

# Mathematik I für Studierende der Geophysik/Ozeanographie, Meteorologie und Physik Vorlesungsskript<sup>1</sup>

Christoph Schweigert  
Fachbereich Mathematik  
Universität Hamburg  
[www.math.uni-hamburg.de/home/schweigert](http://www.math.uni-hamburg.de/home/schweigert)

Hamburg, Wintersemester 2012/13

---

<sup>1</sup>Version vom 7. Januar 2013, basierend auf Versionen der Vorjahre. Ursprüngliche Version: Vicente Cortés, 2006/07

# Inhaltsverzeichnis I

## 1 Grundlagen

- Vorneweg
- Mengen und Quantoren
- Mathematische Aussagen und Beweise
- Abbildungen

## 2 Reelle und komplexe Zahlen

- Natürliche und ganze Zahlen
- Reelle Zahlen und Körperaxiome
- Ordnungsaxiome
- Der Absolutbetrag
- Komplexe Zahlen
- Der Betrag einer komplexen Zahl

## 3 Konvergenz von Folgen und Vollständigkeit

- Konvergenz von Folgen
- Konvergenz und Beschränktheit
- Konvergenz und algebraische Operationen

## Inhaltsverzeichnis II

- Konvergenz und Ordnungsrelation
- Cauchy-Folgen
- Das Vollständigkeitsaxiom
- Der Satz von Bolzano-Weierstraß
- Folgerungen aus dem Vollständigkeitsaxiom
- Die Quadratwurzel einer positiven Zahl
- Approximation reeller Zahlen durch rationale Zahlen
- Vollständigkeit von  $\mathbb{C}$
- Die geometrische Folge

### 4 Konvergenz von Reihen

- Absolute Konvergenz
- Cauchysches Konvergenzkriterium
- Majorantenkriterium
- Die geometrische Reihe
- Quotientenkriterium
- Die Exponentialreihe

## Inhaltsverzeichnis III

- Cauchy-Produkt von Reihen
- Sinus und Cosinus
- Polarkoordinaten
- Weitere Konvergenzkriterien für Reihen
- Umordnungen von Reihen

### 5 Eigenschaften reeller Punktmenge

- Abzählbare und überabzählbare Mengen
- Infimum und Supremum
- Limes superior und Limes inferior

### 6 Stetigkeit

- Definition und Beispiele
- Verkettung stetiger Funktionen
- Stetigkeit der Exponentialfunktion
- Stetige Funktionen auf beschränkten abgeschlossenen Intervallen

### 7 Streng monotone Funktionen

- Trigonometrische Funktionen

## Inhaltsverzeichnis IV

- Umkehrfunktionen streng monotoner Funktionen
- Exponential- und Logarithmusfunktion
- Exponentialfunktion und Logarithmus zur Basis  $a$
- Exponentielles und Logarithmisches Wachstum
- Arcusfunktionen und Polarkoordinaten
- Einheitswurzeln

### 8 Differentialrechnung

- Ableitung
- Geometrische und kinematische Interpretation
- Beispiele
- Affine Approximation
- Ableitungsregeln
- Ableitung der Umkehrfunktion und Kettenregel
- Lokale Extrema
- Mittelwertsatz und Folgerungen
- Grenzwertbestimmung nach L'Hospital

## Inhaltsverzeichnis V

- Monotonie und Ableitung
- Höhere Ableitungen
- Taylor-Entwicklung
- Funktionen mit Werten in  $\mathbb{R}^n$

### 9 Integralrechnung

- Treppenfunktionen
- Ober- und Unterintegral
- Integrierbare Funktionen
- Riemannsche Summen
- Fundamentalsatz der Differential- und Integralrechnung
- Integration durch Substitution
- Partielle Integration

# Einige Bemerkungen vorneweg

## Semesterbetrieb

- Zentraler Kommunikationsweg ist die Webseite  
<http://www.math.uni-hamburg.de/master/lehrexport/physik/mfp1.html>  
Diese Seite sollten Sie mindestens 2x wöchentlich aufmerksam lesen, insbesondere bevor Sie Fragen per e-Mail stellen.
- Probleme bei STINE kann nur klären:  
<https://support.rrz.uni-hamburg.de/stine/>
- Fragen im Übungsbetrieb, die *nicht* dadurch abgedeckt sind:  
Dr. Ralf Holtkamp  
(<http://www.math.uni-hamburg.de/home/holtkamp/>)
- Für jede Modulprüfung müssen Sie sich bis 3 Tage vor der Prüfung über StINE oder per Formular direkt im Studienbüro Physik anmelden. Nur dann kann die Prüfung absolviert und bewertet werden. Nachträgliche Anmeldungen und Abmeldungen sind **ausgeschlossen**.

# Aussagen

Wir werden zwei undefinierte Begriffe verwenden: "Aussage" und "Menge". Das Folgende ist also gar keine Definition.

## Definition

Eine **Aussage** ist ein sprachliches Gebilde, dem man auf sinnvolle Weise einen Wahrheitswert, und zwar entweder "w(ahr)" oder "f(alsch)", zuordnen kann.

- Die Aussage " $1^2 + 2^2 = 3^2$ " hat Wahrheitswert (f).
- Quadrate reeller Zahlen sind nie negativ. (w)
- " $x^2 = 2$ " mit einem bestimmten  $x$  oder durch Ergänzung "Es gibt ein  $x \in \mathbb{Q}$  mit  $x^2 = 2$  zu einer Aussage.

# 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

Implikation  $\Rightarrow$

$A$	$B$	$A \Rightarrow B$
w	w	w
w	f	f
f	w	w
f	f	w

“Wenn Ptolemäus recht hat, ist die Erde eine Scheibe.” ist wahr.

“Äquivalenz”  $\Leftrightarrow$

$A$	$B$	$A \Leftrightarrow B$
w	w	w
w	f	f
f	w	f
f	f	w

# Tautologien

## Definition

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

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

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

A	B	$\neg A$	$\neg B$	$A \Rightarrow B$	$\neg B \Rightarrow \neg A$
w	w	f	f	w	w

ÜA: Vervollständigen Sie die Wahrheitstafel!

# Mengen

Der zweite Begriff, den wir naiv verwenden, ist der einer Menge. Auch das Folgende ist *keine* Definition:

## Definition (Georg Cantor, 1845 - 1918)

Eine **Menge** ist eine Zusammenfassung bestimmter wohlunterschiedener Objekte unserer Anschauung oder unseres Denkens - welche die **Elemente** der Menge genannt werden - zu einem Ganzen.

- Zentrale Eigenschaft einer Mengen  $M$  ist, dass für jedes Objekt  $x$  “unserer Anschauung” die Beziehung “Enthalten sein”  $x \in M$  eine *Aussage* ist, die wahr oder falsch sein kann.
- Man kann Mengen durch Auflistung angeben. Allerdings ist

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

*keine* Definition der natürlichen Zahlen. (Je nach Autor gehört die 0 zu  $\mathbb{N}$  oder nicht!)

# Teilmengen

## Definition

Zwei Mengen  $M, N$  heißen gleich, wenn  $(x \in M) \Leftrightarrow (x \in N)$  gilt.

Seien  $M, N$  Mengen. Dann heißt  $N$  **Teilmenge** (oder **Untermenge**) einer Menge  $M$  genau dann, wenn

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

d.h. wenn jedes Element von  $N$  auch Element von  $M$  ist. Man schreibt  $N \subset M$ .

Insbesondere gilt  $\emptyset \subset M$  und  $M \subset M$  für jede Menge  $M$ . Aus  $x \in A$  folgt  $\{x\} \subset A$ . Die Menge  $\{x\}$  sollte aber nicht mit dem Element  $x$  verwechselt werden: eine Schachtel mit Hut ist etwas anderes als ein Hut!

## Definition

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

# Mengenoperationen

Wir schreiben:

| oder : für "für die gilt"

:= (bzw. :  $\iff$ ) für „nach Definition gleich(bedeutend)“

## 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**,
- $M \times N = \{(p, q) \mid p \in M \text{ und } q \in N\}$  das **kartesische Produkt**

von  $M$  und  $N$ .

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

### Beispiele:

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

# 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),
- $M \cap N \subset M \subset M \cup N$ ,
- $(M \cap N) \cap L = M \cap (N \cap L)$  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).

## 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  $N \subset M$ . 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:

- Die geraden natürlichen Zahlen  $2\mathbb{N} \subset \mathbb{N}$  sollte man nicht mit ... auflisten:

$$2\mathbb{N} = \{0, 2, 4, 6, \dots\}$$

sondern so schreiben:

$$2\mathbb{N} = \left\{ x \in \mathbb{N} \mid \frac{x}{2} \in \mathbb{N} \right\}.$$

Es gilt  $x \in 2\mathbb{N} \iff \exists y \in \mathbb{N} : x = 2y$ .

# Mathematische Aussagen und Beweise

Mathematische Aussagen werden in der Regel mit **Theorem**, **Satz**, **Lemma** (Hilfssatz) und **Folgerung** (Korollar) bezeichnet und müssen bewiesen werden.

Es gibt eine Vielzahl von Beweismethoden, z.B.

- **direkter Beweis** (z.B. Umformen von Gleichungen),  
z.B. „ $4x = 12 \Rightarrow x = 3$ “
- **indirekter Beweis** (auch Beweis durch Widerspruch): beruht auf der Kontraposition  $(A \Rightarrow B) \iff (\neg B \Rightarrow \neg A)$ , wobei  $A$  die Voraussetzung ist und  $B$  die Konklusion.  
Z.B. „ $\exists$  unendlich viele Primzahlen“ kann man indirekt beweisen.
- Aussagen, die von einer natürlichen Zahl abhängen, kann man mit vollständiger Induktion beweisen. Dies beruht auf der folgenden Eigenschaft natürlicher Zahlen: sei  $M \subset \mathbb{N}$  eine Teilmenge mit  $0 \in M$  und mit  $n \in M$  liegt auch der Nachfolger  $n' := n + 1 \in M$ . Dann ist  $M = \mathbb{N}$ .

# 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  $n \in \mathbb{N}$  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  $n = 1$ : nämlich gilt  $\sum_{i=1}^1 i = 1 = \frac{1}{2} \cdot 1 \cdot (1+1)$

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

und beweist dann die **IB:**

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

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



# Ein weiteres Beispiel

## Satz (Geometrische Summenformel)

*Es gilt die Formel*

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

**Beweis.** ÜA



## 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  $f(0)$ .
- 2 Eine Vorschrift  $F_n$ , die für jedes  $n$  in  $\mathbb{N}$  die Zahl  $f(n)$  aus den Zahlen  $f(0), \dots, f(n-1)$  berechnet, d.h.

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

### Beispiele:

- 1 **Potenzen:**  $x^0 := 1$ ,  $x^n := x \cdot x^{n-1}$
- 2 **Fibonacci Zahlen:**  
 $x_0 := 0$ ,  $x_1 := 1$ ,  $x_n := x_{n-1} + x_{n-2}$  ( $n \geq 2$ )
- 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:**  $0! := 1$ ,  $n! := (n-1)! \cdot n$

## 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  $n!$ .

**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 können  $O_1$  auch zweiter, dritter,  $\dots$ ,  $n + 1$ -ter Stelle anordnen und haben wieder  $n!$  Anordnungen der restlichen Objekte. Die Gesamtzahl aller Anordnungen von  $O_1, \dots, O_{n+1}$  ist also  $n! \cdot (n + 1) = (n + 1)!$ . □

# Abbildungen

## Definition

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

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

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

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

- $f(X) := \{f(a) \mid a \in X\}$  heißt **Bild** der Menge  $X \subset A$  unter  $f$ .
- $f^{-1}(Y) := \{a \in A \mid f(a) \in Y\}$  heißt **Urbild** der Menge  $Y \subset B$  unter  $f$ .

## Definition

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

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

## Definition (Eigenschaften von Abbildungen)

Sei  $f : A \rightarrow B$  eine Abbildung (zwischen den Mengen  $A$  und  $B$ ).

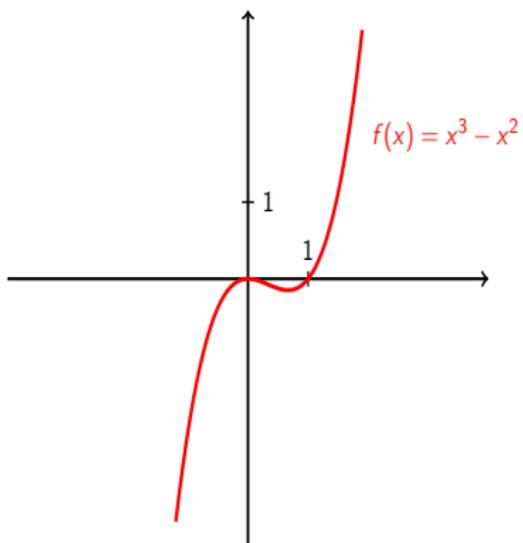
$$\begin{aligned} f \text{ hei\u00dft } \mathbf{surjektiv} & : \iff f(A) = B \\ & \iff \forall b \in B \exists a \in A : f(a) = b \end{aligned}$$

$$\begin{aligned} f \text{ ist } \mathbf{injektiv} & : \iff \text{jedes Bild hat genau ein Urbild} \\ & \iff \forall b \in f(A) \exists! a \in A : f(a) = b \\ & \iff \forall x, y \in A : f(x) = f(y) \Rightarrow x = y \end{aligned}$$

$$f \text{ ist } \mathbf{bijektiv} : \iff f \text{ ist injektiv } \mathbf{und} \text{ surjektiv}$$

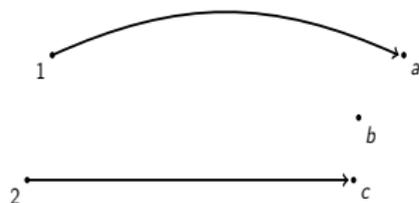
## Beispiele.

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

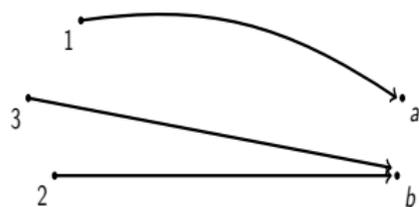


## Veranschaulichung

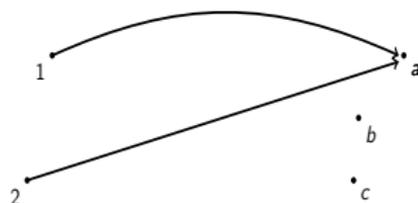
Injektive Abbildung einer zweielementigen Menge auf eine Menge mit drei Elementen



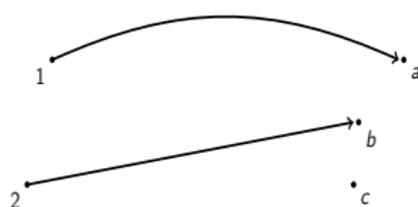
Surjektive Abbildung



Nicht injektive Abbildung

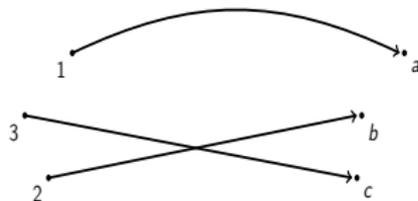


Nicht surjektive Abbildung

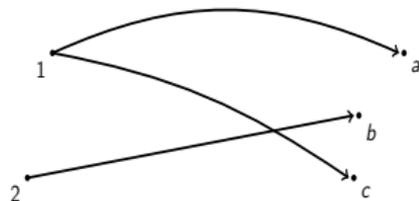


# Veranschaulichung II

Bijektive Abbildung



Keine Abbildung

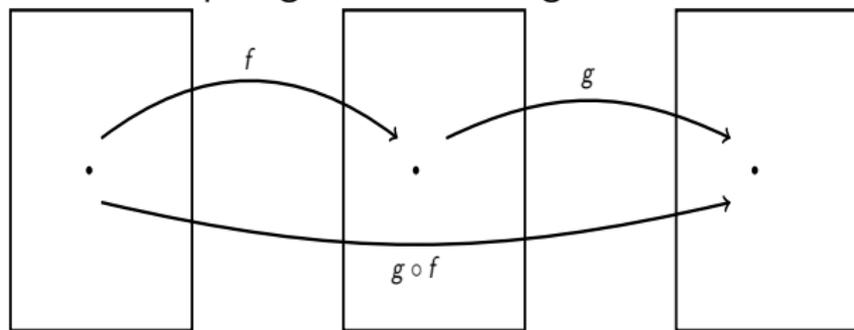


# Komposition von Abbildungen

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

$$g \circ f : A \rightarrow C, \quad a \mapsto g \circ f(a) := g(f(a))$$

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



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

## Lemma

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

$$g \circ f = \text{Id}_A \quad \text{und} \quad f \circ g = \text{Id}_B$$

**Beweis.** ÜA □

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

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

# Reelle und komplexe Zahlen

## Übersicht:

- Wiederholung: Natürliche, ganze und rationale Zahlen
- Reelle Zahlen
  - ▶ Körperaxiome (Rechenregeln)
  - ▶ Ordnungsaxiome
- Komplexe Zahlen
  - ▶ Real- und Imaginärteil, Konjugation
  - ▶ Betrag und Abstand

## 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:

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

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

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

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**

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

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.

## Warum reelle Zahlen?

Schon den Griechen in der Antike war bekannt, dass es irrationale Größen gibt, die zum Beispiel auf natürliche Art und Weise als Längen in einfachen geometrischen Figuren auftauchen. Die **reellen Zahlen**  $\mathbb{R}$  beheben dieses Manko der rationalen Zahlen.

### Beispiel

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.

Widerspruch zur Teilerfremdheit! □

# 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

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

- und Multiplikation

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

welche die folgenden Axiome erfüllen:

## A) Axiome der Addition

- **Kommutativgesetz:**

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

- **Assoziativgesetz:**

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

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

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

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

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

## M) Axiome der Multiplikation

- **Kommutativgesetz:**

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

- **Assoziativgesetz:**

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

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

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

- 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:

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

### 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. ( $\rightarrow$  Computer)
- Natürliche und ganze Zahlen mit  $+$  und  $\cdot$  sind **kein** Körper

## Folgerungen aus den Körperaxiomen

### Satz

Sei  $\mathbb{K}$  ein Körper (z.B.  $\mathbb{K} = \mathbb{R}$ ). Für alle  $a, b \in \mathbb{K}$  ist die Gleichung

$$a + x = b$$

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

**Beweis.** Das folgt allein aus A), also daraus, dass  $(\mathbb{K}, +)$  eine (abelsche) Gruppe ist. Die folgenden Aussagen sind äquivalent:

$$x + a = b$$

$$(x + a) + (-a) = b + (-a) \quad \text{wg. } \exists \text{ Inverses bez. Addition}$$

$$x + (a + (-a)) = b + (-a) \quad \text{wg. Assoziativität der Addition}$$

$$x + 0 = b + (-a) \quad \text{wg. invers bezüglich } +$$

$$x = b + (-a) \quad \text{da } 0 \text{ neutral bezüglich } +$$



## Satz

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.

**Beweis.** Das folgt ganz analog aus M), also der Tatsache, dass  $\mathbb{K} \setminus \{0\}$  eine Gruppe bezüglich der Multiplikation ist. Die Lösung ist

$$x = \frac{b}{a} := a^{-1}b.$$



## Weitere Folgerungen aus den Körperaxiomen

### Satz

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

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

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

( $\Rightarrow$ ) Es genügt zu zeigen:  $x \neq 0$  und  $xy = 0 \Rightarrow y = 0$ .

Da  $x \neq 0$ , hat die Gleichung  $xy = 0$  die eindeutige Lsg.

$$y = x^{-1} \cdot 0 = 0.$$

□

Weitere Folgerungen (ÜA):  $-0 = 0, -(-x) = x, (-x)y = -(xy),$   
 $(-x)(-y) = xy, 1^{-1} = 1, \forall x \neq 0$  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**  $\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\} \subset \mathbb{R}$ . Sie bilden keinen Körper (keine multiplikative Inverse, außer für  $\pm 1$ ).
- Der kleinste Körper, der  $\mathbb{Z}$  enthält, ist der **Körper der rationalen Zahlen**  $\mathbb{Q} = \{\frac{p}{q} \mid p \in \mathbb{Z}, q \in \mathbb{N}, q \neq 0\}$ .
- 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.

## 0) Ordnungsaxiome

In  $\mathbb{R}$  ist die Teilmenge  $\mathbb{R}_+ \subset \mathbb{R}$  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  $R \subset X \times X$ . Z.B. ist Gleichheit die Diagonale  $\{(x, x) \mid x \in X\} \subset X \times X$

Auf den reellen Zahlen sind somit die Relationen

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

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 die die folgenden Axiome gelten sollen:

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

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

(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  $nx := x + \dots + x$  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

## Satz

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

(i) Die Relation  $<$  auf  $\mathbb{R}$  ist **transitiv**:

$$x < y \text{ und } y < z \implies x < z$$

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

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

(iv)  $x < y$  und  $a < 0 \implies ax > ay$ .

**Beweis.** Wir beweisen exemplarisch die Transitivität. Wir schreiben

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

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

□

## Satz

- (i)  $x^2 > 0$  für alle  $x \neq 0$ ,
- (ii)  $x > 0 \implies x^{-1} > 0$ ,
- (iii)  $0 < x < y \implies x^{-1} > y^{-1} > 0$ .

## Beweis.

- (i) Aus  $x > 0$  folgt  $x^2 = x \cdot x > 0$ , wg. (O2). Aus  $x < 0$  folgt  $-x > 0$  und somit ebenfalls  $x^2 = (-x)^2 > 0$ .
- (ii) Für  $x > 0$  gilt wegen (i):

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

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

## Definition

- Das **offene Intervall** ist die Teilmenge

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

schließt also die Eckpunkte 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

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

## 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)  $|x - y| \geq ||x| - |y||$ .

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

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

# Abstand

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

## Satz

Die Abbildung  $d : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}_+ \cup \{0\}$  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  $|\cdot|$ .  $\square$

# 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

$$\mathbb{R}^2 = \{(x, y) \mid x, y \in \mathbb{R}\} \text{ („kartesische Ebene“)}$$

- 2 versehen mit der **Addition** durch Addition der Komponenten:

$$+ : \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}^2$$

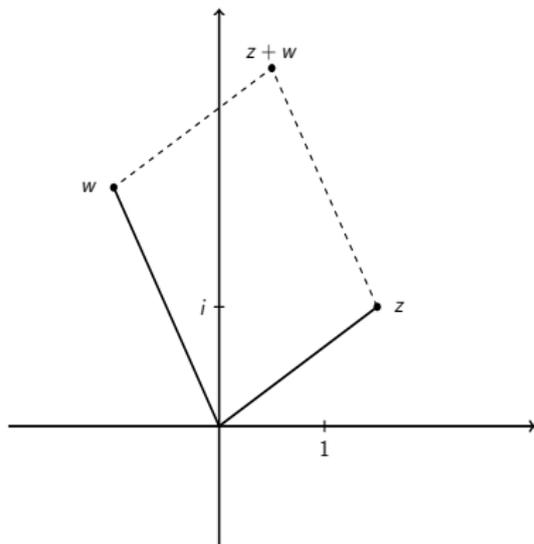
$$(x, y) + (u, v) := (x + u, y + v)$$

- 3 und der **Multiplikation**

$$\cdot : \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}^2$$

$$(x, y) \cdot (u, v) := (xu - yv, xv + yu).$$

# Veranschaulichung der Addition



[Animation]

## 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:

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

## 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:

$$(x, y) \cdot \left( \frac{x}{x^2 + y^2}, -\frac{y}{x^2 + y^2} \right) = \left( \frac{x^2}{x^2 + y^2} - \frac{y(-y)}{x^2 + y^2}, \frac{x(-y)}{x^2 + y^2} + \frac{yx}{x^2 + y^2} \right) = (1, 0).$$

D) **Distributivgesetz**: Auch durch explizites Nachrechnen, ÜA. □

## 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:

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

- 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*

$$z^n + c_{n-1}z^{n-1} + \dots + c_1z + c_0 = 0,$$

*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: spalte Linearfaktoren ab.

# 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

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

## 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

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

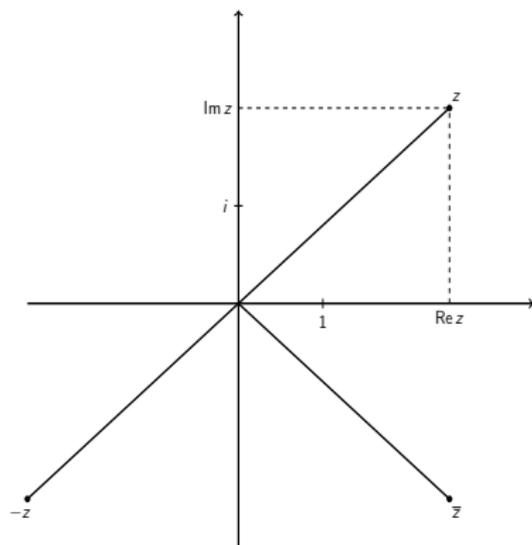
Für eine reelle Zahl  $x \in \mathbb{R} \subset \mathbb{C}$  ergibt sich als Betrag

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

genau der Absolutbetrag von  $x$ . Außerdem gilt:

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

## 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)  $|z - w| \geq ||z| - |w||$ .

**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

$$\operatorname{Re} z = x \leq \sqrt{x^2 + y^2} = |z|.$$

$$\operatorname{Im} z = y \leq \sqrt{x^2 + y^2} = |z|.$$

Somit

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

□

## Metrische Räume

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

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

Er hat wiederum 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)

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  $d(x, x) = 0 \forall x \in X$ .

## Definition

Sei  $X$  eine Menge (etwa  $\mathbb{C}, \mathbb{R}$ ). Eine **Folge** von Elementen in  $X$  ist eine Abbildung  $\mathbb{N} \rightarrow X, n \mapsto x_n$ .

Man schreibt dafür auch  $(x_n)_{n \in \mathbb{N}} = (x_0, x_1, \dots)$  oder einfach  $(x_n)$ .

$n$  heißt der **Index** von  $x_n$ . Wir lassen auch Folgen

$(x_n)_{n \geq n_0} = (x_{n_0}, x_{n_0+1}, \dots)$  zu, die erst ab einem gewissen Index  $n_0 \in \mathbb{N}$  definiert sind.

## Definition

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

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

Schreibweise:  $\lim_{n \rightarrow \infty} x_n = \ell$  oder auch  $x_n \xrightarrow{n \rightarrow \infty} \ell$

**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

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

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.  $\frac{1}{N} < \varepsilon$ .

Für alle  $n \geq N$  gilt dann  $|x_n - 0| = \frac{1}{n} \leq \frac{1}{N} < \varepsilon$ . □

Weitere Folgen mit unterschiedlichem Konvergenzverhalten als Animation auf der Webseite.

# 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  $l_1, l_2 \in X$  hat, d.h.  $l_1 \neq l_2$ . Setze  $\varepsilon := \frac{d(l_1, l_2)}{2} > 0$ .
- Wegen der Konvergenz existieren natürliche Zahlen  $N_1$  und  $N_2$  mit

$$d(x_n, l_1) < \varepsilon \quad \text{für alle } n \geq N_1,$$

$$d(x_n, l_2) < \varepsilon \quad \text{für alle } n \geq N_2.$$

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

$$2\varepsilon = d(l_1, l_2) \leq d(l_2, x_n) + d(x_n, l_1) < 2\varepsilon.$$

- D.h.  $2\varepsilon < 2\varepsilon$ , **Widerspruch!** Also muss  $l_1 = l_2$  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

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

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

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

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

## Einfache ÜA:

Eine Folge  $(x_n)$  reeller Zahlen ist genau dann beschränkt, 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  $\ell = \lim_{n \rightarrow \infty} x_n$ . Dann existiert (etwa zu  $\varepsilon = 1$ ) eine natürliche Zahl  $N \in \mathbb{N}$ , so dass

$$d(x_n, \ell) = |x_n - \ell| < 1 \quad \text{für alle } n \geq N.$$

- 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  $n \in \mathbb{N}$ :

$$|x_n| \leq \max\{|x_1|, |x_2|, \dots, |x_{N-1}|, |\ell| + 1\} =: K.$$

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:  $\lim_{n \rightarrow \infty} (x_n) = \ell$ .

Für  $\varepsilon := \frac{1}{2}$  existiert dann ein  $N \in \mathbb{N}$ :  $|x_n - \ell| < \frac{1}{2} \quad \forall n \geq N$ .

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  $(x_n) = (1, -1, 1, -1, 1, \dots)$ .

□

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)  $\lim_{n \rightarrow \infty} (x_n + y_n) = \lim_{n \rightarrow \infty} x_n + \lim_{n \rightarrow \infty} y_n$
- (ii)  $\lim_{n \rightarrow \infty} (x_n y_n) = \left( \lim_{n \rightarrow \infty} x_n \right) \cdot \left( \lim_{n \rightarrow \infty} y_n \right).$

**Beweis.** Sei  $x = \lim_{n \rightarrow \infty} x_n$  und  $y = \lim_{n \rightarrow \infty} y_n$ .

(i) Es gibt zu jedem  $\varepsilon > 0$  natürliche Zahlen  $N_1, N_2 \in \mathbb{N}$ , so dass

$$|x_n - x| < \varepsilon/2 \quad \text{für alle } n \geq N_1,$$

$$|y_n - y| < \varepsilon/2 \quad \text{für alle } n \geq N_2.$$

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

$$|x_n - x| < \frac{\varepsilon}{2K} \quad \text{für alle } n \geq N_1,$$

$$|y_n - y| < \frac{\varepsilon}{2K} \quad \text{für alle } n \geq N_2.$$

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

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

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



## 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

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

**Beweis.** ÜA



# Konvergenz reeller Folgen und Ordnungsrelation

## Satz

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

$$\lim_{n \rightarrow \infty} x_n \leq \lim_{n \rightarrow \infty} y_n.$$

**Beweis.** Übung □

## Achtung:

Für konvergente Folgen mit  $x_n < y_n$  gilt nicht unbedingt

$$\lim_{n \rightarrow \infty} x_n < \lim_{n \rightarrow \infty} y_n.$$

### Gegenbeispiel:

$x_n := 0$  die konstante Folge,  $y_n := \frac{1}{n}$ . Dann gilt in der Tat

$$\lim_{n \rightarrow \infty} x_n = 0 \leq \lim_{n \rightarrow \infty} y_n = 0,$$

**aber** keine strikte Ungleichung.

# 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

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

## Satz

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

**Beweis.** Sei  $\ell = \lim_{n \rightarrow \infty} x_n$ . 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. Dies charakterisiert sogar  $\mathbb{R}$ : jeder vollständige archimedisch angeordnete Körper  $\mathbb{K}$  ist isomorph zu  $\mathbb{R}$ , also  $\mathbb{K} \cong \mathbb{R}$ .

# 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  $(x_n)$ , wenn es eine streng monoton wachsende Folge  $\varphi : \mathbb{N} \rightarrow \mathbb{N}$  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.** Einfache Übung



# 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

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

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

- (i) das Intervall  $[A_k, B_k]$  unendlich viele Glieder der Folge  $(x_n)$  enthält,
- (ii)  $[A_k, B_k] \subset [A_{k-1}, B_{k-1}]$ , 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  $k \in \{0, \dots, \ell\}$  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.

⇒ 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,  $y_0, \dots, y_k$  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.  $\square$

## Beweis der Behauptung.

- Zum Beweis der Behauptung benutzen wir folgende Tatsache:

$$\lim_{n \rightarrow \infty} \frac{1}{2^n} = 0 \quad (\text{denn für } n \geq 1 \text{ gilt } 0 < \frac{1}{2^n} < \frac{1}{n} \xrightarrow{n \rightarrow \infty} 0).$$

- 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  $\ell := \lim_{k \rightarrow \infty} x_{n_k}$  konvergiert.

## Weiter im Beweis:

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

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

- 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

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

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:

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

## Weiter im Beweis:

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

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

- 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  $\ell^2 = a$ .



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



# 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. □

## 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

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

$|\cdot|$  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

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

- 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

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



## 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

$$\left| \lim_{n \rightarrow \infty} z_n \right| = \lim_{n \rightarrow \infty} |z_n|.$$

(Das folgt aus  $||z_n| - |z|| \leq |z_n - z|$  mit  $z = \lim_{n \rightarrow \infty} z_n$ .)

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  $z_n = (-1)^n$ .

- Es gilt jedoch:  $\lim_{n \rightarrow \infty} |z_n| = 0 \implies \lim_{n \rightarrow \infty} z_n = 0$ .

## 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  $\lim_{n \rightarrow \infty} z^n = 0$ .*
- 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  $n = 2, 3, \dots$  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  $n + 1$ :

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



## Folgerung (aus der Bernoullischen Ungleichung)

Es sei  $t \in \mathbb{R}$ ,  $t > 0$ . Weiter seien  $K \in \mathbb{R}$  und  $\epsilon \in \mathbb{R}_+$  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  $t^n < \epsilon$ .

**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  $0 < t < 1$  wende (i) auf  $\frac{1}{t} > 1$  an. □

**Beweis des Satzes.** Falls  $|z| < 1$  findet man für jedes  $\epsilon > 0$  ein  $N$  mit  $|z^n| = |z|^n < \epsilon$  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. □

# 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

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

## 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 Die **harmonische Reihe**  $s_n = \sum_{k=1}^n \frac{1}{k}$  ist nicht konvergent, die **alternierende harmonische Reihe**

$s_n = \sum_{k=1}^n \frac{(-1)^k}{k}$  schon, und zwar gegen  $-\ln 2$  (wie wir später sehen werden).

## 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 einfach, 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

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

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  $\left| \sum_{k=0}^n z_k \right| \leq \sum_{k=0}^n |z_k| \leq \sum_{k=0}^n a_k$  durch Grenzübergang, der ja Ungleichungen erhält. □

## 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  $z$ )

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

Sei nun  $|z| < 1$ . Dann gilt  $\lim_{n \rightarrow \infty} z^{n+1} = 0$ .

Wegen (1) konvergiert dann die geometrische Reihe:

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

**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.**

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

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

### 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

$$|a_{k+1}| \leq t|a_k| \quad \text{für alle } k. \quad (2)$$

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**.

Beachte, 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

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



## Definition

Die **Exponentialfunktion** ist die durch die Exponentialreihe definierte Abbildung

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

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

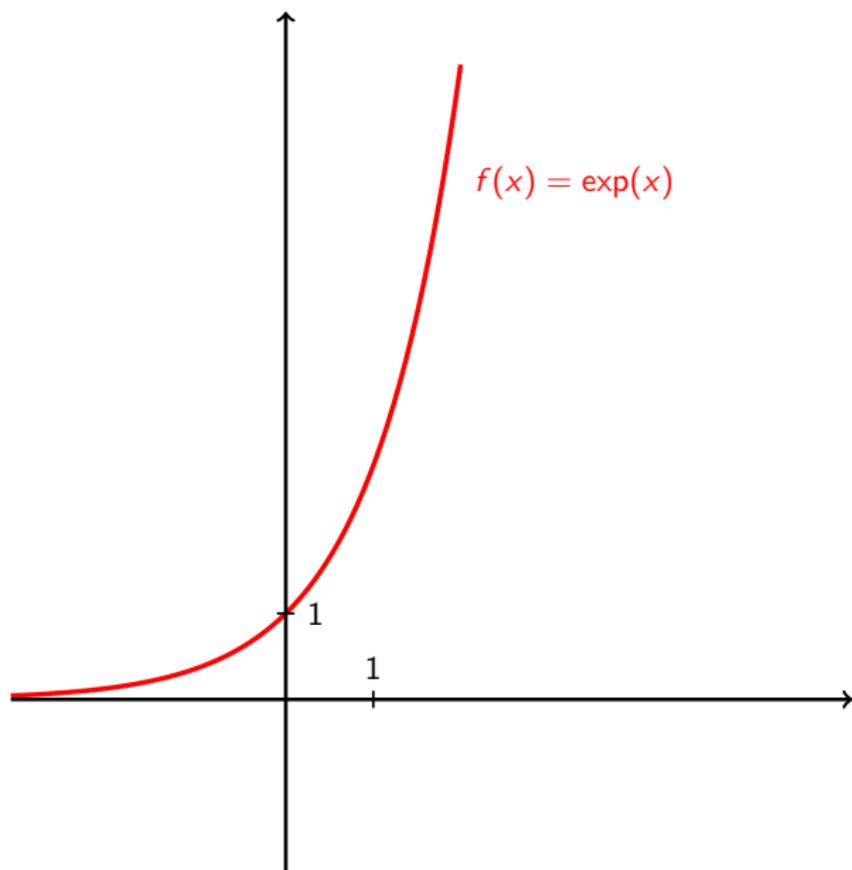
## Definition

Man definiert die **Eulersche Zahl**

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

$$e \simeq 2,7182818284590452353602874713526624977572470936999595 \dots$$

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



Seien  $p_n = \sum_{k=0}^n a_k$  und  $q_n = \sum_{k=0}^n b_k$  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

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

und

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

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

### Definition (Cauchy Produkt)

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

## 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

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

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

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

## Weiter im Beweis

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

$$\begin{aligned}
 \left| \sum_{k=0}^n c_k - \left(\sum_{k=0}^n a_k\right) \cdot \left(\sum_{k=0}^n b_k\right) \right| &= \left| \sum_{0 \leq i, j \leq n, i+j \leq n} a_i b_j - \sum_{0 \leq i, j \leq n} a_i b_j \right| \\
 &= \left| \sum_{0 \leq i, j \leq n, i+j > n} a_i b_j \right| \leq \sum_{0 \leq i, j \leq n, i+j > n} |a_i| |b_j| \\
 &\leq \sum_{k=n+1}^{\infty} |c_k| = \underbrace{\sum_{k=0}^{\infty} |c_k| - \sum_{k=0}^n |c_k|}_{\lim_{n \rightarrow \infty} (\quad) = 0}
 \end{aligned}$$

# 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  $n$ . □

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:

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

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

$$\exp(z) \exp(-z) = \exp(z - z) = \exp(0) = 1. \quad \square$$

## Bemerkungen:

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

$$\lim_{n \rightarrow \infty} \bar{z}_n = \overline{\lim_{n \rightarrow \infty} z_n}.$$

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

## Definition (Sinus und Cosinus)

Für  $x \in \mathbb{C}$  definiert man

$$\cos x := \operatorname{Re} \exp(ix) \quad \text{und}$$

$$\sin x := \operatorname{Im} \exp(ix), \quad x \in \mathbb{C}.$$

## 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.

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

$$\begin{aligned} (\cos x)^2 + (\sin x)^2 &= |\exp(ix)|^2 = \exp(ix) \overline{\exp(ix)} \\ &= \exp(ix) \exp(-ix) \stackrel{\text{Funktionalgl.}}{=} \exp(ix - ix) \\ &= \exp(0) = 1. \quad \square \end{aligned}$$

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

- (Absolut konvergente Potenzreihenentwicklung)

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

- (Symmetrie, resp. Antisymmetrie)

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

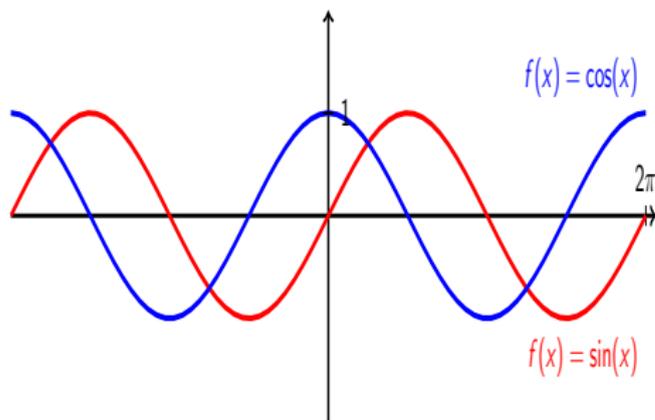
Cosinus ist gerade, sinus ist ungerade Funktion.

# Additionstheoreme

(Additionstheoreme)

$$\cos(x + y) = \cos x \cos y - \sin x \sin y,$$

$$\sin(x + y) = \sin x \cos y + \cos x \sin y.$$



## Ausblick: Ebene Polarkoordinaten

- Wir werden später noch sehen, dass die Funktion  $\mathbb{R} \ni \varphi \mapsto \exp(i\varphi) \in S^1 \subset \mathbb{C}$ , die Kreislinie periodisch (gegen den Uhrzeigersinn) durchläuft.
- Die Periode ist  $2\pi$ , wobei die reelle Zahl  $\pi = 3,14159\dots$  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  $S^1 = \{z \in \mathbb{C} \mid |z| = 1\}$ .
- Demensprechend hat jede komplexe Zahl  $z \in \mathbb{C}^* := \mathbb{C} \setminus \{0\}$  eine eindeutige Darstellung der Form

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

- 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.** Berechne  $z^{20}$  für  $z = (1 + i)$ .

- Die Polarkoordinaten von  $z$  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  $\epsilon > 0$  ein  $N \in \mathbb{N}$ , so dass  $|z_n| = \left| \sum_{k=n}^n z_k \right| < \epsilon$  für alle  $n \geq N$ .  $\square$

### Bemerkung

Das in der Proposition formulierte Konvergenzkriterium ist notwendig aber nicht hinreichend. Beispielsweise ist, wie wir noch sehen werden, die sogenannte **harmonische Reihe**

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

divergent, obwohl die Folge  $(\frac{1}{n})_{n \in \mathbb{N}}$  eine Nullfolge ist.

## Beispiel.

Die **alternierende harmonische Reihe**

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

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

### Satz (Leibnizkriterium für alternierende Reihen)

Sei  $a_k \geq 0$  ( $k = 0, 1, 2, \dots$ ) 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 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  $x = \lim_{n \rightarrow \infty} x_n$  und  $y = \lim_{n \rightarrow \infty} y_n$ .
- Wegen  $\lim_{n \rightarrow \infty} (y_n - x_n) = \lim_{n \rightarrow \infty} (-a_{2n+1}) = 0$ , gilt  $x = y$ .
- Daraus folgt  $\lim_{n \rightarrow \infty} s_n = \lim_{n \rightarrow \infty} s_{2n} = \lim_{n \rightarrow \infty} s_{2n+1}$ . □

Zunächst ein Hilfssatz:

### Satz

Es sei  $(a_n)_{n \in \mathbb{N}}$  eine monoton fallende Folge positiver 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:

$$s_n = a_1 + a_2 + \dots + a_n,$$

$$t_k = a_1 + 2a_2 + \dots + 2^k a_{2^k}.$$

## Weiter im Beweis.

Für  $n \leq 2^k$  gilt wegen  $a_n \geq a_{n+1}$ :

$$\begin{aligned} s_n &\leq a_1 + (a_2 + a_3) + (a_4 + a_5 + a_6 + a_7) + \dots + \overbrace{(a_{2^k} + \dots + a_{2^{k+1}-1})}^{2^k \text{ viele}} \\ &\leq a_1 + 2a_2 + 4a_4 + \dots + 2^k a_{2^k} \\ &= t_k \end{aligned}$$

Andererseits gilt für  $n \geq 2^k$ :

$$\begin{aligned} s_n &\geq a_1 + a_2 + (a_3 + a_4) + (a_5 + a_6 + a_7 + a_8) + \dots + (a_{2^{k-1}+1} + \dots + a_{2^k}) \\ &\geq \frac{1}{2}a_1 + a_2 + 2a_4 + 4a_8 \dots + 2^{k-1}a_{2^k} \\ &= \frac{1}{2}t_k \end{aligned}$$

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

## Folgerung

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

Insbesondere konvergiert die harmonische Reihe  $\sum_{k=1}^n \frac{1}{k}$  nicht.

**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:

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

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

# 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  $\mathbb{N} \rightarrow \mathbb{N}$  gegeben durch  $n \rightarrow k_n$ ). 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.

## Beispiel.

$$S_1 = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \dots - \frac{1}{2j} + \frac{1}{2j+1} - \dots,$$

$$S_2 = 1 + \frac{1}{3} - \frac{1}{2} + \frac{1}{5} + \frac{1}{7} - \frac{1}{4} + \dots + \frac{1}{4j-3} + \frac{1}{4j-1} - \frac{1}{2j} + \dots,$$

Die Reihe  $S_2$ , in der auf jeweils zwei positive ungerade Glieder ein negatives gerades folgt, ist eine Umordnung von  $S_1$ . Nach dem Leibnizkriterium konvergiert die alternierende harmonische Reihe  $S_1$ .

**Weiter im Beispiel.**  $S_2$  kann man schreiben als  $\sum_{k=1}^n (-1)^{k+1} a_k$  mit

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

D.h.  $a_k$  ist monoton fallend, denn

$$a_{2j+1} < \frac{1}{2j} = a_{2j} < \frac{1}{2j-1} < a_{2j-1}$$

Daher konvergiert auch die Reihe  $S_2$  nach dem Leibnizkriterium. **Aber:**

$$S_1 < 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} = \frac{47}{60}$$

denn: was in  $S_1$  folgt sind Summanden  $-\frac{1}{2j} + \frac{1}{2j+1} < 0$ ;

$$S_2 > 1 + \frac{1}{3} - \frac{1}{2} + \frac{1}{5} + \frac{1}{7} - \frac{1}{4} = \frac{47}{60} + \frac{1}{7},$$

denn alle folgenden Summanden erfüllen  $\frac{1}{4j-3} + \frac{1}{4j-1} - \frac{1}{2j} > 0$ .

## Wichtige Folgerung

In allgemeinen Reihen kommt es auf die Reihenfolge der Summationen an.

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  $\epsilon > 0$  ein  $N \in \mathbb{N}$  mit

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

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  $0, 1, \dots, N$  in der Menge  $k_0, k_1, \dots, k_p$  enthalten sind. Es folgt für  $n \geq p$ , dass in der Differenz der Partialsummen  $s_n - s'_n$  sich die Werte  $a_0, \dots, a_N$  gegenseitig aufheben, also

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

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

Bemerkenswerterweise gilt:

### 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

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

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

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

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

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

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

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

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

$$P_0 + \dots + P_{m_0} - Q_0 - \dots - Q_{k_0} + P_{m_0+1} + \dots + P_{m_1} > s,$$

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

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

$$|P_0 + \dots + P_{m_0} - Q_0 - \dots + P_{m_n} - s| < P_{m_n},$$

$$|P_0 + \dots + P_{m_0} - Q_0 - \dots - Q_{k_n} - s| < Q_{k_n}$$

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.

□

# Abzählbare Mengen

## Definition

Eine nicht-leere Menge  $A$  heißt **endlich**, wenn es ein  $n \in \mathbb{N}$  und eine surjektive Abbildung  $\varphi : \{0, 1, \dots, n\} \rightarrow A$  gibt.

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

## Beispiele.

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

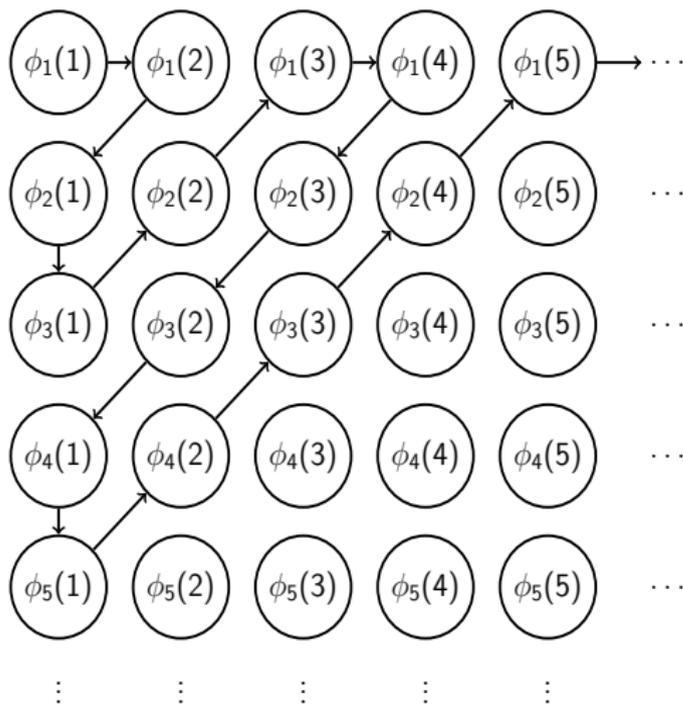
## Satz

Die Vereinigung von abzählbar vielen abzählbaren Mengen ist abzählbar.

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

## Diagonalverfahren

$M = \cup_{i \in \mathbb{N}} M_i$  mit Surjektionen  $\phi_i : \mathbb{N} \rightarrow M_i$  für alle  $i \in \mathbb{N}$ . Wir konstruieren  $\phi : \mathbb{N} \rightarrow M$  durch



## Abzählbarkeit von $\mathbb{Q}$

### Folgerung

$\mathbb{Q}$  ist abzählbar.

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

### Folgerung

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

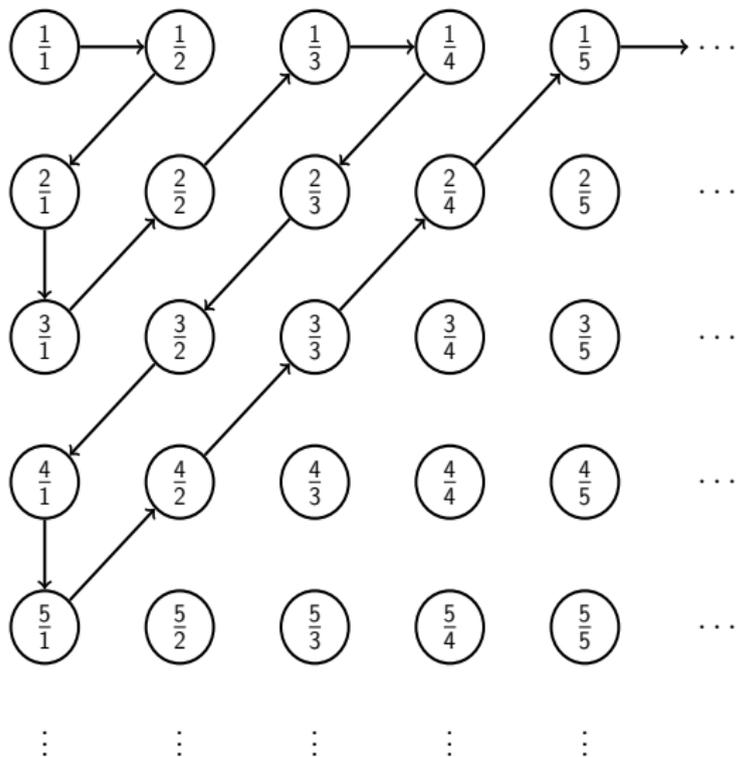
### Beweis.

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 von abzählbar vielen abzählbaren Mengen.

Diagonalverfahren für  $\mathbb{Q}$ 

# Überabzählbarkeit von $\mathbb{R}$

## Satz

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

**Beweis.** Man nimmt an, die Menge  $\mathbb{R}$  reellen Zahlen sei abzählbar,  $\mathbb{R} = \{x_0, x_1, x_2, \dots\}$ , und konstruiert eine reelle Zahl  $x$  mit  $x \neq x_n$  für alle  $n \in \mathbb{N}$ .

- 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  $I_k \subset I_{k-1}$  der Länge  $\frac{1}{3^k}$  mit  $x_0, \dots, x_n \notin I_n$ .

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

# Supremumseigenschaft der reellen Zahlen

## Definition (Supremum und Infimum)

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

$$x \leq s \quad (\text{bzw. } s \leq x) \quad \text{für alle } x \in M$$

nennt man eine **obere** (bzw. **untere**) **Schranke** für  $M$ . Eine obere Schranke  $s \in \mathbb{R}$  einer Menge  $M \subset \mathbb{R}$  heißt **Supremum** von  $M$ , falls  $s$  die kleinste obere Schranke ist, d.h. für jede obere Schranke  $s'$  von  $M$  gilt  $s' \geq s$ . 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  $M \subset \mathbb{R}$  besitzt ein Supremum.*

*Jede nach unten beschränkte, nicht-leere Teilmenge  $M \subset \mathbb{R}$  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.)

# 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  $a_0 \leq b_0, \dots, a_n \leq b_n$  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  $d \in M$  mit  $d > m$  und wir setzen dann  $[a_{n+1}, b_{n+1}] := [d, b_n] \subset [m, b_n]$ .

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  $c = \lim_{n \rightarrow \infty} b_n$  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  $c = \lim_{n \rightarrow \infty} a_n$  gäbe es dann  $N \in \mathbb{N}$ , mit  $b < a_n$  für alle  $n \geq N$ .
- Das widerspricht der Tatsache, dass  $a_n \in M$  und  $c$  obere Schranke ist.

# Maximum und Minimum

## Definition

- Sei  $M \subset \mathbb{R}$  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  $M \subset \mathbb{R}$ , falls  $\inf M \in M$  gilt,  $\min M := \inf M$  auch das **Minimum** der Menge  $M$ .

## Notation

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

# Limes superior, Limes inferior

## Definition

Sei  $(x_n)_{n \in \mathbb{N}}$  eine Folge reeller Zahlen. Dann ist  $\sup\{x_k \mid k \geq n\}$  entweder immer gleich  $\infty$  oder bildet eine beschränkte monoton fallende Folge. Daher definiert man

- $\limsup_{n \rightarrow \infty} x_n := \lim_{n \rightarrow \infty} (\sup\{x_k \mid k \geq n\}) \in \mathbb{R} \cup \{\infty\}$

mit  $\infty := \lim_{n \rightarrow \infty}(\infty)$ . Ähnlich ist  $\inf\{x_k \mid k \geq n\}$  entweder immer gleich  $-\infty$  oder bildet eine beschränkte monoton wachsende Folge. Daher definiert man

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

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

**Bemerkung.** Eine Folge  $(x_n)_{n \in \mathbb{N}}$  reeller Zahlen konvergiert genau dann gegen  $\ell \in \mathbb{R}$ , wenn  $\limsup_{n \rightarrow \infty} x_n = \liminf_{n \rightarrow \infty} x_n = \ell$ .

# Stetigkeit

## Definition (Folgenkriterium der Stetigkeit)

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

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

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

Man kann diese Definition offensichtlich auf die metrischen Räume  $\mathbb{R}$  und  $\mathbb{C}$  anwenden.

## Erste Beispiele.

Sei  $D \subset \mathbb{C}$  und  $f : D \rightarrow \mathbb{C}$ ,  $z \mapsto c$  eine konstante Funktion. Dann ist  $f$  offensichtlich stetig.

Für  $x$  in  $\mathbb{R}$  bezeichne  $\lfloor x \rfloor$  die größte ganze Zahl  $\leq x$ . Dann ist die Funktion  $\mathbb{R} \rightarrow \mathbb{R}$ ,  $x \mapsto \lfloor x \rfloor$  genau in den Punkten  $p \in \mathbb{R} \setminus \mathbb{Z}$  stetig.

## Definition (Grenzwert einer Funktion)

Sei  $f : D \rightarrow \mathbb{C}$  eine Funktion. Wir schreiben

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

falls es

- (i) erstens eine Folge  $(z_n) \in D \setminus \{p\}$  gibt, die gegen  $p$  konvergiert und
- (ii) zweitens  $\lim_{n \rightarrow \infty} f(z_n) = q$  für **jede** solche Folge  $(z_n)$  gilt.

Man nennt  $q$  den **Grenzwert** der Funktion  $f(z)$  für  $z$  gegen  $p$ .

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

## Bemerkung

Mit obiger Notation gilt dann:

$$f \text{ ist stetig in } p \in D \iff \begin{cases} \lim_{z \rightarrow p} f(z) \text{ existiert} \\ \lim_{z \rightarrow p} f(z) = f(p). \end{cases}$$

## Aus den Rechenregeln für Grenzwerte von Folgen folgt:

- (i) Die Funktion  $z \mapsto z$  ist stetig.
- (ii) Die Funktionen  $z \mapsto \bar{z}, |z|, \operatorname{Re} z, \operatorname{Im} z$  sind stetig.
- (iii) Die Summe  $f + g$  und das Produkt  $f \cdot g$  in  $p \in D$  stetiger Funktionen  $f, g : D \rightarrow \mathbb{C}$  sind stetig in  $p$ . Jedes Polynom in einer stetigen Funktion  $f$ , also jede Funktion der Form  $a_n f^n + a_{n-1} f^{n-1} \dots + a_0$  ist stetig.
- (iv)  $\frac{1}{f}$  ist stetig in  $p \in D$ , falls  $f : D \rightarrow \mathbb{C}^* := \mathbb{C} \setminus \{0\} \subset \mathbb{C}$  stetig in  $p$  ist.
- (v) Jede **rationale Funktion**

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

mit  $a_0, \dots, a_n, b_0, \dots, b_m \in \mathbb{C}$ , ist stetig auf dem Komplement der Polstellen:  $D := \{z \in \mathbb{C} \mid b_m z^m + \dots + b_0 \neq 0\} \subset \mathbb{C}$ .

- (vi) Die Funktionen  $\bar{f}, |f|, \operatorname{Re} f, \operatorname{Im} f$  sind stetig in  $p \in D$ , wenn  $f : D \rightarrow \mathbb{C}$  stetig in  $p$  ist.

## Satz (Verkettung stetiger Funktionen ist stetig)

Sei  $f : D \rightarrow \mathbb{C}$  stetig im Punkt  $p \in D \subset \mathbb{C}$ . Gelte  $f(D) \subset E \subset \mathbb{C}$  und sei  $g : E \rightarrow \mathbb{C}$  stetig in  $f(p)$ . Dann ist  $g \circ f : D \rightarrow \mathbb{C}$  stetig im Punkt  $p$ .

### Beweis.

- Sei  $(z_n)_{n \in \mathbb{N}} \subset D$  eine beliebige Folge mit der Eigenschaft

$$\lim_{n \rightarrow \infty} z_n = p.$$

- Wir erhalten eine Folge  $(f(z_n))_{n \in \mathbb{N}} \subset E$ . Wegen der Stetigkeit von  $f$  im Punkt  $p$  folgt

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

- Aus der Stetigkeit von  $g$  im Punkt  $f(p)$  folgt schließlich

$$\lim_{n \rightarrow \infty} g(f(z_n)) = g(f(p)) = (g \circ f)(p),$$

d.h.  $g \circ f$  ist stetig.



## Satz

Die Exponentialfunktion  $\exp : \mathbb{C} \rightarrow \mathbb{C}$  ist stetig.

## Beweis.

- Es genügt zu zeigen, dass  $\exp$  stetig in 0 ist. In der Tat:  
Sei  $(z_n)$  eine konvergente Folge und  $p = \lim_{n \rightarrow \infty} z_n$ . Dann konvergiert auch  $\exp(z_n) = \exp(z_n - p) \exp(p)$  gegen  $\exp(0) \exp(p) = \exp(p)$ , falls  $\exp$  stetig in 0 ist.

- Wir zeigen nun die Stetigkeit im Nullpunkt. Für  $|z| < 1$  gilt:

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

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



## Folgerung

Die Funktionen  $\sin, \cos : \mathbb{C} \rightarrow \mathbb{C}$  sind stetig.

### Beweis.

- Aus der Stetigkeit der Exponentialfunktion folgt die Stetigkeit der Funktionen  $\mathbb{C} \rightarrow \mathbb{C}$ ,  $z \mapsto \exp(iz)$  und  $z \mapsto \exp(-iz)$ .
- Die Stetigkeit von Sinus und Cosinus 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.



## Definition (Hyperbolische Funktionen)

Sei  $x \in \mathbb{C}$ . Die durch

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

definierten Funktionen  $\sinh : \mathbb{C} \rightarrow \mathbb{C}$  und  $\cosh : \mathbb{C} \rightarrow \mathbb{C}$  heißen **Sinus hyperbolicus** bzw. **Cosinus hyperbolicus**.

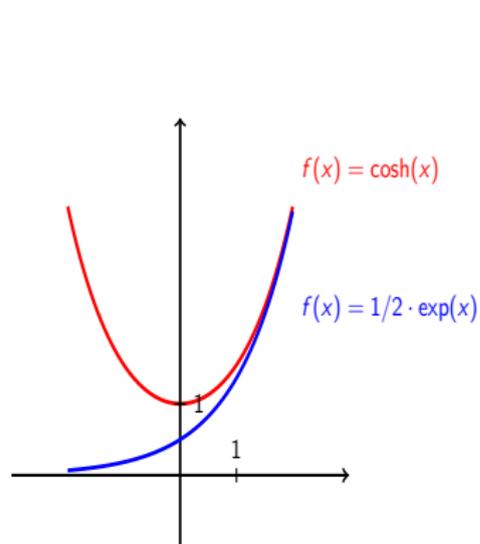
## Satz

Die Funktionen  $\sinh, \cosh : \mathbb{C} \rightarrow \mathbb{C}$  sind stetig.

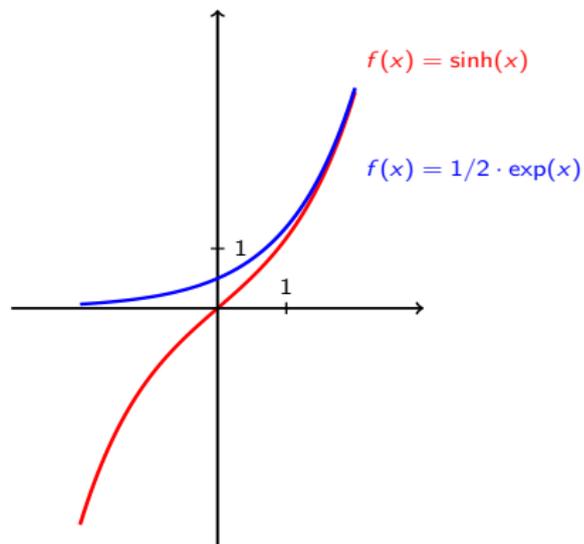
### Beweis.

Das folgt aus der Stetigkeit der Exponentialfunktion. □

## Bilder im Reellen



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



Ungerade Funktion.

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

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

$$\cosh x = \sum_{k=0}^{\infty} \frac{x^{2k}}{(2k)!} = 1 + \frac{x^2}{2} + \frac{x^4}{4!} + \dots,$$

$$\sinh x = \sum_{k=0}^{\infty} \frac{x^{2k+1}}{(2k+1)!} = x + \frac{x^3}{3!} + \frac{x^5}{5!} + \dots$$

- (Symmetrie, bzw. Antisymmetrie)

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

- (Additionstheoreme)

$$\cosh(x+y) = \cosh x \cosh y + \sinh x \sinh y,$$

$$\sinh(x+y) = \sinh x \cosh y + \cosh x \sinh y.$$

## Zwischenwerteigenschaft stetiger reeller Funktionen

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

### Satz (Nullstellensatz von Bolzano)

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

### Folgerung (Zwischenwertsatz)

Sei  $g : [a, b] \rightarrow \mathbb{R}$  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)  $[a_n, b_n] \subset [a_{n-1}, b_{n-1}]$  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  $[a_0, b_0], [a_1, b_1], \dots, [a_n, b_n]$  mit (i)–(iii) bereits konstruiert. Wir konstruieren  $[a_{n+1}, b_{n+1}]$ .
- Sei  $m = \frac{1}{2}(a_n + b_n)$ . Wir setzen:

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

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

## 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

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

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

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

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

und somit  $f(c) = 0$ .



## Beispiel

Die Vollständigkeit von  $\mathbb{R}$  ist wesentlich:

Für stetige Funktionen  $f : [a, b] \subset \mathbb{Q} \rightarrow \mathbb{R}$  gilt der Zwischensatz nicht.

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

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

ist stetig,

besitzt aber keine Fortsetzung zu einer stetigen Funktion  $\tilde{f} : \mathbb{R} \rightarrow \mathbb{R}$ .

D.h. es existiert keine stetige Funktion  $\tilde{f} : \mathbb{R} \rightarrow \mathbb{R}$  mit  $\tilde{f}(q) = f(q)$

$\forall q \in \mathbb{Q}$ .

# Infimum und Supremum einer reellwertigen Funktion

## Definition

- Sei  $f : D \rightarrow \mathbb{R}$  eine reellwertige Funktion auf einer Menge  $D$  (z.B.  $D \subset \mathbb{C}$ ).
- Die Funktion  $f$  heißt **beschränkt**, falls die Menge  $f(D) \subset \mathbb{R}$  beschränkt ist;  
 $f$  heißt **nach unten** (bzw. **nach oben**) **beschränkt**, falls  $f(D)$  nach unten (bzw. nach oben) beschränkt ist.
- Wir setzen

$$\inf f := \inf f(D) \in \mathbb{R} \cup \{-\infty\}$$

$$\sup f := \sup f(D) \in \mathbb{R} \cup \{\infty\}.$$

- 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  $f : [a, b] \rightarrow \mathbb{R}$  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 zeigen, dass  $f$  nach oben beschränkt ist und das Maximum annimmt. Die andere Aussage folgt dann, indem man  $-f$  betrachtet.

(1) Wir beweisen indirekt, dass die Menge  $f([a, b])$  nach oben beschränkt ist.

- ▶ Wäre  $f([a, b])$  nach oben unbeschränkt, so gäbe es eine Folge  $x_n \in [a, b]$ , so dass die Folge  $f(x_n)$  monoton wachsend und unbeschränkt ist.
- ▶ Nach Bolzano-Weierstraß gibt es eine konvergente Teilfolge  $(x_{n_k})_{k \in \mathbb{N}}$ ;  $\ell := \lim_{k \rightarrow \infty} x_{n_k} \in [a, b]$ .
- ▶ Stetigkeit liefert nun  $f(\ell) = \lim_{k \rightarrow \infty} f(x_{n_k})$ . Das ist unmöglich, denn jede Teilfolge einer monoton wachsenden unbeschränkten Folge ist monoton wachsend und unbeschränkt, und somit nicht konvergent.

Weiter im Beweis:

(2) Also ist das Bild  $B := f([a, b]) = \{f(x) | x \in [a, b]\}$  nach oben beschränkt.

- ▶ Dann existiert eine Folge  $x_n \in [a, b]$ , so dass  $f(x_n) \in B$  gegen  $\sup B = \sup f$  konvergiert (vgl. Existenzbeweis für das Supremum).
- ▶ Nach Bolzano-Weierstraß gibt es eine konvergente Teilfolge  $(x_{n_k})_{k \in \mathbb{N}}$ . Wir setzen  $\ell := \lim_{k \rightarrow \infty} x_{n_k} \in [a, b]$ .
- ▶ Stetigkeit liefert  $f(\ell) = \lim_{k \rightarrow \infty} f(x_{n_k}) = \lim_{n \rightarrow \infty} f(x_n) = \sup f$ , d.h.  $f$  nimmt an der Stelle  $\ell$  ihr Maximum an. □

### Folgerung

Sei  $f : [a, b] \rightarrow \mathbb{R}$  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  $f : D \rightarrow Y$  heißt **stetig in  $p$** , wenn es zu jedem  $\epsilon > 0$  ein  $\delta > 0$  gibt, derart, dass gilt

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

[Applet]

## Definition

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

$$f(x) \leq f(x') \quad (\text{bzw.} \quad f(x) < f(x'))$$

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  $\{x \in \mathbb{R} \mid x \geq 0\}$  und streng monoton fallend auf  $\{x \in \mathbb{R} \mid x \leq 0\}$ .

## Die Zahl $\pi$ und trigonometrische Funktionen

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

### Definition

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

**Beweis.** Die Cosinusfunktion 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  $|x| \leq 7$ :

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

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

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

- Wir haben zu zeigen  $0 \leq x < y \leq 2 \implies \cos x - \cos y > 0$ .
- 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  $y = \alpha + \beta$
- Dann folgt aus den Additionstheoremen:  
$$\cos x - \cos y = \cos(\alpha - \beta) - \cos(\alpha + \beta) =$$
$$\cos \alpha \cos \beta + \sin \alpha \sin \beta - (\cos \alpha \cos \beta - \sin \alpha \sin \beta) = 2 \sin \alpha \sin \beta.$$
- D.h.  $\cos x - \cos y > 0$  für alle  $0 \leq x < y \leq 2$ , falls  $\sin z > 0$  für alle  $z \in (0, 2)$ .
- Das folgt aus:

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



## 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

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

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

## Folgerung

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

**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)  $\exp(z + i2\pi) = \exp z$  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:  
 $\exp(z + in\frac{\pi}{2}) = \exp(z)i^n$ ,  $n = 4, 2, 1$ . □

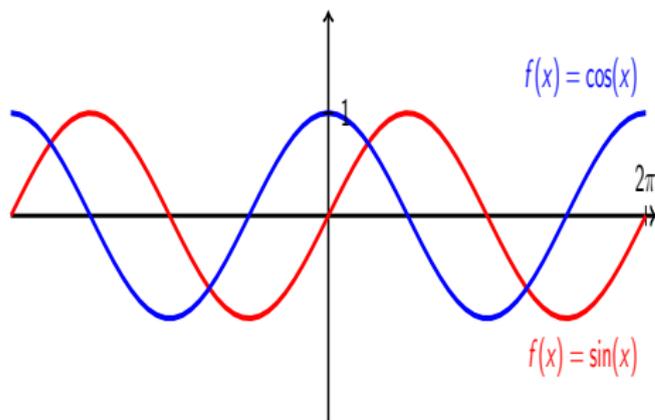
## Folgerung

Die Funktionen Sinus und Cosinus 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 Cosinusfunktion um  $\pi/2$  nach rechts.
- Die Cosinusfunktion 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  $[0, \frac{\pi}{2}]$ .





## Folgerung

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

# Umkehrfunktionen

## Definition

Sei  $D \subset \mathbb{C}$  und  $f : D \rightarrow \mathbb{C}$  eine Funktion. Eine Funktion  $g : f(D) \rightarrow D$  heißt **Umkehrfunktion** von  $f$ , falls  $g \circ f = \text{Id}_D$ , d.h.  $g(f(z)) = z \forall z \in D$ .

## Bemerkungen:

- (i) Eine Umkehrfunktion existiert genau dann, wenn  $f : D \rightarrow \mathbb{C}$  injektiv ist.
- (ii) Für  $f : D \rightarrow \mathbb{R}$  und  $g : f(D) \rightarrow \mathbb{R}$  folgt aus  $g(f(x)) = x$  für alle  $x \in D$  auch  $f(g(y)) = y$  für alle  $y \in f(D)$ .
- (iii) Wenn  $f : D \rightarrow \mathbb{R}$  eine Umkehrfunktion  $g : f(D) \rightarrow D \subset \mathbb{R}$  besitzt, dann schreiben wir diese als  $g = f^{-1}$ . **Vorsicht** bei Funktionen mit Werten in einem Körper:  $f^{-1}(x) \neq f(x)^{-1}$ .
- (iv) Hat  $f : D \subset \mathbb{R} \rightarrow \mathbb{R}$  eine Umkehrfunktion  $f^{-1}$ , so erhält man den Graphen von  $f^{-1}$  durch Spiegelung des Graphen von  $f$  an der Gerade  $\{(x, x) \mid x \in \mathbb{R}\} \subset \mathbb{R}^2$ ,  
 $\text{graph}(f) := \{(x, f(x)) \mid x \in D\}$  und  $\text{graph}(f^{-1}) = \{(f(x), x) \mid x \in D\}$

# Umkehrfunktionen streng monotoner Funktionen

*Erinnerung:* Sei  $D \subset \mathbb{R}$ . Eine Funktion  $f : D \rightarrow \mathbb{R}$  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  $D \subset \mathbb{R}$ . Jede streng monotone Funktion  $f : D \rightarrow \mathbb{R}$  besitzt eine Umkehrfunktion  $f^{-1} : f(D) \rightarrow D$ . 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  
 $x < x' \implies f(x) < f(x')$ .
- Es gilt sogar  $x < x' \iff f(x) < f(x')$ , denn  $x \geq x' \implies f(x) \geq f(x')$ .
- Die Substitution  $y = f(x)$  und  $y' = f(x')$  liefert

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

- Also ist  $g$  streng monoton wachsend.



## Satz

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

### Beweis.

- Wegen der Stetigkeit von  $f$  gilt  $f([a, b]) = [\min f, \max f]$ .
- Da die Funktion  $f$  streng monoton ist, besitzt sie eine streng monotone Umkehrfunktion  $f^{-1} : [\min f, \max f] \rightarrow [a, b]$ .
- 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,  $y = \lim_{n \rightarrow \infty} y_n$ .
- 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  $\epsilon > 0$ , 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  $x' \in [a, b] \setminus \{x\}$  konvergiert (Bolzano-Weierstraß).
- Aus der Stetigkeit von  $f$  erhalten wir nun

$$y_{n_k} = f(x_{n_k}) \xrightarrow{k \rightarrow \infty} f(x') \neq f(x) = y$$

( $f$  str. mon.)

Im Widerspruch zu der Voraussetzung  $\lim_{n \rightarrow \infty} y_n = y$ . □

## Beispiel 1: Potenz- und Wurzelfunktionen

### Satz

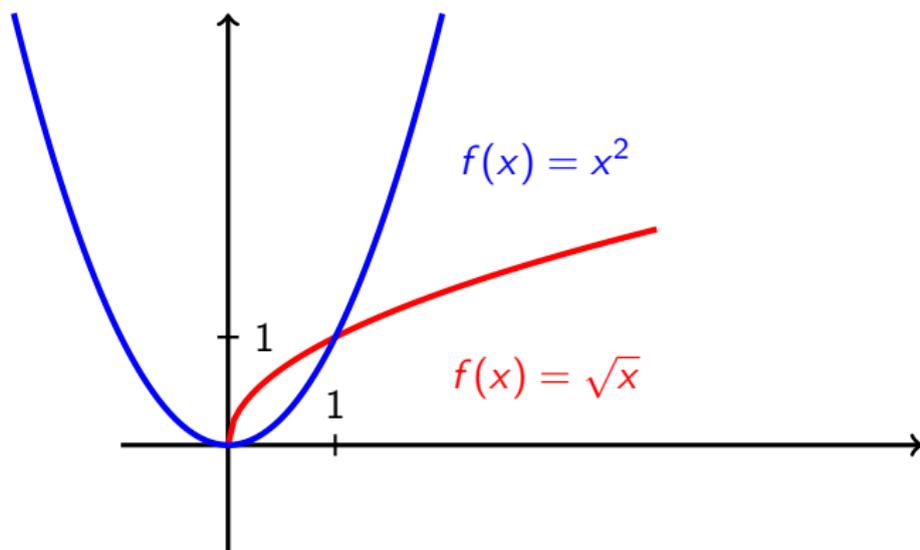
Sei  $k \in \mathbb{N}, k \geq 1$ . Die Potenzfunktion  $x \mapsto f(x) = x^k$  definiert eine stetige und streng monoton wachsende Funktion  $f : [0, \infty) \rightarrow \mathbb{R}$ .

Die Umkehrfunktion  $f^{-1} : [0, \infty) \rightarrow \mathbb{R}, f^{-1}(x) =: \sqrt[k]{x} =: x^{\frac{1}{k}}$ , ist ebenfalls stetig und streng monoton wachsend.

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

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

# Die Wurzelfunktion



## Beispiel 2: Exponential- und Logarithmusfunktion

### Satz

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

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

erfüllt die Gleichung

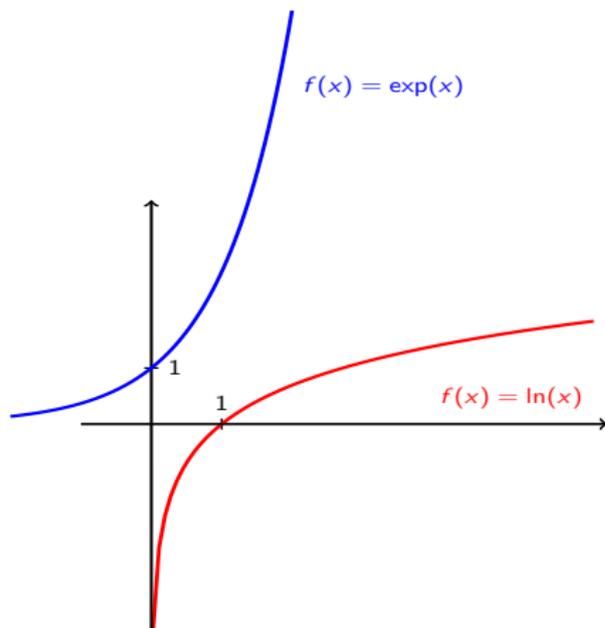
$$\ln(xy) = \ln(x) + \ln(y)$$

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  $\exp(\mathbb{R}) \subset \mathbb{R}_+$ :

- ▶ Für alle  $x \geq 0$  ist  $\exp(x) = 1 + x + \frac{x^2}{2!} + \dots \geq 1 > 0$ .
- ▶ Wegen  $\exp(x)\exp(-x) = \exp(0) = 1$ , folgt daraus  $\exp(-x) = \frac{1}{\exp(x)} > 0$ .

②  $\exp : \mathbb{R} \rightarrow \mathbb{R}$  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  $\exp(n) = 1 + n + \frac{n^2}{2!} + \dots \geq 1 + n$
- ▶ und somit  $\exp(-n) = (\exp n)^{-1} \leq \frac{1}{1+n}$ .
- ▶ Damit folgt mit dem Zwischenwertsatz:

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

## Schluss des Beweises

Wir haben gezeigt, dass  $\exp : \mathbb{R} \rightarrow \mathbb{R}_+$  stetig, streng monoton wachsend und surjektiv ist. Daher existiert eine stetige, streng monoton wachsende Umkehrfunktion  $\ln = \exp^{-1} : \mathbb{R}_+ \rightarrow \mathbb{R}$ .

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

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

heißt *Exponentialfunktion zur Basis  $a$* .

Es gilt:

- $\exp_a(x) = \exp(x) \forall x \in \mathbb{R}$ , 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

Die Funktion  $\exp_a : \mathbb{R} \rightarrow \mathbb{R}$  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)  $\exp_a\left(\frac{1}{n}\right) = a^{\frac{1}{n}} = \sqrt[n]{a}$  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  $x \in \mathbb{R}$  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)  $(\frac{1}{a})^x = (a^{-1})^x \stackrel{(ii)}{=} a^{-x}$ .



## Satz

- (i)  $\exp_1(x) \equiv 1$  für alle  $x \in \mathbb{R}$ .
- (ii) Falls  $a > 1$ , so ist die Funktion  $\exp_a : \mathbb{R} \rightarrow \mathbb{R}$  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  $\ln : \mathbb{R}_+ \rightarrow \mathbb{R}$  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  $\log_a := \exp_a^{-1} : \mathbb{R}_+ \rightarrow \mathbb{R}$  der Exponentialfunktion  $\exp_a : \mathbb{R} \rightarrow \mathbb{R}_+$  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  $\exp_a\left(\frac{\ln x}{\ln a}\right) = \exp\left(\frac{\ln x}{\ln a} \ln a\right) = \exp(\ln x) = x$ .  $\square$

## Hinweis:

Sie finden unter

<http://www.math.uni-hamburg.de/master/lehreexport/physik/visualisierung/>

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

Zur Zeit:

- Komplexe Zahlen
- Plotter für Folgen
- Plotter für Funktionen
- Das Epsilon-Delta-Kriterium für Stetigkeit
- Der Grenzwert von  $\sin(x)/x$  für  $x \rightarrow 0$ .
- Differenzenquotient

Die Seite wird weiter ergänzt.

# Funktionalgleichung für die Exponentialfunktionen

## Theorem

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

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

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

## Beweis.

- Wäre  $f(0) = 0$ , so folgte  $f(x) = f(x + 0) \stackrel{(*)}{=} f(x) \cdot f(0) = 0$  für alle  $x \in \mathbb{R}$ . Aus  $f(0) = f(0 + 0) = f(0)^2$  folgt daher  $f(0) = 1$ .
- Ebenso folgt  $a := f(1) \stackrel{(*)}{=} f\left(\frac{1}{2}\right)^2 \geq 0$  und damit  $a > 0$ .
- Wegen der Funktionalgleichung  $(*)$  ist  $f$  nicht nur in  $0$ , sondern überall stetig:

$$f(x_n) = f(x_n - x + x) \stackrel{(*)}{=} f(x_n - x)f(x) \xrightarrow[x_n \rightarrow x]{\text{stetig in } 0} f(0)f(x) \stackrel{(*)}{=} f(x).$$

## Weiter im Beweis:

Es bleibt zu zeigen, dass  $f(x) = a^x$  für alle  $x \in \mathbb{R}$ . Zwei Schritte:

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  $f(n) \stackrel{(*)}{=} f(1) \cdots f(1) = a^n$  und
- ▶ Aus  $1 = f(n - n) \stackrel{(*)}{=} a^n f(-n)$  folgt  $f(-n) = a^{-n}$ .

② Sei  $\frac{p}{q} \in \mathbb{Q}$  mit  $p \in \mathbb{Z}$  und  $q \in \mathbb{N}$ :

- ▶  $f\left(\frac{p}{q}\right)^q = \underbrace{f\left(\frac{p}{q}\right) \cdots f\left(\frac{p}{q}\right)}_{q \text{ Faktoren}} \stackrel{(*)}{=} f(p) = a^p > 0$ .
- ▶ Diese Rechnung zeigt, dass  $f\left(\frac{p}{q}\right) = \sqrt[q]{a^p} = a^{\frac{p}{q}}$ .

Sei nun  $x \in \mathbb{R}$  und  $x_n \in \mathbb{Q}$  eine Folge mit Grenzwert  $x$ .  
Aus der Stetigkeit von  $f$  und  $\exp_a$  folgt nun

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



# Uneigentliche Grenzwerte

## Definition

- Sei  $(x_n)$  eine Folge reeller Zahlen. Man schreibt  $\lim_{n \rightarrow \infty} x_n = +\infty$ , falls es für jedes  $K > 0$  ein  $N = N(K) \in \mathbb{N}$  gibt, so dass  $x_n \geq K$  für alle  $n > N = N(K)$ .
- Man schreibt  $\lim_{n \rightarrow \infty} x_n = -\infty$ , falls  $\lim_{n \rightarrow \infty} (-x_n) = +\infty$ .
- In beiden Fällen nennt man  $x_n$  **bestimmt divergent**.
- Analog definiert man uneigentliche Grenzwerte von Funktionen,  $\lim_{x \rightarrow a} f(x) = +\infty$ , wenn für **jede** Folge  $x_n \rightarrow a$  die Folge  $f(x_n)$  bestimmt gegen  $+\infty$  divergenz ist.

## Beispiele:

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

# Exponentielles Wachstum

## Satz

Für alle  $k \in \mathbb{N}$  gilt

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

**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

$$\frac{\exp x}{x^k} \geq \frac{x}{(k+1)!} \xrightarrow{x \rightarrow +\infty} +\infty.$$

□

## Bemerkung

Die Exponentialfunktion wächst also schneller als jede Potenzfunktion.

# Logarithmisches Wachstum

## Satz

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

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

## Beweis.

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

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

□

## 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 Cosinusfunktion 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

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

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

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

### Definition

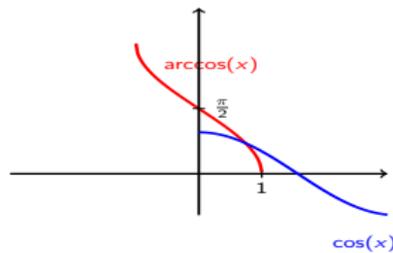
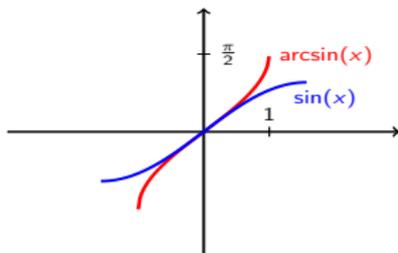
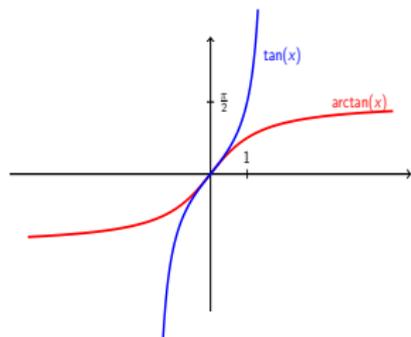
Die (streng monotonen und stetigen) Umkehrfunktionen der trigonometrischen Funktionen heißen *Arcus-Cosinus*, *Arcus-Sinus* und *Arcus-Tangens*:

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

# Arcusfunktionen

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

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



# 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  $\operatorname{Im} z \geq 0$  zu zeigen, denn falls  $\operatorname{Im} z \leq 0$  ist, so gilt  $\operatorname{Im} \bar{z} \geq 0$ , und falls  $\bar{z} = r \exp(i\varphi)$ , so muss  $z = r \exp(-i\varphi)$  gelten.

**Weiter im Beweis.**

Sei also nun  $\operatorname{Im} z \geq 0$ . Wir schreiben

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

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

Die Multiplikation von komplexen Zahlen, die in Polarkoordinaten gegeben sind, ist 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

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

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  $k = 0, 1, \dots, n - 1$ , 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  $p(z) = a_n z^n + \dots a_1 z + a_0$  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

$$\sqrt[n]{r} \exp\left(i \frac{\varphi + k2\pi}{n}\right),$$

dabei ist  $k = 0, \dots, n-1$  und  $c = r \exp(i\varphi)$ .

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

$$z^n + c_{n-1}z^{n-1} + \dots + c_1z + c_0 = 0$$

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

**Beweis.**

Wir schreiben  $P(z) = z^n + c_{n-1}z^{n-1} + \dots + c_1z + c_0$  und setzen

$$\mu = \inf_{z \in \mathbb{C}} |P(z)|.$$

Für  $|z| = R$  gilt wegen der Dreiecksungleichung

$$|P(z)| \geq R^n(1 - |c_{n-1}|R^{-1} - \dots - |c_0|R^{-n}). \quad (4)$$

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  $|P(z)| : \mathbb{C} \rightarrow \mathbb{R}_+$  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

$$Q(z) = 1 + b_k z^k + b_{k+1} z^{k+1} + \dots + b_n z^n \quad \text{mit } b_k \neq 0.$$

Wir betrachten nun die komplexe Zahl  $-\frac{|b_k|}{b_k} \in S^1$  und ihre  $k$ -te Wurzel

$$e^{i\theta} = \sqrt[k]{-\frac{|b_k|}{b_k}} \in S^1.$$

Weiter im Beweis.

Sei nun  $r > 0$  so klein gewählt, dass  $r^k |b_k| < 1$  gilt. Dann gilt

$$\begin{aligned} |Q(re^{i\theta})| &\leq |1 + b_k r^k \underbrace{e^{ik\theta}}_{=-\frac{|b_k|}{b_k}}| + |b_{k+1} r^{k+1} e^{i(k+1)\theta}| + \dots + |b_n r^n e^{in\theta}| \\ &= 1 - r^k (|b_k| - r|b_{k+1}| - \dots - r^{n-k}|b_n|) \end{aligned}$$

Für hinreichend kleines  $r_0$  ist  $(|b_k| - r_0|b_{k+1}| - \dots - r_0^{n-k}|b_n|) > 0$ , und somit

$$|Q(r_0 e^{i\theta})| < 1.$$

So erhalten wir einen Widerspruch zu  $|Q(z)| \geq 1$  und somit zur Annahme  $\mu \neq 0$ . Deshalb muss  $P(z_0) = \mu = 0$  gelten.  $\square$

# Differentialrechnung

## Definition (Differenzenquotient)

Sei  $D \subset \mathbb{R}$ ,  $f : D \rightarrow \mathbb{R}$  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.

- Jedes Element in einem (offenen oder abgeschlossenen) Intervall ist ein Häufungspunkt des Intervalls.
- Schreiben wir im Folgenden “ $\lim_{\xi \rightarrow x}$ ” meinen wir “ $\lim_{\xi \rightarrow x, \xi \neq x}$ ”.

## Definition (Ableitung)

Sei  $D \subset \mathbb{R}$ ,  $f : D \rightarrow \mathbb{R}$  eine Funktion und  $x \in D$  ein Häufungspunkt von  $D$ .

- Man sagt, dass die Funktion  $f$  im Punkt  $x$  differenzierbar ist, wenn der Grenzwert

$$f'(x) := \lim_{\xi \rightarrow x} \frac{f(\xi) - f(x)}{\xi - x}$$

existiert.

- Der Grenzwert  $f'(x)$  heißt **Ableitung** der Funktion  $f$  an der Stelle  $x$ .
- Die Funktion  $f$  heißt **differenzierbar**, wenn sie in allen Punkten  $x \in D$  differenzierbar ist.
- Ist  $f$  differenzierbar, so heißt die Funktion  $f' : x \mapsto f'(x)$  **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 in der Regel mit  $t$  bezeichnet.

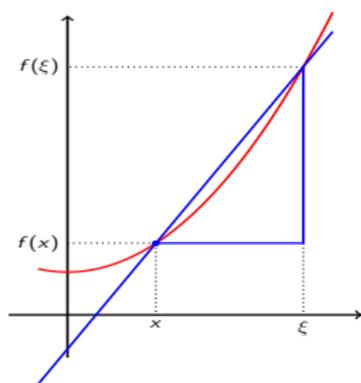
Die heute benutzte Definition der Ableitung wurde nach Vorarbeiten von Cauchy schließlich von Weierstraß Ende des 19. Jahrhunderts formuliert.

# Geometrische Interpretation: Sekante

- Der Differenzenquotient

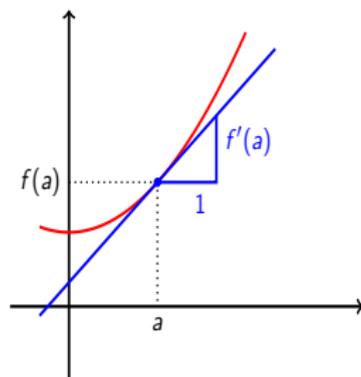
$$\frac{f(\xi) - f(x)}{\xi - x}$$

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  $\xi \rightarrow x$  gegen eine Grenzgerade, die sogenannte Tangente [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  $I \subset \mathbb{R}$  ein Intervall und  $f : I \rightarrow \mathbb{R}$  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  $x_i : \mathbb{R} \rightarrow \mathbb{R}$ ,  $i = 1, 2, 3$ , 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  $f : \mathbb{R} \rightarrow \mathbb{R}$ ,  $f(x) = c$ , gilt

$$f'(x) = \lim_{\xi \rightarrow x} \frac{c - c}{\xi - x} = \lim_{\xi \rightarrow x} 0 = 0.$$

(ii) Für jede **lineare** Funktion  $f : \mathbb{R} \rightarrow \mathbb{R}$ ,  $f(x) = ax$ , gilt

$$f'(x) = \lim_{\xi \rightarrow x} \frac{a\xi - ax}{\xi - x} = \lim_{\xi \rightarrow x} \frac{a(\xi - x)}{\xi - x} = \lim_{\xi \rightarrow x} a = a.$$

Für jede **affine** Funktion  $f(x) := ax + b$  mit  $a, b \in \mathbb{R}$  gilt

$$f'(x) = \lim_{\xi \rightarrow x} \frac{a\xi + b - (ax + b)}{\xi - x} = \lim_{\xi \rightarrow x} a = a.$$

(iii) Für  $f : \mathbb{R} \rightarrow \mathbb{R}$ ,  $f(x) = x^2$ , gilt mit der binomischen Formel:

$$f'(x) = \lim_{\xi \rightarrow x} \frac{\xi^2 - x^2}{\xi - x} = \lim_{\xi \rightarrow x} \frac{(\xi - x)(\xi + x)}{\xi - x} = \lim_{\xi \rightarrow x} (\xi + x) = 2x.$$

## Weitere Beispiele

(iv) Für  $f : \mathbb{R}^* := \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R}$ ,  $f(x) = \frac{1}{x}$ , gilt

$$f'(x) = \lim_{\xi \rightarrow x} \frac{\frac{1}{\xi} - \frac{1}{x}}{\xi - x} = \lim_{\xi \rightarrow x} \frac{\frac{x-\xi}{x\xi}}{\xi - x} = - \lim_{\xi \rightarrow x} \frac{1}{\xi x} = -\frac{1}{x^2}.$$

(v) Es gilt  $\exp' = \exp$ : Für  $h := \xi - x$  gilt

$$\frac{\exp \xi - \exp x}{\xi - x} = \frac{\exp(x+h) - \exp x}{h} = \exp x \frac{\exp h - 1}{h}$$

und  $\lim_{h \rightarrow 0} \frac{\exp(h)-1}{h} = 1$ , denn

$$\begin{aligned} \left| \frac{\exp h - 1}{h} - 1 \right| &= \left| \frac{\frac{h^2}{2!} + \frac{h^3}{3!} + \dots}{h} \right| \leq \frac{|h|}{2!} + \frac{|h|^2}{3!} + \dots \\ &\leq |h| + \frac{|h|^2}{2!} + \dots = \exp(|h|) - 1 \xrightarrow{h \rightarrow 0} 0. \end{aligned}$$

wegen der Stetigkeit der Exponentialfunktion.

## Noch mehr Beispiele

(vi) Es gilt  $\sin' = \cos$ , denn aus dem Additionstheorem für  $\sin$  folgt:

$$\begin{aligned} \frac{\sin(x+h) - \sin x}{h} &= \frac{\sin x \cos h + \cos x \sin h - \sin x}{h} \\ &= \sin x \frac{\cos h - 1}{h} + \cos x \frac{\sin h}{h}, \end{aligned}$$

und  $\lim_{h \rightarrow 0} \frac{\cos h - 1}{h} = 0$ , sowie  $\lim_{h \rightarrow 0} \frac{\sin h}{h} = 1$ .

Das folgt aus:

$$\begin{aligned} \left| \frac{\cos h - 1}{h} \right| &\leq \frac{|h|}{2!} + \frac{|h|^3}{4!} + \dots \leq |h| + \frac{|h|^2}{2!} + \dots \leq e^{|h|} - 1 \\ \left| \frac{\sin h}{h} - 1 \right| &\leq \frac{|h|^2}{3!} + \frac{|h|^4}{5!} + \dots \leq \frac{|h|^2}{2!} + \frac{|h|^4}{4!} + \dots \leq e^{|h|} - 1. \end{aligned}$$

(vii)  $\cos' = -\sin$  (ÜA).

### Beispiel (stetig aber nicht differenzierbar)

Die stetige 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 rechtsseitige Ableitung

$$\lim_{n \rightarrow \infty} \frac{f(x_n^+) - f(0)}{x_n^+ - 0} = \lim_{n \rightarrow \infty} \frac{\frac{1}{n}}{\frac{1}{n}} = \lim_{n \rightarrow \infty} 1 = 1$$

und andererseits

$$\lim_{n \rightarrow \infty} \frac{f(x_n^-) - f(0)}{x_n^- - 0} = \lim_{n \rightarrow \infty} \frac{\frac{1}{n}}{-\frac{1}{n}} = -1.$$

Also existiert der Grenzwert nicht und die Funktion  $f$  in 0 nicht differenzierbar. □

## Satz (Differenzierbare Funktionen sind stetig)

Sei  $f : D \rightarrow \mathbb{R}$  in  $a \in D \subset \mathbb{R}$  differenzierbar. Dann ist  $f$  in  $a$  stetig.

**Beweis.** Es gilt

$$|f(x) - f(a)| = \left| \frac{f(x) - f(a)}{x - a} \right| |x - a|$$

und somit

$$\lim_{x \rightarrow a} |f(x) - f(a)| = |f'(a)| \cdot \lim_{x \rightarrow a} |x - a| = 0$$

D.h.  $\lim_{x \rightarrow a} f(x) = f(a)$  und damit ist  $f$  stetig. □

## Satz (Affine Approximation)

Sei  $f : D \rightarrow \mathbb{R}$  eine in  $a \in D \subset \mathbb{R}$  differenzierbare Funktion und  $h : D \rightarrow \mathbb{R}$  die affine Funktion  $h(x) := f(a) + f'(a)(x - a)$ . Dann gilt

$$\lim_{x \rightarrow a} \frac{f(x) - h(x)}{x - a} = 0.$$

**Beweis.**  $\frac{f(x) - h(x)}{x - a} = \frac{f(x) - f(a)}{x - a} - f'(a) \xrightarrow{x \rightarrow a} 0.$  □

## Beispiel:

Die Wurzelfunktion  $f : x \mapsto \sqrt[n]{x}$  für  $n \in \mathbb{N}, n \geq 2$ , ist nicht differenzierbar in  $a = 0$ :

- 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

$$\frac{\sqrt[n]{x} - f'(0)x}{x} = \frac{1}{\sqrt[n]{x^{n-1}}} - f'(0).$$

- Das ist unbeschränkt für  $x \rightarrow 0$ , was im Widerspruch zum Satz steht.

Umgekehrt gilt:

### Satz

Sei  $f : D \rightarrow \mathbb{R}$  eine Funktion,  $a \in D$  ein Häufungspunkt von  $D$ . Wenn es eine **affine** Funktion  $h(x) = c(x - a) + f(a)$  mit  $c \in \mathbb{R}$  gibt, so dass

$$(*) \quad \lim_{x \rightarrow a} \frac{f(x) - h(x)}{x - a} = 0$$

gilt, so ist  $f$  in  $a$  differenzierbar, und die affine Funktion ist  $h(x) = f(a) + f'(a)(x - a)$ .

**Beweis.** Es gilt:

$$0 = \lim_{x \rightarrow a} \frac{f(x) - h(x)}{x - a} = \lim_{x \rightarrow a} \left( \frac{f(x) - f(a)}{x - a} - c \right) = \lim_{x \rightarrow a} \left( \frac{f(x) - f(a)}{x - a} \right) - c,$$

d.h.  $f$  differenzierbar in  $a$  mit  $f'(a) = c$ . □

## Definition (Affine Approximation/lineare Approximation)

Sei  $f : D \rightarrow \mathbb{R}$  in  $a \in D$  differenzierbar,  $D \subset \mathbb{R}$ .

- Die affine Funktion  $h(x) = f(a) + f'(a)(x - a)$  heißt **affine Approximation** von  $f$  in  $a$ .
- Manchmal spricht man auch von **linearer Approximation**.
- Die Funktion  $R(a, x) := f(x) - h(x) = f(x) - f'(a)(x - a) - f(a)$  heißt **Restglied**.

Für  $x \rightarrow a$  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 (Ableitungsregeln)

Die Funktionen  $f, g : D \rightarrow \mathbb{R}$  seien in  $a \in D \subset \mathbb{R}$  differenzierbar. Dann gilt:

- (i) Sind  $\lambda$  und  $\mu$  reelle Zahlen, dann ist die Funktion  $\lambda \cdot f + \mu \cdot g : D \rightarrow \mathbb{R}$  in  $a$  differenzierbar und

$$(\lambda \cdot f + \mu \cdot g)'(a) = \lambda \cdot f'(a) + \mu \cdot g'(a). \quad (\text{Linearitat})$$

- (ii) Die Funktion  $f \cdot g : D \rightarrow \mathbb{R}$  ist in  $a$  differenzierbar und

$$(fg)'(a) = f'(a)g(a) + f(a)g'(a). \quad (\text{Leibnizregel/Produktregel})$$

- (iii) Wenn  $g(D) \subset \mathbb{R}^*$ , dann ist die Funktion  $\frac{f}{g} : D \rightarrow \mathbb{R}$  in  $a$  differenzierbar und

$$\left(\frac{f}{g}\right)'(a) = \frac{f'(a)g(a) - f(a)g'(a)}{g(a)^2}. \quad (\text{Quotientenregel, NAZ-ZAN})$$

**Beweis.**

(i)

$$\begin{aligned} & \frac{(\lambda f + \mu g)(a+h) - (\lambda f + \mu g)(a)}{h} \\ &= \lambda \frac{f(a+h) - f(a)}{h} + \mu \frac{g(a+h) - g(a)}{h} \xrightarrow{h \rightarrow 0} \lambda \cdot f'(a) + \mu \cdot g'(a) \end{aligned}$$

(ii)

$$\begin{aligned} & \frac{(fg)(a+h) - (fg)(a)}{h} \\ &= \frac{f(a+h)g(a+h) - f(a)g(a+h)}{h} + \frac{f(a)g(a+h) - f(a)g(a)}{h} \\ &= \frac{f(a+h) - f(a)}{h} g(a+h) + f(a) \frac{g(a+h) - g(a)}{h} \\ & \xrightarrow{h \rightarrow 0} f'(a)g(a) + f(a)g'(a) \end{aligned}$$

da  $g$  als differenzierbare Funktion stetig ist.

**Weiter im Beweis.**

(iii) Wir betrachten zunächst den Spezialfall  $f = 1$ :

$$\begin{aligned} \frac{\frac{1}{g(a+h)} - \frac{1}{g(a)}}{h} &= \frac{g(a) - g(a+h)}{hg(a+h)g(a)} \\ &= - \underbrace{\frac{g(a+h) - g(a)}{h}}_{\xrightarrow{h \rightarrow 0} g'(a)} \cdot \underbrace{\frac{1}{g(a+h)g(a)}}_{\xrightarrow{h \rightarrow 0} 1/g(a)^2} \xrightarrow{h \rightarrow 0} - \frac{g'(a)}{g(a)^2} \quad (h \rightarrow 0), \end{aligned}$$

d.h.  $(\frac{1}{g})'(a) = -\frac{g'(a)}{g(a)^2}$ . Mit der Produktregel (ii) erhalten wir dann für beliebiges  $f$ :

$$\begin{aligned} \left(\frac{f}{g}\right)'(a) &= \left(f \cdot \frac{1}{g}\right)'(a) = \frac{f'(a)}{g(a)} + f(a) \left(\frac{1}{g}\right)'(a) \\ &= \frac{f'(a)g(a)}{g(a)^2} - \frac{f(a)g'(a)}{g(a)^2}. \end{aligned}$$



## Beispiele

- (i) Eine einfache Induktion mit Hilfe der Produktregel (ii) ergibt

$$(x^n)' = nx^{n-1} \quad \text{für alle } n \in \mathbb{N}.$$

Insbesondere sind polynomiale Funktionen differenzierbar.

- (ii) Die Regel  $(1/g)' = -g'/g^2$  mit  $g(x) = x^n$  liefert dann

$$(x^{-n})' = -\frac{nx^{n-1}}{x^{2n}} = -nx^{-n-1} \quad (x \neq 0) \quad \text{für alle } n \in \mathbb{N}.$$

- (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  $f : [a, b] \rightarrow \mathbb{R}$  eine stetige streng monotone Funktion und  $g = f^{-1} : [c, d] \rightarrow \mathbb{R}$  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  $y_n \in [c, d] \setminus \{y\}$  eine Folge, die gegen  $y$  konvergiert.

Da  $g$  stetig ist, konvergiert die Folge  $x_n := g(y_n) \in [a, b] \setminus \{x\}$  gegen  $x := g(y)$ .

Somit

$$\lim_{n \rightarrow \infty} \frac{g(y_n) - g(y)}{y_n - y} = \lim_{n \rightarrow \infty} \frac{x_n - x}{f(x_n) - f(x)} = \frac{1}{f'(x)},$$

d.h.  $g'(y) = 1/f'(x)$ . □

## 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  $g = f^{-1} : [c, d] \rightarrow \mathbb{R}$  die Regel

$$g' = \frac{1}{f' \circ g}.$$

## Beispiele

$$(i) \ln'(x) = \frac{1}{\exp'(\ln x)} = \frac{1}{\exp(\ln x)} = \frac{1}{x} \quad \text{für } x > 0.$$

(ii)  $\arcsin = (\sin |_{[-\frac{\pi}{2}, \frac{\pi}{2}]})^{-1} : [-1, 1] \rightarrow \mathbb{R}$  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  $\arccos = (\cos |_{[0, \pi]})^{-1} : [-1, 1] \rightarrow \mathbb{R}$  für  $x \in (-1, 1)$  die Ableitung

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

$$(iv) \arctan'(x) = \frac{1}{\tan'(\arctan x)} = \frac{1}{1 + (\tan(\arctan x))^2} = \frac{1}{1 + x^2},$$

denn  $\tan' = 1 + \tan^2$ .

## Satz (Kettenregel)

Sei  $f : D \rightarrow \mathbb{R}$  an der Stelle  $x \in D \subset \mathbb{R}$  differenzierbar,  $f(D) \subset E \subset \mathbb{R}$  und sei  $g : E \rightarrow \mathbb{R}$  an der Stelle  $y := f(x)$  differenzierbar.

Dann ist die Verkettung  $g \circ f : D \rightarrow \mathbb{R}$  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

$$h : E \rightarrow \mathbb{R}$$

$$\eta \mapsto \begin{cases} \frac{g(\eta) - g(y)}{\eta - y}, & \text{falls } \eta \in E \setminus \{y\} \\ g'(y), & \text{falls } \eta = y. \end{cases}$$

D.h.  $h$  ist die stetige Fortsetzung des Differenzenquotienten von  $g$  in  $y = f(x)$ . Für alle  $\eta \in E$  gilt dann

$$g(\eta) - g(y) = h(\eta) \cdot (\eta - y). \quad (5)$$

Sei nun  $\xi \in D \setminus \{x\}$  beliebig und  $\eta = f(\xi)$ . Aus (5) folgt dann:

$$\frac{(g \circ f)(\xi) - (g \circ f)(x)}{\xi - x} \stackrel{(1)}{=} h(f(\xi)) \cdot \frac{f(\xi) - f(x)}{\xi - x}$$

$$\xrightarrow{\xi \rightarrow x} h(f(x)) \cdot f'(x) = g'(f(x)) \cdot f'(x)$$

da  $h$  stetig. □

## Beispiele

- (i) Sei  $f : \mathbb{R} \rightarrow \mathbb{R}$  differenzierbar,  $a, b \in \mathbb{R}$  und  $g : \mathbb{R} \rightarrow \mathbb{R}$  definiert durch  $g(x) := f(ax + b)$ .

Dann ist  $g$  differenzierbar und

$$g'(x) = a \cdot f'(ax + b).$$

- (ii) Sei  $f : \mathbb{R}_+ \rightarrow \mathbb{R}$ ,  $f(x) = x^\alpha$ , 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  $f : D \rightarrow \mathbb{R}$  eine Funktion und  $z \in D \subset \mathbb{R}$ .

- 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**.
- 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  $f : (a, b) \rightarrow \mathbb{R}$  sei an der Stelle  $x \in (a, b)$  differenzierbar und nehme dort ein lokales Extremum an. Dann gilt  $f'(x) = 0$ .

### Beweis.

Wir nehmen z.B. an, 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" für " $\lim_{\xi \nearrow x}$ " und "lim" für " $\lim_{\xi \searrow x}$ ", 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  $f'(x) = 0$ .



## Beispiele

- (i) **Achtung:** Die Bedingung  $f'(x) = 0$  ist nur notwendig, aber nicht hinreichend für die Existenz eines lokalen Extremums.

Die Funktion  $f : \mathbb{R} \rightarrow \mathbb{R}$ ,  $f(x) = x^3$  erfüllt  $f'(0) = 0$ , hat aber an der Stelle 0 **kein lokales Extremum**.

- (ii) Die Funktion  $f : \mathbb{R} \rightarrow \mathbb{R}$ ,  
 $f(x) = (x+2)(x+1)x^2$ , ist nach oben unbeschränkt ( $\sup f = +\infty$ ) und hat daher **kein (globales) Maximum**.

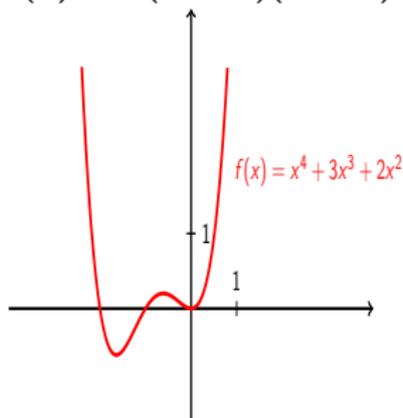
Sie nimmt ihr **(globales) Minimum**  $\min f$  an einer Stelle  $a \in (-2, -1)$  an.

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  $\min f = f(a)$ .

- (iii) Die Funktion  $f : [-1, 2] \rightarrow \mathbb{R}$ ,  $f(x) = x^2$ , 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.

$$f(x) = (x + 2)(x + 1)x^2$$



## Satz (Satz von Rolle)

Sei  $f : [a, b] \rightarrow \mathbb{R}$  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  $f : [a, b] \rightarrow \mathbb{R}$  eine stetige, auf  $(a, b)$  differenzierbare Funktion.  
Dann existiert  $\xi \in (a, b)$  mit

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

[Applet]

### **Beweis.**

Wir betrachten die Hilfsfunktion  $g : [a, b] \rightarrow \mathbb{R}$ ,

$$g(x) := f(x) - \frac{f(b) - f(a)}{b - a}(x - a).$$

$g$  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 (Schränkensatz)

Unter den Voraussetzungen des Mittelwertsatzes gelte zusätzlich für die Ableitungsfunktion

$$m \leq f'(\xi) \leq M \quad \text{für alle } \xi \in (a, b). \quad (6)$$

Dann gilt

$$m(y - x) \leq f(y) - f(x) \leq M(y - x) \quad (7)$$

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  $f : [a, b] \rightarrow \mathbb{R}$  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.  $f = \text{const.}$   $\square$

### 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  $f : \mathbb{R} \rightarrow \mathbb{R}$  eine Lösung der Differentialgleichung

$$f' = c \cdot f. \quad (8)$$

Dann gilt  $f(x) = f(0) \cdot e^{cx}$  (für alle  $x \in \mathbb{R}$ ).

**Beweis.** Wir betrachten die differenzierbare Funktion  $g : \mathbb{R} \rightarrow \mathbb{R}$ ,  
 $g(x) := f(x)e^{-cx}$ .

Ableiten liefert mit der Produktregel:

$$g'(x) = f'(x)e^{-cx} + f(x)(e^{-cx})' = f'(x)e^{-cx} - c \cdot f(x)e^{-cx} \stackrel{(8)}{=} 0.$$

Somit  $g = \text{const} = g(0) = f(0)$ , d.h.  $f(x) = g(x)e^{cx} = f(0)e^{cx}$ .  $\square$

# Grenzwertbestimmung nach L'Hospital/Bernoulli

## Satz (Verallgemeinerter Mittelwertsatz)

Seien  $f, g : [a, b] \rightarrow \mathbb{R}$  stetige Funktionen, die differenzierbar auf dem offenen Intervall  $(a, b)$  sind. Ist  $g'(x) \neq 0$  für alle  $x \in (a, b)$ , so existiert ein  $x_0 \in (a, b)$  mit

$$\frac{f(a) - f(b)}{g(a) - g(b)} = \frac{f'(x_0)}{g'(x_0)}.$$

**Beweis.** Betrachte  $h(x) = f(x) - rg(x)$  mit  $r \in \mathbb{R}$ , so dass  $h(b) = h(a)$  gilt und der Satz von Rolle anwendbar ist. □

## Folgerung (Regel von L'Hospital)

Seien  $f, g : (a, b) \rightarrow \mathbb{R}$  differenzierbar mit  $\lim_{x \searrow a} f(x) = \lim_{x \searrow a} g(x) = 0$ .

**Falls**  $\lim_{x \searrow a} \frac{f'(x)}{g'(x)}$  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 der L'Hospital'schen Regel

Da nach Voraussetzung der Grenzwert  $\lim_{x \searrow a} \frac{f'(x)}{g'(x)}$  existiert, gibt es ein  $\delta > 0$ , so dass  $g'(x) \neq 0$  für alle  $x \in (a, a + \delta)$ . Sei nun  $x \in (a, a + \delta)$ . Man definiert dann

$$\tilde{f} : [a, x] \rightarrow \mathbb{R}, \quad \tilde{f}(\xi) = \begin{cases} 0, & \text{falls } \xi = a \\ f(\xi), & \text{falls } \xi \in (a, x], \end{cases}$$

und  $g$  setzt man ebenso zu  $\tilde{g}$  fort.  $\tilde{f}$  und  $\tilde{g}$  erfüllen dann die Voraussetzungen des verallgemeinerten Mittelwertsatzes, und man findet einen Zwischenwert  $t \in (a, x)$  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

$$\lim_{n \rightarrow \infty} \frac{f(x_n)}{g(x_n)} = \lim_{n \rightarrow \infty} \frac{f'(t_n)}{g'(t_n)} = \lim_{x \searrow a} \frac{f'(x)}{g'(x)}.$$

## Beispiele

$$(i) \lim_{x \rightarrow 0} \frac{\sin x}{x} = \lim_{x \rightarrow 0} \frac{\cos x}{1} = 1. \quad (ii) \lim_{x \rightarrow 0} \frac{1 - \cos x}{x^2} = \lim_{x \rightarrow 0} \frac{\sin x}{2x} = \frac{1}{2}.$$

## Bemerkung

Analog erhält man, dass  $\lim_{x \rightarrow \infty} \frac{f(x)}{g(x)} = \lim_{x \rightarrow \infty} \frac{f'(x)}{g'(x)}$ , falls

- $f$  und  $g$  differenzierbar sind für hinreichend große  $x$ ,
- $\lim_{x \rightarrow \infty} g(x) = \lim_{x \rightarrow \infty} f(x) = \infty$  oder  $\lim_{x \rightarrow \infty} f(x) = \lim_{x \rightarrow \infty} g(x) = 0$ ,
- $\lim_{x \rightarrow \infty} \frac{f'(x)}{g'(x)}$  existiert.

- **Beispiel:**  $\lim_{x \rightarrow \infty} \frac{\ln x}{x^\alpha} = \lim_{x \rightarrow \infty} \frac{1/x}{\alpha x^{\alpha-1}} = \lim_{x \rightarrow \infty} \frac{1}{\alpha x^\alpha} = 0$  für  $\alpha > 0$ .

Der Logarithmus wächst also **langsamer** als jede Potenz.

- **Achtung:**  $\lim_{x \rightarrow \infty} \frac{f(x)}{g(x)}$  kann existieren, auch wenn  $\lim_{x \rightarrow \infty} \frac{f'(x)}{g'(x)}$  nicht existiert.

**Bsp.:** Für  $f(x) = \sin x + 2x$ ,  $g(x) = \cos x + 2x$  ist  $\lim_{x \rightarrow \infty} \frac{f'(x)}{g'(x)}$

unbestimmt divergent, aber  $\lim_{x \rightarrow \infty} \frac{f(x)}{g(x)} = \lim_{x \rightarrow \infty} \left(1 + \frac{\sin x - \cos x}{g(x)}\right) = 1$ .

## Satz (Monotonie und Ableitung)

Sei die Funktion  $f : [a, b] \rightarrow \mathbb{R}$  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 z.B. 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 nehmen wir z.B. an, dass  $f$  monoton wachsend ist.

Dann gilt für alle  $x \in (a, b)$ :

$$f'(x) = \lim_{\xi \searrow x} \frac{f(\xi) - f(x)}{\xi - x} \geq 0.$$



## Folgerung (lokale Extrema)

Sei  $f : (a, b) \rightarrow \mathbb{R}$  differenzierbar,  $x \in (a, b)$  und  $\varepsilon > 0$  derart, dass  $x \in (a + \varepsilon, b - \varepsilon)$ . Es gelte weiterhin

$$f'(\xi) \leq 0 \quad (\text{bzw.} \quad \geq 0) \quad (9)$$

für alle  $\xi \in (x - \varepsilon, x)$  und

$$f'(\xi) \geq 0 \quad (\text{bzw.} \quad \leq 0) \quad (10)$$

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).  $\square$

## Beispiele:

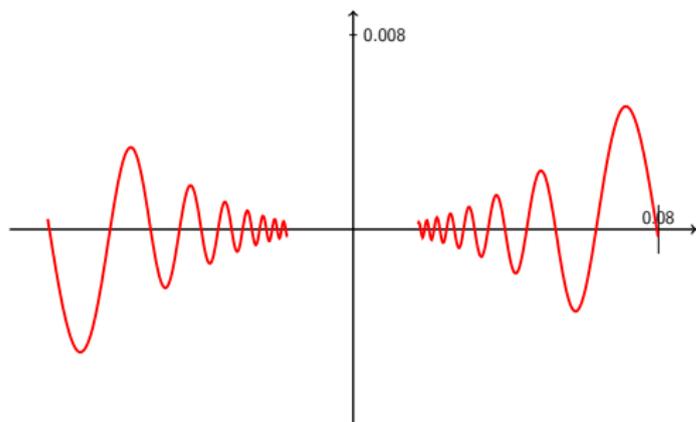
(i) Die stetige Funktion  $f(x) = \begin{cases} 0 & x = 0 \\ x^2 \sin(1/x) & x \neq 0 \end{cases}$

ist differenzierbar, da  $f'(0) = \lim_{x \rightarrow 0} x \sin(1/x) = 0$  (vgl. ÜA)

und für  $x \neq 0$  die Ableitung  $f'(x) = 2x \sin(1/x) - \cos(1/x)$  ist.

Weil der Grenzwert  $\lim_{x \rightarrow 0} f'(x)$  **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*  $f(x) = \sum_{n=1}^{\infty} \frac{2^n \sin(2^n x)}{3^n}$ .

[Graphik zeigen]

## 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  $k \in \mathbb{N}$  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  $f : (a, b) \rightarrow \mathbb{R}$  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 z.B. den Fall  $f''(x) > 0$ .

Wegen

$$0 < f''(x) = \lim_{\xi \rightarrow x} \frac{f'(\xi) - f'(x)}{\xi - x}$$

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 Minimum hat.  $\square$

## Beispiele

- (i) Für  $f : \mathbb{R} \rightarrow \mathbb{R}$ ,  $f(x) = x^2$ , 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  $f : \mathbb{R} \rightarrow \mathbb{R}$ ,  $f(x) = x^4$ , hat ebenfalls an der Stelle 0 ihr (isoliertes) globales Minimum. In diesem Fall gilt jedoch  $f''(0) = 0$ .

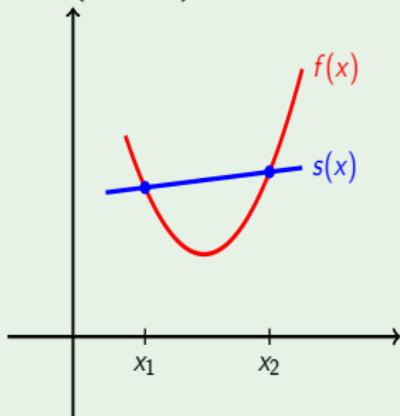
## Konvexe Funktionen

- Sei  $f : I := [a, b] \rightarrow \mathbb{R}$  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

$$f((1-t)x_1 + tx_2) \leq (1-t)f(x_1) + tf(x_2) \quad \text{für alle } t \in (0,1).$$

(Setze  $t := \frac{x-x_1}{x_2-x_1}$ , dann  $1-t = \frac{x_2-x}{x_2-x_1}$  und  $(1-t)x_1 + tx_2 = x$ .)

- Für eine zweimal differenzierbare Funktion  $f : I \rightarrow \mathbb{R}$  gilt dann:

$$f \text{ ist konvex} \iff f''(x) \geq 0 \quad \forall x \in I.$$

(Beweis: siehe z.B. Forster)

## Theorem (Taylor-Entwicklung)

Es sei  $f : [a, b] \rightarrow \mathbb{R}$  eine  $k$ -mal differenzierbare Funktion.

Zu  $x_0 \in [a, b]$  definieren wir das Polynom  $(k - 1)$ -ten Grades

$$T_{k-1}(x) := f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 \\ + \dots + \frac{f^{(k-1)}(x_0)}{(k-1)!}(x - x_0)^{k-1}.$$

Dann gibt es zu **jedem**  $x \in [a, b]$  ein  $\xi$  zwischen  $x_0$  und  $x$ , so dass

$$f(x) = T_{k-1}(x) + \frac{f^{(k)}(\xi)}{k!}(x - x_0)^k$$

*gilt*

**Bemerkung:** Für  $k = 1$  ist dies die Aussage des Mittelwertsatzes.

**Beweis.** Fixiere  $x \in [a, b]$  und bestimme  $M \in \mathbb{R}$  durch die Gleichung

$$f(x) = T_{k-1}(x) + M(x - x_0)^k.$$

Wir definieren nun eine Funktion  $g : [a, b] \rightarrow \mathbb{R}$  durch

$$g(t) := f(t) - T_{k-1}(t) - M(t - x_0)^k$$

Für  $t \in [a, b]$  ist die  $k$ -te Ableitung

$$g^{(k)}(t) = f^{(k)}(t) - k!M.$$

Wir müssen also die Existenz eines  $\xi$  zwischen  $x_0$  und  $x$  mit  $g^{(k)}(\xi) = 0$  nachweisen, denn dann gilt  $M = f^{(k)}(\xi)/k!$  für dieses  $\xi$ .

Zunächst gilt  $T_{k-1}^{(n)}(x_0) = f^{(n)}(x_0)$  für  $n = 0, \dots, k-1$ , und damit

$$0 = g(x_0) = g'(x_0) = \dots = g^{(k-1)}(x_0).$$

Per Definition von  $M$  ist aber auch  $0 = g(x)$ . Nach dem **Satz von Rolle** gibt es also ein  $x_1$  zwischen  $x_0$  und  $x$  mit  $g'(x_1) = 0$ .

Somit gibt es wiederum nach dem Satz von Rolle ein  $x_2$  zwischen  $x_0$  und  $x_1$  mit  $g''(x_2) = 0$ .

So fortfahrend, erhalten wir  $x_3, \dots, x_{k-1}$  und folgern so die Existenz des gesuchten  $\xi = x_k$ .

## Bemerkung

- Der Satz besagt, dass eine  $k$ -mal differenzierbare Funktion durch ein Polynom vom Grad  $k$  approximiert werden kann, und der dabei auftretende Fehler von der  $k$ -ten Ableitung abhängt.

$\frac{f^{(k)}(\xi)}{k!}(x - x_0)^k$  bezeichnet man als **Lagrangesches Restglied**.

- Ist  $f$  unendlich oft differenzierbar, so bezeichnet man

$$f(x)|_{x=x_0} := \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k = f(x_0) + f'(x_0)(x - x_0) + \dots$$

als **Taylor-Reihe** oder **Taylor-Entwicklung** von  $f$  an der Stelle  $x_0$ .

## Beispiele

- Die Taylor-Reihe kann divergent sein für ein  $x \neq x_0$ .

**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  $x > 1$ .

- Ein Beispiel für eine überall konvergente Taylorentwicklung:

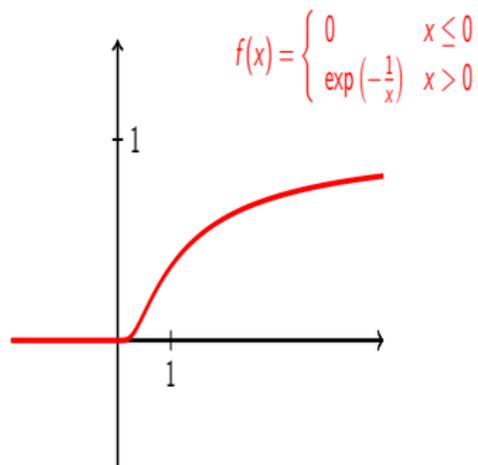
$$\sin(x)|_{x=0} = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} x^{2k+1}$$

## Beispiele

- Ist die Taylorreihe konvergent, muss sie nicht gegen  $f(x)$  konvergieren. Z.B.:

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

hat keine gegen  $f$  konvergierende Taylor-Entwicklung in  $x = 0$ , da  $f^{(n)}(0) = 0$  für alle  $n \in \mathbb{N}$ .



## Verallgemeinerung

Wir werden sehen, dass

$$d(x, y) := \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad \text{für } x, y \in \mathbb{R}^n$$

eine Abstandsfunktion definiert, die  $\mathbb{R}^n$  zum metrischen Raum macht.

Damit wissen wir, wann Funktionen  $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$  stetig sind.

Eine Funktion  $f : \mathbb{R} \rightarrow \mathbb{R}^n$  kann durch  $n$  reellwertige  
Komponentenfunktionen beschrieben werden:

$$t \mapsto f(t) = (f_1(t), f_2(t), \dots, f_n(t)) .$$

### Satz

*Eine Funktion  $f : \mathbb{R} \rightarrow \mathbb{R}^n$  ist genau dann stetig, wenn alle Funktionen  $f_i(t)$  stetig sind. Insbesondere ist eine Funktion  $f : \mathbb{R} \rightarrow \mathbb{C}$  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  $f : \mathbb{R} \rightarrow \mathbb{R}^n$  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  $f' : \mathbb{R} \rightarrow \mathbb{R}^n$ . 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*

$$h : \mathbb{R} \rightarrow \mathbb{R}^n \quad \text{mit} \quad h(t) = f'(a)(t - a) + f(a)$$

*gilt*

$$\lim_{t \rightarrow a} \frac{d(f(t), h(t))}{t - a} = 0$$

## 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

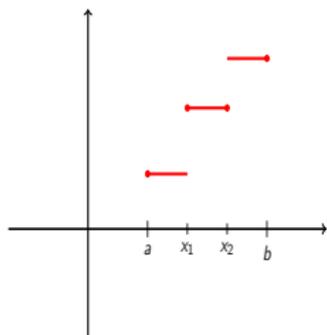
$$\frac{d(f(t), h(a))^2}{(t - a)^2} = \sum_{i=1}^n \left( \frac{f_i(t) - f_i(a)}{t - a} - f'_i(a) \right)^2 \xrightarrow{t \rightarrow a} 0$$



# Integralrechnung

## Definition (Treppenfunktion)

Eine Funktion  $\varphi : [a, b] \rightarrow \mathbb{R}$  heißt **Treppenfunktion**, wenn es eine Unterteilung  $Z : a = x_0 < x_1 < \dots < x_n = b$  des Intervalls  $[a, b]$  gibt, so dass  $\varphi$  auf jedem der offenen Teilintervalle  $(x_{i-1}, x_i)$  konstant ist,  $i = 1, 2, \dots, n$ .



## Bemerkung (ÜA)

Seien  $\varphi, \psi : [a, b] \rightarrow \mathbb{R}$  Treppenfunktionen. Dann sind  $\varphi + \psi$  und  $\varphi \cdot \psi : [a, b] \rightarrow \mathbb{R}$  Treppenfunktionen.

## Definition (Integral einer Treppenfunktion)

Sei  $\varphi : [a, b] \rightarrow \mathbb{R}$  eine Treppenfunktion und  $a = x_0 < x_1 < \dots < x_n = b$  eine Unterteilung des Intervalls  $[a, b]$ , so dass  $\varphi|_{(x_{i-1}, x_i)} = c_i = \text{const}$ ,  $i = 1, 2, \dots, n$ .

Man definiert dann 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  $c \in (a, b)$

$$\int_a^b \varphi(x) dx = \int_a^c \varphi(x) dx + \int_c^b \varphi(x) dx.$$

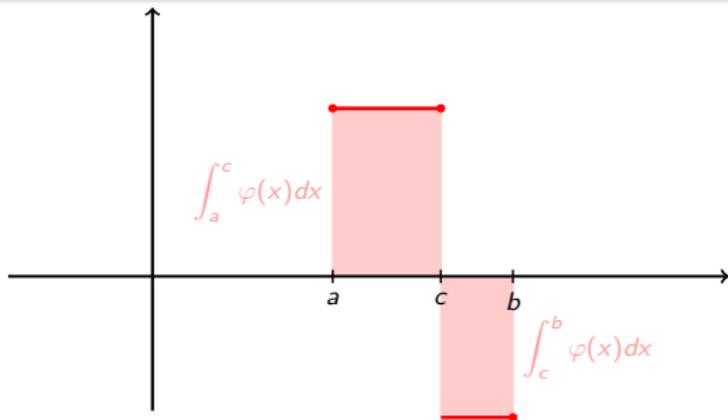
## Geometrische Interpretation

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  $\varphi, \psi : [a, b] \rightarrow \mathbb{R}$  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)  $\varphi \leq \psi \implies \int_a^b \varphi(x) dx \leq \int_a^b \psi(x) dx.$  (Monotonie)

**Beweis.** Durch Verfeinerung finden wir eine Unterteilung

$a = x_0 < x_1 < \dots < x_n = b$ , 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)

$$\begin{aligned}\int_a^b (\lambda\varphi)(x) dx &= \sum_{i=1}^n (\lambda c_i)(x_i - x_{i-1}) \\ &= \lambda \sum_{i=1}^n c_i(x_i - x_{i-1}) = \lambda \int_a^b \varphi(x) dx.\end{aligned}$$

(iii) Aus  $\varphi \leq \psi$  folgt  $c_i \leq d_i$  für alle  $i \in \{1, 2, \dots, n\}$ . Damit ist dann

$$\begin{aligned}\int_a^b \varphi(x) dx &= \sum_{i=1}^n c_i(x_i - x_{i-1}) \\ &\leq \sum_{i=1}^n d_i(x_i - x_{i-1}) = \int_a^b \psi(x) dx.\end{aligned}$$



## Definition (Ober- und Unterintegral)

Sei  $f : [a, b] \rightarrow \mathbb{R}$  eine beschränkte Funktion. Das **Oberintegral** von  $f$  ist die Zahl

$$\int_a^{*b} f(x) dx := \inf \left\{ \int_a^b \varphi(x) dx \mid \varphi \text{ Treppenfkt. mit } \varphi \geq f \right\}.$$

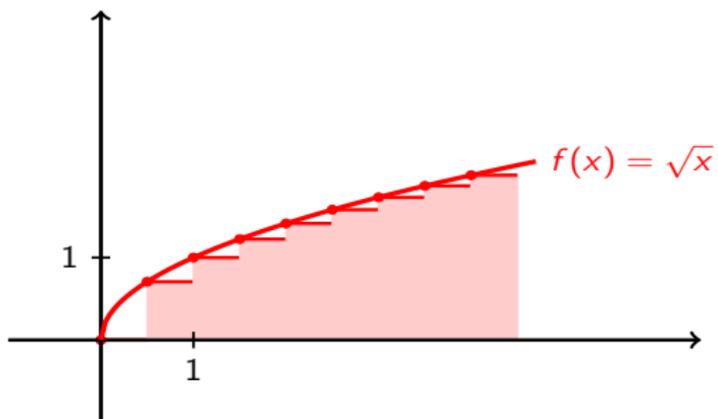
Das **Unterintegral** von  $f$  ist die Zahl

$$\int_{*a}^b f(x) dx := \sup \left\{ \int_a^b \varphi(x) dx \mid \varphi \text{ Treppenfkt. mit } \varphi \leq f \right\}.$$

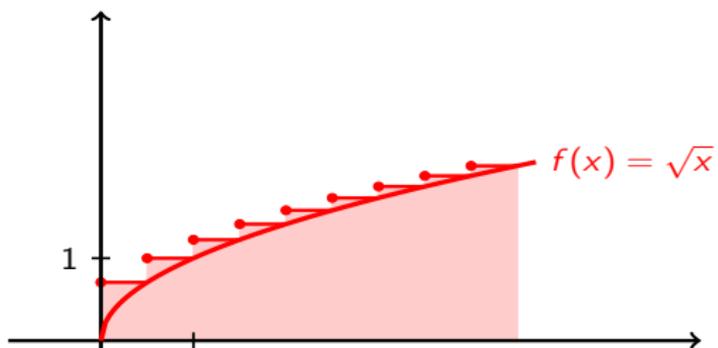
## Bemerkungen (ÜA)

- $\int_{*a}^b f(x) dx \leq \int_a^{*b} f(x) dx$ .
- $\int_{*a}^b \varphi(x) dx = \int_a^{*b} \varphi(x) dx = \int_a^b \varphi(x) dx$  für jede Treppenfunktion

## Untersumme



## Obersumme



## Beispiel

Sei  $f : [0, 1] \rightarrow \mathbb{R}$  definiert durch

$$f(x) = \begin{cases} 1, & \text{wenn } x \in \mathbb{Q} \\ 0, & \text{sonst.} \end{cases}$$

Dann gilt

$$\int_{*0}^1 f(x) dx = 0 \quad \text{und} \quad \int_0^{*1} f(x) dx = 1.$$

Ober- und Unterintegral stimmen also nicht notwendigerweise überein.

## Satz (Eigenschaften des Ober- und Unterintegrals)

Es seien  $f, g : [a, b] \rightarrow \mathbb{R}$  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  $\int^* f = \int_a^* f(x) dx,$  usw.)

**Beweis.** Für jede Teilmenge  $D \subset \mathbb{R}$  gilt  $\sup D = -\inf(-D)$ , wobei  $-D := \{-x \mid x \in D\}$ .

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  $\int^*(f+g) \leq \int^* f + \int^* g$ :

$$\begin{aligned}\int^*(f+g) &= \inf\{\int \xi \mid \xi \geq f+g \text{ und } \xi \text{ Treppenfunktion}\} \\ &\leq \inf\{\int \xi \mid \xi = \varphi + \psi, \varphi \geq f, \psi \geq g \text{ und } \varphi, \psi \text{ Treppenfunktionen}\} \\ &= \inf\{\int \varphi + \int \psi \mid \varphi \geq f, \psi \geq g \text{ Treppenfunktionen}\} \\ &= \inf\{\int \varphi \mid \varphi \geq f \text{ Treppenfkt.}\} + \inf\{\int \psi \mid \psi \geq g \text{ Treppenfunktion}\} \\ &= \int^* f + \int^* g.\end{aligned}$$

(ii) Wir zeigen  $\int^*(\lambda f) = \lambda \int^* f$  für  $\lambda > 0$ :

$$\begin{aligned}\int^*(\lambda f) &= \inf\{\int \varphi \mid \varphi \geq \lambda f \text{ Treppenfkt.}\} \\ &= \inf\{\int \varphi \mid \frac{1}{\lambda} \varphi \geq f \text{ Treppenfkt.}\} \\ &= \inf\{\lambda \int \psi \mid \psi \geq f \text{ Treppenfkt.}\} \quad [\text{mit } \psi = \frac{1}{\lambda} \varphi] \\ &= \lambda \inf\{\int \psi \mid \psi \geq f \text{ Treppenfkt.}\} = \lambda \int^* f.\end{aligned}$$



## Definition (Integrierbare Funktion)

Eine beschränkte Funktion  $f : [a, b] \rightarrow \mathbb{R}$  heißt **integrierbar** (genauer: Riemann-integrierbar), wenn

$$\int_a^{*b} f(x) dx = \int_{*a}^b f(x) dx.$$

Man definiert dann das **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.

Wir setzen  $\int_b^a = -\int_a^b$  und  $\int_a^a = 0$ .

## Satz (Linearität und Monotonie des Riemann-Integrals)

Es seien  $f, g : [a, b] \rightarrow \mathbb{R}$  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)  $f \leq g \implies \int_a^b f(x) dx \leq \int_a^b g(x) dx.$

**Beweis.**

(i) Aus

$$\begin{aligned} \int^*(f + g) &\leq \int^* f + \int^* g = \int f + \int g \\ &= \int_* f + \int_* g \leq \int_*(f + g) \leq \int^*(f + g) \end{aligned}$$

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

$$\begin{aligned} \int f &= \int_* f = \sup\left\{\int \varphi \mid \varphi \leq f \text{ Treppenfkt.}\right\} \\ &\leq \sup\left\{\int \varphi \mid \varphi \leq g \text{ Treppenfkt.}\right\} = \int_* g = \int g. \end{aligned}$$



## Bemerkungen (ÜA)

- (i) Sei  $a < b < c$ . Dann ist  $f : [a, c] \rightarrow \mathbb{R}$  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**  $f_{\pm} : [a, b] \rightarrow \mathbb{R}_+ \cup \{0\}$  einer Fkt.  $f : [a, b] \rightarrow \mathbb{R}$  sind definiert als  $f_+(x) := \max\{f(x), 0\}$  bzw.  $f_-(x) := -\min\{f(x), 0\}$ ,  $x \in [a, b]$ .

Es gilt:  $f$  integrierbar  $\implies f_+, f_-$  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 $\mathbb{R}^n$

Wir haben auch die folgende Verallgemeinerung: sei  $f : [a, b] \rightarrow \mathbb{R}^n$  eine Funktion mit Komponentenfunktionen

$$f(t) = (f_1(t), \dots, f_n(t)) ,$$

so dass jede Komponentenfunktion  $f_i : [a, b] \rightarrow \mathbb{R}$  integrierbar ist. Dann setzen wir

$$\int_a^b f(t) = \left( \int_a^b f_1, \dots, \int_a^b f_n \right) \in \mathbb{R}^n .$$

Wir nennen dann  $f$  (Riemann-)integrierbar.

Insbesondere ist nimmt das Integral einer komplexwertigen Funktion  $f : [a, b] \rightarrow \mathbb{C}$  seine Werte in den komplexen Zahlen an.

## Theorem (Integrierbarkeit stetiger Funktionen)

Jede stetige Funktion  $f : [a, b] \rightarrow \mathbb{R}$  ist integrierbar.

### Beweis.

- Es genügt zu zeigen, dass es für jedes  $\varepsilon > 0$  zwei Treppenfunktionen  $\varphi, \psi : [a, b] \rightarrow \mathbb{R}$  gibt mit  $\varphi \leq f \leq \psi$  und  $0 \leq \int \psi - \int \varphi < \varepsilon$ .
- Dazu benutzen wir folgenden Hilfssatz:

### Lemma

Jede stetige Funktion  $f : [a, b] \rightarrow \mathbb{R}$  ist auf dem **abgeschlossenen** Intervall  $[a, b]$  **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  $f : \mathbb{R} \rightarrow \mathbb{R}$ ,  $f(x) = x^2$  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  $\lim_{n \rightarrow \infty} |x_n - y_n| = 0$  gilt auch  $\lim_{k \rightarrow \infty} y_{n_k} = \ell$ .
- Die Stetigkeit von  $f$  liefert dann

$$\lim_{k \rightarrow \infty} f(x_{n_k}) = f(\ell) = \lim_{k \rightarrow \infty} f(y_{n_k}),$$

- 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  $x_i = a + i \frac{(b-a)}{n}$ ,  $i = 0, 1, \dots, n$  des Intervalls  $[a, b]$ .  
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

$$d_i - c_i < \frac{\varepsilon}{b-a} \quad \text{für alle } n \geq N \text{ und alle } i = 1, \dots, n.$$

- Wir definieren nun zwei Treppenfunktionen

$$\varphi|_{[x_{i-1}, x_i]} := c_i \quad \text{und} \quad \psi|_{[x_{i-1}, x_i]} := d_i, \quad i = 1, \dots, n$$

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  $f : [a, b] \rightarrow \mathbb{R}$  ist integrierbar.

### Definition (Riemannsche Summe)

Sei  $f : [a, b] \rightarrow \mathbb{R}$  eine Funktion,

- $a = x_0 < x_1 < \dots < x_n = b$  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

$$\sum_{i=1}^n f(\xi_i)(x_i - x_{i-1}).$$

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  $f : [a, b] \rightarrow \mathbb{R}$  approximiert die Riemannsche Summe 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  $\max_{1 \leq i \leq n} (x_i - x_{i-1}) < \delta$ .

[Animation]

### Bemerkung

Der Satz gilt sogar für beliebige (Riemann-) integrierbare Funktionen  $f : [a, b] \rightarrow \mathbb{R}$ , vgl. Forster.

**Beweis.**

- Wegen der (gleichmäßigen) Stetigkeit von  $f : [a, b] \rightarrow \mathbb{R}$  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  $\max_{1 \leq i \leq n} (x_i - x_{i-1}) < \delta$  gilt dann:

$$|f(x) - f(\xi_i)| < \frac{\varepsilon}{b-a} \text{ für alle } x \in [x_{i-1}, x_i].$$

- Damit erhalten wir

$$\begin{aligned} \left| \int_a^b f(x) dx - \sum_{i=1}^n f(\xi_i)(x_i - x_{i-1}) \right| &= \left| \sum_{i=1}^n \int_{x_{i-1}}^{x_i} (f(x) - f(\xi_i)) dx \right| \\ &\leq \sum_{i=1}^n \int_{x_{i-1}}^{x_i} \underbrace{|f(x) - f(\xi_i)|}_{< \frac{\varepsilon}{b-a}} dx < \sum_{i=1}^n (x_i - x_{i-1}) \frac{\varepsilon}{b-a} = \varepsilon. \end{aligned}$$



## Beispiel: Berechnung des Integrals $\int_0^b x dx$

- Wähle eine äquidistante Unterteilung  $x_i := \frac{ib}{n}$ ,  $i = 1, \dots, n$ .
- Wähle als Stützstellen  $\xi_i := x_i$ .
- Die zugehörige Riemannsche Summe ist dann

$$\begin{aligned}\sum_{i=1}^n \frac{ib}{n} \frac{b}{n} &= \frac{b^2}{n^2} \sum_{i=1}^n i \\ &= \frac{b^2}{n^2} \frac{n(n+1)}{2} = \frac{b^2}{2} \left(1 + \frac{1}{n}\right).\end{aligned}$$

- Also erhält man  $\int_0^b x dx$  als Grenzwert der Riemannschen Summe für  $n \rightarrow \infty$ :

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

Dies ist in der Tat die Fläche des gleichschenkligen rechtwinkligen Dreiecks unter dem Graphen von  $f(x) = x$ .

**Verabredung:** Im Folgenden sei  $I \subset \mathbb{R}$  immer ein Intervall.

### Definition (Stammfunktion)

Sei  $f : I \rightarrow \mathbb{R}$  eine Funktion. Eine **Stammfunktion** von  $f$  ist eine differenzierbare Funktion  $F : I \rightarrow \mathbb{R}$ , so dass

$$F' = f.$$

Stammfunktionen  $F_1$  und  $F_2$  zu gegebenem  $f$  unterscheiden sich auf einem Intervall durch eine reelle Konstante:  $(F_1 - F_2)' = f - f = 0$ , d.h.

$$F_1 - F_2 \equiv c \in \mathbb{R}.$$

### Theorem (Fundamentalsatz der Differential- und Integralrechnung)

Sei  $f : I \rightarrow \mathbb{R}$  eine stetige Funktion und  $x_0 \in I$ .

Die durch

$$F(x) := \int_{x_0}^x f(t) dt$$

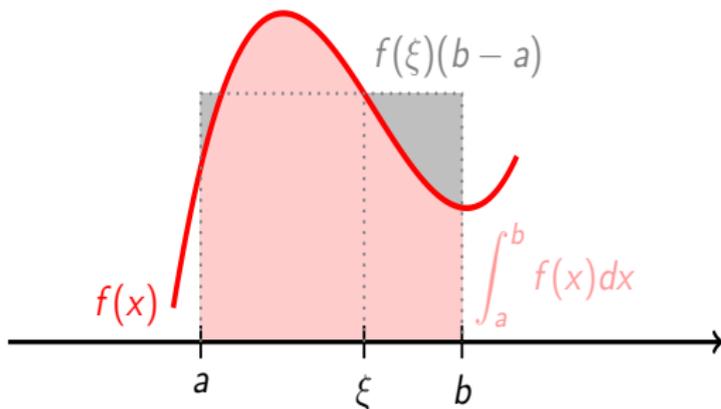
definierte Funktion  $F : I \rightarrow \mathbb{R}$  ist eine Stammfunktion von  $f$ .

**Beweis des Fundamentalsatzes.** Wir benutzen folgendes Lemma:

Lemma (Mittelwertsatz der Integralrechnung)

Sei  $f : [a, b] \rightarrow \mathbb{R}$  eine stetige Funktion. Dann existiert  $\xi \in [a, b]$ , so dass

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



### Lemma (Mittelwertsatz der Integralrechnung)

Sei  $f : [a, b] \rightarrow \mathbb{R}$  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  $f(\xi) = \frac{\int_a^b f(x) dx}{b - a}$ .



## Weiter im Beweis des Fundamentalsatzes

- Sei  $0 \neq h \in \mathbb{R}$  derart, dass  $x, x + h \in I$ . Dank des Lemmas 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  $F$ :

$$\begin{aligned} \frac{F(x+h) - F(x)}{h} &= \frac{\int_{x_0}^{x+h} f(t) dt - \int_{x_0}^x f(t) dt}{h} \\ &= \frac{\int_x^{x+h} f(t) dt}{h} \\ &= f(\xi) \xrightarrow{h \rightarrow 0} f(x), \end{aligned}$$

d.h.  $F'(x) = f(x)$ .



## Folgerung

Sei  $f : I \rightarrow \mathbb{R}$  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.** 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

$$F \Big|_a^b := F(b) - F(a).$$

## Beispiele

- $\int_0^{\pi} \sin(x) dx = -\cos(x)|_0^{\pi} = -\cos(\pi) + \cos(0) = -(-1) + 1 = 2.$

- Ist  $\alpha \neq -1$ , so gilt  $\int_a^b x^{\alpha} dx = \frac{1}{\alpha+1} x^{\alpha+1}|_a^b$ , 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  $\int_a^b \frac{1}{x} dx = \ln(x)|_a^b$

und für  $a < b < 0$  gilt  $\int_a^b \frac{1}{x} dx = \ln(-x)|_a^b.$

D.h. für  $0 \notin [a, b]$  ergibt sich also  $\int_a^b \frac{1}{x} dx = \ln(|x|)|_a^b$

## Theorem (Substitutionsregel)

Sei  $f : I \rightarrow \mathbb{R}$  eine stetige Funktion und  $\varphi : [a, b] \rightarrow \mathbb{R}$  eine stetig differenzierbare Funktion mit  $\varphi([a, b]) \subset I$ . Dann gilt

$$\int_a^b f(\varphi(t))\varphi'(t)dt = \int_{\varphi(a)}^{\varphi(b)} f(x)dx.$$

### Beweis.

- Sei  $F : I \rightarrow \mathbb{R}$  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

$$\begin{aligned} \int_a^b f(\varphi(t))\varphi'(t)dt &= \int_a^b (F \circ \varphi)'(t)dt \\ &= F \circ \varphi \Big|_a^b = F \Big|_{\varphi(a)}^{\varphi(b)} = \int_{\varphi(a)}^{\varphi(b)} f(x)dx. \end{aligned}$$

**Schreibweise:** Definiert man

$$d\varphi(t) := \varphi'(t)dt = \frac{d\varphi}{dt} dt ,$$

so kann man die Substitutionsregel schreiben als

$$\int_a^b f(\varphi(t))d\varphi(t) = \int_{\varphi(a)}^{\varphi(b)} f(x)dx,$$

d.h.  $x$  wird durch  $\varphi(t)$  ersetzt.

## Beispiele

- Sei  $c \neq 0$  und  $d \in \mathbb{R}$ ,  $f$  stetig. Dann ist

$$\int_a^b f(ct + d)dt = \frac{1}{c} \int_{ca+d}^{cb+d} f(x)dx.$$

- 

$$\begin{aligned} \int_0^2 \frac{2t}{t^2+1} dt & \stackrel{(x=t^2)}{=} \int_0^4 \frac{dx}{x+1} = \ln(x+1) \Big|_0^4 \\ & = \ln 5 - \ln 1 = \ln 5 \end{aligned}$$

## Beispiel: Fläche des Kreises vom Radius 1

Wir stellen den Halbkreis als Graphen der stetigen Funktion  $f : [-1, 1] \rightarrow \mathbb{R}$ ,  $f(x) = \sqrt{1 - x^2}$  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

$$\begin{aligned} 2 \int_{-1}^1 \sqrt{1 - x^2} dx &= \int_{-\pi/2}^{\pi/2} \cos^2(t) dt = \int_{-\pi/2}^{\pi/2} \cos 2t dt + \int_{-\pi/2}^{\pi/2} dt \\ &= \frac{1}{2} \int_{-\pi}^{\pi} \cos s ds + \int_{-\pi/2}^{\pi/2} dt = \frac{1}{2} \sin s \Big|_{-\pi}^{\pi} + t \Big|_{-\pi/2}^{\pi/2} = 0 + \pi = \pi. \end{aligned}$$

## Theorem (Partielle Integration)

Es seien  $f, g : [a, b] \rightarrow \mathbb{R}$  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:

$$(fg)|_a^b \stackrel{\text{Fundamentalsatz}}{=} \int_a^b (fg)'(x)dx = \int_a^b f'(x)g(x)dx + \int_a^b f(x)g'(x)dx.$$

□

**Kurzschreibweise für die partielle Integration:**

$$\int fdg = fg - \int gdf.$$

## Beispiele

- Sei  $a, b > 0$ . Dann ist

$$\int_a^b \ln x \, dx = \int_a^b 1 \cdot \ln x \, dx = x \ln x \Big|_a^b - \int_a^b \underbrace{x \cdot \frac{1}{x}}_{=1} dx = x(\ln x - 1) \Big|_a^b$$

Hierbei ist

$$\begin{array}{ll} f(x) = x & f'(x) = 1 \\ g(x) = \ln x & g'(x) = 1/x \end{array}$$

- Also ist  $x(\ln x - 1)$  eine Stammfunktion des Logarithmus.

## Beispiele

- $x \arctan x - \frac{1}{2} \ln(x^2 + 1)$  ist eine Stammfunktion des  $\arctan$ , denn:

$$\begin{aligned} \int_a^b \arctan x \, dx &= x \arctan x \Big|_a^b - \int_a^b x \arctan' x \, dx \\ &= x \arctan x \Big|_a^b - \int_a^b \frac{x}{1+x^2} dx \end{aligned}$$

Mittels der Substitution  $t = x^2$  hatten wir gesehen, dass

$$\int_a^b \frac{x}{x^2+1} dx = \frac{1}{2} \int_{a^2}^{b^2} \frac{1}{t+1} dt = \frac{1}{2} \ln(t+1) \Big|_{a^2}^{b^2} = \frac{1}{2} \ln(x^2+1) \Big|_a^b.$$

$$\text{Damit ist } \int_a^b \arctan x \, dx = x \arctan x - \frac{1}{2} \ln(x^2+1) \Big|_a^b.$$

# Mathematik I für Studierende der Geophysik/Ozeanographie, Meteorologie und Physik Vorlesungsskript<sup>1</sup>

Christoph Schweigert  
Fachbereich Mathematik  
Universität Hamburg  
[www.math.uni-hamburg.de/home/schweigert](http://www.math.uni-hamburg.de/home/schweigert)

Hamburg, Wintersemester 2012/13

---

<sup>1</sup>Version vom 10. April 2013, basierend auf Versionen der Vorjahre. Ursprüngliche Version: Vicente Cortés, 2006/07

# Inhaltsverzeichnis I

- 1 Vektorräume
  - Definition und Beispiele
  - Unterräume
  - Lineare Unabhängigkeit
  - Erzeugendensysteme, Basen
  - Austauschsätze von Steinitz
  - Dimension
  - Gaußscher Algorithmus
- 2 Lineare Abbildungen
  - Definition und Eigenschaften
  - Lineare Abbildungen und Matrizen
  - Rang einer linearen Abbildung
  - Der Lösungsraum eines linearen Gleichungssystems
  - Direkte Summe von Unterräumen
- 3 Gruppen
  - Gruppen und Gruppenhomomorphismen

## Inhaltsverzeichnis II

- Die symmetrische Gruppe

4

### Determinante

- Determinante einer  $2 \times 2$ -Matrix
- Charakterisierung der Determinante
- Explizite Formel für die Determinante
- Determinantenmultiplikationssatz
- Orientierung
- Transponierte Matrix
- Matrixinversion und Determinanten

# 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

$$+ : V \times V \rightarrow V, \quad (v, w) \mapsto v + w,$$

genannt **Addition**, und mit einer Abbildung

$$\mathbb{K} \times V \rightarrow V, \quad (\lambda, v) \mapsto \lambda \cdot v (=:\lambda v),$$

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**

$$\mathbb{K}^n = \{(x_1, \dots, x_n) \mid x_1, \dots, x_n \in \mathbb{K}\}$$

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  $\text{Abb}(X, \mathbb{K})$  die Menge der Funktionen  $f : X \rightarrow \mathbb{K}$ .

$\text{Abb}(X, \mathbb{K})$  ist mit der punktweise definierten Addition und skalaren Multiplikation

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

$$(\lambda f)(x) = \lambda f(x) \quad (f, g \in \text{Abb}(X, \mathbb{K}), \quad \lambda \in \mathbb{K}, \quad x \in X)$$

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  $U \subset V$ , 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

$$+ : U \times U \rightarrow U$$

und eine skalare Multiplikation

$$\cdot : \mathbb{K} \times U \rightarrow U,$$

die  $U$  zu einem Vektorraum machen, denn wegen  $-u = (-1)u$  liegt auch das additive Inverse von  $u \in U$  in  $U$ .

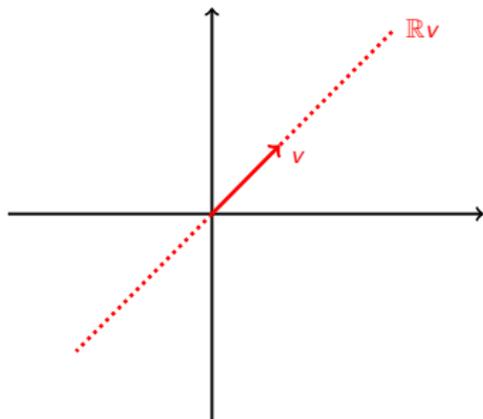
## Beispiele von Unterräumen

Sei  $V$  ein Vektorraum.

- 1)  $\{0\}$  und  $V$  selbst sind Untervektorräume von  $V$ .
- 2) Für jeden Vektor  $v \in V$  ist  $\mathbb{K}v = \{\lambda v \mid \lambda \in \mathbb{K}\} \subset V$  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  $V$ .

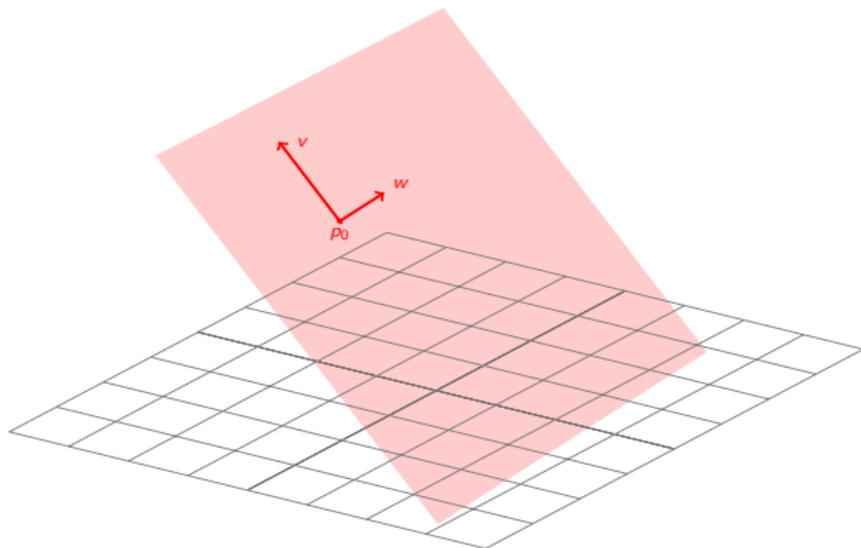
## Beispiele von Unterräumen

3) Für jedes Paar von Vektoren  $v, w$  ist

$$\mathbb{K}v + \mathbb{K}w = \{\lambda v + \mu w \mid \lambda, \mu \in \mathbb{K}\} \subset V$$

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  $I \subset \mathbb{R}$  ein Intervall und  $k \in \mathbb{N}$ ,  $k \geq 1$ . Die folgenden Mengen sind jeweils Unterräume von  $\text{Abb}(I, \mathbb{R})$ :

$$\begin{aligned} C^k(I, \mathbb{R}) &:= \{f : I \rightarrow \mathbb{R} \mid f \text{ } k\text{-mal stetig differenzierbar}\} \\ &\subset \text{Diff}(I, \mathbb{R}) := \{f : I \rightarrow \mathbb{R} \mid f \text{ differenzierbar}\} \\ &\subset C(I, \mathbb{R}) := C^0(I, \mathbb{R}) := \{f : I \rightarrow \mathbb{R} \mid f \text{ stetig}\} \end{aligned}$$

- 5) Sei  $V$  ein Vektorraum und  $U_j \subset V$  durch  $j \in J$  indizierte Unterräume, wobei  $J$  eine beliebige Menge ist.

Der Durchschnitt  $U = \bigcap_{j \in J} U_j \subset V$  ist ein Unterraum (ÜA).

Beispielsweise ist

$$C^\infty(I, \mathbb{R}) := \bigcap_{k \in \mathbb{N}} C^k(I, \mathbb{R})$$

ein Unterraum von  $\text{Abb}(I, \mathbb{R})$ . 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

$$a_0 + a_1x + a_2x^2 + \cdots + a_nx^n,$$

wobei  $a_0, \dots, a_n \in \mathbb{K}$  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  $a_0 + a_1x + \cdots + a_nx^n \in \mathbb{K}[x]$  definiert durch Einsetzen eine Funktion

$$P : \mathbb{K} \rightarrow \mathbb{K}, \quad \lambda \mapsto P(\lambda) = a_0 + a_1\lambda + \cdots + a_n\lambda^n.$$

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  $a_0 + \cdots + a_nx^n \mapsto P$  injektiv (ÜA) und  $\mathbb{K}[x]$  wird auf diese Weise zu einem Untervektorraum  $\mathbb{K}[x] \subset \text{Abb}(\mathbb{K}, \mathbb{K})$ . Wir haben die Inklusion  $\mathbb{R}[x] \subset C^\infty(\mathbb{R}, \mathbb{R}) \subset \text{Abb}(\mathbb{R}, \mathbb{R})$ .

Wir kommen zu einem weiteren zentralen Begriff dieser Vorlesung.

### Definition (Lineare Unabhängigkeit)

- (i) Sei  $V$  ein Vektorraum und  $v_1, \dots, v_r \in V$  endlich viele Vektoren. Man sagt, dass  $v \in V$  eine **Linearkombination** der Vektoren  $v_1, \dots, v_r$  ist, wenn es  $\lambda_j \in \mathbb{K}$  gibt, so dass

$$v = \sum_{i=1}^r \lambda_i v_i.$$

Die Familie von Vektoren  $v_1, \dots, v_r$  heißt **linear unabhängig**, wenn

$$\sum_{i=1}^r \lambda_i v_i = 0 \implies \lambda_1 = \dots = \lambda_r = 0,$$

d.h. wenn der Nullvektor  $0 \in V$  sich nur in trivialer Weise als Linearkombination der Vektoren der Familie darstellen lässt.

(ii) Sei nun  $(v_j)_{j \in J}$  eine durch eine unendliche Menge  $J$  indizierte Familie von Vektoren  $v_j \in V$ .

Man sagt, dass  $v \in V$  eine **Linearkombination** der Vektoren  $(v_j)_{j \in J}$  ist, wenn es **eine endliche** Teilmenge  $J_0 \subset J$  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  $J_0 \subset J$  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

$$e_i := (0, \dots, 0, 1, 0, \dots, 0) \in \mathbb{K}^n.$$

(i-te Stelle)

Die Vektoren  $(e_1, \dots, e_n)$  sind linear unabhängig. In der Tat:

$$0 = \sum_{i=1}^n \lambda_i e_i = (\lambda_1, \dots, \lambda_n) \implies \lambda_1 = \dots = \lambda_n = 0.$$

- 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.
- 4) Für jeden Vektorraum  $V$  ist die zu Grunde liegende Menge  $V$  ein (sehr großes) Erzeugendensystem.

## 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

$$\text{span}\{v_j | j \in J\} := \{v | v \text{ ist eine Linearkombination der } (v_j)_{j \in J}\}.$$

**Beispiel:** Sei  $v_1 = (1, 1, 0)$  und  $v_2 = (1, -1, 0) \in \mathbb{R}^3$ . Dann ist

$$\begin{aligned} \text{span}\{v_1, v_2\} &= \{\lambda(1, 1, 0) + \mu(1, -1, 0) \mid \lambda, \mu \in \mathbb{R}\} \\ &= \{(\lambda, \lambda, 0) + (\mu, -\mu, 0) \mid \lambda, \mu \in \mathbb{R}\} \\ &= \{(\lambda + \mu, \lambda - \mu, 0) \mid \lambda, \mu \in \mathbb{R}\} \\ &= \{(x, y, 0) \mid x, y \in \mathbb{R}\} \\ &= \text{span}\{e_1, e_2\} \end{aligned}$$

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  $v \in \text{span}\{v_j | j \in J\}$  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  $J_0 \subset J$  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  $J'_0 \subset J$ .

**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

$$\begin{aligned} \lambda_j - \lambda'_j &= 0, & \text{wenn } j \in J_0 \cap J'_0 \\ \lambda_j &= 0, & \text{wenn } j \in J_0 \setminus J_0 \cap J'_0 \\ \lambda'_j &= 0, & \text{wenn } j \in J'_0 \setminus J_0 \cap J'_0. \end{aligned}$$

Also stimmen die beiden Darstellungen

$$v = \sum_{j \in J_0} \lambda_j v_j = \sum_{j \in J_0 \cap J'_0} \lambda_j v_j \quad \text{und} \quad v = \sum_{j \in J'_0} \lambda'_j v_j = \sum_{j \in J_0 \cap J'_0} \lambda'_j v_j$$

überein. □

## Definition (Erzeugendensysteme, Basen)

Sei  $V$  ein Vektorraum. Eine Familie von Vektoren  $(v_j)_{j \in J}$  heißt **Erzeugendensystem** von  $V$ , falls  $V = \text{span}\{v_j \mid j \in J\}$ .

Eine **Basis** von  $V$  ist ein linear unabhängiges Erzeugendensystem.

## Beispiele

- 1)  $(e_1, \dots, e_n)$  ist eine Basis von  $\mathbb{K}^n$ , die sogenannte **kanonische Basis**.

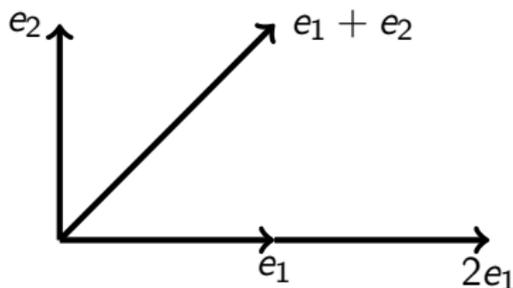
**Warnung:** Im allgemeinen hat ein Vektorraum keine *kanonische* Basis!

- 2)  $(1, x, x^2, \dots)$  ist eine Basis von  $\mathbb{K}[x]$ .

## Ein Beispiel

- 3)  $(e_1, e_2, e_1 + e_2, 2e_1)$  ist ein Erzeugendensystem 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  $\mathbb{K} = \mathbb{R}$ :



## Ansage:

- Bitte melden Sie sich für die Klausur über STINE oder über Ihr Studienbüro an!
- Bitte melden Sie sich rechtzeitig beim Büro für die Belange von Studierenden mit Behinderung oder chronischer Erkrankung, <http://www.verwaltung.uni-hamburg.de/behinderung/> wenn Sie glauben, dass sie nicht am normalen Klausurablauf teilnehmen können.  
Uns muss über Ihr Studienbüro eine offizielle Anerkennung durch den Prüfungsausschuss vorliegen!

## 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  $J_0 \subset J$  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  $J_0 \supset J$  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  $\text{span}\{v_j | j \in J\}$  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  $V = \text{span}\{v_j | j \in J\}$ .

$(iii) \Rightarrow (i)$  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.

Dann hat jeder der Vektoren  $v_i$ ,  $i \in J \setminus J_0$ , eine Darstellung  $(*) v_i = \sum_{j \in J_0} \lambda_j v_j$ . Nach (iii) ist das die eindeutige Darstellung als Linearkombination der  $(v_j)_{j \in J}$ .

Wäre nun  $i \notin J_0$ , so wären  $v_i = v_i$  und  $(*)$  zwei verschiedene Darstellungen. Widerspruch zur Eindeutigkeit, also  $J = J_0$ .

**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  $J_0 \subset J$  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  $(v_j)_{j \in J_0}$ ,  $J_0 \supsetneq J$ , eine Familie, die die  $(v_j)_{j \in J}$  enthält und  $j_0 \in J_0 \setminus J$ .

Da  $(v_j)_{j \in J}$  ein Erzeugendensystem ist, gibt es eine Darstellung

$$v_{j_0} = \sum_{j \in J} \lambda_j v_j.$$

Das zeigt, dass  $(v_j)_{j \in J_0}$  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  $v \in V \setminus \text{span}\{v_j | j \in J\}$ . 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 finde  $v = -\sum (\lambda_j / \mu) v_j$ , 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  $(v_1, \dots, v_n)$  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  $(v_1, \dots, v_{k-1}, w, v_{k+1}, \dots, v_n)$  eine Basis von  $V$ .

**Beweis.**

Übungsaufgabe. □

## Satz (Steinitz'scher 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  $\varphi : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$  (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  $r \rightarrow r + 1$  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  $\varphi : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ , 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  $(w_1, \dots, w_{r+1}, v_{\varphi(r+2)}, \dots, v_{\varphi(n)})$ . □

## 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  $(v_1, \dots, v_n)$  eine Basis. Aus dem Austauschsatz folgt, dass jede linear unabhängige Familie höchstens  $n$  Elemente hat.
- (ii) Sei  $(w_1, \dots, w_r)$  eine zweite Basis. Aus dem Austauschsatz folgt, wie gesagt,  $r \leq n$ . Nach Vertauschen der Rollen der beiden Basen folgt ebenso  $n \leq r$ .



## Definition (Dimension)

Sei  $V$  ein Vektorraum über  $\mathbb{K}$ .

Die *Dimension* von  $V$  ist die Zahl

$$\dim V = \dim_{\mathbb{K}} V := \begin{cases} 0, & \text{falls } V = \{0\} \\ n, & \text{falls } V \text{ eine endliche Basis} \\ & (v_1, \dots, v_n) \text{ hat} \\ \infty & \text{sonst} \end{cases}$$

## 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)  $\dim_{\mathbb{K}} \text{Abb}(X, \mathbb{K}) = \text{card}(X)$ , wobei

$$\text{card}(X) := \begin{cases} n, & \text{falls } X \text{ aus } n \text{ Elementen besteht} \\ \infty, & \text{falls } X \text{ unendlich ist.} \end{cases}$$

- 5) Für den Vektorraum der Polynome gilt  $\dim_{\mathbb{K}} \mathbb{K}[X] = \infty$ .
- 6) Sei  $U \subset V$  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. □

## Basis von Unterräumen aus Erzeugendensystemen

Sei  $V = \mathbb{K}^n$  und  $v_1, \dots, v_m \in V$  und

$$U = \text{span}\{v_1, \dots, v_m\} \subset V$$

der von  $(v_1, \dots, v_m)$  aufgespannte Unterraum.

Wir wollen nun eine aus dem Erzeugendensystem  $\{v_1, \dots, v_m\}$  eine Basis des Unterraumes  $U$  bestimmen. Dazu benutzt man den **Gaußschen Algorithmus** (Gaußsches Eliminationsverfahren). Dieser beruht auf den folgenden Fakten, von denen sich (ii) und (iii) aus dem Austauschatz ergeben:

- (i)  $\text{span}\{v_1, \dots, v_m\} = \text{span}\{v_{\varphi(1)}, \dots, v_{\varphi(m)}\}$   
für jede Permutation  $\varphi : \{1, \dots, m\} \rightarrow \{1, \dots, m\}$ ,
- (ii)  $\text{span}\{\lambda v_1, v_2, \dots, v_m\} = \text{span}\{v_1, \dots, v_m\}$   
für alle  $\lambda \in \mathbb{K}^* := \mathbb{K} \setminus \{0\}$
- (iii)  $\text{span}\{v_1, \dots, v_{i-1}, v_i + \lambda v_1, v_{i+1}, \dots, v_m\} = \text{span}\{v_1, \dots, v_m\}$  für  $i \neq 1$ .

Dazu schreibt man die Komponenten der Vektoren

$$v_i = (a_{i1}, a_{i2}, \dots, a_{in}) \in \mathbb{K}^n$$

als Zeilen einer **Matrix**

$$A := \begin{pmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} =: (a_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}}$$

Mit Hilfe des Gaußschen Algorithmus, d.h. durch Anwendung **elementarer Zeilenoperationen**, die auf **(i)**, **(ii)** und **(iii)** beruhen, überführen wir nun  $A$  in eine Matrix

$$B = \begin{pmatrix} b_{11} & \dots & b_{1n} \\ \vdots & & \vdots \\ b_{m1} & \dots & b_{mn} \end{pmatrix} =: (b_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}},$$

deren erste  $k \leq m$  Zeilen  $w_i = (b_{i1}, \dots, b_{in})$ ,  $i = 1, \dots, k$ , **eine Basis von  $U$  bilden** und  $b_{ij} = 0$  für alle weiteren Zeilen,  $k < i \leq m$ ,  $j = 1, \dots, n$ .

# 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  $1 \leq j_1 < j_2 < \dots < j_k \leq n$  so daß die Matrixeinträge  $b_{ij}$  die folgenden Eigenschaften haben:

- (1)  $b_{1j_1} = b_{2j_2} = \dots = b_{kj_k} = 1$ ,
- (2)  $b_{ij} = 0$  für alle  $i = 1, \dots, k$  und  $1 \leq j < j_i$ ,
- (3)  $b_{ij} = 0$  für alle  $i = k + 1, \dots, m$  und  $j = 1, \dots, n$ .

Eine Matrix  $B = (b_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}}$  in **Zeilenstufenform** sieht so aus:

$$\begin{pmatrix} 0 & \begin{matrix} j_1-1 \\ \dots \end{matrix} \text{ viele} & 0 & \mathbf{1} & b_{1j_1+1} & \dots & b_{1n} \\ 0 & \begin{matrix} j_2-1 \\ \dots \end{matrix} \text{ viele} & & 0 & \mathbf{1} & b_{2j_2+1} & \dots & b_{2n} \\ \vdots & & & & & & & \vdots \\ 0 & \begin{matrix} j_i-1 \\ \dots \end{matrix} \text{ viele} & & 0 & \mathbf{1} & b_{ij_i+1} & \dots & b_{in} \\ \vdots & & & & & & & \vdots \\ 0 & \begin{matrix} j_k-1 \\ \dots \end{matrix} \text{ viele} & & & 0 & \mathbf{1} & b_{kj_k+1} \dots & b_{kn} \\ 0 & & & \dots & & & & 0 \\ \vdots & & & & & & & \vdots \\ 0 & & & \dots & & & & 0 \end{pmatrix}$$

## Konklusion

Die Vektoren, die aus den ersten  $k$  Zeilen von  $B$  bestehen,

$$\begin{aligned} w_1 &= (0 \quad \overset{j_1-1 \text{ viele}}{\dots} \quad 0, 1, b_{1j_1+1}, \dots, b_{1n}) \\ &\vdots \\ w_k &= (0 \quad \overset{j_k-1 \text{ viele}}{\dots} \quad 0, 1, b_{kj_k+1}, \dots, b_{kn}) \end{aligned}$$

bilden eine Basis von  $U = \text{span}\{v_1, \dots, v_m\}$ , 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.  $(w_1, \dots, w_k)$  sind ein Erzeugendensystem von  $U$ .
- Die Vektoren  $(w_1, \dots, w_k)$  sind linear unabhängig, da

$$0 = \sum_{i=1}^k \lambda_i w_i = (\dots, \lambda_1, \dots, \lambda_2 + \lambda_1 \cdot (\dots), \dots, \text{etc.})$$

impliziert  $\lambda_i = 0$  für  $i = 1, \dots, k$ .

**Zahlenbeispiel.**

Sei z.B.  $\mathbb{K} = \mathbb{Q}$ ,  $= \mathbb{R}$  oder  $= \mathbb{C}$  und  $V = \mathbb{K}^5$  und ein Erzeugendensystem des Unterraums gegeben durch

$$v_1 = (0, 0, 2, 1, 0)$$

$$v_2 = (0, 1, 0, 2, 1)$$

$$v_3 = (0, 2, 1, 1, 1)$$

$$v_4 = (0, 4, 4, 3, 2)$$

Vertauschen der ersten beiden Zeilen ändert nach (i) nicht die lineare Hülle der Zeilenvektoren:

$$A = \begin{pmatrix} 0 & 0 & 2 & 1 & 0 \\ 0 & 1 & 0 & 2 & 1 \\ 0 & 2 & 1 & 1 & 1 \\ 0 & 4 & 4 & 3 & 2 \end{pmatrix} \xrightarrow{(i)} \begin{pmatrix} 0 & 1 & 0 & 2 & 1 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 2 & 1 & 1 & 1 \\ 0 & 4 & 4 & 3 & 2 \end{pmatrix}$$

**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:

Weiter im Zahlenbeispiel:

$$\begin{pmatrix} 0 & 1 & 0 & 2 & 1 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 2 & 1 & 1 & 1 \\ 0 & 4 & 4 & 3 & 2 \end{pmatrix} \xrightarrow{-2 \times I + III} \begin{pmatrix} 0 & 1 & 0 & 2 & 1 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 1 & -3 & -1 \\ 0 & 4 & 4 & 3 & 2 \end{pmatrix}$$

$$\xrightarrow{-4 \times I + IV} \begin{pmatrix} 0 & 1 & 0 & 2 & 1 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 1 & -3 & -1 \\ 0 & 0 & 4 & -5 & -2 \end{pmatrix}$$

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:

**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

$$w_1 = (0, 1, 0, 2, 1)$$

$$w_2 = (0, 0, 1, \frac{1}{2}, 0)$$

$$w_3 = (0, 0, 0, 1, \frac{2}{7})$$

Somit ist  $(w_1, w_2, w_3)$  eine Basis von  $U = \text{span}\{v_1, v_2, v_3, v_4\} \subset \mathbb{K}^5$ .

**Bemerkung:**

Auf den letzten Schritt könnte man auch verzichten, um eine Basis zu erhalten.

**Bemerkung:**

Der Gaußsche Algorithmus kann auch zum Lösen **inhomogener linearer Gleichungssysteme**

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

verwendet werden. Hierbei sind die  $a_{ij}, b_j \in \mathbb{K}$  gegeben. Gesucht sind die **Unbekannten**  $x_1, x_2, \dots, x_n \in \mathbb{K}$ .

Dazu wendet man den Algorithmus an auf die  $m \times (n+1)$ -Matrix

$$(A|b) = \begin{pmatrix} a_{11} & \cdots & a_{1n} & b_1 \\ & \vdots & \vdots & \\ a_{m1} & \cdots & a_{mn} & b_m \end{pmatrix}$$

(mehr dazu später).

# Kapitel 11

## Lineare Abbildungen

# Lineare Abbildungen

## Definition

$V, W$  seien Vektorräume über einem Körper  $\mathbb{K}$ .

Eine Abbildung  $F : V \rightarrow W$  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  $F : V \rightarrow W$  heißt auch **Isomorphismus** von Vektorräumen.

Die Vektorräume  $V$  und  $W$  heißen **isomorph**, wenn es einen Isomorphismus  $F : V \rightarrow W$  von Vektorräumen gibt.

## Grundlegende Eigenschaften:

- 1 Für jede Menge  $X$  und jeden Vektorraum  $V$  bilden die Abbildungen  $X \rightarrow V$  einen Vektorraum, den wir mit  $\text{Abb}(X, V)$  bezeichnen. Die Operationen sind  $(\lambda \cdot f)(x) := \lambda \cdot f(x)$  und  $(f + g)(x) := f(x) + g(x)$ .
- 2 Seien  $V, W$  Vektorräume. Die Menge aller linearen Abbildungen  $V \rightarrow W$  ist ein Unterraum von  $\text{Abb}(V, W)$ , den wir mit  $L(V, W)$  bezeichnen. Dass dies ein Unterraum ist, heißt, dass  $\lambda f$  und  $f + g$  linear sind, falls  $f$  und  $g$  linear sind.
- 3 Eine lineare Abbildung ist bestimmt durch ihre Werte auf Basisvektoren: sind  $F, G \in L(V, W)$  und  $(v_1, \dots, v_n)$  ist eine Basis von  $V$  und gilt  $F(v_i) = G(v_i)$  für  $i = 1, \dots, n$ , 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

$$F(v) = F\left(\sum_{i=1}^n \lambda_i v_i\right) = \sum_{i=1}^n \lambda_i F(v_i) = \sum_{i=1}^n \lambda_i G(v_i) = G(v).$$

# 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  $F : V \rightarrow W$  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

$$F(v) = \sum_{i=1}^n \lambda_i F(v_i) = \sum_{i=1}^n \lambda_i w_i$$

annehmen. Man rechnet leicht nach, dass dies eine lineare Abbildung definiert. □

## Weitere wichtige Eigenschaften linearer Abbildungen

- 1 Sei  $F : U \rightarrow V$  linear und  $G : V \rightarrow W$  linear. Dann ist auch die Komposition  $G \circ F : U \rightarrow W$  linear.
- 2 Sei  $F : V \rightarrow W$  linear. Für jeden Unterraum  $V' \subset V$  ist das Bild  $F(V') \subset W$  ein Unterraum. Ebenso ist das Urbild  $F^{-1}(W') \subset V$  ein Unterraum für jeden Unterraum  $W' \subset W$ .
- 3 Insbesondere sind für jede lineare Abbildung  $F$  ihr Bild  $\text{im } F := F(V) \subset W$  und ihr Kern  $\text{ker } F := F^{-1}(0) = \{v \in V \mid F(v) = 0\} \subset V$  jeweils Unterräume.
- 4  $F$  ist genau dann injektiv, wenn  $\text{ker } F = \{0\}$ .
- 5 Eine lineare Abbildung  $F : V \rightarrow W$  ist ein Isomorphismus genau dann, wenn  $\text{im } F = W$  und  $\text{ker } F = \{0\}$ .
- 6 Sei  $F : V \rightarrow W$  ein Isomorphismus. Dann ist  $F^{-1} : W \rightarrow V$  linear.

## Satz

Es seien  $V, W$  Vektorräume,  $F : V \rightarrow W$  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  $F : V \rightarrow W$  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

$$\begin{aligned}\varphi : L(V, \mathbb{K}) &\rightarrow \mathbb{K}^n \\ F &\mapsto (F(v_1), F(v_2), \dots, F(v_n))\end{aligned}$$

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  $(a_1, \dots, a_n) \in \mathbb{K}^n$  beliebig. Dann definieren wir die Abbildung  $F$  durch die Werte  $F(v_i) := a_i$  auf den Basisvektoren  $(v_i)$ .



# Der Dualraum

## Definition

Sei  $V$  ein Vektorraum der Dimension  $n < \infty$ . Der  $n$  dimensionale 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  $(e_1, \dots, e_n)$  von  $\mathbb{K}^n$  duale Basis die kanonische Basis von  $(\mathbb{K}^n)^*$ . I.a. 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

$$\text{Mat}(m, n, \mathbb{K}) := \left\{ A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \mid a_{ij} \in \mathbb{K} \right\}$$

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!

## 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  $\text{Mat}(m, n, \mathbb{K})$ . Wir ordnen dabei einer linearen Abbildung  $F \in L(\mathbb{K}^n, \mathbb{K}^m)$  eine  $m \times n$ -Matrix zu durch

$$\varphi : F \mapsto (F(e_1) \ \dots \ F(e_n)),$$

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 von  $\mathbb{K}^n$ .

Somit ist  $\varphi(F) = A$ . □

# 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 von  $\mathbb{K}^n$ .

Daher ist für  $x = \sum_{j=1}^n x_j e_j$

$$\begin{aligned} F(x) &= F\left(\sum_{j=1}^n x_j e_j\right) = \sum_{j=1}^n x_j F(e_j) \\ &= \sum_{j=1}^n x_j \sum_{i=1}^m a_{ij} e_i = \sum_{i=1}^m \left(\sum_{j=1}^n a_{ij} x_j\right) e_i = Ax \end{aligned}$$

D.h. das Bild von  $x$  unter  $F$  ist gleich dem Produkt  $Ax$ .

# Produkt zweier Matrizen

## Definition

Seien nun  $A = (a_{ij})_{i,j} \in \text{Mat}(m, n, \mathbb{K})$  und  $B = (b_{jk})_{j,k} \in \text{Mat}(n, p, \mathbb{K})$ .

Das **Produkt**  $C = AB$  von  $A$  und  $B$  ist die Matrix

$C = (c_{ik})_{\substack{i=1 \dots m \\ k=1 \dots p}} \in \text{Mat}(m, p, \mathbb{K})$  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

### Definition

Seien nun  $A = (a_{ij})_{i,j} \in \text{Mat}(m, n, \mathbb{K})$  und  $B = (b_{jk})_{j,k} \in \text{Mat}(n, p, \mathbb{K})$ .

Das **Produkt**  $C = AB$  von  $A$  und  $B$  ist die Matrix

$$C = (c_{ik})_{\substack{i=1 \dots m \\ k=1 \dots p}} \in \text{Mat}(m, p, \mathbb{K}) \text{ mit } c_{ik} = \sum_{j=1}^n a_{ij} b_{jk}.$$

Diese Matrix entspricht der Matrix der Verkettung der zu  $B$  und  $A$  gehörenden linearen Abbildungen  $\mathbb{K}^p \rightarrow \mathbb{K}^n$  und  $\mathbb{K}^n \rightarrow \mathbb{K}^m$ , denn für

$x = \sum_{k=1}^p x_k e_k \in \mathbb{K}^p$  ist

$$\begin{aligned} A(Bx) &= A\left(\sum_{j=1}^n \sum_{k=1}^p b_{jk} x_k e_j\right) = \sum_{k=1}^p \sum_{j=1}^n b_{jk} x_k A(e_j) \\ &= \sum_{k=1}^p \sum_{j=1}^n b_{jk} x_k \sum_{i=1}^m a_{ij} e_i = \sum_{i=1}^m \sum_{k=1}^p \left(\sum_{j=1}^n a_{ij} b_{jk}\right) x_k e_i \end{aligned}$$

Also  $A(Bx) = Cx$  für  $C = (c_{ik})_{i,k}$  mit  $c_{ik} = \sum_{j=1}^n a_{ij} b_{jk}$ .

## Darstellende Matrix einer linearen Abbildung

$B = (v_1, \dots, v_n)$  bzw.  $B' = (w_1, \dots, w_m)$  seien geordnete Basen der endlich-dimensionalen  $\mathbb{K}$ -Vektorräume  $V$  bzw.  $W$ ;  $\phi_B : \mathbb{K}^n \rightarrow V$  und  $\phi_{B'} : \mathbb{K}^m \rightarrow W$  seien die zugehörigen Isomorphismen definiert durch  $\phi_B(e_i) = v_i$ ,  $\phi_{B'}(e_j) = w_j$ , wobei  $i = 1, \dots, n$ ,  $j = 1, \dots, m$  und  $(e_i)_i$  die entsprechende kanonische Basis ist.

### Definition

Sei  $F \in L(V, W)$ . Die Matrix  $A \in \text{Mat}(m, n, \mathbb{K})$ , 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  $\phi_B = \text{Id}_V$ ,  $\phi_{B'} = \text{Id}_W$  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:

$$F(v_j) = \sum_{i=1}^m a_{ij} w_i, \quad j = 1, \dots, n. \quad (1)$$

**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 Anwenden von  $\phi_{B'}$  in die äquivalente Gleichung (1) über, da  $\phi_B(e_j) = v_j$  und  $\phi_{B'}(e_i) = w_i$ . □

## 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  $V$ ,  $\text{End}(V) := L(V, V)$ .

Den Vektorraum der quadratischen  $(n \times n)$ -Matrizen bezeichnet man mit  $\text{Mat}(n, \mathbb{K}) := \text{Mat}(n, n, \mathbb{K})$ .

Für die darstellende Matrix eines Endomorphismus  $F \in \text{End}(V)$  bzgl. **einer** Basis  $B$  von  $V$  schreibt man  $M_B(F) := M_B^B(F)$ .

## Beispiel

Sei  $F \in \text{End}(\mathbb{R}^2)$  gegeben durch  $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \in \text{Mat}(2, \mathbb{R})$ . 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 von linearen Abbildungen und Matrizen

## Definition (Rang einer linearen Abbildung)

Der **Rang** einer linearen Abbildung  $F : V \rightarrow W$  ist die nicht-negative ganze Zahl

$$\operatorname{rg}(F) := \dim \operatorname{im} F.$$

## Bemerkung

Sei  $A = (a_{ij})_{i,j} \in \operatorname{Mat}(m, n, \mathbb{K})$  eine Matrix.

- Die  $n$  Spalten von  $A$  sind gegeben durch  $A(e_1), \dots, A(e_n) \in \mathbb{K}^m$  und die Dimension  $\operatorname{rg}(A)$  des von den Spalten von  $A$  erzeugten Unterraumes von  $\mathbb{K}^m$  nennt man den **Spaltenrang** von  $A$ .
- Die Dimension des von den  $m$  Zeilen  $(a_{i1}, \dots, a_{in}) \in \mathbb{K}^n$  erzeugten Unterraumes von  $\mathbb{K}^n$  heißt **Zeilenrang** von  $A$ .
- Wir werden später sehen, dass Zeilenrang=Spaltenrang. Damit kann man  $\operatorname{rg}(A)$  mit dem **Gaußschen Algorithmus** bestimmen.

## Satz

Seien  $V, W$  endlich-dimensionale Vektorräume und  $B, B'$  Basen von  $V$  bzw.  $W$ . Dann gilt für jede lineare Abbildung  $F : V \rightarrow W$

$$\operatorname{rg}(F) = \operatorname{rg}(M_B^{B'}(F)).$$

**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.

Also  $\operatorname{rg}(F) = \operatorname{rg}(F \circ \phi_B) = \operatorname{rg}(M_B^{B'}(F))$ . □

## Bemerkung

Man kann also den Rang einer jeden linearen Abbildung zwischen endlich-dimensionalen Vektorräumen mit Hilfe des Gaußschen Algorithmus bestimmen, indem man den Rang der darstellenden Matrix bezüglich irgend zweier Basen bestimmt.

## Rang und Dimension

### Satz (Dimensionsformel)

Seien  $V, W$  Vektorräume,  $\dim V$  endlich und  $F \in L(V, W)$ . Dann gilt

$$\operatorname{rg}(F) + \dim \ker(F) = \dim V.$$

### 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  $(u_1, \dots, u_k)$  eine Basis von  $\ker(F)$ ,  $(w_1, \dots, w_r)$  eine Basis des Bildes  $F(V)$ . Wähle Urbilder  $v_1, \dots, v_r \in V$ , so dass  $F(v_i) = w_i$ .

### Behauptung:

$(u_1, \dots, u_k, v_1, \dots, v_r)$  ist eine Basis von  $V$ .

Aus der Behauptung folgt die Aussage des Satzes:  $\dim V = k + r$ .  $\square$

## Beweis der Behauptung.

- 1) Wir zeigen zuerst, dass  $(u_1, \dots, u_k, v_1, \dots, v_r)$  linear unabhängig ist.

Aus  $0 = \sum_{i=1}^k \lambda_i u_i + \sum_{j=1}^r \mu_j v_j$  folgt

$$0 = F\left(\sum_{i=1}^k \lambda_i u_i + \sum_{j=1}^r \mu_j v_j\right) \stackrel{u_i \in \text{Ker}(F)}{=} \sum_{j=1}^r \mu_j w_j$$

und somit  $\mu_1 = \dots = \mu_r = 0$ , da die Familie  $(w_j)$  linear unabhängig ist.

Also  $0 = \sum_{i=1}^k \lambda_i u_i$ , woraus  $\lambda_1 = \dots = \lambda_k = 0$  folgt, wegen der linearen Unabhängigkeit von  $(u_i)$  als Basis von  $\ker F$ .

- 2) Nun zeigen wir, dass  $\text{span}\{u_1, \dots, u_k, v_1, \dots, v_r\} = V$ .

Sei  $v \in V$ . Wir schreiben  $F(v) = \sum_{j=1}^r \mu_j w_j$ .

Dann gilt  $v - \sum_{j=1}^r \mu_j v_j \in \ker(F) = \text{span}\{u_1, \dots, u_k\}$

und somit  $v \in \text{span}\{u_1, \dots, u_k, v_1, \dots, v_r\}$ . □

## Folgerung

Seien  $V, W$  Vektorräume,  $\dim V < \infty$  und  $F \in L(V, W)$ . Dann gilt:

- (i)  $F$  ist genau dann surjektiv, wenn  $\operatorname{rg}(F) = \dim W$ ,
- (ii)  $F$  ist genau dann injektiv, wenn  $\operatorname{rg}(F) = \dim V$ ,
- (iii)  $F$  ist genau dann bijektiv, wenn  $\operatorname{rg}(F) = \dim V = \dim W$ .

## Beweis.

- (i)  $F$  surjektiv  $\iff F(V) = W \iff \dim F(V) = \dim W$ .
- (ii)  $F$  injektiv  $\iff \ker(F) = \{0\} \stackrel{(\text{Dim.formel})}{\iff} \operatorname{rg}(F) = \dim V$ .
- (iii) folgt aus (i-ii). □

## Folgerung

Seien  $V, W$  Vektorräume mit  $\dim V = \dim W < \infty$ .

Dann gilt: Eine lineare Abbildung  $F : V \rightarrow W$  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

$$\dim(V) = \operatorname{rg}(F) + \dim(\ker(F)) = \dim(W).$$

Daraus folgen die Äquivalenzen. □

# 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  $x_1, x_2, \dots, x_n \in \mathbb{K}$  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

$A = (a_{ij}) \in \text{Mat}(m, n, \mathbb{K})$  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 Gleichungssystems  $Ax = 0$ ,  $A \in \text{Mat}(m, n, \mathbb{K})$ , ist genau der Kern der zu  $A$  gehörenden linearen Abbildung, also insbesondere ein Unterraum  $U \subset \mathbb{K}^n$ .
- (ii) Die Dimension von  $U$  beträgt  $n - r$ , wobei  $r = \text{rg}(A)$ .

**Beweis.** (i) folgt aus der Definition und (ii) aus der Dimensionsformel.  $\square$

## Satz

Sei  $A \in \text{Mat}(m, n, \mathbb{K})$  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  $Ae_j$  der kanonischen Basis. Der Beweis beruht nun darauf, dass wir jede **Zeilen** 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

$$\alpha_i(e_j) := a_{ij}, \quad \text{d.h.} \quad Ax = \begin{pmatrix} \alpha_1(x) \\ \vdots \\ \alpha_m(x) \end{pmatrix}.$$

Die  $a_{ij}$  sind also die Komponenten der Linearform  $\alpha_i$  in der kanonischen dualen Basis von  $(\mathbb{K}^n)^*$ . Sei  $V^* := \text{span} \{ \alpha_1, \dots, \alpha_m \} \subset (\mathbb{K}^n)^*$  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

$$\{x \in \mathbb{K}^n \mid L(x) = 0 \forall L \in V^*\}$$



## 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

$$\begin{array}{rcl}
 x_{j_1} & + & \sum_{l=j_1+1}^n b_{1l}x_l = 0 \\
 & & \vdots \\
 x_{j_i} & + & \sum_{l=j_i+1}^n b_{il}x_l = 0 \\
 & & \vdots \\
 x_{j_k} & + & \sum_{l=j_k+1}^n b_{kl}x_l = 0
 \end{array}$$

- 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  $k = \text{Zeilenrang}(A) = \text{Spaltenrang}(B) = \text{Spaltenrang}(A)$ .

## Satz

Wir betrachten ein *inhomogenes* lineares Gleichungssystem

$$Ax = b, \quad \text{mit } A \in \text{Mat}(m, n, \mathbb{K}). \quad (2)$$

Es bezeichne  $U'$  die Lösungsmenge von (2).

Falls  $U' \neq \emptyset$ , so ist  $U'$  von der Form

$$U' = x_0 + U = \{x_0 + u \mid u \in U\},$$

wobei  $x_0 \in U'$  eine (spezielle) Lösung von (2) ist und  $U \subset \mathbb{K}^n$  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  $u \in U$  finde  $A(x_0 + u) = Ax_0 + Au = b + 0 = b$ . Also  $x_0 + U \subset U'$ .

Sei  $v \in U'$  eine Lösung von (2). Dann gilt

$A(v - x_0) = Av - Ax_0 = b - b = 0$ , also  $v \in x_0 + U$  und somit

$U' \subset x_0 + U$ .



## Definition

Ein **affiner Unterraum** eines Vektorraums  $V$  ist eine Teilmenge der Form  $v + U$ , wobei  $U \subset V$  ein Untervektorraum ist und  $v \in V$ .

## Bemerkung

Demnach ist der Lösungsraum eines inhomogenen linearen Gleichungssystems  $Ax = b$ ,  $A \in \text{Mat}(m, n, \mathbb{K})$ , 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  $0 = c_{k+1} = \dots = c_m$ .
- Die restlichen  $k$  Gleichungen löst man, indem man wieder die  $n - k$  Unbekannten  $x_j$  mit  $j \notin \{j_1, \dots, j_k\}$  frei wählt und die Lösungen  $x_{j_1}, \dots, x_{j_k}$  bestimmt.

**Zahlenbeispiel:** Wir wollen das inhomogene lineare Gleichungssystem

$$x_1 + 2x_2 + 3x_3 = 0$$

$$4x_1 + 5x_2 + 6x_3 = 3$$

$$7x_1 + 8x_2 + 9x_3 = 6$$

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

$$(A|b) = \begin{pmatrix} 1 & 2 & 3 & 0 \\ 4 & 5 & 6 & 3 \\ 7 & 8 & 9 & 6 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 2 & 3 & 0 \\ 0 & -3 & -6 & 3 \\ 0 & -6 & -12 & 6 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 2 & 3 & 0 \\ 0 & -3 & -6 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\rightarrow \begin{pmatrix} 1 & 2 & 3 & 0 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix} = (B|c).$$

Da  $c_3 = 0$ , ist  $Ax = b$  lösbar. Das zugehörige Gleichungssystem ist:

$$x_1 + 2x_2 + 3x_3 = 0$$

$$x_2 + 2x_3 = -1.$$

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 von Unterräumen

## 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** von  $U$  und  $V$ .

- Für die Unterräume

$$U \times \{0\} = \{(u, 0) \in U \times V \mid u \in U\} \subset U \oplus V$$

$$\{0\} \times V = \{(0, v) \in U \times V \mid v \in V\} \subset U \oplus V$$

gilt dann  $(U \times \{0\}) \cap (\{0\} \times V) = \{0\}$  und

$$U \oplus V = \{u + v \mid u \in U \times \{0\}, v \in \{0\} \times V\}$$

- Es ist  $\dim(U \oplus V) = \dim U + \dim V$ : Ist  $(u_1, \dots, u_m)$  eine Basis von  $U$  und  $(v_1, \dots, v_n)$  eine Basis von  $V$ , so ist  $((u_i, 0), (0, v_j))$  mit  $i = 1, \dots, m, j = 1, \dots, n$  eine Basis von  $U \oplus V$ .

# Innere Summe von Unterräumen

## Definition

$W, W'$  seien Unterräume eines Vektorraums  $V$ .

Die (innere) **Summe**

$$W + W' := \{w + w' \mid w \in W, w' \in W'\} \subset V$$

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  $F : W \oplus W' \rightarrow W + W'$ ,

$$(w, w') \mapsto F(w, w') := w - w'.$$

Offenbar ist  $F$  surjektiv:  $F(W \oplus W') = W + W'$  und daher  $\text{rg}(F) = \dim(W + W')$ . Ferner ist

$$\ker(F) = \{(w, w) \mid w \in W \cap W'\} \cong W \cap W'.$$

Es ist  $\dim(W \oplus W') = \dim W + \dim W'$ .

Daher folgt aus der Dimensionsformel für lineare Abbildungen

$$\begin{aligned} \text{rg}(F) &= \dim(W + W') = \dim(W \oplus W') - \dim \ker(F) \\ &= \dim(W \oplus W') - \dim(W \cap W') \\ &= \dim W + \dim W' - \dim(W \cap W'). \end{aligned}$$



# 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  $W \subset V$  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

$$U \oplus_{\text{äußer}} V = (U \times \{0\}) \oplus_{\text{inner}} (\{0\} \times V). \quad (3)$$

Durch die injektiven linearen Abbildungen

$$\begin{array}{ll} U & \rightarrow U \oplus_{\text{äußer}} V \quad \text{bzw.} \quad V & \rightarrow U \oplus_{\text{äußer}} V \\ u & \mapsto (u, 0) & v & \mapsto (0, v) \end{array}$$

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

$$v = w + w', \quad w \in W, \quad w' \in W'. \quad (4)$$

## Beweis.

' $\Rightarrow$ ' Aus  $V = W + W'$  folgt, dass jeder Vektor  $v \in V$  eine Darstellung der Form (4) 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 (4).

' $\Leftarrow$ ' Wenn umgekehrt jeder Vektor  $v \in V$  eine eindeutige Darstellung (4) hat, so gilt  $V = W + W'$  und aus  $0 = w + (-w)$  mit  $w \in W \cap W'$ , folgt wg. der Eindeutigkeit von (4)  $w = 0$ , d.h.  $W \cap W' = \{0\}$ .



## Satz

Sei  $V$  ein endlichdimensionaler Vektorraum. Dann gilt:

- a) Jeder Unterraum  $W \subset V$  besitzt ein Komplement.
- b)  $W, W' \subset V$  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  $(w_1, \dots, w_k)$  eine beliebige Basis des Unterraums  $W$ . Nach dem Austauschsatz von Steinitz können wir diese zu einer Basis  $(w_1, \dots, w_k, w'_1, \dots, w'_l)$  von  $V$  ergänzen.  
 $W' := \text{span}\{w'_1, \dots, w'_l\}$  ist dann ein Komplement zu  $W$ .
- b) Die Äquivalenz der Eigenschaften (i)-(iii) folgt aus der Dimensionsformel

$$\dim(W + W') = \dim W + \dim W' - \dim W \cap W'$$

# Kapitel 12

## Gruppen

# Gruppen

## Definition (Gruppe)

Eine **Gruppe**  $(G, \cdot)$  ist eine Menge  $G$  zusammen mit einer assoziativen Verknüpfung  $\cdot : G \times G \rightarrow G$ , so dass

(i) es ein **neutrales Element**  $e \in G$  gibt mit

$$e \cdot a = a \cdot e = a \quad \text{für alle } a \in G \quad \text{und}$$

(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) Die Bijektionen  $\varphi : X \rightarrow X$  einer Menge  $X$  in sich bilden eine Gruppe, die mit  $\text{Bij}(X)$  bezeichnet wird. Die Verknüpfung ist die Verkettung, das neutrale Element ist die identische Abbildung  $\text{Id}_X$  und das Inverse einer Bijektion ist ihre Umkehrabbildung.
- (ii) Im Fall einer endlichen Menge  $X$  nennt man die Bijektionen  $\sigma \in \text{Bij}(X)$  auch **Permutationen** von  $X$ . Die Permutationsgruppe  $\text{Bij}(X)$  ist dann eine endliche Gruppe mit  $n!$  Elementen, wobei  $n = \text{card}(X)$  die Anzahl der Elemente von  $X$  ist.
- (iii) Unter der  $n$ -ten **symmetrischen Gruppe** versteht man die Permutationsgruppe  $S_n := \text{Bij}(\{1, 2, \dots, n\})$ .
- (iv) Symmetrien von Objekten werden oft durch Gruppen beschrieben.
- (v) Sei  $V$  ein Vektorraum. Die Isomorphismen  $V \rightarrow V$  heißen auch **Automorphismen** des Vektorraums  $V$  und bilden die sogenannte Automorphismengruppe  $\text{Aut}(V)$  von  $V$ .

## weitere Beispiele von Gruppen

(vi) Im Fall eines endlichdimensionalen  $\mathbb{K}$ -Vektorraums  $V$  spricht man auch von der **allgemeinen linearen Gruppe** über  $\mathbb{K}$  und schreibt dafür  $GL(V) := \text{Aut}(V)$ .

Sei  $V = \mathbb{K}^n$ . Dann entspricht  $GL(V)$  gerade einer Gruppe von Matrizen, genauer der Gruppe der **invertierbaren**  $n \times n$ -Matrizen mit Einträgen aus  $\mathbb{K}$ , die auch mit  $GL(n, \mathbb{K})$  bezeichnet wird.

Es gilt

$$\begin{aligned} GL(n, \mathbb{K}) &= \{A \in \text{Mat}(n, \mathbb{K}) \mid \text{rg}(A) = n\} \\ &= \{A \in \text{Mat}(n, \mathbb{K}) \mid \ker(A) = \{0\}\}, \end{aligned}$$

wobei wir hier für den Lösungsraum von  $Ax = 0$  kurz  $\ker(A)$  schreiben. Diese Gruppe ist nicht kommutativ.

(vii)  $GL(1, \mathbb{K}) = (\mathbb{K}^* := \mathbb{K} \setminus \{0\}, \cdot)$  ist eine kommutative Gruppe.

(viii) Für jeden Vektorraum  $V$  ist durch  $(V, +)$  eine kommutative Gruppe gegeben.

## Noch mehr Beispiele von Gruppen

- (ix) Für jeden Körper  $\mathbb{K}$  sind durch  $(\mathbb{K}, +)$  und  $(\mathbb{K} \setminus \{0\}, \cdot)$  kommutative Gruppen gegeben.
- (x) Die Kreislinie  $S^1 := \{z \in \mathbb{C} \mid |z| = 1\}$  ist mit der Multiplikation komplexer Zahlen eine kommutative Gruppe. (Diese Gruppe spielt eine zentrale Rolle als Eichgruppe in der Elektrodynamik.)
- (xi) 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:  $(gh)(h^{-1}g^{-1}) = g(hh^{-1})g^{-1} = gg^{-1} = e$ .

Wenn ich nämlich morgens erst Strümpfe, dann Schuhe anziehe, so ziehe ich abends erst die Schuhe aus, dann die Strümpfe!

## Definition (Gruppenhomomorphismen)

Eine Abbildung  $\varphi : G \rightarrow H$  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  $\varphi : G \rightarrow H$  gibt.

Eine **Untergruppe** einer Gruppe  $G$  ist eine nicht-leere Teilmenge  $H \subset G$ , 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  $H \subset G$  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  $\varphi : G \rightarrow H$  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  $\varphi : G \rightarrow H$  ist wieder ein Isomorphismus von Gruppen.
- Das Bild  $\varphi(K) \subset H$  einer Untergruppe  $K \subset G$  unter einem Gruppenhomomorphismus  $\varphi : G \rightarrow H$  ist eine Untergruppe von  $H$
- Das Urbild  $\varphi^{-1}(K) \subset G$  einer Untergruppe  $K \subset H$  unter einem Gruppenhomomorphismus  $\varphi : G \rightarrow H$  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,  $\text{Aut}(G) \subset \text{Bij}(G)$ .
- Die Gruppen  $(\mathbb{R}, +)$  und  $(\mathbb{R}_+, \cdot)$  sind isomorph. Die Exponentialfunktion  $\exp : \mathbb{R} \rightarrow \mathbb{R}_+$  ist ein Isomorphismus von Gruppen, und damit auch ihre Umkehrfunktion  $\ln : \mathbb{R}_+ \rightarrow \mathbb{R}$ .
- Die Abbildung  $t \mapsto \exp(it)$  definiert einen surjektiven aber nicht injektiven Gruppenhomomorphismus  $\varphi : \mathbb{R} \rightarrow S^1$ . Sein Kern ist die Untergruppe  $\ker \varphi = 2\pi\mathbb{Z}$  von  $\mathbb{R}$ .

## 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:

$$\begin{aligned}\tau_{12} \circ \tau_{13}(1) &= \tau_{12}(3) = 3 \\ \tau_{13} \circ \tau_{12}(1) &= \tau_{13}(2) = 2\end{aligned}$$

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  $S_3 \rightarrow S_n$  ( $n \geq 3$ ). Das Bild ist eine nicht kommutative Untergruppe von  $S_n$ . Daher ist auch  $S_n$  für  $n \geq 3$  nicht kommutativ.  $\square$

## Satz

Für alle  $\sigma \in S_n$  ( $n \geq 2$ ) gibt es Transpositionen  $\tau_1, \dots, \tau_k \in S_n$ , so dass  $\sigma = \tau_1 \circ \dots \circ \tau_k$ . Die Darstellung von  $\sigma$  als Produkt ist nicht eindeutig.

### Beweis.

Für jede Transposition  $\tau$  gilt  $\tau^{-1} = \tau$  und somit  $\text{Id} = \tau \circ \tau$ .

Sei daher von nun an  $\text{Id} \neq \sigma \in S_n$ . 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  $\sigma_1 \neq \text{Id}$ , 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  $\tau_2 := \tau_{i_2 \sigma_1(i_2)}$  und  $\sigma_2 := \tau_2 \circ \sigma_1$ .

Durch Fortsetzen dieses Iterationsverfahren erhalten wir nach endlich vielen (genauer: nach  $k \leq n - 1$ ) Schritten  $\sigma_k = \tau_k \circ \dots \circ \tau_1 \circ \sigma = \text{Id}$  und somit  $\sigma = \tau_1 \circ \dots \circ \tau_k$ . □

## 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

$\varepsilon : S_n \rightarrow \{-1, 1\}$  ist ein Gruppenhomomorphismus in die Untergruppe  $\{-1, 1\} \subset \mathbb{R}^*$ .

**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

$$\begin{aligned} \prod_{i < j} f_{ij} &= \prod_{\substack{i < j \\ \tau(i) < \tau(j)}} f_{ij} \prod_{\substack{i < j \\ \tau(i) > \tau(j)}} f_{ij} \\ &\stackrel{(*)}{=} \prod_{\substack{i < j \\ \tau(i) < \tau(j)}} f_{ij} \prod_{\substack{i > j \\ \tau(i) < \tau(j)}} f_{ij} = \prod_{\tau(i) < \tau(j)} f_{ij} = \varepsilon(\sigma). \end{aligned}$$

Damit folgt  $\prod_{i < j} f_{ij} = \varepsilon(\sigma)$  und die Behauptung. Dieser Gruppenhomomorphismus spielt eine wichtige Rolle bei Fermionen.

## 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  $\varepsilon(\sigma) = \varepsilon(\tau_1 \circ \cdots \circ \tau_k) = \varepsilon(\tau_1) \cdots \varepsilon(\tau_k) = (-1)^k$ . □

## Beispiel

Ein **Zykel**  $(i_1, \dots, i_r)$  der Länge  $r$  ist ein  $r$ -Tupel von paarweise verschiedenen Zahlen aus  $\{1, \dots, n\}$  und wird aufgefasst als **zyklische Permutation**  $i_1 \mapsto i_2 \mapsto \dots \mapsto i_r \mapsto i_1$ . 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

$$\langle \zeta \rangle := \{\zeta, \zeta^2, \dots, \zeta^n = \text{Id}\} \subset S_n.$$

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  $S_3$ .

# Kapitel 13

## Determinanten

## Determinante einer $2 \times 2$ -Matrix

### Bemerkung:

Es sei  $A \in \text{Mat}(n, \mathbb{K})$  eine **quadratische** Matrix und  $b \in \mathbb{K}^n$ . Das durch  $Ax = b$  gegebene inhomogene lineare Gleichungssystem sei eindeutig lösbar.

Gesucht: Lösungsformel für die  $n$  Unbekannten  $x_1, x_2, \dots, x_n$ .

Im Spezialfall  $n = 2$ :

$$a_{11}x_1 + a_{12}x_2 = b_1$$

$$a_{21}x_1 + a_{22}x_2 = b_2$$

Mit elementaren Zeilenumformungen findet man im Fall, dass eine eindeutige Lösung existiert ( $\text{rg}(A) = n = 2$ ) die folgende Regel:

## Cramersche Regel für $n = 2$ :

$$x_1 = \frac{b_1 a_{22} - a_{12} b_2}{a_{11} a_{22} - a_{12} a_{21}}$$
$$x_2 = \frac{a_{11} b_2 - b_1 a_{21}}{a_{11} a_{22} - a_{12} a_{21}}$$

Dass dies wirklich eine Lösung ist, verifiziert man leicht durch Einsetzen.

### Definition

Der Ausdruck  $\text{Det}(A) := a_{11} a_{22} - a_{12} a_{21} \in \mathbb{K}$  heißt **Determinante** der  $(2 \times 2)$ -Matrix  $A$ .

# Determinante

## Definition (Determinante)

Eine Abbildung  $\det : \text{Mat}(n, \mathbb{K}) \rightarrow \mathbb{K}$  heißt **Determinante**, wenn sie folgende Eigenschaften hat:

(D1)  $\det$  ist linear in **jeder** Spalte, d.h.

$$\begin{aligned} & \det(a_1 \cdots a_{i-1} \quad \lambda a_i + \mu b_i \quad a_{i+1} \cdots a_n) \\ &= \lambda \det(a_1 \cdots a_i \cdots a_n) + \mu \det(a_1 \cdots a_{i-1} \quad b_i \quad a_{i+1} \cdots a_n) \end{aligned}$$

für alle Spaltenvektoren  $a_1, \dots, a_n, b_i \in \mathbb{K}^n$  und  $\lambda, \mu \in \mathbb{K}$ .

(D2)  $\det$  ist **alternierend**, d.h.  $\det A = 0$ , falls zwei Spalten von  $A$  übereinstimmen.

(D3)  $\det$  ist folgendermaßen normiert:  $\det(\mathbf{1}_n) = 1$ , wobei  $\mathbf{1}_n := (e_1 \cdots e_n)$  die Einheitsmatrix ist.

Wir werden zeigen, dass es **genau eine** Abbildung  $\det : \text{Mat}(n, \mathbb{K}) \rightarrow \mathbb{K}$  mit den Eigenschaften (D1)-(D3) gibt.

Zunächst zeigen wir, dass (D1)-(D3) eine Reihe weiterer Rechenregeln nach sich ziehen.

## Satz

Sei  $A = (a_1 \cdots a_n) \in \text{Mat}(n, \mathbb{K})$ . Aus (D1)-(D2) folgt für jede Permutation  $\sigma \in S_n$ :

$$(*) \quad \det(a_{\sigma(1)} \cdots a_{\sigma(n)}) = \varepsilon(\sigma) \det A.$$

**Beweis.** Aus (D1)-(D2) erhalten wir für  $i < j$

$$\begin{aligned} 0 &= \det(\cdots a_i + a_j \cdots a_i + a_j \cdots) \\ &= \underbrace{\det(\cdots a_i \cdots a_i \cdots)}_{=0} + \det(\cdots a_i \cdots a_j \cdots) \\ &\quad + \det(\cdots a_j \cdots a_i \cdots) + \underbrace{\det(\cdots a_j \cdots a_j \cdots)}_{=0} \\ &= \det(\cdots a_i \cdots a_j \cdots) + \det(\cdots a_j \cdots a_i \cdots). \end{aligned}$$

Hierbei ist die  $i$ -te und  $j$ -te Spalte angegeben. Die drei Auslassungspunkte stehen für  $a_1 \cdots a_{i-1}$ , sowie  $a_{i+1} \cdots a_{j-1}$  und  $a_{j+1} \cdots a_n$ .

Das beweist (\*) für jede Transposition  $\sigma = \tau_{ij}$ .

Der allgemeine Fall folgt nun daraus, dass sich jede Permutation als Produkt von Transpositionen schreiben lässt. □

## Satz

Sei  $A = (a_1 \cdots a_n) \in \text{Mat}(n, \mathbb{K})$  und  $\lambda \in \mathbb{K}$ . Aus (D1)-(D2) folgt, dass Addition des  $\lambda$ -fachen der  $j$ -ten Spalte von  $A$  zur  $i$ -ten Spalte von  $A$  ( $i \neq j$ ) den Wert der Determinante nicht ändert:

$$\det(a_1 \cdots a_{i-1} \ a_i + \lambda a_j \ a_{i+1} \cdots a_n) = \det(A).$$

**Beweis.** Aus der Linearität in der  $i$ -ten Spalte folgt:

$$\begin{aligned} & \det(a_1 \cdots a_{i-1} \ a_i + \lambda a_j \ a_{i+1} \cdots a_n) \\ \stackrel{(D1)}{=} & \det A + \lambda \det(\cdots a_j \cdots a_j \cdots) \stackrel{(D2)}{=} \det A. \end{aligned}$$

□

## Folgerung

Sei  $A = (a_1 \cdots a_n) \in \text{Mat}(n, \mathbb{K})$ . Aus (D1)-(D2) folgt: Falls  $\text{rg}(A) \neq n$ , so ist  $\det(A) = 0$ .

**Beweis.**  $\text{rg}(A) \neq n$  bedeutet, dass die Dimension des Bildes von  $A$  kleiner als  $n$  ist. D.h. aber, dass sich wenigstens eine der Spalten  $a_i$  als Linearkombination der anderen schreiben lässt, dass für ein  $i$  gilt

$$a_i = \sum_{j \neq i} \lambda_j a_j.$$

Daraus ergibt sich durch Zuaddieren die Spalten  $j \neq i$

$$\det A = \det(a_1 \cdots a_{i-1} \underbrace{a_i - \sum_{j \neq i} \lambda_j a_j}_{=0} a_{i+1} \cdots a_n) = 0.$$



## Berechnung der Determinante mittels Spaltenumformungen

### Satz

Jede invertierbare Matrix  $A \in \text{GL}(n, \mathbb{K})$  lässt sich durch wiederholtes Anwenden folgender zwei **Spaltenumformungen** in eine Diagonalmatrix  $A' = \text{diag}(\lambda_1, \dots, \lambda_n) := (\lambda_1 e_1 \cdots \lambda_n e_n)$  überführen:

- (S1) Vertauschen von zwei Spalten bei gleichzeitiger Multiplikation einer der beiden mit  $-1$  und
- (S2) Addition eines Vielfachen einer Spalte zu einer anderen Spalte.

### Folgerung

Für  $A$  invertierbar und  $A'$  wie im Satz gilt  $\det A = \det A' = \lambda_1 \cdots \lambda_n \neq 0$ .

### Beweis der Folgerung

Die Spaltenumformungen (S1)-(S2) ändern nicht den Wert der Determinante. Somit erhalten wir aus dem Satz:

$$\det A = \det A' \stackrel{(D1)}{=} \lambda_1 \cdots \lambda_n \det(\mathbf{1}_n) \stackrel{(D3)}{=} \lambda_1 \cdots \lambda_n.$$



**Beweis des Satzes:** Da  $A$  invertierbar ist, ist  $\text{rg}(A) = n$ . Daher können wir durch Anwendung des Gaußalgorithmus auf die Spalten (statt auf die Zeilen), mittels (S1)-(S2) die Matrix  $A$  in untere Dreiecksgestalt bringen:

$$\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ * & & \lambda_n \end{pmatrix}.$$

Diese Matrix hat Rang  $n$ , da die Umformungen (S1)-(S2) den Rang nicht ändern.

Also ist das Produkt  $\lambda_1 \cdots \lambda_n \neq 0$ .

Durch Umformungen vom Typ (S2) können wir deswegen alle Matrixeinträge unterhalb der Diagonalen eliminieren und erhalten die gewünschte Diagonalgestalt  $A' = \text{diag}(\lambda_1, \dots, \lambda_n)$ . □

Der Gaußalgorithmus liefert auch:

### Satz/ÜA

Für Blockdreiecksmatrizen gilt

$$\det \begin{pmatrix} A & B \\ 0 & D \end{pmatrix} = \det \begin{pmatrix} A & 0 \\ C & D \end{pmatrix} = \det A \det D.$$

Hierbei ist  $A$  eine  $(n \times n)$ -Matrix,  $D$  eine  $(m \times m)$ -Matrix,  $B$  eine  $(n \times m)$ -Matrix und  $C$  eine  $(m \times n)$ -Matrix.

Wir fassen einige der bisherigen Ergebnisse zusammen:

### Folgerung

*Es gibt höchstens eine Abbildung  $\det : \text{Mat}(n, \mathbb{K}) \rightarrow \mathbb{K}$  mit den Eigenschaften (D1)-(D3).*

*Unter der Voraussetzung, dass eine solche Abbildung existiert, gilt:*

*$A \in \text{Mat}(n, \mathbb{K})$  ist genau dann invertierbar, wenn  $\det A \neq 0$ .*

# Explizite Formel für die Determinante

## Satz (Explizite Formel für die Determinante)

*Es gibt genau eine Abbildung  $\det : \text{Mat}(n, \mathbb{K}) \rightarrow \mathbb{K}$  mit den Eigenschaften (D1)-(D3).*

*Die Determinante einer  $(n \times n)$ -Matrix  $A = (a_{ij})_{i,j}$  ist durch folgende Formel gegeben:*

$$\text{Det } A := \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(n)n}.$$

**Beweis.** Da wir bereits wissen, dass es *höchstens* eine Abbildung mit den Eigenschaften (D1)-(D3) gibt, genügt es (D1)-(D3) für die Abbildung  $\text{Det}$  nachzuweisen, um  $\text{Det}A = \det A$  zu zeigen.

**Beweis der expliziten Formel für die Determinante:**

(D1):

$$\begin{aligned} & \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{\sigma(1)1} \cdots (\lambda a_{\sigma(i)i} + \mu b_{\sigma(i)i}) \cdots a_{\sigma(n)n} \\ = & \lambda \operatorname{Det} A + \mu \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{\sigma(1)1} \cdots b_{\sigma(i)i} \cdots a_{\sigma(n)n}. \end{aligned}$$

## Weiter Beweis der expliziten Formel für die Determinante:

(D2): Es genügt zu zeigen, dass beim Vertauschen der Spalten

$$\underbrace{\text{Det}(a_{\tau(1)} \cdots a_{\tau(n)})}_{A_{\tau} :=} = -\text{Det} A$$

für jede Transposition  $\tau$  gilt. Denn im Falle  $A = A_{\tau}$ , der in (D2) betrachtet wird, folgt dann  $\text{Det} A = \text{Det} A_{\tau} = -\text{Det} A$  und somit  $\text{Det} A = 0$ , also (D2).

Mit  $\sigma' := \sigma\tau = \sigma \circ \tau$  haben wir  $\sigma = \sigma'\tau$ ,  $\varepsilon(\sigma') = -\varepsilon(\sigma)$  und somit

$$\begin{aligned} \text{Det}(A_{\tau}) &= \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{\sigma(1)\tau(1)} \cdots a_{\sigma(n)\tau(n)} \\ &= - \sum_{\sigma' \in S_n} \varepsilon(\sigma') a_{\sigma'(\tau(1))\tau(1)} \cdots a_{\sigma'(\tau(n))\tau(n)} \\ &= - \sum_{\sigma' \in S_n} \varepsilon(\sigma') a_{\sigma'(1)1} \cdots a_{\sigma'(n)n} = -\text{Det} A \end{aligned}$$

**Weiter im Beweis der expliziten Formel für die Determinante:**

(D3): Sei nun  $A = \mathbf{1}_n$ , d.h.

$$a_{ij} = \delta_{ij} := \begin{cases} 1, & \text{falls } i = j \\ 0 & \text{sonst} \end{cases} \quad (\delta_{ij} \text{ heißt Kroneckersymbol}).$$

Daraus folgt

$$\varepsilon(\sigma)\delta_{\sigma(1)1} \cdots \delta_{\sigma(n)n} = \begin{cases} 1, & \text{falls } \sigma = \text{Id} \\ 0 & \text{sonst} \end{cases}$$

und somit

$$\text{Det}(A) = \text{Det}(\mathbf{1}_n) = \sum_{\sigma \in S_n} \varepsilon(\sigma)\delta_{\sigma(1)1} \cdots \delta_{\sigma(n)n} = 1.$$

□

## Berechnung von Determinanten

Die Berechnung der Determinante einer Matrix  $A \in \text{Mat}(n, \mathbb{K})$  mittels der Formel

$$\det A = \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(n)n}$$

ist wegen des schnellen Wachstums von  $\text{card}(S_n) = n!$  für große  $n$  sehr aufwändig.

Beachten Sie, dass  $n!$  schneller wächst als jede Exponentialfunktion  $a^n$ ,  $a > 1$ .

Wir werden später (im nächsten Semester) effizientere Formeln zur Berechnung der Determinante kennenlernen.

## Regel von Sarrus:

Wir erhalten für  $n = 3$  direkt aus der Formel:

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32},$$

dabei entsprechen die ersten drei Terme den zyklischen Permutationen  $[123]$ ,  $[231]$  und  $[312]$ , die letzten drei den Transpositionen  $\tau_{13}$ ,  $\tau_{12}$  und  $\tau_{23}$ .

Man schreibt den 1. und 2. Spaltenvektor nochmal hinter die Matrix: Für die ersten drei Terme betrachtet man die Diagonale und ihre Parallelen (positives Vorzeichen), für die letzten drei die Nebendiagonale und ihre Parallelen (negatives Vorzeichen).

## Satz (Determinantenmultiplikationssatz)

Aus D1-D3 folgt für alle  $A, B \in \text{Mat}(n, \mathbb{K})$ :

$$\det(AB) = \det(A) \det(B).$$

### Beweis.

**1. Fall** Falls  $\text{rg}(A) < n$ , so ist  $\text{rg}(AB) < n$  und somit  $\det(AB) = 0 = \det(A) \det(B) = 0 \cdot \det(B)$ .

**2. Fall** Falls  $\text{rg}(B) < n$ , so ist  $\ker(B) \neq 0$  und somit  $\ker(AB) \neq 0$ . Also ebenfalls  $\det(AB) = 0 = \det(A) \det(B) = \det(A) \cdot 0$ .

**3. Fall** Seien also  $A, B$  invertierbar. Wir halten  $B$  mit  $\det B \neq 0$  fest und zeigen, dass die Funktion

$$\widetilde{\det} A := \frac{\det(A \cdot B)}{\det B}$$

eine die Axiome (D1)-(D3) erfüllt ist, woraus wegen der Eindeutigkeit einer solchen Funktion  $\widetilde{\det} A = \det A$  und somit die Behauptung folgt.



(D1a) Es folgt  $\tilde{A}B = \Delta_{\lambda,i}(A \cdot B)$ , d.h. auch  $\tilde{A}B$  entsteht aus  $AB$  durch Multiplikation der  $i$ -ten Zeile mit  $\lambda$ . Nach dem Axiom (D1) für die Determinantenfunktion  $\det$  folgt

$$\det(\tilde{A}B) = \lambda \det AB$$

und somit

$$\widetilde{\det A} = \frac{\det(\tilde{A}B)}{\det B} = \lambda \frac{\det AB}{\det B} = \lambda \widetilde{\det A}.$$

(D1a) Um die Additivität von  $\widetilde{\det}$  zu zeigen, drücken wir die Matrix  $A$  durch Zeilenvektoren  $a_j \in K^n$  und die Matrix  $B$  durch Spaltenvektoren  $b_j \in K^n$  aus.

$$A = \begin{pmatrix} (a_1)^t \\ \vdots \\ (a_n)^t \end{pmatrix} \quad B = (b_1, \dots, b_n) ,$$

Es gelte für die  $i$ -te Zeile  $a_i = a'_i + a''_i$ . Dann ist

$$\begin{aligned} A \cdot B &= \begin{pmatrix} (a_1)^t b_1 & \dots & (a_1)^t b_n \\ \vdots & & \vdots \\ (a_n)^t b_1 & \dots & (a_n)^t b_n \end{pmatrix} \\ &= ((a'_i + a''_i)^t b_1 \dots (a'_i + a''_i)^t b_n) \end{aligned}$$

und somit für die  $i$ -te Zeile  $((a'_i + a''_i)^t b_1 \dots (a'_i + a''_i)^t b_n)$ . Aus der Additivität von  $\det$  folgt  $\det AB = \det A'B + \det A''B$  und daraus die von  $\widetilde{\det}$ :

$$\widetilde{\det} A = \frac{\det AB}{\det B} = \widetilde{\det} A' + \widetilde{\det} A'' .$$

## Folgerung

(i)

$$\det : \text{GL}(n, \mathbb{K}) \rightarrow (\mathbb{K} \setminus \{0\}, \cdot)$$

ist ein Gruppenhomomorphismus.

(ii) Insbesondere gilt für alle  $A \in \text{GL}(n, \mathbb{K})$ :

$$\det(A^{-1}) = (\det A)^{-1}$$

(iii)

$$\text{SL}(n, \mathbb{K}) := \{A \in \text{GL}(n, \mathbb{K}) \mid \det A = 1\} \subset \text{GL}(n, \mathbb{K}).$$

ist eine Untergruppe, die sogenannte **spezielle lineare Gruppe**.

**Beweis.** (i-ii) sind klar. (iii) folgt, da  $\text{SL}(n, \mathbb{K}) = \det^{-1}(1)$  ein Kern ist.  $\square$

**Achtung:** Für  $n \geq 2$  gilt **nicht**  $\det(A + B) = \det(A) + \det(B)$ !

Gegenbeispiel  $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$  und  $B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ .

## Definition (Determinante eines Endomorphismus)

Sei  $V$  ein  $\mathbb{K}$ -Vektorraum der Dimension  $n$ ,  $F \in \text{End}(V)$  ein Endomorphismus und  $A = M_B(F) \in \text{Mat}(n, \mathbb{K})$  seine darstellende Matrix bezüglich einer beliebigen Basis  $B$  von  $V$ . Wir definieren die Determinante von  $F$  als:

$$\det F := \det A.$$

**Behauptung:** Dies ist wohldefiniert, d.h. die Definition hängt nicht von der Wahl der Basis  $B$  ab.

**Beweis.** Sei  $B'$  eine weitere Basis von  $V$ ,  $A' := M_{B'}(F)$ . Dann entspricht  $A$  einer Selbstabbildung  $\phi_B^{-1} \circ F \circ \phi_B$  von  $\mathbb{R}^n$  und  $A'$  einer Abbildung  $\phi_{B'}^{-1} \circ F \circ \phi_{B'}$ .

Letztere schreibt sich auch als

$$\underbrace{\phi_{B'}^{-1} \circ \phi_B}_{=: S} \circ \underbrace{\phi_B^{-1} \circ F \circ \phi_B}_{= A} \circ \underbrace{\phi_B^{-1} \circ \phi_{B'}}_{=: S^{-1}}$$

d.h.  $A' = SAS^{-1}$  mit der invertiblen  $n \times n$ -Matrix  $S$ .

Also ist  $\det A' = \det(SAS^{-1}) = \det S \det A (\det S)^{-1} = \det A$ . □

# Organisatorisches

Alle Informationen erhalten Sie weiterhin über unsere Webseite:

<http://www.math.uni-hamburg.de/master/lehreexport/physik/mfp2.html>

Dies betrifft:

- Prüfungstermine
- Regeln für die Klausur, einschließlich erlaubter Hilfsmittel (keine)
- Bonusregelung
- Anmeldungen zu Modul, zur Prüfung und zu den Übungen (Nachmeldung über die Studienbüros)

Bitte lesen Sie regelmäßig, mindestens zwei Mal wöchentlich, die Webseite.

## Tutorien Physik

Di 12.00-14.00    Sitzungszimmer AP    Tobias Kroker + Keno Riechers

Fr 12.00-14.00    Bibliothek AP                    Niklas Mann + Matthias Peschke

Beginn: 09.04. bzw. zweite Sitzung 19.04.2013

Die Räumlichkeiten befinden sich in der Jungiusstraße 11, Eingang A.

# Orientierung

## Definition

Sei  $V$  ein endlichdimensionaler **reeller** Vektorraum. Zwei (geordnete) Basen  $B = (b_1, \dots, b_n)$  und  $B' = (b'_1, \dots, b'_n)$  heißen **gleich orientiert**, wenn der **Basiswechsel**  $F = \phi_{B'} \circ \phi_B^{-1} : V \rightarrow V$  positive Determinante hat.

## Bemerkungen

- Beachte, dass  $F(b_i) = \phi_{B'} \phi_B^{-1}(b_i) = \phi_{B'}(e_i) = b'_i$ .
- Die darstellende Matrix von  $F$  bezüglich der Basis  $B$  ist gegeben durch die Abbildung  $\phi_B^{-1} \circ (\phi_{B'} \circ \phi_B^{-1}) \circ \phi_B = \phi_B^{-1} \circ \phi_{B'} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ . Insbesondere:

$$\det F = \det(\phi_B^{-1} \circ \phi_{B'}).$$

# Äquivalenzrelation

## Definition

- Eine **Relation** auf einer Menge  $X$  ist eine Teilmenge  $R \subset X \times X$ . Wir schreiben auch  $xRy$ , genau dann, wenn  $(x, y) \in R$  gilt.
- Eine Relation  $\sim$  auf  $X$  heißt **Äquivalenzrelation**, wenn sie folgende Eigenschaften hat:
  - (i)  $x \sim x$  für alle  $x \in X$ , (reflexiv)
  - (ii)  $x \sim y \implies y \sim x$  (symmetrisch)
  - (iii)  $x \sim y$  und  $y \sim z \implies x \sim z$  (transitiv)

Jedes Element  $x \in X$  definiert eine **Äquivalenzklasse**

$$[x] := \{y \in X \mid y \sim x\}.$$

Die Menge der Äquivalenzklassen wird mit  $X / \sim$  bezeichnet.

## Beispiel

$X$  sei die Menge der Schüler einer Schule,  $\sim$  die Äquivalenzrelation "in die gleiche Klasse gehen". Dann ist  $X / \sim$  die Menge der Klassen (mit der Stundenplanmacher arbeiten).

# Quotientenraum

## Beispiel

Sei  $V$  ein Vektorraum und  $U \subset V$  ein Unterraum. Dann definieren wir eine Äquivalenzrelation  $\sim$  auf  $V$  wie folgt:

$$x \sim y \iff x - y \in U.$$

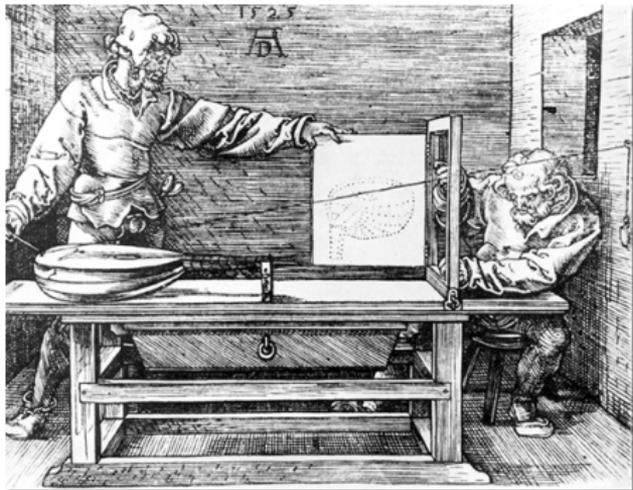
Also  $R = \{(x, y) \mid x, y \in V \text{ und } x - y \in U\} \subset V \times V$ .

$V/U := \{[x] \mid x \in V\}$  ist (in kanonischer Weise) ein Vektorraum, wobei die Abbildung  $x \mapsto [x]$  eine lineare Abbildung ist.

$V/U$  heißt **Quotientenvektorraum** von  $V$  nach  $U$ .

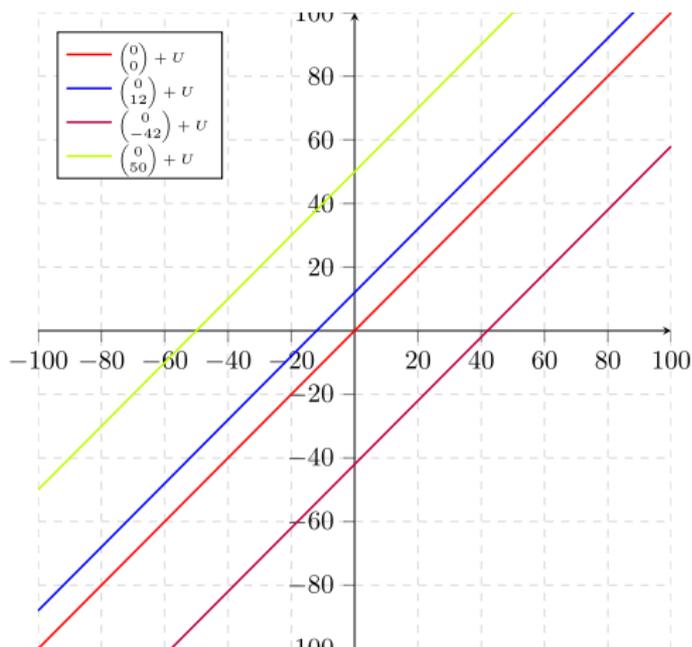
# Dürer

Auf einem Bild von Albrecht Dürer (1471-1528) sieht man zwei Männer, die sich sehr bemühen, die Äquivalenzklassen bezüglich eines eindimensionalen Unterraums von  $\mathbb{R}^3$  zu verstehen.



## Ein mathematisches Bild

Hier sehen Sie einige Äquivalenzklassen von  $\mathbb{R}^2$  nach dem in hellem Rot gezeichneten ein-dimensionalen Unterraum:



# Orientierung

## Satz

Sei  $V$  ein reeller Vektorraum der Dimension  $n \in \mathbb{N}$ . Wir schreiben  $B \sim B'$ , falls  $B$  und  $B'$  gleich orientierte (geordnete) Basen von  $V$  sind.

- (a) Die Relation “ $\sim$ ” ist eine Äquivalenzrelation auf der Menge der Basen von  $V$ .
- (b) Jede geordnete Basis  $B$  definiert eine Äquivalenzklasse (ihre “Händigkeit”)

$$[B] := \{B' \text{ Basis von } V \mid B' \sim B\}$$

und es gibt genau zwei Äquivalenzklassen.

## Beweis.

- a) ist eine einfache Übungsaufgabe.
- b) Sei  $B = (b_1, \dots, b_n)$  eine geordnete Basis. Jede geordnete Basis ist entweder gleich orientiert zu  $B$  oder zu  $(-b_1, b_2, \dots, b_n)$ . Daher gibt es genau zwei Äquivalenzklassen.

# Orientierung

## Definition

Sei  $V$  ein reeller Vektorraum der Dimension  $n \in \mathbb{N}$ . Eine **Orientierung** von  $V$  ist eine Äquivalenzklasse gleich orientierter geordneter Basen von  $V$ .

## Bemerkungen

- Jede Basis  $B = (b_1, \dots, b_n)$  definiert eine Orientierung  $[B]$ . Wie wir gesehen haben, gibt es genau zwei Orientierungen  $[B]$  und  $[B]^{op} = [(-b_1, b_2, \dots, b_n)]$ . Letztere heißt die zu  $[B]$  **entgegengesetzte** (oder **umgekehrte**) Orientierung.
- Automorphismen  $F \in GL(V)$  mit  $\det F > 0$  heißen **orientierungserhaltend**, denn  $FB = (Fb_1, \dots, Fb_n) \in [B]$  für jede Basis  $B = (b_1, \dots, b_n)$ . Ein Automorphismus  $F \in GL(V)$  mit  $\det F < 0$  heißt **orientierungsumkehrend**, denn  $FB = (Fb_1, \dots, Fb_n) \in [B]^{op}$  für jede Basis  $B$ .

# Orientierung

## Beispiele I

- Die kanonische Orientierung des  $\mathbb{R}^n$  ist die durch die kanonische Basis definierte Orientierung  $[(e_1, \dots, e_n)]$ . Die Basen  $(e_2, e_3, e_1)$  und  $(-e_2, e_1, e_3)$  definieren auch die **kanonische Orientierung** des  $\mathbb{R}^3$ .
- Die durch die Basis  $(e_2, e_1, e_3)$  definierte Orientierung ist zur durch die kanonische Basis definierten Orientierung  $[(e_1, e_2, e_3)]$  entgegengesetzt. Es gilt  $[(e_2, e_1, e_3)] = [(-e_1, e_2, e_3)]$ .
- Die Drehung (um die  $z$ -Achse, entgegen dem Uhrzeigersinn um den Winkel  $\varphi$ )

$$D = \begin{pmatrix} \cos \varphi & -\sin \varphi & 0 \\ \sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

ist wegen  $\det D = 1 > 0$  orientierungserhaltend.

# Orientierung

## Beispiele II

Die Spiegelung (an der  $(x, y)$ -Ebene)

$$S = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

ist wegen  $\det S = -1 < 0$  orientierungsumkehrend.

### Definition (Transponierte Matrix)

Sei  $A = (a_{ij})_{i,j}$  eine  $(m \times n)$ -Matrix.

Die  $(n \times m)$ -Matrix  $A^t$  mit den Einträgen

$$(A^t)_{ij} := a_{ji}$$

heißt die zu  $A$  **transponierte Matrix**.

### Bemerkung

Die Abbildung  $\alpha : \text{Mat}(n, \mathbb{K}) \ni A \mapsto A^t \in \text{Mat}(n, \mathbb{K})$ , ist eine **Involution** auf der Menge  $\text{Mat}(n, \mathbb{K})$ , d.h.  $\alpha \circ \alpha = \text{Id}$ .

## Satz

Für alle  $A \in \text{Mat}(m, n, \mathbb{K})$  und  $B \in \text{Mat}(n, r, \mathbb{K})$  gilt

$$(AB)^t = B^t A^t.$$

**Beweis.** Für  $C = AB$  und  $D = B^t A^t$  gilt

$$(C^t)_{ki} = c_{ik} = \sum_{j=1}^n a_{ij} b_{jk} = \sum_{j=1}^n (B^t)_{kj} (A^t)_{ji} = (D)_{ki}.$$



## Folgerung

(i) *Die Involution*

$$\begin{aligned} \mathrm{GL}(n, \mathbb{K}) &\rightarrow \mathrm{GL}(n, \mathbb{K}) \\ A &\mapsto (A^{-1})^t \end{aligned}$$

*ist ein Gruppenautomorphismus von  $\mathrm{GL}(n, \mathbb{K})$ .*

(ii) *Für alle  $A \in \mathrm{GL}(n, \mathbb{K})$  gilt  $(A^{-1})^t = (A^t)^{-1}$ .*

**Beweis.** Übung.



## Satz

Für jede Matrix  $A = (a_{ij})_{i,j} \in \text{Mat}(n, \mathbb{K})$  gilt

$$\begin{aligned} \det A &= \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(n)n} \\ &= \sum_{\tau \in S_n} \varepsilon(\tau) a_{1\tau(1)} \cdots a_{n\tau(n)} = \det A^t. \end{aligned}$$

**Beweis.** Die Substitution  $\tau = \sigma^{-1}$  liefert wegen  $\varepsilon(\tau) = \varepsilon(\sigma)$  die Behauptung. □

## Folgerung

(i) *Ebenso wie durch die Eigenschaften*

(D1) *det ist linear in jeder Spalte,*

(D2) *det ist alternierend bzgl. der Spalten, d.h.  $\det A = 0$ , falls zwei Spalten von  $A$  übereinstimmen,*

(D3) *det ist normiert:  $\det(\mathbf{1}_n) = 1$*

*lässt sich  $\det : \text{Mat}(n, \mathbb{K}) \rightarrow \mathbb{K}$  durch die folgenden Eigenschaften charakterisieren:*

(D1') *det ist linear in jeder Zeile,*

(D2') *det ist alternierend bzgl. der Zeilen, d.h.  $\det A = 0$ , falls zwei Zeilen von  $A$  übereinstimmen,*

(D3')  *$\det(\mathbf{1}_n) = 1$ .*

## Folgerung

- (ii) *Addition des  $\lambda$ -fachen der  $j$ -ten Zeile der Matrix  $A$  zur  $i$ -ten Zeile ( $i \neq j$ ) ändert nicht den Wert der Determinante.*
- (iii) *Vertauschen der  $i$ -ten und  $j$ -ten Zeile ( $i \neq j$ ) ändert die Determinante nur um ein Vorzeichen.*
- (iv) *Insbesondere lässt sich die Determinante durch Anwendung des Gaußalgorithmus auf die Zeilen der Matrix berechnen (bei Buchführung über die Zeilenvertauschungen!).*

## Matrixinversion mittels Gaußalgorithmus

Wir wollen die Inverse  $A^{-1}$  zu  $A \in GL(n, \mathbb{K})$  bestimmen. Dazu erinnern wir uns an:

- $i$ -te Spalte der Matrix  $A^{-1}$  ist  $A^{-1}e_i$ , wobei  $e_i$  der  $i$ -te Vektor der kanonischen Basis ist.
- Nun ist  $\mathbf{x} := A^{-1}e_i$ , also die  $i$ -te Spalte von  $A^{-1}$ , die eindeutige Lösung von des Gleichungssystems  $A\mathbf{x} = e_i$ .
- D.h. wir können  $\mathbf{x}$  bestimmen, indem wir den Gaußschen Algorithmus auf  $(A|e_i)$  anwenden bis wir  $(\mathbf{1}_n|\mathbf{x})$  erhalten.

Somit haben wir die  $i$ -te Spalte  $\mathbf{x} = A^{-1}e_i$  von  $A^{-1}$  ermittelt. Das Verfahren auf alle Basisvektoren  $e_i$  gleichzeitig angewandt

$$(A \mid \mathbf{1}_n) \xrightarrow{\text{Gaußscher Alg.}} (\mathbf{1}_n \mid A')$$

liefert die inverse Matrix  $A^{-1} = A'$ .

## Ein Zahlenbeispiel:

Wir wollen  $A = \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 2 \\ 1 & 1 & -1 \end{pmatrix}$  invertieren (das geht, da  $\det A = -3$ ):

$$\left( \begin{array}{ccc|ccc} 1 & 1 & 2 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 & 1 & 0 \\ 1 & 1 & -1 & 0 & 0 & 1 \end{array} \right) \xrightarrow{III-I} \left( \begin{array}{ccc|ccc} 1 & 1 & 2 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 & 1 & 0 \\ 0 & 0 & -3 & -1 & 0 & 1 \end{array} \right) \xrightarrow{I-II, III \cdot (-1/3)}$$

$$\left( \begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & -1 & 0 \\ 0 & 1 & 2 & 0 & 1 & 0 \\ 0 & 0 & 1 & \frac{1}{3} & 0 & -\frac{1}{3} \end{array} \right) \xrightarrow{II-2 \cdot III} \left( \begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & -\frac{2}{3} & 1 & \frac{2}{3} \\ 0 & 0 & 1 & \frac{1}{3} & 0 & -\frac{1}{3} \end{array} \right),$$

$$\text{d.h. } A^{-1} = \begin{pmatrix} 1 & -1 & 0 \\ -\frac{2}{3} & 1 & \frac{2}{3} \\ \frac{1}{3} & 0 & -\frac{1}{3} \end{pmatrix}.$$

Als explizite Probe rechne man  $AA^{-1} = \mathbf{1}_3$  nach.

# Berechnung der Inversen mittels Determinanten

Dazu benötigt man den Begriff der Streichungsmatrix.

## Definition (Streichungsmatrix)

Sei  $A = (a_{kl})_{k,l=1\dots n} \in \text{Mat}(n, \mathbb{K})$  und  $1 \leq i, j \leq n$  zwei Indizes. Die Matrix  $A_{ij} \in \text{Mat}(n-1, \mathbb{K})$ , die aus  $A$  durch Streichen der  $i$ -ten Zeile und  $j$ -ten Spalte entsteht, heißt  $(i, j)$ -te **Streichungsmatrix** von  $A$ .

## Lemma

- (i) die  $(i, j)$ -te Streichungsmatrix von  $A^t$  stimmt mit der Matrix  $(A_{ji})^t$  überein.
- (ii)  $(-1)^{i+j} \det A_{ij} = \det(\mathbf{a}_1 \cdots \mathbf{a}_{j-1} \mathbf{e}_i \mathbf{a}_{j+1} \cdots \mathbf{a}_n)$ .

**Beweis.** (i) ist klar.

**Beweis von (ii).** Die  $(n \times n)$ -Matrix  $(\mathbf{a}_1 \cdots \mathbf{a}_{j-1} \mathbf{e}_j \mathbf{a}_{j+1} \cdots \mathbf{a}_n)$  lässt sich durch Addition von Vielfachen der  $j$ -ten Spalte zu den anderen Spalten in folgende Form bringen:

$$B_{ij} := \begin{pmatrix} a_{11} & \cdots & a_{1j-1} & 0 & a_{1j+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{i-11} & \cdots & a_{i-1j-1} & 0 & a_{i-1j+1} & \cdots & a_{i-1n} \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ a_{i+11} & \cdots & a_{i+1j-1} & 0 & a_{i+1j+1} & \cdots & a_{i+1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nj-1} & 0 & a_{nj+1} & \cdots & a_{nn} \end{pmatrix}$$

Die Matrix  $B_{ij}$  lässt sich durch  $i - 1$  Zeilen- und  $j - 1$  Spaltenvertauschungen in die Blockdiagonalgestalt  $\begin{pmatrix} 1 & 0 \\ 0 & A_{ij} \end{pmatrix}$  bringen.

Also  $\det(\mathbf{a}_1 \cdots \mathbf{a}_{j-1} \mathbf{e}_j \mathbf{a}_{j+1} \cdots \mathbf{a}_n) =$

$$\det B_{ij} = (-1)^{(i-1)+(j-1)} \det A_{ij} = (-1)^{i+j} \det A_{ij}.$$

□

## Satz

Sei  $A = (a_{ij})_{i,j} \in \text{Mat}(n, \mathbb{K})$  und  $\tilde{A} = (\tilde{a}_{ij})_{i,j}$  definiert durch

$$\tilde{a}_{ij} := (-1)^{i+j} \det A_{ji} \stackrel{\text{Lemma}}{=} \det(\mathbf{a}_1 \cdots \mathbf{a}_{i-1} \mathbf{e}_j \mathbf{a}_{i+1} \cdots \mathbf{a}_n).$$

Dann gilt  $\tilde{A}A = A\tilde{A} = (\det A)\mathbf{1}_n$ .

**Beweis.** Wir zeigen zuerst  $\tilde{A}A = (\det A)\mathbf{1}_n$ :

$$\begin{aligned} \sum_j \tilde{a}_{ij} a_{jk} &\stackrel{\text{Lemma}}{=} \sum_j a_{jk} \det(\mathbf{a}_1 \cdots \mathbf{a}_{i-1} \mathbf{e}_j \mathbf{a}_{i+1} \cdots \mathbf{a}_n) \\ &= \det(\mathbf{a}_1 \cdots \mathbf{a}_{i-1} \underbrace{\sum_j a_{jk} \mathbf{e}_j}_{=\mathbf{a}_k} \mathbf{a}_{i+1} \cdots \mathbf{a}_n) = \delta_{ik} \det A. \end{aligned}$$

**Weiter im Beweis:** Ist  $\det(A) \neq 0$ , so erhalten wir aus  $\tilde{A}A = (\det A)\mathbf{1}_n$  auch  $A\tilde{A} = \det(A)\mathbf{1}_n$ , und zwar durch Multiplikation von  $\tilde{A}A = (\det A)\mathbf{1}_n$  von links mit  $A$  und von rechts mit  $A^{-1}$ .

Ist aber  $0 = \det(A) = \det(A^t)$ , so erhält man

$$0 = (\widetilde{A^t})A^t \stackrel{\text{Lemma}(i)}{=} (\tilde{A})^t A^t = (A\tilde{A})^t$$

und damit auch  $A\tilde{A} = 0$ . □

## Folgerung (Entwicklungssatz von Laplace)

Für jede Matrix  $A \in \text{Mat}(n, \mathbb{K})$  gilt:

(Z) (Entwicklung nach der  $i$ -ten Zeile)

$$\det A = \sum_{j=1}^n (-1)^{i+j} a_{ij} \det A_{ij}$$

(S) (Entwicklung nach der  $j$ -ten Spalte)

$$\det A = \sum_{i=1}^n (-1)^{i+j} a_{ij} \det A_{ij}.$$

**Beweis.** Auf der rechten Seite von (Z) bzw (S) steht, siehe Lemma (ii), genau der  $i$ -te bzw.  $j$ -te Diagonaleintrag der Matrix  $\tilde{A}A = (\det A)\mathbf{1}_n$  bzw. der Matrix  $\tilde{A}A = (\det A)\mathbf{1}_n$ . □

## Folgerung (Cramersche Regel)

Sei  $A \in \text{GL}(n, \mathbb{K})$  und  $\mathbf{b} \in \mathbb{K}^n$ . Dann gilt:

- (i)  $A^{-1} = \frac{1}{\det A} \tilde{A}$ .
- (ii) Die eindeutig bestimmte Lösung  $\mathbf{x} = (x_i)_{i=1 \dots n}$  der Gleichung  $A\mathbf{x} = \mathbf{b}$  berechnet sich nach der folgenden **Cramerschen Regel**

$$x_i = \frac{1}{\det A} \det(\mathbf{a}_1 \cdots \mathbf{a}_{i-1} \mathbf{b} \mathbf{a}_{i+1} \cdots \mathbf{a}_n).$$

### Beweis.

- (i) folgt sofort aus  $\det A \neq 0$  und  $\tilde{A}A = (\det A)\mathbf{1}_n$ .
- (ii) Mit  $\tilde{a}_{ij} = \det(\mathbf{a}_1 \cdots \mathbf{a}_{i-1} \mathbf{e}_j \mathbf{a}_{i+1} \cdots \mathbf{a}_n)$  und  $\mathbf{x} = A^{-1}\mathbf{b}$  ist

$$\begin{aligned} x_i &= \frac{1}{\det A} \sum_j \tilde{a}_{ij} b_j = \frac{1}{\det A} \det(\mathbf{a}_1 \cdots \mathbf{a}_{i-1} \sum_j b_j \mathbf{e}_j \mathbf{a}_{i+1} \cdots \mathbf{a}_n) \\ &= \frac{1}{\det A} \det(\mathbf{a}_1 \cdots \mathbf{a}_{i-1} \mathbf{b} \mathbf{a}_{i+1} \cdots \mathbf{a}_n) \quad \square \end{aligned}$$

## Cramersche Regel für $n = 2$ :

$$x_1 = \frac{b_1 a_{22} - a_{12} b_2}{a_{11} a_{22} - a_{12} a_{21}}$$

$$x_2 = \frac{a_{11} b_2 - b_1 a_{21}}{a_{11} a_{22} - a_{12} a_{21}}$$

Und (vgl. Spezialfälle:  $\mathbf{b} = \mathbf{e}_1$  bzw.  $\mathbf{b} = \mathbf{e}_2$ ):

$$A^{-1} = \frac{1}{a_{11} a_{22} - a_{12} a_{21}} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}.$$

## Ein Zahlenbeispiel für den Entwicklungssatz von Laplace:

Wir wollen die Determinante von  $A = \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 2 \\ \pi & 1 & -1 \end{pmatrix}$  durch Entwicklung nach der ersten Spalte berechnen:

$$\begin{aligned} \det \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 2 \\ \pi & 1 & -1 \end{pmatrix} &= 1 \cdot \det A_{11} - 0 \cdot \det A_{21} + \pi \cdot \det A_{31} \\ &= 1 \cdot \det \begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix} + \pi \cdot \det \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} = -3 + \pi \cdot 0 = -3 \end{aligned}$$