# Mathematik I für Studierende der Geophysik/Ozeanographie, Meteorologie und Physik

## Vorlesungsskript

Ralf Holtkamp
Fachbereich Mathematik
Universität Hamburg
www.math.uni-hamburg.de/home/holtkamp

Hamburg, 2023/24



[0] Version vom 24. Januar 2024, basierend auf Versionen der Vorjahre. Ursprüngliche Version: Vicente Cortés, 2006/07

# Inhaltsverzeichnis I





# Inhaltsverzeichnis II

## Inhaltsverzeichnis III





## Inhaltsverzeichnis IV

# Inhaltsverzeichnis V





# Inhaltsverzeichnis VI

# Kapitel 1

Grundlagen





## Mathematische Aussagen

Einer (mathematischen) Aussage kann man stets auf sinnvolle Weise einen Wahrheitswert, und zwar entweder "w(ahr)" oder "f(alsch)", zuordnen. Mathematische Aussagen, deren Wahrheit wir durch Beweise eingesehen haben, bezeichnen wir (je nach Wichtigkeit) in der Regel mit Theorem, Satz, Bemerkung/Rechenregel, Lemma (Hilfssatz) und Korollar (Folgerung).

- Die Aussage "$1^2 + 2^2 = 3^2$" hat Wahrheitswert (f).
- Quadrate reeller Zahlen sind nie negativ. (w)
- "$x^2 = 2$" wird erst mit einem bestimmten $x$ (oder z.B. durch Ergänzung von "Es gibt eine rationale Zahl $x$ mit") zu einer Aussage.
- Dieser Satz ist falsch. (?!)

# Wahrheitstafeln

Aus gegebenen Aussagen $A, B$ kann man neue Aussagen gewinnen, deren Wahrheitswert durch eine Wahrheitstafel definiert ist:

**Verneinung $\neg$**

| $A$ | $\neg A$ |
|---|---|
| w | f |
| f | w |

**“und” $\wedge$**

| $A$ | $B$ | $A \wedge B$ |
|---|---|---|
| w | w | w |
| w | f | f |
| f | w | f |
| f | f | f |

**“logisches oder” $\vee$**

| $A$ | $B$ | $A \vee B$ |
|---|---|---|
| w | w | w |
| w | f | w |
| f | w | w |
| f | f | f |





# Wahrheitstafeln

**Implikation $\Rightarrow$**

| $A$ | $B$ | $A \Rightarrow B$ |
|---|---|---|
| w | w | w |
| w | f | f |
| f | w | w |
| f | f | w |

“Wenn Ptolemäus recht hat, so dreht sich die Sonne um die Erde.” ist wahr.

**“Äquivalenz” $\iff$**

| $A$ | $B$ | $A \Leftrightarrow B$ |
|---|---|---|
| w | w | w |
| w | f | f |
| f | w | f |
| f | f | w |

# Tautologien

Zusammengesetzte Aussagen, die auf Grund ihrer logischen Struktur unabhängig vom Wahrheitswert der Teilaussagen wahr sind, heißen Tautologien.

Kontraposition: $(A \Rightarrow B) \iff (\neg B \Rightarrow \neg A)$

## Übung

Dies ist eine Tautologie, wie man anhand der Wahrheitstafel überprüft:

| $A$ | $B$ | $\neg A$ | $\neg B$ | $A \Rightarrow B$ | $\neg B \Rightarrow \neg A$ | $(A \Rightarrow B) \iff (\neg B \Rightarrow \neg A)$ |
|---|---|---|---|---|---|---|
| w | w | f | f | w | w | w |

Vervollständigen Sie die Wahrheitstafel!





# Mengen

Sie kennen die Menge der natürlichen Zahlen

$$\mathbb{N} = \{0, 1, 2, 3, 4, \ldots\}.$$

Achtung: Für uns gehört die 0 zu $\mathbb{N}$.

Für $\{1, 2, 3, 4, \ldots\}$ können wir kurz $\mathbb{N}^*$ schreiben.

## Mengen

Wir geben Mengen durch Beschreibung ihrer Elemente an, im einfachsten Fall durch Auflistung $\{a, b, c, \ldots\}$ ihrer Elemente.
Wichtig ist die Beziehung $\in$, d.h. “Enthalten sein”: Ist $M$ eine (bekannte) Menge, so ist für jedes Objekt $a$ die Beziehung $a \in M$ eine (wahre oder falsche) Aussage.
Es sind zwei Mengen $M, N$ gleich, wenn für alle Objekte $a$ gilt:
$(a \in M) \Leftrightarrow (a \in N)$.

Gilt $a = b = 1$, so ist $\{a, b\} = \{1\}$.
Die Menge der geraden natürlichen Zahlen ist:

$$2\mathbb{N} = \{m \in \mathbb{N} \mid m/2 \in \mathbb{N}\}.$$

Wir schreiben dabei | oder : für “für die gilt”.
Wir schreiben manchmal:
$:=$ für „nach Definition gleich“
$:\iff$ für „nach Definition gleichbedeutend.“

$$2\mathbb{N} := \{m \in \mathbb{N} \mid m/2 \in \mathbb{N}\}.$$





## Leere Menge, Teilmengen, Potenzmenge

Die Menge ohne Elemente bezeichnet man als die leere Menge $\emptyset$.

Seien $M, N$ Mengen. Dann heißt $N$ Teilmenge (oder Untermenge) der Menge $M$ genau dann, wenn jedes Element von $N$ auch Element von $M$ ist, d.h. wenn immer gilt:

$$(a \in N) \Rightarrow (a \in M).$$

Man schreibt $N \subseteq M$.
Mit $N \subsetneq M$ ist $(N \subseteq M) \wedge (N \neq M)$ gemeint.

Es gelten $M \subseteq M$ und $\emptyset \subseteq M$ für jede Menge $M$. Eine Menge $M$ ist nicht leer, wenn $\emptyset \subsetneq M$ gilt.

# Teilmengen, Potenzmenge

- Aus $a \in M$ folgt $\{a\} \subseteq M$. Die *Menge* $\{a\}$ sollte aber nicht mit dem Element $a$ verwechselt werden: eine Schachtel mit Hut ist etwas anderes als ein Hut!
- Ist $M$ eine Menge, dann nennt man die Menge aller Teilmengen von $M$ auch die Potenzmenge $\mathcal{P}(M)$.
  Beispiel: $\mathcal{P}(\{0,1\}) = \{\emptyset, \{0\}, \{1\}, \{0,1\}\}$.
- Achtung: Wir gehen hier auf Axiome der Mengenlehre nicht näher ein: Nicht durch alle "Beschreibungen" lassen sich sinnvoll Mengen bilden. (Können wir beispielsweise bilden: die Menge aller Mengen, die Menge aller Mengen, die sich selbst als Element enthalten, und die Menge aller Mengen, die sich nicht selbst als Element enthalten ?)





# Quantoren

Oft hat man für jedes Element einer gegebenen Menge eine Aussage. Es ist bequem, ein Symbol dafür zu haben, dass alle diese Aussagen mit "und" verknüpft werden. Beispiel:
Sei $N \subseteq M$. Dann gilt für alle $x \in N$ dass $x \in M$. Wir schreiben:
$\forall x \in N : x \in M$ und meinen, dass das eine wahre Aussage ist.
**Quantoren**:

| | | |
|---|---|---|
| $\forall$ | für | „für alle" (Allquantor) |
| $\exists$ | für | „es existiert ein" (Existenzquantor) |
| $\exists!$ | für | „es existiert genau ein" |

Beispiele:

- Es gilt $m \in 2\mathbb{N} \iff \exists n \in \mathbb{N} : m = 2n$.
- Es gilt $\forall m \in 2\mathbb{N} : \ (2m)/4 \in \mathbb{N}$.
- Es gilt nicht $\exists! n \in \mathbb{N} : \ n^2 = n$.

## Mengenoperationen

Im Folgenden erklären wir Operationen, durch die aus Mengen neue Mengen gebildet werden:

### Definition (Mengenoperationen)

*Es seien $M$, $N$ zwei beliebige Mengen, dann heißen*

- $M \cap N := \{p \mid p \in M \ \wedge \ p \in N\}$ *der Durchschnitt,*
- $M \cup N := \{p \mid p \in M \ \vee \ p \in N\}$ *die Vereinigung,*
- $M \setminus N := \{p \mid p \in M \ \wedge \ p \notin N\}$ *die Differenz,*
- $M \times N := \{(p, q) \mid p \in M \ \wedge \ q \in N\}$ *das kartesische Produkt*

*von $M$ und $N$.*

Hierbei bezeichnet $(p, q) \in M \times N$ das geordnete Paar.

**Beispiele:**

- $\mathbb{N}^* = \mathbb{N} \setminus \{0\}$.
- $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R} = \{(x, y) \mid x, y \in \mathbb{R}\}$, oder allgemeiner:
- $\mathbb{R}^n = \mathbb{R} \times \ldots \times \mathbb{R} = \{(x_1, \ldots, x_n) \mid x_i \in \mathbb{R}\}$.





## Einige Rechenregeln für Mengen

### Rechenregeln für Mengen

Seien $M, N, L$ beliebige Mengen. Dann gilt

- $M \cup M = M = M \cap M$ (Idempotenz),
- $M \cup N = N \cup M$ und $M \cap N = N \cap M$ (Kommutativität),
- $M \cap N \subseteq M \subseteq M \cup N$,
- $(M \cap N) \cap L = M \cap (N \cap L)$ und $(M \cup N) \cup L = M \cup (N \cup L)$ (Assoziativität),
- $(M \cup N) \cap L = (M \cap L) \cup (N \cap L)$ und $(M \cap N) \cup L = (M \cup L) \cap (N \cup L)$ (Distributivität).

# Symmetrische Differenz

**Übung**

Wir definieren für Teilmengen $A, B$ einer Menge $M$ ihre symmetrische Differenz durch

$$A \triangle B := (A \setminus B) \cup (B \setminus A).$$

Für beliebige Teilmengen $A, B, C \subseteq M$ gilt dann

$$A \cap (B \triangle C) = (A \cap B) \triangle (A \cap C).$$





# Beweise

Es gibt eine Vielzahl von Beweismethoden, z.B.

- direkter Beweis (z.B. Umformen von Gleichungen), z.B. „ $4x = 12 \Rightarrow x = 3$“
- indirekter Beweis: beruht auf der Kontraposition $(A \Rightarrow B) \iff (\neg B \Rightarrow \neg A)$, wobei $A$ die Voraussetzung ist und $B$ die Konklusion.
  (auch Beweis durch Widerspruch)
- Z.B. „$\exists$ unendlich viele Primzahlen“ kann man indirekt beweisen, in dem man zunächst annimmt, dass es nur endlich viele Primzahlen gibt, und dann einen Widerspruch (zu bereits eingesehenen Sätzen) herbeiführt.

# Beweis durch vollständige Induktion

**Prinzip der vollständigen Induktion**

Es sei $n_0 \in \mathbb{N}$ und $A(n)$ für jede natürliche Zahl $n \in \mathbb{N}$ mit $n \geq n_0$ eine Aussage. Es gelte
i) $A(n_0)$ ist wahr,

ii) für alle $n \geq n_0$ gilt die Implikation: $A(n) \Rightarrow A(n+1)$.

Dann ist $A(n)$ wahr für alle $n \in \mathbb{N}$ mit $n \geq n_0$.

Möchte man nun nachweisen, dass $A(n)$ für alle $n \in \mathbb{N}$ mit $n \geq n_0$ gilt, so muss man folgendes beweisen:

1. Induktionsanfang: $A(n_0)$ ist richtig,
2. Induktionsschritt: $A(n)$ für ein $n \geq n_0$ impliziert $A(n+1)$.
   Dazu nimmt man an, dass $A(n)$ gilt (Induktionvoraussetzung), und beweist die Aussage $A(n+1)$ (Induktionsbehauptung)





# Ein einfaches Beispiel

**Satz**

*Die Summe der natürlichen Zahlen von* 1 *bis* $n$ *ist* $\frac{1}{2}\, n(n+1)$, *für* $n \geq 1$,

$$\sum_{i=1}^{n} i \;=\; \frac{1}{2}\, n(n+1).$$

**Beweis.** (mit vollständiger Induktion nach $n$)

IA: Die Aussage gilt für $n = 1$, weil $\sum_{i=1}^{1} i = 1 = \frac{1}{2} \cdot 1 \cdot (1+1)$

IS: Man nimmt an, es gelte IV: $\sum_{i=1}^{n} i = \frac{1}{2}\, n(n+1)$

und beweist dann die IB:

$$\sum_{i=1}^{n+1} i = \sum_{i=1}^{n} i + (n+1) \overset{IV}{=} \frac{1}{2} n(n+1) + (n+1) = \frac{1}{2}(n+1)(n+2)$$

Damit folgt die Aussage für $n+1$ aus der Richtigkeit für $n$. $\square$

## Ein weiteres Beispiel

### Satz (Geometrische Summenformel)

*Es gilt die Formel*

$$\sum_{k=0}^{n} x^k = \frac{1-x^{n+1}}{1-x} \qquad x \neq 1,\ x \in \mathbb{R} \quad \textit{beliebig}$$

**Beweis.** Übung $\square$





### Satz

*Für alle $n \in \mathbb{N}$, $n \geq 1$, gilt:*
*Die Anzahl der möglichen Anordnungen (=Reihenfolgen) von $n$ verschiedenen Objekten $o_1, \dots, o_n$ ist $1 \cdot 2 \cdot \ldots \cdot n = n!$.*

**Beweis.** Beweis mit vollständiger Induktion:

IA: Für ein Objekt gibt es genau eine Anordnung.

IS: Wir betrachten nun die möglichen Anordnungen von $n+1$ Objekten. Ordnen wir ein gegebenes Objekt, etwa $o_1$ an erster Stelle an, so gibt es nach Induktionsvoraussetzung $n!$ mögliche Anordnungen der restlichen $n$ Objekte. Aber wir könnten $o_1$ auch an die zweite, dritte,..., $n+1$-te Stelle setzen und haben jeweils wieder $n!$ Anordnungen der restlichen Objekte. Die Gesamtzahl aller Anordnungen von $o_1, \dots, o_{n+1}$ ist also $n! \cdot (n+1) = (n+1)!$. $\square$

# Rekursive Definitionen

Eng verwandt mit dem Beweis durch vollständige Induktion ist die Konstruktion durch vollständige Induktion, auch *rekursive Definition* genannt. Jeder natürlichen Zahl $n$ soll eine Zahl $f(n)$ zugeordnet werden. Dies ist möglich durch:

1. Die Angabe von $f(0)$ oder $f(0)$ und $f(1)$ (oder jedenfalls nur endlich vieler Zahlen), und
2. einer Vorschrift $F_n$, die für jedes $n$ in $\mathbb{N}$ die Zahl $f(n)$ jeweils aus den Zahlen $f(0), \ldots, f(n-1)$ berechnet, d.h.

$$f(n) \ = \ F_n(f(0), \ldots, f(n-1))$$

**Beispiele:**

1. Potenzen: $x^0 \ := 1 \ , \quad x^n \ := \ x \cdot x^{n-1}$
2. Fibonacci Zahlen:
   $x_0 \ := \ 0, \quad x_1 \ := \ 1, \quad x_n \ := \ x_{n-1} + x_{n-2} \ (n \geq 2)$
3. Summe $\sum_{k=1}^{n} a_k$: $\sum_{k=1}^{0} a_k := 0$, $\sum_{k=1}^{n} a_k := a_n + \sum_{k=1}^{n-1} a_k$.
4. Fakultät: $0! \ := \ 1, \quad n! \ := \ (n-1)! \cdot n$





# Eigenschaften natürlicher Zahlen

In der Menge $\mathbb{N}$ der natürlichen Zahlen lassen sich auf bekannte Art und Weise Additionen und Multiplikationen ausführen. Führt man die natürlichen Zahlen mittels der **Peano Axiome** (G. Peano, 1889) ein, so werden Addition und Multiplikation rekursiv definiert:

$$m + 0 := m, \quad m + n' := (m+n)'$$

wobei $n'$ der Nachfolger von $n$ ist, sowie

$$m \cdot 0 := 0, \quad m \cdot n' := m \cdot n + m$$

Dann kann man die folgenden **Rechenregeln** beweisen:

- Kommutativität: $a + b = b + a$ und $a \cdot b = b \cdot a$ für alle $a, b \in \mathbb{N}$.
- Assoziativität: für alle $a, b, c \in \mathbb{N}$ gilt:
  $a + (b + c) = (a + b) + c$ und $a \cdot (b \cdot c) = (a \cdot b) \cdot c$
- Distributivität: für alle $a, b, c \in \mathbb{N}$ gilt:
  $a \cdot (b + c) = a \cdot b + a \cdot c$

## Ganze und rationale Zahlen

Ausgehend von den natürlichen Zahlen $\mathbb{N}$ konstruiert man die ganzen Zahlen

$$\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$$

um für alle Gleichungen $x + a = b$ mit $a, b \in \mathbb{N}$ Lösungen zu haben, und damit dann die rationalen Zahlen

$$\mathbb{Q} = \left\{ \frac{a}{b} \mid a, b \in \mathbb{Z}, b \neq 0 \right\} ,$$

um auch für möglichst viele Gleichungen $ax = b$ mit $a, b \in \mathbb{Z}$ Lösungen zu haben. Dann beweist man die üblichen Rechenregeln.

**Verabredung:**

Wir setzen die Kenntnis der elementaren Eigenschaften und Rechenregeln der ganzen und rationalen Zahlen als gegeben voraus.





## Abbildungen

**Definition**

Eine Abbildung $f$ zwischen zwei Mengen $A$, $B$ ist eine Vorschrift, die jedem Element $a \in A$ <u>genau ein</u> Element $f(a) \in B$ zuordnet.

$$\begin{aligned} f \ : \ A &\longrightarrow B \\ a &\mapsto f(a) \end{aligned}$$

- Für $a \in A$ heißt $b := f(a)$ das Bild von $a$ unter $f$.
- Für $b \in B$ heißt $a \in A$ ein Urbild von $b$ unter $f$ $:\iff \ f(a) = b$.

Seien $X \subseteq A$ und $Y \subseteq B$ zwei beliebige Teilmengen.

- $f(X) := \{f(a) \mid a \in X\}$ heißt Bild der Menge $X \subseteq A$ unter $f$. Es heißt $\operatorname{im} f := f(A)$ das Bild von $f$.
- $f^{-1}(Y) := \{a \in A \mid f(a) \in Y\}$ heißt Urbild der Menge $Y \subseteq B$ unter $f$.

## Definition

Sei $f : A \to B$ eine Abbildung zwischen den Mengen $A$ und $B$. Dann ist der Graph $\Gamma_f \subseteq A \times B$ der Abbildung definiert als die Menge

$$\Gamma_f \;:=\; \{(a, f(a)) \mid a \in A\}\ .$$

## Definition (Eigenschaften von Abbildungen)

Sei $f\colon A \to B$ eine Abbildung (zwischen den Mengen $A$ und $B$).

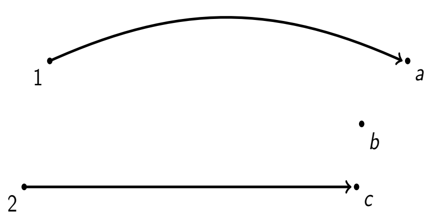
$$\begin{aligned}
f \text{ heißt surjektiv } &:\iff f(A) = B \\
&\iff \forall\, b \in B\ \exists\, a \in A:\ f(a) = b \\[1ex]
f \text{ ist injektiv } &:\iff \text{jedes Bild hat genau ein Urbild} \\
&\iff \forall\, b \in f(A)\ \exists!\, a \in A:\ f(a) = b \\
&\iff \forall\, x, y \in A : f(x) = f(y) \Rightarrow x = y \\[1ex]
f \text{ ist bijektiv } &:\iff f \text{ ist injektiv } \textbf{und} \text{ surjektiv}
\end{aligned}$$



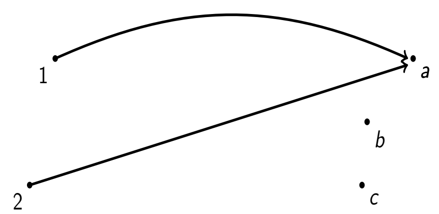


# Veranschaulichung

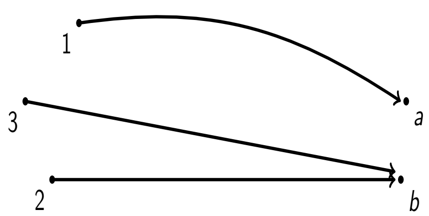
Injektive Abbildung einer zweielementigen Menge auf eine Menge mit drei Elementen



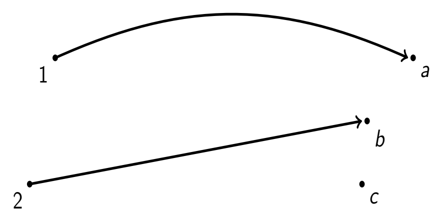

Nicht injektive Abbildung




Surjektive Abbildung

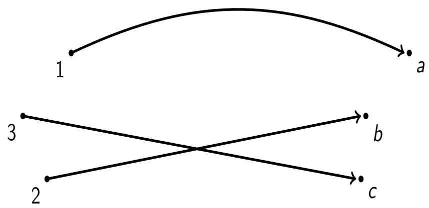



Nicht surjektive Abbildung

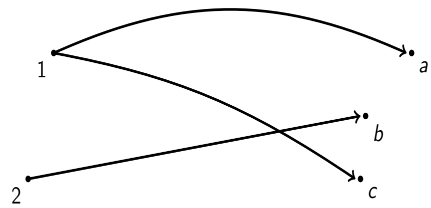


# Veranschaulichung II

Bijektive Abbildung




Keine Abbildung

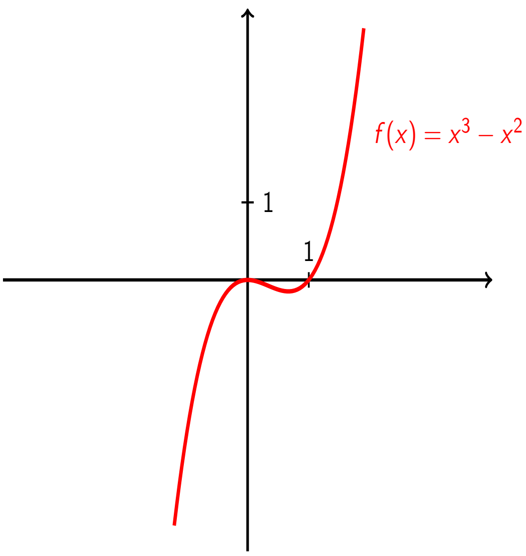







**Weitere Beispiele.**

- Die *Identitätsabbildung* oder *Identität* einer Menge $M$, definiert durch $\mathrm{Id}_M : M \to M$, $a \mapsto a$ ist bijektiv. Der Graph $\Gamma_{Id_M}$ ist die Diagonale.
- Eine andere bijektive Abbildung ist $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^3$.
- Die Abbildung $f : \mathbb{N} \to \mathbb{N}, \quad f(n) := 2 \cdot n$ ist injektiv aber *nicht* surjektiv. Die Abbildung $\tilde{f} : \mathbb{Q} \to \mathbb{Q}$ mit gleicher Vorschrift ist bijektiv.
- Die Abbildung $f : \mathbb{Q} \to \mathbb{Q}, \quad f(x) := x^2(x-1)$ ist surjektiv aber *nicht* injektiv, da $f(1) = f(0) = 0$.

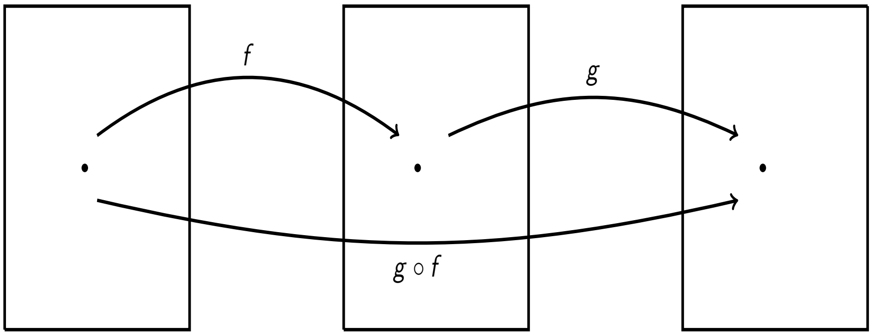


# Komposition von Abbildungen

**Bezeichnung.** Seien $f : A \to B$ und $g : B \to C$ zwei Abbildungen. Als Verkettung, Hintereinanderausführung oder auch Komposition der Abbildungen $f$ und $g$ bezeichnet man die Abbildung

$$g \circ f : A \to C, \qquad a \mapsto g \circ f(a) := g(f(a))$$

Die Verknüpfung von Abbildungen ist assoziativ: $f \circ (g \circ h) = (f \circ g) \circ h$.








**Lemma**

*Eine Abbildung $f : A \to B$ ist genau dann bijektiv, falls es eine Abbildung $g : B \to A$ gibt mit:*

$$g \circ f = \mathrm{Id}_A \qquad \textit{und} \qquad f \circ g = \mathrm{Id}_B$$

**Beweis.** Übung $\square$

**Bezeichnung.** Die Abbildung $g : B \to A$, die nach dem Lemma zu einer bijektiven Abbildung $f : A \to B$ existiert, nennt man die Umkehrabbildung zu $f$.

Warnung: eine Umkehrabbildung erbt nicht automatisch alle guten Eigenschaften einer Abbildung!

# Kapitel 2

## Reelle und komplexe Zahlen





# Warum reelle Zahlen?

Schon den Griechen in der Antike war bekannt, dass selbst einfache Konstruktionen auf Größen führen, die nicht als Verhältnis von Vielfachen einer Einheit (mit anderen Worten als rationale Zahlen) beschrieben werden können. Die Zahlmenge $\mathbb{Q}$ hat also "Lücken".
Die reellen Zahlen $\mathbb{R}$ beheben dieses Manko der rationalen Zahlen.

### Satz

*Die Länge der Diagonale eines Quadrats mit Seitenlängen gleich* 1 *ist irrational, d.h.* $\sqrt{2} \notin \mathbb{Q}$.

**Beweis.** Aus dem Satz von Pythagoras folgt, dass die Länge der Diagonale gleich $\sqrt{2}$ ist.
Wir zeigen $\sqrt{2} \notin \mathbb{Q}$ mittels Beweis durch Widerspruch:
Wir nehmen an $\sqrt{2} \in \mathbb{Q}$, d.h. $\sqrt{2} = \frac{p}{q}$, o.B.d.A. $p, q \in \mathbb{N}$ teilerfremd.
Die Gleichung $p^2 = 2q^2$ zeigt dann, dass $p^2$ und damit $p$ gerade sind.
Somit ist $p^2$ durch 4 teilbar, und damit sind $q^2$ und damit auch $q$ gerade.
Dies ist ein Widerspruch zur Teilerfremdheit! $\square$

## Körperaxiome

Die reellen Zahlen werden mit dem Symbol $\mathbb{R}$ bezeichnet.
Wir setzen ihre Existenz und die Kenntnis ihrer elementaren Eigenschaften und Rechenregeln voraus!
Abstrakt betrachtet sind die reellen Zahlen $\mathbb{R}$ ein Beispiel für einen Körper.
Ein Körper $\mathbb{K}$ ist eine Menge mit zwei Abbildungen,

- Addition

$$\begin{aligned} + \; : \; \mathbb{K} \times \mathbb{K} &\to \mathbb{K} \\ (x, y) &\mapsto x + y \end{aligned}$$

- und Multiplikation

$$\begin{aligned} \cdot \; : \; \mathbb{K} \times \mathbb{K} &\to \mathbb{K} \\ (x, y) &\mapsto x \cdot y = xy, \end{aligned}$$

welche die folgenden Axiome erfüllen:





## A) Axiome der Addition

- Kommutativgesetz:

$$x + y = y + x \quad \text{für alle} \quad x, y \in \mathbb{K}.$$

- Assoziativgesetz:

$$(x + y) + z = x + (y + z) \quad \text{für alle} \quad x, y, z \in \mathbb{K}.$$

- Es gibt ein neutrales Element $0 \in \mathbb{K}$, so dass

$$x + 0 = x \quad \text{für alle} \quad x \in \mathbb{K}.$$

  (Daraus folgt: $\exists!$ neutrales Element:
  Sei $0'$ ein weiteres neutrales Element, dann gilt
  $0 = 0 + 0' = 0' + 0 = 0'$.)

- Es gibt zu jedem $x \in \mathbb{K}$ ein additives Inverses $-x \in \mathbb{K}$, so dass

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

  (Wiederum folgt: $\forall x \in \mathbb{K} \quad \exists!$ additives Inverses)

Die Axiome A) fasst man auch in der Aussage zusammen, dass $(\mathbb{K}, +)$ eine abelsche Gruppe ist.

# M) Axiome der Multiplikation

- Kommutativgesetz:

$$xy = yx \quad \text{für alle} \quad x, y \in \mathbb{K}.$$

- Assoziativgesetz:

$$(xy)z = x(yz) \quad \text{für alle} \quad x, y, z \in \mathbb{K}.$$

- Es gibt ein neutrales Element $1 \in \mathbb{K} \setminus \{0\}$, so dass

$$x \cdot 1 = x \quad \text{für alle} \quad x \in \mathbb{K}.$$

- Es gibt zu jedem $x \in \mathbb{K}$ mit $x \neq 0$ ein Inverses $x^{-1} \in \mathbb{K}$, so dass

$$x \cdot x^{-1} = 1.$$

(Wie im Fall der Addition folgt die Eindeutigkeit von 1 und $x^{-1}$.)

Die Axiome M) fasst man auch in der Aussage zusammen, dass $(\mathbb{K} \setminus \{0\}, \cdot)$ eine abelsche Gruppe ist.





# D) Distributivgesetz

- Das Distributivgesetz drückt die Verträglichkeit von Addition und Multiplikation aus:

$$x(y + z) = xy + xz \quad \text{für alle} \quad x, y, z \in \mathbb{K}.$$

**Definition**

Eine Menge $\mathbb{K}$ mit zwei Abbildungen $+$ und $\cdot$, die die Axiome A), M) und D) erfüllen, heißt Körper.

**Einige Körper**

- Die reellen Zahlen $\mathbb{R}$ mit $+$ und $\cdot$ sind ein Körper
- Die rationalen Zahlen $\mathbb{Q}$ mit $+$ und $\cdot$ sind ein Körper
- Die Menge $\{0, 1\}$ mit $+$ und $\cdot$ derart, dass $0 = 0 \cdot 0 = 0 \cdot 1 = 1 \cdot 0 = 0 + 0 = 1 + 1$ und $1 = 1 \cdot 1 = 1 + 0 = 0 + 1$ ist ebenfalls ein Körper, der mit $\mathbb{F}_2$ bezeichnet wird.
- Natürliche und ganze Zahlen mit $+$ und $\cdot$ sind **kein** Körper

# Folgerungen aus den Körperaxiomen

Durch direkte Umformungen folgt aus aus A):

## Lemma

*Sei $\mathbb{K}$ ein Körper. Für alle $a, b \in \mathbb{K}$ ist die Gleichung*

$$a + x = b$$

*eindeutig lösbar. Die Lösung ist $x = b - a := b + (-a)$.*

Analog folgt aus M):

## Lemma

*Sei $\mathbb{K}$ ein Körper. Für alle $a \in \mathbb{K} \setminus \{0\}$ und $b \in \mathbb{K}$ ist die Gleichung*

$$ax = b$$

*eindeutig lösbar. Die Lösung ist $x = \frac{b}{a} := a^{-1}b$.*





# Weitere Folgerungen aus den Körperaxiomen

## Satz (Nullteilerfreiheit)

*Sei $\mathbb{K}$ ein Körper (z.B. $\mathbb{K} = \mathbb{R}$) und $x, y \in \mathbb{K}$. Dann gilt:*
*$xy = 0$ genau dann, wenn $x = 0$ oder $y = 0$.*

**Beweis.**

($\Leftarrow$) Es genügt zu zeigen: $x \cdot 0 = 0$ (wg. Kommutativität M). Es gilt

$$x \cdot 0 \overset{\text{(A)}}{=} x(0 + 0) \overset{\text{(D)}}{=} x \cdot 0 + x \cdot 0$$

Abziehen von $x \cdot 0$ auf beiden Seiten liefert $0 = x \cdot 0$.

($\Rightarrow$) Es genügt zu zeigen: $x \neq 0$ und $xy = 0 \Rightarrow y = 0$. Wir benutzen das vorhergehende Lemma: Da $x \neq 0$, hat die Gleichung $xy = 0$ die eindeutige Lösung $y = x^{-1} \cdot 0 = 0$. □

Weitere Folgerungen: $-0 = 0$,$-(-x) = x$, $(-x)y = -(xy)$,
$(-x)(-y) = xy$, $1^{-1} = 1$, $\forall x \neq 0$ ist $x^{-1} \neq 0$ und $(x^{-1})^{-1} = x$, etc.

### Konvention

Es ist üblich, natürliche, ganze und rationale Zahlen als Teilmengen der reellen Zahlen aufzufassen.

- Die natürlichen Zahlen $\mathbb{N} = \{0, 1, 2, 3, \dots\}$ fassen wir als Teilmenge von $\mathbb{R}$ auf, indem wir $n \in \mathbb{N}$ mit der reellen Zahl $1 + \dots + 1$ ($n$ Summanden) identifizieren.
- Durch Hinzunahme der additiven Inversen erhält man die ganzen Zahlen $\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\} \subseteq \mathbb{R}$. Sie bilden keinen Körper (keine multiplikative Inverse, außer für $\pm 1$).
- Der kleinste Körper, der $\mathbb{Z}$ enthält, ist der Körper der rationalen Zahlen $\mathbb{Q} = \{\frac{p}{q} \mid p \in \mathbb{Z}, \quad q \in \mathbb{N}, q \neq 0\}$.
- In jedem Körper $\mathbb{K}$ können wir die Elemente 1, $1+1$, $1+1+1$ etc, sowie ihre additiven Inversen betrachten. Aber daraus folgt nicht, dass $\mathbb{K}$ die ganzen Zahlen $\mathbb{Z}$ enthält! Gegenbeispiel: der Körper $\mathbb{F}_2$ mit zwei Elementen.





## O) Ordnungsaxiome

In $\mathbb{R}$ ist die Teilmenge $\mathbb{R}_+ \subseteq \mathbb{R}$ der positiven reellen Zahlen ausgezeichnet. Wir schreiben $x > 0$, wenn $x \in \mathbb{R}_+$.

### Vereinbarungen

- $x < 0 :\iff -x > 0$ (in dem Fall heißt $x$ negativ )
- $x > y :\iff x - y > 0$,
- $x \geq y :\iff x > y$ oder $x = y$,
- $x < y :\iff y > x$, sowie $x \leq y :\iff y \geq x$.

Eine Relation auf einer Menge $X$ ist eine Teilmenge $R \subseteq X \times X$.
Zum Beispiel wird die Relation “Gleichheit” durch die Diagonale $\{(x, x) \mid x \in X\} \subseteq X \times X$ beschrieben.
Auf den reellen Zahlen sind somit die Relationen

$$> \quad \text{durch} \quad \{(x, y) \mid x > y\} \subseteq \mathbb{R} \times \mathbb{R}$$

und $<$ (sowie $\geq$ und $\leq$) definiert. (Insbesondere sind $>$ und $<$ definiert auf den rationalen Zahlen $\mathbb{Q}$, wobei: $\frac{n}{m} > 0 \iff n \cdot m > 0$.)

## $\mathbb{R}$ als archimedisch geordneter Körper

Ein Körper $\mathbb{K}$ sei mit einer Relation $<$ versehen, für welche die folgenden Axiome gelten sollen:

(O1) $\forall x \in \mathbb{K}$ gilt entweder

$$x > 0 \quad \text{oder} \quad x = 0 \quad \text{oder} - x > 0.$$

(O2) Aus $x > 0$ und $y > 0$ folgt $x + y > 0$ und $x \cdot y > 0$ (Verträglichkeit mit der Addition und Multiplikation).

(O3) Archimedisches Axiom: $\forall x > 0, y > 0$ gibt es $n \in \mathbb{N}$ mit $nx > y$. (Hierbei ist $nx := x + \ldots + x$ mit $n$ Summanden)

**Definition**

*Ein Körper $\mathbb{K}$, der (O1), (O2) und (O3) erfüllt, heißt archimedisch angeordneter Körper.*

$\mathbb{R}$ und $\mathbb{Q}$ sind Beispiele für archimedisch angeordnete Körper.
Wir setzen hier die Existenz von $\mathbb{R}$ mit Gültigkeit der obigen Axiome als gegeben voraus.





## Folgerungen aus den Ordnungsaxiomen

**Rechenregeln**

Es gilt ($a, x, y, z$ aus $\mathbb{R}$ oder $\mathbb{Q}$):

(i) $x < y$ und $y < z \Longrightarrow x < z$ (Transitivität von $<$),

(ii) $x < y$ und $a \in \mathbb{R} \Longrightarrow x + a < y + a$,

(iii) $x < y$ und $a > 0 \Longrightarrow ax < ay$,

(iv) $x < y$ und $a < 0 \Longrightarrow ax > ay$,

(v) $x^2 > 0$ für alle $x \neq 0$.

**Beweis.** Wir beweisen exemplarisch die Transitivität von $<$: Wir schreiben

$$z - x = \underbrace{(z - y)}_{>0} + \underbrace{(y - x)}_{>0}$$

Aus der Verträglichkeit (O2) mit der Addition folgt $z - x > 0$, also $x < z$.
$\square$

### Korollar

Es gilt ($x, y$ aus $\mathbb{R}$ oder $\mathbb{Q}$):

(i) $x > 0 \Longrightarrow x^{-1} > 0$,

(ii) $0 < x < y \Longrightarrow x^{-1} > y^{-1} > 0$.

**Beweis.**

(i) Für $x > 0$ gilt wegen Rechenregel (v):

$$x^{-1} = (xx^{-1})x^{-1} = \underbrace{x}_{>0} \cdot \underbrace{(x^{-1})^2}_{>0} > 0.$$

(ii) Sei $0 < x < y$. (i) $\Rightarrow x^{-1} > 0$ und $y^{-1} > 0$. Somit $x^{-1}y^{-1} > 0$. Multiplikation der Ungleichung $x < y$ mit der positiven Zahl $x^{-1}y^{-1}$ liefert schließlich $y^{-1} < x^{-1}$.

□





## Intervalle

Wir benutzen die Ordnungsrelationen $<$ und $\leq$, um für gegebene reelle Zahlen $a < b$ Teilmengen von $\mathbb{R}$ zu beschreiben.

- Das offene Intervall ist die Teilmenge

$$(a, b) := \{x \in \mathbb{R} \mid a < x < b\} ,$$

  schließt also die Endpunkte nicht ein.
- Das abgeschlossene Intervall ist die Teilmenge

$$[a, b] := \{x \in \mathbb{R} \mid a \leq x \leq b\} ,$$

  schließt also die Endpunkte ein.
- Die Teilmengen $(a, b] := \{x \in \mathbb{R} \mid a < x \leq b\}$ und $[a, b) := \{x \in \mathbb{R} \mid a \leq x < b\}$ heißen halboffene Intervalle.

# Der Absolutbetrag

## Definition

*Der Absolutbetrag einer reellen Zahl $x$ ist definiert durch*

$$|x| := \begin{cases} x, & \text{wenn} \quad x \geq 0 \\ -x, & \text{wenn} \quad x < 0 \end{cases}$$

## Satz

*Für alle $x, y \in \mathbb{R}$ gilt:*

(i) $|x| \geq 0$ *und es ist dabei* $|x| = 0$ *genau dann, wenn* $x = 0$,

(ii) $|x + y| \leq |x| + |y|$ *(Dreiecksungleichung)*,

(iii) $|xy| = |x||y|$ *und*

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

**Beweis.** (i)–(iii) sind einfache ÜA. (iv) folgt aus der Dreiecksungl. (ii):

$|x| = |x - y + y| \overset{(ii)}{\leq} |x - y| + |y|$, dasselbe für $|y|$, $\Rightarrow (iv)$. □





# Abstand

Den Absolutbetrag benutzt man, um den Abstand zweier reeller Zahlen zu definieren:

## Satz

*Die Abbildung* $d : \mathbb{R} \times \mathbb{R} \to \mathbb{R}_+ \cup \{0\}$ *definiert durch* $d(x, y) := |x - y|$ *heißt Abstand von $x$ und $y$ und erfüllt die folgenden Eigenschaften:*

- $d(x, y) = 0 \iff x = y$,
- $d(x, y) = d(y, x)$ *(Symmetrie)*,
- $d(x, y) \leq d(x, z) + d(z, y)$ *(Dreiecksungleichung)*

*für alle* $x, y, z \in \mathbb{R}$.

**Beweis.** Folgt aus den Eigenschaften des Betrages. Insbesondere folgt die Dreiecksungleichung für den Abstand $d$ aus der für den Betrag $|\ .\ |$. □

# Komplexe Zahlen

Wir definieren nun den Körper der komplexen Zahlen. Dazu brauchen wir eine Menge, eine Addition und eine Multiplikation.

1. Wir betrachten die Menge aller Paare reeller Zahlen $\mathbb{R}^2 = \{(x,y) | x, y \in \mathbb{R}\}$ („kartesische Ebene")
2. versehen mit der Addition durch Addition der Komponenten:

$$\begin{aligned} + \;:\; \mathbb{R}^2 \times \mathbb{R}^2 &\to \mathbb{R}^2 \\ (x,y) + (u,v) &:= (x+u, y+v) \end{aligned}$$

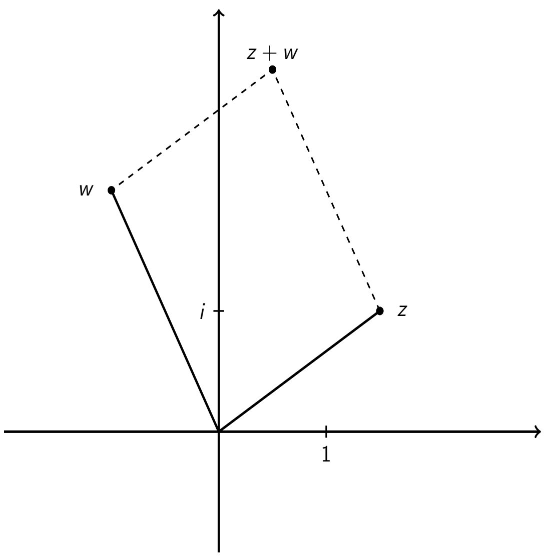
3. und der Multiplikation

$$\begin{aligned} \cdot \;:\; \mathbb{R}^2 \times \mathbb{R}^2 &\to \mathbb{R}^2 \\ (x,y) \cdot (u,v) &:= (xu - yv, xv + yu). \end{aligned}$$





# Veranschaulichung der Addition




Sie finden unter

http://www.math.uni-hamburg.de/master/lehrexport/physik/visualisierung/

Sage-Programme, die Sachverhalte veranschaulichen, die in der Vorlesung erklärt wurden.

### Theorem

*Das Tripel $(\mathbb{R}^2, +, \cdot)$ ist ein Körper. Dieser wird der Körper der komplexen Zahlen genannt und mit $\mathbb{C}$ bezeichnet.*

**Beweis.** Wir müssen die Axiome A), M) und D) nachweisen:

A) Kommutativ- und Assoziativgesetz für $(\mathbb{R}^2, +)$ folgen direkt aus den entsprechenden Axiomen für $(\mathbb{R}, +)$, da die Addition komponentenweise erklärt ist. Das neutrale Element in $(\mathbb{R}^2, +)$ ist $(0, 0)$ und $-(x, y) = (-x, -y)$.

M) Wir überprüfen zuerst das Assoziativgesetz der Multiplikation:

$$\begin{aligned}
&(x, y) \cdot ((u, v) \cdot (u', v')) \\
&\quad = (x, y) \cdot (uu' - vv', uv' + vu') \\
&\quad = (x(uu' - vv') - y(uv' + vu'), x(uv' + vu') + y(uu' - vv')) \\
&\quad = ((xu - yv)u' - (xv + yu)v', (xu - yv)v' + (xv + yu)u') \\
&\quad = (xu - yv, xv + yu) \cdot (u', v') \\
&\quad = ((x, y) \cdot (u, v)) \cdot (u', v')
\end{aligned}$$





## Weiter im Beweis:

M) Das Kommutativgesetz für die Multiplikation in $\mathbb{R}^2$ folgt aus der Invarianz des Ausdrucks $(xu - yv, xv + yu)$ unter Vertauschung der Paare $(x, y)$ und $(u, v)$.

Das neutrale Element der Multiplikation ist $(1, 0)$.

In der Tat: $(x, y) \cdot (1, 0) = (x \cdot 1 - y \cdot 0, x \cdot 0 + y \cdot 1) = (x, y)$.

Das multiplikative Inverse zu $(x, y) \neq (0, 0)$ ist

$$(x, y)^{-1} = \left( \frac{x}{x^2 + y^2}, -\frac{y}{x^2 + y^2} \right).$$

In der Tat:
$(x, y) \cdot (\frac{x}{x^2+y^2}, -\frac{y}{x^2+y^2}) = (\frac{x^2}{x^2+y^2} - \frac{y(-y)}{x^2+y^2}, \frac{x(-y)}{x^2+y^2} + \frac{yx}{x^2+y^2}) = (1, 0).$

D) Distributivgesetz: Auch durch explizites Nachrechnen. $\square$

# Konventionen

- Wir identifizieren reelle Zahlen $x$ mit Paaren $(x, 0) \in \mathbb{R}^2$.
  - Dies ist mit beiden Verknüpfungen in $\mathbb{R}$ und $\mathbb{R}^2$ verträglich.
  - Insbesondere können wir reelle und komplexe Zahlen miteinander multiplizieren: $x \cdot (u, v) := (x, 0) \cdot (u, v) = (xu, xv)$.
- Führt man dann die imaginäre Einheit

  $$i := (0, 1) \in \mathbb{R}^2$$

  ein, lässt sich *jedes* Paar $(x, y) \in \mathbb{R}^2$ in der Form $(x, y) = x + iy$ schreiben.
- Es gilt $i^2 = -1$, denn
  $i^2 = (0, 1) \cdot (0, 1) = (0 \cdot 0 - 1 \cdot 1, 0 \cdot 1 + 1 \cdot 0) = (-1, 0) = -1.$
  Daher ist diese Schreibweise mit den Verknüpfungen verträglich:

  $$\begin{aligned}(x+iy)(u+iv) &= xu + i^2yv + i(xv+yu) \\ &= xu - yv + i(xv+yu) \quad = \quad (x,y)\cdot(u,v)\end{aligned}$$

- Es ist $i^{-1} = -i$. Wir schreiben $\mathbb{C} = \{z = x + iy \mid x, y \in \mathbb{R}\}$.
- Auf dem Körper $\mathbb{C}$ kann *keine* Ordnung definiert werden (ÜA).





# Noch einmal das multiplikative Inverse

Sei $z = x + iy \in \mathbb{C}$ und $z \neq 0$. Dann ist

$$x^2 + y^2 \neq 0$$

da nicht die beiden reellen Zahlen $x$ und $y$ gleich Null sein können. Rechne:

$$\begin{aligned}\frac{1}{z} &= \frac{1}{x+iy} = \frac{x-iy}{(x+iy)(x-iy)} \\ &= \frac{x-iy}{x^2+y^2} = \frac{x}{x^2+y^2} - \frac{y}{x^2+y^2}i\end{aligned}$$

Das hatten wir oben schon gesehen. Wir haben zunächst erweitert, dann die binomische Formel im Nenner angewandt mit $i^2 = -1$.

## Satz

*$i$ und $-i$ sind die einzigen Lösungen der Gleichung $z^2 + 1 = 0$.*

**Beweis.** Sei $z^2 = -1$ für $z = x + iy \implies x^2 - y^2 = -1$ und $xy = 0$.
Also $x = 0$ und $y^2 = 1$, d.h. $y = \pm 1$. □

Vereinbarung: $\sqrt{-1} := i = (0, 1)$.

## Satz (Fundamentalsatz der Algebra)

*Jede polynomiale Gleichung mit komplexen Koeffizienten $c_k$ der Form*

$$z^n + c_{n-1}z^{n-1} + \ldots + c_1 z + c_0 = 0,$$

*mit $n \geq 1$, hat mindestens eine komplexe Lösung.*

**Beweis.** Der Beweis wird später gegeben. □

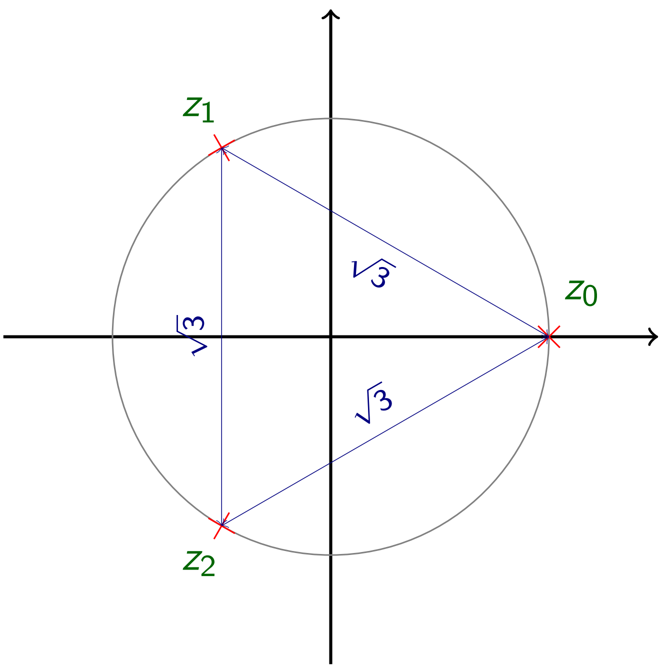
Folge: Jede polynomiale Gleichung vom Grad $n$ hat genau $n$ komplexe Lösungen (mit Vielfachheiten gerechnet).





## Übung

Die Verbindungsstrecken der drei komplexen Nullstellen von $z^3 - 1$ bilden ein gleichseitiges Dreieck.



# Real- und Imaginärteil, Konjugation

## Definition

- *Sei $z = x + iy \in \mathbb{C}$ eine komplexe Zahl. Dann heißen die reellen Zahlen $x = \operatorname{Re} z$ und $y = \operatorname{Im} z$ Real- bzw. Imaginärteil von $z$.*
- *Die komplexe Zahl $\bar{z} := x - iy$ heißt die zu $z = x + iy \in \mathbb{C}$ komplex konjugierte Zahl.*
- *Komplexe Zahlen $z \in \mathbb{C}$ mit $\operatorname{Re} z = 0$ heißen rein imaginär.*

Es gilt

$$x = \operatorname{Re} z = \frac{1}{2}(z + \bar{z}) \quad \text{und} \quad y = \operatorname{Im} z = \frac{1}{2i}(z - \bar{z})$$

sowie

$$\begin{aligned} \overline{z \cdot w} &= \bar{z} \cdot \bar{w} \quad \forall\, z, w \in \mathbb{C}. \\ \overline{z + w} &= \bar{z} + \bar{w} \quad \forall\, z, w \in \mathbb{C}. \\ \bar{\bar{z}} &= z \quad \forall\, z \in \mathbb{C}. \end{aligned}$$





# Der Betrag einer komplexen Zahl

Sei $z = x + iy \in \mathbb{C}$ (wie immer $x, y \in \mathbb{R}$).
$\Rightarrow z\bar{z} = (x + iy)(x - iy) = x^2 + y^2$ ist eine positive reelle Zahl oder Null.

## Definition

*Wir definieren den Betrag der komplexen Zahl $z$ durch*

$$|z| := \sqrt{z\bar{z}} = \sqrt{(x + iy)(x - iy)} = \sqrt{x^2 + y^2} \in \mathbb{R}_+ \cup \{0\}.$$

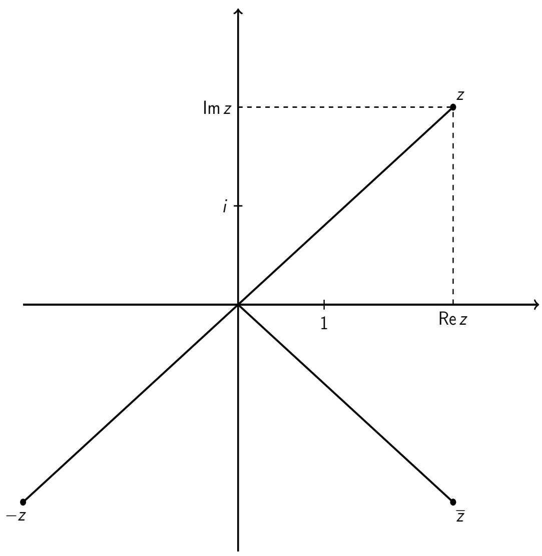
Für eine reelle Zahl $x \in \mathbb{R} \subseteq \mathbb{C}$ ergibt sich als Betrag

$$|x| = \sqrt{x^2} = \begin{cases} x, & \text{wenn} \quad x \geq 0 \\ -x, & \text{wenn} \quad x < 0 \end{cases}$$

genau der Absolutbetrag von $x$. Außerdem gilt:

$$|\mathrm{Re} z| \leq |z|, \quad |\mathrm{Im} z| \leq |z|, \quad \text{sowie} \quad |z| = |\bar{z}|.$$

# Veranschaulichung:




Geometrisch:
Konjugation ist Spiegelung an der reellen Achse.
Der Betrag ist der Abstand von $z$ zum Ursprung.





**Satz**

*Für alle $z, w \in \mathbb{C}$ gilt:*

(i) $|z| \geq 0$ *und* $|z| = 0$ *genau dann, wenn* $z = 0$,
(ii) $|z + w| \leq |z| + |w|$ (*Dreiecksungleichung*),
(iii) $|zw| = |z||w|$
(iv) $||z| - |w|| \leq |z - w|$.

**Beweis.** (i) ist klar und (iv) folgt mit den gleichen Argumenten wie im reellen Fall aus (ii).
(iii) folgt aus der Gleichung $|zw|^2 = zw\overline{zw} = zw\bar{z}\bar{w} = |z|^2|w|^2$ durch Wurzelziehen. (Hierbei haben wir benutzt, dass $\overline{zw} = \bar{z}\bar{w}$.)

(ii) Für alle $z = x + iy \in \mathbb{C}$ gilt

$$\begin{aligned} \mathrm{Re}\, z &= x \leq \sqrt{x^2 + y^2} = |z|. \\ \mathrm{Im}\, z &= y \leq \sqrt{x^2 + y^2} = |z|. \end{aligned}$$

Somit

$$\begin{aligned} |z + w|^2 &= |z|^2 + |w|^2 + z\bar{w} + \bar{z}w \\ &= |z|^2 + |w|^2 + z\bar{w} + \overline{z\bar{w}} \\ &= |z|^2 + |w|^2 + 2\mathrm{Re}\,(z\bar{w}) \\ &\leq |z|^2 + |w|^2 + 2|z\bar{w}| \\ &\overset{(iii)}{=} |z|^2 + |w|^2 + 2|z||w| = (|z| + |w|)^2. \quad \square \end{aligned}$$

$\square$





## Metrische Räume

Wieder definiert der Betrag einen Abstand, diesmal von komplexen Zahlen,

$$d \;:\; \mathbb{C} \times \mathbb{C} \longrightarrow \mathbb{R}_+ \cup \{0\}, \qquad d(z, w) := |z - w| \;.$$

Wie bei den reellen Zahlen hat er die Eigenschaften

- $d(z, w) \geq 0$ und $d(z, w) = 0 \iff z = w$,
- $d(z, w) = d(w, z)$,
- $d(z, w) \leq d(z, v) + d(v, w)$ (Dreiecksungleichung)

**Definition**

*Eine Menge $X$ mit einer Abstandsfunktion*

$$d \;:\; X \times X \longrightarrow \mathbb{R}_+ \cup \{0\},$$

*die die drei oben genannten Eigenschaften hat, nennt man einen metrischen Raum.*

Beispiele sind $\mathbb{C}$ und $\mathbb{R}$, aber auch jede Menge $X$ mit der Abstandsfunktion $d(x, y) = 1\ \forall x, y \in X$, falls $x \neq y$ und $d(x, x) = 0\ \forall x \in X$.

# Kapitel 3

## Konvergenz von Folgen und Vollständigkeit





### Definition

*Sei $X$ eine Menge (z.B. $\mathbb{C}$ oder $\mathbb{R}$).*
*Eine Folge von Elementen in $X$ ist eine Abbildung $\mathbb{N} \to X$, $n \mapsto x_n$.*
*Man schreibt dafür auch $(x_n)_{n\in\mathbb{N}} = (x_0, x_1, \ldots)$ oder einfach $(x_n)$.*

$n$ heißt der Index von $x_n$. Wir lassen auch Folgen zu, die erst ab einem gewissen Index $n_0 \in \mathbb{N}$ definiert sind: $(x_n)_{n\geq n_0} = (x_{n_0}, x_{n_0+1}, \ldots)$

### Definition

*Eine Folge $(x_n)$ mit Werten in einem metrischen Raum $(X, d)$ heißt konvergent mit dem Grenzwert („Limes“) $\ell \in X$, wenn es für jedes $\varepsilon \in \mathbb{R}$ mit $\varepsilon > 0$ einen Wert $N = N(\varepsilon) \in \mathbb{N}$ gibt, so dass*

$$d(x_n, \ell) < \varepsilon \quad \text{für alle Indizes} \quad n \geq N(\varepsilon).$$

*Schreibweise:* $\lim\limits_{n\to\infty} x_n = \ell$ *oder auch* $x_n \overset{n\to\infty}{\longrightarrow} \ell$

## Beachte:

Für Folgen reeller oder komplexer Zahlen gilt $d(x_n, \ell) = |x_n - \ell|$.

**Grundlegendes Beispiel**

Die Folge reeller Zahlen $(\frac{1}{n})_{n \geq 1}$ konvergiert gegen 0.

**Beweis.** Bezeichne $x_n := \frac{1}{n}$. Wir müssen für jedes reelle $\varepsilon > 0$ einen Index $N = N(\varepsilon)$ finden, so dass gilt

$$\forall n \geq N(\varepsilon): \quad d(x_n, 0) = |x_n - 0| < \varepsilon.$$

Sei also $\varepsilon > 0$ eine beliebige positive reelle Zahl. Nach dem Archimedischen Axiom (O3) gibt es $N = N(\varepsilon) \in \mathbb{N}$ mit $N \cdot \varepsilon > 1$.
D.h. $\frac{1}{N} < \varepsilon$.
Mit dieser Wahl für $N$ gilt für alle $n \geq N$: $|x_n - 0| = \frac{1}{n} \leq \frac{1}{N} < \varepsilon$. □

[Weitere Folgen mit unterschiedlichem Konvergenzverhalten als Animation]





## Eindeutigkeit des Grenzwertes

**Satz**

*Eine konvergente Folge $(x_n)$ in einem metrischen Raum $(X, d)$ hat genau einen Grenzwert.*

**Beweis.** (Indirekter Beweis)

- Wir nehmen an, dass $(x_n)$ zwei verschiedene Grenzwerte $\ell_1, \ell_2 \in X$ hat, d.h. $\ell_1 \neq \ell_2$. Setze $\varepsilon := \frac{d(\ell_1,\ell_2)}{2} > 0$.
- Wegen der Konvergenz existieren natürliche Zahlen $N_1$ und $N_2$ mit

$$\begin{aligned} d(x_n, \ell_1) &< \varepsilon \quad \text{für alle} \quad n \geq N_1, \\ d(x_n, \ell_2) &< \varepsilon \quad \text{für alle} \quad n \geq N_2. \end{aligned}$$

- Daraus folgt nach der Dreiecksungleichung für alle $n \geq \max\{N_1, N_2\}$:

$$2\varepsilon = d(\ell_1, \ell_2) \leq d(\ell_2, x_n) + d(x_n, \ell_1) < 2\varepsilon.$$

- D.h. $2\varepsilon < 2\varepsilon$, Widerspruch! Also muss $\ell_1 = \ell_2$ gelten. □

## Beschränkte Folgen reeller Zahlen

### Definition

(i) *Eine Folge* $(x_n)$ *reeller Zahlen heißt nach oben beschränkt, wenn es* $K \in \mathbb{R}$ *gibt derart, dass*

$$x_n \leq K \quad \text{für alle} \quad n.$$

(ii) *Sie heißt nach unten beschränkt, wenn es* $K \in \mathbb{R}$ *gibt derart, dass* $K \leq x_n \quad$ *für alle* $\quad n$.

(iii) *Eine Folge* $(x_n)$ *reeller oder komplexer Zahlen heißt beschränkt, wenn die Folge* $(|x_n|)$ *ihrer Beträge nach oben beschränkt ist:*

$$|x_n| \leq K \quad \text{für alle} \quad n.$$

### Einfache ÜA:

Eine Folge $(x_n)$ reeller Zahlen ist genau dann beschränkt, wenn sie nach oben und nach unten beschränkt ist.





## Konvergenz und Beschränktheit

### Satz

*Jede konvergente Folge* $(x_n)_{n\in\mathbb{N}}$ *komplexer Zahlen ist beschränkt.*

**Beweis.** Wir müssen ein $K \in \mathbb{R}$ finden, welches die Folge $(|x_n|)$ beschränkt: $|x_n| \leq K$ für alle $n$.

- Sei $\ell = \lim_{n\to\infty} x_n$. Dann existiert (etwa zu $\varepsilon = 1$) eine natürliche Zahl $N \in \mathbb{N}$, so dass
$$d(x_n, \ell) = |x_n - \ell| < 1 \quad \text{für alle} \quad n \geq N.$$
- Daraus folgt für alle $n \geq N$ (wieder mit der Dreiecksungleichung):
$$|x_n| = |x_n - \ell + \ell| \leq |x_n - \ell| + |\ell| \leq |\ell| + 1.$$
- und somit für alle $n \in \mathbb{N}$:
$$|x_n| \leq \max\{|x_0|, |x_1|, \ldots, |x_{N-1}|, |\ell| + 1\} =: K.$$

Man beachte, dass das Maximum hier über eine *endliche* Menge genommen wird. □

### Bemerkungen

(i) Die Umkehrung des letzten Satzes gilt nicht:
Die Folge $x_n = (-1)^n$ ist beschränkt aber nicht konvergent.

**Beweis.** Widerspruchsannahme: $\lim\limits_{n\to\infty}(x_n) = \ell$.
Für $\varepsilon := \frac{1}{2}$ existiert dann ein $N \in \mathbb{N}$: $|x_n - \ell| < \frac{1}{2}\ \forall n \geq N$.
Für *alle* $n \geq N$ und $m \geq N$ hat man wegen der Dreiecksungleichung

$$|x_n - x_m| = |x_n - \ell + \ell - x_m| \leq |x_n - \ell| + |x_m - \ell| < 1.$$

Mit $m = n+1$ ist das ein Widerspruch zu $(x_n) = (1, -1, 1, -1, 1, \ldots)$.
□

ii) Aus dem Satz folgt, dass unbeschränkte Folgen nicht konvergent sind.
Z.B. ist die Folge $x_n = n$ unbeschränkt und somit nicht konvergent.





## Konvergenz und algebraische Operationen auf $\mathbb{R}$ oder $\mathbb{C}$

### Satz

*$(x_n)$ und $(y_n)$ seien konvergente Folgen reeller oder komplexer Zahlen. Dann gilt:*

(i) $\lim\limits_{n\to\infty}(x_n + y_n) = \lim\limits_{n\to\infty} x_n + \lim\limits_{n\to\infty} y_n$

(ii) $\lim\limits_{n\to\infty}(x_n y_n) = (\lim\limits_{n\to\infty} x_n) \cdot (\lim\limits_{n\to\infty} y_n)$.

**Beweis.** Sei $x = \lim\limits_{n\to\infty} x_n$ und $y = \lim\limits_{n\to\infty} y_n$.
(i) Es gibt zu jedem $\varepsilon > 0$ natürliche Zahlen $N_1, N_2 \in \mathbb{N}$, so dass

$$\begin{aligned} |x_n - x| &< \varepsilon/2 \quad \text{für alle} \quad n \geq N_1, \\ |y_n - y| &< \varepsilon/2 \quad \text{für alle} \quad n \geq N_2. \end{aligned}$$

Dann gilt aber auch für alle $n \geq \max\{N_1, N_2\}$ mit der Dreiecksungleichung:

$$|x_n + y_n - (x+y)| = |x_n - x + y_n - y| \leq |x_n - x| + |y_n - y| < \varepsilon.$$

**Weiter im Beweis:**
(ii) Die Folge $(x_n)$ ist als konvergente Folge beschränkt.
Daher können wir $K > |y| \geq 0$ wählen, so dass $|x_n| \leq K$ für alle $n$.
Für jedes vorgegebene $\varepsilon > 0$ existieren $N_1, N_2 \in \mathbb{N}$ mit

$$\begin{aligned} |x_n - x| &< \frac{\varepsilon}{2K} \quad \text{für alle} \quad n \geq N_1, \\ |y_n - y| &< \frac{\varepsilon}{2K} \quad \text{für alle} \quad n \geq N_2. \end{aligned}$$

Somit gilt nun für alle $n \geq \max\{N_1, N_2\}$:

$$\begin{aligned} |x_n y_n - xy| &= |x_n y_n \underbrace{- x_n y + x_n y}_{=0} - xy| \\ &\leq \underbrace{|x_n|}_{\leq K} \underbrace{|y_n - y|}_{< \frac{\varepsilon}{2K}} + \underbrace{|x_n - x|}_{< \frac{\varepsilon}{2K}} \underbrace{|y|}_{< K} \; < \; \varepsilon. \end{aligned}$$

Also $\lim_{n\to\infty} (x_n y_n) = xy$. □





**Satz**

*Sei $(x_n)$ eine konvergente Folge reeller oder komplexer Zahlen mit Grenzwert $x \neq 0$.*
*Dann gibt es $n_0 \in \mathbb{N}$, so dass $x_n \neq 0$ für alle $n \geq n_0$ und*

$$\lim_{n\to\infty} \left(\frac{1}{x_n}\right)_{n \geq n_0} = \frac{1}{x}.$$

**Beweis.** ÜA □

# Konvergenz reeller Folgen und Ordnungsrelation

## Satz

- *Sei $(x_n)$ eine konvergente Folge reeller Zahlen mit $x_n \geq 0$. Dann gilt* $\lim\limits_{n\to\infty} x_n \geq 0$.
- $(x_n)$, $(y_n)$ *seien konvergente Folgen reeller Zahlen mit $x_n \leq y_n$. Dann ist*
$$\lim_{n\to\infty} x_n \leq \lim_{n\to\infty} y_n.$$

**Beweis.** Übung ☐

## Achtung:

Für konvergente Folgen mit $x_n < y_n$ gilt nicht unbedingt $\lim\limits_{n\to\infty} x_n < \lim\limits_{n\to\infty} y_n$.
Gegenbeispiel:
$x_n := 0$ die konstante Folge, $y_n := \frac{1}{n}$. Dann gilt in der Tat $\lim_{n\to\infty} x_n = 0 \leq \lim_{n\to\infty} y_n = 0$,
aber keine strikte Ungleichung.





## Übung

Ist $(x_n)_{n\in\mathbb{N}}$ eine Folge reeller Zahlen mit $x_n \geq 0$ für alle $n$ und $\lim\limits_{n\to\infty} x_n = x$, so ist
$$\lim_{n\to\infty} \sqrt{x_n} = \sqrt{x}.$$

# Cauchy-Folgen

## Definition

*Eine Folge* $(x_n)$ *mit Werten in einem metrischen Raum* $(X, d)$ *heißt* *Cauchy-Folge, wenn es für alle* $\varepsilon > 0$ *ein* $N \in \mathbb{N}$ *gibt, so dass*

$$d(x_n, x_m) < \varepsilon \quad \text{für alle} \quad n, m \geq N.$$

## Satz

*Jede konvergente Folge* $(x_n)_{n\in\mathbb{N}}$ *in einem metrischen Raum ist eine Cauchy-Folge.*

**Beweis.** Sei $\ell = \lim\limits_{n\to\infty} x_n$. Für alle $n, m \geq N(\frac{\varepsilon}{2})$ gilt wegen der Dreiecksungleichung:

$$d(x_n, x_m) \leq \underbrace{d(x_n, \ell)}_{<\frac{\varepsilon}{2}} + \underbrace{d(\ell, x_m)}_{<\frac{\varepsilon}{2}} < \varepsilon.$$

□





# Das Vollständigkeitsaxiom für die reellen Zahlen

Für den metrischen Raum, der den reellen Zahlen $\mathbb{R}$ zu Grunde liegt, gilt sogar die Umkehrung:

## Vollständigkeitsaxiom

$\mathbb{R}$ ist vollständig, d.h. jede Cauchy-Folge reeller Zahlen konvergiert gegen eine reelle Zahl.

## Bemerkungen

(i) Wegen der Vollständigkeit der reellen Zahlen konvergiert jede Cauchy-Folge rationaler Zahlen gegen eine reelle Zahl, diese kann jedoch irrational sein (wie wir später sehen werden). Somit ist der Körper $\mathbb{Q}$ der rationalen Zahlen nicht vollständig.

(ii) Der Körper $\mathbb{R}$ ist also ein vollständiger archimedisch angeordneter Körper, und jeder andere solche Körper $K$ ist isomorph zu $\mathbb{R}$ (d.h., dass es eine Bijektion zwischen beiden Körpern gibt, welche verträglich ist mit Addition, Multiplikation, Null und Eins).

# Monotone Folgen reeller Zahlen

### Definition

- *Eine Folge* $(x_n)$ reeller *Zahlen heißt monoton wachsend (bzw. streng monoton wachsend) wenn* $x_n \leq x_{n+1}$ *(bzw.* $x_n < x_{n+1}$*) für alle* $n \in \mathbb{N}$.
- *Analog definiert man „monoton fallend" und „streng monoton fallend".*

**Bemerkung.** Ist $(x_n)$ eine monoton wachsende (bzw. fallende) Folge reeller Zahlen mit dem Grenzwert $\ell$. Dann gilt $x_n \leq \ell$ (bzw. $\ell \leq x_n$) für alle $n \in \mathbb{N}$.





# Teilfolgen

### Definition

*Eine Folge* $(y_n)$ *heißt Teilfolge einer Folge* $(x_m)$*, wenn es eine streng monoton wachsende Folge* $\varphi : \mathbb{N} \to \mathbb{N}$ *gibt mit* $y_n = x_{\varphi(n)}$ *für alle* $n \in \mathbb{N}$.

### Beispiele

- Sei $x_n = \frac{1}{n}$. Die Folge $(x_n)$ ist konvergent. Betrachte für $\varphi(n) = n^2$ die Teilfolge $(y_n)$ mit $y_n := x_{n^2} = \frac{1}{n^2}$. Sie ist ebenfalls konvergent.
- Die Folge mit $x_n = (-1)^n$ ist nicht konvergent, aber die Teilfolgen $(y_n)$, $(z_n)$ mit $y_n := x_{2n} = 1$ und $z_n := x_{2n+1} = -1$ sind konvergent.

### Es gilt:

Ist $(x_n)$ eine konvergente Folge mit dem Grenzwert $\ell$, dann konvergiert jede Teilfolge von $(x_n)$ ebenfalls gegen $\ell$.

**Beweis.** Übung $\square$

# Der Satz von Bolzano-Weierstraß

## Theorem (Bolzano-Weierstraß)

*Jede beschränkte Folge reeller Zahlen besitzt eine konvergente Teilfolge.*

**Beispiel.** $x_n = (-1)^n$ ist beschränkt aber nicht konvergent. $y_n$ und $z_n$ aus dem obigen Beispiel sind konvergente Teilfolgen.

**Beweis.** Sei $(x_n)$ eine beschränkte reelle Folge, d.h. $\exists\ A, B \in \mathbb{R}$ mit

$$A \leq x_n \leq B \quad \text{für alle} \quad n \in \mathbb{N}.$$

Wir konstruieren rekursiv eine Folge von Intervallen $[A_k, B_k] \subseteq \mathbb{R}$ $(k \in \mathbb{N})$, derart dass

(i) das Invervall $[A_k, B_k]$ unendlich viele Glieder der Folge $(x_n)$ enthält,

(ii) $[A_k, B_k] \subseteq [A_{k-1}, B_{k-1}]$, falls $k \geq 1$ und

(iii) $B_k - A_k = \frac{1}{2^k}(B - A)$.

Die Intervalle sind also mit abnehmender Länge geschachtelt.





## Weiter im Beweis des Satzes von Bolzano-Weierstraß

- Wir setzen $A_0 := A$, $B_0 := B$ und nehmen an, die Intervalle $[A_k, B_k]$ mit den obigen Eigenschaften seien für $k \in \{0, \ldots, \ell\}$ bereits konstruiert.
- Wir definieren dann das $\ell + 1$-te Intervall $[A_{\ell+1}, B_{\ell+1}]$ wie folgt:
  - Sei $M := \frac{1}{2}(A_\ell + B_\ell)$ der Mittelpunkt des $\ell$-ten Intervalls.
  - Wir setzen $[A_{\ell+1}, B_{\ell+1}] := [A_\ell, M]$, also der linken Hälfte falls $[A_\ell, M]$ unendlich viele Glieder der Folge $(x_n)$ enthält und $[A_{\ell+1}, B_{\ell+1}] := [M, B_\ell]$ sonst.

$\Longrightarrow$ Das Intervall $[A_{\ell+1}, B_{\ell+1}]$ erfüllt (i)–(iii).

Als nächstes konstruieren wir, wieder rekursiv, eine Teilfolge $(y_k) = (x_{n_k})$ der vorgegebenen Folge $(x_n)$ mit $y_k \in [A_k, B_k]$.

- Wir setzen $y_0 := x_0$ und nehmen an, $y_0, \ldots, y_k$ seien schon konstruiert.
- Da $[A_{k+1}, B_{k+1}]$ unendlich viele Glieder der Folge $(x_n)$ enthält, gibt es eine natürliche Zahl $n_{k+1} > n_k$ mit $x_{n_{k+1}} \in [A_{k+1}, B_{k+1}]$.
- Wir setzen $y_{k+1} := x_{n_{k+1}}$.

## Ende des Beweises des Satzes von Bolzano-Weierstraß

**Behauptung**

Die Teilfolge $(y_k)$ ist eine Cauchy-Folge.

Aufgrund des Vollständigkeitsaxioms für $\mathbb{R}$ ist die Teilfolge $(y_k)$ auch konvergent, somit folgt aus der Behauptung nun der Satz von B-W. □

**Beweis der Behauptung.**

- Zum Beweis der Behauptung benutzen wir folgende Tatsache:

$$\lim_{n\to\infty} \frac{1}{2^n} = 0 \quad (\text{denn für } n \geq 1 \text{ gilt } 0 < \tfrac{1}{2^n} < \tfrac{1}{n} \overset{n\to\infty}{\longrightarrow} 0).$$

- Daher gibt es für alle $\varepsilon > 0$ ein $N \in \mathbb{N}$, so dass $\frac{1}{2^N}(B-A) < \varepsilon$.
- Für alle $k, \ell \geq N$ gilt $y_k, y_\ell \in [A_N, B_N]$ und somit

$$|y_k - y_\ell| \leq B_N - A_N = \frac{1}{2^N}(B-A) < \varepsilon. \quad \square$$





# Weitere Folgerungen aus dem Vollständigkeitsaxiom

## Konvergenz monotoner beschränkter Folgen

**Theorem**

*Jede monoton wachsende, nach oben beschränkte reelle Zahlenfolge $(x_n)$ konvergiert.*

*Ebenso konvergiert jede monoton fallende, nach unten beschränkte Folge.*

**Beweis.**

- Da die Folge $(x_n)$ monoton wachsend und nach oben beschränkt ist, ist sie beschränkt.
- Nach dem Satz von Bolzano-Weierstraß existiert also eine konvergente Teilfolge $(x_{n_k})_{k\in\mathbb{N}}$.
- Wir zeigen, dass die ganze Folge $(x_n)$ gegen $\ell := \lim_{k\to\infty} x_{n_k}$ konvergiert.

**Weiter im Beweis:**

- Wegen $\lim\limits_{k\to\infty} x_{n_k} = \ell$ gibt es zu jedem $\varepsilon > 0$ ein $k_0 \in \mathbb{N}$, so dass

$$|x_{n_k} - \ell| < \varepsilon \quad \text{für alle} \quad k \geq k_0.$$

- Wir setzen $N := n_{k_0}$.
- Für jedes $n \geq N = n_{k_0}$ gilt

$$x_{n_{k_0}} \leq x_n \leq \ell \ ,$$

  da die Folge$(x_n)$ monoton wachsend ist.
- und somit $|x_n - \ell| \leq |x_{n_{k_0}} - \ell| < \varepsilon$.

□





# Anwendung: Die Quadratwurzel einer positiven Zahl

**Theorem**

*Sei $a > 0$ eine positive reelle Zahl.*

(i) *Dann hat die Gleichung $x^2 = a$ genau eine positive Lösung. (Diese wird mit $\sqrt{a}$ bezeichnet.)*

(ii) *Sei $b > 0$ eine weitere positive reelle Zahl und $(x_n)$ die durch $x_0 := b$ und*

$$x_{n+1} := \frac{1}{2}(x_n + \frac{a}{x_n}) \quad (n \in \mathbb{N})$$

*rekursiv definierte reelle Folge. Dann konvergiert $(x_n)$ gegen $\sqrt{a}$.*

**Beweis.** (i) Wir zeigen zunächst, dass es höchstens eine Lösung gibt: Seien $x$ und $y$ zwei verschiedene positive Lösungen von $x^2 = a$. Dann gilt

$$0 = x^2 - y^2 = \underbrace{(x-y)}_{\neq 0}(x+y),$$

woraus $y = -x$ folgt. D.h. es kann höchstens eine positive Lösung geben.

**Beweis von (ii):**

1) Ein einfaches Induktionsargument zeigt, dass $x_n > 0$ für alle $n \in \mathbb{N}$.
2) Wir zeigen, dass $x_n^2 \geq a$ für alle $n \geq 1$ durch algebraische Umformungen:

$$\begin{aligned} x_n^2 - a &= \frac{1}{4}(x_{n-1} + \frac{a}{x_{n-1}})^2 - a \\ &= \frac{1}{4}(x_{n-1}^2 + 2a + \frac{a^2}{x_{n-1}^2}) - a \\ &= \frac{1}{4}(x_{n-1}^2 - 2a + \frac{a^2}{x_{n-1}^2}) \\ &= \frac{1}{4}(x_{n-1} - \frac{a}{x_{n-1}})^2 \geq 0 \end{aligned}$$





**Weiter im Beweis:**

3) Die Folge $(x_n)_{n\geq 1}$ ist monoton fallend:

$$\begin{aligned} x_n - x_{n+1} &= x_n - \frac{1}{2}(x_n + \frac{a}{x_n}) = \frac{1}{2}(x_n - \frac{a}{x_n}) \\ &= \underbrace{\frac{1}{2x_n}}_{>0}\underbrace{(x_n^2 - a)}_{\geq 0} \geq 0 \end{aligned}$$

- Die Folge $(x_n)_{n\geq 1}$ ist also monoton fallend und durch 0 nach unten beschränkt.
- Somit konvergiert die Folge $(x_n)$ gegen eine reelle Zahl $\ell \geq 0$.

**Weiter im Beweis:**

- Wir haben gezeigt, dass die Folge $(x_n)$ gegen eine Zahl $\ell \geq 0$ konvergiert.
- Die Rekursionsformel $x_{n+1} = \frac{1}{2}(x_n + \frac{a}{x_n})$ impliziert

$$2x_{n+1}x_n = x_n^2 + a.$$

Nun konvergiert sowohl die Folge $(x_{n+1})$ als auch $(x_n)$ gegen $\ell$. Rechenregeln für konvergente Folgen ergeben:

$$2\ell^2 = \ell^2 + a$$

und somit $\ell^2 = a$. $\square$





**Anmerkung.**

Obiges Verfahren zur Berechnung der Quadratwurzel ist eine Anwendung des Newton-Verfahrens zur Berechnung der Nullstelle einer stetigen reellen Funktion auf die Funktion $f\colon \mathbb{R} \to \mathbb{R}$ gegeben durch $f(x) = x^2 - a$.
Eine andere gegen $\sqrt{a}$ konvergierende Folge ist $x_{n+1} = x_n(3 - x_n^2/a)/2$. Sie ergibt sich bei der Berechnung der Nullstelle der Funktion $f(x) = 1 - a/x^2$ und konvergiert deutlich schneller.

# Unvollständigkeit von $\mathbb{Q}$

## Folgerung

*Der Körper $\mathbb{Q}$ der rationalen Zahlen ist nicht vollständig.*

**Beweis.**

- Wie wir gesehen haben, konvergiert die durch $x_0 := b > 0$ und $x_{n+1} := \frac{1}{2}(x_n + \frac{2}{x_n})$, $n \in \mathbb{N}$, rekursiv definierte Folge gegen $\sqrt{2}$.
- Insbesondere ist $(x_n)$ eine Cauchy-Folge.
- Wenn wir $b$ rational wählen, so sind alle Folgenglieder rational.
- Dann ist $(x_n)$ eine Cauchy-Folge von rationalen Zahlen, die gegen die reelle Zahl $\sqrt{2}$ konvergiert. Wir hatten bereits gesehen, dass diese irrational ist, $\sqrt{2} \notin \mathbb{Q}$.

$\square$





# Approximation reeller Zahlen durch rationale Zahlen

## Satz

*Für jede reelle Zahl $x$ gibt es eine monoton wachsende Folge rationaler Zahlen, die gegen $x$ konvergiert.*

**Beweis.**
Definiere $x_n := \max\{\frac{k}{2^n} \mid k \in \mathbb{Z},\ k \leq 2^n x\}$.
Die Folge $(x_n)_{n\in\mathbb{N}}$ ist monoton wachsend und durch $x$ nach oben beschränkt, also konvergent.
Wegen

$$|x - x_n| = \frac{1}{2^n}|2^n x - \max_{k\leq 2^n x} k| < \frac{1}{2^n}|(\max_{k\leq 2^n x} k) + 1 - \max_{k\leq 2^n x} k| = \frac{1}{2^n}$$

ist $x$ der Grenzwert. $\square$

**Bemerkung.**
Ebenso kann man $x$ durch fallende Folgen approximieren (ersetze Maximum durch Minimum).

# Vollständigkeit von $\mathbb{C}$

Erinnerung: Der Betrag einer komplexen Zahl $z = x + iy \in \mathbb{C}$ ist

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

$|\,.\,|$ erfüllt dieselben Eigenschaften wie der Absolutbetrag in $\mathbb{R}$ und liefert eine Abstandsfunktion auf $\mathbb{C}$, wodurch $\mathbb{C}$ zum metrischen Raum wird.
Es gilt dann:

**Satz**

*Der Körper $\mathbb{C}$ ist als metrischer Raum vollständig, d.h. jede Cauchy-Folge komplexer Zahlen konvergiert gegen eine komplexe Zahl.*





## Vollständigkeit von $\mathbb{C}$ (Beweis)

**Beweis.**

- Sei $(z_n = x_n + iy_n)$ eine komplexe Cauchy-Folge.
- Dann sind $(x_n)$ und $(y_n)$ reelle Cauchy-Folgen, denn

$$\begin{array}{rclcl} |x_n - x_m| & = & |\mathrm{Re}(z_n - z_m)| & \leq & |z_n - z_m| \quad \text{und} \\ |y_n - y_m| & = & |\mathrm{Im}(z_n - z_m)| & \leq & |z_n - z_m|. \end{array}$$

- Da $\mathbb{R}$ vollständig ist, konvergieren die reellen Folgen $(x_n)$ und $(y_n)$ gegen reelle Zahlen $x$ bzw. $y$.
- Wir setzen $z := x + iy$.
- Die Folge $(z_n)$ konvergiert gegen $z$, denn wegen der Dreiecksungleichung gilt

$$\begin{array}{rcl} |z_n - z| & = & |x_n - x + i(y_n - y)| \\ & \leq & |x_n - x| + |i(y_n - y)| \quad = \quad |x_n - x| + |y_n - y|. \end{array}$$

$\square$

## Bemerkung:

- Für jede konvergente Folge $(z_n = x_n + iy_n)$ konvergieren die Folgen $(x_n)$ und $(y_n)$ der Real- und Imaginärteile.
- Es konvergiert auch die Folge $(|z_n|)$ der Beträge, und es gilt

$$|\lim_{n\to\infty} z_n| = \lim_{n\to\infty} |z_n|.$$

  (Das folgt aus $||z_n| - |z|| \leq |z_n - z|$ mit $z = \lim_{n\to\infty} z_n$.)
  Die Umkehrung, dass aus der Konvergenz von $|z_n|$ die von $z_n$ folgt, gilt nicht:
  z. B. muss eine Folge komplexer Zahlen mit Betrag 1 nicht konvergieren, etwa $z_n = i^n$.
- Es gilt jedoch: $\lim_{n\to\infty} |z_n| = 0 \implies (z_n)$ konvergiert und $\lim_{n\to\infty} z_n = 0$.





## Geometrische Folge

Eine Folge von Potenzen $(z^n)$ einer gegebenen komplexen (oder reellen) Zahl $z$ heißt geometrische Folge. Für sie gilt:

**Satz**

*Sei $z$ eine komplexe oder reelle Zahl.*

1. *Ist $|z| < 1$, dann ist $(z^n)$ konvergent mit $\lim_{n\to\infty} z^n = 0$.*
2. *Ist $|z| > 1$, dann ist die Folge $(z^n)$ nicht beschränkt und damit nicht konvergent.*

Bemerkung:
Für $|z| = 1$ kann die Folge $(z^n)$ sowohl konvergent sein, nämlich für $z = 1$, als auch nicht konvergent, etwa für $z = -1$.

## Schritt 1 des Beweises

### Lemma (Bernoullische Ungleichung)

*Sei $x \in \mathbb{R}$ mit $x > -1$, $x \neq 0$. Für alle $n = 2, 3, ..$ gilt dann*

$$(1+x)^n > 1 + nx.$$

**Beweis.** (mittels Induktion)
Für $n = 2$ ist die Aussage richtig: $(1+x)^2 = 1 + 2x + x^2 > 1 + 2x$.
Gilt die Aussage für $n$, so gilt sie ebenso für $n+1$:

$$\begin{array}{rcll} (1+x)^{n+1} & = & (1+x)^n(1+x) & \\ & > & (1+nx)(1+x) & (IV) \text{ und } 1+x>0 \\ & = & 1+(n+1)x+nx^2 & > \quad 1+(n+1)x. \end{array}$$

□





### Folgerung (aus der Bernoullischen Ungleichung)

*Es sei $t \in \mathbb{R}$, $t > 0$. Weiter seien $K \in \mathbb{R}$ und $\varepsilon \in \mathbb{R}_+$ beliebig. Dann gilt*

(i) *Ist $t > 1$, dann gibt es ein $n \in \mathbb{N}$ so, dass $t^n > K$.*

(ii) *Ist $t < 1$, dann gibt es ein $n \in \mathbb{N}$ so, dass $t^n < \varepsilon$.*

**Beweis.** Folgt aus der Bernoullischen Ungleichung und (O3):

(i) Sei $t > 1$. Schreibe $t = 1 + x$ mit $x > 0$.
Sei $K \in \mathbb{R}$; ohne Einschränkung $K > 1$.
Wegen (O3) finden wir ein $n \in \mathbb{N}$, so dass $nx > K - 1$.
Die Bernoullische Ungleichung und $x > 0$ ergeben dann:
$t^n = (1+x)^n > 1 + nx > K$.

(ii) Für $0 < t < 1$ wenden wir (i) mit $K = \frac{1}{\varepsilon}$ auf $\frac{1}{t} > 1$ an. □

**Beweis des Satzes.** Falls $|z| < 1$ findet man für jedes $\varepsilon > 0$ ein $N$ mit $|z^n| = |z|^n < \varepsilon$ für alle $n \geq N$.
Für $|z| > 1$ findet man zu jedem $K$ ein $n$ mit $|z^n| > K$, d.h. $(z^n)$ ist nicht beschränkt. □

# Binomialkoeffizienten

Für natürliche Zahlen $n$ und $k$ mit $0 \leq k \leq n$ definiert man den Binomialkoeffizienten:

$$\binom{n}{k} = \frac{n!}{k! \cdot (n-k)!} = \frac{n(n-1)\dots(n-k+1)}{1 \cdot 2 \cdot \dots \cdot k},$$

man sagt dazu auch $n$ *über* $k$. Man hat $\binom{n}{0} = 1$ und $\binom{n}{n} = 1$.

**Satz (Binomischer Lehrsatz)**

*Seien $z, w$ beliebige komplexe Zahlen und sei $n$ eine natürliche Zahl, dann gilt:*

$$(z + w)^n = \sum_{k=0}^{n} \binom{n}{k} z^{n-k} w^k$$

**Beweis.** Man zeigt dies mit vollständiger Induktion nach $n$. □





# Kapitel 4

Konvergenz von Reihen

# Konvergenz von Reihen

## Definition

*Eine* Reihe *(komplexer Zahlen) ist eine Folge* $(s_n)_{n\in\mathbb{N}}$ *mit Folgengliedern*

$$s_n = \sum_{k=0}^{n} z_k,$$

*wobei* $(z_k)_{k\in\mathbb{N}}$ *eine Folge (komplexer Zahlen) ist.*
*Die Zahlen* $z_k$ *heißen* Glieder *der Reihe.*
*Die endlichen Summen* $s_n \in \mathbb{C}$ *heißen* Partialsummen *der Reihe.*
*Die Reihe* $(s_n)$ *heißt* konvergent*, falls die Folge der Partialsummen* $(s_n)$ *konvergent ist. Ihr Grenzwert wird mit*

$$\sum_{k=0}^{\infty} z_k := \lim_{n\to\infty} s_n \quad \text{bezeichnet.}$$





## Definition

*Eine Reihe komplexer Zahlen*

$$\left(\sum_{k=0}^{n} z_k\right)_{n\in\mathbb{N}}$$

*heißt* absolut konvergent*, wenn die Reihe der Absolutbeträge*

$$\left(\sum_{k=0}^{n} |z_k|\right)_{n\in\mathbb{N}}$$

*konvergent ist.*

Nicht jede konvergente Reihe ist absolut konvergent:

Zum Beispiel ist die alternierende harmonische Reihe $s_n = \sum_{k=1}^{n} \frac{(-1)^{k-1}}{k}$ konvergent (und zwar gegen $\ln 2$ wie wir später sehen werden). Aber es gilt:

### Satz

*Die Reihe* $\sum_{k=1}^{n} \frac{1}{k^p}$ *konvergiert für* $p > 1$ *und konvergiert nicht für* $p \leq 1$.
*Insbesondere (Fall* $p = 1$*) konvergiert die harmonische Reihe*

$$1 + \frac{1}{2} + \frac{1}{3} + \cdots$$

*nicht.*

Der Beweis wird später gegeben.





### Satz (Cauchysches Konvergenzkriterium)

*Eine Reihe* $(\sum_{k=0}^{n} z_k)_{n\in\mathbb{N}}$ *komplexer Zahlen konvergiert genau dann, wenn es zu jedem* $\varepsilon > 0$ *ein* $N \in \mathbb{N}$ *gibt, so dass*

$$\left| \sum_{k=m}^{n} z_k \right| < \varepsilon$$

*für alle* $n \geq m \geq N$.

**Beweis.** Die Bedingung des Satzes besagt, dass die Folge der Partialsummen $(s_n)$ eine Cauchy-Folge ist, denn für $m \leq n$

$$\sum_{k=m}^{n} z_k = \sum_{k=0}^{n} z_k - \sum_{k=0}^{m-1} z_k = s_n - s_{m-1}.$$

□

## Satz ("Absolute Konvergenz impliziert Konvergenz")

*Jede absolut konvergente Reihe komplexer Zahlen ist konvergent.*

**Beweis.** Wegen der Vollständigkeit von $\mathbb{C}$ genügt es, zu überprüfen, dass die Folge $(\sum_{k=0}^{n} z_k)_{n\in\mathbb{N}}$ der Partialsummen das Cauchy-Kriterium erfüllt.

Dies ist aber nach Voraussetzung erfüllt für die Reihe $(\sum_{k=0}^{n} |z_k|)_{n\in\mathbb{N}}$.

Daher gibt es zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$ mit

$$\sum_{k=m}^{n} |z_k| < \varepsilon \quad \text{für alle} \quad n \geq m \geq N.$$

Mit der Dreiecksungleichung folgt daraus für alle $n \geq m \geq N$:

$$\left|\sum_{k=m}^{n} z_k\right| \leq \sum_{k=m}^{n} |z_k| < \varepsilon.$$

Aus dem Cauchy-Kriterium folgt die Konvergenz der Reihe $(\sum_{k=0}^{n} z_k)_{n\in\mathbb{N}}$. $\square$





## Satz (Majorantenkriterium)

*Sei* $(\sum_{k=0}^{n} z_k)_{n\in\mathbb{N}}$ *eine Reihe, und* $a_k \geq 0$ *eine Folge reeller Zahlen mit*

- $|z_k| \leq a_k$ *für alle* $k \in \mathbb{N}$ *und*
- $(\sum_{k=0}^{n} a_k)_{n\in\mathbb{N}}$ *ist eine konvergente Reihe.*

*Dann ist die Reihe* $(\sum_{k=0}^{n} z_k)_{n\in\mathbb{N}}$ *absolut konvergent und*

$$\left|\sum_{k=0}^{\infty} z_k\right| \leq \sum_{k=0}^{\infty} |z_k| \leq \sum_{k=0}^{\infty} a_k.$$

**Beweis.** Wegen $\sum_{k=m}^{n} |z_k| \leq \sum_{k=m}^{n} a_k$ folgt aus der Konvergenz von $(\sum_{k=0}^{n} a_k)$ mit dem Cauchy-Kriterium die absolute Konvergenz.

Die Ungleichungskette folgt dann aus $|\sum_{k=0}^{n} z_k| \leq \sum_{k=0}^{n} |z_k| \leq \sum_{k=0}^{n} a_k$ durch Grenzübergang, der Ungleichungen erhält. $\square$

## Definition (geometrische Reihe)

*Sei $z \in \mathbb{C}$. Die Reihe $(\sum_{k=0}^{n} z^k)_{n \in \mathbb{N}}$ heißt geometrische Reihe.*

## Satz

*Für alle $z \in \mathbb{C}$ mit $|z| < 1$ gilt $\sum_{k=0}^{\infty} z^k = \frac{1}{1-z}$.*

**Beweis.** Für alle $z \neq 1$ gilt (in Verallgemeinerung der Summenformel für reelle $z$)

$$\sum_{k=0}^{n} z^k = 1 + z + z^2 + \cdots + z^n = \frac{1 - z^{n+1}}{1 - z}. \qquad (1)$$

Sei nun $|z| < 1$. Dann gilt $\lim_{n \to \infty} z^{n+1} = 0$.
Wegen (1) konvergiert dann die geometrische Reihe:

$$\sum_{k=0}^{\infty} z^k = 1 + z + z^2 + \cdots = \lim_{n \to \infty} \frac{1 - z^{n+1}}{1 - z} = \frac{1}{1 - z}.$$



□



**Bemerkung.**
Die Konvergenz der geometrischen Reihe ist für $|z| < 1$ absolut:

$$\sum_{k=0}^{\infty} |z^k| = \sum_{k=0}^{\infty} |z|^k = \frac{1}{1 - |z|}.$$

**Beispiele.**

$$z = \frac{1}{2}: \quad \sum_{k=0}^{\infty} \frac{1}{2^k} = 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \ldots = \frac{1}{1 - \frac{1}{2}} = 2.$$

$$z = -\frac{1}{2}: \quad \sum_{k=0}^{\infty} \left(-\frac{1}{2}\right)^k = 1 - \frac{1}{2} + \frac{1}{4} - \frac{1}{8} + \ldots = \frac{1}{1 + \frac{1}{2}} = \frac{2}{3}.$$

## Satz (Quotientenkriterium für die Konvergenz von Reihen)

*Sei $(a_k)_{k\geq 0}$ eine Folge komplexer Zahlen. Es gebe $0 \leq t < 1$, so dass*

$$|a_{k+1}| \leq t|a_k| \quad \textit{für alle} \quad k. \tag{2}$$

*Dann ist die Reihe $(\sum_{k=0}^{n} a_k)$ absolut konvergent und es gilt*

$$\sum_{k=0}^{\infty} |a_k| \leq \frac{|a_0|}{1-t}.$$

**Beweis.** Aus (2) erhält man durch Induktion $|a_k| \leq t^k|a_0|$.
Demnach gilt $\sum_{k=0}^{n} |a_k| \leq |a_0| \sum_{k=0}^{n} t^k$.
Die absolute Konvergenz folgt nun aus dem Majorantenkriterium und der Konvergenz der geometrischen Reihe $\sum_{k=0}^{\infty} t^k = \frac{1}{1-t}$. □





## Definition (Exponentialreihe)

*Sei $z \in \mathbb{C}$ gegeben. Die Reihe*

$$\left(\sum_{k=0}^{n} \frac{z^k}{k!}\right)_{n\geq 0}$$

*heißt Exponentialreihe.*

Die $n$-te Partialsumme der Exponentialreihe ist also

$$s_n = \sum_{k=0}^{n} \frac{z^k}{k!} = 1 + z + \frac{z^2}{2} + \frac{z^3}{6} + \frac{z^4}{24} + \cdots + \frac{z^n}{n!}.$$

## Satz

*Die Exponentialreihe ist für jedes $z \in \mathbb{C}$ absolut konvergent.*

**Beweis.**

- Finde zu gegebenen $z \in \mathbb{C}$ ein $N \in \mathbb{N}$, so dass $\frac{|z|}{N+1} \leq \frac{1}{2}$.
- Für alle $k \geq N$ gilt dann

$$\left|\frac{z^{k+1}}{(k+1)!}\right| = \frac{|z|}{k+1} \cdot \left|\frac{z^k}{k!}\right| \leq \frac{|z|}{N+1} \cdot \left|\frac{z^k}{k!}\right| \leq \frac{1}{2}\left|\frac{z^k}{k!}\right|$$

- Nach dem Quotientenkriterium ist also die Reihe $(\sum_{k=N}^{n} \frac{z^k}{k!})_{n \geq N}$ absolut konvergent und daher auch die Reihe

$$\sum_{k=0}^{n} \frac{z^k}{k!} = \sum_{k=0}^{N-1} \frac{z^k}{k!} + \sum_{k=N}^{n} \frac{z^k}{k!}.$$

$\square$





## Definition

*Die Exponentialfunktion ist die durch die Exponentialreihe definierte Abbildung*

$$\exp : \mathbb{C} \to \mathbb{C}, \quad z \mapsto \exp(z) := \sum_{k=0}^{\infty} \frac{z^k}{k!}.$$

Es gilt offenbar $\exp(0) = 1$.

## Definition

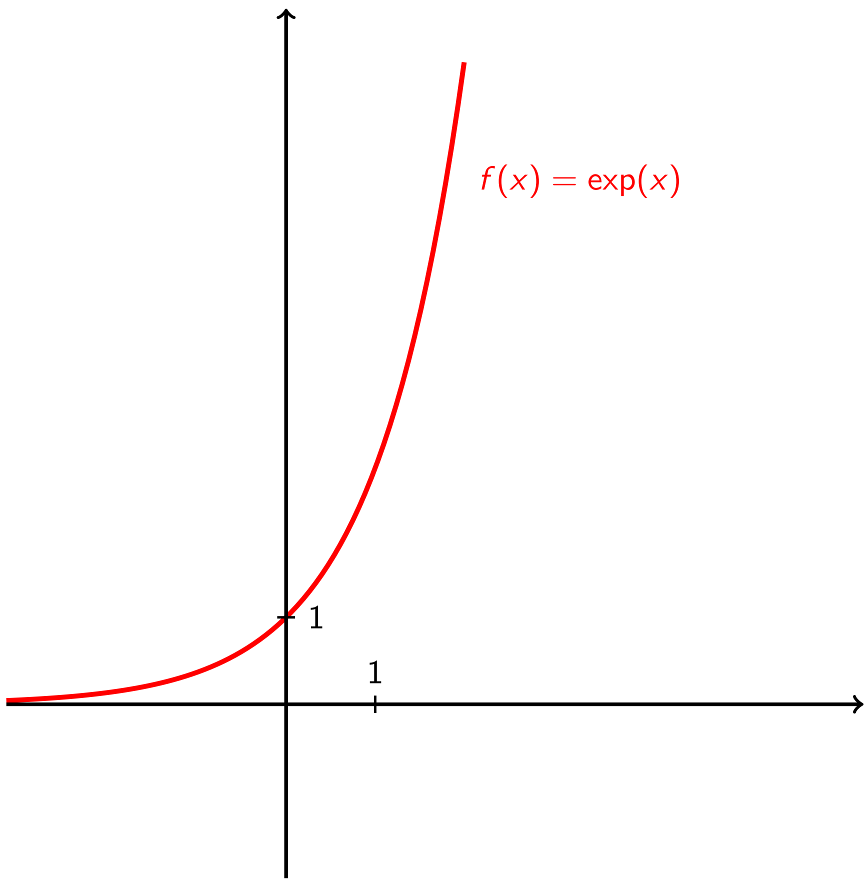
*Man definiert die Eulersche Zahl*

$$e := \exp(1) \in \mathbb{R} .$$

$$e \simeq 2,71828182845904523536028747135266249775724709369995\ldots$$

Diese Zahl ist nicht rational, ja nicht einmal Nullstelle irgendeines Polynoms mit rationalen Koeffizienten.

reeller Graph:








Seien $p_n = \sum\limits_{k=0}^{n} a_k$ und $q_n = \sum\limits_{k=0}^{n} b_k$ zwei (absolut) konvergente Reihen, d.h. die Folgen der Partialsummen (der Beträge) sind konvergent.
Wegen der Rechenregeln für Folgen ist auch deren Produkt konvergent, d.h. die Folgen

$$\begin{aligned} (\sum_{k=0}^{n} a_k) \cdot (\sum_{k=0}^{n} b_k) &= \sum_{0 \leq i,j \leq n} a_i \cdot b_j \\ \text{und} \qquad (\sum_{k=0}^{n} |a_k|) \cdot (\sum_{k=0}^{n} |b_k|) &= \sum_{0 \leq i,j \leq n} |a_i| \cdot |b_j| \end{aligned}$$

sind konvergent, aber nicht formal als Reihen gegeben. Man definiert:

**Definition (Cauchy Produkt)**

*Das Cauchy-Produkt zweier Reihen $(\sum\limits_{k=0}^{n} a_k)$ und $(\sum\limits_{k=0}^{n} b_k)$ ist die Reihe $(\sum\limits_{k=0}^{n} c_k)$ mit den Gliedern $c_k := \sum\limits_{j=0}^{k} a_j b_{k-j}$.*

## Satz

*Das Cauchy-Produkt* $(\sum_{k=0}^{n} c_k)$ *zweier absolut konvergenter Reihen* $(\sum_{k=0}^{n} a_k)$ *und* $(\sum_{k=0}^{n} b_k)$ *ist absolut konvergent mit dem Grenzwert*

$$\sum_{k=0}^{\infty} c_k = (\sum_{k=0}^{\infty} a_k) \cdot (\sum_{k=0}^{\infty} b_k).$$

**Beweis.** Die absolute Konvergenz der Reihe $(\sum_{k=0}^{n} c_k)$ folgt aus der Konvergenz beschränkter monotoner Folgen vermöge folgender Abschätzung:

$$\begin{aligned} \sum_{k=0}^{n} |c_k| &\leq \sum_{0\leq i,j\leq n,\ i+j\leq n} |a_i||b_j| \leq \sum_{0\leq i,j\leq n} |a_i||b_j| \\ &= (\sum_{i=0}^{n} |a_i|) \cdot (\sum_{j=0}^{n} |b_j|) \leq (\sum_{i=0}^{\infty} |a_i|) \cdot (\sum_{j=0}^{\infty} |b_j|). \end{aligned}$$





### Weiter im Beweis

- Bleibt zu zeigen, dass $\sum_{k=0}^{\infty} c_k = (\sum_{k=0}^{\infty} a_k) \cdot (\sum_{k=0}^{\infty} b_k)$.
- Es genügt zu zeigen, dass $\sum_{k=0}^{n} c_k - (\sum_{k=0}^{n} a_k) \cdot (\sum_{k=0}^{n} b_k)$ gegen Null konvergiert.
- Wir schätzen diese Folge gegen eine konvergente Folge ab:

$$\left| \sum_{k=0}^{n} c_k - (\sum_{k=0}^{n} a_k) \cdot (\sum_{k=0}^{n} b_k) \right| = \left| \sum_{0\leq i,j\leq n,\ i+j\leq n} a_i b_j - \sum_{0\leq i,j\leq n} a_i b_j \right|$$

$$= \left| \sum_{0\leq i,j\leq n,\ i+j>n} a_i b_j \right| \leq \underbrace{\sum_{k=0}^{\infty} |c_k| - \sum_{k=0}^{n} |c_k|}_{\lim_{n\to\infty}(\qquad)=0}$$

$\square$

Eine Folgerung aus dem Satz über das Cauchy-Produkt und dem binomischen Lehrsatz ist die Funktionalgleichung der Exponentialfunktion:

**Satz**

*Für alle $z, w \in \mathbb{C}$ gilt*

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

**Beweis.** Mit Hilfe des binomischen Lehrsatzes ergibt sich:

$$\begin{aligned}
\exp(z+w) &= \sum_{n=0}^{\infty} \frac{(z+w)^n}{n!} = \sum_{n=0}^{\infty}\sum_{k=0}^{n} \frac{z^{n-k}}{(n-k)!}\frac{w^k}{k!} \\
&\overset{\substack{\text{Satz über das}\\ \text{Cauchy-Produkt}}}{=} \sum_{n=0}^{\infty} \frac{z^n}{n!} \cdot \sum_{n=0}^{\infty} \frac{w^n}{n!} = \exp(z)\exp(w).
\end{aligned}$$

Hier geht die absolute Konvergenz der Exponentialreihe ein. □





Weitere Folgerungen

**Satz**

*Für alle $z \in \mathbb{C}$ gilt* $\exp(z) \neq 0$.

**Beweis.**
Die Behauptung folgt sofort aus
$\exp(z)\exp(-z) = \exp(z-z) = \exp(0) = 1$. □

**Bemerkungen:**

- Beachte, $\exp(n) = \exp(1+\cdots+1) = \exp(1)\cdots\exp(1) = e^n$.
  Darum schreibt man oft auch $e^z$ anstelle von $\exp(z)$.
- Für jede konvergente Folge $(z_n)$ konvergiert die Folge $(\overline{z_n})$ und es gilt

$$\lim_{n\to\infty} \overline{z_n} = \overline{\lim_{n\to\infty} z_n}.$$

  Insbesondere gilt $\exp(\bar{z}) = \overline{\exp(z)}$.

### Definition (Sinus und Kosinus)

*Für $z \in \mathbb{C}$ definiert man*

$$\cos z := \frac{1}{2}(\exp(iz) + \exp(-iz)) \quad \textit{und}$$
$$\sin z := \frac{1}{2i}(\exp(iz) - \exp(-iz)), \quad z \in \mathbb{C}.$$

Für $z \in \mathbb{C}$ haben wir also absolut konvergente Potenzreihenentwicklungen

$$\begin{aligned}
\cos z &= \frac{1}{2}(\exp(iz) + \exp(-iz)) \\
&= \sum_{k=0}^{\infty} (-1)^k \frac{z^{2k}}{(2k)!} = 1 - \frac{z^2}{2} + \frac{z^4}{4!} \cdots, \\
\sin z &= \frac{1}{2i}(\exp(iz) - \exp(-iz)) \\
&= \sum_{k=0}^{\infty} (-1)^k \frac{z^{2k+1}}{(2k+1)!} = z - \frac{z^3}{3!} + \frac{z^5}{5!} \cdots
\end{aligned}$$





## Einige Eigenschaften von Sinus und Kosinus

Einfache Folgerungen der Eigenschaften der Exponentialfunktion sind folgende Eigenschaften von Sinus und Kosinus, die für alle $x \in \mathbb{R}$ gelten:

Symmetrie, resp. Antisymmetrie

$$\cos(-x) = \cos x, \quad \sin(-x) = -\sin x.$$

Kosinus ist eine gerade, Sinus ist eine ungerade Funktion.

### Satz

*Für alle $x \in \mathbb{R}$ gilt $(\cos x)^2 + (\sin x)^2 = 1$ und somit*

$$\exp(ix) = \cos x + i \sin x \;\in\; S^1$$

*wobei $S^1 := \{z \in \mathbb{C} \mid |z| = 1\}$ den Kreis vom Radius 1 in $\mathbb{C}$ bezeichne.*

# Eigenschaften von Sinus und Kosinus

**Beweis.** Unmittelbar aus den Definitionen folgern wir

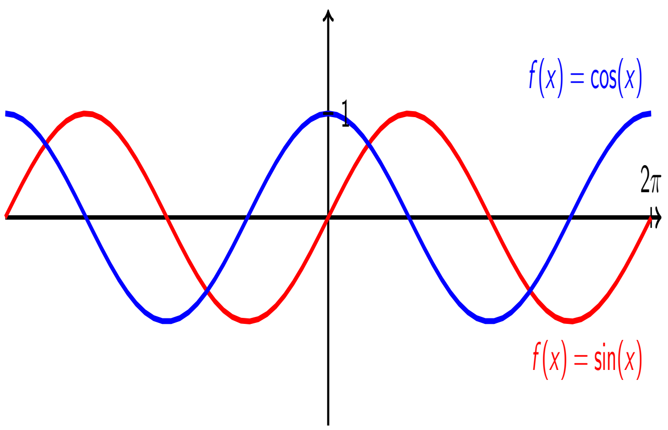
$$(\cos x)^2 + (\sin x)^2 = (\cos x + i \sin x)(\cos x - i \sin x) = \exp(ix)\exp(-ix)$$
$$\overset{\textit{Funktionalgl.}}{=} \exp(ix - ix) = \exp(0) = 1. \quad \square$$

Additionstheoreme (ÜA)

$$\begin{aligned} \cos(x+y) &= \cos x \cos y - \sin x \sin y, \\ \sin(x+y) &= \sin x \cos y + \cos x \sin y. \end{aligned}$$







## Ausblick: Ebene Polarkoordinaten

- Wir werden später noch sehen, dass die Funktion $\mathbb{R} \ni \varphi \mapsto \exp(i\varphi) \in S^1 \subseteq \mathbb{C}$, die Kreislinie periodisch (gegen den Uhrzeigersinn) durchläuft.
- Die Periode ist $2\pi$, wobei die reelle Zahl $\pi = 3,14159\ldots$ noch zu definieren ist, d.h. $\exp(2\pi i) = 1$.
- Der Parameter $\varphi$ lässt sich als Bogenlänge des entsprechenden Kreisbogens interpretieren.
- Insbesondere ist die Periode $2\pi$ genau die Länge der Einheitskreislinie $S^1 = \{z \in \mathbb{C} | |z| = 1\}$.
- Demensprechend hat jede komplexe Zahl $z \in \mathbb{C}^* := \mathbb{C} \setminus \{0\}$ eine eindeutige Darstellung der Form

$$z = r\exp(i\varphi), \quad \text{wobei} \quad r > 0 \quad \text{und} \quad \varphi \in [0, 2\pi).$$

- Man nennt $r = |z|$ und $\varphi$ die Polarkoordinaten von $z$
  Die reelle Zahl $\arg z := \varphi$ heißt Argument von $z$.





## Multiplikation in Polarkoordinaten

- Die Polarkoordinatendarstellung erleichtert die *Multiplikation* komplexer Zahlen:

$$r\exp(i\varphi) \cdot r'\exp(i\varphi') = rr'\exp(i(\varphi + \varphi')).$$

**Beispiel.** Wir berechnen $(1+i)^{20}$.

- Die Polarkoordinaten von $z = 1 + i$ sind $r = \sqrt{2}$ und $\varphi = \frac{\pi}{4}$,
- Also

$$z^{20} = \sqrt{2}^{20}\exp(i20\frac{\pi}{4}) = 2^{10}\exp(i5\pi) = 1024\exp(i\pi) = -1024.$$

- Hierbei haben wir benutzt, dass $\cos\frac{\pi}{4} = \sin\frac{\pi}{4} = \frac{\sqrt{2}}{2}$ und $\exp(i\pi) = -1$. (Das folgt aus geometrischen Überlegungen am Kreis.)

# Weitere Konvergenzkriterien für Reihen

### Satz

*Die Glieder $z_k$ einer konvergenten Reihe $\sum_{k=0}^{\infty} z_k$ bilden eine Nullfolge.*

**Beweis.** Nach dem Cauchy-Kriterium gibt es zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$, so dass $|z_n| = |\sum_{k=n}^{n} z_k| < \varepsilon$ für alle $n \geq N$. □

### Bemerkung

Das in der Proposition formulierte Konvergenzkriterium ist notwendig aber nicht hinreichend. Beispielsweise konvergiert ja die harmonische Reihe nicht, obwohl die Folge $(\frac{1}{k})_{k\geq 1}$ eine Nullfolge ist.





**Beispiel.**
Die alternierende harmonische Reihe

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} \cdots$$

konvergiert (und zwar gegen $\ln 2$, wie wir noch sehen werden). Die Konvergenz ergibt sich aus dem Leibnizkriterium:

### Satz (Leibnizkriterium für alternierende Reihen)

*Sei $a_k \geq 0$ $(k = 0, 1, 2, \ldots)$ eine monoton fallende Nullfolge reeller Zahlen. Dann konvergiert die Reihe $\sum_{k=0}^{\infty} (-1)^k a_k$.*

Beachte: der Satz gilt nicht ohne die Voraussetzung „monoton fallend“, wie man am Beispiel $(a_k)$ mit $a_k := \begin{cases} 1/l, & k = 2l \\ 0, & \text{sonst} \end{cases}$ sieht, wo (im Wesentlichen) wieder die harmonische Reihe vorliegt.

**Beweis.**

- Wir betrachten die Partialsummen $s_n = \sum_{k=0}^{n} (-1)^k a_k$ und setzen $x_n := s_{2n}$ und $y_n := s_{2n+1}$.
- Die Folge $(x_n)$ ist monoton fallend, denn $x_{n+1} - x_n = a_{2n+2} - a_{2n+1} \leq 0$.
- Die Folge $(y_n)$ ist monoton wachsend, denn $y_{n+1} - y_n = -a_{2n+3} + a_{2n+2} \geq 0$.
- Außerdem gilt $y_0 \leq y_n \leq x_n \leq x_0$, denn $y_n - x_n = -a_{2n+1} \leq 0$.
- Also existieren $x = \lim_{n\to\infty} x_n$ und $y = \lim_{n\to\infty} y_n$.
- Wegen $\lim_{n\to\infty} (y_n - x_n) = \lim_{n\to\infty} (-a_{2n+1}) = 0$, gilt $x = y$.
- Daraus folgt $\lim_{n\to\infty} s_n = \lim_{n\to\infty} s_{2n} = \lim_{n\to\infty} s_{2n+1}$. □





# Umordnungssatz

## Definition (Umordnung)

*Sei $(k_n)_{n\in\mathbb{N}}$ eine Folge, in der jede natürliche Zahl genau einmal vorkommt (d.h. wir haben eine Bijektion $\mathbb{N} \to \mathbb{N}$ gegeben durch $n \mapsto k_n$). Weiter sei $\sum_{n=0}^{\infty} a_n$ eine Reihe. Die Reihe $\sum_{n=0}^{\infty} a_{k_n}$ wird eine Umordnung von $\sum_{n=0}^{\infty} a_n$ genannt.*

## Wichtig:

Selbst bei konvergenten Reihen kommt es auf die Reihenfolge der Summationen an, Reihen sind als Folgen von Partialsummen definiert. Das Grenzverhalten kann sich bei einer Umordnung ändern.

**Beispiel.**

$$S_1 = \overbrace{1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5}}^{\approx 0{,}78333} + \ldots - \frac{1}{2j} + \frac{1}{2j+1} - \ldots = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n},$$

$$S_2 = \underbrace{1 + \frac{1}{3} - \frac{1}{2} + \frac{1}{5} + \frac{1}{7} - \frac{1}{4}}_{\approx 0{,}92619} + \ldots + \frac{1}{4j-3} + \frac{1}{4j-1} - \frac{1}{2j} + \ldots$$

Die Reihe $S_2$, in der auf jeweils zwei positive Glieder ein negatives mit geradem Nenner folgt, ist eine Umordnung der alternierenden harmonischen Reihe $S_1$. Man schreibt $S_2$ als $\sum_{k=1}^{n} (-1)^{k+1} a_k$ mit

$$a_{2j-1} = \frac{1}{4j-3} + \frac{1}{4j-1} \text{ und } a_{2j} = \frac{1}{2j} \quad \text{für } j = 1, 2, \ldots$$

und zeigt, dass beide Reihen nach dem Leibnizkriterium konvergieren.
Aber:
Da bei $S_1$ alle $-\frac{1}{2j} + \frac{1}{2j+1} < 0$ sind, während bei $S_2$ jeweils $\frac{1}{4j-3} + \frac{1}{4j-1} - \frac{1}{2j} > 0$ gilt, sind die Grenzwerte verschieden.





Für absolut konvergente Reihen gilt jedoch:

**Satz**

*Ist $\sum_{n=0}^{\infty} a_n$ eine Reihe komplexer Zahlen, die absolut konvergiert, dann konvergiert jede Umordnung von $\sum_{n=0}^{\infty} a_n$, und alle Umordnungen konvergieren gegen denselben Wert.*

**Beweis.** Da $\sum a_n$ absolut konvergent ist, findet man zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$ mit

$$\left| \sum_{n=0}^{N} |a_n| - \sum_{n=0}^{\infty} |a_n| \right| = \sum_{n=N+1}^{\infty} |a_n| < \varepsilon.$$

Es sei $\sum a_{k_n}$ eine Umordnung mit Partialsummen $s'_n$.
Sei nun $p \in \mathbb{N}$ so groß gewählt, dass die endlich vielen Zahlen $0, 1, \ldots, N$ in der Menge $k_0, k_1, \ldots, k_p$ enthalten sind. Es folgt für $n \geq p$, dass in der Differenz der Partialsummen $s_n - s'_n$ sich die Werte $a_0, \ldots, a_N$ gegenseitig aufheben, also

$$|s_n - s'_n| = \left| \sum_{i=0}^{n} (a_i - a_{k_i}) \right| \leq \sum_{i=N+1}^{\infty} |a_i| < \varepsilon.$$

Darum konvergiert $(s'_n)_{n\in\mathbb{N}}$ gegen den gleichen Wert wie $(s_n)_{n\in\mathbb{N}}$. □





Ergänzung:

### Satz (Riemannscher Umordnungssatz)

*Sei $\sum a_n$ eine konvergente Reihe reeller Zahlen, die jedoch nicht absolut konvergiert. Weiter sei $s \in \mathbb{R}$ beliebig vorgegeben. Dann existiert eine Umordnung $\sum a_{k_n}$ der Reihe $\sum a_n$ mit Wert*

$$\sum_{n=0}^{\infty} a_{k_n} = s.$$

**Beweis.** Es seien $(P_n)_{n\in\mathbb{N}}$ die positiven Glieder in $\sum a_n$ in der Reihenfolge, in der sie auftreten, und $(Q_n)_{n\in\mathbb{N}}$ die Absolutbeträge der negativen Glieder in der Reihenfolge des Auftretens.
Wir konstruieren nun Folgen $(m_n)_{n\in\mathbb{N}}$ und $(k_n)_{n\in\mathbb{N}}$, so dass die Reihe

$$P_0 + \ldots + P_{m_0} - Q_0 - \ldots - Q_{k_0} + P_{m_0+1} + \ldots + P_{m_1} - Q_{k_0+1} - \ldots - Q_{k_1} + \ldots$$

gegen den vorgegebenen Wert $s$ konvergiert. Dabei benutzen wir:
**Zentrale Hilfsbehauptung:** Die Summen $\sum P_n$ und $\sum Q_n$ sind divergent.

**Beweis der Hilfsbehauptung.**
Setze $p_n := \frac{|a_n|+a_n}{2}$ und $q_n := \frac{|a_n|-a_n}{2}$. Dann sind $p_n \geq 0$ und $q_n \geq 0$ und es gilt

$$a_n = p_n - q_n, \text{ und } |a_n| = p_n + q_n.$$

Würden beide Reihen $\sum p_n$ und $\sum q_n$ konvergieren, dann würde entgegen der Annahme auch

$$\sum |a_n| = \sum (p_n + q_n) \overset{Wdspruchann.}{=} \sum p_n + \sum q_n$$

konvergieren, also unsere Reihe absolut konvergieren. Andererseits folgt aus der Konvergenz von

$$\sum a_n = \sum (p_n - q_n) \overset{!}{=} \sum p_n - \sum q_n,$$

dass entweder beide Reihen $\sum p_n$ und $\sum q_n$ gleichzeitig divergieren oder konvergieren. Also divergieren beide Reihen.
Da nun aber (bis auf Glieder gleich Null) gilt, dass $\sum P_n = \sum p_n$ und $\sum Q_n = \sum q_n$, folgt die Hilfsbehauptung. □





**Weiter im Beweis.** Es seien $m_0$, $k_0$ die kleinsten Zahlen mit

$$P_0 + \ldots + P_{m_0} > s \text{ und } P_0 + \ldots + P_{m_0} - Q_0 - \ldots - Q_{k_0} < s.$$

Weiter seien nun $m_1$ und $k_1$ die kleinsten Zahlen mit

$$P_0 + \ldots + P_{m_0} - Q_0 - \ldots - Q_{k_0} + P_{m_0+1} + \ldots + P_{m_1} > s,$$
$$P_0 + \ldots - Q_0 - \ldots + P_{m_0+1} + \ldots + P_{m_1} - Q_{k_0+1} - \ldots - Q_{k_1} < s.$$

So fortfahrend konstruiert man $m_n$ und $k_n$.
Man beachte: Wegen der Divergenz der Reihen $\sum P_n$ und $\sum Q_n$ bricht dieses Konstruktionsverfahren nicht ab. Wegen

$$|P_0 + \ldots + P_{m_0} - Q_0 - \ldots + P_{m_n} - s| \leq P_{m_n},$$
$$|P_0 + \ldots + P_{m_0} - Q_0 - \ldots - Q_{k_n} - s| \leq Q_{k_n}$$

konvergiert die Folge der Partialsummen der so konstruierten Reihen, da aufgrund der Konvergenz von $\sum a_n$ die Folgen $(P_n)_{n\in\mathbb{N}}$ und $(Q_n)_{n\in\mathbb{N}}$, die ja aus Folgengliedern der Form $|a_n|$ bestehen, beide gegen 0 konvergieren. □

Zunächst zeigen wir jetzt noch:

**Lemma (Verdichtungssatz)**

*Es sei $(a_n)_{n\in\mathbb{N}}$ eine monoton fallende Folge nichtnegativer reeller Zahlen. Dann konvergiert die Reihe $\sum_{n=1}^{\infty} a_n$ genau dann, wenn die Reihe*

$$\sum_{k=0}^{\infty} 2^k a_{2^k} = a_1 + 2a_2 + 4a_4 + 8a_8 + \dots$$

*konvergiert.*

**Beweis.**
Die Folge der Partialsummen ist für beide Folgen monoton wachsend. Darum genügt es deren Beschränktheit zu untersuchen um die Konvergenz zu beweisen. Wir betrachten daher die jeweiligen Partialsummen:

$$\begin{aligned} s_n &= a_1 + a_2 + \dots + a_n, \\ t_k &= a_1 + 2a_2 + \dots + 2^k a_{2^k}. \end{aligned}$$





**Weiter im Beweis.**
Für $n \le 2^k$ gilt wegen $a_n \ge a_{n+1}$:

$$\begin{aligned} s_n &\le a_1 + a_2 + a_3 + \underbrace{a_4 + a_5 + a_6 + a_7}_{\le 4a_4} + \dots + \underbrace{a_{2^k} + \dots + a_{2^{k+1}-1}}_{2^k \text{ Glieder}} \\ &\le a_1 + 2a_2 + 4a_4 + \dots + 2^k a_{2^k} \\ &= t_k \end{aligned}$$

Andererseits gilt für $n \ge 2^k$:

$$\begin{aligned} s_n &\ge a_1 + a_2 + a_3 + a_4 + \underbrace{a_5 + a_6 + a_7 + a_8}_{\ge 4a_8} + \dots + \underbrace{a_{2^{k-1}+1} + \dots + a_{2^k}}_{2^{k-1} \text{ Glieder}} \\ &\ge \frac{1}{2} a_1 + a_2 + 2a_4 + 4a_8 \dots + 2^{k-1} a_{2^k} \\ &= \frac{1}{2} t_k \end{aligned}$$

Somit sind die Folgen $(s_n)_{n\in\mathbb{N}}$ und $(t_k)_{n\in\mathbb{N}}$ entweder beide beschränkt oder beide nicht beschränkt. □

Mit dem vorhergehenden Lemma zeigen wir nun, wie angekündigt:

### Satz

*Die Reihe* $\sum_{k=1}^{n} \frac{1}{k^p}$ *konvergiert für* $p > 1$ *und konvergiert nicht für* $p \leq 1$.

**Beweis.**
Gilt $p \leq 0$, dann bilden die Glieder der Reihe keine gegen Null konvergente Folge und somit divergiert die Reihe.
Falls $p > 0$ so bilden die Glieder der Reihe eine monoton fallende Folge reeller Zahlen, wegen des vorigen Satzes betrachten wir die Reihe:

$$\sum_{k=0}^{\infty} 2^k \frac{1}{(2^k)^p} = \sum_{k=0}^{\infty} 2^{(1-p)k}. \tag{3}$$

Die geometrische Reihe (3) konvergiert genau dann, wenn $2^{1-p} < 1$, also genau dann, wenn $1 - p < 0$. Aus dem vorhergehenden Satz folgt nun die Behauptung. □





# Kapitel 5

## Eigenschaften reeller Punktmengen

## Mächtigkeit

### Definition

- *Eine nicht-leere Menge $A$ heißt endlich, wenn es ein $n \in \mathbb{N}$ und eine surjektive Abbildung $\varphi : \{0, 1, \ldots n\} \to A$ gibt.*
- *Zwei Mengen $M, N$ heißen gleichmächtig, wenn es eine bijektive Abbildung $\varphi : M \to N$ gibt.*

### Beispiel

- Die Menge $\{0, 1, 2, 3, 4\}$ und jede beliebige fünf-elementige Menge sind gleichmächtig.
- Die Mengen $\{0, 1, 2, \ldots\}$ und $\{1, 2, \ldots\}$ sind gleichmächtig, denn die Abbildung $n \mapsto n + 1$ liefert eine Bijektion zwischen diesen beiden Mengen.





## Abzählbare Mengen

### Definition

*Eine nicht-leere Menge $A$ heißt abzählbar, wenn es eine surjektive Abbildung $\varphi : \mathbb{N} \to A$ gibt. Ansonsten heißt $A$ überabzählbar.*

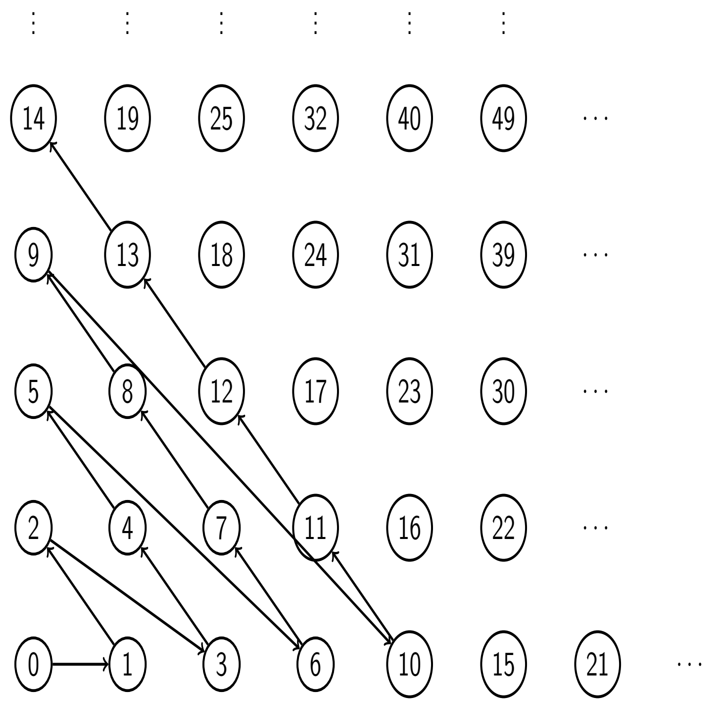
**Beispiele.**

- Jede endliche Menge $A = \{a_0, a_1, \ldots, a_N\}$ ist insbesondere abzählbar. Eine Surjektion $\varphi : \mathbb{N} \to A$ ist $(\varphi(n))_{n \in \mathbb{N}} = (a_0, a_1, \ldots, a_N, a_N, a_N, \ldots)$.
- Die Menge $\mathbb{Z}$ ist abzählbar. Eine Bijektion, und damit eine Surjektion, $\varphi : \mathbb{N} \to \mathbb{Z}$ ist gegeben durch $\varphi(0) = 0$, $\varphi(2k - 1) = k$ und $\varphi(2k) = -k$ für $k = 1, 2, \ldots$ D.h. $(\varphi(n))_{n \in \mathbb{N}} = (0, 1, -1, 2, -2, \ldots)$.

## Satz

*Es existiert eine bijektive Abbildung $\Psi : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$.*
*Insbesondere ist das Produkt $\mathbb{N} \times \mathbb{N}$ abzählbar.*

**Beweis.** Das beweist man mit einem (einfachen) Diagonalverfahren.







Explizit kann man die Abbildung $\Psi : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ auch durch folgende Formel beschreiben:

$$\Psi(n, m) = \frac{(n+m)(n+m+1)}{2} + m.$$

Man kann nun nachrechnen, dass diese Abbildung sowohl injektiv als auch surjektiv, also bijektiv ist. $\square$

## Folgerung

*Eine abzählbare Vereinigung von abzählbaren Mengen ist abzählbar.*

**Beweis.** Sei $M = \cup_{i \in \mathbb{N}} M_i$ mit Surjektionen $\rho_i : \mathbb{N} \to M_i$ für alle $i \in \mathbb{N}$. Wir wählen eine Bijektion $\Phi : \mathbb{N} \to \mathbb{N} \times \mathbb{N}$ (z.B. die Umkehrung der oben konstruierten Abbildung $\Psi : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$) und schreiben $\Phi_1 : \mathbb{N} \to \mathbb{N}$ und $\Phi_2 : \mathbb{N} \to \mathbb{N}$ für die beiden Komponenten von $\Phi$. Dann wird durch

$$\rho(n) = \rho_{\Phi_1(n)}(\Phi_2(n))$$

die gewünschte Surjektion $\rho : \mathbb{N} \to \cup_{i \in \mathbb{N}} M_i$ definiert. $\square$

# Abzählbarkeit von $\mathbb{Q}$

## Folgerung

*$\mathbb{Q}$ ist abzählbar.*

**Beweis.** $\mathbb{Q} = \mathbb{Z} \cup \{\frac{n}{2} : n \in \mathbb{Z}\} \cup \{\frac{n}{3} : n \in \mathbb{Z}\} \cup \ldots$. □

## Folgerung

*$\mathbb{Q}^n = \{(x_1, x_2, \ldots, x_n) | x_1, x_2, \ldots, x_n \in \mathbb{Q}\}$ ist abzählbar für alle $n \geq 1$.*

**Beweis.**
Beweis durch Induktion nach $n$.

- Der Induktionsanfang ist die Abzählbarkeit von $\mathbb{Q}$.
- Aus der Abzählbarkeit von $\mathbb{Q}^n$ folgt die von $\mathbb{Q}^{n+1}$, denn

$$\mathbb{Q}^{n+1} = \bigcup_{y \in \mathbb{Q}} \{(x, y) | x \in \mathbb{Q}^n\}.$$

  ist eine Vereinigung abzählbar vieler abzählbarer Mengen.

□





# Überabzählbarkeit von $\mathbb{R}$

## Satz

*Die Menge $\mathbb{R}$ der reellen Zahlen ist nicht abzählbar.*

**Beweis.** Wir müssen zeigen: Es gibt keine surjektive Abbildung $\phi : \mathbb{N} \to \mathbb{R}$, d.h. keine Folge von reellen Zahlen enthält alle reelle Zahlen. Sei also $\{x_0, x_1, x_2, \ldots\}$ eine beliebige Folge in $\mathbb{R}$ gegeben. Wir werden nun eine reelle Zahl $x$ mit $x \neq x_n$ für alle $n \in \mathbb{N}$ konstruieren.

- Es sei $I_0$ das abgeschlossene Intervall $[x_0 + 1, x_0 + 2]$, d.h. $x_0 \notin I_0$.
- Wir unterteilen $I_0$ in drei gleich große abgeschlossene Intervalle und nennen eines davon $I_1$, wobei die Bedingung $x_1 \notin I_1$ gelten soll.
- Durch Fortsetzung dieses rekursiven Verfahrens erhalten wir eine Intervallschachtelung von kleiner werdenden abgeschlossenen Intervallen $I_k \subseteq I_{k-1}$ der Länge $\frac{1}{3^k}$ mit $x_0, \ldots, x_n \notin I_n$.

Die Folge der oberen oder unteren Intervallgrenzen ist dann eine Cauchy-Folge und konvergiert gegen eine eindeutig bestimmte reelle Zahl $x \in \cap_{n \in \mathbb{N}} I_n$. Daraus folgt aber $x \neq x_n$ für alle $n \in \mathbb{N}$. □

# Supremumseigenschaft der reellen Zahlen

## Definition (Supremum und Infimum)

*Sei $M \subseteq \mathbb{R}$ eine Menge reeller Zahlen. Jede reelle Zahl $s$ mit*

$$x \le s \qquad (\text{bzw. } s \le x) \qquad \text{für alle } x \in M$$

*nennt man eine obere (bzw. untere) Schranke für $M$. Eine obere Schranke $s \in \mathbb{R}$ einer Menge $M \subseteq \mathbb{R}$ heißt Supremum von $M$, falls $s$ die kleinste obere Schranke ist, d.h. für jede obere Schranke $s'$ von $M$ gilt $s' \ge s$.*
*Entsprechend definiert man das Infimum als die größte untere Schranke.*
*Bezeichnung: $s = \sup M$ bzw. $s = \inf M$.*

**Bemerkung.** Es gibt höchstens ein Supremum bzw. Infimum.





## Satz (Supremumseigenschaft der reellen Zahlen)

*Jede nach oben beschränkte, nicht-leere Teilmenge $M \subseteq \mathbb{R}$ besitzt ein Supremum.*
*Jede nach unten beschränkte, nicht-leere Teilmenge $M \subseteq \mathbb{R}$ besitzt ein Infimum.*

## Die rationalen Zahlen erfüllen die Supremumseigenschaft nicht.

Betrachte $M = [0, \sqrt{2}] \cap \mathbb{Q}$. $M$ ist beschränkt, hat aber kein Supremum in $\mathbb{Q}$, da die reelle Zahl $\sqrt{2}$ von oben durch rationale Zahlen approximiert werden kann.

**Beweis der Supremumseigenschaft für $\mathbb{R}$.** Wir beweisen nur die Existenz des Supremums. (Die Existenz des Infimums beweist man ähnlich.) Dabei verwenden wir die Vollständigkeit von $\mathbb{R}$.

## Beweis der Supremumseigenschaft für $\mathbb{R}$

Wir konstruieren rekursiv $a_n \leq b_n$, $n \in \mathbb{N}$ mit

(i) $a_n \in M$,
(ii) $b_n$ ist eine obere Schranke für $M$,
(iii) $a_n \leq a_{n+1}$, also monoton wachsend
(iv) $b_{n+1} \leq b_n$ , also monoton fallend, und
(v) $b_n - a_n \leq \frac{1}{2^n}(b_0 - a_0)$.

- Wir beginnen mit $a_0 \in M$ und einer beliebigen oberen Schranke $b_0$ von $M$.
- Ausgehend von $a_0 \leq b_0, \ldots, a_n \leq b_n$ mit (i-v) konstruieren wir $a_{n+1} \leq b_{n+1}$: Sei dazu $m = \frac{1}{2}(a_n + b_n)$.
  - Falls $m$ eine obere Schranke von $M$ ist, sei $[a_{n+1}, b_{n+1}] := [a_n, m]$.
  - Wenn nicht, gibt es $d \in M \cap [m, b_n]$ mit $d > m$ und wir setzen dann $[a_{n+1}, b_{n+1}] := [d, b_n] \subseteq [m, b_n]$.

  In beiden Fällen gilt $a_{n+1} \leq b_{n+1}$ und die Eigenschaften (i-v) sind erfüllt.





Weiter im Beweis.

Für die $a_n$ und $b_n$ gilt dann:

- Die monoton wachsende Folge $(a_n)$ ist durch $b_0$ nach oben beschränkt.
- Die monoton fallende Folge $(b_n)$ ist durch $a_0$ nach unten beschränkt.
- Wegen (v) können wir schließen, dass $(a_n)$ und $(b_n)$ gegen *denselben* Grenzwert $c \in \mathbb{R}$ konvergieren.

**Behauptung.** *Es gilt $c = \sup M$.*

**Beweis.** 1) $c$ ist eine obere Schranke von $M$:

- Annahme: $\exists\ a \in M$ mit $a > c$.
- Wegen $c = \lim_{n\to\infty} b_n$ gäbe es dann $N \in \mathbb{N}$, mit $a > b_n$ für alle $n \geq N$.
- Das widerspricht der Definition von $b_n$ als einer oberen Schranke für $M$.

2) $c$ ist die *kleinste* obere Schranke von $M$:

- Annahme: es gibt eine kleinere obere Schranke $b < c$ von $M$.
- Wegen $c = \lim_{n\to\infty} a_n$ gäbe es dann $N \in \mathbb{N}$, mit $b < a_n$ für alle $n \geq N$.
- Das widerspricht der Tatsache, dass $a_n \in M$. $\square$

# Maximum und Minimum

## Definition

- *Sei $M \subseteq \mathbb{R}$ eine nach oben beschränkte nicht-leere Teilmenge von $\mathbb{R}$.*
- *Falls $\sup M \in M$ gilt, heißt $\sup M$ auch das Maximum der Menge $M$. Wir schreiben dann auch $\max M$ statt $\sup M$.*
- *Analog nennen wir für eine nach unten beschränkte Menge $M \subseteq \mathbb{R}$, falls $\inf M \in M$ gilt, $\min M := \inf M$ auch das Minimum der Menge $M$.*

## Notation

- Falls $\emptyset \neq M \subseteq \mathbb{R}$ nicht nach unten beschränkt ist, so schreiben wir $\inf M := -\infty$.
- Falls $\emptyset \neq M \subseteq \mathbb{R}$ nicht nach oben beschränkt ist, so schreiben wir $\sup M := \infty$.





# Limes inferior, Limes superior

## Definition

*Sei $(x_n)_{n\in\mathbb{N}}$ eine Folge reeller Zahlen. Dann ist die Folge $(\inf\{x_k \mid k \geq n\})_{n\in\mathbb{N}}$ entweder konstant $-\infty$ (Fall 1) oder nach oben unbeschränkt (Fall 2) oder eine beschränkte monoton wachsende Folge (Fall 3). Man definiert*

$$\liminf_{n\to\infty} x_n := \lim_{n\to\infty} (\inf\{x_k \mid k \geq n\}) \in \mathbb{R} \cup \{-\infty, \infty\},$$

*wobei im 3.Fall eine reelle Zahl erhalten wird, während wir im 1.Fall $-\infty = \lim_{n\to\infty}(-\infty)$, und im 2.Fall $\infty$ erhalten.*
*Analog definiert man*

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

*Es heißt $\liminf_{n\to\infty} x_n$ (bzw. $\limsup_{n\to\infty} x_n$) der Limes inferior (bzw. Limes superior) der Folge $(x_n)_{n\in\mathbb{N}}$.*

**Beispiel:** $x_n = (-1)^n \left(1 + \frac{1}{2^n}\right)$

$$\sup\{x_k \mid k \geq n\} = \max\{x_k \mid k \geq n\} = \begin{cases} 1 + \frac{1}{2^n}, & \text{falls } n \text{ gerade} \\ 1 + \frac{1}{2^{n+1}}, & \text{falls } n \text{ ungerade} \end{cases}$$

$$\inf\{x_k \mid k \geq n\} = \min\{x_k \mid k \geq n\} = \begin{cases} -\left(1 + \frac{1}{2^n}\right), & \text{falls } n \text{ ungerade} \\ -\left(1 + \frac{1}{2^{n+1}}\right), & \text{falls } n \text{ gerade} \end{cases}$$

Damit gilt:

$$\limsup_{n\to\infty} x_n = \lim_{n\to\infty} (\sup\{x_k \mid k \geq n\}) = 1,$$

$$\liminf_{n\to\infty} x_n = \lim_{n\to\infty} (\inf\{x_k \mid k \geq n\}) = -1.$$





**Bemerkungen.**

- Eine Folge $(x_n)_{n\in\mathbb{N}}$ reeller Zahlen konvergiert genau dann gegen $\ell \in \mathbb{R}$, wenn $\limsup\limits_{n\to\infty} x_n = \liminf\limits_{n\to\infty} x_n = \ell$.
- Ist $(x_n)_{n\in\mathbb{N}}$ eine reelle Folge mit $\limsup\limits_{n\to\infty} x_n = -\infty$, so existiert zu jeder reellen Zahl $K \in \mathbb{R}$ ein Index $N \in \mathbb{N}$ mit $x_n \leq K$ für alle $n \geq N$. Wir sagen dann, dass die Folge bestimmt gegen $-\infty$ divergiert und schreiben $\lim_{n\to\infty} x_n = -\infty$.
- Anlog definieren wir die Schreibweise $\lim\limits_{n\to\infty} x_n = \infty$, und sagen in diesem Fall, dass die Folge bestimmt gegen $\infty$ divergiert.
- Ist $(x_n)_{n\in\mathbb{N}}$ eine Folge reeller Zahlen, so existieren Teilfolgen $(x_{n_k})_{k\in\mathbb{N}}$ und $(x_{n_\ell})_{\ell\in\mathbb{N}}$ mit

$$\lim_{k\to\infty} x_{n_k} = \liminf_{n\to\infty} x_n \quad \text{und} \quad \lim_{\ell\to\infty} x_{n_\ell} = \limsup_{n\to\infty} x_n.$$

# Kapitel 6

## Stetigkeit





# Stetigkeit

### Definition (Folgenkriterium der Stetigkeit)

*Sei $D$ eine Teilmenge eines metrischen Raums $X$ und $p \in D$, und sei $Y$ ein metrischer Raum. Eine Abbildung $f : D \to Y$ heißt stetig in $p \in D$, wenn für jede gegen $p$ konvergierende Folge $(z_n)_{n\in\mathbb{N}}$ aus $D$ gilt:*

$$\lim_{n\to\infty} f(z_n) = f(p).$$

*Eine Funktion $f : D \to Y$ heißt stetig, wenn $f$ in allen Punkten $p \in D$ stetig ist.*

Man kann diese Definition insbesondere in den metrischen Räumen $\mathbb{R}$ und $\mathbb{C}$ anwenden.

### Erste Beispiele.

Sei $D \subseteq X$ und $f : D \to \mathbb{C}$, eine konstante Funktion, d.h. $f(x) = c$ für alle $x \in D$ und ein $c \in \mathbb{C}$. Dann ist $f$ stetig.

## Erste Beispiele (Fortsetzung)

Für jeden metrischen Raum $X$ ist die Identität $\mathrm{id}_X : X \to X$, $\mathrm{id}_X(x) = x$, eine stetige Abbildung.
Für $x$ in $\mathbb{R}$ bezeichne $\lfloor x \rfloor$ die größte ganze Zahl $\leq x$. Dann ist die Funktion $\mathbb{R} \to \mathbb{R}$, $x \mapsto \lfloor x \rfloor$ genau in den Punkten $p \in \mathbb{R} \setminus \mathbb{Z}$ stetig.

## Definition (Grenzwert einer Funktion)

*Sei $X$ ein metrischer Raum, $D \subseteq X$, $f : D \to \mathbb{C}$ eine Funktion und $p \in X$. Wir schreiben*

$$\lim_{z \to p} f(z) = q,$$

*falls es*

(i) *eine Folge $(z_n) \in D \setminus \{p\}$ gibt, die gegen $p$ konvergiert und*

(ii) $\lim\limits_{n \to \infty} f(z_n) = q$ *für jede solche Folge $(z_n)$ gilt.*

*Man nennt $q$ den Grenzwert der Funktion $f$ an der Stelle $p \in X$.*

**Achtung:** Wir setzen hier nicht voraus, dass $p \in D$ gilt!





## Bemerkung

Mit obiger Notation gilt dann für Intervalle $I \subseteq \mathbb{R}$:

$$f : I \to \mathbb{C} \quad \text{ist stetig in} \quad p \in I \quad \iff \begin{cases} \lim\limits_{z \to p} f(z) \text{ existiert} \\ \lim\limits_{z \to p} f(z) = f(p). \end{cases}$$

**Mit den Rechenregeln für Grenzwerte von Folgen sehen wir:**

(i) Die Funktionen $z \mapsto \bar{z}, |z|, \mathrm{Re}\, z, \mathrm{Im}\, z$ sind stetig.

(ii) Die Summe $f + g$ und das Produkt $f \cdot g$ in $p \in D$ stetiger Funktionen $f, g : D \to \mathbb{C}$ sind stetig in $p$.

(iii) $\frac{1}{f}$ ist stetig in $p \in D$, falls $f : D \to \mathbb{C}^* := \mathbb{C} \setminus \{0\} \subseteq \mathbb{C}$ stetig in $p$ ist.

(iv) Jede rationale Funktion

$$f(z) = \frac{a_n z^n + \cdots + a_0}{b_m z^m + \cdots + b_0}$$

mit $a_0, \ldots, a_n, b_0, \ldots, b_m \in \mathbb{C}$, ist stetig auf dem Definitionsbereich $D := \{z \in \mathbb{C} | b_m z^m + \cdots + b_0 \neq 0\} \subseteq \mathbb{C}$.

## Satz (Verkettung stetiger Funktionen ist stetig)

*Sei $f\colon D \to \mathbb{C}$ stetig im Punkt $p \in D \subseteq X$. Gelte $f(D) \subseteq E \subseteq \mathbb{C}$ und sei $g\colon E \to \mathbb{C}$ stetig in $f(p)$. Dann ist $g \circ f\colon D \to \mathbb{C}$ stetig im Punkt $p$.*

**Beweis.**

- Sei $(x_n)_{n\in\mathbb{N}} \subseteq D$ eine beliebige Folge mit der Eigenschaft
$$\lim_{n\to\infty} x_n = p.$$
- Wir erhalten eine Folge $(f(x_n))_{n\in\mathbb{N}} \subseteq E$. Wegen der Stetigkeit von $f$ im Punkt $p$ konvergiert diese Folge in $E$ und es gilt
$$\lim_{n\to\infty} f(x_n) = f(p).$$
- Aus der Stetigkeit von $g$ im Punkt $f(p)$ folgt schließlich
$$\lim_{n\to\infty} g(f(x_n)) = g(f(p)) = (g \circ f)(p),$$
d.h. $g \circ f$ ist stetig.

□



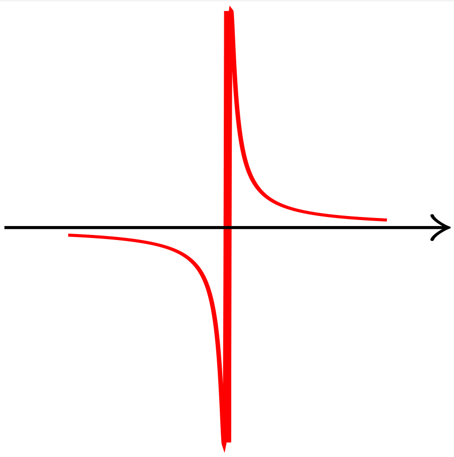


## Beispiele

(i) Jedes Polynom in einer stetigen Funktion $f$, also jede Funktion der Form $a_n f^n + a_{n-1} f^{n-1} \ldots + a_0$ ist stetig.

(ii) Die Funktionen $\bar{f}, |f|, \mathrm{Re}\, f, \mathrm{Im}\, f$ sind stetig in $p \in D$, wenn $f\colon D \to \mathbb{C}$ stetig in $p$ ist.

(iii) Die Funktion $f : \mathbb{R} \setminus \{0\} \to \mathbb{R}$, $f(x) = \sin(\frac{1}{x})$ ist als Verkettung stetiger Funktionen stetig, besitzt aber keine stetige Fortsetzung auf ganz $\mathbb{R}$.



## Satz

*Die Exponentialfunktion* $\exp\colon \mathbb{C} \to \mathbb{C}$ *ist stetig.*

**Beweis.**

- Es genügt zu zeigen, dass exp stetig in 0 ist. In der Tat:
  Sei $(z_n)$ eine konvergente Folge und $p = \lim_{n\to\infty} z_n$. Ist exp stetig in 0, dann konvergiert auch $\exp(z_n) = \exp(z_n - p)\exp(p)$ gegen $\exp(0)\exp(p) = \exp(p)$.
- Wir zeigen nun die Stetigkeit im Nullpunkt. Für $|z| < 1$ gilt:

$$\begin{aligned} |\exp(z) - 1| &= |z + \frac{z^2}{2!} + \frac{z^3}{3!} + \cdots| \le |z|(1 + \frac{|z|}{2!} + \frac{|z|^2}{3!} + \cdots) \\ &\le |z|(1 + \frac{1}{2!} + \frac{1}{3!} + \cdots) = |z| \cdot (e - 1) \end{aligned}$$

- Also erfüllt jede Nullfolge $(z_n)$ ab einem hinreichend großen Folgenglied die Ungleichung $|\exp(z_n) - 1| \le |z_n| \cdot (e-1)$, woraus folgt $\lim_{n\to\infty} \exp(z_n) = 1$.

$\square$





## Folgerung

*Die Funktionen* $\sin, \cos\colon \mathbb{C} \to \mathbb{C}$ *sind stetig.*

**Beweis.**

- Aus der Stetigkeit der Exponentialfunktion folgt die Stetigkeit der Funktionen $\mathbb{C} \to \mathbb{C}$, $z \mapsto \exp(iz)$ und $z \mapsto \exp(-iz)$.
- Die Stetigkeit von Sinus und Kosinus folgt nun aus der Definition:

$$\begin{aligned} \cos(z) &= \frac{1}{2}(\exp(iz) + \exp(-iz)) \\ \sin(z) &= \frac{1}{2i}(\exp(iz) - \exp(-iz)) \end{aligned}$$

  als Verkettung stetiger Funktionen.

$\square$

### Definition (Hyperbolische Funktionen)

*Sei $z \in \mathbb{C}$. Die durch*

$$\begin{aligned}\cosh(z) &= \frac{1}{2}(\exp(z) + \exp(-z)) \\ \sinh(z) &= \frac{1}{2}(\exp(z) - \exp(-z))\end{aligned}$$

*definierten Funktionen $\sinh\colon \mathbb{C} \to \mathbb{C}$ und $\cosh\colon \mathbb{C} \to \mathbb{C}$ heißen Sinus hyperbolicus bzw. Kosinus hyperbolicus.*

### Satz

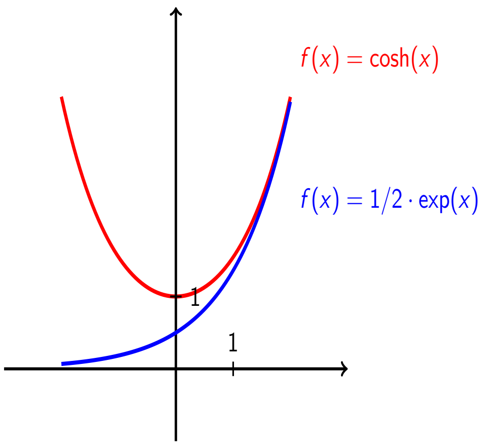
*Die Funktionen $\sinh, \cosh\colon \mathbb{C} \to \mathbb{C}$ sind stetig.*

**Beweis.**
Das folgt aus der Stetigkeit der Exponentialfunktion. □

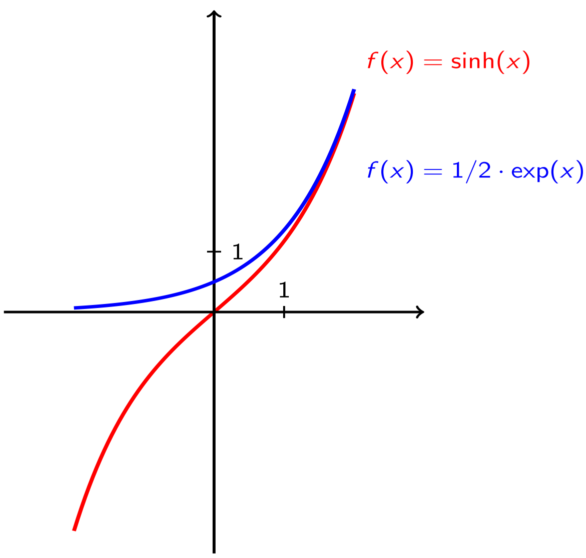




## Bilder im Reellen




Gerade Funktion, $\cosh(x) \geq 1$.




Ungerade Funktion.

Eigenschaften von Sinus hyperbolicus und Kosinus hyperbolicus, die für alle $z \in \mathbb{C}$ gelten:

- $\cosh^2(z) - \sinh^2(z) = 1$.
- (absolut konvergente Potenzreihenentwicklung).

$$\begin{aligned} \cosh z &= \sum_{k=0}^{\infty} \frac{z^{2k}}{(2k)!} = 1 + \frac{z^2}{2} + \frac{z^4}{4!} + \cdots, \\ \sinh z &= \sum_{k=0}^{\infty} \frac{z^{2k+1}}{(2k+1)!} = z + \frac{z^3}{3!} + \frac{z^5}{5!} + \cdots \end{aligned}$$

- (Symmetrie, bzw. Antisymmetrie)

$$\cosh(-z) = \cosh z, \quad \sinh(-z) = -\sinh z.$$

- (Additionstheoreme)

$$\begin{aligned} \cosh(z+w) &= \cosh z \cosh w + \sinh z \sinh w, \\ \sinh(z+w) &= \sinh z \cosh w + \cosh z \sinh w. \end{aligned}$$





## Zwischenwerteigenschaft stetiger reeller Funktionen

Im Folgenden sei $a < b$. Wir betrachten reellwertige stetige Funktionen auf dem abgeschlossenen Intervall $[a,b] = \{x \in \mathbb{R} | a \leq x \leq b\}$.

### Satz (Nullstellensatz von Bolzano)

*Sei $f : [a,b] \to \mathbb{R}$ eine stetige Funktion, so dass $f(a)f(b) < 0$. Dann existiert $c \in [a,b]$ mit $f(c) = 0$.*

### Satz (Zwischenwertsatz)

*Sei $g : [a,b] \to \mathbb{R}$ eine stetige Funktion und $d$ eine reelle Zahl zwischen $g(a)$ und $g(b)$. Dann gibt es $c \in [a,b]$ mit $g(c) = d$.*

**Beweis.** (Reduktion des Zwischenwertsatzes auf den Nullstellensatz)
Falls $g(a) < d < g(b)$ oder $g(b) < d < g(a)$, so erfüllt $f(x) := g(x) - d$ die Voraussetzungen des Nullstellensatzes.
Daher existiert $c \in [a,b]$ mit $0 = f(c) = g(c) - d$ und somit $g(c) = d$. $\square$

Beweis des Nullstellensatzes.

- Durch Ersetzen von $f$ durch $-f$, falls notwendig, können wir ohne Einschränkung annehmen, dass $f(a) < f(b)$ gilt.
- Wir konstruieren rekursiv Intervalle $[a_n, b_n]$, so dass
  (i) $[a_n, b_n] \subseteq [a_{n-1}, b_{n-1}]$ für alle $n \in \mathbb{N}$,
  (ii) $b_n - a_n = 2^{-n}(b-a)$,
  (iii) $f(a_n) \leq 0$ und $f(b_n) \geq 0$.
- Wir beginnen mit $[a_0, b_0] := [a, b]$.
- Seien $[a_0, b_0], [a_1, b_1], \ldots, [a_n, b_n]$ mit (i)–(iii) bereits konstruiert. Wir konstruieren $[a_{n+1}, b_{n+1}]$.
- Sei $m = \frac{1}{2}(a_n + b_n)$. Wir setzen:

$$[a_{n+1}, b_{n+1}] := \begin{cases} [a_n, m], & \text{falls} \quad f(m) \geq 0 \\ [m, b_n], & \text{falls} \quad f(m) < 0. \end{cases}$$

- Die Eigenschaften (i) bis (iii) sind dann erfüllt.



Stetigkeit | Stetige Funktionen auf beschränkten abgeschlossenen Intervallen

Weiter im Beweis des Nullstellensatzes:

- $(a_n)$ ist monoton wachsend, $(b_n)$ fallend und beide Folgen sind durch $a$ nach unten und durch $b$ nach oben beschränkt.
- Wegen (ii) können wir schließen, dass

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n =: c.$$

- Die Stetigkeit von $f$ ermöglicht den Grenzübergang in den Ungleichungen $f(a_n) \leq 0$ und $f(b_n) \geq 0$ und liefert somit

$$f(c) \stackrel{\text{stetig}}{=} \lim_{n\to\infty} f(a_n) \leq 0$$
$$f(c) \stackrel{\text{stetig}}{=} \lim_{n\to\infty} f(b_n) \geq 0$$

und somit $f(c) = 0$.

$\square$

## Beispiel

Die Vollständigkeit von $\mathbb{R}$ ist wesentlich:
Für stetige Funktionen $f : [a, b] \subseteq \mathbb{Q} \to \mathbb{R}$ gilt der Zwischensatz nicht.

ÜA: Die Funktion $f : \mathbb{Q} \to \mathbb{R}$, die definiert ist durch

$$f(x) := \begin{cases} -1, & \text{falls} \quad x < \sqrt{2} \\ 1, & \text{falls} \quad x > \sqrt{2} \end{cases}$$

ist stetig,
besitzt aber keine Fortsetzung zu einer stetigen Funktion $\tilde{f} : \mathbb{R} \to \mathbb{R}$.
D.h. es existiert keine stetige Funktion $\tilde{f} : \mathbb{R} \to \mathbb{R}$ mit $\tilde{f}(q) = f(q)$ $\forall\, q \in \mathbb{Q}$.





## Infimum und Supremum einer reellwertigen Funktion

### Definition

- *Sei $f : D \to \mathbb{R}$ eine reellwertige Funktion auf einer Menge $D$ (z.B. $D \subseteq \mathbb{C}$).*
- *Die Funktion $f$ heißt beschränkt, falls die Menge $f(D) \subseteq \mathbb{R}$ beschränkt ist;*

  *$f$ heißt nach unten (bzw. nach oben) beschränkt, falls $f(D)$ nach unten (bzw. nach oben) beschränkt ist.*
- *Wir setzen*

$$\begin{aligned} \inf f &:= \inf f(D) \in \mathbb{R} \cup \{-\infty\} \\ \sup f &:= \sup f(D) \in \mathbb{R} \cup \{\infty\}. \end{aligned}$$

- *Falls $\inf f \in f(D)$ bzw. $\sup f \in f(D)$, so schreibt man stattdessen auch $\min f$ bzw. $\max f$.*
- *Man sagt dann, dass die Funktion ihr Minimum bzw. Maximum annimmt.*

# Minimum-Maximum-Eigenschaft

## Theorem

*Jede stetige Funktion $f: [a, b] \to \mathbb{R}$ auf einem abgeschlossenen Intervall ist beschränkt und nimmt ihr Minimum und Maximum an,*
*d.h. es gibt $x_{min} \in [a, b]$ und $x_{max} \in [a, b]$ mit $f(x_{min}) = \min f$ und $f(x_{max}) = \max f$.*

**Beweis.** Wir beweisen indirekt, dass die Bildmenge $B := f([a, b])$ ihr Supremum enthält. Die andere Aussage folgt dann aus $\inf f = -\sup(-f)$.
Sei dazu $s := \sup B \in \mathbb{R} \cup \{\infty\}$.

- Im Beweis der Supremumseigenschaft der reellen Zahlen haben wir gesehen: Falls $s \neq \infty$, dann existiert eine Folge $b_n \in B$ mit $\lim_{n\to\infty} b_n = s$.
- Im Fall $s = \infty$ finden wir eine Folge $b_n \in B$ mit $b_n > n$. Nach Konvention gilt dann ebenfalls $\lim_{n\to\infty} b_n = s$.
- Wir wählen nun $x_n \in [a, b]$ mit $f(x_n) = b_n$.





Weiter im Beweis:

- Nach dem Satz von Bolzano-Weierstraß gibt es eine konvergente Teilfolge $(x_{n_k})_{k\in\mathbb{N}}$. Wir setzen $\ell := \lim\limits_{k\to\infty} x_{n_k} \in [a, b]$.
- Stetigkeit liefert nun $f(\ell) = \lim\limits_{k\to\infty} f(x_{n_k}) = \lim_{k\to\infty} b_{n_k} = s$.
- Also folgt $s \in B \subseteq \mathbb{R}$, d.h. die Bildmenge ist nach oben beschränkt und enthält ihr Supremum.

□

## Folgerung

*Sei $f: [a, b] \to \mathbb{R}$ stetig. Dann gilt $f([a, b]) = [\min f, \max f]$.*

**Beweis.** Das folgt aus dem vorherigen Satz und der Zwischenwerteigenschaft. □

# $\epsilon$-$\delta$-Definition der Stetigkeit

Wir erwähnen noch, dass man oft folgende äquivalente (ÜA!) Definition der Stetigkeit verwendet:

## $\epsilon$-$\delta$-Definition der Stetigkeit

Sei $D$ eine Teilmenge eines metrischen Raums $(X, d)$ und sei $p \in D$. Sei $(Y, d')$ ein weiterer metrischer Raum. Eine Funktion $f : D \to Y$ heißt stetig in $p$, wenn es zu jedem $\epsilon > 0$ ein $\delta > 0$ gibt, derart, dass gilt

$$d'(f(z), f(p)) < \epsilon \quad \text{für alle } z \in D \quad \text{mit } d(z, p) < \delta.$$





# Kapitel 7

Streng monotone Funktionen

### Definition

*Sei $D \subseteq \mathbb{R}$. Eine Funktion $f : D \to \mathbb{R}$ heißt monoton wachsend (bzw. streng monoton wachsend), falls*

$$f(x) \le f(x') \quad (\textit{bzw.} \quad f(x) < f(x'))$$

*für alle $x, x' \in D$ mit $x < x'$.*
*(Die Begriffe 'monoton fallend' und 'streng monoton fallend' werden analog definiert.)*

### Beispiel: Potenzfunktionen

Sei $k \in \mathbb{N}$. Die Funktion $x \mapsto x^k$ heißt Potenzfunktion.

- Ist $k$ ungerade, so ist die Potenzfunktion streng monoton wachsend auf ganz $\mathbb{R}$.
- Ist $k$ gerade, so ist die Potenzfunktion streng monoton wachsend auf $\{x \in \mathbb{R} \mid x \ge 0\}$ und streng monoton fallend auf $\{x \in \mathbb{R} \mid x \le 0\}$.





## Die Zahl $\pi$ und trigonometrische Funktionen

Wiederholung: Die Kosinusfunktion ist definiert als

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

### Theorem

*Die Kosinusfunktion hat im Intervall $[0, 2]$ genau eine Nullstelle $c$.*

### Definition

*Man definiert so die Zahl $\pi$ durch $\pi := 2c$.*

**Beweis.** Die Kosinusfunktion ist stetig und $\cos(0) = 1 > 0$. Wir zeigen:

1. $\cos(2) < 0$. Die Existenz einer Nullstelle $c$ folgt dann aus dem Nullstellensatz von Bolzano.
2. cos ist auf dem Intervall $[0, 2]$ streng monoton fallend. Somit ist $c$ die einzige Nullstelle in $(0, 2)$.

## Beweis des Theorems, Schritt 1: $\cos(2) < 0$

- Die folgende Abschätzung gilt für $|x| \le 7$:

$$
\begin{aligned}
\cos x &= 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots \\
&= 1 - \left(\frac{x^2}{2!} - \frac{x^4}{4!}\right) - \left(\frac{x^6}{6!} - \frac{x^8}{8!}\right) - \cdots \\
&= 1 - \frac{x^2}{2!}\left(1 - \frac{x^2}{3\cdot 4}\right) - \frac{x^6}{6!}\underbrace{\left(1 - \frac{x^2}{7\cdot 8}\right)}_{\ge 0 \text{ für } |x|\le 7,} - \cdots \quad \text{alle weiteren auch} \\
&< 1 - \frac{x^2}{2!}\left(1 - \frac{x^2}{3\cdot 4}\right).
\end{aligned}
$$

- Also $\cos 2 < 1 - 2(1 - \frac{1}{3}) = 1 - \frac{4}{3} = -\frac{1}{3} < 0$.





## Beweis des Theorems, Schritt 2: cos monoton fallend auf $[0, 2]$

- Wir haben zu zeigen $0 \le x < y \le 2 \Longrightarrow \cos x - \cos y > 0$.
- Man betrachtet $\alpha = \frac{x+y}{2} \in (0, 2)$ und $\beta = \frac{y-x}{2} \in (0, 1]$.
  Dann gilt $x = \alpha - \beta$ und $y = \alpha + \beta$.
- Aus den Additionstheoremen folgt nun:
  $\cos x - \cos y = \cos(\alpha - \beta) - \cos(\alpha + \beta) =$
  $\cos\alpha\cos\beta + \sin\alpha\sin\beta - (\cos\alpha\cos\beta - \sin\alpha\sin\beta) = 2\sin\alpha\sin\beta$.
- D.h. $\cos x - \cos y > 0$ für alle $0 \le x < y \le 2$ , falls $\sin z > 0$ für alle $z \in (0, 2)$.
- Das folgt aus:

$$
\begin{aligned}
\sin z &= \left(z - \frac{z^3}{3!}\right) + \left(\frac{z^5}{5!} - \frac{z^7}{7!}\right) + \cdots \\
&= z\left(1 - \frac{z^2}{2\cdot 3}\right) + \frac{z^5}{5!}\left(1 - \frac{z^2}{6\cdot 7}\right) + \cdots
\end{aligned}
$$

□

## Satz (Euler)

*Es ist* $\sin(\frac{\pi}{2}) = 1$ *und damit* $\exp(i\frac{\pi}{2}) = i$.

**Beweis.** Nach Definition von $\pi$ gilt $\cos\frac{\pi}{2} = 0$. Es folgt

$$(\sin\frac{\pi}{2})^2 = 1 - (\cos\frac{\pi}{2})^2 = 1$$

und somit $\sin\frac{\pi}{2} = 1$, denn $\sin > 0$ auf $(0, 2)$. □

## Folgerung

$$\exp(i\pi) = -1, \quad \exp(i\frac{3\pi}{2}) = -i \quad \textit{und} \quad \exp(i2\pi) = 1.$$

**Beweis.**
Das folgt aus $\exp(in\frac{\pi}{2}) = (\exp(i\frac{\pi}{2}))^n = i^n$ für $n = 2, 3$ und $4$. □





## Folgerung

(i) $\exp(z + 2\pi i) = \exp z$ *für alle* $z \in \mathbb{C}$,

(ii) $\cos(x + 2\pi) = \cos x$, $\sin(x + 2\pi) = \sin x$,

(iii) $\cos(x + \pi) = -\cos x$, $\sin(x + \pi) = -\sin x$ *und*

(iv) $\cos(x + \frac{\pi}{2}) = -\sin x$, $\sin(x + \frac{\pi}{2}) = \cos x$ *für alle* $x \in \mathbb{R}$.

**Beweis.** Aus der Funktionalgleichung für die Exponentialfunktion folgt:

$$\exp(z + in\frac{\pi}{2}) = \exp(z)i^n.$$

Für $n = 4$ erhalten wir (i). Hieraus erhalten wir mit $\cos(x) = Re\exp(ix)$ auch (ii).
Für $n = 2$ erhalten wir (iii), und für $n = 1$ erhalten wir (iv). □

## Folgerung

*Die Funktionen Sinus und Kosinus sind vollständig durch die Einschränkung* $\cos|_{[0,\frac{\pi}{2}]}$ *bestimmt.*

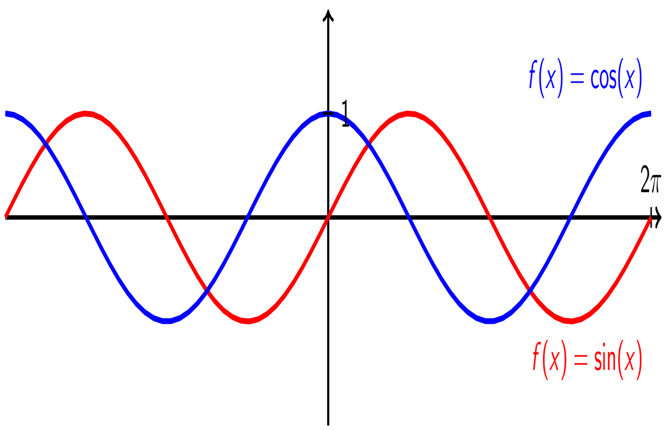
**Beweis.**

- Wegen $\cos x = \sin(x + \frac{\pi}{2})$ erhält man den Graphen der Sinusfunktion durch Verschiebung des Graphen der Kosinusfunktion um $\pi/2$ nach rechts.
- Die Kosinusfunktion ist vollständig durch die Einschränkung $\cos|_{[0,\frac{\pi}{2}]}$ bestimmt:
  - Wegen (iii) genügt es cos auf einem Intervall der Länge $\pi$ zu kennen, z.B. auf $[-\frac{\pi}{2}, \frac{\pi}{2}]$.
  - Wegen der Symmetrie $\cos(x) = \cos(-x)$, genügt $[0, \frac{\pi}{2}]$.

$\square$







### Folgerung

- *Die Funktionen* sin, cos *und* $\mathbb{R} \ni x \mapsto \exp(ix)$ *sind periodisch mit Periode* $2\pi$*. Es gilt* $\sin, \cos\colon \mathbb{R} \to [-1, 1]$.
- $\cos x = 0 \iff x = \frac{\pi}{2} + n\pi,\ n \in \mathbb{Z}$.
- $\sin x = 0 \iff x = n\pi,\ n \in \mathbb{Z}$.
- $\exp(ix) = 1 \iff x = n2\pi,\ n \in \mathbb{Z}$.
- *Die Kosinusfunktion ist auf dem Intervall* $[0, \pi]$ *streng monoton fallend und auf dem Intervall* $[\pi, 2\pi]$ *streng monoton wachsend.*
- *Die Sinusfunktion ist auf dem Intervall* $[-\frac{\pi}{2}, \frac{\pi}{2}]$ *streng monoton wachsend und auf dem Intervall* $[\frac{\pi}{2}, \frac{3\pi}{2}]$ *streng monoton fallend.*
- *Die Tangensfunktion* $\tan := \frac{\sin}{\cos}$ *(definiert dort, wo* $\cos \neq 0$*) ist auf dem Intervall* $(-\frac{\pi}{2}, \frac{\pi}{2})$ *streng monoton wachsend und* $\tan(-\frac{\pi}{2}, \frac{\pi}{2}) = \mathbb{R}$.
- *Die Kotangensfunktion* $\cot := \frac{\cos}{\sin}$ *(definiert dort, wo* $\sin \neq 0$*) ist auf dem Intervall* $(0, \pi)$ *streng monoton fallend und* $\cot(0, \pi) = \mathbb{R}$.





## Umkehrfunktionen

### Definition

*Sei* $D \subseteq X$ *und* $f\colon D \to Y$ *eine Funktion. Eine Funktion* $g\colon f(D) \to D$ *heißt Umkehrfunktion von* $f$*, falls* $g \circ f = Id_D$*, d.h.* $g(f(z)) = z\ \forall z \in D$.

### Bemerkungen:

(i) Eine Umkehrfunktion existiert genau dann, wenn $f\colon D \to Y$ injektiv ist.

(ii) Für $D \subseteq \mathbb{R}$, $f\colon D \to \mathbb{R}$ und $g\colon f(D) \to \mathbb{R}$ folgt aus $g(f(x)) = x$ für alle $x \in D$ auch $f(g(y)) = y$ für alle $y \in f(D)$.

(iii) Wenn $f\colon D \to \mathbb{R}$ eine Umkehrfunktion $g\colon f(D) \to D \subseteq \mathbb{R}$ besitzt, dann schreiben wir diese als $g = f^{-1}$. Vorsicht: Bei Funktionen mit Werten in einem Körper ist im Allgemeinen $f^{-1}(x) \neq f(x)^{-1} (= \frac{1}{f(x)})$.

(iv) Hat $f\colon D \subseteq \mathbb{R} \to \mathbb{R}$ eine Umkehrfunktion $f^{-1}$, so erhält man den Graphen von $f^{-1}$ durch Spiegelung des Graphen von $f$ an der Diagonalen $\{(x, x) \mid x \in \mathbb{R}\} \subseteq \mathbb{R}^2$,
$graph(f) := \{(x, f(x)) \mid x \in D\}$ und $graph(f^{-1}) = \{(f(x), x) \mid x \in D\}$

# Umkehrfunktionen streng monotoner Funktionen

*Erinnerung:* Sei $D \subseteq \mathbb{R}$. Eine Funktion $f: D \to \mathbb{R}$ heißt streng monoton wachsend/fallend, falls

$$f(x) < f(y) \quad \text{bzw.} \quad f(x) > f(y)$$

für alle $x, y \in D$ mit $x < y$.

**Satz**

*Sei $D \subseteq \mathbb{R}$. Jede streng monotone Funktion $f: D \to \mathbb{R}$ besitzt eine Umkehrfunktion $f^{-1}: f(D) \to D$. Die Umkehrfunktion ist wieder streng monoton, und zwar wachsend, wenn $f$ wachsend ist und fallend, wenn $f$ fallend ist.*





**Beweis.**

- Sei $y \in f(D)$, d.h. es existiert ein $x \in D$ mit $f(x) = y$.
  Da $f$ streng monoton ist, ist dieses $x$ eindeutig durch $y$ bestimmt: Gäbe es ein weiteres $x' \in D$ mit $f(x') = y$, dann wäre $x < x'$ oder $x > x'$ und damit $f(x) \neq f(x')$ wegen der strengen Monotonie von $f$.
- Wir definieren $g(y) := x$.
- Um zu zeigen, dass $g$ streng monoton ist, nehmen wir z.B. an, dass $f$ streng monoton wachsend ist, also
  $x < x' \Longrightarrow f(x) < f(x')$.
- Es gilt sogar $x < x' \Longleftrightarrow f(x) < f(x')$, denn $x \geq x' \Longrightarrow f(x) \geq f(x')$.
- Die Substitution $y = f(x)$ und $y' = f(x')$ liefert

$$g(y) < g(y') \Longleftrightarrow y < y'.$$

- Also ist $g$ streng monoton wachsend.

$\square$

## Satz

*Sei $f\colon [a,b] \to \mathbb{R}$ eine stetige streng monoton wachsende (bzw. streng monoton fallende) Funktion. Dann ist die streng monotone Umkehrfunktion $f^{-1}\colon [\min f, \max f] \to [a,b]$ auch stetig.*

**Beweis.**

- Wegen der Stetigkeit von $f$ gilt $f([a,b]) = [\min f, \max f]$.
- Da die Funkton $f$ streng monoton ist, besitzt sie eine streng monotone Umkehrfunktion $f^{-1}\colon [\min f, \max f] \to [a,b]$.
- Ist $f$ streng monoton wachsend, dann gilt $\min f = f(a)$ und $\max f = f(b)$.
- Ist $f$ fallend, so ist $\min f = f(b)$ und $\max f = f(a)$.





### Weiter im Beweis: Stetigkeit der Umkehrfunktion $f^{-1}$

- Sei $y_n \in [\min f, \max f]$ eine konvergente Folge, $y = \lim\limits_{n\to\infty} y_n$.
- Wir beweisen durch Widerspruch, dass die Folge der Urbilder $x_n := f^{-1}(y_n) \in [a,b]$ gegen $x := f^{-1}(y)$ konvergiert.
- Wenn $(x_n)$ nicht gegen $x$ konvergieren würde, so gäbe es ein $\varepsilon > 0$, so dass für alle $N \in \mathbb{N}$ eine natürliche Zahl $n \geq N$ existiert mit $|x_n - x| \geq \epsilon$.
- Daher kann man eine Teilfolge $(x_{n_k})_{k\in\mathbb{N}}$ konstruieren mit $|x_{n_k} - x| \geq \epsilon$ für alle $k$.
- Da $x_{n_k} \in [a,b]$, können wir durch Übergang zu einer noch feineren Teilfolge annehmen, dass $(x_{n_k})$ gegen $x' \in [a,b]$ konvergiert (Bolzano-Weierstraß). Nach Konstruktion gilt $x' \neq x$.
- Aus der Stetigkeit von $f$ erhalten wir nun

$$y_{n_k} = f(x_{n_k}) \underset{k\to\infty}{\longrightarrow} f(x') \underset{(f \text{ str. mon.})}{\neq} f(x) = y$$

Im Widerspruch zu der Voraussetzung $\lim\limits_{n\to\infty} y_n = y$. □

# Beispiel 1: Potenz- und Wurzelfunktionen

## Satz

*Sei $k \in \mathbb{N}, k \geq 1$. Die Potenzfunktion $x \mapsto f(x) = x^k$ definiert eine stetige und streng monoton wachsende Funktion $f : [0, \infty) \to \mathbb{R}$.*
*Die Umkehrfunktion $f^{-1} : [0, \infty) \to \mathbb{R}$, $f^{-1}(x) =: \sqrt[k]{x} =: x^{\frac{1}{k}}$, ist ebenfalls stetig und streng monoton wachsend.*

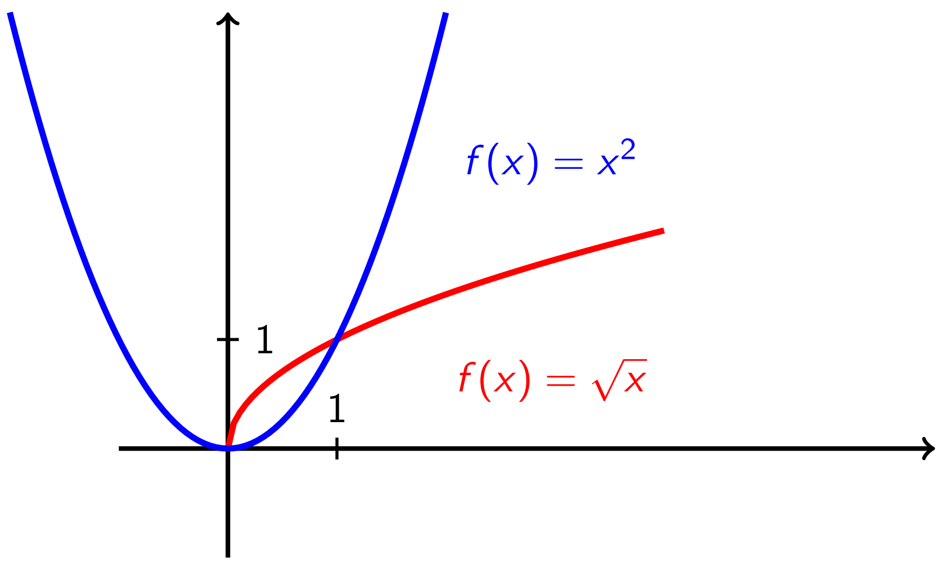
**Beweis.** Das folgt durch Anwendung des vorhergehenden Satzes auf die Einschränkung $f|_{[0,n]} : [0, n] \to [0, n^k] \subseteq \mathbb{R}$, $n = 1, 2, \ldots$. □

Für ungerades $k$ ist die stetige Funktion $x \mapsto x^k$ streng monoton wachsend auf ganz $\mathbb{R}$, ebenso wie die Umkehrfunktion $f^{-1} : \mathbb{R} \to \mathbb{R}$, $f^{-1}(x) =: \sqrt[k]{x} =: x^{\frac{1}{k}}$.





# Die Wurzelfunktion



# Beispiel 2: Exponential- und Logarithmusfunktion

**Satz**

- *Die reelle Exponentialfunktion* $\exp\colon \mathbb{R} \to \mathbb{R}$ *ist stetig, streng monoton wachsend und erfüllt* $\exp(\mathbb{R}) = \mathbb{R}_+ := \{x \in \mathbb{R} \mid x > 0\}$.
- *Die Umkehrfunktion*

$$\ln := \exp^{-1}\colon \mathbb{R}_+ \to \mathbb{R} \quad \textit{(natürlicher Logarithmus)}$$

*erfüllt die Gleichung*

$$\ln(xy) = \ln(x) + \ln(y)$$

*für alle* $x, y \in \mathbb{R}_+$.

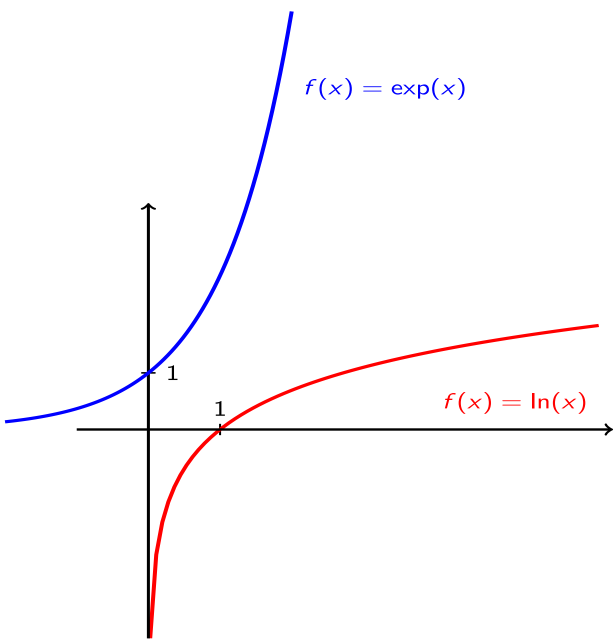
**Bemerkung**

Der Beweis liefert zwar die Existenz der Funktion ln, aber keine Berechnungsvorschrift!





# Die Logarithmusfunktion



**Beweis.** Die Exponentialfunktion ist stetig. Wir beweisen nun

1. $\exp(\mathbb{R}) \subseteq \mathbb{R}_+$:
   - Für alle $x \geq 0$ ist $\exp(x) = 1 + x + \frac{x^2}{2!} + \cdots \geq 1 > 0$.
   - Wegen $\exp(x)\exp(-x) = \exp(0) = 1$, folgt daraus $\exp(-x) = \frac{1}{\exp(x)} > 0$.
2. $\exp\colon \mathbb{R} \to \mathbb{R}$ streng monoton wachsend:
   - Für $x < y$ haben wir $\exp(y - x) > 1$ und somit:

$$\exp(y) = \exp(y - x + x) = \exp(y - x)\exp(x) > \exp(x).$$

3. $\exp(\mathbb{R}) = \mathbb{R}_+$:
   - Für alle $n \in \mathbb{N}$ gilt $\exp(n) = 1 + n + \frac{n^2}{2!} + \cdots \quad \geq 1 + n$
   - und somit $\exp(-n) = (\exp n)^{-1} \leq \frac{1}{1+n}$.
   - Damit folgt mit dem Zwischenwertsatz:

$$\exp(\mathbb{R}) = \bigcup_{n\in\mathbb{N}} \exp([-n, n]) \supset \bigcup_{n\in\mathbb{N}} [\frac{1}{1+n}, 1 + n] = \mathbb{R}_+.$$





## Schluss des Beweises

Wir haben gezeigt, dass $\exp\colon \mathbb{R} \to \mathbb{R}_+$ stetig, streng monoton wachsend und surjektiv ist. Daher existiert eine stetige, streng monoton wachsende Umkehrfunktion $\ln = \exp^{-1}\colon \mathbb{R}_+ \to \mathbb{R}$.
Es bleibt noch die Funktionalgleichung für den Logarithmus zu beweisen:

- Seien $x, x' \in \mathbb{R}_+$. Wir schreiben $x = \exp(y)$ und $x' = \exp(y')$, wobei $y = \ln(x)$, $y' = \ln(x') \in \mathbb{R}$.
- Die Funktionalgleichung der Exponentialfunktion liefert dann

$$xx' = \exp(y)\exp(y') = \exp(y + y').$$

- Anwendung des Logarithmus ergibt:

$$\ln(xx') = \ln(\exp(y + y')) = y + y' = \ln(x) + \ln(x').$$

□

## Beispiel 3: Exponentialfunktion zur Basis $a$

Wir hatten gesehen, dass

$$\exp(n) = \exp(1)^n = e^n.$$

Wir wollen dies nun verallgemeinern, indem wir $e$ durch eine beliebige reelle Zahl $a > 0$ ersetzen.

### Definition

*Sei $a > 0$. Die Funktion*

$$\begin{aligned} \exp_a : \mathbb{R} &\to \mathbb{R} \\ x &\mapsto \exp(x \ln(a)). \end{aligned}$$

*heißt Exponentialfunktion zur Basis a.*

### Es gilt:

- $\exp_e(x) = \exp(x)\ \forall x \in \mathbb{R}$, denn $\ln(e) = \ln \exp(1) = 1$.
- Für $n \in \mathbb{N}$ gilt $\exp_a(n) = \exp(n \ln(a)) = (\exp(\ln(a)))^n = a^n$.





### Satz

*Sei $a > 0$. Die Funktion $\exp_a : \mathbb{R} \to \mathbb{R}$ ist stetig und erfüllt:*

(i) $\exp_a(x + y) = \exp_a(x) \exp_a(y)$ *für alle $x, y \in \mathbb{R}$,*

(ii) $\exp_a(n) = a^n$, $\exp_a(-n) = \frac{1}{a^n}$ *für alle $n \in \mathbb{N}$ und*

(iii) $\exp_a(\frac{1}{n}) = a^{\frac{1}{n}} = \sqrt[n]{a}$ *für alle $n \in \mathbb{N}$. (n-te Wurzel von a.)*

**Beweis.**

- $\exp_a$ ist als Verkettung $g \circ f$ der stetigen Funktionen $g = \exp$ und $x \mapsto f(x) = x \ln(a)$ stetig.
- (i)–(iii) sind Folgerungen aus der Funktionalgleichung von exp.

□

### Definition

*Für $a > 0$ und $x \in \mathbb{R}$ setzen wir*

$$a^x := \exp_a(x).$$

### Satz

*Für alle $a, b \in \mathbb{R}_+$, $x, y \in \mathbb{R}$ gilt:*

(i) $a^{x+y} = a^x a^y$,

(ii) $(a^x)^y = a^{xy}$,

(iii) $a^x b^x = (ab)^x$,

(iv) $(\frac{1}{a})^x = a^{-x}$.

**Beweis.**

(i) ist die Funktionalgleichung von $\exp_a$.

(ii) $(a^x)^y = \exp(y \ln(a^x)) = \exp(y \ln(\exp(x \ln a))) = \exp(yx \ln a) = a^{yx} = a^{xy}$.

(iii) $a^x b^x = \exp(x \ln a) \exp(x \ln b) = \exp(x \ln a + x \ln b) = \exp(x(\ln a + \ln b)) = \exp(x(\ln(ab)) = (ab)^x$.

(iv) $(\frac{1}{a})^x = (a^{-1})^x \underset{(ii)}{=} a^{-x}$.

□





### Satz

i) $\exp_1(x) \equiv 1$, $x \in \mathbb{R}$.

ii) *Falls $a > 1$, so ist die Funktion $\exp_a : \mathbb{R} \to \mathbb{R}$ streng monoton wachsend.*

iii) *Falls $0 < a < 1$, so ist sie streng monoton fallend.*

iv) *Für $0 < a \neq 1$ gilt $\exp_a(\mathbb{R}) = \mathbb{R}_+$.*

**Beweis.**

i)–iii) Es ist $\ln 1 = \ln \exp(0) = 0$ und damit $\exp_1(x) = \exp(x \cdot 0) = 1$. Außerdem ist $\ln : \mathbb{R}_+ \to \mathbb{R}$ streng monoton wachsend. Daher gilt $\ln(a) > 0$ für $a > 1$ und $\ln(a) < 0$ für $0 < a < 1$. Entsprechend ist $x \mapsto \exp_a(x) = \exp(x \ln a)$ streng monoton wachsend bzw. fallend.

iv) Da $\ln a \neq 0$, haben wir $\exp_a(\mathbb{R}) = \exp(\mathbb{R}) = \mathbb{R}_+$. □

### Definition

*Sei* $0 < a \neq 1$*. Die Umkehrfunktion* $\log_a := \exp_a^{-1} \colon \mathbb{R}_+ \to \mathbb{R}$ *der Exponentialunktion* $\exp_a \colon \mathbb{R} \to \mathbb{R}_+$ *zur Basis a heißt Logarithmusfunktion zur Basis a.*

### Satz

*Es gilt*

$$\log_a = \frac{\ln}{\ln a}$$

*(insbesondere also* $\log_e = \ln$*).*

**Beweis.** Für alle $x \in \mathbb{R}_+$ gilt $\exp_a(\frac{\ln x}{\ln a}) = \exp(\frac{\ln x}{\ln a} \ln a) = \exp(\ln x) = x$. $\square$





## Charakterisierung der Exponentialfunktionen

### Theorem

*Sei* $f \colon \mathbb{R} \to \mathbb{R}$ *eine Funktion, die stetig in* $0 \in \mathbb{R}$ *ist, nicht konstant null ist und die die Funktionalgleichung*

$$(*) \quad f(x+y) = f(x)f(y) \quad \textit{für alle} \quad x, y \in \mathbb{R}$$

*erfüllt. Dann gilt* $f(1) =: a > 0$ *und* $f = \exp_a$.

**Beweis.** Wir zeigen zuerst, dass $f(\mathbb{R}) \subseteq (0, \infty)$ und $f(0) = 1$.

- Wegen $f(x) = f(\frac{x}{2} + \frac{x}{2}) = f(\frac{x}{2})^2$ gilt $f(x) \geq 0$ für alle $x \in \mathbb{R}$.

  Außerdem folgt aus $f(x) = f(x+0) \overset{(*)}{=} f(x) \cdot f(0)$ auch $f(0) = 1$, da sonst $f(x) = 0$ für *alle* $x \in \mathbb{R}$ im Widerspruch zur Voraussetzung.
- Wäre $f(x) = 0$ für *ein* $x \in \mathbb{R}$, so hätten wir wegen $f(x) = f(\frac{x}{n})^n$ auch $f(\frac{x}{n}) = 0$ für alle $n \in \mathbb{N}$, also wegen der Stetigkeit in Null auch $f(0) = 0$, im Widerspruch zum gerade Gezeigten.
- Also gilt $f(x) > 0$ für alle $x \in \mathbb{R}$, insbesondere $a = f(1) > 0$.

Weiter im Beweis:

Die Stetigkeit von $f$ zeigt man genau wie im Beweis für die Exponentialfunktion:
Konvergiert die Folge $(x_n)$ gegen $x$, so folgt aus

$$f(x_n) = f(x + x_n - x) \overset{(*)}{=} f(x)f(x_n - x)$$

und der Stetigkeit von $f$ in 0 auch die Konvergenz der Folge $(f(x_n))$ gegen $f(x)$.
Es bleibt zu zeigen, dass $f(x) = a^x$ für alle $x \in \mathbb{R}$, wobei $a = f(1)$.

1. $f(n) = a^n$ für alle $n \in \mathbb{Z}$:
   - Wir wissen schon: $f(0) = 1 = a^0$.
   - Für $n \in \mathbb{N}$ haben wir $f(n) \overset{(*)}{=} f(1) \cdots f(1) = a^n$ und
   - Aus $1 = f(n - n) \overset{(*)}{=} a^n f(-n)$ folgt $f(-n) = a^{-n}$.
2. Sei $\frac{p}{q} \in \mathbb{Q}$ mit $p \in \mathbb{Z}$ und $q \in \mathbb{N}$:
   - $f\left(\frac{p}{q}\right)^q = \underbrace{f\left(\frac{p}{q}\right) \cdots f\left(\frac{p}{q}\right)}_{\text{q Faktoren}} \overset{(*)}{=} f(p) = a^p > 0$. Also $f(\frac{p}{q}) = \sqrt[q]{a^p} = a^{\frac{p}{q}}$.





Sei nun $x \in \mathbb{R}$ und $(x_n)$ eine Folge in $\mathbb{Q}$ mit Grenzwert $x$.
Aus der Stetigkeit von $f$ und $\exp_a$ und der Übereinstimmung der Funktionen für Argumente aus $\mathbb{Q}$ folgt nun

$$f(x) = \lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} a^{x_n} = a^x.$$

$\square$

**Bemerkung**

Der Schluss des Beweises enthält folgende allgemeine Aussage: zwei *stetige* Funktionen $f : \mathbb{R} \to \mathbb{R}$ und $g : \mathbb{R} \to \mathbb{R}$, welche für alle $x \in \mathbb{Q}$ übereinstimmen, sind gleich.

# Uneigentliche Grenzwerte

## Definition

- *Für eine Folge* $(x_n)$ *reeller Zahlen hatten wir vereinbart:* $\lim_{n\to\infty} x_n = +\infty$, *falls* $\liminf_{n\to\infty} x_n = \infty$, *d.h. falls es für jedes* $K > 0$ *ein* $N = N(K) \in \mathbb{N}$ *gibt, so dass* $x_n \geq K$ *für alle* $n > N = N(K)$.
- *Genauso haben wir* $\lim_{n\to\infty} x_n = -\infty$, *falls* $\limsup_{n\to\infty} x_n = -\infty$.
- *In beiden Fällen nennt man* $x_n$ *bestimmt divergent.*
- *Analog definiert man nun uneigentliche Grenzwerte von Funktionen:* $\lim_{x\to a} f(x) = +\infty$, *wenn für jede Folge* $x_n \to a$ *mit* $x_n \neq a$ *die Folge* $f(x_n)$ *bestimmt gegen* $+\infty$ *divergent ist. Divergenz gegen* $-\infty$ *definiert man genauso.*

## Beispiele:

$\lim_{x\to 1} \frac{1}{(x-1)^2} = \infty$, $\lim_{x\to 0} \ln x = -\infty$, $\lim_{x\to 0} \frac{1}{x}$ existiert nicht.





# Mehr uneigentliche Grenzwerte von Funktionen

## Definition

- *Ist die Funktion* $f$ *auf einem halboffenen Intervall* $[c, \infty)$ *definiert, so sagen wir* $\lim_{x\to\infty} f(x) = a \in \mathbb{R}$, *falls für jede bestimmt gegen* $\infty$ *divergente Folge von Argumenten* $x_n$ *die Folge* $f(x_n)$ *gegen* $a$ *konvergiert.*
- *Divergieren die Funktionswerte* $f(x_n)$ *für jede solche Folge* $(x_n)$ *bestimmt gegen* $+\infty$, *so sagen wir* $\lim_{x\to\infty} f(x) = +\infty$.
- *Analog definiert man uneigentliche Grenzwerte einer auf einem Intervall der Form* $(-\infty, c]$ *definierten Funktion bei* $-\infty$.

## Beispiel:

$\lim_{x\to-\infty} e^x = 0$, $\lim_{x\to+\infty} e^x = +\infty$, $\lim_{x\to+\infty} \ln x = +\infty$.

# Exponentielles Wachstum

## Satz

*Für alle $k \in \mathbb{N}$ gilt*

$$\lim_{x\to+\infty} \frac{\exp x}{x^k} = +\infty.$$

**Beweis.** Für alle $x \geq 0$ gilt $\exp x = \sum_{n=0}^{\infty} \frac{x^n}{n!} \geq \frac{x^{k+1}}{(k+1)!}$ und somit

$$\frac{\exp x}{x^k} \geq \frac{x}{(k+1)!} \underset{x\to+\infty}{\longrightarrow} +\infty.$$

$\square$

## Bemerkung

Die Exponentialfunktion wächst also schneller als jede Potenzfunktion.





# Logarithmisches Wachstum

## Satz

*Für alle $k \in \mathbb{N}, k \geq 1$, gilt*

$$\lim_{x\to+\infty} \frac{\sqrt[k]{x}}{\ln x} = +\infty.$$

**Beweis.**

- Für alle $x > 1$ gilt $x = \exp(ky)$, mit $y = \frac{\ln x}{k} > 0$.
- Also $\frac{\sqrt[k]{x}}{\ln x} = \frac{\sqrt[k]{\exp(ky)}}{\ln x} = \frac{\sqrt[k]{(\exp y)^k}}{\ln x} = \frac{\exp y}{ky}$.
- Mit $x \to +\infty$ geht auch $y \to +\infty$:

$$\lim_{x\to+\infty} \frac{\sqrt[k]{x}}{\ln x} = \lim_{y\to+\infty} \frac{\exp y}{ky} = +\infty.$$

$\square$

## Bemerkung

Der Logarithmus wächst also langsamer als jede Wurzelfunktion.

# Beispiel 4: Trigonometrische Funktionen und Arcusfunktionen

**Wiederholung (Trigonometrische Funktionen im Reellen).**

- Die Funktionen sin, cos und $\mathbb{R} \ni x \mapsto \exp(ix)$ sind periodisch mit Periode $2\pi$.
- $\cos x = 0 \iff x = \frac{\pi}{2} + n\pi,\ n \in \mathbb{Z}$.
- $\sin x = 0 \iff x = n\pi,\ n \in \mathbb{Z}$.
- $\exp(ix) = 1 \iff x = n2\pi,\ n \in \mathbb{Z}$.
- Die Kosinusfunktion ist auf dem Intervall $[0, \pi]$ streng monoton fallend.
- Die Sinusfunktion ist auf dem Intervall $[-\frac{\pi}{2}, \frac{\pi}{2}]$ streng monoton wachsend.
- Die Tangensfunktion $\tan := \frac{\sin}{\cos}$ (definiert dort wo $\cos \neq 0$) ist auf dem Intervall $(-\frac{\pi}{2}, \frac{\pi}{2})$ streng monoton wachsend und $\tan\left((-\frac{\pi}{2}, \frac{\pi}{2})\right) = \mathbb{R}$. Der Tangens ist periodisch mit Periode $\pi$.





# Arcusfunktionen

Durch Einschränkung der trigonometrischen Funktionen auf die Intervalle, wo sie streng monoton sind, erhalten wir deren Umkehrfunktionen. Wir schreiben z.B.

$$\cos|_{[0,\pi]} : [0, \pi] \to [-1, 1]$$

für die Funktion, die wir durch Einschränkung des Kosinus' auf das Intervall $[0, \pi]$ erhalten.

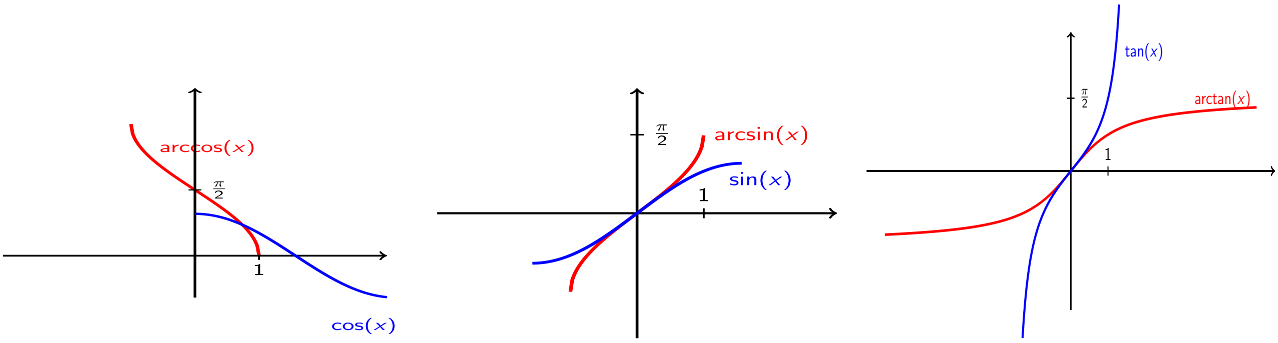
### Definition

*Die (streng montonen und stetigen) Umkehrfunktionen der trigonometrischen Funktionen heißen Arcus-Kosinus, Arcus-Sinus und Arcus-Tangens:*

- $\arccos := (\cos|_{[0,\pi]})^{-1} : [-1, 1] \to \mathbb{R}$,
- $\arcsin := (\sin|_{[-\frac{\pi}{2},\frac{\pi}{2}]})^{-1} : [-1, 1] \to \mathbb{R}$ *und*
- $\arctan := (\tan|_{(-\frac{\pi}{2},\frac{\pi}{2})})^{-1} : \mathbb{R} \to \mathbb{R}$.

# Arcusfunktionen

- $\arccos := (\cos|_{[0,\pi]})^{-1} \colon [-1,1] \to \mathbb{R}$,
- $\arcsin := (\sin|_{[-\frac{\pi}{2},\frac{\pi}{2}]})^{-1} \colon [-1,1] \to \mathbb{R}$ und
- $\arctan := (\tan|_{(-\frac{\pi}{2},\frac{\pi}{2})})^{-1} \colon \mathbb{R} \to \mathbb{R}$








# Polarkoordinaten

**Satz**

*Jede komplexe Zahl $z \neq 0$ besitzt eine Darstellung*

$$z = r\exp(i\varphi),$$

*mit $r = |z| > 0$ und $\varphi \in \mathbb{R}$; dabei ist $\varphi$ bis auf die Addition eines ganzen Vielfachen von $2\pi$ bestimmt.*

**Definition (Polarkoordinaten)**

*Das Paar $(r,\varphi)$ heißt die Polarkoordinaten von $z = r\exp(i\varphi)$ und $\arg(z) = \varphi$, für $\varphi \in [0, 2\pi)$, das Argument von $z$.*

**Beweis.** Es genügt, die Aussage für $z$ mit $\mathrm{Im} z \geq 0$ zu zeigen, denn falls $\mathrm{Im} z \leq 0$ ist, so gilt $\mathrm{Im}\bar{z} \geq 0$, und falls $\bar{z} = r\exp(i\varphi)$, so muss $z = r\exp(-i\varphi)$ gelten.

**Weiter im Beweis.**
Sei also nun $\mathrm{Im} z \geq 0$. Wir schreiben

$$\frac{z}{|z|} = \xi + i\eta, \qquad (\xi, \eta \in \mathbb{R}).$$

Es gilt dann $\xi^2 + \eta^2 = 1$ und $\eta \geq 0$. Es sei

$$\varphi = \arccos(\xi),$$

dann ist $\varphi \in [0, \pi]$ und weiter

$$\cos(\varphi) = \xi \text{ und } \sin(\varphi) = \eta \geq 0.$$

Also erhalten wir eine Darstellung

$$z = |z| \exp(i\varphi).$$

Wegen der Periodizität der komplexen Exponentialfunktion folgt die Eindeutigkeit dieser Darstellung bis auf ganze Vielfache von $2\pi$. □





### Bemerkungen

Wie wir bereits zuvor erwähnt haben, ist die Multiplikation von komplexen Zahlen, die in Polarkoordinaten gegeben sind, sehr einfach:
Sind nämlich $z = r\exp(i\varphi)$ und $w = t\exp(i\rho)$ gegeben so ist

$$zw = (rt)\exp(i(\varphi + \rho)).$$

Die Absolutbeträge multiplizieren sich, die Argumente addieren sich.
Ebenso leicht kann man Wurzeln aus einer komplexen Zahl ziehen:
Ist nämlich $z = r\exp(i\varphi)$, dann ist

$$w = \sqrt[k]{r}\exp(\frac{i\varphi}{k})$$

eine $k$-te Wurzel von $z$, d.h. $w^k = z$.
Gibt es außer $w$ weitere $k$-te Wurzeln von $z$?

### Satz

*Die Gleichung $z^n = 1$, $n \in \mathbb{N}$ besitzt genau die $n$ Lösungen*

$$\zeta^k = \exp\left(k\frac{2\pi i}{n}\right) = \cos(k2\pi/n) + i\sin(k2\pi/n),$$

*mit $k = 0, 1, \ldots, n-1$, diese werden die n-ten Einheitswurzeln genannt.*

**Beweis.** Offensichtlich erfüllen diese $n$ verschiedenen Zahlen die Gleichung $z^n = 1$. Es gibt keine weiteren Lösungen, denn es gilt:

### Lemma

*Sei $p(z) = a_n z^n + \ldots a_1 z + a_0$ ein Polynom. Dann hat $p$ höchstens $n$ verschiedene Nullstellen.*

**Beweis.** Dies gilt, da ein Polynom für jede Nullstelle $\lambda$ durch den Linearfaktor $(z - \lambda)$ teilbar ist. □





### Folgerung

*Für $0 \neq c \in \mathbb{C}$ mit $c = r\exp(i\varphi)$ hat die Gleichung $z^n = c$ mit $n \in \mathbb{N}$ genau die $n$ verschiedenen Lösungen*

$$\sqrt[n]{r}\exp\left(i\frac{\varphi + k2\pi}{n}\right),$$

*wobei $k \in \{0, \ldots, n-1\}$.*

Die Existenz einer Lösung von $z^n = c$ ist ein Spezialfall des Fundamentalsatzes der Algebra:

### Satz (Fundamentalsatz der Algebra)

*Jede polynomiale Gleichung mit komplexen Koeffizienten $c_k$ der Form*

$$z^n + c_{n-1}z^{n-1} + \ldots + c_1 z + c_0 = 0$$

*mit $n \geq 1$ hat mindestens eine komplexe Lösung.*

**Beweis.**
Wir schreiben $P(z) = z^n + c_{n-1}z^{n-1} + \ldots + c_1 z + c_0$ und setzen

$$\mu = \inf_{z \in \mathbb{C}} |P(z)|.$$

Für $|z| = R$ gilt wegen der Dreiecksungleichung

$$|P(z)| \geq R^n(1 - |c_{n-1}|R^{-1} - \ldots - |c_0|R^{-n}). \qquad (4)$$

Für große $R$ strebt die rechte Seite von (4) gegen $\infty$.
Also gibt es ein $R_0$, so dass $|P(z)| > \mu$ für alle $z$ mit $|z| > R_0$.

**Behauptung (ÜA)**

Analog zu stetigen reellen Funktionen auf abgeschlossenen Intervallen, nimmt die stetige Funktion $|P(z)|\colon \mathbb{C} \to \mathbb{R}_+$ auf der abgeschlossenen Kreisscheibe um 0 mit Radius $R_0$ ein Minimum an, d.h. es gibt ein $z_0$ mit $|z_0| \leq R_0$, so dass

$$|P(z_0)| = \mu.$$





Weiter im Beweis.

Wir zeigen nun per Widerspruch, dass $\mu = 0$.
Sei also $P(z_0) \neq 0$. Wir definieren das Polynom $Q$ durch

$$Q(z) := P(z + z_0)/P(z_0).$$

Wegen $P(z_0) = \min |P|$ gilt $|Q(z)| \geq 1$ für alle $z$. Außerdem ist $Q(0) = 1$.
Deswegen gibt es ein $k$ mit $1 \leq k \leq n$ so dass

$$Q(z) = 1 + b_k z^k + b_{k+1} z^{k+1} + \ldots + b_n z^n \quad \text{mit } b_k \neq 0.$$

Sei nun $e^{i\theta}$ eine $k$-te Wurzel der komplexen Zahl $-\frac{|b_k|}{b_k} \in S^1$, d.h. $\theta$ löst die Gleichung

$$e^{ik\theta} = -\frac{|b_k|}{b_k}.$$

Weiter im Beweis.

Wir wählen jetzt $\rho > 0$ so klein, dass $\rho^k|b_k| < 1$. Dann gilt für $0 < r < \rho$

$$|Q(re^{i\theta})| \leq |1 + b_k r^k \underbrace{e^{ik\theta}}_{=-\frac{|b_k|}{b_k}}| + |b_{k+1}r^{k+1}e^{i(k+1)\theta}| + \dots + |b_n r^n e^{in\theta}|$$

$$= 1 - r^k(|b_k| - r|b_{k+1}| - \dots - r^{n-k}|b_n|)$$

Für hinreichend kleines $r$ ist $(|b_k| - r|b_{k+1}| - \dots - r^{n-k}|b_n|) > 0$, und somit

$$|Q(re^{i\theta})| < 1.$$

So erhalten wir einen Widerspruch zu $|Q(z)| \geq 1$ für alle $z \in \mathbb{C}$, und somit zur Annahme $\mu \neq 0$. Deshalb muss $P(z_0) = \mu = 0$ gelten. □





Ergänzung: Für die im Beweis benutzte Übungsaufgabe nutzt man zunächst:

**Lemma ( Bolzano-Weierstraß für Folgen in $\mathbb{C}$)**

Jede beschränkte Folge $(z_n)_n$ komplexer Zahlen enthält eine in $\mathbb{C}$ konvergente Teilfolge $(z_{n_k})_k$.

**Beweis.**

Die Folge $(\mathrm{Re}(z_n))_n$ ist eine beschränkte Folge reeller Zahlen. Nach dem bereits gezeigten Satz von Bolzano-Weierstraß für Folgen in $\mathbb{R}$ hat sie eine konvergente Teilfolge $(\mathrm{Re}(z_{n_k}))_k$ mit Grenzwert $c \in \mathbb{R}$
Im zweiten Schritt betrachten wir nun die beschränkte reelle Folge $(\mathrm{Im}(z_{n_k}))_k$. Nach Bolzano-Weierstraß existiert wiederum eine Teilfolge $(\mathrm{Im}(z_{n_{k_m}}))_m$ die gegen ein $d \in \mathbb{R}$ konvergiert.
Nun haben wir für die Teilfolge $(z_{n_{k_m}})_m = \left(\mathrm{Re}(z_{n_{k_m}}) + i\,\mathrm{Im}(z_{n_{k_m}})\right)_m$ von $(z_n)_n$ die Konvergenz gegen $c + di \in \mathbb{C}$.

□

# Kapitel 8

Differentialrechnung





## Differentialrechnung

### Definition (Differenzenquotient)

*Sei $D \subseteq \mathbb{R}$, $f\colon D \to \mathbb{R}$ eine Funktion und $x \in D$.*
*Die auf $D \setminus \{x\}$ definierte Funktion*

$$\xi \mapsto \frac{f(\xi) - f(x)}{\xi - x}$$

*heißt Differenzenquotient von $f$ an der Stelle $x$.*

### Definition (Häufungspunkt)

*Ein Punkt $x \in \mathbb{R}$ heißt Häufungspunkt von $D$, wenn es eine Folge $x_n \in D \setminus \{x\}$ gibt, die gegen $x$ konvergiert.*

- Jeder Punkt in einem (offenen oder abgeschlossenen) Intervall ist ein Häufungspunkt des Intervalls.
- Erinnerung: Wir meinen "$\lim\limits_{\xi \to x,\ \xi \neq x}$", wenn wir "$\lim\limits_{\xi \to x}$" schreiben.

## Definition (Ableitung)

*Sei $D \subseteq \mathbb{R}$, $f: D \to \mathbb{R}$ eine Funktion und $x \in D$ ein Häufungspunkt von $D$.*

- *Wir sagen, dass die Funktion $f$ im Punkt $x$ differenzierbar ist, wenn der Grenzwert*

$$f'(x) := \lim_{\xi \to x} \frac{f(\xi) - f(x)}{\xi - x}$$

  *existiert.*
- *Der Grenzwert $f'(x)$ heißt Ableitung der Funktion $f$ an der Stelle $x$.*
- *Die Funktion $f: D \to \mathbb{R}$ heißt differenzierbar, wenn sie in allen Punkten $x \in D$ differenzierbar ist.*
- *Ist $f$ differenzierbar, so heißt die Funktion $f': D \to \mathbb{R}$ Ableitung von $f$.*





## Differentialquotient und zeitliche Ableitung

- Nach Leibniz (1646-1716) schreibt man auch

$$\frac{df}{dx}$$

  statt $f'$.
- Nach Newton (1643-1727) schreibt man auch

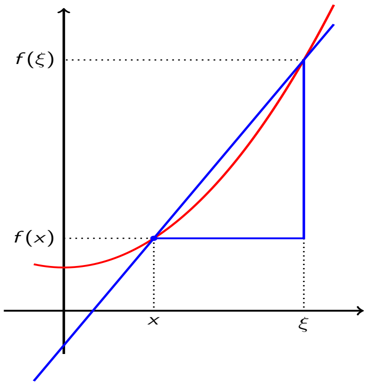
$$\dot{f}$$

  statt $f'$, und nennt die Variable, von der $f$ abhängt, “Zeit”. Die Zeitvariable wird oft mit $t$ bezeichnet.

Die heute übliche Definition der Ableitung wurde nach Vorarbeiten von Cauchy schließlich von Weierstraß Ende des 19. Jahrhunderts formuliert.

# Geometrische Interpretation: Sekante

- Der Differenzenquotient
$$\frac{f(\xi) - f(x)}{\xi - x}$$
ist genau die Steigung der Geraden durch die Punkte $(x, f(x))$ und $(\xi, f(\xi))$.

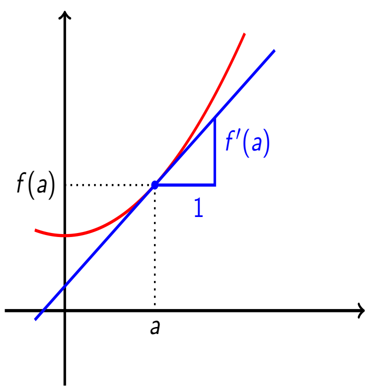



- Diese Gerade nennt man die Sekante (="die Schneidende").





# Geometrische Interpretation: Tangente




- Wenn $f$ in $x$ differenzierbar ist, dann strebt die Sekante für $\xi \to x$ gegen eine Grenzgerade, die sogenannte Tangente [s. Animation]:
- Die Tangente (="die Berührende") an den Graphen von $f$ im Punkt $p = (x, f(x))$ ist die Gerade durch $p$ mit Steigung $f'(x)$.

## Kinematische Interpretation

- Sei $I \subseteq \mathbb{R}$ ein Intervall und $f\colon I \to \mathbb{R}$ eine Funktion.
- Wir können $t \mapsto f(t)$ als die Bewegung eines Punktes im eindimensionalen Raum $\mathbb{R}$ auffassen.
- Der Differenzenquotient $\frac{f(t)-f(t_0)}{t-t_0}$ ist dann die Durchschnittsgeschwindigkeit zwischen $t_0$ und $t$, und der Grenzwert $f'(t_0)$ ist die Geschwindigkeit zum Zeitpunkt $t_0$.
- Die Bewegung eines Punktes $x(t) := (x_1(t), x_2(t), x_3(t))$ im dreidimensionalen Raum $\mathbb{R}^3$ wird entsprechend durch drei Funktionen $x_i\colon \mathbb{R} \to \mathbb{R}$, $i = 1, 2, 3$, beschrieben.
- Die Geschwindigkeit $v(t)$ zum Zeitpunkt $t$ hat dementsprechend drei Komponenten:
$$v(t) := \dot{x}(t) := (\dot{x}_1(t), \dot{x}_2(t), \dot{x}_3(t)).$$
- Nochmaliges Ableiten liefert die Beschleunigung
$$a(t) := \dot{v}(t) = \ddot{x}(t) \ .$$





## Beispiele

(i) Für jede konstante Funktion $f\colon \mathbb{R} \to \mathbb{R}$, $f(x) = c$, gilt
$$f'(x) = \lim_{\xi \to x} \frac{c - c}{\xi - x} = \lim_{\xi \to x} 0 = 0.$$

(ii) Für jede lineare Funktion $f\colon \mathbb{R} \to \mathbb{R}$, $f(x) = ax$, gilt
$$f'(x) = \lim_{\xi \to x} \frac{a\xi - ax}{\xi - x} = \lim_{\xi \to x} \frac{a(\xi - x)}{\xi - x} = \lim_{\xi \to x} a = a.$$

Für jede affine Funktion $f(x) := ax + b$ mit $a, b \in \mathbb{R}$ gilt
$$f'(x) = \lim_{\xi \to x} \frac{a\xi + b - (ax + b)}{\xi - x} = \lim_{\xi \to x} a = a.$$

(iii) Für $f\colon \mathbb{R} \to \mathbb{R}$, $f(x) = x^2$, gilt mit der binomischen Formel:
$$f'(x) = \lim_{\xi \to x} \frac{\xi^2 - x^2}{\xi - x} = \lim_{\xi \to x} \frac{(\xi - x)(\xi + x)}{\xi - x} = \lim_{\xi \to x} (\xi + x) = 2x.$$

# Weitere Beispiele

(iv) Für $f\colon \mathbb{R}^* := \mathbb{R}\setminus\{0\} \to \mathbb{R}$, $f(x) = \frac{1}{x}$, gilt

$$f'(x) = \lim_{\xi\to x} \frac{\frac{1}{\xi} - \frac{1}{x}}{\xi - x} = \lim_{\xi\to x} \frac{\frac{x-\xi}{x\xi}}{\xi - x} = -\lim_{\xi\to x} \frac{1}{\xi x} = -\frac{1}{x^2}.$$

(v) Es gilt $\exp' = \exp$: Für $h := \xi - x$ gilt

$$\frac{\exp\xi - \exp x}{\xi - x} = \frac{\exp(x+h) - \exp x}{h} = \exp x \frac{\exp h - 1}{h}$$

.

Andererseits haben wir $\lim\limits_{h\to 0} \frac{\exp(h)-1}{h} = 1$, denn

$$\left|\frac{\exp h - 1}{h} - 1\right| = \left|\frac{\frac{h^2}{2!} + \frac{h^3}{3!} + \dots}{h}\right| \le \frac{|h|}{2!} + \frac{|h|^2}{3!} + \dots$$

$$\le |h| + \frac{|h|^2}{2!} + \dots = \exp(|h|) - 1 \xrightarrow[h\to 0]{} 0.$$

wegen der Stetigkeit der Exponentialfunktion.





# Noch mehr Beispiele

(vi) Es gilt $\sin' = \cos$, denn aus dem Additionstheorem für sin folgt:

$$\begin{aligned} \frac{\sin(x+h) - \sin x}{h} &= \frac{\sin x \cos h + \cos x \sin h - \sin x}{h} \\ &= \sin x \frac{\cos h - 1}{h} + \cos x \frac{\sin h}{h}. \end{aligned}$$

Andererseits haben wir $\lim\limits_{h\to 0} \frac{\cos h - 1}{h} = 0$, sowie $\lim\limits_{h\to 0} \frac{\sin h}{h} = 1$.

Das folgt aus:

$$\begin{aligned} \left|\frac{\cos h - 1}{h}\right| &\le \frac{|h|}{2!} + \frac{|h|^3}{4!} + \dots \le |h| + \frac{|h|^2}{2!} + \dots \le e^{|h|} - 1 \\ \left|\frac{\sin h}{h} - 1\right| &\le \frac{|h|^2}{3!} + \frac{|h|^4}{5!} + \dots \le \frac{|h|^2}{2!} + \frac{|h|^4}{4!} + \dots \le e^{|h|} - 1. \end{aligned}$$

(vii) $\cos' = -\sin$ (ÜA).

## Beispiel (stetig aber nicht differenzierbar)

Die Funktion $f(x) := |x|$, $x \in \mathbb{R}$, ist im Nullpunkt nicht differenzierbar, aber sehr wohl stetig.

**Beweis.**
Wir betrachten die beiden Nullfolgen $x_n^+ := \frac{1}{n}$ und $x_n^- := -\frac{1}{n}$.
Einerseits gilt für die erste Folge

$$\lim_{n\to\infty} \frac{f(x_n^+) - f(0)}{x_n^+ - 0} = \lim_{n\to\infty} \frac{\frac{1}{n}}{\frac{1}{n}} = \lim_{n\to\infty} 1 = 1,$$

und andererseits für die zweite Folge

$$\lim_{n\to\infty} \frac{f(x_n^-) - f(0)}{x_n^- - 0} = \lim_{n\to\infty} \frac{\frac{1}{n}}{-\frac{1}{n}} = -1.$$

Also existiert der Grenzwert der Differenzenquotienten von $f$ in 0 nicht, und die Funktion ist demnach in 0 nicht differenzierbar. $\square$





Umgekehrt haben wir aber folgende Aussage:

## Satz (Differenzierbare Funktionen sind stetig)

*Sei $f : D \to \mathbb{R}$ in $a \in D \subseteq \mathbb{R}$ differenzierbar. Dann ist $f$ im Punkt $a$ stetig.*

**Beweis.** Es gilt

$$|f(x) - f(a)| = \left| \frac{f(x) - f(a)}{x - a} \right| |x - a|$$

und somit

$$\lim_{x\to a} |f(x) - f(a)| = |f'(a)| \cdot \lim_{x\to a} |x - a| = 0$$

D.h. $\lim_{x\to a} f(x) = f(a)$ und damit ist $f$ stetig. $\square$

# Affine Approximation

## Satz (Affine Approximation)

*Sei $f: D \to \mathbb{R}$ eine in $a \in D \subseteq \mathbb{R}$ differenzierbare Funktion und $h: D \to \mathbb{R}$ die affine Funktion $h(x) := f(a) + f'(a)(x-a)$. Dann gilt*

$$\lim_{x \to a} \frac{f(x) - h(x)}{x - a} = 0.$$

**Beweis.** $\frac{f(x)-h(x)}{x-a} = \frac{f(x)-f(a)}{x-a} - f'(a) \underset{x \to a}{\longrightarrow} 0.$ □

## Beispiel:

Die Wurzelfunktionen $f: x \mapsto \sqrt[n]{x}$, $n \in \mathbb{N}$ sind für $n \geq 2$ in $a = 0$ nicht differenzierbar:

- Wir nehmen an, $f$ ist differenzierbar in 0 mit $f'(0)$ als Ableitung.
- Die zugehörige affine Funktion wäre $h(x) = f'(0)x$ und somit $\frac{f(x)-h(x)}{x-a} = \frac{\sqrt[n]{x}-f'(0)x}{x} = \frac{1}{\sqrt[n]{x^{n-1}}} - f'(0).$
- Dieser Ausdruck ist für $x \to 0$ unbeschränkt.





## Definition (Affine Approximation/lineare Approximation)

*Sei $f: D \to \mathbb{R}$ in $a \in D$ differenzierbar, $D \subseteq \mathbb{R}$.*

- *Die affine Funktion $h(x) = f(a) + f'(a)(x-a)$ heißt affine Approximation von $f$ in $a$.*
- *Manchmal spricht man auch einfach von linearer Approximation.*
- *Die Funktion $R(a,x) := f(x) - h(x) = f(x) - f'(a)(x-a) - f(a)$ heißt Restglied.*

Für $x \to a$ geht das Restglied schneller gegen 0 als $x - a$. Das heißt, dass sogar auch noch

$$\frac{R(a,x)}{x-a}$$

gegen Null geht.

## Satz

*Sei $f\colon D \to \mathbb{R}$ eine stetige Funktion, $a \in D$ ein Häufungspunkt von $D$. Gibt es eine* affine *Funktion $h(x) = c(x-a) + d$ mit $c, d \in \mathbb{R}$, so dass*

$$(*) \qquad \lim_{x\to a} \frac{f(x) - h(x)}{x - a} = 0$$

*gilt, so ist $f$ in $a$ differenzierbar, und es gilt $c = f'(a)$ und $d = f(a)$.*

**Beweis.** Nach Voraussetzung gilt

$$0 = \lim_{x\to a} \frac{f(x) - h(x)}{x - a} = \lim_{x\to a} \left( \frac{f(x) - d}{x - a} - c \right) = \lim_{x\to a} \left( \frac{f(x) - d}{x - a} \right) - c.$$

Damit der Grenzwert endlich ist, muss nun $d = \lim_{x\to a} f(x)$ gelten.Dann folgt aber aus der Stetigkeit von $f$ auch $d = f(a)$, woraus wir wiederum schließen, dass $f$ in $a$ differenzierbar ist mit Ableitung $f'(a) = c$. $\square$





## Satz (Ableitungsregeln)

*Die Funktionen $f, g\colon D \to \mathbb{R}$ seien in $a \in D \subseteq \mathbb{R}$ differenzierbar. Dann gilt:*

(i) *Sind $\lambda$ und $\mu$ reelle Zahlen, dann ist die Funktion $\lambda \cdot f + \mu \cdot g\colon D \to \mathbb{R}$ im Punkt $a$ differenzierbar und*

$$(\lambda \cdot f + \mu \cdot g)'(a) = \lambda \cdot f'(a) + \mu \cdot g'(a). \qquad \textit{(Linearität)}$$

(ii) *Die Funktion $f \cdot g\colon D \to \mathbb{R}$ ist im Punkt $a$ differenzierbar und*

$$(f \cdot g)'(a) = f'(a)g(a) + f(a)g'(a). \qquad \textit{(Leibnizregel/Produktregel)}$$

(iii) *Wenn $g(a) \neq 0$, dann ist die Funktion $\frac{f}{g}\colon D \setminus g^{-1}(0) \to \mathbb{R}$ im Punkt $a$ differenzierbar und*

$$\left(\frac{f}{g}\right)'(a) = \frac{f'(a)g(a) - f(a)g'(a)}{g(a)^2}. \qquad \textit{(Quotientenregel)}$$

**Beweis.**
(i)

$$\frac{(\lambda f+\mu g)(a+h)-(\lambda f+\mu g)(a)}{h}$$
$$=\lambda\frac{f(a+h)-f(a)}{h}+\mu\frac{g(a+h)-g(a)}{h}\xrightarrow[h\to 0]{}\lambda\cdot f'(a)+\mu\cdot g'(a)$$

(ii)

$$\frac{(fg)(a+h)-(fg)(a)}{h}$$
$$=\frac{f(a+h)g(a+h)-f(a)g(a+h)}{h}+\frac{f(a)g(a+h)-f(a)g(a)}{h}$$
$$=\frac{f(a+h)-f(a)}{h}g(a+h)+f(a)\frac{g(a+h)-g(a)}{h}$$
$$\xrightarrow[h\to 0]{}f'(a)g(a)+f(a)g'(a)$$

da $g$ als differenzierbare Funktion stetig ist.





**Weiter im Beweis.**
(iii) Wir betrachten zunächst den Spezialfall $f=1$:

$$\frac{\frac{1}{g(a+h)}-\frac{1}{g(a)}}{h}=\frac{g(a)-g(a+h)}{hg(a+h)g(a)}$$
$$=-\underbrace{\frac{g(a+h)-g(a)}{h}}_{\xrightarrow{h\to 0}g'(a)}\cdot\underbrace{\frac{1}{g(a+h)g(a)}}_{\xrightarrow{h\to 0}1/g(a)^2}\xrightarrow{h\to 0}-\frac{g'(a)}{g(a)^2}\ ,$$

d.h. $(\frac{1}{g})'(a)=-\frac{g'(a)}{g(a)^2}$. Mit der Produktregel (ii) erhalten wir dann für beliebiges $f$:

$$\left(\frac{f}{g}\right)'(a)=\left(f\cdot\frac{1}{g}\right)'(a)=\frac{f'(a)}{g(a)}+f(a)\left(\frac{1}{g}\right)'(a)$$
$$=\frac{f'(a)g(a)}{g(a)^2}-\frac{f(a)g'(a)}{g(a)^2}.$$

$\square$

## Beispiele

(i) Eine einfache Induktion mit Hilfe der Produktregel (ii) ergibt

$$(x^n)' = nx^{n-1} \quad \text{für alle} \quad n \in \mathbb{N}.$$

Insbesondere sind polynomiale Funktionen differenzierbar.

(ii) Die Regel $(1/g)' = -g'/g^2$ mit $g(x) = x^n$ liefert dann

$$(x^{-n})' = -\frac{nx^{n-1}}{x^{2n}} = -nx^{-n-1} \quad (x \neq 0) \quad \text{für alle} \quad n \in \mathbb{N}, n \geq 1.$$

(iii) Die Quotientenregel liefert

$$\tan' = \frac{\sin' \cos - \sin \cos'}{\cos^2} = \frac{\cos^2 + \sin^2}{\cos^2} = \frac{1}{\cos^2} = 1 + \tan^2.$$





### Satz (Ableitung der Umkehrfunktion)

*Sei $f : [a, b] \to \mathbb{R}$ eine stetige streng monotone Funktion und $g = f^{-1} : [c, d] \to \mathbb{R}$ sei die zugehörige Umkehrfunktion. Wenn $f$ in $x \in [a, b]$ differenzierbar ist und $f'(x) \neq 0$ gilt,dann ist $g$ in $y := f(x)$ differenzierbar und es gilt*

$$g'(y) = \frac{1}{f'(x)} = \frac{1}{f'(g(y))}.$$

**Beweis.** Sei $(y_n)$ eine Folge in $[c, d] \setminus \{y\}$, die gegen $y$ konvergiert. Da $g$ stetig ist, konvergiert die Folge $(x_n := g(y_n))$ aus $[a, b] \setminus \{x\}$ gegen $x := g(y)$.
Somit gilt

$$\lim_{n\to\infty} \frac{g(y_n) - g(y)}{y_n - y} = \lim_{n\to\infty} \frac{x_n - x}{f(x_n) - f(x)} = \frac{1}{f'(x)},$$

d.h. $g'(y) = 1/f'(x)$. $\square$

## Bemerkung

Wenn im obigen Satz $f$ überall differenzierbar ist und $f'(x) \neq 0$ für alle $x \in [a, b]$, so erhalten wir für die Ableitung der Umkehrfunktion $g = f^{-1} \colon [c, d] \to \mathbb{R}$ die Regel

$$g' = \frac{1}{f' \circ g}.$$





## Beispiele

(i) $\ln'(x) = \dfrac{1}{\exp'(\ln x)} = \dfrac{1}{\exp(\ln x)} = \dfrac{1}{x}$ für $x > 0$.

(ii) $\arcsin = (\sin|_{[-\frac{\pi}{2},\frac{\pi}{2}]})^{-1} \colon [-1, 1] \to \mathbb{R}$ hat für $x \in (-1, 1)$ die Ableitung

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

denn $\cos = \sqrt{1 - \sin^2}$ auf $[-\frac{\pi}{2}, \frac{\pi}{2}]$.

(iii) Ebenso hat $\arccos = (\cos|_{[0,\pi]})^{-1} \colon [-1, 1] \to \mathbb{R}$ für $x \in (-1, 1)$ die Ableitung

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

(iv) $\arctan'(x) = \dfrac{1}{\tan'(\arctan x)} = \dfrac{1}{1 + (\tan(\arctan x))^2} = \dfrac{1}{1 + x^2}$, denn $\tan' = 1 + \tan^2$.

## Satz (Kettenregel)

*Sei $f\colon D \to \mathbb{R}$ an der Stelle $x \in D \subseteq \mathbb{R}$ differenzierbar, $f(D) \subseteq E \subseteq \mathbb{R}$ und sei $g\colon E \to \mathbb{R}$ an der Stelle $y := f(x)$ differenzierbar.*
*Dann ist die Verkettung $g \circ f\colon D \to \mathbb{R}$ an der Stelle $x$ differenzierbar und es gilt*

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





**Beweis.** Sei $x$ fixiert und $y = f(x)$. Wir betrachten die folgende Funktion

$$\begin{aligned} h\colon E &\to \mathbb{R} \\ \eta &\mapsto \begin{cases} \frac{g(\eta)-g(y)}{\eta-y}, & \text{falls} \quad \eta \in E \setminus \{y\} \\ g'(y), & \text{falls} \quad \eta = y. \end{cases} \end{aligned}$$

D.h. $h$ ist die stetige Fortsetzung des Differenzenquotienten von $g$ in $y = f(x)$. Für alle $\eta \in E$ gilt dann

$$g(\eta) - g(y) = h(\eta) \cdot (\eta - y). \tag{5}$$

Sei nun $\xi \in D \setminus \{x\}$ beliebig und $\eta = f(\xi)$. Aus (5) folgt dann:

$$\begin{aligned} \frac{(g \circ f)(\xi) - (g \circ f)(x)}{\xi - x} &= h(f(\xi)) \cdot \frac{f(\xi) - f(x)}{\xi - x} \\ &\xrightarrow[\xi \to x]{} h(f(x)) \cdot f'(x) \;=\; g'(f(x)) \cdot f'(x) \end{aligned}$$

da $h$ stetig ist. $\square$

## Beispiele

(i) Sei $f\colon \mathbb{R} \to \mathbb{R}$ differenzierbar, $a, b \in \mathbb{R}$ und $g\colon \mathbb{R} \to \mathbb{R}$ definiert durch $g(x) := f(ax+b)$.
Dann ist $g$ differenzierbar und

$$g'(x) = a \cdot f'(ax+b).$$

(ii) Sei $f\colon \mathbb{R}_+ \to \mathbb{R}$, $f(x) = x^\alpha$, wobei $\alpha \in \mathbb{R}$.
Dann ist

$$f'(x) = \alpha x^{\alpha-1},$$

denn aus $f(x) = \exp(\alpha \ln x)$ folgt mit der Kettenregel

$$f'(x) = \exp'(\alpha \ln x)(\alpha \ln x)' = \underbrace{\exp(\alpha \ln x)}_{x^\alpha} \frac{\alpha}{x} = \alpha x^{\alpha-1}.$$





### Definition (Lokale Extrema)

*Sei $f\colon D \to \mathbb{R}$ eine Funktion und $z \in D \subseteq \mathbb{R}$.*

- *Man sagt, dass $f$ in $z$ ein lokales Maximum (bzw. ein lokales Minimum) annimmt, falls es ein $\varepsilon > 0$ gibt, derart dass*

  $$f(z) \geq f(\zeta) \quad (\text{bzw.} \quad f(z) \leq f(\zeta))$$

  *für alle $\zeta \in D$ mit $|\zeta - z| < \varepsilon$.*
- *Im Gegensatz dazu bezeichnet man $\max f$ und $\min f$ als globales Minimum bzw. Maximum.*
- *Lokale (bzw. globale) Minima und Maxima bezeichnet man auch als lokale (bzw. globale) Extrema (falls sie existieren).*
- *Man spricht von einem isolierten lokalen Extremum, falls zusätzlich $f(\zeta) \neq f(z)$ gilt, für alle $\zeta \in D \setminus \{z\}$ mit $|\zeta - z| < \varepsilon$.*

### Satz

*Die Funktion $f\colon (a,b) \to \mathbb{R}$ sei an der Stelle $x \in (a,b)$ differenzierbar und nehme dort ein lokales Extremum an. Dann gilt $f'(x) = 0$.*

**Beweis.**

Wir betrachten den Fall, dass in $x$ ein lokales Maximum vorliegt.

Dann gibt es $\varepsilon > 0$, so dass $f(x) - f(\xi) \geq 0$ für alle $\xi$ mit $|\xi - x| < \varepsilon$.

Schreibt man "$\lim\limits_{\xi \nearrow x}$" für "$\lim\limits_{\xi \to x, \xi < x}$" und "$\lim\limits_{\xi \searrow x}$" für "$\lim\limits_{\xi \to x, \xi > x}$", dann ist

$$f'(x) = \lim_{\xi \nearrow x} \frac{f(\xi) - f(x)}{\xi - x} \geq 0$$

und ebenso

$$f'(x) = \lim_{\xi \searrow x} \frac{f(\xi) - f(x)}{\xi - x} \leq 0 \, .$$

Daraus folgt $f'(x) = 0$. Lokale Minima behandelt man analog. $\square$



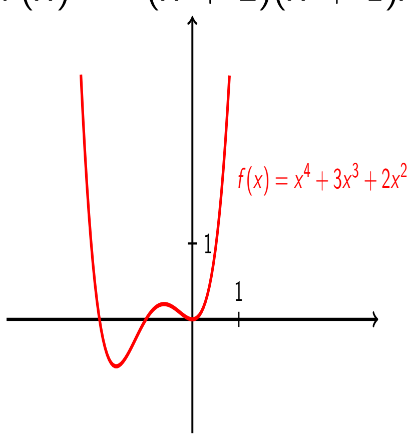


## Beispiele

(i) Achtung: Die Bedingung $f'(x) = 0$ ist nur notwendig, aber nicht hinreichend für die Existenz eines lokalen Extremums.
Die Funktion $f\colon \mathbb{R} \to \mathbb{R}$, $f(x) = x^3$ erfüllt $f'(0) = 0$, hat aber an der Stelle 0 kein lokales Extremum.

(ii) Die Funktion $f\colon \mathbb{R} \to \mathbb{R}$, $f(x) = (x+2)(x+1)x^2$, ist nach oben unbeschränkt ($\sup f = +\infty$) und hat daher kein (globales) Maximum.
Sie nimmt ihr (globales) Minimum $\min f$ an einer Stelle $a \in (-2,-1)$ an.
Sie hat zwei weitere lokale Extrema: ein lokales Maximum an einer Stelle $b \in (-1,0)$ und ein lokales Minimum bei 0.
ÜA: Berechnen Sie $a$, $b$ und $\min f = f(a)$.

$f(x) = (x+2)(x+1)x^2$




(iii) Die Funktion $f\colon [-1,2] \to \mathbb{R}$, $f(x) = x^2$, nimmt an der Stelle 0 ihr Minimum $\min f = 0$, an der Stelle 2 ihr Maximum $\max f = 4$ und an der Stelle $-1$ ein lokales (Rand-)Maximum an.

## Satz (Satz von Rolle)

*Sei $f\colon [a,b] \to \mathbb{R}$ eine stetige Funktion mit $f(a) = f(b)$.*
*Falls $f$ auf $(a,b)$ differenzierbar ist, so existiert ein $\xi \in (a,b)$ mit $f'(\xi) = 0$.*

**Beweis.** Falls $f$ konstant ist, so gilt $f' \equiv 0$ und der Satz ist erfüllt.

Falls $f$ nicht konstant ist, so existiert $x \in (a,b)$ mit $f(x) > f(a) = f(b)$ oder $f(x) < f(a) = f(b)$.

Im ersten Fall existiert wegen der Stetigkeit von $f$ ein $\xi \in (a,b)$, also im offenen Intervall, mit $f(\xi) = \max f$, im zweiten Fall $\xi \in (a,b)$ mit $f(\xi) = \min f$.

In beiden Fällen ist $f'(\xi) = 0$. □





## Satz (Mittelwertsatz)

*Sei $f\colon [a,b] \to \mathbb{R}$ eine stetige, auf $(a,b)$ differenzierbare Funktion.*
*Dann existiert $\xi \in (a,b)$ mit*

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

[s. Animation]
**Beweis.**
Wir betrachten die Hilfsfunktion $g\colon [a,b] \to \mathbb{R}$,

$$g(x) := f(x) - \frac{f(b) - f(a)}{b - a}(x - a).$$

$g$ erfüllt die Voraussetzungen des Satzes von Rolle, insbesondere $g(a) = g(b) = f(a)$.

Somit existiert $\xi \in (a,b)$ mit $0 = g'(\xi) = f'(\xi) - \frac{f(b)-f(a)}{b-a}$. □

## Folgerung (Schrankensatz)

*Unter den Voraussetzungen des Mittelwertsatzes gelte zusätzlich für die Ableitungsfunktion*

$$m \leq f'(\xi) \leq M \quad \text{für alle} \quad \xi \in (a, b). \tag{6}$$

*Dann gilt*

$$m(y-x) \leq f(y) - f(x) \leq M(y-x) \tag{7}$$

*für alle* $x, y \in [a, b]$ *mit* $x \leq y$.

**Beweis.**
Für $x = y$ ist nichts zu zeigen. Für $x < y$ gibt es nach dem Mittelwertsatz ein $\xi_0 \in (x, y)$ mit $f'(\xi_0) = \frac{f(y)-f(x)}{y-x}$.
Multiplikation der Ungleichung (6) für $\xi = \xi_0$ mit $y - x > 0$ ergibt dann (7). □





## Folgerung

*Sei* $f : [a, b] \to \mathbb{R}$ *stetig, auf* $(a, b)$ *differenzierbar und* $f'(x) = 0$ *für alle* $x \in (a, b)$. *Dann ist* $f$ *konstant.*

**Beweis.**

- Die Funktion erfüllt die Voraussetzungen des Schrankensatzes mit $m = M = 0$.
- Also $f(y) = f(x)$ für alle $x, y \in [a, b]$ mit $x \leq y$, d.h. $f = const$. □

## Bemerkung

- Dieser Satz ist sehr hilfreich beim Studium der Eindeutigkeit von Lösungen von Differentialgleichungen bei gegebenen Anfangsbedingungen.
- Er besagt, dass die Differentialgleichung $f' = 0$ genau eine Lösung mit der Anfangsbedingung $f(x_0) = c$ hat, nämlich die konstante Funktion $f \equiv c$. (Hierbei ist $x_0 \in [a, b]$ und $c \in \mathbb{R}$.)

### Satz (Charakterisierung von exp durch eine Differentialgleichung)

*Sei $c \in \mathbb{R}$ und $f : \mathbb{R} \to \mathbb{R}$ eine Lösung der Differentialgleichung*

$$f' = c \cdot f. \tag{8}$$

*Dann gilt $f(x) = f(0) \cdot e^{cx}$ (für alle $x \in \mathbb{R}$).*

**Beweis.** Wir betrachten die differenzierbare Funktion $g : \mathbb{R} \to \mathbb{R}$, $g(x) := f(x)e^{-cx}$.

Ableiten liefert mit der Produktregel:

$$g'(x) \;=\; f'(x)e^{-cx} + f(x)\left(e^{-cx}\right)' \;=\; f'(x)e^{-cx} - c \cdot f(x)e^{-cx} \overset{(8)}{=} 0.$$

Somit $g = const = g(0) = f(0)$, d.h. $f(x) = g(x)e^{cx} = f(0)e^{cx}$. $\square$





## Grenzwertbestimmung nach L'Hospital/Bernoulli

### Satz (Verallgemeinerter Mittelwertsatz)

*Seien $f, g : [a, b] \to \mathbb{R}$ stetige Funktionen, die differenzierbar auf dem offenen Intervall $(a, b)$ sind. Dann existiert ein $\xi \in (a, b)$ mit*

$$(f(b) - f(a))g'(\xi) = (g(b) - g(a))f'(\xi).$$

**Beweis.** Betrachte $h : [a, b] \to \mathbb{R}$, definiert als

$$h(y) = [f(b) - f(a)]g(y) - [g(b) - g(a)]f(y),$$

so dass $h(a) = f(b)g(a) - f(a)g(b) = h(b)$ gilt.

Also können wir den Satz von Rolle anwenden, und erhalten $\xi \in (a, b)$ mit $h'(\xi) = 0$. Dies ist aber gleichbedeutend mit der Behauptung. $\square$

### Bemerkung

Im Spezialfall $g(x) = x$ erhalten wir den Mittelwertsatz in der ursprünglichen Form.

## Folgerung (Regel von L'Hospital)

*Seien* $f, g\colon (a,b) \to \mathbb{R}$ *differenzierbar mit* $\lim_{x\searrow a} f(x) = \lim_{x\searrow a} g(x) = 0$.

*Falls* $g'(x) \neq 0$ *für* $x$ *nahe* $a$ *und* $\lim_{x\searrow a} \frac{f'(x)}{g'(x)}$ *existiert, so gilt*

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

*Dieselbe Aussage gilt für* $\lim_{x\nearrow b}$ *und damit für beidseitige Grenzwerte.*

**Beweis.** Nach Voraussetzung gibt es ein $\delta > 0$, so dass $g'(x) \neq 0$ für alle $x \in (a, a+\delta)$. Wir definieren

$$\tilde{f}\ :\ [a, a+\delta) \to \mathbb{R}, \quad \tilde{f}(\xi) = \begin{cases} 0, & \text{falls } \xi = a \\ f(\xi), & \text{falls } \xi \in (a, a+\delta), \end{cases}$$

und setzen $g$ ebenso zu $\tilde{g}$ fort.





Für jedes $x \in (a, a+\delta)$ erfüllen die Funktionen $\tilde{f}$ und $\tilde{g}$ dann auf dem Intervall $[a, x]$ die Voraussetzungen des verallgemeinerten Mittelwertsatzes. Wir finden also zu jeden $x \in (a, a+\delta)$ einen Zwischenwert $t = t(x) \in (a, x)$ mit

$$\frac{f(x)}{g(x)} = \frac{\tilde{f}(x) - \tilde{f}(a)}{\tilde{g}(x) - \tilde{g}(a)} = \frac{f'(t)}{g'(t)}.$$

Ist nun $(x_n)$ eine beliebige Folge mit $x_n \searrow a$, so gilt auch für die Folge der Zwischenwerte $t_n \searrow a$ und damit

$$\lim_{n\to\infty} \frac{f(x_n)}{g(x_n)} = \lim_{n\to\infty} \frac{f'(t_n)}{g'(t_n)} = \lim_{x\searrow a} \frac{f'(x)}{g'(x)}.$$

□

## Bemerkung

Analoge Aussagen gibt es auch im Fall $\lim_{x\searrow a} f(x) = \lim_{x\searrow a} g(x) \in \{-\infty, \infty\}$.

### Beispiele

(i) $\lim\limits_{x\to 0} \dfrac{\sin x}{x} = \lim\limits_{x\to 0} \dfrac{\cos x}{1} = 1.$ (ii) $\lim\limits_{x\to 0} \dfrac{1-\cos x}{x^2} = \lim\limits_{x\to 0} \dfrac{\sin x}{2x} = \dfrac{1}{2}.$

### Bemerkung

Analog erhält man, dass $\lim\limits_{x\to\infty} \dfrac{f(x)}{g(x)} = \lim\limits_{x\to\infty} \dfrac{f'(x)}{g'(x)}$, falls

- $f$ und $g$ differenzierbar sind für hinreichend große $x$, mit $g' \neq 0$,
- $\lim\limits_{x\to\infty} g(x) = \lim\limits_{x\to\infty} f(x) \in \{-\infty, 0, \infty\}$, und
- $\lim\limits_{x\to\infty} \frac{f'(x)}{g'(x)}$ existiert.

- Beispiel: Für jedes $\alpha > 0$ gilt

$$\lim_{x\to\infty} \frac{\ln x}{x^\alpha} = \lim_{x\to\infty} \frac{1/x}{\alpha x^{\alpha-1}} = \lim_{x\to\infty} \frac{1}{\alpha x^\alpha} = 0.$$

Der Logarithmus wächst also langsamer als jede positive Potenz.





- Achtung: $\lim \frac{f(x)}{g(x)}$ kann existieren, auch wenn $\lim \frac{f'(x)}{g'(x)}$ nicht existiert.
  Bsp.: Für $f(x) = \sin x + 2x$, $g(x) = \cos x + 2x$ ist $\lim\limits_{x\to\infty} \frac{f'(x)}{g'(x)}$ unbestimmt divergent, aber $\lim\limits_{x\to\infty} \frac{f(x)}{g(x)} = \lim\limits_{x\to\infty} (1 + \frac{\sin x - \cos x}{g(x)}) = 1.$
- Die Regel von L'Hospital gilt nicht für komplexwertige Funktionen.

## Satz (Monotonie und Ableitung)

*Sei die Funktion $f : [a, b] \to \mathbb{R}$ stetig und auf $(a, b)$ differenzierbar.*

(i) *Falls $f'(x) \geq 0$ (bzw. $f'(x) > 0$) für alle $x \in (a, b)$, so ist $f$ auf $[a, b]$ monoton wachsend (bzw. streng monoton wachsend).*

(ii) *Falls $f'(x) \leq 0$ (bzw. $f'(x) < 0$) für alle $x \in (a, b)$, so ist $f$ auf $[a, b]$ monoton fallend (bzw. streng monoton fallend).*

(iii) *Umgekehrt gilt: $f$ monoton wachsend (bzw. monoton fallend) impliziert $f'(x) \geq 0$ (bzw. $f'(x) \leq 0$) für alle $x \in (a, b)$.*

## Bemerkung

- Aus dem streng monotonen Wachstum von $f$ folgt nur $f' \geq 0$ und nicht die strikte Ungleichung.
- Beispielsweise ist die Funktion $f(x) = x^3$ streng monoton wachsend, aber $f'(0) = 0$.





## Beweis des Satzes:

(i-ii) Wir betrachten den Fall $f' > 0$ auf $(a, b)$ und zeigen, dass $f$ streng monoton wachsend ist.

Wäre $f$ nicht streng monoton wachsend, so gäbe es $a \leq x < y \leq b$ mit $f(x) \geq f(y)$.

Nach dem Mittelwertsatz gibt es dann $\xi \in (x, y)$ mit

$$f'(\xi) = \frac{f(y) - f(x)}{y - x} \leq 0,$$

im Widerspruch zur Annahme $f' > 0$ auf $(a, b)$.

(iii) Für die Umkehrung betrachten wir ebenfalls den Fall, dass $f$ monoton wachsend ist. Dann gilt für alle $x \in (a, b)$:

$$f'(x) = \lim_{\xi \searrow x} \frac{f(\xi) - f(x)}{\xi - x} \geq 0,$$

weil die Differenzenquotienten nichtnegativ sind.

$\square$

## Folgerung (lokale Extrema)

*Sei $f : (a, b) \to \mathbb{R}$ differenzierbar, $x \in (a, b)$ und $\varepsilon > 0$ derart, dass $x \in (a + \varepsilon, b - \varepsilon)$. Es gelte weiterhin*

$$f'(\xi) \leq 0 \quad (\textit{bzw.} \quad \geq 0) \tag{9}$$

*für alle $\xi \in (x - \varepsilon, x)$ und*

$$f'(\xi) \geq 0 \quad (\textit{bzw.} \quad \leq 0) \tag{10}$$

*für alle $\xi \in (x, x + \varepsilon)$.*

*Dann hat $f$ ein lokales Minimum (bzw. Maximum) an der Stelle $x$.*

*Ersetzt man die Ungleichungen (9) und (10) durch strikte Ungleichungen, $f'(\xi) < 0$ (bzw. $> 0$) usw., so folgt, dass das lokale Extremum isoliert ist.*

**Beweis.** Aus (9) und (10) folgt, dass $f$ auf $[x - \varepsilon, x]$ monoton fallend (bzw. wachsend) und auf $[x, x + \varepsilon]$ monoton wachsend (bzw. fallend) ist. Also hat $f$ an der Stelle $x$ ein lokales Minimum (bzw. Maximum). □

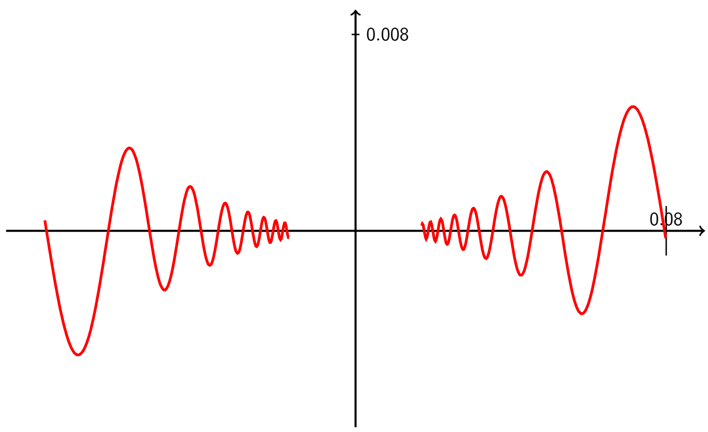




## Beispiele:

(i) Die stetige Funktion $f(x) = \begin{cases} 0 & x = 0 \\ x^2 \sin(1/x) & x \neq 0 \end{cases}$

ist differenzierbar, da $f'(0) = \lim_{x\to 0} x \sin(1/x) = 0$ (vgl. Übungen) und für $x \neq 0$ die Ableitung $f'(x) = 2x \sin(1/x) - \cos(1/x)$ ist.
Weil der Grenzwert $\lim_{x\to 0} f'(x)$ nicht existiert, ist $f'(x)$ nicht stetig!
Man sagt, $f$ ist nicht stetig-differenzierbar.
Die Ableitung kann also in $x = 0$ nicht differenzierbar sein.



(ii) Es gibt stetige aber nirgendwo differenzierbare Funktionen, z.B. die *Weierstraß-Funktion* $f(x) = \sum_{n=1}^{\infty} \frac{2^n \sin(2^n x)}{3^n}$.



Differentialrechnung Höhere Ableitungen

## Definition (Höhere Ableitungen)

*Eine differenzierbare Funktion $f$ heißt zweimal differenzierbar, wenn $f'$ differenzierbar ist.*

*Die Ableitung $f'' := (f')'$ von $f'$ heißt zweite Ableitung von $f$.*

*Allgemein definiert man rekursiv $f^{(0)} := f$ und für alle $k \geq 1$ die $k$-te Ableitung $f^{(k)}$ von $f$ als*

$$f^{(k)} := (f^{(k-1)})',$$

*falls $f^{(k-1)}$ existiert und differenzierbar ist. Man sagt dann, dass $f$ $k$-mal differenzierbar ist.*
*Falls zusätzlich $f^{(k)}$ stetig ist, so heißt $f$ $k$-mal stetig differenzierbar.*

**Bemerkung.** Jede $k$-mal differenzierbare Funktion ist $(k$-1$)$-mal stetig differenzierbar $(k \in \mathbb{N})$.

**Satz (Zweite Ableitung und isolierte lokale Extrema)**

*Sei $f : (a, b) \to \mathbb{R}$ zweimal differenzierbar und $x \in (a, b)$, derart dass $f'(x) = 0$ und $f''(x) > 0$ (bzw. $f''(x) < 0$).*

*Dann hat $f$ ein isoliertes lokales Minimum (bzw. Maximum) an der Stelle $x$.*

**Beweis.**
Wir betrachten den Fall $f''(x) > 0$.

Wegen

$$0 < f''(x) = \lim_{\xi \to x} \frac{f'(\xi) - f'(x)}{\xi - x}$$

gibt es $\varepsilon > 0$, so dass $\frac{f'(\xi)-f'(x)}{\xi - x} > 0$ für alle $\xi$ mit $|x - \xi| < \varepsilon$.
Also gilt dass $f'(\xi) < f'(x) = 0$ für alle $\xi \in (x - \varepsilon, x)$ und $f'(\xi) > f'(x) = 0$ für alle $\xi \in (x, x + \varepsilon)$.

Wir folgern, dass $f$ an der Stelle $x$ ein isoliertes lokales Mininimum hat. □





## Beispiele

(i) Für $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^2$, gilt $f'(0) = 0$ und $f''(0) = 2 > 0$. Also hat $f$ an der Stelle 0 ein isoliertes lokales Minimum. Das ist auch das globale Minimum der Funktion.

(ii) Für $f(x) = x^3$ gilt $f''(0) = 0$; $x$ ist auch kein lokales Extremum, sondern ein Sattelpunkt.

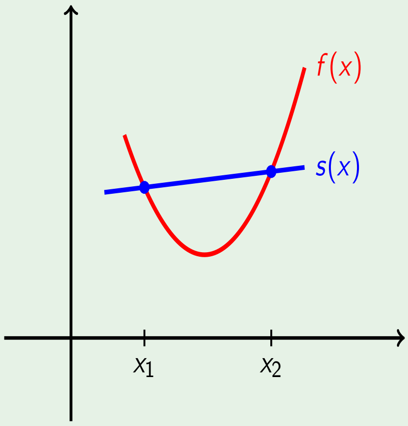
(iii) Die Bedingung $f''(x) \neq 0$ ist hinreichend für die Existenz eines lokalen Extremums, aber nicht notwendig:
Die Funktion $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^4$, hat ebenfalls an der Stelle 0 ihr (isoliertes) globales Minimum. In diesem Fall gilt jedoch $f''(0) = 0$.

## Konvexe Funktionen

- Sei $f : I := [a, b] \to \mathbb{R}$ eine Funktion. Sei $a \leq x_1 < x_2 \leq b$; dann ist

$$s(x) := f(x_1)\frac{x_2 - x}{x_2 - x_1} + f(x_2)\frac{x - x_1}{x_2 - x_1} = \frac{f(x_2) - f(x_1)}{x_2 - x_1}x + \frac{f(x_1)x_2 - f(x_2)x_1}{x_2 - x_1}$$

  die Sekante durch $(x_1, f(x_1))$ und $(x_2, f(x_2))$.
- $f$ heißt konvex, wenn für alle $x_1, x_2 \in I$ mit $x_1 < x_2$ und alle $x \in (x_1, x_2)$ gilt, dass $f(x) \leq s(x)$.








# Konvexität

## Konvexe Funktionen

- Die Konvexitätsbedingung

$$f(x) \leq s(x) := \frac{f(x_2) - f(x_1)}{x_2 - x_1}x + \frac{f(x_1)x_2 - f(x_2)x_1}{x_2 - x_1}$$

  ist äquivalent zu der Ungleichung

$$f((1-t)x_1 + tx_2) \leq (1-t)f(x_1) + tf(x_2) \quad \text{für alle} \quad t \in (0,1).$$

  (Setze $t := \frac{x - x_1}{x_2 - x_1}$. Dann ist $1 - t = \frac{x_2 - x}{x_2 - x_1}$ und $(1-t)x_1 + tx_2 = x$.)

# Konvexität

**Satz**

*Sei $f : I \to \mathbb{R}$ eine zweimal differenzierbare Funktion auf einem Intervall $I \subseteq \mathbb{R}$. $f$ ist genau dann konvex, wenn $f'' \geq 0$.*

**Beweis.**

"$\Longrightarrow$" Wir nehmen an, es gäbe $x_0 \in I$ mit $f''(x_0) < 0$ und führen diese Annahme zum Widerspruch.

Sei $g(x) = f(x_0) + f'(x_0)(x - x_0)$ die affine Approximation von $f$ an der Stelle $x_0$.

Die Hilfsfunktion $h := f - g$ erfüllt $h(x_0) = 0$ und nimmt an der Stelle $x_0$ ein isoliertes lokales Maximum an, denn
$h'(x_0) = f'(x_0) - g'(x_0) = f'(x_0) - f'(x_0) = 0$ und
$h''(x_0) = f''(x_0) - g''(x_0) = f''(x_0) - 0 = f''(x_0) < 0$.

Also gibt es $x_1, x_2 \in I$ mit $x_1 < x_0 < x_2$, so dass $f(x_1) < g(x_1)$ und $f(x_2) < g(x_2)$.





# Konvexität

**Weiter im Beweis:**

Sei nun $t = \frac{x - x_1}{x_2 - x_1} \in (0, 1)$ die eindeutige Zahl mit $x_0 = (1 - t)x_1 + tx_2$.Dann gilt

$$f(x_0) = g(x_0) = (1 - t)g(x_1) + tg(x_2) > (1 - t)f(x_1) + t\,f(x_2),$$

im Widerspruch zur Konvexität von $f$.

"$\Longleftarrow$" Wir zeigen nun umgekehrt, dass aus $f'' \geq 0$ die Konvexität von $f$ folgt.
Seien $x_1, x_2 \in I$, $x_1 < x_2$, $x = (1 - t)x_1 + tx_2$, $t \in (0, 1)$.
Wegen des Mittelwertsatzes gibt es $\xi_1 \in (x_1, x)$ und $\xi_2 \in (x, x_2)$, so dass

$$\frac{f(x) - f(x_1)}{x - x_1} = f'(\xi_1) \underset{(f'' \geq 0)}{\leq} f'(\xi_2) = \frac{f(x_2) - f(x)}{x_2 - x}$$

Also

$$f(x) \leq f(x_1)\frac{x_2 - x}{x_2 - x_1} + f(x_2)\frac{x - x_1}{x_2 - x_1}.$$

Das ist genau die Konvexität von $f$. $\square$

# Taylorentwicklung

## Theorem

*Es sei $f : [a, b] \to \mathbb{R}$ eine $k$-mal differenzierbare Funktion.*
*Zu $x_0 \in [a, b]$ definieren wir das Polynom $(k-1)$-ten Grades*

$$T_{k-1}(x) = T_{k-1}(f, x_0)(x) := \sum_{r=0}^{k-1} \frac{f^{(r)}(x_0)}{r!}(x - x_0)^r$$
$$= f(x_0) + f'(x_0)(x - x_0) + \tfrac{f''(x_0)}{2!}(x - x_0)^2 + \ldots + \tfrac{f^{(k-1)}(x_0)}{(k-1)!}(x - x_0)^{k-1}.$$

*Dann gibt es zu jedem $x \in [a, b]$ ein $\xi$ zwischen $x_0$ und $x$, so dass*

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

**Bemerkung:** Für $k = 1$ ist dies die Aussage des Mittelwertsatzes.





**Beweis.** Wir fixieren $x \in [a, b]$, $x \neq x_0$ und betrachten die Funktion

$$g(t) = f(x) - \sum_{r=0}^{k-1} \frac{(x - t)^r}{r!} f^{(r)}(t).$$

Dann gilt $g(x_0) = f(x) - T_{k-1}(x)$ und $g(x) = 0$.
Außerdem ist $g$ differenzierbar mit

$$g'(t) = -\left( -\sum_{r=1}^{k-1} \frac{(x - t)^{r-1}}{(r-1)!} f^{(r)}(t) + \sum_{r=0}^{k-1} \frac{(x - t)^r}{r!} f^{(r+1)}(t) \right)$$
$$= -\frac{(x - t)^{k-1}}{(k-1)!} f^{(k)}(t).$$

Wir wenden nun auf dem Intervall mit Grenzen $x$ und $x_0$ den verallgemeinerten Mittelwertsatz auf $g$ und die Funktion $h(t) = (x - t)^k$ an und erhalten ein $\xi$ zwischen $x_0$ und $x$ mit

$$h'(\xi)(g(x) - g(x_0)) = g'(\xi)(h(x) - h(x_0)).$$

Nun gilt aber

$$h'(\xi)(g(x) - g(x_0)) = k(x-\xi)^{k-1}(f(x) - T_{k-1}(x))$$

und

$$g'(\xi)(h(x) - h(x_0)) = \frac{(x-\xi)^{k-1}}{(k-1)!} f^{(k)}(\xi)(x - x_0)^k.$$

Also erhalten wir für das oben gefundene $\xi$ zwischen $x_0$ und $x$

$$f(x) - T_{k-1}(x) = \frac{(x-x_0)^k}{k!} f^{(k)}(\xi)$$

wie behauptet. □

## Bemerkungen

- Der Satz besagt, dass eine $k$-mal differenzierbare Funktion durch ein Polynom vom Grad $k-1$ approximiert werden kann, und der dabei auftretende Fehler von der $k$-ten Ableitung abhängt.



Differentialrechnung Taylorentwicklung

## Bemerkungen (Fortsetzung)

- Den Ausdruck $R_k = \frac{f^{(k)}(\xi)}{k!}(x-x_0)^k$ bezeichnet man als Restglied in Lagrange-Form.
- (ÜA) Durch Anwendung des Mittelwertsatzes statt des verallgemeinerten Mittelwertsatzes auf die Hilfsfunktion $g$ mit Beweis erhält man das Restglied in Cauchy Form:

$$R_k = \frac{(x-x_0)(x-\eta)^{k-1}}{(k-1)!} f^{(k)}(\eta) \quad \text{für ein } \eta \text{ zwischen } x_0 \text{ und } x.$$

- Abschätzungen für das Restglied erlauben Aussagen zur Genauigkeit der Approximation von $f$ durch $T_{k-1}(f, x_0)$.

**Definition (Taylorreihe)**

*Ist $f$ unendlich oft differenzierbar, so bezeichnet man*

$$\begin{aligned} f(x)\big|_{x=x_0} &:= T(f,x_0)(x) \\ &:= \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k \\ &= f(x_0) + f'(x_0)(x-x_0) + \dots \end{aligned}$$

*als Taylorentwicklung beziehungsweise Taylorreihe von $f$ an der Stelle $x_0$.*





## Beispiele

- Ein Beispiel für eine überall konvergente Taylorentwicklung ist

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

Hier gilt $T(\sin,0)(x) = \sin(x)$ für alle $x \in \mathbb{R}$.

- Die Taylorreihe kann divergent sein für gewisse $x \neq x_0$.
Beispiel: Betrachte die Taylorentwicklung von $f(x) = \ln(x+1)$ an der Stelle $x = 0$. Es ist $f'(x) = \frac{1}{x+1}$ und $f^{(k)}(x) = \frac{(-1)^{k-1}(k-1)!}{(x+1)^k}$.
Damit ist die Taylorreihe an $x = 0$ gegeben durch

$$\ln(x+1)|_{x=0} = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} x^k.$$

Diese Reihe ist konvergent für $-1 < x \leq 1$ und divergent für $|x| > 1$.

Für die Beziehung von $T(f,0)$ zu $f$ betrachten wir hier den Fall $0 < x \leq 1$ und benutzen die Beschreibung des Restglieds $R_n(x)$ als

$$R_n(x) = \frac{f^{(n)}(\xi)}{n!}x^n \quad \text{für ein } \xi = \xi(x) \in (0,x).$$

Mit $f^{(k)}(x) = \frac{(-1)^{k-1}(k-1)!}{(1+x)^k}$ erhalten wir dann

$$|R_n(x)| \leq \left( \sup_{\xi\in(0,x)} \frac{(n-1)!}{n!(1+\xi)^n} \right) \cdot x^n = \frac{x^n}{n} \xrightarrow{n\to\infty} 0.$$

Im Spezialfall $x = 1$ folgt

$$\ln(2) = T(f,1) = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} \pm \cdots = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k}.$$

ÜA: Wie argumentiert man im Fall $-1 < x < 0$?



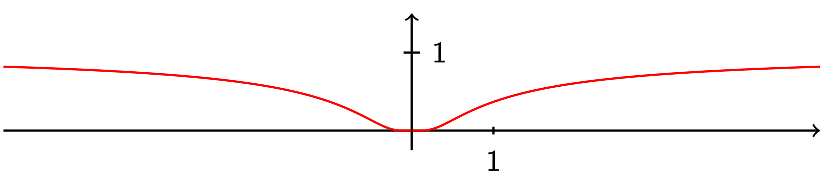


## Weitere Beispiele

- Ist die Taylorreihe konvergent, muss sie nicht gegen $f(x)$ konvergieren. Beispiel: Die Funktion

$$f(x) = \begin{cases} 0 & x = 0 \\ \exp(-1/|x|) & x \neq 0. \end{cases}$$

  ist überall unendlich oft differenzierbar, mit $f^{(n)}(0) = 0$ für alle $n \in \mathbb{N}$. Insbesondere stimmt die Taylorreihe $T(f,0) = 0$ nur in 0 mit $f$ überein.



# Verallgemeinerung

Wir werden sehen, dass

$$d(x,y) := \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2} \quad \text{für} \quad x,y \in \mathbb{R}^n$$

eine Abstandsfunktion definiert, die $\mathbb{R}^n$ zum metrischen Raum macht. Damit wissen wir, wann Funktionen $f : \mathbb{R}^n \to \mathbb{R}^m$ stetig sind.
Eine Funktion $f : \mathbb{R} \to \mathbb{R}^n$ kann durch $n$ reellwertige Komponentenfunktionen beschrieben werden:

$$t \mapsto f(t) = (f_1(t), f_2(t), \dots, f_n(t)) \ .$$

## Satz

*Eine Funktion $f : \mathbb{R} \to \mathbb{R}^n$ ist genau dann stetig, wenn alle Funktionen $f_i(t)$ stetig sind. Insbesondere ist eine Funktion $f : \mathbb{R} \to \mathbb{C}$ genau dann stetig, wenn Real- und Imaginärteil stetig sind. Dies folgt aus*

$$|x_i - y_i| = \sqrt{(x_i - y_i)^2} \leq d(x,y) \leq \sqrt{n} \cdot \max |x_i - y_i|$$





# Verallgemeinerung II

Sei $f : \mathbb{R} \to \mathbb{R}^n$ eine Funktion. Wir nehmen an, dass alle Komponentenfunktionen $f_i$ differenzierbar sind. Wir setzen dann

$$f'(t) := (f_1'(t), \dots, f_n'(t)) \ .$$

Dies definiert eine Ableitungsfunktion $f' : \mathbb{R} \to \mathbb{R}^n$. Insbesondere können wir komplexwertige Funktionen einer reellen Variable differenzieren.

## Satz

*Auch diese Ableitungsfunktion approximiert die Funktion $f$: für $a \in \mathbb{R}$ und die affine Funktion*

$$h : \mathbb{R} \to \mathbb{R}^n \quad \text{mit} \quad h(t) = f'(a)(t-a) + f(a)$$

*gilt*

$$\lim_{t \to a} \frac{d(f(t), h(t))}{t-a} = 0$$

# Verallgemeinerung III

**Beweis.** Aus

$$d(f(t), h(t))^2 = \sum_{i=1}^{n} (f_i(t) - h_i(t))^2 = \sum_{i=1}^{n} (f_i(t) - f_i'(a)(t-a) - f_i(a))^2$$

folgt

$$\frac{d(f(t), h(a))^2}{(t-a)^2} = \sum_{i=1}^{n} \left( \frac{f_i(t) - f_i(a)}{t-a} - f_i'(a) \right)^2 \xrightarrow{t \to a} 0$$

□





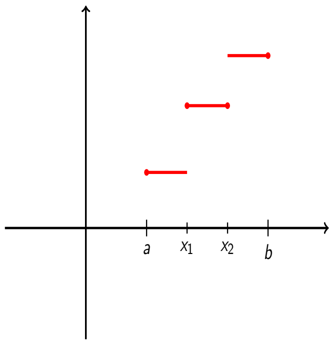
# Kapitel 9

Integralrechnung

# Integralrechnung

## Definition (Treppenfunktion)

*Eine Funktion $\varphi\colon [a,b] \to \mathbb{R}$ heißt Treppenfunktion, wenn es eine Unterteilung $Z: a = x_0 < x_1 < \cdots < x_n = b$ des Intervalls $[a,b]$ gibt, so dass $\varphi$ auf jedem der offenen Teilintervalle $(x_{i-1}, x_i)$ konstant ist, $i = 1, 2, \ldots, n$.*




**Bemerkung (ÜA)**

Seien $\varphi, \psi\colon [a,b] \to \mathbb{R}$ Treppenfunktionen. Dann sind $\varphi + \psi$ und $\varphi \cdot \psi\colon [a,b] \to \mathbb{R}$ Treppenfunktionen.





## Definition (Integral einer Treppenfunktion)

*Sei $\varphi\colon [a,b] \to \mathbb{R}$ eine Treppenfunktion und $a = x_0 < x_1 < \cdots < x_n = b$ eine Unterteilung des Intervalls $[a,b]$, so dass $\varphi|_{(x_{i-1},x_i)} = c_i = const$, $i = 1, 2, \ldots, n$.*

*Wir definieren das Integral der Treppenfunktion $\varphi$ durch*

$$\int_a^b \varphi(x)dx := \sum_{i=1}^{n} c_i(x_i - x_{i-1}).$$

**Bemerkung (ÜA).**

Überlegen Sie sich, dass diese Definition nicht von der Wahl der Unterteilung abhängt und dass für alle $c \in (a,b)$ gilt:

$$\int_a^b \varphi(x)dx = \int_a^c \varphi(x)dx + \int_c^b \varphi(x)dx.$$

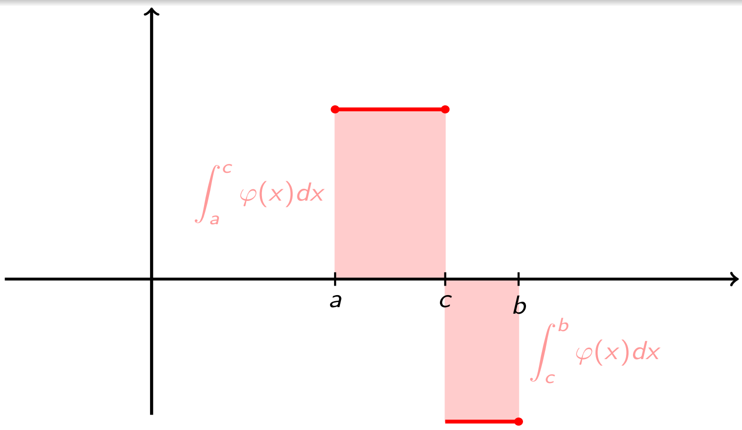
## Geometrische Interpretation

Sei $F$ das zwischen der $x$-Achse und dem Graphen der Funktion $\varphi$ liegende Gebiet.

$F$ ist eine endliche Vereinigung von Rechtecken. Sei $A(F)$ der Flächeninhalt von $F$. Wenn $\varphi \geq 0$, so ist $\int_a^b \varphi(x)dx = A(F) \geq 0$.

Wenn $\varphi \leq 0$, so ist $\int_a^b \varphi(x)dx = -A(F) \leq 0$.

D.h. die Fläche oberhalb der $x$-Achse trägt positiv, die unterhalb der $x$-Achse negativ zum Integral bei.








## Satz (Linearität und Monotonie des Integrals von Treppenfunktionen)

*Es seien $\varphi, \psi \colon [a, b] \to \mathbb{R}$ Treppenfunktionen und $\lambda \in \mathbb{R}$. Dann gilt:*

(i) $\int_a^b (\varphi + \psi)(x)dx = \int_a^b \varphi(x)dx + \int_a^b \psi(x)dx.$

(ii) $\int_a^b (\lambda\varphi)(x)dx = \lambda \int_a^b \varphi(x)dx.$

(iii) $\varphi \leq \psi \Longrightarrow \int_a^b \varphi(x)dx \leq \int_a^b \psi(x)dx.$ *(Monotonie)*

**Beweis.** Durch Verfeinerung finden wir ein Unterteilung $a = x_0 < x_1 < \cdots < x_n = b$, so dass $\varphi|_{(x_{i-1},x_i)} = c_i$ und $\psi|_{(x_{i-1},x_i)} = d_i$. Damit erhalten wir (i):

$$\begin{aligned}
\int_a^b (\varphi + \psi)(x)dx &= \sum_{i=1}^n (c_i + d_i)(x_i - x_{i-1}) \\
&= \sum_{i=1}^n c_i(x_i - x_{i-1}) + \sum_{i=1}^n d_i(x_i - x_{i-1}) \\
&= \int_a^b \varphi(x)dx + \int_a^b \psi(x)dx.
\end{aligned}$$

**Weiter im Beweis:**

(ii)

$$\begin{aligned}\int_a^b (\lambda\varphi)(x)dx &= \sum_{i=1}^n (\lambda c_i)(x_i - x_{i-1}) \\ &= \lambda \sum_{i=1}^n c_i(x_i - x_{i-1}) \quad = \quad \lambda \int_a^b \varphi(x)dx.\end{aligned}$$

(iii) Aus $\varphi \le \psi$ folgt $c_i \le d_i$ für alle $i \in \{1, 2, \ldots, n\}$. Damit ist dann

$$\begin{aligned}\int_a^b \varphi(x)dx &= \sum_{i=1}^n c_i(x_i - x_{i-1}) \\ &\le \sum_{i=1}^n d_i(x_i - x_{i-1}) \quad = \quad \int_a^b \psi(x)dx.\end{aligned}$$

□





**Definition (Ober- und Unterintegral)**

*Sei $f : [a, b] \to \mathbb{R}$ eine beschränkte Funktion. Das Oberintegral von $f$ ist die Zahl*

$$\int_a^{*b} f(x)dx := \inf \left\{ \int_a^b \varphi(x)dx \;\middle|\; \varphi \quad \textit{Treppenfkt. mit} \quad \varphi \ge f \right\}.$$

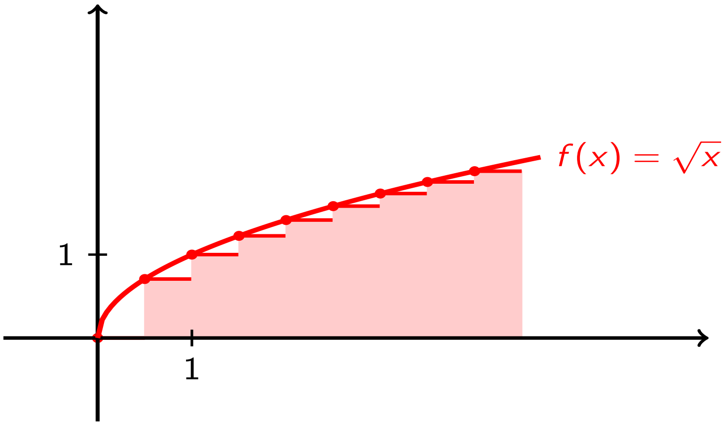
*Das Unterintegral von $f$ ist die Zahl*

$$\int_{*a}^{b} f(x)dx := \sup \left\{ \int_a^b \varphi(x)dx \;\middle|\; \varphi \quad \textit{Treppenfkt. mit} \quad \varphi \le f \right\}.$$

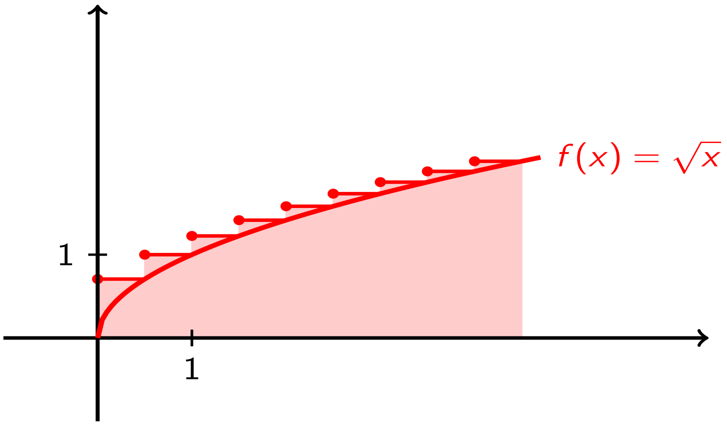
**Bemerkungen (ÜA)**

- $\int_{*a}^{b} f(x)dx \le \int_a^{*b} f(x)dx$.
- $\int_{*a}^{b} \varphi(x)dx = \int_a^{*b} \varphi(x)dx = \int_a^b \varphi(x)dx$ für jede Treppenfunktion

Untersumme




Obersumme








**Beispiel**

Sei $f\colon [0,1] \to \mathbb{R}$ definiert durch

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

Dann gilt

$$\int_{*0}^{1} f(x)dx = 0 \quad \text{und} \quad \int_{0}^{*1} f(x)dx = 1.$$

Ober- und Unterinteral stimmen also nicht notwendigerweise überein.

## Satz (Eigenschaften des Ober- und Unterintegrals)

*Es seien $f, g\colon [a.b] \to \mathbb{R}$ beschränkte Funktionen. Dann gilt:*

(i) $\int^*(f+g) \leq \int^* f + \int^* g$,

(ii) $\int^*(\lambda f) = \lambda \int^* f$ *für alle* $\lambda \geq 0$,

(iii) $\int_*(f+g) \geq \int_* f + \int_* g$,

(iv) $\int_*(\lambda f) = \lambda \int_* f$ *für alle* $\lambda \geq 0$,

(v) $\int^*(\lambda f) = \lambda \int_* f$ *und* $\int_*(\lambda f) = \lambda \int^* f$ *für alle* $\lambda \leq 0$.

*(Hierbei ist* $\int^* f = \int_a^{*b} f(x)dx$, *usw.)*

**Beweis.** Für jede Teilmenge $D \subseteq \mathbb{R}$ gilt $\sup D = -\inf(-D)$, wobei $-D := \{-x | x \in D\}$.

Insbesondere gilt $\int_* f = -\int^*(-f)$.

(iii)–(v) folgt damit aus (i) und (ii).





**Weiter im Beweis:**

Wir beweisen nun ($i$) und ($ii$).

(i) Wir zeigen $\int^*(f+g) \leq \int^* f + \int^* g$:

$$\begin{aligned}
\int^*(f+g) &= \inf\{\textstyle\int \xi | \xi \geq f+g \text{ und } \xi \text{ Treppenfunktion}\}\\
&\leq \inf\{\textstyle\int \xi | \xi = \varphi + \psi, \varphi \geq f, \psi \geq g \text{ und } \varphi, \psi \text{ Treppenfunktionen}\}\\
&= \inf\{\textstyle\int \varphi + \int \psi | \varphi \geq f, \psi \geq g \text{ Treppenfunktionen}\}\\
&= \inf\{\textstyle\int \varphi | \varphi \geq f \text{ Treppenfkt.}\} + \inf\{\int \psi | \psi \geq g \text{ Treppenfunktion}\}\\
&= \textstyle\int^* f + \int^* g.
\end{aligned}$$

(ii) Wir zeigen $\int^*(\lambda f) = \lambda \int^* f$ für $\lambda > 0$:

$$\begin{aligned}
\int^*(\lambda f) &= \inf\{\textstyle\int \varphi | \varphi \geq \lambda f \text{ Treppenfkt.}\}\\
&= \inf\{\textstyle\int \varphi | \frac{1}{\lambda}\varphi \geq f \text{ Treppenfkt.}\}\\
&= \inf\{\lambda \textstyle\int \psi | \psi \geq f \text{ Treppenfkt.}\} \qquad [\text{mit } \psi = \frac{1}{\lambda}\varphi]\\
&= \lambda \inf\{\textstyle\int \psi | \psi \geq f \text{ Treppenfkt.}\} = \lambda \int^* f.
\end{aligned}$$

□

## Definition (Integrierbare Funktion)

*Eine beschränkte Funktion $f: [a,b] \to \mathbb{R}$ heißt integrierbar (genauer: Riemann-integrierbar), wenn*

$$\int_a^{*b} f(x)dx = \int_{*a}^b f(x)dx.$$

*Man definiert dann das (Riemann-)Integral von $f$ als die reelle Zahl*

$$\int_a^b f(x)dx := \int_a^{*b} f(x)dx = \int_{*a}^b f(x)dx.$$

**Notation.**
Für $b > a$ setzen wir $\int_b^a f(x)\,dx = -\int_a^b f(x)\,dx$, und außerdem $\int_a^a f(x)\,dx = 0$.





## Satz (Linearität und Monotonie des Riemann-Integrals)

*Es seien $f, g: [a,b] \to \mathbb{R}$ integrierbar und $\lambda \in \mathbb{R}$. Dann sind $f+g$ und $\lambda f$ integrierbar und es gilt:*

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

(ii) $\int_a^b (\lambda f)(x)dx = \lambda \int_a^b f(x)dx$ *und*

(iii) $f \leq g \Longrightarrow \int_a^b f(x)dx \leq \int_a^b g(x)dx$.

**Beweis.**

(i) Aus

$$\begin{aligned}\int^*(f+g) \leq \int^* f + \int^* g \quad &= \int f + \int g \\ = \int_* f + \int_* g \quad &\leq \int_*(f+g) \leq \int^*(f+g)\end{aligned}$$

folgt

$$\int^*(f+g) = \int_*(f+g) = \int f + \int g.$$

**Weiter im Beweis:**

(ii) Für $\lambda > 0$ gilt

$$\int^*(\lambda f) = \lambda \int^* f = \lambda \int f = \lambda \int_* f = \int_*(\lambda f).$$

Für $\lambda < 0$ gilt

$$\int^*(\lambda f) = \lambda \int_* f = \lambda \int f = \lambda \int^* f = \int_*(\lambda f).$$

In beiden Fällen ist also $\lambda f$ integrierbar und

$$\int(\lambda f) = \lambda \int f.$$

(iii) Aus $f \leq g$ folgern wir

$$\int f = \int_* f = \sup\{\int \varphi | \varphi \leq f \text{ Treppenfkt.}\}$$
$$\leq \sup\{\int \varphi | \varphi \leq g \text{ Treppenfkt.}\} = \int_* g = \int g.$$

$\square$





## Bemerkungen (ÜA)

(i) Sei $a < b < c$. Dann ist $f : [a, c] \to \mathbb{R}$ integrierbar genau dann, wenn $f|_{[a,b]}$ und $f|_{[b,c]}$ integrierbar sind.

Es gilt dann

$$\int_a^c f = \int_a^b f + \int_b^c f.$$

(ii) Positiv- und Negativteil $f_\pm : [a, b] \to \mathbb{R}_+ \cup \{0\}$ einer Fkt. $f : [a, b] \to \mathbb{R}$ sind definiert als $f_+(x) := \max\{f(x), 0\}$ bzw. $f_-(x) := -\min\{f(x), 0\}$, $x \in [a, b]$.

Es gilt: $f$ integrierbar $\Longrightarrow$ $f_+$, $f_-$ und $|f|$ integrierbar.

(iii) Aus der Monotonie des Integrals folgt

$$\left|\int_a^b f\right| \leq \int_a^b |f|$$

für jede integrierbare Funktion $f$.

# Integral von Funktionen mit Werten in $\mathbb{R}^n$ oder $\mathbb{C}$

Wir haben auch die folgende Verallgemeinerung: Sei $f\colon [a,b] \to \mathbb{R}^n$ eine Funktion mit Komponentenfunktionen

$$f(t) = (f_1(t), \ldots, f_n(t)) \ ,$$

so dass jede Komponentenfunktion $f_i : [a,b] \to \mathbb{R}$ integrierbar ist. Dann setzen wir

$$\int_a^b f(t) = (\int_a^b f_1, \ldots, \int_a^b f_n) \in \mathbb{R}^n.$$

Wir nennen dann $f$ (Riemann-)integrierbar.

Insbesondere nimmt das Integral einer komplexwertigen Funktion $f\colon [a,b] \to \mathbb{C}$ seine Werte in den komplexen Zahlen an.





## Theorem (Integrierbarkeit stetiger Funktionen)

*Jede stetige Funktion $f\colon [a,b] \to \mathbb{R}$ ist integrierbar.*

**Beweis.**

- Es genügt zu zeigen, dass es für jedes $\varepsilon > 0$ zwei Treppenfunktionen $\varphi, \psi\colon [a,b] \to \mathbb{R}$ gibt mit $\varphi \le f \le \psi$ und $0 \le \int \psi - \int \varphi < \varepsilon$.
- Dazu benutzen wir folgenden Hilfssatz:

## Lemma

*Jede auf einem abgeschlossenen Intervall $[a,b]$ definierte und dort stetige Funktion $f\colon [a,b] \to \mathbb{R}$ ist gleichmäßig stetig, d.h.*

*Für jedes $\varepsilon > 0$ existiert ein $\delta > 0$, so dass $|f(x) - f(y)| < \varepsilon$ für alle Paare $x, y \in [a,b]$ mit $|x - y| < \delta$.*

## Beispiel

Die Funktion $f\colon \mathbb{R} \to \mathbb{R}$, $f(x) = x^2$ ist nicht gleichmäßig stetig, ihre Einschränkung auf jedes abgeschlossene Intervall $[a,b]$ schon.

**Beweis des Lemmas.**

- Wir nehmen an, $f$ sei nicht gleichmäßig stetig und führen das zum Widerspruch.
- Wir nehmen also an, es gibt ein $\varepsilon > 0$ und Folgen $x_n, y_n \in [a, b]$ mit
$$|x_n - y_n| < \frac{1}{n}, \quad \text{aber} \quad |f(x_n) - f(y_n)| \geq \varepsilon.$$
- Nach Bolzano-Weierstraß existiert eine Teilfolge $(x_{n_k})_{k\in\mathbb{N}}$, die gegen ein $\ell \in [a, b]$ konvergiert.
- Wegen $\lim_{n\to\infty} |x_n - y_n| = 0$ gilt auch $\lim_{k\to\infty} y_{n_k} = \ell$.
- Die Stetigkeit von $f$ liefert dann
$$\lim_{k\to\infty} f(x_{n_k}) = f(\ell) = \lim_{k\to\infty} f(y_{n_k}),$$
- im Widerspruch zu $|f(x_n) - f(y_n)| \geq \varepsilon$ für alle $n \in \mathbb{N}$. □





**Beweis des Theorems**

- Zu $n \in \mathbb{N}$ betrachten wir die äquidistante Unterteilung
$x_i = a + i\frac{(b-a)}{n}, \quad i = 0, 1, \ldots, n$ des Intervals $[a, b]$.
Mit wachsendem $n$ werden diese Unterteilungen immer feiner.
- Wir setzen $c_i := \min f|_{[x_{i-1},x_i]}$ und $d_i := \max f|_{[x_{i-1},x_i]}$.
- Nach dem Lemma gibt es zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$, so dass
$$d_i - c_i < \frac{\varepsilon}{b-a} \quad \text{für alle } n \geq N \text{ und alle } i = 1, \ldots n.$$
- Wir definieren nun zwei Treppenfunktionen
$$\varphi|_{[x_{i-1},x_i)} := c_i \quad \text{und} \quad \psi|_{[x_{i-1},x_i)} := d_i, \quad i = 1, \ldots, n$$
und $\varphi(b) := \psi(b) := f(b)$.
- Offenbar gilt dann $\varphi \leq f \leq \psi$ und
$$\int_a^b \psi(x)dx - \int_a^b \varphi(x)dx = \sum_{i=1}^n \underbrace{(d_i - c_i)}_{<\frac{\varepsilon}{b-a}} \frac{b-a}{n} < \varepsilon. \quad \square$$

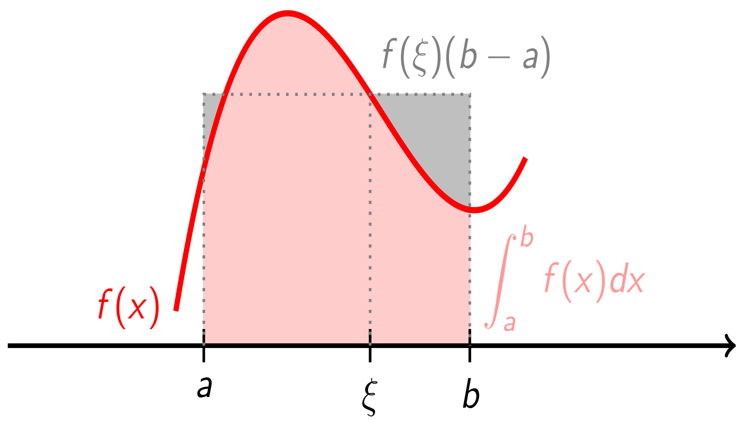
## Übungsaufgabe

Jede monotone Funktion $f : [a, b] \to \mathbb{R}$ ist integrierbar.

## Satz (Mittelwertsatz der Integralrechnung)

*Sei $f : [a, b] \to \mathbb{R}$ eine stetige Funktion. Dann existiert $\xi \in [a, b]$, so dass*

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








**Beweis.**

- Aus $\min f \leq f \leq \max f$ folgt wg. der Monotonie des Integrals
$$(b-a)\min f \leq \int_a^b f(x)dx \leq (b-a)\max f.$$
- Der Zwischenwertsatz für die stetige Funktion $f$ liefert dann die Existenz von $\xi \in [a, b]$ mit $f(\xi) = \frac{\int_a^b f(x)dx}{b-a}$.

$\square$

### Satz (Integrierbarkeit von Produkten)

*Es seien $f, g\colon [a,b] \to \mathbb{R}$ integrierbare Funktionen und $p \in [1, \infty)$. Dann gilt:*

(i) *Die Funktion $|f|^p$ ist integrierbar.*

(ii) *Die Funktion $f \cdot g$ ist integrierbar.*

**Beweis.**
(i): Nach dem Satz (und Bemerkungen) über Linearität und Monotonie können wir annehmen, dass $0 \leq f \leq 1$ gilt.
Zu $\varepsilon > 0$ wählen wir Treppenfunktionen $\varphi, \psi\colon [a,b] \to \mathbb{R}$ mit $0 \leq \varphi \leq f \leq \psi \leq 1$, so dass

$$\int_a^b (\psi - \varphi)(x)dx \leq \frac{\varepsilon}{p}.$$





## Weiter im Beweis:

Die Funktion $x \mapsto x^p$ hat die Ableitung $px^{p-1}$, welche auf dem Intervall $[0,1]$ durch $p$ beschränkt ist. Aus $0 \leq \varphi^p \leq \psi^p \leq 1$ und dem Schrankensatz (Folgerung des Mittelwertsatzes der Differentialrechnung) erhalten wir nun

$$\psi^p - \varphi^p \leq p(\psi - \varphi),$$

und somit gilt

$$\int_a^b (\psi^p - \varphi^p)(x)dx \leq p \int_a^b (\psi - \varphi)(x)dx \leq \varepsilon.$$

Wegen $0 \leq \varphi^p \leq f^p \leq \psi^p$ ist also $|f|^p = f^p$ integrierbar.

(ii) folgt aus (i) wg. $4f \cdot g = (f+g)^2 - (f-g)^2$. $\square$

## Definition (Riemannsche Summe)

*Sei $f : [a, b] \to \mathbb{R}$ eine Funktion,*

- *$a = x_0 < x_1 < \cdots < x_n = b$ eine Unterteilung des Intervalls $[a, b]$ und*
- *$\xi_i \in [x_{i-1}, x_i]$ beliebige Zahlen in diesen Teilintervallen.*

*Die zugehörige **Riemannsche Summe** ist die Zahl*

$$\textstyle\sum_{i=1}^{n} f(\xi_i)(x_i - x_{i-1}).$$

*Die $\xi_i$ heißen **Stützstellen**.*

## Bemerkung

Die Riemannsche Summe kann aufgefasst werden als das Integral einer Treppenfunktion mit Werten gegeben durch Funktionswerte an den Stützstellen.





## Satz (Approximation durch Riemannsche Summen)

*Für jede stetige Funktion $f : [a, b] \to \mathbb{R}$ approximieren Riemannsche Summen in folgendem Sinne das Integral:*

*Für jedes $\varepsilon > 0$ gibt es ein $\delta > 0$, so dass*

$$\left| \int_a^b f(x)dx - \sum_{i=1}^{n} f(\xi_i)(x_i - x_{i-1}) \right| < \varepsilon$$

*für jede Riemannsche Summe mit $\max\limits_{1 \le i \le n} (x_i - x_{i-1}) < \delta$.*

[Animation]

## Bemerkung

Der Satz gilt sogar für beliebige (Riemann-) integrierbare Funktionen $f : [a, b] \to \mathbb{R}$, vgl. Forster.

**Beweis.**

- Wegen der (gleichmäßigen) Stetigkeit von $f : [a, b] \to \mathbb{R}$ gibt es zu jedem $\varepsilon > 0$ ein $\delta > 0$, so dass
$$|f(x) - f(y)| < \frac{\varepsilon}{b-a} \text{ für alle } x, y \text{ mit } |x - y| < \delta.$$
- Für Riemannsche Summen mit $\max\limits_{1 \le i \le n} (x_i - x_{i-1}) < \delta$ gilt dann: $|f(x) - f(\xi_i)| < \frac{\varepsilon}{b-a}$ für alle $x \in [x_{i-1}, x_i]$.
- Damit erhalten wir
$$\left| \int_a^b f(x)dx - \sum_{i=1}^n f(\xi_i)(x_i - x_{i-1}) \right| = \left| \sum_{i=1}^n \int_{x_{i-1}}^{x_i} (f(x) - f(\xi_i))dx \right|$$
$$\le \sum_{i=1}^n \int_{x_{i-1}}^{x_i} \underbrace{|f(x) - f(\xi_i)|}_{< \frac{\varepsilon}{b-a}} dx < \sum_{i=1}^n (x_i - x_{i-1}) \frac{\varepsilon}{b-a} = \varepsilon.$$

$\square$





## Beispiel: Berechnung des Integrals $\int_0^b x \, dx$

- Wähle eine äquidistante Unterteilung $x_i := \frac{ib}{n}$, $i = 1, \ldots n$.
- Wähle als Stützstellen $\xi_i := x_i$.
- Die zugehörige Riemannsche Summe ist dann
$$\sum_{i=1}^n \frac{ib}{n}\frac{b}{n} = \frac{b^2}{n^2} \sum_{i=1}^n i$$
$$= \frac{b^2}{n^2} \frac{n(n+1)}{2} = \frac{b^2}{2}\left(1 + \frac{1}{n}\right).$$
- Wir erhalten $\int_0^b x \, dx$ als Grenzwert der Riemannschen Summe für $n \to \infty$:
$$\int_0^b x \, dx = \frac{b^2}{2}.$$
Dies ist in der Tat die Fläche des gleichschenkligen rechtwinkligen Dreiecks unter dem Graphen von $f(x) = x$.

**Verabredung:** Im Folgenden sei $I \subseteq \mathbb{R}$ immer ein Intervall.

### Definition (Stammfunktion)

*Sei $f : I \to \mathbb{R}$ eine Funktion. Eine Stammfunktion von $f$ ist eine differenzierbare Funktion $F : I \to \mathbb{R}$, so dass*

$$F' = f.$$

Stammfunktionen $F_1$ und $F_2$ zu gegebenem $f$ unterscheiden sich auf einem Intervall $I$ durch eine reelle Konstante: aus $(F_1 - F_2)' = f - f = 0$ folgt $F_1 - F_2 \equiv c \in \mathbb{R}$.

### Theorem (Fundamentalsatz der Differential- und Integralrechnung)

*Sei $f : I \to \mathbb{R}$ eine stetige Funktion und $x_0 \in I$.*

*Die durch*

$$F(x) := \int_{x_0}^{x} f(t)dt$$

*definierte Funktion $F : I \to \mathbb{R}$ ist eine Stammfunktion von $f$.*





**Beweis des Fundamentalsatzes.**
Wir benutzen den Mittelwertsatz der Integralrechnung.

- Sei $0 \neq h \in \mathbb{R}$ derart, dass $x, x+h \in I$. Dank des Satzes gilt

  $$\int_{x}^{x+h} f(t)dt = f(\xi)h,$$

  wobei $\xi \in [x, x+h]$, falls $h > 0$ und $\xi \in [x+h, x]$, falls $h < 0$.
- Damit erhalten wir für den Differenzenquotienten von $F$:

$$\begin{aligned} \frac{F(x+h) - F(x)}{h} &= \frac{\int_{x_0}^{x+h} f(t)dt - \int_{x_0}^{x} f(t)dt}{h} \\ &= \frac{\int_{x}^{x+h} f(t)dt}{h} \\ &= f(\xi) \xrightarrow{h \to 0} f(x), \end{aligned}$$

d.h. $F'(x) = f(x)$. $\square$

## Folgerung

*Sei $f : I \to \mathbb{R}$ stetig und $F$ eine Stammfunktion von $f$.*

*Dann gilt für alle $a, b \in I$:*

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

**Beweis.** Sei $x_0 \in [a, b]$. Sowohl $F$ als auch $\int_{x_0}^x f(t)dt$ sind Stammfunktionen von $f$, d.h. $F(x) = \int_{x_0}^x f(t)dt + c$ und somit

$$F(b) - F(a) = \int_{x_0}^b f(t)dt + c - \left(\int_{x_0}^a f(t)dt + c\right) = \int_a^b f(t)dt.$$

□

## Notation für die Differenz von Randtermen

$$F|_a^b := F(b) - F(a).$$





## Beispiele

- $\int\limits_0^\pi \sin(x)dx = -\cos(x)|_0^\pi = -\cos(\pi) + \cos(0) = -(-1) + 1 = 2.$
- Ist $\alpha \neq -1$, so gilt $\int\limits_a^b x^\alpha dx = \frac{1}{\alpha+1} x^{\alpha+1}|_a^b$, wobei
  - $a, b \in \mathbb{R}$ beliebig, falls $\alpha \in \mathbb{N}$,
  - $0 \notin [a, b]$, falls $\alpha \in \mathbb{Z}$ und $\alpha \leq -2$, und
  - $0 < a < b$, falls $\alpha \notin \mathbb{Z}$.
- Für $0 < a < b$ erhält man $\int\limits_a^b \frac{1}{x} dx = \ln(x)|_a^b$

  und für $a < b < 0$ gilt $\int\limits_a^b \frac{1}{x} dx = \ln(-x)|_a^b$.

  D.h. für $0 \notin [a, b]$ ergibt sich also $\int\limits_a^b \frac{1}{x} dx = \ln(|x|)|_a^b$

## Theorem (Substitutionsregel)

*Sei $f: I \to \mathbb{R}$ eine stetige Funktion und $\varphi: [a,b] \to \mathbb{R}$ eine stetig differenzierbare Funktion mit $\varphi([a,b]) \subseteq I$. Dann gilt*

$$\int_a^b f(\varphi(t))\varphi'(t)dt = \int_{\varphi(a)}^{\varphi(b)} f(x)dx.$$

**Beweis.**

- Sei $F: I \to \mathbb{R}$ eine Stammfunktion von $f$.
- Nach der Kettenregel gilt für alle $t \in [a,b]$

$$(F \circ \varphi)'(t) = F'(\varphi(t))\varphi'(t) = f(\varphi(t))\varphi'(t).$$

- Somit

$$\begin{aligned} \int_a^b f(\varphi(t))\varphi'(t)dt &= \int_a^b (F \circ \varphi)'(t)dt \\ &= F \circ \varphi\big|_a^b = F\big|_{\varphi(a)}^{\varphi(b)} = \int_{\varphi(a)}^{\varphi(b)} f(x)dx. \end{aligned}$$

$\square$





## Schreibweise:

Definiert man

$$d\varphi := \varphi'(t)dt = \frac{d\varphi}{dt}dt \ ,$$

so kann man die Substitutionsregel verkürzt schreiben als

$$\int_a^b f \circ \varphi \ d\varphi = \int_{\varphi(a)}^{\varphi(b)} f \ dx.$$

## Beispiele

- Sei $c \neq 0$ und $d \in \mathbb{R}$, $f$ stetig. Dann ist

$$\int_a^b f(ct+d)dt = \frac{1}{c}\int_{ca+d}^{cb+d} f(x)dx.$$

-

$$\int_0^2 \frac{2t}{t^2+1}dt \underset{(x=t^2)}{=} \int_0^4 \frac{dx}{x+1} = \ln(x+1)|_0^4$$
$$= \ln 5 - \ln 1 = \ln 5$$

Hier ist $f(x) = \frac{1}{x+1}$, $\varphi(t) = t^2$ und damit $\varphi'(t) = 2t$.





## Beispiel: Fläche des Kreises vom Radius 1

Wir stellen den Halbkreis als Graphen der stetigen Funktion $f\colon [-1,1] \to \mathbb{R}$, $f(x) = \sqrt{1-x^2}$ dar und müssen $\int_{-1}^1 \sqrt{1-x^2}dx$ berechnen.
Die Substitution $x = \varphi(t) = \sin(t)$ ergibt

$$\int_{-1}^1 \sqrt{1-x^2}dx = \int_{-\pi/2}^{\pi/2} \sqrt{1-\sin^2 t}\; d\sin(t) = \int_{-\pi/2}^{\pi/2} \cos^2(t)dt.$$

Es ist aber $\cos^2 t = \left(\frac{e^{it}+e^{-it}}{2}\right)^2 = \frac{e^{2it}+e^{-2it}}{4} + \frac{1}{2} = \frac{1}{2}(\cos 2t + 1)$
Damit ist die Fläche des Kreises vom Radius 1 gegeben durch

$$2\int_{-1}^1 \sqrt{1-x^2}dx = 2\int_{-\pi/2}^{\pi/2} \cos^2(t)dt = \int_{-\pi/2}^{\pi/2} \cos 2t\; dt + \int_{-\pi/2}^{\pi/2} dt$$

$$= \frac{1}{2}\int_{-\pi}^{\pi} \cos s\; ds + \int_{-\pi/2}^{\pi/2} dt = \frac{1}{2}\sin s|_{-\pi}^{\pi} + t|_{-\pi/2}^{\pi/2} = 0 + \pi = \pi.$$

## Theorem (Partielle Integration)

*Es seien $f, g\colon [a,b] \to \mathbb{R}$ stetig differenzierbar. Dann gilt*

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

**Beweis.** Folgt aus der Produktregel $(fg)' = f'g + fg'$ durch Integration:

$$(fg)|_a^b \overset{\textit{Fundamentalsatz}}{=} \int_a^b (fg)'(x)dx \;=\; \int_a^b f'(x)g(x)dx + \int_a^b f(x)g'(x)dx.$$

□

**Kurzschreibweise für die partielle Integration:**

$$\int f\,dg = fg - \int g\,df.$$





## Beispiele

- Sei $a, b > 0$. Dann ist
  $$\int_a^b \ln x\,dx = \int_a^b 1 \cdot \ln x\,dx = x \ln x|_a^b - \int_a^b \underbrace{x \cdot \frac{1}{x}}_{=1} dx = x(\ln x - 1)|_a^b$$
  Hierbei ist
  $$\begin{array}{rclcrcl} f(x) & = & x & \quad & f'(x) & = & 1 \\ g(x) & = & \ln x & & g'(x) & = & 1/x \end{array}$$
- In der Tat rechnet man nach, dass $x(\ln x - 1)$ eine Stammfunktion des Logarithmus ist:
  $$(x(\ln x - 1))' = \ln x - 1 + x \cdot \frac{1}{x} = \ln x.$$

## Beispiele

- $x \arctan x - \frac{1}{2}\ln(x^2+1)$ ist eine Stammfunktion des arctan, denn:

$$\begin{aligned}\int_a^b \arctan x \; dx &= x \arctan x|_a^b - \int_a^b x \arctan' x \; dx \\ &= x \arctan x|_a^b - \int_a^b \frac{x}{1+x^2} dx\end{aligned}$$

  Mittels der Substitution $t = x^2$ hatten wir gesehen, dass $\int_a^b \frac{x}{x^2+1} dx = \frac{1}{2}\int_{a^2}^{b^2} \frac{1}{t+1} dt = \frac{1}{2}\ln(t+1)|_{a^2}^{b^2} = \frac{1}{2}\ln(x^2+1)|_a^b$.
  Damit ist $\int_a^b \arctan x \; dx = \left(x \arctan x - \frac{1}{2}\ln(x^2+1)\right)\Big|_a^b$.
- Wieder war partielle Integration ein Mittel, *um die Stammfunktion zu finden*. Um das Ergebnis zu überprüfen, leitet man einfach ab (ÜA).



Integralrechnung Partielle Integration

## Übung

Für alle $m, n \in \mathbb{Z}$ ist

- $\int\limits_0^{2\pi} e^{imx} e^{inx} \mathrm{d}\, x = \begin{cases} 2\pi, & n = -m, \\ 0, & \text{sonst}, \end{cases}$
- $\int\limits_0^{2\pi} \sin(mx)\cos(nx)\, \mathrm{d}\, x = 0,$
- $\int\limits_0^{2\pi} \sin(mx)\sin(nx)\, \mathrm{d}\, x = \begin{cases} -\pi, & n = -m, \\ \pi, & n = m, \\ 0, & \text{sonst}, \end{cases}$

### Übung

- $\int_0^{2\pi} \cos(mx)\cos(nx)\mathrm{d}\,x = \begin{cases} 2\pi & \text{wenn } n = m = 0, \\ \pi & \text{wenn } 0 \neq n = -m \text{ oder } 0 \neq n = m, \\ 0 & \text{sonst.} \end{cases}$





## Uneigentliche Integrale

Wir erweitern den Riemannschen Integralbegriff noch in zwei Richtungen.

### Definition (Integral über unbeschränkte Integrationsbereiche)

*Sei $f : [a, \infty) \to \mathbb{R}$ eine Funktion, so dass für jedes $x > a$ die Einschränkung $f|_{[a,x]}$ beschränkt und integrierbar ist. Wir nennen $f$ uneigentlich (Riemann-)integrierbar auf $[a, \infty)$, falls der Grenzwert*

$$\int_a^\infty f(t)\,dt := \lim_{x\to\infty} \int_a^x f(t)\,dt$$

*existiert. Analog verfährt man mit $f : (-\infty, b] \to \mathbb{R}$.*
*Eine Funktion $f : \mathbb{R} \to \mathbb{R}$ heißt uneigentlich integrierbar auf $\mathbb{R}$, falls für $c \in \mathbb{R}$ die beiden Integrale*

$$\int_{-\infty}^c f(t)\,dt \quad \text{und} \quad \int_c^\infty f(t)\,dt$$

*existieren.*

# Beispiel

- Für welche $s \in (0, \infty)$ existiert das Integral

$$\int_1^\infty \frac{1}{t^s}\,dt\,?$$

Für $s \neq 1$ haben wir

$$\int_1^R t^{-s}\,dt = \frac{1}{-s+1} t^{-s+1}\Big|_1^R = \frac{1}{1-s}\left(R^{1-s}-1\right),$$

und für $s = 1$ gilt

$$\int_1^R t^{-1}\,dt = \ln(t)\Big|_1^R = \ln(R).$$

Somit ergibt sich für $s > 1$ der Grenzwert $\frac{1}{s-1}$, und für $s \leq 1$ existiert das Integral über $[1, \infty)$ nicht.





## Definition (Integrale unbeschränkter Funktionen)

*Sei $f : (a, b) \to \mathbb{R}$ nicht unbedingt beschränkt, aber für jedes Teilintervall $[\alpha, \beta] \subseteq (a, b)$ integrierbar. Falls dann für $c \in (a, b)$ die Grenzwerte*

$$\lim_{\alpha \searrow a} \int_\alpha^c f(t)\,dt \quad \textit{und} \quad \lim_{\beta \nearrow b} \int_c^\beta f(t)\,dt$$

*existieren, so nennen wir $f$ über das Intervall $(a, b)$ (eventuell uneigentlich) integrierbar, und setzen*

$$\int_a^b f(t)\,dt := \lim_{\alpha \searrow a} \lim_{\beta \nearrow b} \int_\alpha^\beta f(t)\,dt.$$

# Beispiel

- Für welche $s \in (0, \infty)$ existiert das Integral $\displaystyle\int_0^1 \frac{1}{t^s}\,dt$?

  Für $s \neq 1$ haben wir

$$\int_\varepsilon^1 t^{-s}\,dt = \frac{1}{-s+1} t^{-s+1}\Big|_\varepsilon^1 = \frac{1}{1-s}\left(1-\varepsilon^{1-s}\right),$$

  und für $s = 1$ gilt

$$\int_\varepsilon^1 t^{-1}\,dt = \ln(t)\Big|_\varepsilon^1 = -\ln(\varepsilon).$$

  Somit ergibt sich für $s < 1$ der Grenzwert $\frac{1}{1-s}$, und für $s \geq 1$ existiert das Integral über $(0, 1)$ nicht.

## Folgerung

*Das Integral* $\displaystyle\int_0^\infty \frac{1}{t^s}\,dt$ *existiert für kein* $s > 0$.





# Kapitel 10

## Vektorräume

# Lineare Algebra

## Definition (Vektorraum)

*Sei $\mathbb{K}$ ein Körper, z.B. $\mathbb{K} = \mathbb{Q}$, $\mathbb{R}$ oder $\mathbb{C}$.*

*Ein Vektorraum über $\mathbb{K}$ ist eine Menge $V$ zusammen mit einer Verknüpfung*

$$+\colon V \times V \to V, \quad (v, w) \mapsto v + w,$$

*genannt Addition, und mit einer Abbildung*

$$\mathbb{K} \times V \to V, \quad (\lambda, v) \mapsto \lambda \cdot v \ (=: \lambda v),$$

*genannt skalare Multiplikation, so dass*





1) $(V, +)$ eine kommutative Gruppe ist, d.h. es gilt:
   (i) $v + w = w + v$ für alle $v, w \in V$ (Kommutativgesetz),
   (ii) $(u + v) + w = u + (v + w)$ für alle $u, v, w \in V$ (Assoziativgesetz),
   (iii) es gibt ein Element $0 \in V$, so dass $v + 0 = v$ für alle $v$ und
   (iv) zu jedem $v \in V$ existiert $-v$, so dass $v + (-v) = 0$.
   (Wie im Fall von $(\mathbb{R}, +)$ zeigt man die Eindeutigkeit des neutralen Elements 0 und des additiven Inversen $-v$ von $v$.)
2) Für alle $\lambda, \mu \in \mathbb{K}$, $v, w \in V$ gilt:
   (i) $(\lambda\mu) \cdot v = \lambda \cdot (\mu \cdot v)$,
   (ii) $1 \cdot v = v$,
   (iii) $(\lambda + \mu) \cdot v = \lambda \cdot v + \mu \cdot v$,
   (iv) $\lambda \cdot (v + w) = \lambda \cdot v + \lambda \cdot w$.

## Definition

*Die Elemente von $\mathbb{K}$ heißen Skalare, die von $V$ Vektoren. Das neutrale Element $0 \in V$ heißt der Nullvektor von $V$.*

Man unterscheide immer $0 \in \mathbb{K}$ und den Nullvektor $0 \in V$, auch wenn wir die gleiche Notation verwenden.

## Beispiele von Vektorräumen

1) Der kartesische Raum

$\mathbb{K}^n = \{(x_1, \dots, x_n) | x_1, \dots, x_n \in \mathbb{K}\}$

mit der Addition

$$(x_1, \dots, x_n) + (y_1, \dots, y_n) := (x_1 + y_1, \dots, x_n + y_n)$$

und der skalaren Multiplikation

$$\lambda \cdot (x_1, \dots, x_n) := (\lambda x_1, \dots, \lambda x_n)$$

ist ein Vektorraum über $\mathbb{K}$.

Das neutrale Element 0 der kommutativen Gruppe $(\mathbb{K}^n, +)$ ist der Vektor $(0, \dots, 0)$ und das additive Inverse $-(x_1, \dots, x_n)$ des Vektors $(x_1, \dots, x_n)$ ist $(-x_1, \dots, -x_n)$.





2) Funktionenräume

Sei $X$ eine Menge und $\mathrm{Abb}(X, \mathbb{K})$ die Menge der Funktionen $f : X \to \mathbb{K}$.

$\mathrm{Abb}(X, \mathbb{K})$ ist mit der punktweise definierten Addition und skalaren Multiplikation

$$\begin{aligned} (f+g)(x) &= f(x) + g(x) \\ (\lambda f)(x) &= \lambda f(x) \quad (f, g \in \mathrm{Abb}(X, \mathbb{K}), \quad \lambda \in \mathbb{K}, \quad x \in X) \end{aligned}$$

ein Vektorraum über $\mathbb{K}$.

**Übungsaufgabe**

Sei $V$ ein Vektorraum über $\mathbb{K}$. Für alle $v \in V$, $\lambda \in \mathbb{K}$ gilt

(i) $\lambda v = 0 \iff \lambda = 0$ oder $v = 0$,

(ii) $-v = (-1) \cdot v$.

## Definition

*Ein Unterraum (genauer: ein Untervektorraum) eines Vektorraumes $V$ ist eine nicht leere Teilmenge $U \subseteq V$, so dass*

(i) $v + w \in U$ *für alle* $v, w \in U$ *und*

(ii) $\lambda v \in U$ *für alle* $\lambda \in \mathbb{K}, v \in U$.

**Bemerkung**

Wegen (i) und (ii) induzieren die Addition und die skalare Multiplikation in $V$ eine Addition

$$+\colon U \times U \to U$$

und eine skalare Multiplikation

$$\cdot\colon \mathbb{K} \times U \to U,$$

die $U$ zu einem Vektorraum machen, denn wegen $-u = (-1)u$ liegt auch das additive Inverse von $u \in U$ in $U$.



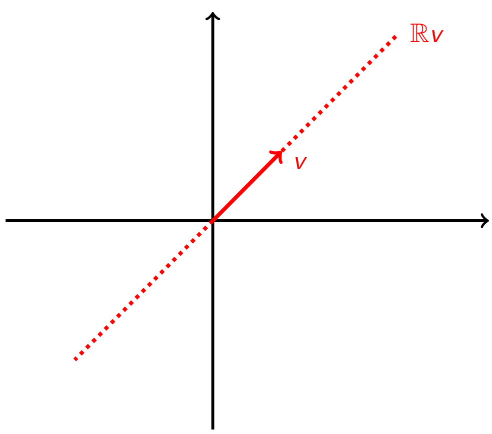


**Beispiele von Unterräumen**

Sei $V$ ein Vektorraum.

1) $\{0\}$ und $V$ selbst sind Untervektorräume von $V$.
2) Für jeden Vektor $v \in V$ ist $\mathbb{K}v = \{\lambda v | \lambda \in \mathbb{K}\} \subseteq V$ der kleinste Unterraum, der $v$ enthält.
   Wenn $v \neq 0$, so ist $\mathbb{K}v \neq \{0\}$ und heißt die von $v$ erzeugte Gerade.

   Für $V = \mathbb{R}^2$:



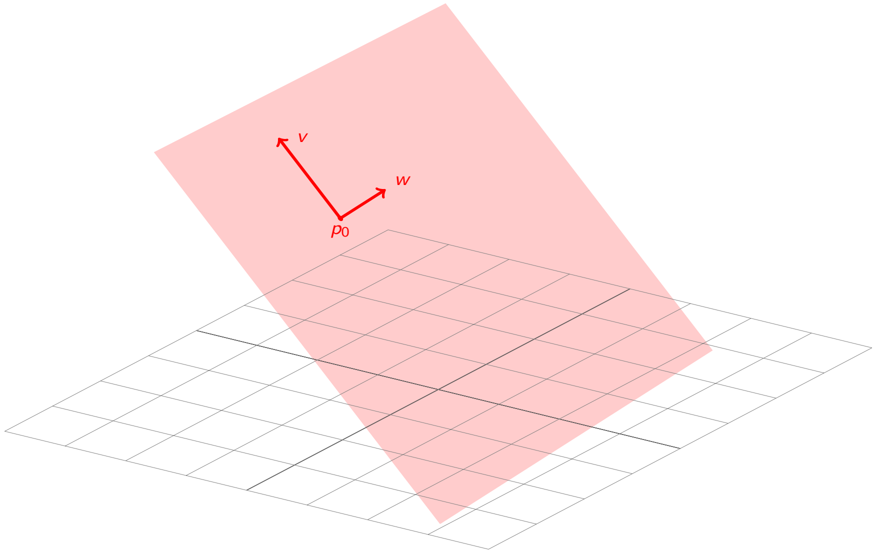

   Beachte: die (affine) Gerade $w + \mathbb{K}v$ ist im allgemeinen *kein* Unterraum des Vektorraums $V$.

## Beispiele von Unterräumen

3) Für jedes Paar von Vektoren $v, w$ ist

$$\mathbb{K}v + \mathbb{K}w = \{\lambda v + \mu w | \lambda, \mu \in \mathbb{K}\} \subseteq V$$

der kleinste Unterraum, der $v$ und $w$ enthält.
Wenn $v \neq 0$ und $w \notin \mathbb{K}v$, so heißt $\mathbb{K}v + \mathbb{K}w \supsetneq \mathbb{K}v \supsetneq \{0\}$ die von $v$ und $w$ erzeugte Ebene.








4) Sei $I \subseteq \mathbb{R}$ ein Intervall und $k \in \mathbb{N}$, $k \geq 1$. Die folgenden Mengen sind jeweils Unterräume von $\mathrm{Abb}(I, \mathbb{R})$:

$$\begin{aligned} C^k(I, \mathbb{R}) &:= \{f : I \to \mathbb{R} | f \quad k\text{-mal stetig differenzierbar}\} \\ &\subseteq Diff(I, \mathbb{R}) := \{f : I \to \mathbb{R} | f \quad \text{differenzierbar}\} \\ &\subseteq C(I, \mathbb{R}) := C^0(I, \mathbb{R}) := \{f : I \to \mathbb{R} | f \quad \text{stetig}\} \end{aligned}$$

5) Sei $V$ ein Vektorraum und $U_j \subseteq V$ durch $j \in J$ indizierte Unterräume, wobei $J$ eine beliebige Menge ist.
Der Durchschnitt $U = \cap_{j \in J} U_j \subseteq V$ ist ein Unterraum (ÜA).
Beispielsweise ist

$$C^\infty(I, \mathbb{R}) := \cap_{k \in \mathbb{N}} C^k(I, \mathbb{R})$$

ein Unterraum von $\mathrm{Abb}(I, \mathbb{R})$. Die Funktionen in $C^\infty(I, \mathbb{R})$ heißen unendlich oft differenzierbar oder glatt.

6) Ein Polynom in der Variablen $x$ mit Koeffizienten aus einem Körper $\mathbb{K}$ ist ein Ausdruck der Form

$$a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n,$$

wobei $a_0, \ldots, a_n \in \mathbb{K}$ und $n \in \mathbb{N}$.

Die Menge aller solchen Polynome wird mit $\mathbb{K}[x]$ bezeichnet und bildet in offensichtlicher Weise (ÜA) einen Vektorraum.

Jedes Polynom $a_0 + a_1 x + \cdots + a_n x^n \in \mathbb{K}[x]$ definiert durch Einsetzen eine Funktion

$$P \colon \mathbb{K} \to \mathbb{K}, \quad \lambda \mapsto P(\lambda) = a_0 + a_1 \lambda + \cdots + a_n \lambda^n.$$

Solche Funktionen $P$ heißen polynomial.

Wenn der Körper $\mathbb{K}$ unendlich ist (z.B. $\mathbb{K} = \mathbb{Q}, \mathbb{R}, \mathbb{C}$), so ist diese Zuordnung $a_0 + \cdots + a_n x^n \mapsto P$ injektiv (ÜA) und $\mathbb{K}[x]$ wird auf diese Weise zu einem Untervektorraum $\mathbb{K}[x] \subseteq \mathrm{Abb}(\mathbb{K}, \mathbb{K})$. Wir haben die Inklusion $\mathbb{R}[x] \subseteq C^\infty(\mathbb{R}, \mathbb{R}) \subseteq \mathrm{Abb}(\mathbb{R}, \mathbb{R})$.





### Definition (Lineare Unabhängigkeit)

(i) *Sei $V$ ein Vektorraum und $v_1, \ldots, v_r \in V$ endlich viele Vektoren.*

*Man sagt, dass $v \in V$ eine Linearkombination der Vektoren $v_1, \ldots, v_r$ ist, wenn es $\lambda_i \in \mathbb{K}$ gibt, so dass*

$$v = \sum_{i=1}^{r} \lambda_i v_i.$$

*Die (endliche) Folge der Vektoren $v_1, \ldots, v_r$ heißt linear unabhängig, wenn*

$$\sum_{i=1}^{r} \lambda_i v_i = 0 \implies \lambda_1 = \cdots = \lambda_r = 0,$$

*d.h. wenn der Nullvektor $0 \in V$ sich nur in trivialer Weise als Linearkombination der Vektoren $v_1, \ldots, v_r$ darstellen lässt.*

(ii) Sei nun $(v_j)_{j\in J}$ eine durch eine unendliche Menge $J$ indizierte Familie von Vektoren $v_j \in V$. (Der Begriff Familie verallgemeinert den der Folge oder endlichen Folge, die Indexmenge muss hier nicht abzählbar sein.)

Man sagt, dass $v \in V$ eine Linearkombination der Vektoren $(v_j)_{j\in J}$ ist, wenn es eine endliche Teilmenge $J_0 \subseteq J$ gibt, so dass $v$ eine Linearkombination der $(v_j)_{j\in J_0}$ ist.

Die Familie von Vektoren $(v_j)_{j\in J}$ heißt linear unabhängig, wenn die Vektoren $(v_j)_{j\in J_0}$ für jede endliche Teilmenge $J_0 \subseteq J$ linear unabhängig sind.

Man sagt auch, die Vektoren $(v_j)_{j\in J}$ seien linear unabhängig. Man beachte, dass dies eine Aussage über die Familie als Ganzes ist, nicht über das Verhältnis von Vektoren in der Familie. Daher sagen wir auch nicht, die Vektoren seien "voneinander" linear unabhängig.





**Beispiele**

1) Wir setzen

$$e_i := (0, \dots, 0, 1, 0, \dots, 0) \in \mathbb{K}^n.$$
$$\text{(i-te Stelle)}$$

Die Vektoren $(e_1, \dots, e_n)$ sind linear unabhängig. In der Tat:

$$0 = \sum_{i=1}^{n} \lambda_i e_i = (\lambda_1, \dots, \lambda_n) \Longrightarrow \lambda_1 = \dots = \lambda_n = 0.$$

2) Die Monome $(1, x, x^2, x^3, \dots)$ sind linear unabhängig in $\mathbb{K}[x]$.

3) Eine Familie $(v)$ von Vektoren, die aus nur einem Vektor $v \in V$ besteht, ist genau dann linear unabhängig, wenn $v \neq 0$ gilt.

## Definition (Lineare Hülle)

*Sei $V$ ein Vektorraum und $(v_j)_{j\in J}$ eine Familie von Vektoren.*
*Die lineare Hülle der Familie $(v_j)_{j\in J}$ ist der Unterraum*

$$\operatorname{span}\{v_j|j\in J\} := \{v|v \text{ ist eine Linearkombination der } (v_j)_{j\in J}\}.$$

**Beispiel:** Sei $v_1 = (1,1,0)$ und $v_2 = (1,-1,0) \in \mathbb{R}^3$. Dann ist

$$\begin{aligned}
\operatorname{span}\{v_1, v_2\} &= \{\lambda(1,1,0) + \mu(1,-1,0) \mid \lambda, \mu \in \mathbb{R}\} \\
&= \{(\lambda,\lambda,0) + (\mu,-\mu,0) \mid \lambda, \mu \in \mathbb{R}\} \\
&= \{(\lambda+\mu, \lambda-\mu, 0) \mid \lambda, \mu \in \mathbb{R}\} \\
&= \{(x,y,0) \mid x, y \in \mathbb{R}\} \\
&= \operatorname{span}\{e_1, e_2\}
\end{aligned}$$

für $e_1 = (1,0,0)$ und $e_2 = (0,1,0) \in \mathbb{R}^3$.





## Satz

*Die Familie $(v_j)_{j\in J}$ von Vektoren $v_j \in V$ sei linear unabhängig.*
*Dann hat jeder Vektor $v \in \operatorname{span}\{v_j|j\in J\}$ eine eindeutige Darstellung*

$$v = \sum_{j\in J} \lambda_j v_j, \quad \lambda_j \in \mathbb{K},$$

*wobei $\lambda_j = 0$ für fast alle $j \in J$, d.h. für alle bis auf endlich viele $j \in J$.*

**Beweis.** Sei $v$ in der linearen Hülle vorgegeben.

Nach Definition der linearen Hülle gibt es eine endliche Teilmenge $J_0 \subseteq J$ und eine Darstellung

$$v = \sum_{j\in J_0} \lambda_j v_j.$$

Sei $v = \sum_{j\in J_0'} \lambda_j' v_j$ eine weitere solche Darstellung mit endlichem $J_0' \subseteq J$.

**Weiter im Beweis:**

Dann gilt

$$\begin{aligned} 0 &= v - v = \sum_{j \in J_0} \lambda_j v_j - \sum_{j \in J_0'} \lambda_j' v_j \\ &= \sum_{j \in J_0 \cap J_0'} (\lambda_j - \lambda_j') v_j + \sum_{j \in J_0 \setminus J_0 \cap J_0'} \lambda_j v_j - \sum_{j \in J_0' \setminus J_0 \cap J_0'} \lambda_j' v_j. \end{aligned}$$

Wegen der linearen Unabhängigkeit der $v_j$, $j \in J$, und der Endlichkeit der Menge $J_0 \cup J_0'$ folgt

$$\begin{aligned} \lambda_j - \lambda_j' &= 0, \quad \text{wenn} \quad j \in J_0 \cap J_0' \\ \lambda_j &= 0, \quad \text{wenn} \quad j \in J_0 \setminus J_0 \cap J_0' \\ \lambda_j' &= 0, \quad \text{wenn} \quad j \in J_0' \setminus J_0 \cap J_0'. \end{aligned}$$

Also stimmen die beiden Darstellungen
$v = \sum_{j \in J_0} \lambda_j v_j = \sum_{j \in J_0 \cap J_0'} \lambda_j v_j$ und $v = \sum_{j \in J_0'} \lambda_j' v_j = \sum_{j \in J_0 \cap J_0'} \lambda_j' v_j$
überein. □





**Definition (Erzeugendensysteme, Basen)**

*Sei $V$ ein Vektorraum. Eine Familie von Vektoren $(v_j)_{j \in J}$ heißt Erzeugendensystem von $V$, falls $V = \mathrm{span}\{v_j | j \in J\}$.*

*Eine Basis von $V$ ist ein linear unabhängiges Erzeugendensystem.*

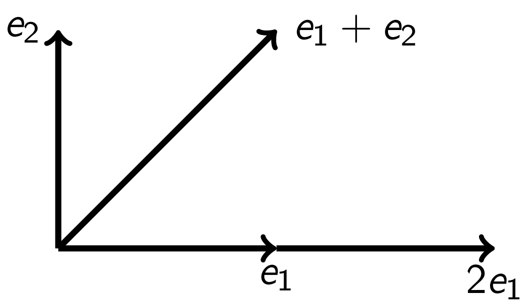
**Beispiele**

1) $(e_1, \ldots, e_n)$ ist eine Basis von $\mathbb{K}^n$, die sogenannte kanonische Basis.
   Warnung: Im allgemeinen hat ein Vektorraum keine *kanonische* Basis!
2) $(1, x, x^2, \ldots)$ ist eine Basis von $\mathbb{K}[x]$.
3) Für jeden Vektorraum $V$ ist die zu Grunde liegende Menge $V$ ein (sehr großes) Erzeugendensystem.

## Ein Beispiel

3) $(e_1, e_2, e_1 + e_2, 2e_1)$ ist ein Erzeugendensytem von $\mathbb{K}^2$, aus dem wir folgende Basen auswählen können: $(e_1, e_2)$, $(e_1, e_1 + e_2)$ und $(e_2, e_1 + e_2)$. Falls in $\mathbb{K}$ die Gleichung $0 \neq 2$ $(:= 1 + 1)$ gilt, so sind $(e_2, 2e_1)$ und $(e_1 + e_2, 2e_1)$ ebenfalls Basen.

Bild für $\mathbb{K} = \mathbb{R}$:






Vektorräume | Erzeugendensysteme, Basen

### Satz (Charakterisierung von Basen)

*Sei $V \neq 0$ ein Vektorraum. Eine Familie $(v_j)_{j\in J}$ von Vektoren von $V$ ist genau dann eine Basis, wenn sie eine der folgenden Bedingungen erfüllt.*

(i) *$(v_j)_{j\in J}$ ist ein minimales Erzeugendensystem, d.h. ist $J_0 \subseteq J$ eine Teilmenge, für die auch die Familie $(v_j)_{j\in J_0}$ ein Erzeugendensystem ist, so ist $J_0 = J$.*

(ii) *$(v_j)_{j\in J}$ ist eine maximale linear unabhängige Familie, d.h. ist $J_0 \supseteq J$ eine Obermenge von $J$, für die auch die Familie $(v_j)_{j\in J_0}$ linear unabhängig ist, so ist $J_0 = J$.*

(iii) *Jeder Vektor $v \in V$ besitzt genau eine Darstellung*

$$v = \sum_{j\in J} \lambda_j v_j,$$

*wobei $\lambda_j = 0$ für fast alle $j \in J$.*

**Beweis.**
Sei (0) die Aussage, dass $(v_j)_{j\in J}$ eine Basis bildet. Wir zeigen $(0) \Rightarrow (iii) \Rightarrow (i) \Rightarrow (ii) \Rightarrow (0)$.

$(0) \Rightarrow (iii)$ Wir wissen bereits, dass jeder Vektor aus $\mathrm{span}\{v_j | j \in J\}$ eine eindeutige Darstellung als (endliche) Linearkombination der linear unabhängigen Vektoren $(v_j)_{j\in J}$ hat. Da $(v_j)_{j\in J}$ auch ein Erzeugendensystem ist, gilt das für jeden Vektor aus $V = \mathrm{span}\{v_j | j \in J\}$.

$(iii) \Rightarrow (i)$ Aus (iii) folgt offensichtlich, dass $(v_j)_{j\in J}$ ein Erzeugendensystem ist.

Um die Minimalität zu zeigen, nehmen wir an, dass die echte Teilfamilie $(v_j)_{j\in J_0}$, $J_0 \subsetneq J$, schon ein Erzeugendensystem ist.

Wählen wir $i \in J \setminus J_0$, dann gibt es eine Darstellung (*) $v_i = \sum_{j\in J_0} \lambda_j v_j$, aber auch $v_i = v_i$ ist eine Darstellung. Dies widerspricht der Eindeutigkeit in (iii). Also muss $J = J_0$ gelten.





**Weiter im Beweis:** $(i) \Rightarrow (ii)$

- Wir zeigen zuerst, dass aus (i) die lineare Unabhängigkeit von $(v_j)_{j\in J}$ folgt.
  Sei $\sum_{j\in J_0} \lambda_j v_j = 0$, wobei $J_0 \subseteq J$ endlich ist.
  Wenn für ein $j_0 \in J_0$ der Koeffizient $\lambda_{j_0} \neq 0$ wäre, so wäre $v_{j_0} = -\sum_{j\in J_0\setminus\{j_0\}} \frac{\lambda_j}{\lambda_{j_0}} v_j$ und somit $(v_j)_{j\in J\setminus\{j_0\}}$ ein Erzeugendensystem, im Widerspruch zur angenommenen Minimalität in (i).
  Also $\lambda_{j_0} = 0$ für alle $j_0 \in J_0$. Das zeigt die lineare Unabhängigkeit von $(v_j)_{j\in J}$.
- Als Nächstes zeigen wir die Maximalität der linear unabhängigen Familie $(v_j)_{j\in J}$.
  Sei also $(v_j)_{j\in J_1}$, $J_1 \supsetneq J$, eine Familie, die die $(v_j)_{j\in J}$ enthält und $j_1 \in J_1 \setminus J$.
  Da $(v_j)_{j\in J}$ ein Erzeugendensystem ist, gibt es eine Darstellung $v_{j_1} = \sum_{j\in J} \lambda_j v_j$.
  Das zeigt, dass $(v_j)_{j\in J_1}$ linear abhängig ist. Somit ist die linear unabhängige Familie $(v_j)_{j\in J}$ maximal.

**Ende des Beweises**

$(ii) \Rightarrow (0)$ Sei $(v_j)_{j \in J}$ eine maximale linear unabhängige Familie. Wir zeigen, dass $(v_j)_{j \in J}$ ein Erzeugendensystem und somit eine Basis ist.
Wäre es kein Erzeugendensystem, so gäbe es $v \in V \setminus \mathrm{span}\{v_j | j \in J\}$.
Die Familie $(v, (v_j)_{j \in J})$ wäre dann linear unabhängig: sei $0 = \mu v + \sum \lambda_j v_j$ eine Darstellung des Nullvektors. Ist $\mu = 0$, so folgt $\lambda_j = 0$ für alle $j$ aus der linearen Unabhängigkeit von $(v_j)_{j \in J}$. Ist $\mu \neq 0$, so erhalten wir $v = -\sum (\lambda_j / \mu) v_j$, im Widerspruch zur Annahme über $v$.
Damit haben wir einen Widerspruch zur Maximalität von $(v_j)_{j \in J}$. □

**Folgerung (Basisauswahlsatz)**

*Jedes endliche Erzeugendensystem eines Vektorraums $V$ enthält eine Basis.*

**Beweis.** Jedes endliche Erzeugendensystem enthält ein minimales Erzeugendensystem. □





**Lemma (Austauschlemma)**

*Sei $(v_1, \ldots, v_n)$ eine Basis von $V$ und*

$$w = \sum_{i=1}^{n} \lambda_i v_i$$

*eine Linearkombination.*

*Wenn $\lambda_k \neq 0$ gilt, so ist auch $(v_1, \ldots, v_{k-1}, w, v_{k+1}, \ldots, v_n)$ eine Basis von $V$.*

**Beweis.**
Übungsaufgabe. □

### Satz (Steinitzscher Austauschsatz)

*Sei $V$ ein Vektorraum, $(v_1, \dots, v_n)$ eine Basis und die Familie $(w_1, \dots, w_r)$ linear unabhängig.*

*Dann gilt $n \geq r$ und es gibt eine Bijektion $\varphi\colon \{1, \dots, n\} \to \{1, \dots, n\}$ (eine Permutation von $\{1, \dots, n\}$), so dass die Familie $(w_1, \dots, w_r, v_{\varphi(r+1)}, \dots, v_{\varphi(n)})$ eine Basis ist.*

**Beweis.**

Beweis durch Induktion nach $r$:

Für $r = 0$ ist nichts zu zeigen. Wir führen den Induktionsschritt $r \to r + 1$ aus.

Sei also die Familie $(w_1, \dots, w_{r+1})$ linear unabhängig und $(v_1, \dots, v_n)$ eine Basis.

Nach Induktionsvoraussetzung gilt $r \leq n$ und es gibt eine Permutation $\varphi\colon \{1, \dots, n\} \to \{1, \dots, n\}$, so dass $(w_1, \dots, w_r, v_{\varphi(r+1)}, \dots, v_{\varphi(n)})$ eine Basis ist.





**Weiter im Beweis:**

Wir können daher schreiben

$$w_{r+1} = \sum_{i=1}^{r} \lambda_i w_i + \sum_{j=r+1}^{n} \lambda_j v_{\varphi(j)}.$$

Da die Familie $(w_1, \dots, w_{r+1})$ linear unabhängig ist, existiert $j_0 \in \{r + 1, \dots, n\}$ mit $\lambda_{j_0} \neq 0$. Insbesondere ist $r + 1 \leq n$.

Wir können (ggf. nach Abänderung von $\varphi$) annehmen, dass $j_0 = r + 1$.

Nach dem Austauschlemma können wir dann in der Basis $(w_1, \dots, w_r, v_{\varphi(r+1)}, \dots, v_{\varphi(n)})$ den Vektor $v_{\varphi(j_0)} = v_{\varphi(r+1)}$ durch $w_{r+1}$ ersetzen und erhalten die Basis $(w_1, \dots, w_{r+1}, v_{\varphi(r+2)}, \dots, v_{\varphi(n)})$. $\square$

### Theorem

*Sei $V$ ein Vektorraum.*

(i) *Wenn $V$ eine endliche Basis besitzt, dann ist jede Basis von $V$ endlich.*

(ii) *Alle endlichen Basen von $V$ haben die gleiche Anzahl von Elementen.*

**Beweis.**

(i) Sei $(v_1, \ldots, v_n)$ eine Basis. Aus dem Austauschsatz folgt, dass jede linear unabhängige Familie höchstens $n$ Elemente hat.

(ii) Sei $(w_1, \ldots, w_r)$ eine zweite Basis. Aus dem Austauschsatz folgt, wie gesagt, $r \leq n$. Nach Vertauschen der Rollen der beiden Basen folgt ebenso $n \leq r$.

$\square$





### Definition (Dimension)

*Sei $V$ ein Vektorraum über $\mathbb{K}$.*

*Die Dimension von $V$ ist die Zahl*

$$\dim V = \dim_{\mathbb{K}} V := \begin{cases} 0, & \text{falls } V = \{0\} \\ n, & \text{falls } V \text{ eine endliche Basis } (v_1, \ldots, v_n) \text{ hat} \\ \infty & \text{sonst} \end{cases}$$

### Bemerkung

Die leere Familie ist die Basis des Nullvektorraums $V = \{0\}$.

**Beispiele/Übungsaufgaben**

1) $\dim_{\mathbb{K}} \mathbb{K}^n = n$.
2) Jeder komplexe Vektorraum $V$ kann als reeller Vektorraum aufgefasst werden und $\dim_{\mathbb{R}} V = 2 \dim_{\mathbb{C}} V$: ist $(b_k)$ eine Basis über $\mathbb{C}$, so ist die Familie $(b_k, ib_k)$ eine Basis über $\mathbb{R}$. Insbesondere gilt $\dim_{\mathbb{R}} \mathbb{C}^n = 2 \dim_{\mathbb{C}} \mathbb{C}^n = 2n$.
3) $\dim_{\mathbb{Q}} \mathbb{R} = \infty$ (Hinweis: Das folgt aus der Überabzählbarkeit von $\mathbb{R}$).
4) $\dim_{\mathbb{K}} \mathrm{Abb}(X, \mathbb{K}) = \mathrm{card}(X)$, wobei

$$\mathrm{card}(X) := \begin{cases} n, & \text{falls } X \text{ aus n Elementen besteht} \\ \infty, & \text{falls } X \text{ unendlich ist.} \end{cases}$$

5) Für den Vektorraum der Polynome gilt $\dim_{\mathbb{K}} \mathbb{K}[x] = \infty$.
6) Sei $U \subseteq V$ ein Unterraum. Dann gilt $\dim U \leq \dim V$. Falls $V$ endlichdimensional ist, so gilt $\dim U = \dim V$ genau dann, wenn $U = V$.





**Folgerung**

*Sei $V$ ein Vektorraum der Dimension $n \in \mathbb{N}$.*

(i) *Jede linear unabhängige Familie von Vektoren von $V$ hat* höchstens *$n$ Elemente.*

(ii) *Eine linear unabhängige Familie von Vektoren von $V$ ist genau dann eine Basis, wenn sie $n$ Elemente hat.*

(iii) *Jedes Erzeugendensystem von $V$ hat* mindestens *$n$ Elemente.*

(iv) *Ein Erzeugendensystem von $V$ ist genau dann eine Basis, wenn es $n$ Elemente hat.*

**Beweis.**

(i-ii) folgt aus dem Austauschsatz.

(iii) folgt daraus, dass man aus jedem endlichen Erzeugendensystem eine Basis auswählen kann.

(iv) Ein Erzeugendensystem mit $n$ Elementen ist wegen (iii) minimal und somit eine Basis. □

# Kapitel 11

Lineare Abbildungen





# Lineare Abbildungen

## Definition

*$V$, $W$ seien Vektorräume über demselben Körper $\mathbb{K}$.*

*Eine Abbildung $F\colon V \to W$ heißt linear (genauer: $\mathbb{K}$-linear), wenn*

$$\begin{aligned} F(v+w) &= F(v)+F(w) \quad \text{und} \\ F(\lambda v) &= \lambda F(v) \end{aligned}$$

*für alle $v, w \in V$, $\lambda \in \mathbb{K}$.*

*Eine bijektive lineare Abbildung $F\colon V \to W$ heißt auch Isomorphismus von Vektorräumen.*

*Die Vektorräume $V$ und $W$ heißen isomorph, wenn es einen Isomorphismus $F\colon V \to W$ von Vektorräumen gibt.*

Grundlegende Eigenschaften:

- Für jede Menge $X$ und jeden Vektorraum $V$ bilden die Abbildungen $X \to V$ einen Vektorraum, den wir mit $\mathrm{Abb}(X, V)$ bezeichnen. Die Operationen sind $(\lambda \cdot f)(x) := \lambda \cdot f(x)$ und $(f + g)(x) := f(x) + g(x)$.

### Definition (Raum der linearen Abbildungen)

*Seien $V$, $W$ Vektorräume über $\mathbb{K}$. Die Teilmenge $L(V, W) \subseteq \mathrm{Abb}(V, W)$ aller linearen Abbildungen $V \to W$ ist ein Unterraum von $\mathrm{Abb}(V, W)$, d.h. für $\lambda \in \mathbb{K}$ und $f, g \in L(V, W)$ sind $\lambda f$ und $f + g$ auch in $L(V, W)$.*

- Eine lineare Abbildung ist bestimmt durch ihre Werte auf Basisvektoren: sind $F, G \in L(V, W)$ und $(v_1, \ldots, v_n)$ ist eine Basis von $V$ und gilt $F(v_i) = G(v_i)$ für $i = 1, \ldots, n$, dann ist $F = G$. Beweis: jedes $v \in V$ hat eine eindeutige Darstellung $v = \sum_{i=1}^{n} \lambda_i v_i$. Da $F$ und $G$ linear sind, gilt

$$F(v) = F\left(\sum_{i=1}^{n} \lambda_i v_i\right) = \sum_{i=1}^{n} \lambda_i F(v_i) = \sum_{i=1}^{n} \lambda_i G(v_i) = G(v).$$





# Beispiele

- Die identische Abbildung eines jeden Vektorraums ist linear.
- Spiegelungen im $\mathbb{R}^2$ an Ursprungsgeraden und allgemeiner Spiegelungen im $\mathbb{R}^n$ an Hyperebenen ($n-1$-dimensionalen Unterräumen) sind $\mathbb{R}$-linear.
- Die Abbildung, die alle Elemente eines Vektorraums auf den Nullvektor abbildet, ist linear.
- $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^2$ ist natürlich nicht linear. Was ist mit $x \mapsto ax + b$ mit $a, b \in \mathbb{R}$?
- $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x_1, x_2) = 3x_2 - x_1$ ist $\mathbb{R}$-linear.
- Die Ableitung definiert eine $\mathbb{R}$-lineare Abbildung $A : C^1(\mathbb{R}, \mathbb{R}) \to C^0(\mathbb{R}, \mathbb{R})$, $A(f) = f'$.
- Für jedes $x_0 \in [a, b]$ definiert das bestimmte Integral eine $\mathbb{R}$-lineare Abbildung $I_{x_0} : C^0([a, b], \mathbb{R}) \to C^1([a, b], \mathbb{R})$, wobei

$$I_{x_0}(f)(x) = \int_{x_0}^{x} f(t)\, dt.$$

# Definition von linearen Abbildungen auf einer Basis

**Satz**

*Seien $V, W$ Vektorräume über $\mathbb{K}$. Es sei $\dim V < \infty$ und $(v_1, \dots, v_n)$ eine Basis von $V$. Für jedes n-Tupel $(w_1, \dots w_n)$ von Vektoren in $W$ gibt es genau eine lineare Abbildung $F\colon V \to W$ mit $F(v_i) = w_i$.*

**Beweis.** Jeder Vektor $v \in V$ lässt sich eindeutig in der Form $v = \sum_{i=1}^{n} \lambda_i v_i$ schreiben. Wegen der Linearität muss $F$ auf $v$ den Wert

$$F(v) = \sum_{i=1}^{n} \lambda_i F(v_i) = \sum_{i=1}^{n} \lambda_i w_i$$

annnehmen. Man rechnet leicht nach, dass dies eine lineare Abbildung definiert. $\square$



Lineare Abbildungen | Definition und Eigenschaften

## Weitere wichtige Eigenschaften linearer Abbildungen (ÜA)

1. Seien $F\colon U \to V$ linear und $G\colon V \to W$ linear. Dann ist auch die Komposition $G \circ F\colon U \to W$ linear.
2. Sei $F\colon V \to W$ linear. Für jeden Unterraum $V' \subseteq V$ ist das Bild $F(V') \subseteq W$ ein Unterraum. Ebenso ist das Urbild $F^{-1}(W') \subseteq V$ ein Unterraum für jeden Unterraum $W' \subseteq W$.
3. Insbesondere sind für jede lineare Abbildung $F$ ihr Bild $\mathrm{im}\, F := F(V) \subseteq W$ und ihr Kern $\ker F := F^{-1}(0) = \{v \in V | F(v) = 0\} \subseteq V$ jeweils Unterräume.
4. $F$ ist genau dann injektiv, wenn $\ker F = \{0\}$.
5. Eine lineare Abbildung $F\colon V \to W$ ist ein Isomorphismus genau dann, wenn $\mathrm{im}\, F = W$ und $\ker F = \{0\}$.
6. Sei $F\colon V \to W$ ein Isomorphismus. Dann ist $F^{-1}\colon W \to V$ linear.

## Satz

*Es seien $V$, $W$ Vektorräume, $F\colon V \to W$ linear und $(v_1, \dots, v_n)$ eine Basis von $V$. Dann gilt*

(i) *$F$ ist surjektiv $\iff$ $(F(v_i))_i$ ist ein Erzeugendensystem von $W$.*

(ii) *$F$ ist injektiv $\iff$ $(F(v_i))_i$ ist linear unabhängig.*

(iii) *$F$ ist ein Isomorphismus $\iff$ $(F(v_i))_i$ ist eine Basis von $W$.*

**Beweis:** Übungsaufgabe □

## Folgerung

*Zwei endlich-dimensionale Vektorräume $V$ und $W$ sind isomorph $\iff$* $\dim(V) = \dim(W)$.

**Beweis.** Sei $(v_1, \dots, v_n)$ eine Basis von $V$.

'$\Rightarrow$' Wegen (iii) ist für jeden Isomorphismus $(F(v_i))_i$ eine Basis.

'$\Leftarrow$' Sei umgekehrt $\dim W = \dim V = n$. Dann existiert eine Basis $(w_1, \dots, w_n)$ von $W$ mit $n$ Elementen. Die lineare Abbildung $F\colon V \to W$ definiert durch $F(v_i) = w_i$ ist ein Isomorphismus. □





## Satz

*Sei $V$ ein endlich-dimensionaler Vektorraum über $\mathbb{K}$ und $(v_1, \dots, v_n)$ eine Basis von $V$. Dann definiert die Zuordnung*

$$\begin{array}{rrcl} \varphi: & L(V,\mathbb{K}) & \to & \mathbb{K}^n \\ & F & \mapsto & (F(v_1), F(v_2), \dots, F(v_n)) \end{array}$$

*einen Isomorphismus. Insbesondere gilt* $\dim(V) = \dim(L(V,\mathbb{K}))$.

**Beweis.**

- $\varphi$ ist offensichtlich linear.
- $\varphi : F \mapsto (F(v_i))_i$ ist injektiv, denn aus $0 = \varphi(F)$ folgt $F(v_1) = F(v_2) = \dots = F(v_n) = 0$ und damit wegen der Linearität $F(v) = 0$ für jeden Vektor $v = \sum_{i=1}^n \lambda_i v_i \in V$, also $F = 0$.
- $\varphi$ ist auch surjektiv: Sei $a = (a_1, \dots, a_n) \in \mathbb{K}^n$ beliebig. Wenn wir die Abbildung $F$ durch die Werte $F(v_i) := a_i$ auf den Basisvektoren $(v_i)$ definieren, so ist $F$ linear mit $\varphi(F) = a$.

□

# Der Dualraum

## Definition

Sei $V$ ein Vektorraum über dem Körper $\mathbb{K}$. Der Vektorraum

$$V^* := L(V, \mathbb{K})$$

heißt Dualraum zu $V$ oder auch Raum der Linearformen.

## Duale Basen

- Ist $(v_1, \dots, v_n)$ eine Basis von $V$, so bilden die Linearformen $v_i^* \in L(V, \mathbb{K}) = V^*$ mit

$$v_i^*(v_j) = \begin{cases} 1, & \text{falls } i = j, \\ 0, & \text{falls } i \neq j \end{cases}$$

  eine Basis von $V^*$ Diese Basis heißt die zu $(v_1, \dots, v_n)$ duale Basis.
- Ist $V = \mathbb{K}^n$, so heißt die zur kanonischen Basis $(e_1, \dots, e_n)$ des $\mathbb{K}^n$ duale Basis die kanonische Basis des $(\mathbb{K}^n)^*$. Im Allgemeinen gibt es aber keine kanonische Isomorphie $V \cong V^*$.





# Der Vektorraum der $(m \times n)$-Matrizen

Die Menge der Matrizen mit $m$ Zeilen und $n$ Spalten, definiert durch

$$\mathrm{Mat}(m, n, \mathbb{K}) := \left\{ A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \middle| \; a_{ij} \in \mathbb{K} \right\}$$

bildet mit den Operationen

$$\begin{aligned} \lambda(a_{ij})_{\substack{i=1,\dots m \\ j=1,\dots n}} &:= (\lambda a_{ij})_{\substack{i=1,\dots m \\ j=1,\dots n}} \quad \text{und} \\ (a_{ij})_{\substack{i=1,\dots m \\ j=1,\dots n}} + (b_{ij})_{\substack{i=1,\dots m \\ j=1,\dots n}} &:= (a_{ij} + b_{ij})_{\substack{i=1,\dots m \\ j=1,\dots n}} \end{aligned}$$

einen Vektorraum, der offensichtlich isomorph ist zum Vektorraum $\mathbb{K}^{nm}$. Man beachte, dass nur die Addition von Matrizen gleicher Gestalt definiert ist!

Vereinbarung: Die Elemente des $\mathbb{K}^n$ schreiben wir (ab jetzt) in der Regel als Spaltenvektoren.

### Satz

*Der $\mathbb{K}^n$ sei versehen mit der kanonischen Basis $(e_j)_{j=1,\dots n}$. Der Vektorraum $L(\mathbb{K}^n, \mathbb{K}^m)$ ist in kanonischer Weise isomorph zu $\mathrm{Mat}(m, n, \mathbb{K})$.*
*Wir ordnen dabei einer linearen Abbildung $F \in L(\mathbb{K}^n, \mathbb{K}^m)$ eine $m \times n$-Matrix zu durch*

$$\varphi \colon F \mapsto (F(e_1) \ \dots \ F(e_n)),$$

*wobei die $F(e_j) \in \mathbb{K}^m$ als Spaltenvektoren zu verstehen sind.*

**Beweis.** Die Abbildung $\varphi$ ist linear und injektiv, denn: $F(e_j) = 0$ für alle $j = 1, \dots, n$ impliziert $F = 0$.





**Weiter im Beweis:**
Die lineare Abbildung $\varphi$ ist aber auch surjektiv:
Sei $A = (a_{ij})_{\substack{i=1,\dots m \\ j=1,\dots n}}$ eine beliebige Matrix.
Dazu definiert man $F \in L(\mathbb{K}^n, \mathbb{K}^m)$ mittels

$$F(e_j) := \sum_{i=1}^{m} a_{ij} e_i.$$

Damit ist die $j$-te Spalte von $\varphi(F)$,

$$a_j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix} = \sum_{i=1}^{m} a_{ij} e_i \in \mathbb{K}^m,$$

das Bild $F(e_j)$ des $j$-ten Basisvektors $e_j$ der kanonischen Basis des $\mathbb{K}^n$.
Somit ist $\varphi(F) = A$. $\square$

# Produkt einer Matrix mit einem Vektor

## Definition

Das Produkt einer $(m \times n)$-Matrix $A = (a_{ij})_{\substack{i=1,\dots m \\ j=1,\dots n}}$ mit einem

$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{K}^n$ ist der Vektor in $\mathbb{K}^m$ gegeben durch

$$Ax = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} := \begin{pmatrix} y_1 := \sum_{j=1}^n a_{1j} x_j \\ \vdots \\ y_m := \sum_{j=1}^n a_{mj} x_j \end{pmatrix} \in \mathbb{K}^m$$

Man beachte, dass eine $m \times n$-Matrix nur mit einem Vektor im $\mathbb{K}^n$, nicht aber mit Vektoren in $\mathbb{K}^r$ mit $r \neq n$ multipliziert werden kann.





Sei nun $F \in L(\mathbb{K}^n, \mathbb{K}^m)$ eine lineare Abbildung und $A = (a_{ij})_{\substack{i=1,\dots m \\ j=1,\dots n}}$ die kanonisch zugeordnete Matrix, d.h. die $j$-te Spalte der Matrix $A$ ist genau das Bild $F(e_j)$ des $j$-ten Basisvektors $e_j$ der kanonischen Basis des $\mathbb{K}^n$.

Dann gilt für $x = \sum\limits_{j=1}^n x_j e_j$

$$\begin{aligned} F(x) &= F\Big(\sum_{j=1}^n x_j e_j\Big) = \sum_{j=1}^n x_j F(e_j) \\ &= \sum_{j=1}^n x_j \sum_{i=1}^m a_{ij} e_i = \sum_{i=1}^m \left( \sum_{j=1}^n a_{ij} x_j \right) e_i = Ax. \end{aligned}$$

D.h. das Bild von $x$ unter $F$ ist gleich dem Produkt $Ax$.

# Produkt zweier Matrizen

## Definition

*Seien nun* $A = (a_{ij})_{i,j} \in \mathrm{Mat}(m,n,\mathbb{K})$ *und* $B = (b_{jk})_{j,k} \in \mathrm{Mat}(n,p,\mathbb{K})$. *Das* **Produkt** $C = AB$ *von* $A$ *und* $B$ *ist die Matrix* $C = (c_{ik})_{\substack{i=1\dots m\\ k=1\dots p}} \in \mathrm{Mat}(m,p,\mathbb{K})$ *mit* $c_{ik} = \sum_{j=1}^{n} a_{ij} b_{jk}$.

Man beachte, dass eine $m \times n$-Matrix mit einer $p \times q$-Matrix nur multipliziert werden kann, wenn $n = p$ gilt. Dann ist das Produkt eine $m \times q$-Matrix.





# Produkt zweier Matrizen (Fortsetzung)

## Definition

*Seien nun* $A = (a_{ij})_{i,j} \in \mathrm{Mat}(m,n,\mathbb{K})$ *und* $B = (b_{jk})_{j,k} \in \mathrm{Mat}(n,p,\mathbb{K})$. *Das* **Produkt** $C = AB$ *von* $A$ *und* $B$ *ist die Matrix* $C = (c_{ik})_{\substack{i=1\dots m\\ k=1\dots p}} \in \mathrm{Mat}(m,p,\mathbb{K})$ *mit* $c_{ik} = \sum_{j=1}^{n} a_{ij} b_{jk}$.

Diese Matrix entspricht der Matrix der Verkettung der zu $B$ und $A$ gehörenden linearen Abbildungen $\mathbb{K}^p \to \mathbb{K}^n$ und $\mathbb{K}^n \to \mathbb{K}^m$, denn für $x = \sum_{k=1}^{p} x_k e_k \in \mathbb{K}^p$ ist

$$\begin{aligned} A(Bx) &= A(\sum_{j=1}^{n}\sum_{k=1}^{p} b_{jk}x_k e_j) = \sum_{k=1}^{p}\sum_{j=1}^{n} b_{jk}x_k A(e_j) \\ &= \sum_{k=1}^{p}\sum_{j=1}^{n} b_{jk}x_k \sum_{i=1}^{m} a_{ij}e_i = \sum_{i=1}^{m}\sum_{k=1}^{p}(\sum_{j=1}^{n} a_{ij}b_{jk})x_k e_i \end{aligned}$$

Also $A(Bx) = Cx$ für $C = (c_{ik})_{i,k}$ mit $c_{ik} = \sum_{j=1}^{n} a_{ij} b_{jk}$.

# Spur einer Matrix

## Übung

Durch

$$\mathrm{Tr}\colon \mathrm{Mat}(n,n,\mathbb{K}) \to \mathbb{K},\ A \mapsto \mathrm{Tr}(A) = \sum_{k=1}^{n} a_{k\,k}$$

wird eine lineare Abbildung definiert. Sie heißt die Spur.
Sind $A = (a_{i\,j})_{i\in\{1,\ldots,n\},j\in\{1,\ldots,m\}}$ und $B = (b_{i\,j})_{i\in\{1,\ldots,m\},j\in\{1,\ldots,n\}}$ nicht notwendigerweise quadratische Matrizen, so gilt stets

$$\mathrm{Tr}(AB) = \mathrm{Tr}(BA).$$





# Darstellende Matrix einer linearen Abbildung

$B = (v_1, \ldots, v_n)$ bzw. $B' = (w_1, \ldots, w_m)$ seien geordnete Basen der endlich-dimensionalen $\mathbb{K}$-Vektorräume $V$ bzw. $W$; $\phi_B\colon \mathbb{K}^n \to V$ und $\phi_{B'}\colon \mathbb{K}^m \to W$ seien die zugehörigen Isomorphismen definiert durch $\phi_B(e_i) = v_i$, $\phi_{B'}(e_j) = w_j$, wobei $i = 1, \ldots, n$, $j = 1, \ldots, m$ und $(e_i)_i$ die entsprechende kanonische Basis ist.

## Definition

*Sei $F \in L(V, W)$. Die Matrix $A \in \mathrm{Mat}(m, n, \mathbb{K})$, die bestimmt wird durch*

$$Ae_j = (\phi_{B'}^{-1} \circ F \circ \phi_B)(e_j)$$

*heißt darstellende Matrix der linearen Abbildung $F$ bzgl. der Basen $B$, $B'$ und wird mit $M_B^{B'}(F)$ bezeichnet.*

## Beispiel

Seien $V = \mathbb{K}^n$, $W = \mathbb{K}^m$ und $B$ bzw. $B'$ die zugehörigen *kanonischen* Basen. Dann gilt $\phi_B = \mathrm{Id}_V$, $\phi_{B'} = \mathrm{Id}_W$ und für jede lineare Abbildung $F \in L(\mathbb{K}^n, \mathbb{K}^m)$ ist $M_B^{B'}(F)$ die kanonisch zugeordnete Matrix.

### Satz

*Mit den obigen Bezeichnungen ist die darstellende Matrix $A = (a_{ij})_{i,j} = M_B^{B'}(F)$ durch folgende Gleichung bestimmt:*

$$F(v_j) = \sum_{i=1}^{m} a_{ij} w_i, \quad j = 1, \ldots, n. \tag{11}$$

**Beweis.** Die $A = M_B^{B'}(F)$ definierende Gleichung

$$(\phi_{B'}^{-1} \circ F \circ \phi_B)(e_j) = Ae_j = \sum_{i=1}^{m} a_{ij} e_i$$

geht durch Anwendung von $\phi_{B'}$ in die äquivalente Gleichung (11) über, da $\phi_B(e_j) = v_j$ und $\phi_{B'}(e_i) = w_i$. □



Lineare Abbildungen | Lineare Abbildungen und Matrizen

### Satz

*Seien $V, V', V''$ endlich-dimensionale Vektorräume mit Basen $B, B', B''$ und $F \in L(V, V')$, $G \in L(V', V'')$.*
*Dann gilt $G \circ F \in L(V, V'')$ und $M_B^{B''}(G \circ F) = M_{B'}^{B''}(G) M_B^{B'}(F)$.*

**Beweis.** Die Linearität von $G \circ F$ folgt, wie gesehen, aus der von $F$ und $G$. Die darstellende Matrix $M_B^{B''}(G \circ F)$ ergibt sich aus

$$\phi_{B''}^{-1} \circ (G \circ F) \circ \phi_B = (\phi_{B''}^{-1} \circ G \circ \phi_{B'}) \circ (\phi_{B'}^{-1} \circ F \circ \phi_B).$$

□

# Endomorphismen

## Definition

*Sei $V$ ein Vektorraum. Lineare Abbildungen von $V$ nach $V$ heißen auch Endomorphismen von $V$, $\mathrm{End}(V) := L(V, V)$.*

*Den Vektorraum der quadratischen $(n \times n)$-Matrizen bezeichnet man mit $\mathrm{Mat}(n, \mathbb{K}) := \mathrm{Mat}(n, n, \mathbb{K})$.*

*Für die darstellende Matrix eines Endomorphismus $F \in \mathrm{End}(V)$ bzgl. einer Basis $B$ von $V$ schreibt man $M_B(F) := M_B^B(F)$.*

## Beispiel

Sei $F \in \mathrm{End}(\mathbb{R}^2)$ gegeben durch $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \in \mathrm{Mat}(2, \mathbb{R})$. Betrachte die Basis $B = (b_1, b_2) := (e_1 + e_2, e_1 - e_2)$. Es gilt
$F(b_1) = e_2 + e_1 = b_1$ und $F(b_2) = e_2 - e_1 = -b_2$.
Die darstellende Matrix bzgl. der Basis $B = (b_1, b_2)$ ist daher $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.





# Rang linearer Abbildungen und Matrizen

## Definition (Rang einer linearen Abbildung und einer Matrix)

*Der Rang einer linearen Abbildung $F\colon V \to W$ ist die nicht-negative ganze Zahl*

$$\mathrm{rg}(F) := \dim \mathrm{im}\, F.$$

*Sei $A = (a_{ij})_{i,j} \in \mathrm{Mat}(m, n, \mathbb{K})$ eine Matrix.*
*Die Dimension $\mathrm{rg}(A)$ des von den $n$ Spalten von $A$ erzeugten Unterraumes des $\mathbb{K}^m$ nennt man den Spaltenrang von $A$.*

## Satz

*Seien $V$, $W$ endlich-dimensionale Vektorräume und $B$, $B'$ Basen von $V$ bzw. $W$. Dann gilt für jede lineare Abbildung $F\colon V \to W$*

$$\mathrm{rg}(F) = \mathrm{rg}(M_B^{B'}(F)).$$

**Beweis.** Das Bild von $F \circ \phi_B$ stimmt mit dem von $F$ überein, da $\phi_B$ als Isomorphismus surjektiv ist.
Das Bild von $F \circ \phi_B$ wird durch $\phi_{B'}^{-1}$ isomorph auf das Bild der zu $M_B^{B'}(F)$ gehörenden linearen Abbildung abgebildet.
Für die letztgenannte lineare Abbildung bilden die Bilder der kanonischen Basisvektoren gerade die Spaltenvektoren der Matrix $M_B^{B'}(F)$.
Also $\mathrm{rg}(F) = \mathrm{rg}(F \circ \phi_B) = \mathrm{rg}(M_B^{B'}(F))$. $\square$

**Bemerkung**

Sei $A = (a_{ij})_{i,j} \in \mathrm{Mat}(m, n, \mathbb{K})$ eine Matrix.
Die Dimension des von den $m$ Zeilen $(a_{i1}, \ldots, a_{in}) \in \mathbb{K}^n$ erzeugten Unterraumes des $\mathbb{K}^n$ heißt Zeilenrang von $A$. Wir werden später sehen, dass Zeilenrang=Spaltenrang.





## Rang und Dimension

**Satz (Dimensionsformel, auch Rangsatz)**

*Seien $V$, $W$ Vektorräume,* $\dim V$ *endlich und $F \in L(V, W)$. Dann gilt*

$$\mathrm{rg}(F) + \dim \ker(F) \;=\; \dim V.$$

**Beweis.**
Das Bild jeder Basis von $V$ unter $F$ ist ein Erzeugendensystem des Bildes $F(V)$. Also ist $\dim F(V) \leq \dim V$.
Sei $(u_1, \ldots, u_k)$ eine Basis von $\ker(F)$, $(w_1, \ldots, w_r)$ eine Basis des Bildes $F(V)$. Wir wählen Urbilder $v_1, \ldots, v_r \in V$, so dass $F(v_i) = w_i$.

**Behauptung:**

$(u_1, \ldots, u_k, v_1, \ldots, v_r)$ ist eine Basis von $V$.

Aus der Behauptung folgt die Aussage des Rangsatzes: $\dim V = k + r$.

## Beweis der Behauptung.

1) Wir zeigen zuerst, dass $(u_1, \ldots, u_k, v_1, \ldots, v_r)$ linear unabhängig ist.
   Aus $0 = \sum_{i=1}^{k} \lambda_i u_i + \sum_{j=1}^{r} \mu_j v_j$ folgt

$$0 = F\Big(\sum_{i=1}^{k} \lambda_i u_i + \sum_{j=1}^{r} \mu_j v_j\Big) \overset{u_i \in \ker(F)}{=} \sum_{j=1}^{r} \mu_j w_j$$

   und somit $\mu_1 = \ldots = \mu_r = 0$, da die Familie $(w_j)$ linear unabhängig ist.
   Also $0 = \sum_{i=1}^{k} \lambda_i u_i$, woraus $\lambda_1 = \cdots = \lambda_k = 0$ folgt, wegen der linearen Unabhängigkeit von $(u_i)$ als Basis von $\ker F$.
2) Nun zeigen wir, dass $\mathrm{span}\{u_1, \ldots, u_k, v_1, \ldots, v_r\} = V$.
   Sei $v \in V$. Wir schreiben $F(v) = \sum_{j=1}^{r} \mu_j w_j$.
   Dann gilt $v - \sum_{j=1}^{r} \mu_j v_j \in \ker(F) = \mathrm{span}\{u_1, \ldots, u_k\}$
   und somit $v \in \mathrm{span}\{u_1, \ldots, u_k, v_1, \ldots, v_r\}$. □





Wir halten zwei Folgerungen aus der Dimensionsformel für lineare Abbildungen fest:

## Folgerung

*Seien $V$, $W$ Vektorräume, $\dim V < \infty$ und $F \in L(V, W)$. Dann gilt:*

(i) *$F$ ist genau dann surjektiv, wenn $\mathrm{rg}(F) = \dim W$,*

(ii) *$F$ ist genau dann injektiv, wenn $\mathrm{rg}(F) = \dim V$,*

(iii) *$F$ ist genau dann bijektiv, wenn $\mathrm{rg}(F) = \dim V = \dim W$.*

**Beweis.**

(i) $F$ surjektiv $\iff F(V) = W \iff \dim F(V) = \dim W$.

(ii) $F$ injektiv $\iff \ker(F) = \{0\} \overset{\text{(Dim.formel)}}{\iff} \mathrm{rg}(F) = \dim V$.

(iii) folgt aus (i-ii). □

## Folgerung

*Seien $V$, $W$ Vektorräume mit* $\dim V = \dim W < \infty$.

*Dann gilt: Eine lineare Abbildung $F\colon V \to W$ ist genau dann ein Isomorphismus, wenn eine der folgenden äquivalenten Bedingungen erfüllt ist:*

(i) *$F$ ist injektiv,*

(ii) *$F$ ist surjektiv,*

(iii) *$F$ ist bijektiv.*

**Beweis.** Wegen der Dimensionsformel gilt

$$\dim(V) \;=\; \mathrm{rg}(F) + \dim(\ker(F)) \;=\; \dim(W).$$

Daraus folgen die Äquivalenzen. □





# Transponierte Matrix

## Definition (Transponierte Matrix)

*Sei $A = (a_{ij})_{i,j}$ eine $(m \times n)$-Matrix.*
*Die $(n \times m)$-Matrix $A^{\mathrm{T}}$ mit den Einträgen*

$$(A^{\mathrm{T}})_{ij} := a_{ji}$$

*heißt die zu $A$ transponierte Matrix.*

## Bemerkung

Die Abbildung $\alpha\colon \mathrm{Mat}(n,\mathbb{K}) \ni A \mapsto A^{\mathrm{T}} \in \mathrm{Mat}(n,\mathbb{K})$, ist eine Involution auf der Menge $\mathrm{Mat}(n,\mathbb{K})$, d.h. $\alpha \circ \alpha = \mathrm{Id}$.

### Satz

*Für alle $A \in \mathrm{Mat}(m, n, \mathbb{K})$ und $B \in \mathrm{Mat}(n, r, \mathbb{K})$ gilt*

$$(AB)^{\mathrm{T}} = B^{\mathrm{T}} A^{\mathrm{T}}.$$

**Beweis.** Für $C = AB$ und $D = B^{\mathrm{T}} A^{\mathrm{T}}$ gilt

$$(C^{\mathrm{T}})_{ki} = c_{ik} = \sum_{j=1}^{n} a_{ij} b_{jk} = \sum_{j=1}^{n} (B^{\mathrm{T}})_{kj} (A^{\mathrm{T}})_{ji} = (D)_{ki}.$$

□





## Gaußscher Algorithmus: Einführung

Der Gaußsche Algorithmus (Gaußsches Eliminierungsverfahren) überführt eine Matrix

$$A := \begin{pmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} =: (a_{ij})_{\substack{i=1\dots m \\ j=1,\dots n}}$$

durch Anwendung elementarer Zeilenoperationen in eine Matrix

$$B = \begin{pmatrix} b_{11} & \dots & b_{1n} \\ \vdots & & \vdots \\ b_{m1} & \dots & b_{mn} \end{pmatrix} =: (b_{ij})_{\substack{i=1\dots m \\ j=1,\dots n}} .$$

Analog ließe sich auch eine Variante des Algorithmus betrachten, bei der Spaltenoperationen durchgeführt werden. Wir werden jedoch stets “zeilenweise” vorgehen. (Im Bedarfsfall können wir vorab und nach Abschluss transponieren!)

# Basis von Unterräumen aus Erzeugendensystemen

Der Gaußsche Algorithmus kann zum Lösen linearer Gleichungssysteme

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= c_1 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= c_2 \\
&\vdots \\
a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= c_m
\end{aligned}
$$

verwendet werden. Hierbei sind die $a_{ij}, c_j \in \mathbb{K}$ gegeben. Gesucht sind die n Unbekannten $x_1, x_2, \ldots, x_n \in \mathbb{K}$.
Wir verwenden den Gaußsche Algorithmus zunächst, um aus einem Erzeugendensystem $\{v_1, \ldots, v_m\}$ eines Unterraums von $V = \mathbb{K}^n$ eine Basis des Unterraum zu bestimmen.





Sei $V = \mathbb{K}^n$ $U = \mathrm{span}\{v_1, \ldots, v_m\} \subseteq V$ der betrachtete Unterraum. Wir schreiben die Vektoren

$$v_i = (a_{i1}, a_{i2}, \ldots, a_{in}) \in \mathbb{K}^n$$

als Zeilen (!) in eine Matrix

$$A := \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ a_{21} & \ldots & a_{2n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} =: (a_{ij})_{\substack{i=1\ldots m \\ j=1,\ldots n}} .$$

Wir bemerken, wobei sich (ii) und (iii) aus dem Austauschsatz ergeben:

(i) $\mathrm{span}\{v_1, \ldots, v_m\} = \mathrm{span}\{v_{\varphi(1)}, \ldots, v_{\varphi(m)}\}$ für jede Permutation $\varphi\colon \{1, \ldots, m\} \to \{1, \ldots, m\}$,

(ii) $\mathrm{span}\{\lambda v_1, v_2, \ldots, v_m\} = \mathrm{span}\{v_1, \ldots, v_m\}$ für alle $\lambda \in \mathbb{K}^* := \mathbb{K} \setminus \{0\}$

(iii) $\mathrm{span}\{v_1, \ldots, v_{i-1}, v_i + \lambda v_1, v_{i+1}, \ldots, v_m\} = \mathrm{span}\{v_1, \ldots, v_m\}$ für $i \neq 1$ und alle $\lambda \in \mathbb{K}$.

# Gaußscher Algorithmus

Dieser besteht im iterativen Anwenden der folgenden drei elementaren Zeilenoperationen auf eine Matrix $A = (a_{ij})_{\substack{i=1\dots m \\ j=1,\dots n}}$:

(i) Vertauschen zweier Zeilen,

(ii) Multiplizieren einer Zeile mit einem Skalar $\lambda \in \mathbb{K}^* := \mathbb{K} \setminus \{0\}$,

(iii) Addieren eines Vielfachen einer Zeile zu einer anderen Zeile.

Dies macht man solange, bis man eine Matrix $B = (b_{ij})_{\substack{i=1\dots m \\ j=1,\dots n}}$ erhält, die Zeilenstufenform hat, d.h.

Es existieren ein $k$ mit $0 \leq k \leq m$ und $k$ Indizes $1 \leq j_1 < j_2 < \ldots < j_k \leq n$ so dass die Matrixeinträge $b_{ij}$ die folgenden Eigenschaften haben:

(1) $b_{1j_1} = b_{2j_2} = \ldots = b_{kj_k} = 1$,

(2) $b_{ij} = 0$ für alle $i = 1, \ldots, k$ und $1 \leq j < j_i$,

(3) $b_{ij} = 0$ für alle $i = k+1, \ldots, m$ und $j = 1, \ldots n$.





Eine Matrix $B = (b_{ij})_{\substack{i=1\dots m \\ j=1,\dots n}}$ in Zeilenstufenform sieht so aus:

$$\begin{pmatrix}
0 & \overset{j_1-1 \text{ viele}}{\dots} & 0 & 1 & b_{1j_1+1} & \dots & b_{1n} \\
0 & \overset{j_2-1 \text{ viele}}{\dots} & 0 & 1 & b_{2j_2+1} & \dots & b_{2n} \\
\vdots & & & & & & \vdots \\
0 & \overset{j_i-1 \text{ viele}}{\dots} & 0 & 1 & b_{ij_i+1} & \dots & b_{in} \\
\vdots & & & & & & \vdots \\
0 & \overset{j_k-1 \text{ viele}}{\dots} & 0 & 1 & b_{kj_k+1} & \dots & b_{kn} \\
0 & & & \dots & & & 0 \\
\vdots & & & & & & \vdots \\
0 & & & \dots & & & 0
\end{pmatrix}$$

# Konklusion

Die Vektoren, die aus den ersten $k$ Zeilen von $B$ bestehen,

$$\begin{array}{rcl} w_1 & = & (0 \overset{j_1-1 \text{ viele}}{\ldots} 0, 1, b_{1j_1+1}, \ldots\ldots\ldots, b_{1n}) \\ \vdots & & \vdots \\ w_k & = & (0 \overset{j_k-1 \text{ viele}}{\ldots} 0, 1, b_{kj_k+1}, \ldots, b_{kn}) \end{array}$$

bilden eine Basis von $U = \mathrm{span}\{v_1, \ldots v_m\}$, denn

- Die elementaren Zeilenoperationen (i), (ii) und (iii) haben die lineare Hülle der Vektoren, die sich aus den Zeilen der Matrix ergeben, nicht verändert, d.h. $(w_1, \ldots w_k)$ sind ein Erzeugendensystem von $U$.
- Die Vektoren $(w_1, \ldots w_k)$ sind linear unabhängig, da

$$0 = \sum_{i=1}^{k} \lambda_i w_i = (\ldots, \lambda_1, \ldots, \lambda_2 + \lambda_1 \cdot (\ldots), \ldots, etc.)$$

impliziert $\lambda_i = 0$ für $i = 1, \ldots k$.





**Zahlenbeispiel.**

Sei z.B. $\mathbb{K}$ einer der Körper $\mathbb{Q}$, $\mathbb{R}$ oder $\mathbb{C}$ und $V = \mathbb{K}^5$ und ein Erzeugendensystem des Unterraums $U \subseteq V$ gegeben durch

$$\begin{array}{rcl} v_1 & = & (0, 0, 2, 1, 0) \\ v_2 & = & (0, 1, 0, 2, 1) \\ v_3 & = & (0, 2, 1, 1, 1) \\ v_4 & = & (0, 4, 4, 3, 2) \end{array}$$

Vertauschen der ersten beiden Zeilen ändert nach (i) nicht die lineare Hülle der Zeilenvektoren:

$$A = \begin{pmatrix} 0 & 0 & 2 & 1 & 0 \\ 0 & 1 & 0 & 2 & 1 \\ 0 & 2 & 1 & 1 & 1 \\ 0 & 4 & 4 & 3 & 2 \end{pmatrix} \overset{(i)}{\to} \begin{pmatrix} 0 & 1 & 0 & 2 & 1 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 2 & 1 & 1 & 1 \\ 0 & 4 & 4 & 3 & 2 \end{pmatrix}$$

**Weiter im Zahlenbeispiel:**
Die erste Spalte der Matrix

$$\begin{pmatrix} 0 & 1 & 0 & 2 & 1 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 2 & 1 & 1 & 1 \\ 0 & 4 & 4 & 3 & 2 \end{pmatrix}$$

ist Null.

Die zweite Spalte beginnt mit $w_{1j_1} = w_{12} = 1 \neq 0$.

Daher kann man durch Addition von geeigneten Vielfachen der ersten Zeile zu den anderen Zeilen erreichen, dass alle Einträge unterhalb von $w_{12}$ Null werden:



Lineare Abbildungen | Gaußscher Algorithmus

**Weiter im Zahlenbeispiel:**

$$\begin{pmatrix} 0 & 1 & 0 & 2 & 1 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 2 & 1 & 1 & 1 \\ 0 & 4 & 4 & 3 & 2 \end{pmatrix} \overset{-2\times I+III}{\longrightarrow} \begin{pmatrix} 0 & 1 & 0 & 2 & 1 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 1 & -3 & -1 \\ 0 & 4 & 4 & 3 & 2 \end{pmatrix}$$

$$\overset{-4\times I+IV}{\longrightarrow} \begin{pmatrix} 0 & 1 & 0 & 2 & 1 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 1 & -3 & -1 \\ 0 & 0 & 4 & -5 & -2 \end{pmatrix}$$

Als Nächstes produzieren wir Nullen unterhalb von $w_{2j_2} = w_{23} = 2 \neq 0$ durch Addition von Vielfachen der zweiten Zeile:

## Weiter im Zahlenbeispiel:

$$\begin{pmatrix} 0 & 1 & 0 & 2 & 1 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 1 & -3 & -1 \\ 0 & 0 & 4 & -5 & -2 \end{pmatrix} \xrightarrow[-2\times II+IV]{-\frac{1}{2}\times II+III} \begin{pmatrix} 0 & 1 & 0 & 2 & 1 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & -\frac{7}{2} & -1 \\ 0 & 0 & 0 & -7 & -2 \end{pmatrix}$$

Wir erhalten $w_{3j_3} = w_{34} = -\frac{7}{2} \neq 0$ und $-2 \times III + IV$ liefert:

$$A' = \begin{pmatrix} 0 & 1 & 0 & 2 & 1 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & -\frac{7}{2} & -1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Nun multiplizieren wir noch die zweite Zeile mit $\frac{1}{2}$ und die dritte mit $-\frac{2}{7}$ und erhalten:



Lineare Abbildungen | Gaußscher Algorithmus

## Weiter im Zahlenbeispiel:

$$B = \begin{pmatrix} 0 & 1 & 0 & 2 & 1 \\ 0 & 0 & 1 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 1 & \frac{2}{7} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Also $k = 3$ und

$$\begin{aligned} w_1 &= (0, 1, 0, 2, 1) \\ w_2 &= (0, 0, 1, \frac{1}{2}, 0) \\ w_3 &= (0, 0, 0, 1, \frac{2}{7}) \end{aligned}$$

Somit ist $(w_1, w_2, w_3)$ eine Basis von $U = \mathrm{span}\{v_1, v_2, v_3, v_4\} \subseteq \mathbb{K}^5$.

**Bemerkung:**
Auf den letzten Schritt könnte man auch verzichten, wenn es nur darum geht, eine Basis zu erhalten.

# Lineare Gleichungssysteme

## Definition

*Ein System von Gleichungen der Form*

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

*wobei $a_{ij}, b_j \in \mathbb{K}$ gegeben und $x_1, x_2, \ldots, x_n \in \mathbb{K}$ gesucht sind, bezeichnet man als lineares Gleichungssystem.*

*Ein solches lässt sich in der Form $Ax = b$ schreiben, mit $b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in \mathbb{K}^m$ und*

*$A = (a_{ij}) \in \mathrm{Mat}(m, n, \mathbb{K})$ gegeben und $x \in \mathbb{K}^n$ gesucht.*

*$m$ ist die Anzahl der Gleichungen und $n$ die Anzahl der Unbekannten.*

*Das System heißt homogen, falls $b = 0$ und inhomogen falls $b \neq 0$.*





## Satz

(i) *Die Lösungsmenge eines homogenen linearen Gleichungssytems $Ax = 0$, $A \in \mathrm{Mat}(m, n, \mathbb{K})$, ist genau der Kern der zu $A$ gehörenden linearen Abbildung, also insbesondere ein Unterraum $U \subseteq \mathbb{K}^n$.*

(ii) *Die Dimension von $U$ beträgt $n - r$, wobei $r = \mathrm{rg}(A)$.*

**Beweis.** (i) folgt aus der Definition und (ii) aus der Dimensionsformel. □

## Satz

*Sei $A \in \mathrm{Mat}(m, n, \mathbb{K})$ eine Matrix und $B$ die aus $A$ mittels Gaußschem Algorithmus hervorgegangene Matrix in Zeilenstufenform.*
*Dann haben die linearen homogenen Gleichungssysteme $Ax = 0$ und $Bx = 0$ denselben Lösungsraum.*

**Beweis.** Die Spalten von $A$ sind gegeben durch die Bilder $Ae_i$ der kanonischen Basis. Der Beweis beruht nun darauf, dass wir jede Zeile von $A$ als ein Element des Dualraums $(\mathbb{K}^n)^* = L(\mathbb{K}^n, \mathbb{K})$ interpretieren.
Für $i = 1, \dots m$ definieren wir $\alpha_i \in L(\mathbb{K}^n, \mathbb{K})$ mittels

$$\alpha_i(e_j) := a_{ij}, \quad \text{d.h.} \qquad Ax = \begin{pmatrix} \alpha_1(x) \\ \vdots \\ \alpha_m(x) \end{pmatrix}.$$

Die $a_{ij}$ sind also die Komponenten der Linearform $\alpha_i$ in der kanonischen dualen Basis von $(\mathbb{K}^n)^*$. Sei $V^* := \operatorname{span}\{\alpha_1, \dots, \alpha_m\} \subseteq (\mathbb{K}^n)^*$ der durch die Zeilen von $A$ aufgespannte Unterraum. Mittels des Gaußschen Algorithmus erhält man die Matrix $B$ in Zeilenstufenform, die eine Basis $(\beta_1, \dots, \beta_k)$ von $V^*$ liefert. Also haben beide linearen homogenen Gleichungssysteme $Ax = 0$ und $Bx = 0$ denselben Lösungsraum

$$\{x \in \mathbb{K}^n \mid L(x) = 0 \ \forall \ L \in V^*\}$$

$\square$





**Lösen von homogenen linearen Gleichungssystemen mittels Gauß-Algorithmus**

- Wir wollen das Gleichungssystem $Ax = 0$ lösen. Dazu bringen wir $A$ auf Zeilenstufenform und erhalten $B$ vom Zeilenrang $k$.
- D.h. wir erhalten ein Gleichungssystem der Form

$$\begin{array}{rcccl} x_{j_1} & + & \sum_{l=j_1+1}^{n} b_{1l} x_l & = & 0 \\ & & & \vdots & \\ x_{j_i} & + & \sum_{l=j_i+1}^{n} b_{il} x_l & = & 0 \\ & & & \vdots & \\ x_{j_k} & + & \sum_{l=j_k+1}^{n} b_{kl} x_l & = & 0 \end{array}$$

- Dieses löst man, indem man die Unbekannten $x_j$ mit $j \notin \{j_1, \dots, j_k\}$ frei wählt und die Lösungen $x_{j_1}, \dots, x_{j_k}$ bestimmt.
- D.h. der Raum der Lösungen ist $(n-k)$-dimensional. Damit gilt auch $k = \mathit{Zeilenrang}(A) = \mathit{Spaltenrang}(B) = \mathit{Spaltenrang}(A)$.

## Satz

*Wir betrachten ein inhomogenes lineares Gleichungssytem*

$$Ax = b, \qquad \text{mit } A \in \mathrm{Mat}(m, n, \mathbb{K}). \tag{12}$$

*Es bezeichne $U'$ die Lösungsmenge von (12).*
*Falls $U' \neq \emptyset$, so ist $U'$ von der Form*

$$U' = x_0 + U = \{x_0 + u | u \in U\},$$

*wobei $x_0 \in U'$ eine (beliebige) Lösung von (12) ist und $U \subseteq \mathbb{K}^n$ der Lösungsraum des zugehörigen homogenen Systems $Ax = 0$ ist, d.h. $U = \ker(A)$.*

**Beweis.** Wir nehmen an, dass $U' \neq \emptyset$ gilt und wählen $x_0 \in U'$.
Für $u \in U$ gilt dann $A(x_0 + u) = Ax_0 + Au = b + 0 = b$, also $x_0 + U \subseteq U'$.
Sei umgekehrt $v \in U'$ eine Lösung von (12). Dann gilt
$A(v - x_0) = Av - Ax_0 = b - b = 0$, also $v \in x_0 + U$ und somit
$U' \subseteq x_0 + U$. □





## Definition

*Ein affiner Unterraum eines Vektorraums $V$ ist eine Teilmenge der Form $v + U$, wobei $U \subseteq V$ ein Untervektorraum ist und $v \in V$.*

## Bemerkung

Demnach ist der Lösungsraum eines inhomogenen linearen Gleichungssystems $Ax = b$, $A \in \mathrm{Mat}(m, n, \mathbb{K})$, ein affiner Unterraum von $\mathbb{K}^n$, wenn er nicht die leere Menge ist

Der Lösungsraum ist genau dann ein Untervektorraum von $\mathbb{K}^n$, wenn das Gleichungssystem homogen ist, d.h. wenn $b = 0$.

## Lösen von inhomogenen linearen Gleichungssystemen mittels Gauß-Algorithmus

- Wir wollen $Ax = b$ lösen. Wir bringen die $m \times (n+1)$ Matrix $(A|b)$ auf Zeilenstufenform und erhalten eine Matrix $(B|c)$.
  D.h. wir erhalten ein Gleichungssystem der Form

$$\begin{array}{rcl} x_{j_1} + \sum_{l=j_1+1}^{n} b_{1l} x_l & = & c_1 \\ & \vdots & \\ x_{j_k} + \sum_{l=j_k+1}^{n} b_{kl} x_l & = & c_k \\ 0 & = & c_{k+1} \\ & \vdots & \end{array}$$

- $Ax = b \iff (A|b)\begin{pmatrix} x \\ -1 \end{pmatrix} = 0 \iff (B|c)\begin{pmatrix} x \\ -1 \end{pmatrix} = 0 \iff Bx = c$ und $0 = c_{k+1} = \ldots = c_m$.
- Die restlichen $k$ Gleichungen löst man, indem man wieder die $n-k$ Unbekannten $x_j$ mit $j \notin \{j_1, \ldots, j_k\}$ frei wählt und die Lösungen $x_{j_1}, \ldots, x_{j_k}$ bestimmt.





**Zahlenbeispiel:** Wir wollen das inhomogene lineare Gleichungssystem

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 0 \\ 4x_1 + 5x_2 + 6x_3 &= 3 \\ 7x_1 + 8x_2 + 9x_3 &= 6 \end{aligned}$$

lösen. D.h. $Ax = b$ mit $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$ und $b = \begin{pmatrix} 0 \\ 3 \\ 6 \end{pmatrix}$. Erweiterte Matrix

$$(A|b) = \begin{pmatrix} 1 & 2 & 3 & 0 \\ 4 & 5 & 6 & 3 \\ 7 & 8 & 9 & 6 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & 3 & 0 \\ 0 & -3 & -6 & 3 \\ 0 & -6 & -12 & 6 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & 3 & 0 \\ 0 & -3 & -6 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\to \begin{pmatrix} 1 & 2 & 3 & 0 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix} = (B|c).$$

Da $c_3 = 0$, ist $Ax = b$ lösbar. Das zugehörige Gleichungssystem ist:

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 0 \\ x_2 + 2x_3 &= -1. \end{aligned}$$

D.h. $x_2 = -2x_3 - 1$, $x_1 = -2x_2 - 3x_3 = -2(-2x_3 - 1) - 3x_3 = x_3 + 2$.
Die allgemeine Lösung ist also $x_1 = \lambda + 2$, $x_2 = -2\lambda - 1$, $x_3 = \lambda \in \mathbb{K}$ beliebig.

# Äußere direkte Summe

## Definition

Seien $U, V$ Vektorräume über einem Körper $\mathbb{K}$. Wir versehen das kartesische Produkt $U \times V$ mit der Struktur eines $\mathbb{K}$-Vektorraums durch

$$\lambda(u, v) := (\lambda u, \lambda v) \quad \text{und} \quad (u, v) + (u', v') := (u + u', v + v') \ .$$

Dieser Vektorraum heißt die (äußere) direkte Summe $U \oplus V$ von $U$ und $V$.

Ist $(u_1, \dots, u_m)$ eine Basis von $U$ und $(v_1, \dots, v_n)$ eine Basis von $V$, so ist $((u_1, 0), \dots, (u_m, 0), (0, v_1), \dots (0, v_n))$ eine Basis von $U \oplus V$.
Insbesondere ist $\dim(U \oplus V) = \dim U + \dim V$.





# Innere Summe von Unterräumen

## Definition

*$W, W'$ seien Unterräume eines Vektorraums $V$.*

*Die (innere) Summe*

$$W + W' := \{w + w' | w \in W, w' \in W'\} \subseteq V$$

*ist der kleinste Unterraum von $V$, der $W$ und $W'$ enthält.*

## Satz

*Falls die Unterräume $W$ und $W'$ endlich-dimensional sind, gilt folgende Dimensionsformel*

$$\dim(W + W') = \dim W + \dim W' - \dim(W \cap W').$$

**Beweis der Dimensionsformel für die Summe.**
Wir betrachten die lineare Abbildung $F\colon W \oplus W' \to W + W'$,

$$(w, w') \mapsto F(w, w') := w - w'.$$

Offenbar ist $F$ surjektiv: $F(W \oplus W') = W + W'$ und daher $\mathrm{rg}(F) = \dim(W + W')$. Ferner ist

$$\ker(F) = \{(w, w) | w \in W \cap W'\} \cong W \cap W'.$$

Es ist $\dim(W \oplus W') = \dim W + \dim W'$.
Daher folgt aus der Dimensionsformel für lineare Abbildungen

$$\begin{aligned} \mathrm{rg}(F) &= \dim(W + W') = \dim(W \oplus W') - \dim \ker(F) \\ &= \dim(W \oplus W') - \dim(W \cap W') \\ &= \dim W + \dim W' - \dim(W \cap W'). \end{aligned}$$

$\square$





# Innere direkte Summe von Unterräumen

## Definition (Innere direkte Summe von Unterräumen)

*$W, W'$ seien Unterräume eines Vektorraums $V$.*
*Falls $W \cap W' = \{0\}$ gilt, bezeichnet man die Summe $W + W'$ als direkte Summe und schreibt (mit dem gleichen Symbol wie für die äußere direkte Summe)*

$$W \oplus W'$$

*Zwei Unterräume $W, W'$ heißen komplementär, wenn $V = W \oplus W'$.*

*Wir sagen dann, dass $W'$ ein Komplement zu $W$ (in $V$) ist.*

Bemerkung:
Das Komplement zu einem Unterraum $W \subseteq V$ ist nicht eindeutig.

## Beispiel

- Sei $V = \mathbb{R}^3$, $W = \text{span}\{e_1, e_2\}$ und $W' = \text{span}\{e_1 + e_2, e_3\}$. Die Summe $W + W' = V$ ist nicht direkt, denn $W \cap W' = \mathbb{R} \cdot (e_1 + e_2)$.
- In der äußeren direkten Summe $U \oplus V$ zweier Vektorräume sind die Unterräume $U \times \{0\}$ und $\{0\} \times V$ komplementär und es ist

$$U \oplus_{\text{äußer}} V = (U \times \{0\}) \oplus_{\text{inner}} (\{0\} \times V).$$

Durch die injektiven linearen Abbildungen

$$\begin{array}{rcl} U & \to & U \oplus_{\text{äußer}} V \\ u & \mapsto & (u, 0) \end{array} \quad \text{bzw.} \quad \begin{array}{rcl} V & \to & U \oplus_{\text{äußer}} V \\ v & \mapsto & (0, v) \end{array}$$

können wir $U$ und $V$ mit den Unterräumen $U \times \{0\}$ bzw. $\{0\} \times V$ der äußeren direkten Summe $U \oplus_{\text{äußer}} V$ identifizieren. Dies rechtfertigt, dass wir für innere und äußere direkte Summe das gleiche Symbol benutzen.





## Satz

*$W, W'$ seien Unterräume eines Vektorraums $V$. Es gilt:*

*$V = W \oplus W' \iff$ Jeder Vektor $v \in V$ hat eine eindeutige Darstellung*

$$v = w + w', \quad w \in W, \quad w' \in W'. \tag{13}$$

**Beweis.**

'$\Rightarrow$' Aus $V = W + W'$ folgt, dass jeder Vektor $v \in V$ eine Darstellung der Form (13) hat. Sei $v = w_1 + w_1'$ eine weitere solche Darstellung.

Es folgt $0 = v - v = (w - w_1) + (w' - w_1')$ und daraus $W \ni w - w_1 = -(w' - w_1') \in W'$.

Wegen $W \cap W' = \{0\}$ folgt nun $w - w_1 = w' - w_1' = 0$ und somit die Eindeutigkeit der Darstellung (13).

'$\Leftarrow$' Wenn umgekehrt jeder Vektor $v \in V$ eine eindeutige Darstellung (13) hat, so gilt $V = W + W'$ und aus $0 = w + (-w)$ mit $w \in W \cap W'$, folgt wg. der Eindeutigkeit von (13) $w = 0$, d.h. $W \cap W' = \{0\}$.

□

### Satz

*Sei $V$ ein endlich-dimensionaler Vektorraum. Dann gilt:*

a) *Jeder Unterraum $W \subseteq V$ besitzt ein Komplement.*

b) *$W, W' \subseteq V$ seien Unterräume. Dann sind folgende Aussagen äquivalent:*

   (i) $V = W \oplus W'$,

   (ii) $W' \cap W = \{0\}$ *und* $\dim W + \dim W' = \dim V$,

   (iii) $V = W + W'$ *und* $\dim W + \dim W' = \dim V$.

**Beweis.**

a) Sei $(w_1, \ldots, w_k)$ eine beliebige Basis des Unterraums $W$. Nach dem Austauschsatz von Steinitz können wir diese zu einer Basis $(w_1, \ldots, w_k, w'_1, \ldots, w'_l)$ von $V$ ergänzen.
$W' := \mathrm{span}\{w'_1, \ldots, w'_l\}$ ist dann ein Komplement zu $W$.

b) Die Äquivalenz der Eigenschaften (i)-(iii) folgt aus der Dimensionsformel

$$\dim(W + W') = \dim W + \dim W' - \dim W \cap W'$$

$\square$





# Kapitel 12

## Gruppen

# Gruppen

## Definition (Gruppe)

*Eine Gruppe $(G, \cdot)$ ist eine Menge $G$ zusammen mit einer assoziativen Verknüpfung $\cdot\colon G \times G \to G$, so dass*

(i) *es ein neutrales Element $e \in G$ gibt mit*

$$e \cdot a = a \cdot e = a \quad \text{für alle} \quad a \in G \quad \text{und}$$

(ii) *es zu jedem $a \in G$ ein Inverses $a^{-1}$ gibt, so dass $a \cdot a^{-1} = a^{-1} \cdot a = e$.*

## Bemerkung

Statt (i) und (ii) genügt es, die folgenden Bedingungen zu fordern, aus denen auch bereits die Eindeutigkeit von $e$ und $a^{-1}$ folgt (vgl. etwa Fischer):

(i') Es gibt ein Element $e \in G$ mit $e \cdot a = a$ für alle $a \in G$ und

(ii') Zu jedem $a \in G$ gibt es ein $a^{-1}$, so dass $a^{-1} \cdot a = e$.





**Beispiele von Gruppen**

(i) $(\mathbb{Z}, +), (\mathbb{Q}, +)$ sowie $(\mathbb{Q}\backslash\{0\}, \cdot)$ sind kommutative Gruppen. Für jeden Körper $\mathbb{K}$ sind durch $(\mathbb{K}, +)$ und $(\mathbb{K} \setminus \{0\}, \cdot)$ kommutative Gruppen gegeben.

(ii) Die Bijektionen $\varphi\colon X \to X$ einer Menge $X$ in sich bilden eine Gruppe, die mit $\mathrm{Bij}(X)$ bezeichnet wird. Die Vernüpfung ist die Verkettung, das neutrale Element ist die identische Abbildung $\mathrm{Id}_X$ und das Inverse einer Bijektion ist ihre Umkehrabbildung.

(iii) Im Fall einer endlichen Menge $X$ nennt man die Bijektionen $\sigma \in \mathrm{Bij}(X)$ auch Permutationen von $X$. Die Permutationsgruppe $\mathrm{Bij}(X)$ ist dann eine endliche Gruppe mit $n!$ Elementen, wobei $n = \mathrm{card}(X)$ die Anzahl der Elemente von $X$ ist. Unter der $n$-ten symmetrischen Gruppe versteht man die Permutationsgruppe $S_n := \mathrm{Bij}(\{1, 2, \ldots, n\})$.

(iv) Symmetrien von Objekten werden oft durch Gruppen beschrieben.

(v) Für jeden Vektorraum $V$ ist durch $(V, +)$ eine kommutative Gruppe gegeben.

## weitere Beispiele von Gruppen

(vi) Sei $V$ ein Vektorraum. Die Isomorphismen $V \to V$ heißen auch Automorphismen des Vektorraums $V$ und bilden die sogenannte Automorphismengruppe $\mathrm{Aut}(V)$ von $V$.

(vii) Im Fall eines endlichdimensionalen $\mathbb{K}$-Vektorraums $V$ spricht man auch von der allgemeinen linearen Gruppe über $\mathbb{K}$ und schreibt dafür $\mathrm{GL}(V) := \mathrm{Aut}(V)$.

Sei $V = \mathbb{K}^n$. Dann entspricht $\mathrm{GL}(V)$ gerade einer Gruppe von Matrizen, genauer der Gruppe der invertierbaren $n \times n$-Matrizen mit Einträgen aus $\mathbb{K}$, die auch mit $\mathrm{GL}(n, \mathbb{K})$ bezeichnet wird.

Es gilt

$$\mathrm{GL}(n, \mathbb{K}) \quad = \quad \{A \in \mathrm{Mat}(n, \mathbb{K}) | \mathrm{rg}(A) = n\}.$$

Diese Gruppe ist nicht kommutativ, wenn $n > 1$.

(viii) $\mathrm{GL}(1, \mathbb{K}) = (\mathbb{K}^* := \mathbb{K} \setminus \{0\}, \cdot)$ ist eine kommutative Gruppe.





## Noch mehr Beispiele von Gruppen

(ix) Die Kreislinie $S^1 := \{z \in \mathbb{C} | |z| = 1\}$ ist mit der Multiplikation komplexer Zahlen eine kommutative Gruppe.

(x) Das kartesische Produkt $G_1 \times G_2$ von Gruppen $G_1, G_2$ ist mit der komponentenweisen Verknüpfung wieder eine Gruppe. D.h. $(g_1, g_2) \cdot (h_1, h_2) := (g_1 \cdot h_1, g_2 \cdot h_2)$ definiert eine Gruppenstruktur auf $G_1 \times G_2$.
$G_1 \times G_2$ ist genau dann kommutativ, wenn $G_1$ und $G_2$ kommutativ sind.

**Inverse von Produkten**

In jeder Gruppe gilt $(gh)^{-1} = h^{-1}g^{-1}$. Dies folgt aus der Eindeutigkeit der Inversen: $(gh)(h^{-1}g^{-1}) = g(hh^{-1})g^{-1} = gg^{-1} = e$.

## Definition (Gruppenhomomorphismen)

*Eine Abbildung $\varphi\colon G \to H$ zwischen Gruppen $G$ und $H$ heißt ein Gruppenhomomorphismus, falls*

$$\varphi(a \cdot b) = \varphi(a) \cdot \varphi(b) \text{ für alle } a, b \in G \ .$$

*Ein bijektiver Gruppenhomomorphismus heißt auch Isomorphismus von Gruppen. Ein Isomorphismus einer Gruppe in sich heißt auch Automorphismus.*

*Zwei Gruppen $G$ und $H$ heißen isomorph, falls es einen Isomorphismus von Gruppen $\varphi\colon G \to H$ gibt.*

*Eine Untergruppe einer Gruppe $G$ ist eine nicht-leere Teilmenge $H \subseteq G$, die mit $a, b \in H$ auch $a \cdot b$ und $a^{-1}$ enthält.*

Mit $a \in H$ enthält die Untergruppe auch $a^{-1}$ und damit $a \cdot a^{-1} = e$, das neutrale Element von $G$. Jede Untergruppe $H \subseteq G$ einer Gruppe $G$ ist mit der induzierten Verknüpfung wieder eine Gruppe.





## Übungsaufgaben/Beispiele I

- Jede Gruppe $G$ hat sich selbst und $\{e\}$ als Untergruppen.
- Jeder Gruppenhomomorphismus $\varphi\colon G \to H$ bildet das neutrale Element in $G$ auf das neutrale Element in $H$ ab und erfüllt

  $$\varphi(a^{-1}) = \varphi(a)^{-1}$$

  für alle $a \in G$.
- Das Inverse eines Isomorphismus $\varphi\colon G \to H$ ist wieder ein Isomorphismus von Gruppen.
- Das Bild $\varphi(K) \subseteq H$ einer Untergruppe $K \subseteq G$ unter einem Gruppenhomomorphismus $\varphi\colon G \to H$ ist eine Untergruppe von $H$
- Das Urbild $\varphi^{-1}(K) \subseteq G$ einer Untergruppe $K \subseteq H$ unter einem Gruppenhomomorphismus $\varphi\colon G \to H$ ist eine Untergruppe von $G$.
- Insbesondere ist der Kern $\ker \varphi := \varphi^{-1}(\{e\})$ des Gruppenhomomorphismus $\varphi$ eine Untergruppe.

## Übungsaufgaben/Beispiele II

- Die Automorphismen einer Gruppe $G$ bilden eine Untergruppe der Bijektionen der Menge, die $G$ zu Grunde liegt, $\mathrm{Aut}(G) \subseteq \mathrm{Bij}(G)$.
- Die Gruppen $(\mathbb{R}, +)$ und $(\mathbb{R}_+, \cdot)$ sind isomorph. Die Exponentialfunktion $\exp\colon \mathbb{R} \to \mathbb{R}_+$ ist ein Isomorphismus von Gruppen, und damit auch ihre Umkehrfunktion $\ln\colon \mathbb{R}_+ \to \mathbb{R}$.
- Die Abbildung $t \mapsto \exp(it)$ definiert einen surjektiven aber nicht injektiven Gruppenhomomorphismus $\varphi\colon \mathbb{R} \to S^1$. Sein Kern ist die Untergruppe $\ker\varphi = 2\pi\mathbb{Z}$ von $\mathbb{R}$.
- Ein Beispiel für einen Gruppenautomorphismus von $\mathrm{GL}(n, \mathbb{K})$ wird durch
$$A \quad \mapsto \quad (A^{-1})^{\mathrm{T}}$$
gegeben, und es ist dabei $(A^{-1})^{\mathrm{T}} = (A^{\mathrm{T}})^{-1}$.





## Die symmetrische Gruppe

Eine Permutation $\sigma \in S_n$ kann man durch eine zweizeilige Matrix
$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix}$$
beschreiben. Es genügt, die untere Zeile zu notieren: Beispielsweise schreibt man statt $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 1 & 2 & 4 & 5 \end{pmatrix}$ auch einfach 31245, besser [31245].

Eine Transposition $\tau = \tau_{ij}$ ist eine Permutation, die zwei Zahlen $i < j$ vertauscht und alle anderen Zahlen unverändert lässt. Beispielsweise enthält $S_3$ genau 3 Transpositionen:
$$\tau_{12} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \quad \tau_{13} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix},$$
$$\tau_{23} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$$

Ist $\tau$ eine Transposition, so gilt offensichtlich $\tau^{-1} = \tau$

## Satz

*Die symmetrische Gruppe $S_n$, $n \in \mathbb{N}$, ist nur für $n \leq 2$ kommutativ.*

**Beweis.**
Es ist klar, dass $S_1$ und $S_2$ kommutativ sind. $S_3$ ist nicht kommutativ, denn z.B. ist $\tau_{12} \circ \tau_{13} \neq \tau_{13} \circ \tau_{12}$. Letzteres folgt bereits aus

$$\tau_{12} \circ \tau_{13}(1) = \tau_{12}(3) = 3$$
$$\tau_{13} \circ \tau_{12}(1) = \tau_{13}(2) = 2.$$

Die Abbildung

$$\sigma \mapsto \begin{pmatrix} 1 & 2 & 3 & 4 & \cdots & n \\ \sigma(1) & \sigma(2) & \sigma(3) & 4 & \cdots & n \end{pmatrix}$$

ist ein injektiver Gruppenhomomorphismus $S_3 \to S_n$ ($n \geq 3$). Das Bild ist eine nicht kommutative Untergruppe von $S_n$. Daher ist auch $S_n$ für $n \geq 3$ nicht kommutativ. □





## Satz

*Für alle $\sigma \in S_n$ ($n \geq 2$) gibt es Transpositionen $\tau_1, \ldots \tau_k \in S_n$, so dass $\sigma = \tau_1 \circ \cdots \circ \tau_k$. Die Darstellung von $\sigma$ als Produkt ist nicht eindeutig.*

**Beweis.**
Für jede Transposition $\tau$ gilt $\tau^{-1} = \tau$ und somit $\mathrm{Id} = \tau \circ \tau$.
Sei daher von nun an $\mathrm{Id} \neq \sigma \in S_n$. Dann existiert ein $1 \leq i_1 \leq n$ mit

- $\sigma(i) = i$ für alle $1 \leq i \leq i_1 - 1$ und
- $\sigma(i_1) > i_1$.

Wir setzen $\tau_1 := \tau_{i_1\sigma(i_1)}$ und $\sigma_1 := \tau_1 \circ \sigma$. Dann gilt $\sigma_1(i) = i$ für alle $1 \leq i \leq i_1$.
Falls $\sigma_1 \neq \mathrm{Id}$, so gibt es wieder ein $i_2$, $i_1 < i_2 \leq n$, so dass $\sigma_1(i) = i$ für alle $1 \leq i \leq i_2 - 1$ und $\sigma_1(i_2) > i_2$. Wir setzen $\tau_2 := \tau_{i_2\sigma(i_2)}$ und $\sigma_2 := \tau_2 \circ \sigma_1$.
Durch Fortsetzen dieses Iterationsverfahren erhalten wir nach endlich vielen (genauer: nach $k \leq n - 1$) Schritten $\sigma_k = \tau_k \circ \cdots \circ \tau_1 \circ \sigma = \mathrm{Id}$ und somit $\sigma = \tau_1 \circ \cdots \circ \tau_k$. □

## Definition (Vorzeichen einer Permutation)

*Das Vorzeichen oder Signum $\varepsilon(\sigma)$ einer Permutation $\sigma \in S_n$ ist definiert als*

$$(*) \quad \varepsilon(\sigma) := \prod_{i<j} \frac{\sigma(j) - \sigma(i)}{j - i} \in \{-1, +1\}.$$

*(Man beachte, dass im Zähler und Nenner bis auf das Vorzeichen dasselbe Produkt steht.)*

*Die Permutation heißt gerade, wenn $\varepsilon(\sigma) = +1$ und ungerade, wenn $\varepsilon(\sigma) = -1$.*

## Bemerkung

Die Anzahl der negativen Faktoren im Produkt (*) ist genau die Anzahl der Paare $i < j$ mit $\sigma(i) > \sigma(j)$, den "Fehlstellungen".

Die Permutation $\sigma$ ist also genau dann gerade (bzw. ungerade), wenn diese Anzahl gerade (bzw. ungerade) ist.





## Satz

$\varepsilon\colon S_n \to \{-1, 1\}$ *ist ein Gruppenhomomorphismus in die Untergruppe* $\{-1, 1\} \subseteq \mathbb{R}^*$.

**Beweis.** Wir berechnen für $\sigma, \tau \in S_n$

$$\varepsilon(\sigma \circ \tau) = \prod_{i<j} \frac{\sigma(\tau(j)) - \sigma(\tau(i))}{j - i} = \prod_{i<j} \underbrace{\frac{\sigma(\tau(j)) - \sigma(\tau(i))}{\tau(j) - \tau(i)}}_{f_{ij}} \cdot \underbrace{\prod_{i<j} \frac{\tau(j) - \tau(i)}{j - i}}_{\varepsilon(\tau)}$$

Nun gilt aber (*) $f_{ij} = f_{ji}$ und somit

$$\begin{aligned} \prod_{i<j} f_{ij} &= \prod_{\substack{i<j \\ \tau(i)<\tau(j)}} f_{ij} \prod_{\substack{i<j \\ \tau(i)>\tau(j)}} f_{ij} \\ &\overset{(*)}{=} \prod_{\substack{i<j \\ \tau(i)<\tau(j)}} f_{ij} \prod_{\substack{i>j \\ \tau(i)<\tau(j)}} f_{ij} = \prod_{\tau(i)<\tau(j)} f_{ij} = \varepsilon(\sigma). \end{aligned}$$

Damit folgt $\prod_{i<j} f_{ij} = \varepsilon(\sigma)$ und die Behauptung. $\square$

## Folgerung

*Für $\sigma \in S_n$ gilt $\varepsilon(\sigma) = \varepsilon(\sigma^{-1})$.*

**Beweis.** Da $\varepsilon$ ein Gruppenhomomorphismus ist, gilt $\varepsilon(\sigma^{-1}) = \varepsilon(\sigma)^{-1}$, und in $\{1, -1\}$ ist jedes Element sein eigenes Inverses. $\square$

## Folgerung

*Sei $\sigma \in S_n$ und $\sigma = \tau_1 \circ \cdots \circ \tau_k$ eine Darstellung als Produkt von Transpositionen. Dann gilt*

$$\varepsilon(\sigma) = (-1)^k.$$

**Beweis.**
Für die Transposition $\tau = \tau_{12}$ gilt $\varepsilon(\tau) = -1$, denn $(1, 2)$ ist das einzige Paar $(i, j)$ mit $i < j$ und $\tau(i) > \tau(j)$. Sei $\sigma \in S_n$ mit $\sigma(1) = i$ und $\sigma(2) = j$. Dann gilt $\sigma \circ \tau_{12} \circ \sigma^{-1} = \tau_{ij}$ und somit für jede Transposition

$$\varepsilon(\tau_{ij}) = \varepsilon(\sigma)\varepsilon(\tau_{12})\varepsilon(\sigma^{-1}) = \varepsilon(\tau_{12}) = -1.$$

Also $\varepsilon(\sigma) = \varepsilon(\tau_1 \circ \cdots \circ \tau_k) = \varepsilon(\tau_1) \cdots \varepsilon(\tau_k) = (-1)^k$. $\square$





**Beispiel**
Ein Zykel $(i_1, ..., i_r)$ der Länge $r$ ist ein $r$-Tupel von paarweise verschiedenen Zahlen aus $\{1, \dots, n\}$ und wird aufgefasst als zyklische Permutation $i_1 \mapsto i_2 \mapsto ... \mapsto i_r \mapsto i_1$. Die Kommata lässt man in der Notation oft weg.
Die Permutation

$$\zeta := (1\ 2\ \dots\ n) := (1, 2, \dots, n) = \begin{pmatrix} 1 & 2 & \cdots & n-1 & n \\ 2 & 3 & \cdots & n & 1 \end{pmatrix}$$

ist gleich $\tau_{12} \circ \tau_{23} \circ \cdots \circ \tau_{n-1n}$.

Also ist $\varepsilon(\zeta) = (-1)^{n-1}$.

$\zeta$ erzeugt eine $n$-elementige Untergruppe

$$\langle \zeta \rangle := \{\zeta, \zeta^2, \dots, \zeta^n = \mathrm{Id}\} \subseteq S_n.$$

Die $\zeta^k$, $k = 1, 2, \dots, n-1$, sind zyklische Permutationen der Länge $n$.

Für $n = 3$ ist diese Untergruppe gerade die Untergruppe der geraden Permutationen in $S_3$.