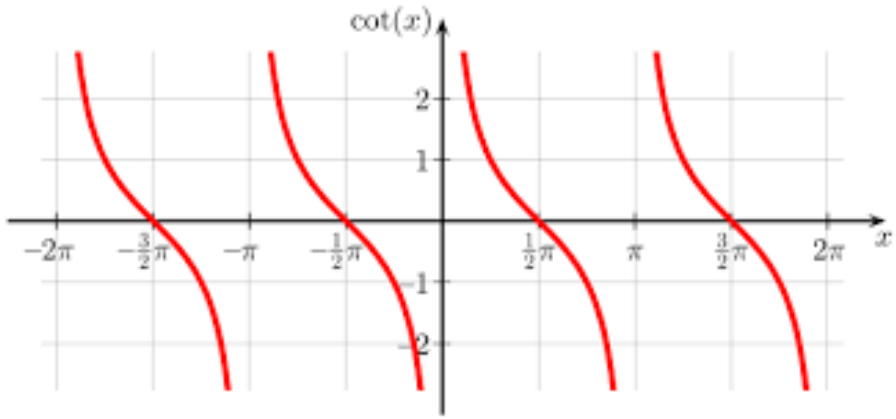
The function

$$\tan\colon \mathbb{R} \setminus \left\{\frac{\pi}{2} + \pi k \mid k \in \mathbb{Z}\right\} \to \mathbb{R}, \quad x \mapsto \tan x = \frac{\sin x}{\cos x}$$

is called **tangent function**. Its reciprocal

$$\cot\colon \mathbb{R} \setminus \{k\pi \mid k \in \mathbb{Z}\} \to \mathbb{R}, \quad x \mapsto \cot x = \frac{1}{\tan x} = \frac{\cos x}{\sin x}$$

is the **cotangent**.



Definition

1. $\tan\colon]-\frac{\pi}{2}, \frac{\pi}{2}[\to \mathbb{R}$ is strictly increasing. The inverse function of tan

$$\arctan\colon \mathbb{R} \ \to \ ]-\frac{\pi}{2}, \frac{\pi}{2}[\ \subset \ \mathbb{R}$$

   is called **arc tangent**.

2. The map

$$\sin\colon \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \ \to \ [-1, 1]$$

   is strictly increasing. The inverse function

$$\arcsin\colon [-1, 1] \ \to \ \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$$

   is called **arc sine**.

Theorem

1. arcsin *is differentiable on* $]-\frac{\pi}{2}, \frac{\pi}{2}[$ *with*

$$\arcsin'(x) = \frac{1}{\sqrt{1-x^2}}.$$

2. arctan *is differentiable with*

$$\arctan'(x) = \frac{1}{1+x^2}.$$

Asymptotic behavior and L' Hospital's rule

Theorem (L'Hospital's rule)

Let $f, g : [a, b] \to \mathbb{R}$ *be continuous functions which are differentiable on* $]a, b]$ *with* $g'(x) \neq 0 \ \forall x \in]a, b[$ *and* $f(a) = g(a) = 0$.
If the limit $\lim_{x \searrow a} \frac{f'(x)}{g'(x)}$ *exists, then the limit* $\lim_{x \searrow a} \frac{f(x)}{g(x)}$ *also exists and*

$$\lim_{x \searrow a} \frac{f(x)}{g(x)} = \lim_{x \searrow a} \frac{f'(x)}{g'(x)}$$

holds.

Example: $f(x) = \sin(x)$, $g(x) = e^x - 1$, $[a, b] = [0, 1]$. The quotient $\frac{f(0)}{g(0)} = \frac{0}{0}$ makes no sense, but

$$\frac{f'(x)}{g'(x)} = \frac{\cos x}{e^x}$$

is continuous in $x = 0$ with

$$\lim_{x \searrow 0} \frac{f'(x)}{g'(x)} = \frac{\cos 0}{e^0} = \frac{1}{1} = 1.$$

Hence

$$\lim_{x \searrow 0} \frac{\sin x}{e^x - 1} = \lim_{x \searrow 0} \frac{\cos x}{e^x} = \frac{1}{1} = 1.$$

There are various variants of L'Hospital's rule. To formulate them we need some notation.

Definition (Convergence to ∞)

Let $f : [a, \infty[\to \mathbb{R}$ be a function. We say, $f(x)$ **approaches** $c \in \mathbb{R}$ for x to ∞, in symbols

$$\lim_{x \to \infty} f(x) = c,$$

if $\forall \varepsilon > 0 \; \exists \; N : \; |f(x) - c| < \varepsilon \; \forall x \geq N$. We say:
$\lim_{x \to \infty} f(x) = \infty$, if

$$\forall M > 0 \; \exists \; N > 0 : \; f(x) > M \; \forall x > N.$$

For $f :]a, b] \to \mathbb{R}$ we write $\lim_{x \searrow a} f(x) = \infty$, if

$$\forall M > 0 \; \exists \; \varepsilon > 0 : \; f(x) > M \; \forall x > a \text{ with } |x - a| < \varepsilon.$$

Similarly we define $\lim_{x \to -\infty} f(x) = c$ or $\lim_{x \nearrow b} f(x) = -\infty$.

Theorem (Variants of L'Hospital's rule)

1. *Let* $f, g :]a, b[\to \mathbb{R}$ *be differentiable functions with* $g'(x) \neq 0 \ \forall x \in]a, b[$ *and*

$$\lim_{x \searrow a} f(x) = \infty = \lim_{x \searrow a} g(x).$$

If the limit $\lim_{x \searrow a} \frac{f'(x)}{g'(x)}$ *exist, then* $\lim_{x \searrow a} \frac{f(x)}{g(x)}$ *exists and*

$$\lim_{x \searrow a} \frac{f(x)}{g(x)} = \lim_{x \searrow a} \frac{f'(x)}{g'(x)}.$$

2. *Let* $f, g : [a, \infty[\to \mathbb{R}$ *be differentiable functions with* $g'(x) \neq 0 \ \forall x \in [a, \infty[$ *and*

$$\lim_{x \to \infty} f(x) = 0 = \lim_{x \to \infty} g(x)$$

or

$$\lim_{x \to \infty} f(x) = \infty = \lim_{x \to \infty} g(x).$$

Theorem (continued)

If the limit $\lim_{x\to\infty} \frac{f'(x)}{g'(x)}$ *exists, then* $\lim_{x\to\infty} \frac{f(x)}{g(x)}$ *exists and*

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

Example: We show for all $n \in \mathbb{N}$

$$\lim_{x\to\infty} \frac{x^n}{e^x} = 0.$$

Since $\lim_{x\to\infty} x^n = \lim_{x\to\infty} e^x = \infty$. We can try to apply L'Hospital's rule.

$$\lim_{x\to\infty} \frac{x^n}{e^x} = \lim_{x\to\infty} \frac{nx^{n-1}}{e^x} = \cdots = n! \lim_{x\to\infty} \frac{1}{e^x} = 0.$$

Hence the exponential function grows faster then any polynomial.

The O-notion will also be used for functions:

Definition (O– and o– notation for functions)

Let $f, g : [a, \infty[\to \mathbb{R}$ be functions. We write

$$f \in O(g) \text{ for } x \to \infty,$$

if $\exists\ c > 0\ \exists\ M$, such that $|f(x)| \leq c \cdot g(x)\ \forall x \geq M$, and say f **lies in big O of** g. We say $f \in o(g)$, f **lies in small o of** g, if $\lim_{x\to\infty} \frac{f(x)}{g(x)} = 0$.

Example:

1. $x^n \in o(e^x)$ for $x \to \infty$ for every $n \in \mathbb{N}$ as shown above.
2. Let $f(x) = a_n x^n + \cdots + a_0 \in \mathbb{R}[x]$ be a polynomial of degree n. Then $f(x) \in O(x^n)$ for $x \to \infty$. More precisely: For every $C = |a_n| + \varepsilon, \varepsilon > 0, \ \exists\ M > 0$, such that

$$|f(x)| \leq C \cdot x^n \ \forall x \geq M.$$

Asymptotic behaviour of rational functions

Let $h(x) = \frac{f(x)}{g(x)}$ be a rational function with polynomials

$$\begin{aligned} f(x) &= a_n x^n + \cdots + a_0 \in \mathbb{R}[x], \\ g(x) &= b_m x^m + \cdots + b_0 \in \mathbb{R}[x], \end{aligned}$$

of degree n and m respectively, i.e., $a_n, b_m \neq 0$. L'Hospital's rule gives

$$\begin{aligned} \lim_{x\to\infty} \frac{f(x)}{g(x)} &= 0, \quad \text{if } n < m, \\ \lim_{x\to\infty} \frac{f(x)}{g(x)} &= \frac{a_n}{b_m}, \quad \text{if } n = m, \\ \lim_{x\to\infty} \frac{f(x)}{g(x)} &= \begin{cases} +\infty, & \text{if } n > m \text{ and } \frac{a_n}{b_m} > 0, \\ -\infty, & \text{if } n > m \text{ and } \frac{a_n}{b_m} < 0. \end{cases} \end{aligned}$$

Then last case can be refined.

Theorem (Division with remainder)

Let $f, g \in \mathbb{R}[x]$ be polynomials in one variable x with real coefficients. Then there exist uniquely determined polynomials $q(x), r(x) \in \mathbb{R}[x]$, such that

$$f(x) = q(x) \cdot g(x) + r(x) \quad \textit{and} \quad \deg r < \deg g.$$

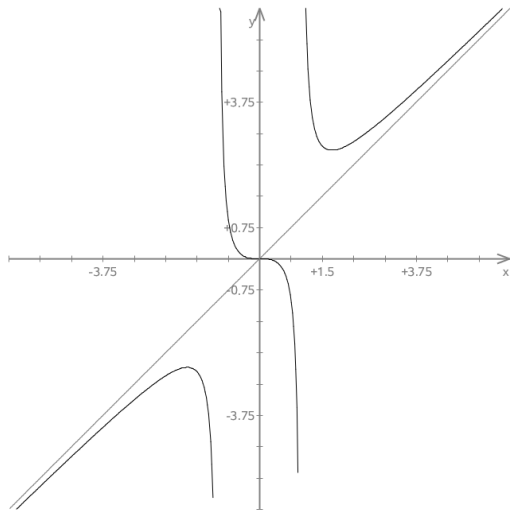
Example: Consider

$$h(x) = \frac{f(x)}{g(x)} = \frac{x^3}{x^2 - 1}.$$

Long division gives

$$\begin{array}{l} x^3 : (x^2 - 1) = x + \frac{x}{x^2-1}, \\ \underline{x^3 - x} \\ \quad x \end{array}$$

i.e., $q(x) = x$, $r(x) = x$. Hence $h(x) \in x + o(1)$, i.e., asymptotically $h(x)$ has the same behaviour as x. Near 0 the function h is very different from $x \mapsto x$. For example, h has poles at $x = \pm 1$ and a saddle point at 0.



In general we have

Corollary

Let $h(x) = \frac{f(x)}{g(x)}$ *be a rational function and*

$$f(x) = q(x) \cdot g(x) + r(x)$$

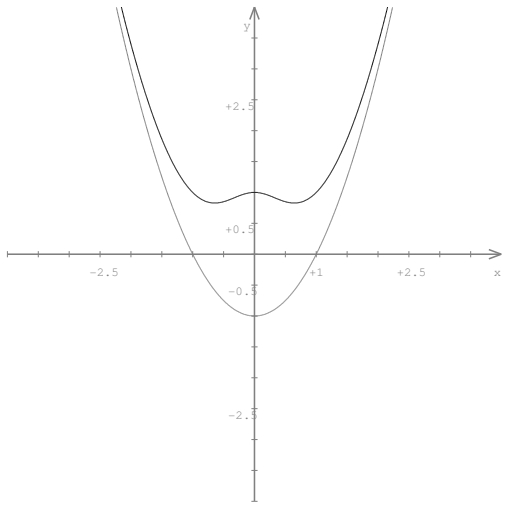
with $\deg r < \deg g$. *Then* h *behaves like* $q(x)$ *for* $x \to \pm\infty$, *more precisely*

$$h(x) \in q(x) + o(1).$$

Example: We consider the rational function $h(x) = \frac{x^4+1}{x^2+1}$. Long division gives

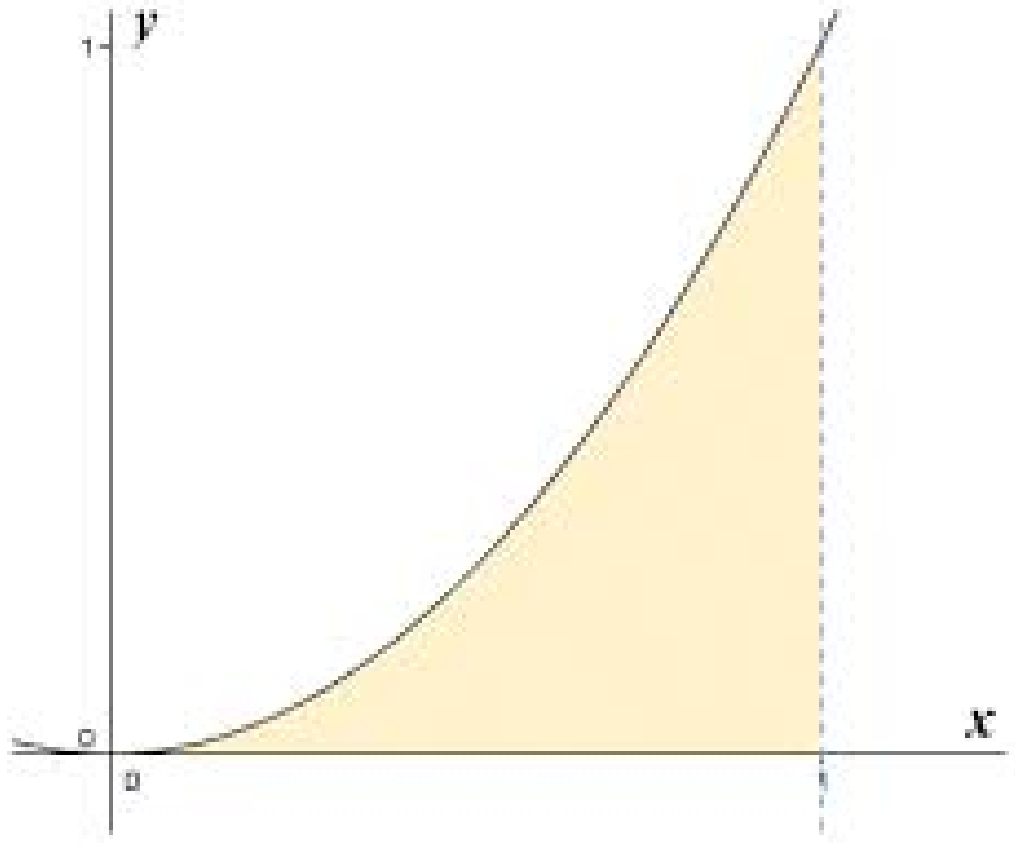
$$
\begin{array}{l}
(x^4+1) : (x^2+1) = x^2 - 1 + \frac{2}{x^2+1}, \\
\underline{x^4 + x^2} \\
\quad -x^2 + 1 \\
\quad \underline{-x^2 - 1} \\
\qquad\qquad 2
\end{array}
$$

i.e., $q = x^2 - 1$ und $r = 2$. Since $x^2 + 1$ has no zero in $\mathbb{R}$ the function $h(x)$ has no poles. To draw the graph, we compute the the extrema of $h(x)$.



Integration

Let $f : [a, b] \to \mathbb{R}$ be a function on a closed interval. We want to compute the area under the graph of f.

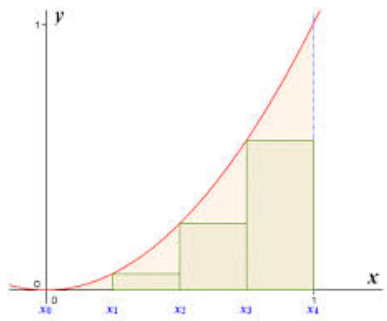


The basic idea is to use an approximation by staircase functions.

Definition

A **staircase function** $\varphi\colon [a,b] \to \mathbb{R}$ is a function, such that there exists a subdivision $a = t_0 < t_1 < \cdots < t_n = b$ of the interval $[a,b]$, such that φ is constant on the open intervals $]t_{i-1}, t_i[$, i.e., for each $i \in \{1,\ldots,n\}$ there is a $c_i \in \mathbb{R}$, such that

$$\varphi|_{]t_{i-1},t_i[}\colon\ ]t_{i-1}, t_i[\to \mathbb{R}, \quad \varphi|_{]t_{i-1},t_i[}(x) = \varphi(x) = c_i.$$



For $\varphi(t_i)$ we require nothing. The **integral of the staircase functions** is

$$\int_a^b \varphi(x)\ dx := \sum_{i=1}^{n} c_i(t_i - t_{i-1}).$$

Definition (continued)

Each sum of this kind is called a **Riemann sum**. Using these, the integral of an arbitrary bounded function $f : [a, b] \to \mathbb{R}$ is defined as follows: The **upper integral** of f ist

$$\int_a^{*b} f := \inf \left\{ \int_a^b \psi \, dx \mid \psi \geq f, \ \psi \text{ a staircase function} \right\},$$

and the **lower integral** is

$$\int_{*a}^b f := \sup \left\{ \int_a^b \varphi \, dx \mid \varphi \leq f, \ \varphi \text{ a staircase function} \right\}.$$

f is integrable (more precisely: **Riemann-integrable**), if $\int_a^{*b} f = \int_{*a}^b f$ holds. If this is the case, then we define the **integral of the bounded function** $f : [a, b] \to \mathbb{R}$ by

$$\int_a^b f(x) \, dx := \int_a^{*b} f = \int_{*a}^b f.$$

Example: Staircase functions are integrable.

Theorem

Monotone functions $f : [a, b] \to \mathbb{R}$ *are integrable.*

Example: Suppose $0 \leq b$. What is $\int_0^b x^2 \, dx$?

Example: The function

$$f : [0,1] \to \mathbb{R},\ f(x) = \begin{cases} 1, & \text{for } x \in \mathbb{Q}, \\ 0, & \text{for } x \notin \mathbb{Q}, \end{cases}$$

is not Riemann integrable. For each pair φ, ψ of staircase functions with $\varphi \leq f \leq \psi$ we have

$$\int_0^1 \varphi(x)\ dx \leq 0, \quad \int_0^1 \psi(x)\ dx \geq 1,$$

since each interval $]t_{i-1}, t_i[$ contains rational and irrational points.

Theorem (Integrability of continuous function)

Let $f : [a, b] \to \mathbb{R}$ be a continuous function on a bounded closed interval. Then f is integrable on $[a, b]$.

For the proof we need more than pointwise continuity.

Definition

A function $f : I \to \mathbb{R}$ is **uniformly continuous** on I, if $\forall \varepsilon > 0$ $\exists\, \delta > 0$, such that

$$|f(x_1) - f(x_0)| < \varepsilon \ \forall\ x_0, x_1 \in I \text{ with } |x_1 - x_0| < \delta.$$

The crucial difference to pointwise continuity, is that δ does not depend on x_0.

Example: The functions

$$f : \mathbb{R}_{>0} \to \mathbb{R},\ f(x) = \frac{1}{x} \text{ and } g : \mathbb{R} \to \mathbb{R},\ g(x) = x^2$$

are continuous but not uniformly continuous.

Theorem

Let $f : [a, b] \to \mathbb{R}$ be a continuous function on a bounded closed interval. Then f is uniformly continuous.

Theorem (Properties of the integrals)

Let $f, g : [a, b] \to \mathbb{R}$ be integrable functions, and let $c \in \mathbb{R}$ be a constant. Then the following holds:

1. *(**Linearity of the integral**) The functions $c \cdot f$ and $f + g$ are integrable with*

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

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

2. *(**Monotonicity of the integral**)*
   $f \leq g \implies \int_a^b f(x)\, dx \leq \int_a^b g(x)\, dx.$
3. *With f also the functions $f_+ := \max(f, 0)$, $f_- := \max(-f, 0)$ and $|f| = f_+ + f_-$ are integrable.*
4. *For $p \in \mathbb{R}_{>0}$ also $|f|^p$ is integrable. In particular the product*

$$f \cdot g = \frac{1}{4}\big(|f + g|^2 - |f - g|^2\big)$$

   is integrable.

Theorem (Mean value theorem of integral calculus)

Let $f, g : [a, b] \to \mathbb{R}$ *be functions,* f *continuous,* g *integrable and* $g(x) \geq 0 \ \forall x$. *Then there exists* $\xi \in [a, b]$, *such that*

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

In particular: $\exists\ \xi \in [a, b]$, *such that*

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

Antiderivative

The key mathematical discovery of Newton (and Leibniz) was that for $f : [a, b] \to \mathbb{R}$ the integrals

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

for any $t \in [a, b]$ can be computed simultaneously using antiderivatives.

Proposition

Let $f : [a, b] \to \mathbb{R}$ be an integrable function. Then f is also integrable over each closed subinterval of $[a, b]$ and

$$\int_a^t f(x)\ dx + \int_t^b f(x)\ dx = \int_a^b f(x)\ dx$$

holds for all $t \in]a, b[$.

Definition

Let $f : [a, b] \to \mathbb{R}$ be an integrable function. We set

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

for interchanged lower and upper boundary points.

Definition

Let $f : I \to \mathbb{R}$ be a continuous function on an interval. A differentiable function $F : I \to \mathbb{R}$ is called a antiderivative of f, if $F' = f$.

Theorem (Fundamental theorem of calculus)

Let $f : I \to \mathbb{R}$ be a continuous function on an interval and let $a \in I$ be a point.

1. *$F : I \to \mathbb{R}$ defined by $F(x) = \int_a^x f(t)\, dt$ is an antiderivative of f.*
2. *If $G : I \to \mathbb{R}$ is an antiderivative of f, then*

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

for any subinterval $[a, b] \subset I$. Two short hand notations are in use

$$G(x)\Big|_a^b := \big[G(x)\big]_a^b := G(b) - G(a).$$

Definition

The **indefinite integral** $\int f(x)\,dx$ denotes an antiderivative of f.

Examples: Our computation of derivatives gives the following basic examples.

1. $\int x^{\alpha}\,dx = \frac{x^{\alpha+1}}{\alpha+1}, \quad \alpha \neq -1.$
2. $\int \frac{1}{x}\,dx = \ln|x|.$
3. $\int e^{x}\,dx = e^{x}.$
4. $\int \sin x\,dx = -\cos x.$
5. $\int \cos x\,dx = \sin x.$
6. $\int \frac{1}{1+x^2}\,dx = \arctan x.$
7. $\int \frac{1}{\sqrt{1+x^2}}\,dx = \arcsin x.$

Each differentiation rule gives a rule for computing integrals. The chain rule gives the following:

Theorem (Substitution rule)

Let $f : I \to \mathbb{R}$ be continuous, $\varphi : [a, b] \to I$ continuously differentiable and $\alpha = \varphi(a), \beta = \varphi(b)$. Then the following holds

$$\int_{\alpha}^{\beta} f(x)\, dx = \int_{a}^{b} f(\varphi(t)) \cdot \varphi'(t)\, dt.$$

Examples: A frequent application of the substitution rule is the following.

1. let $g\colon [a,b] \to \mathbb{R}$ be differentiable with $g(t) \neq 0\ \forall t$. Then

$$\int \frac{g'(t)}{g(t)}\ dt = \ln|g(t)|.$$

   With the notation of the substitution rule we have $f(x) = \frac{1}{x}$ and $\varphi = g$.

2. For the tangent function this gives

$$\int \tan x\ dx = \int \frac{\sin x}{\cos x}\ dx = -\int \frac{-\sin x}{\cos x}\ dx = -\ln|\cos x|.$$

Remark

A way to memorise the substitution rule is to use Leibniz notation:

$$x = \varphi(t), \quad \frac{dx}{dt} = \varphi'(t),$$

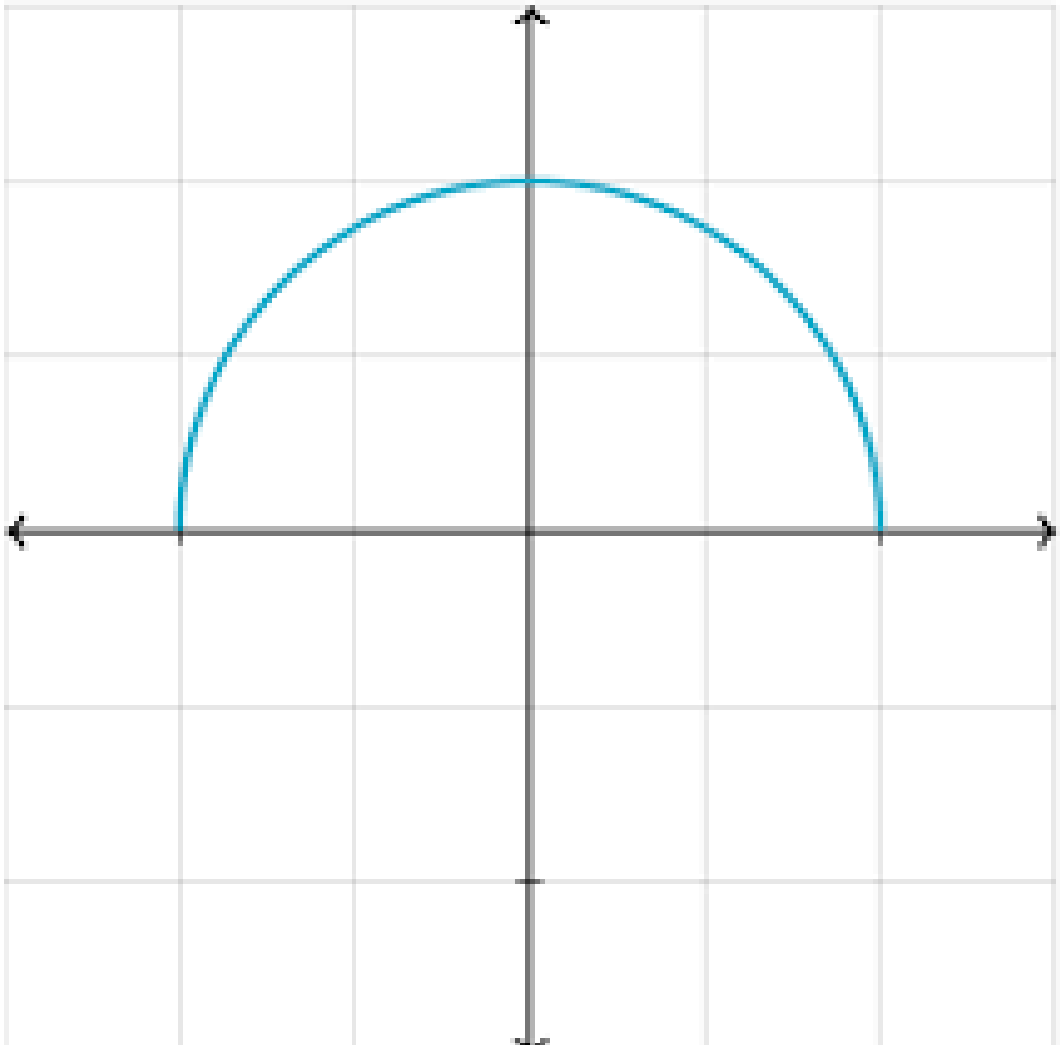
hence "$dx = \varphi'(t)\ dt$". We cannot give a precise definition, what "dx" mathematically is without more theory, but it gives the right result: If we replace x by $\varphi(t)$ and dx by $\varphi'(t)\ dt$ in $F(x) = \int f(x)\ dx$, then we obtain

$$F(\varphi(t)) = \int f(\varphi(t)) \cdot \varphi'(t)\ dt.$$

Example: We prove

$$\frac{\pi}{2} = \int_{-1}^{1} \sqrt{1 - x^2}\, dx$$

establishing that the circle with radius 1 has indeed area π.



From the product rule $(fg)' = f'g + fg'$ one deduces

Theorem (Integration by parts)

Let $f, g : I \to \mathbb{R}$ be differentiable functions. Then

$$\int f'(x)g(x)\,dx = f(x)g(x) - \int f(x)g'(x)\,dx$$

and

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

holds. □

Examples:

1. $$\int_0^{\pi} x \sin x \, dx = (-x \cos x) \Big|_0^{\pi} + \sin x \Big|_0^{\pi} = -\pi \cdot (-1) = \pi,$$

   since in integration by parts we may choose
   $g(x) = x$, $g'(x) = 1$ and $f'(x) = \sin x$, i.e., $f(x) = -\cos x$.
2. To calculate $\int e^{-x} \sin x \, dx$ we use integration by parts twice.
3. We show, that $\int_0^{2\pi} \sin^2 x \, dx = \pi$.
4. Similarly, $\int_0^{2\pi} \sin x \cos x \, dx = \frac{1}{2} \sin^2 x \Big|_0^{2\pi} = 0$.

To give a closed formula for indefinite integrals can be difficult.

Definition

The set of **elementary functions** is the smallest set of functions satisfying

1. x^n, $\sin x$, $\tan x$, e^x and their inverse functions are elementary.
2. Sums, products and quotients of elementary functions are elementary.
3. Compositions of elementary functions are elementary.

Theorem (Liouville's theorem)

Not every elementary function has a elementary antiderivative.

The proof of this theorem is far beyond the material of this course. Explicit examples are the functions e^{-x^2} and $\frac{1}{\sqrt{x^3-x}}$.

On the positive side one has

Theorem

Rational functions have elementary antiderivatives.

Example: We compute

$$\int \frac{1}{1-x^2}\, dx.$$

Improper Integrals

Definition

Let $f : [a, \infty[\to \mathbb{R}$ be a continuous function. We define the **improper integral**

$$\int_a^\infty f(x)\ dx := \lim_{b \to \infty} \int_a^b f(x)\ dx,$$

in case the limit exists. In this case f is called integrable over $[a, \infty[$, and we say that $\int_a^\infty f(x)\ dx$ converges.

Example: For $f(x) = x^{-s}$ for $s \in \mathbb{R}$ the limit

$$\int_1^\infty x^{-s}\, dx = \lim_{b\to\infty} \left(\frac{1}{1-s} x^{1-s} \Big|_1^b\right) = \lim_{b\to\infty} \frac{1}{1-s}(1 - b^{1-s})$$

exists, if and only if $s > 1$. In this case

$$\int_1^\infty x^{-s}\, dx = \frac{1}{1-s}.$$

Theorem (Integral criterion for convergence of series)

Let $f : [0, \infty[\to \mathbb{R}$ be a monotone decreasing non-negative function. The sum $\sum_{n=0}^{\infty} f(n)$ converges if and only if the integral $\int_0^{\infty} f(x)\, dx$ converges.

Corollary

The limit

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$$

exists for $s > 1$.

$\zeta(s)$ is called the **Riemannsche Zetafunktion**. It satisfies the Euler product

$$\prod_{p \text{ prime}} \frac{1}{1 - p^{-s}} = \sum_{n=1}^{\infty} \frac{1}{n^s}.$$

Definition

Let $f :]a, b]$ be a continuous function on a semi-open interval. We define

$$\int_a^b f(x)\ dx := \lim_{t \searrow a} \int_t^b f(x)\ dx,$$

in case the limit exist.

Example:

$$\int_0^1 \frac{1}{\sqrt{x}} dx = \lim_{t \searrow 0} \left[\frac{1}{2} x^{\frac{1}{2}}\right]_t^1 = \frac{1}{2}.$$

More general, $\int_0^1 x^{-s} dx$ converges if $s < 1$, because $\int x^{-s} dx = \frac{1}{1-s} x^{1-s}$. In contrast $\int_0^1 \frac{1}{x}\ dx$ does not converge, because $\ln x \underset{x \to 0}{\to} -\infty$.

Definition

A function $f : \mathbb{R} \to \mathbb{R}$ is integrable over $[-\infty, \infty]$ if the limits

$$\lim_{b\to\infty} \int_0^b f(x)\ dx \quad \text{and} \quad \lim_{a\to-\infty} \int_a^0 f(x)\ dx$$

exist. In this case we write

$$\int_{-\infty}^{\infty} f(x)\ dx := \lim_{a\to\infty} \int_a^0 f(x)\ dx + \lim_{b\to\infty} \int_0^b f(x)\ dx.$$

Examples:

1. $$\int_{-\infty}^{\infty} \frac{1}{1+x^2}\ dx = \lim_{a\to-\infty,\ b\to\infty} \left(\arctan x \Big|_a^b\right) = \pi.$$

2. The limit $\int_{-\infty}^{\infty} e^{-x^2}\ dx$ exists, because $e^{-x^2} \in O\left(\frac{1}{1+x^2}\right)$. In the third semester we will prove

$$\int_{-\infty}^{\infty} e^{-x^2}\ dx = \sqrt{\pi}.$$

Taylor polynomial

Let $f : I \to \mathbb{R}$ be an n–times continuous differentiable function, and let $x_0 \in I$. We want to approximate f near x_0 by a polynomial. The "best" approximation by a linear polynomial is the tangent

$$L(x) = f(x_0) + f'(x_0) \cdot (x - x_0).$$

If we allow higher degree polynomials we get the Taylor polynomials.

Definition

Let $f : I \to \mathbb{R}$ be an n–times continuous differentiable function, and let $x_0 \in I$. Then

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

is called the n–th **Taylor polynomial** of f in x_0.

Apparently, $T^n_{x_0} f$ is the unique polynomial of degree $\leq n$ which has the same derivatives up to order n in x_0 as f.

Theorem (Taylor formula)

Let $f : I \to \mathbb{R}$ be an $(n+1)$–times continuous differentiable function, and let $x_0 \in I$ be a point. Then

$$f(x) = (T^n_{x_0} f)(x) + R_{n+1}(x) \ \left(= \sum_{k=0}^{n} \frac{f^{(k)}(x_0)(x - x_0)^k}{k!} + R_{n+1}(x)\right)$$

with error term

$$R_{n+1}(x) = \int_{x_0}^{x} f^{(n+1)}(t) \frac{(x-t)^n}{n!} \, dt.$$

In details:

$$\begin{aligned} f(x) &= f(x_0) + f'(x_0)(x - x_0) + \frac{f^{(2)}(x_0)}{2}(x - x_0)^2 + \cdots \\ &\quad + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n + \int_{x_0}^{x} f^{(n+1)}(t) \frac{(x-t)^n}{n!} dt. \end{aligned}$$

Theorem (Lagrange's form of the error term)

Let $f : I \to \mathbb{R}$ be an $(n+1)$–times continuous differentiable function, and let $x_0 \in I$ be a point. Then $\exists\, \xi \in [x_0, x]$ if $x > x_0$ or $\xi \in [x, x_0]$ if $x < x_0$ such that

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

Example: $f(x) = \sin x$, $x_0 = 0$.

$$(T_0^{2n+1} \sin)(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots + (-1)^n \frac{x^{2n+1}}{(2n+1)!},$$

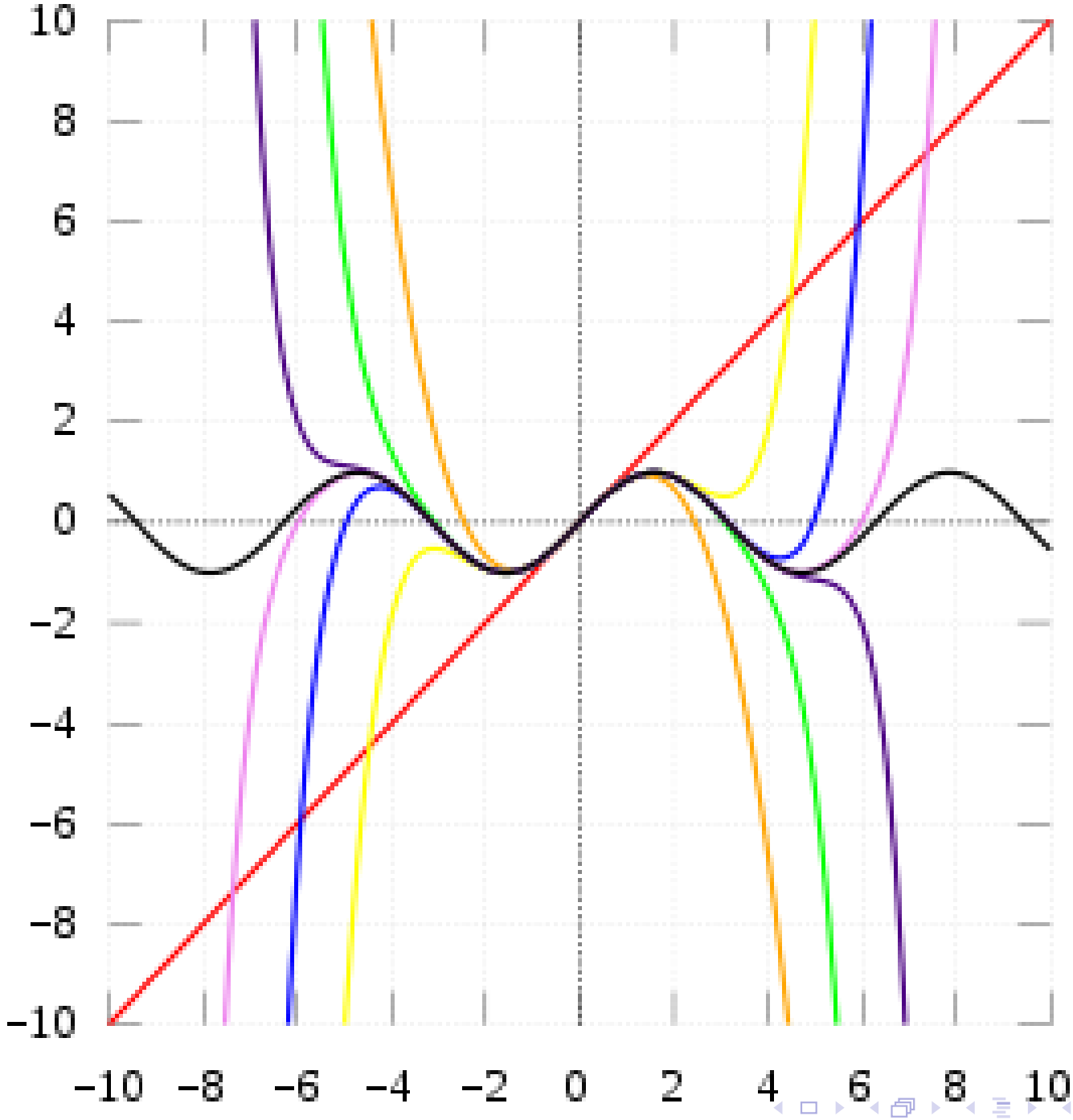
since

$$\sin^{(k)}(0) = \begin{cases} 0, & k \text{ even}, \\ (-1)^{\frac{k-1}{2}}, & k \text{ odd}. \end{cases}$$

We estimate the error

$$|R_{n+1}(x)| \quad = \quad |f^{(n+1)}(\xi)| \cdot \frac{|x|^{n+1}}{(n+1)!} \leq \frac{|x|^{n+1}}{(n+1)!} \leq \frac{R^{n+1}}{(n+1)!}$$

for $|x| \leq R$.


10
8
6
4
2
0
−2
−4
−6
−8
−10
−10 −8 −6 −4 −2 0 2 4 6 8 10

Example: Consider the function $f(x) = (1+x)^\alpha$, for example $\alpha = \frac{1}{2}$. Since

$$f^{(k)}(x) = \alpha \cdot (\alpha - 1) \cdots (\alpha - k + 1) \cdot (1+x)^{\alpha - k},$$

we deduce

$$T_0^n f(x) = \sum_{k=1}^{n} \binom{\alpha}{k} x^k,$$

where for $\alpha \in \mathbb{R}$ the **binomial coefficients** are defined by

$$\binom{\alpha}{k} := \frac{\alpha}{1} \cdot \frac{\alpha - 1}{2} \cdots \frac{\alpha - k + 1}{k}.$$

Definition

Let $f : I \to \mathbb{R}$ be an infinitely often differentiable function, and let $x_0 \in I$ be a point. Then

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

is called the **Taylor series** of f with **expansion point** x_0. The Taylor series $T_{x_0} f$ is a power series in $(x - x_0)$.

Example:

1. $(T_0 \exp)(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!}$ is the defining power series of exp.
2. The function $f(x) = (1+x)^\alpha$ has the Taylor series

$$(T_0 f)(x) = \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k.$$

Question: Does the Taylor series of f converges towards f?
The answer is negative in general:

1. A Taylor series has not necessarily a positive radius R of convergence.
2. Even if $R > 0$, the Taylor series does not necessarily converges to f on $]-R+x_0, x_0+R[$.

Example: We give an example of the second kind. Consider

$$f : \mathbb{R} \to \mathbb{R},\ x \mapsto f(x) = \begin{cases} e^{-\frac{1}{x^2}}, & x \neq 0, \\ 0, & \text{else.} \end{cases}$$

We show: f is ∞–often differentiable and $f^{(n)}(0) = 0\ \forall n$. In particular, the Taylor series is 0, and does not converge to f.

By the answers of the above question the following result is less trivial than might look like on the first glance.

Theorem (Binomial series)

For $|x| < 1$ *the following holds:*

$$(1+x)^{\alpha} = \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k.$$

Uniform convergence

Let

$$f(x) = \sum_{n=0}^{\infty} a_n x^n$$

be a power series with radius of convergence $R > 0$. We want to prove that f is ∞–often differentiable on $]-R, R[$ and that the power series coincides with the Taylor series of f in $x_0 = 0$. Already continuity is not obvious.

Definition

Let $f_n : I \to \mathbb{R}$ be a sequence of function on an interval. The sequence (f_n) **converges** (more precisely: **converges pointwise**), if for every x the sequence $(f_n(x))$ converges. In this case we call

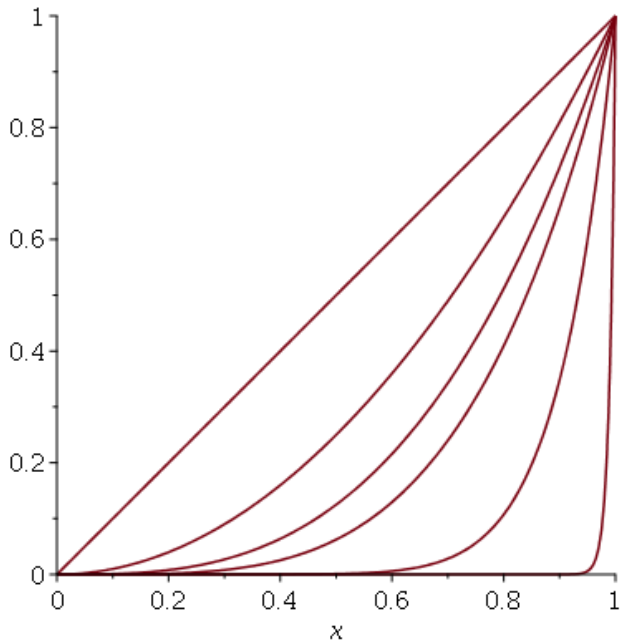
$$f : I \to \mathbb{R}, \ f(x) = \lim_{n\to\infty} f_n(x)$$

the **limit function** and we write: $\lim_{n\to\infty} f_n = f$.

Question: Is $\lim f_n$ continuous, if all f_n continuous?
The answer is no in general. For example, for
$f_n\colon [0,1] \to \mathbb{R},\ f_n(x) = x^n$ the limit $f = \lim f_n$ exists, but

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

f is not continuous. One needs a stronger assumption on the convergence.



Definition

Let $(f_n : I \to \mathbb{R})_{n\in\mathbb{N}}$ be a sequence of functions on an interval I. (f_n) **converges uniformly** to a limit function $f : I \to \mathbb{R}$, if

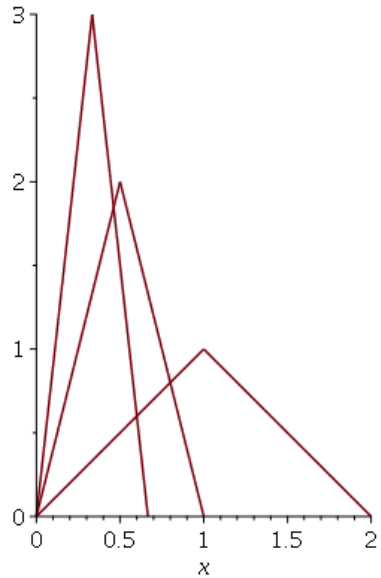
$$\forall \varepsilon > 0 \; \exists \; n_0 : \; |f_n(x) - f(x)| < \varepsilon \; \forall n \geq n_0 \; \forall \; x \in I.$$

Theorem (Uniform limit of continuous functions)

If $(f_n : I \to \mathbb{R})$ is a sequence of continuous functions, which converges uniformly to f, then also f is continuous.

Question: Can one interchange integration with taking the limit?
No, not necessarily.
Example: If $f_n : [0,1] \to \mathbb{R}$ is defined by

$$f(x) = \begin{cases} n^2 x & \text{if } 0 \le x \le \frac{1}{n} \\ n^2(\frac{2}{n} - x) & \text{if } \frac{1}{n} < x \le \frac{2}{n} \\ 0 & \text{else .} \end{cases}$$



Then $\int_0^1 f_n(x)\ dx = 1\ \forall\ n \ge 2$ und $\lim_{n\to\infty} f_n = 0$ pointwise, but

$$\lim_{n\to\infty} \int_0^1 f_n(x)\ dx = 1 \neq 0 = \int_0^1 \lim_{n\to\infty} f_n(x)\ dx.$$

Theorem

Let $f_n : [a, b] \to \mathbb{R}$ be a sequence of continuous functions on a closed bounded interval, which converges uniformly to $f : [a, b] \to \mathbb{R}$. Then

$$\int_a^b f(x)\ dx = \lim_{n\to\infty} \int_a^b f_n(x)\ dx.$$

Remark

For improper integrals one needs additional assumptions as the example

$$f(x) = \begin{cases} \frac{1}{n^2}(n - |x|) & \text{if } |x| < n \\ 0 & \text{else} \end{cases}$$

shows. Apparently $\lim f_n = 0$ is uniform, but

$$\int_{-\infty}^{\infty} f_n(x)\ dx = 1 \neq 0 = \int_{\infty}^{\infty} 0\ dx.$$

Corollary

Let $f_n : [a, b] \to \mathbb{R}$ be a sequence of continuous differentiable functions which converges pointwise towards $f : [a, b] \to \mathbb{R}$. If the sequence of derivatives (f_n') converges uniformly, then f is differentiable and we have

$$f' = \lim_{n \to \infty} f_n'.$$

Application to power series

Theorem

Let $f(x) = \sum_{n=0}^{\infty} a_n x^n$ *be a power series with radius of convergence* $R > 0$, *hence* $f :]-R, R[\to \mathbb{R}$. *The power series*

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

which we obtain by termwise differentiation and integration have the same radius of convergence and converge on $]-R, R[$ *to*

1. $f'(x)$,
2. $\int_0^x f(x)\, dx$.

In particular f *is* ∞*–often differentiable and*

$$f^{(n)}(0) = a_n \cdot n!.$$

Examples:

1. The logarithm $\ln(1+x)$ is the antiderivative of

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

Termwise integration and $\ln(1) = 0$ gives:

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

for $|x| < 1$.

2. arctan is the antiderivative of

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

Integration and $\arctan(0) = 0$ gives:

$$\arctan(x) = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{2n+1} \text{ for } |x| < 1.$$

3. Since

$$\sum_{n=0}^{\infty} nx^n = x \cdot \sum_{n=1}^{\infty} nx^{n-1}.$$

we conclude

$$\sum_{n=0}^{\infty} nx^n = x \cdot \left(\frac{1}{1-x}\right)' = \frac{x}{(1-x)^2} \text{ for } |x| < 1.$$

For example:

$$\sum_{n=1}^{\infty} n\left(\frac{1}{2}\right)^n = \frac{\frac{1}{2}}{(1-\frac{1}{2})^2} = 2.$$

In the above examples 1. and 2. the series converge also for $x = 1$. This suggests

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

$$\sum_{n=0}^{\infty}(-1)^n \frac{1}{2n+1} = \arctan(1) = \frac{\pi}{4}$$

(since $\tan \frac{\pi}{4} = 1$). That this is really true follows from the following theorem.

Theorem (Abel's theorem)

Let $\sum_{n=0}^{\infty} a_n$ be a convergent series. Then the power series $f(x) = \sum_{n=0}^{\infty} a_n x^n$ has radius of convergence ≥ 1 and the limit function $f :]-1, 1[\to \mathbb{R}$ satisfies

$$\lim_{x \to 1} f(x) = \sum_{n=0}^{\infty} a_n.$$

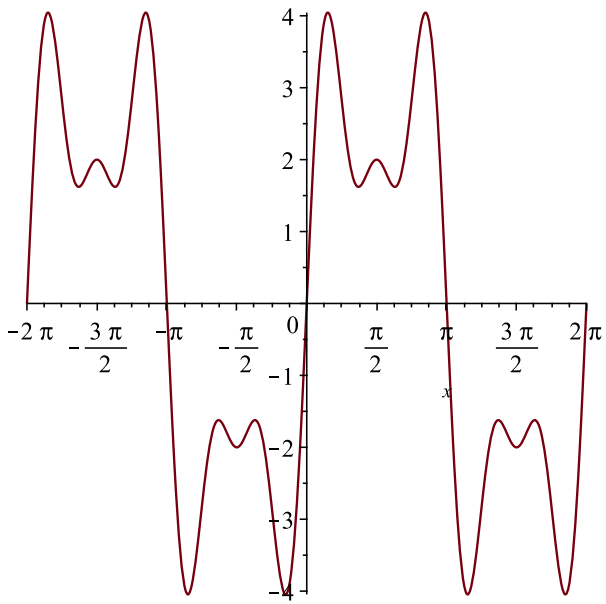
Fourier expansion

Let $f : \mathbb{R} \to \mathbb{R}$ be a periodic function with period L, i.e. $f(x + L) = f(x)$. By replacing x by $\frac{2\pi}{L}x$ we may assume that f is 2π-periodic. Examples of 2π-periodic functions are $\sin(x), \cos(x)$ and more general $\sin(kx), \cos(\ell x)$ for $k, \ell \in \mathbb{N}$.

A sum

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{n} a_k \cos(kx) + b_k \sin(kx)$$

with coefficients $a_0, \ldots, a_n, b_1, \ldots, b_n \in \mathbb{R}$ is called a **Fourier** or **trigonometric polynomial**.



The coefficients can be recovered from f, since

$$a_k = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos(kx) dx \text{ for } k = 0, 1, \ldots$$

and

$$b_k = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin(kx) dx \text{ for } k = 1, 2, \ldots.$$

holds.

Frequently, it is useful to allow complex valued functions

$$f : \mathbb{R} \to \mathbb{C},$$

since

$$\cos x = \frac{1}{2}(e^{ix} + e^{-ix}),\ \sin x = \frac{1}{2}(e^{ix} - e^{-ix}).$$

We can write the trigonometric polynomial f as

$$f(x) = \sum_{k=-n}^{n} c_k e^{ikx}$$

with

$$c_0 = \frac{1}{2}a_0,$$

$$c_k = \frac{1}{2}(a_k - ib_k), c_{-k} = \frac{1}{2}(a_k + ib_k).$$

The coefficients of the trigonometric polynomial

$$f(x) = \sum_{k=-n}^{n} c_k e^{ikx}$$

are given by

$$c_k = \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-ikx} dx \text{ for } k = 0, \pm 1, \pm 2, \ldots, \pm n$$

Definition

Let $f : \mathbb{R} \to \mathbb{C}$ be a 2π-periodic function, which is integrable over $[0, 2\pi]$. Then the numbers

$$c_k = \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-ikx} dx, \quad k \in \mathbb{Z}$$

are called the **Fourier coefficients** of f, and

$$\sum_{k=-\infty}^{\infty} c_k e^{ikx},$$

i.e. the sequence of partial sums

$$s_n(x) = \sum_{k=-n}^{n} c_k e^{ikx},$$

is called the **Fourier series** of f.

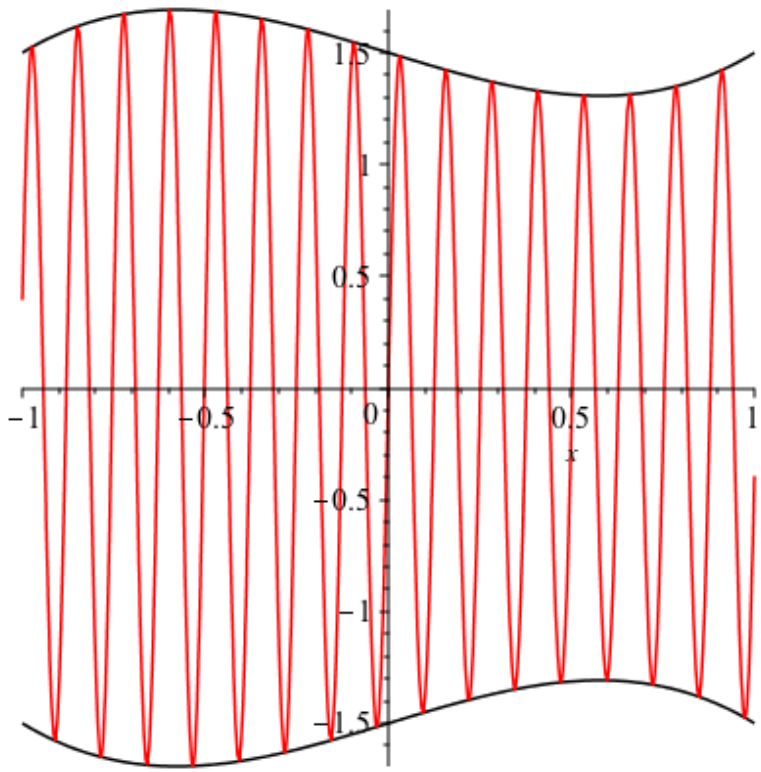
Theorem

Let $f : [a, b] \to \mathbb{R}$ be a continuously differentiable function. Consider

$$F(k) = \int_a^b f(x) \sin(kx) dx.$$

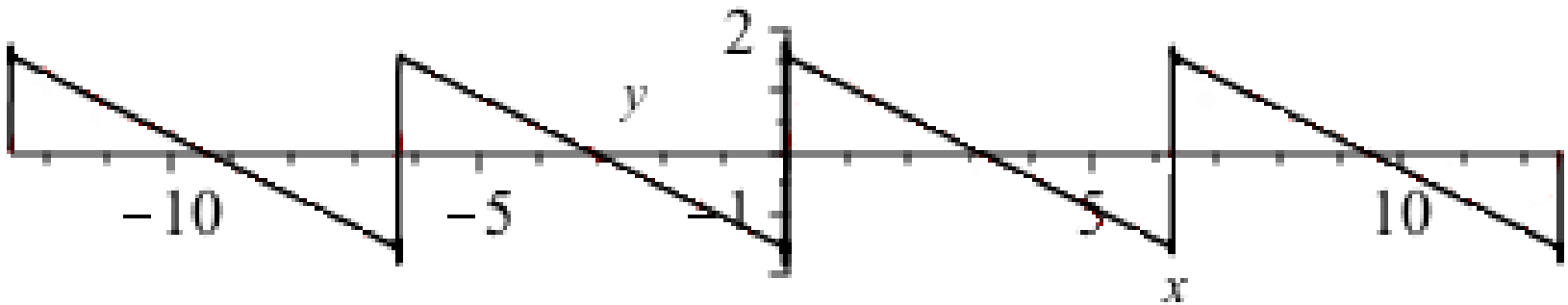
for $k \in \mathbb{R}$. Then

$$\lim_{k \to \pm\infty} F(k) = 0.$$

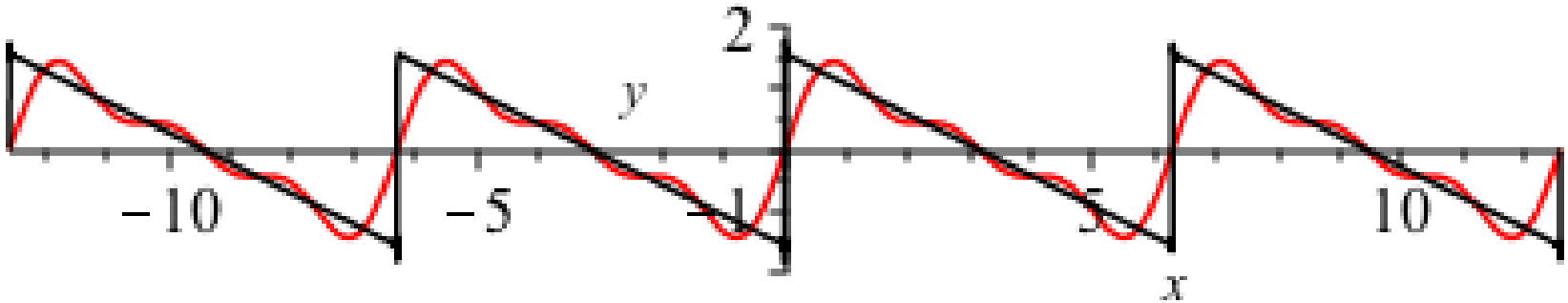


Example Consider the sawtooth function $\sigma : \mathbb{R} \to \mathbb{R}$ defined by

$$\sigma(0) = 0,$$
$$\sigma(x) = \frac{\pi - x}{2} \text{ for } x \in]0, 2\pi[, \text{ and}$$
$$\sigma(x) = \sigma(x + 2\pi n)\ \forall x \in \mathbb{R} \text{ and } n \in \mathbb{Z}.$$



The Fourier series of σ is $s(x) = \sum_{k=1}^{\infty} \frac{\sin(kx}{k}$.



The partial sum $\sum_{k=1}^{3} \frac{\sin(kx}{k}$ in red.

Proposition

$$\sum_{k=1}^{n} \cos(kt) = \frac{\sin((n+\frac{1}{2})t)}{2\sin(\frac{1}{2}t)} - \frac{1}{2}.$$

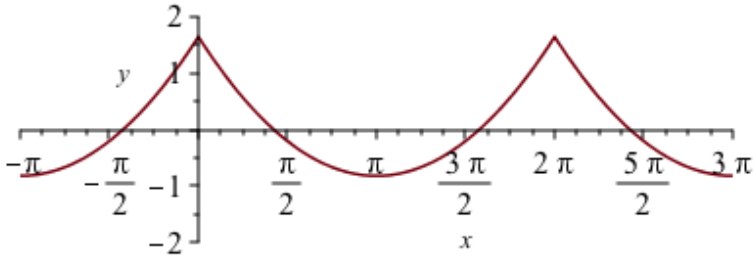
Theorem

$$\sum_{k=1}^{\infty} \frac{\sin(kx)}{k} = \frac{\pi - x}{2} \text{ for } x \in]0, 2\pi[$$

and for $\delta > 0$ *the convergence is uniform on* $[\delta, 2\pi - \delta]$.

As an application we will prove the formula

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}.$$



$$F(x) = \sum_{k=1}^{\infty} \frac{\cos(kx)}{k^2} = (\frac{x-\pi}{2})^2 - \frac{\pi^2}{12}.$$

Theorem

Let $f : \mathbb{R} \to \mathbb{R}$ be continuous 2π-periodic function, which is piecewise continuously differentiable. Then its Fourier series converges uniformly to f.