Mathematik I für Studierende der Physik

Vorlesungsskript[1]

Thomas Leistner
Department Mathematik
Universität Hamburg
www.math.uni-hamburg.de/home/leistner

Hamburg, Wintersemester 2008/09

[1]Version vom 4. Februar 2009

Inhaltsverzeichnis I

Inhaltsverzeichnis II

Inhaltsverzeichnis III

Inhaltsverzeichnis IV

Inhaltsverzeichnis V

Inhaltsverzeichnis VI

Inhaltsverzeichnis VII

Mathematische Kurzschreibweise

Man verwendet in mathematischen Texten die Bezeichnungen

$\Rightarrow$	für	"daraus folgt"	(Implikation)
$\iff$	für	"genau dann, wenn"	(Äquivalenz)
$\forall$	für	"für alle"	(Allquantor)
$\exists$	für	"es existiert ein"	(Existenzquantor)
$\exists!$	für	"es existiert genau ein"	
$\in$	für	"Element in"	
$:$ oder $\mid$	für	"so dass" oder "mit der Eigenschaft"	

Außerdem benutzt man $:\iff$ als definitorisches "genau dann, wenn".

Beispiel:

Seien $\mathbb{N} = \{0, 1, 2, 3, 4, ...\}$ die natürlichen Zahlen.
Die geraden natürlichen Zahlen $2\mathbb{N} = \{0, 2, 4, 6, \ldots\}$ sind gegeben durch

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

Mengen

Definition

- *Eine Menge ist eine Zusammenfassung wohlunterschiedener Objekte, welche Elemente genannt werden (G. Cantor, 1845-1918).*
- *Die Menge ohne Elemente ist die leere Menge:* $\emptyset$.
- *Sei* M *eine Menge.* N *ist Teilmenge von* M $:\iff \quad (\forall x \in N \Rightarrow x \in M)$. *Man schreibt* $N \subset M$.

Man benutzt die mathematische Schreibweise

$$\{x, y, z, \ldots\}$$

um die Objekte $x, y, z, \ldots$ zu einer (ungeordneten) Menge zusammenzufassen. Oder auch

$$N := \{x \in M \mid x \text{ hat die Eigenschaft ... }\}.$$

Beispiel:

$$2\mathbb{N} = \{0, 2, 4, 6, ...\} = \{x \in \mathbb{N} \mid \exists y \in \mathbb{N} : x = 2y\}$$

Mengenoperationen

Definition (Mengenoperationen)

Es seien M, N zwei beliebige Mengen, dann heißen

- $M \cap N = \{p \mid p \in M \text{ und } p \in N\}$ *der* *Durchschnitt*,
- $M \cup N = \{p \mid p \in M \text{ oder } p \in N\}$ *die* *Vereinigung*,
- $M \setminus N = \{p \in M \mid p \notin N\}$ *der* *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:
- $\mathbb{R}^n = \mathbb{R} \times \ldots \times \mathbb{R} = \{(x_1, \ldots, x_n) \mid x_i \in \mathbb{R}\}$.

Einige Rechenregeln für Mengen

Rechenregeln für Mengen

Seien M, N, L beliebige Mengen. Dann gilt

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

Beweis. Übung $\square$

Mathematische Aussagen und Beweise

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

Es gibt eine Vielzahl von Beweismethoden, z.B.

- direkter Beweis (im Wesentlichen Umformen von Gleichungen), z.B. " $4x = 12 \Rightarrow x = 3$"
- indirekter Beweis (auch Beweis durch Widerspruch): beruht auf der logischen Regel ("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.
- Beweis mittels vollständiger Induktion.

Mehr zu Beweisen finden Sie unter `http://www.mathematik.de/mde/information/ landkarte/stichpunkte/beweis.html`

Beweis durch vollständige Induktion

Prinzip der vollständigen Induktion

Seien $A(1), A(2), \ldots, A(n), \ldots$ Aussagen, die für alle natürlichen Zahlen $n \in \mathbb{N}$ formuliert sind, und die die folgenden Eigenschaften erfüllen

i) $A(1)$ 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 $n \in \mathbb{N}$.

Möchte man nun nachweisen, daß $A(n)$ für alle natürlichen Zahlen gilt, so muss man folgendes beweisen:

1. Induktionsanfang: $A(1)$ ist richtig,
2. Induktionsschritt: $A(n)$ 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* $\frac{1}{2}\ n(n+1)$,

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

Beweis. (mit vollständiger Induktion nach n)

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

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. $\square$

Ein weiteres Beispiel

Satz (Geometrische Summenformel)

Es gilt die Formel

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

Beweis. … □

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 $f(1)$.
2. Eine Vorschrift F, die für jedes n in $\mathbb{N}$ die Zahl $f(n+1)$ aus den Zahlen $f(1), \ldots, f(n)$ berechnet, d.h.

$$f(n+1) \;=\; F(f(1), \ldots, f(n))$$

Beispiele:

1. Potenzen: $x^0 \;:= 1\,, \quad x^1 \;:=\; x, \quad x^{n+1} \;:=\; x \cdot x^n$
2. Fakultät: $0! \;:=\; 1, \quad 1! \;:=\; 1, \quad (n+1)! \;:=\; (n+1) \cdot n!$
3. Fibonacci Zahlen: $x_1 \;:=\; 0, \quad x_2 \;:=\; 1, \quad x_{n+1} \;:=\; x_n \,+\, x_{n-1}$
4. Summe $\sum_{k=1}^{n} a_k$: $\sum_{k=1}^{1} a_k := a_1$, $\sum_{k=1}^{n} a_k := a_n + \sum_{k=1}^{n-1} a_k$.

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.

$$\begin{array}{rcrcl} f & : & A & \longrightarrow & B \\ & & a & \longmapsto & f(a) \end{array}$$

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

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

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

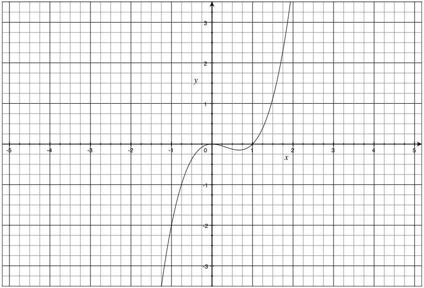
Definition (Eigenschaften von Abbildungen)

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

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

Beispiele.

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



Bezeichnung. Seien $f : A \longrightarrow B$ und $g : B \longrightarrow C$ zwei Abbildungen. Als Verknüpfung, Hintereinanderausführung oder auch Superposition' der Abbildungen f und g bezeichnet man die Abbildung

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

Lemma

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

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

Beweis. ÜA □

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

Mehr zu den Grundlagen (Aussagenlogik, Mengen, etc.)

Skript von E. Bönecke auf

`http://www.math.uni-hamburg.de/home/leistner/mfp1/grund-1.pdf`

und an vielen anderen Stellen im Internet.

Reelle und komplexe Zahlen

- Wiederholung: Natürliche 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 beliebig Additionen und Multiplikationen ausführen. Dabei gelten folgende **Rechenregeln:**

- Kommutativität: $a + b = b + a$ und $a \cdot b = b \cdot a$
- Assoziativität:
  $a + (b + c) = (a + b) + c$ und $a \cdot (b \cdot c) = (a \cdot b) \cdot c$
- Distributivität:
  $a \cdot (b + c) = a \cdot b + a \cdot c$

Bemerkung.

Führt man die natürlichen Zahlen mittels der **Peano Axiome** (G. Peano, 1889) ein, so werden obige Rechenregeln mathematisch beweisbare Aussagen!

Ganze und rationale Zahlen

Ausgehend von den natürlichen Zahlen konstruiert man die ganzen Zahlen

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

und damit dann die rationalen Zahlen

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

und beweist 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. Wegen dem Satz von Phytagoras ist die Länge der Diagonale gleich $\sqrt{2}$.
Zeigen $\sqrt{2} \notin \mathbb{Q}$ mittels Beweis durch Widerspruch:
Wir nehmen an $\sqrt{2} \in \mathbb{Q}$, d.h. $\sqrt{2} = \frac{p}{q}$, $p, q \in \mathbb{N}$ teilerfremd.
Die Gleichung $p^2 = 2q^2$ zeigt dann, dass p gerade ist.
Somit ist p^2 durch 4 teilbar, und damit sind q^2 und q gerade.
Widerspruch zur Teilerfremdheit! $\square$

Körperaxiome

Die reellen Zahlen werden mit dem Symbol $\mathbb{R}$ bezeichnet.
Wir setzten ihre Existenz und die Kenntniss 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,

- Die Addition

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

- und die 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:

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

- Assoziativgesetz:

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

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

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

  (Daraus folgt: $\exists!$ neutrales Element:
  Sei $0'$ ein weiteres, 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: $\forall x \in \mathbb{R} \ \ \exists!$ additives Inverses)

M) Axiome der Multiplikation

- Kommutativgesetz:

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

- Assoziativgesetz:

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

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

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

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

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

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

D) Distributivgesetz

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

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

Definition

Eine Menge $\mathbb{K}$ mit zwei Abbildungen $+$ und $\cdot$, die die die Axiome A), M) und D) erfüllen, heißt Zahlkörper, oder auch nur 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 $0 = 0 \cdot 0 = 0 \cdot 1 = 0 \cdot 1 = 0 + 0 = 1 + 1$ und $1 = 1 + 0 = 0 + 1$ ist ebenfalls ein Körper, der mit $\mathbb{F}_2$ bezeichnet wird.
- Natürliche und ganze Zahlen mit $+$ und $\cdot$ sind **kein** Körper

Folgerungen aus den Körperaxiomen

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.

Beweis. Das folgt aus A). Die Lösung ist $x = b - a := b + (-a)$. $\square$

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 aus M). Die Lösung ist $x = \frac{b}{a} := a^{-1}b$. $\square$

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). Es gilt

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

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 aufassen.

- Die natürlichen Zahlen $\mathbb{N} := \{1, 2, 3, \dots\}$ fassen wir als Teilmenge von $\mathbb{R}$ auf, indem wir $n \in \mathbb{N}$ mit der reellen Zahl $1 + \cdots + 1$ (n Summanden) identifizieren.
- Durch Hinzunahme der additive 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).
- Der kleinste Körper, der $\mathbb{Z}$ enthält ist der Körper der rationalen Zahlen $\mathbb{Q} = \{\frac{p}{q} \mid p \in \mathbb{Z}, \quad q \in \mathbb{N}\}$.

O) Ordnungsaxiome

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

(O1) $\forall x \in \mathbb{R}$ gilt genau eine der folgenden drei Relationen:

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

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

(O3) Archimedisches Axiom: $\forall x > 0 \in \mathbb{R}$ gibt es $n \in \mathbb{N}$ mit $n - x > 0$.

Es gilt:

$\mathbb{R}$ erfüllt die Axiome (O1), (O2) und (O3).

Ebenso wie die reellen Zahlen, setzten wir diese Eigenschaft als gegeben voraus.

$\mathbb{R}$ als archimedisch geordneter Körper

Bezeichnungen

- $x \in \mathbb{R}$ ist negativ $\iff -x > 0$.
- $x > y :\iff x - y > 0$,
- $x \geq y :\iff x > y$ oder $x = y$,
- $x < y :\iff y > x$, sowie $x \leq y :\iff y \geq x$.

Die Relation $x > y :\iff x - y > 0$ definiert eine Ordnungsrelation $>$ auf den reellen Zahlen. $\mathbb{R}$ mit dieser Ordnung ist ein Beipiel für einen archimedisch geordneten Körper.

Folgerungen aus den Ordnungsaxiomen

Satz

Für $a, x, y, z \in \mathbb{R}$ gilt:

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

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

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

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

Beweis. Wir beweisen z.B. (i):

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

□

Satz

(i) $x^2 > 0$ *für alle* $x \neq 0$,

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

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

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

□

Der Absolutbetrag

Definition

Der Absolutbetrag einer reellen Zahl x ist definiert durch

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

Satz

Für alle $x, y \in \mathbb{R}$ gilt:

(i) $|x| \geq 0$ *und* $|x| = 0$, *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 (ii):

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

Abstand

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

Satz

Die Abbildung $d : \mathbb{R} \times \mathbb{R} \to \mathbb{R}_+$ *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 d aus der für $| \, . \, |$. $\square$

Zusammenfassung:

Die reellen Zahlen $\mathbb{R}$ mit zwei Abbildungen $+$ und $\cdot$ erfüllen die Axiome A), M), D) und O) und sind damit ein archimedisch geordneter Körper.

Die rationalen Zahlen $\mathbb{Q}$ sind ein ebenfalls ein archimedischer geordneter Körper. Die Relation ist $\frac{n}{m} > 0 \iff n \cdot m > 0$. (Beweis: Übung.)

Komplexe Zahlen

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

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

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

3. und der Multiplikation

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

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:

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

Weiter im Beweis:

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

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

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

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

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

In der Tat:

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

D) Distributivgesetz: ÜA. $\square$

Konventionen

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

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

  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:

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

- Es ist $i^{-1} = -i$.

Der Körper der komplexen Zahlen wird mit dem Symbol $\mathbb{C}$ bezeichnet:

$$\mathbb{C} := \{z = x + iy \mid x, y \in \mathbb{R}\} = \mathbb{R}^2 .$$

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: $\sqrt{-1} := i.$

Satz (Fundamentalsatz der Algebra)

Jede polynomiale Gleichung mit komplexen Koeffizienten c_k der Form

$$z^n + c_{n-1}z^{n-1} + ... + c_1 z + c_0 = 0$$

hat mindestens eine komplexe Lösung.

Beweis. Der Beweis wird später gegeben. □

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 der 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

$$\overline{z \cdot w} = \bar{z} \cdot \overline{w} \quad \forall\, z, w \in \mathbb{C}.$$

Der Betrag einer komplexen Zahl

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

Definition

Wir definieren den Betrag der komplexen Zahl z durch

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

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

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

genau der Absolutbetrag von x. Ausserdem gilt:

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

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) $|zw| = |z||w|$ *und* *(iv)* $|z + w| \geq ||z| - |w||$.

Beweis. Übung $\square$

Wieder definiert der Betrag einen Abstand von komplexen Zahlen,

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

mit denselben Eigenschaften wie der Abstand von reellen Zahlen:

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

Die Dreiecksungleichung handelt diesmal wirklich von Dreiecken.

Definition

Eine Folge reeller Zahlen ist eine Abbildung $\mathbb{N} \to \mathbb{R}$, $n \mapsto x_n$.
Man schreibt dafür auch $(x_n)_{n\in\mathbb{N}} = (x_1, x_2, \ldots)$ *oder einfach* (x_n).

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

Definition

Eine Folge (x_n) *heißt konvergent mit dem Grenzwert ("Limes")* ℓ, *wenn es für alle* $\varepsilon \in \mathbb{R}$ *mit* $\varepsilon > 0$ *einen Index* $N = N(\varepsilon) \in \mathbb{N}$ *gibt, so dass*

$$|x_n - \ell| < \varepsilon \quad \text{für alle} \quad n \geq N.$$

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

Beachte: $|x_n - \ell|$ kann man auch mit Hilfe des Abstandes schreiben:
$|x_n - \ell| = d(x_n, \ell)$.

Beispiel

Die Folge $x_n = \frac{1}{n}$ konvergiert gegen 0.

Beweis. Wir müssen für jedes reelle $\varepsilon > 0$ einen Index N finden, so daß $|x_n - 0| < \varepsilon \; \forall n \geq N$.
Sei also $\varepsilon > 0$ eine beliebige positive reelle Zahl.
Nach dem Archimedischen Axiom gibt es $N \in \mathbb{N}$ mit $N > \frac{1}{\varepsilon}$. D.h. $\frac{1}{N} < \varepsilon$.
Für alle $n \geq N$ gilt dann $|x_n - 0| = \frac{1}{n} \leq \frac{1}{N} < \varepsilon$. $\square$

Eindeutigkeit des Grenzwertes

Satz

Eine konvergente Folge hat genau einen Grenzwert.

Beweis. (Indirekter Beweis)

- Wir nehmen an, dass (x_n) zwei Grenzwerte ℓ_1, ℓ_2 hat, d.h. $\ell_1 \neq \ell_2$.
- Setze $\varepsilon := \frac{|\ell_2 - \ell_1|}{2} > 0$.
  Wegen der Konvergenz existieren natürliche Zahlen N_1 und N_2 mit

$$\begin{array}{rcccl} |x_n - \ell_1| & < & \varepsilon & \text{für alle} & n \geq N_1, \\ |x_n - \ell_2| & < & \varepsilon & \text{für alle} & n \geq N_2. \end{array}$$

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

$$2\varepsilon = |\ell_2 - \ell_1| = |\ell_2 - x_n + x_n - \ell_1| \leq |\ell_2 - x_n| + |x_n - \ell_1| < 2\varepsilon.$$

- D.h. $2\varepsilon < 2\varepsilon$, ein Widerspruch! Also $\ell_1 = \ell_2$. $\square$

Beschränkte Folgen

Definition

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

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

(ii) *Sie heißt* nach unten beschränkt, *wenn es* $K \in \mathbb{R}$ *gibt derart, dass*

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

(iii) *Eine Folge* (x_n) *heißt* beschränkt, *wenn sie nach oben und nach unten beschränkt ist.*

Einfache ÜA:

Eine Folge (x_n) ist genau dann beschränkt, wenn es $K \in \mathbb{R}$ gibt derart, dass

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

Konvergenz und Beschränktheit

Satz

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

Beweis. Müssen ein $K \in \mathbb{R}$ finden, welches die Folge beschränkt.

- Sei $\ell = \lim\limits_{n\to\infty} x_n$. Dann existiert (zu $\varepsilon = 1$) eine natürliche Zahl $N \in \mathbb{N}$, so dass

$$|x_n - \ell| < 1 \quad \text{für alle} \quad n \geq N.$$

- Daraus folgt für alle $n \geq N$ (wieder mit der Dreicksungleichung):

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

- und somit für alle $n \in \mathbb{N}$:

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

□

Bemerkungen

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

Beweis. Annahme: $\lim_{n\to\infty}(x_n) = \ell$.
Für $\varepsilon := \frac{1}{2}$ existiert eine $N \in \mathbb{N}$: $|x_n - \ell| < \frac{1}{2}\ \forall n \geq N$.
Für *alle* $n \geq N$ und $m \geq N$ hat man

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

Das ist ein Widerspruch zu $(x_n) = (1, -1, 1, -1, 1, \ldots)$. $\square$

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

Satz

(x_n) und (y_n) seien konvergente Folgen. Dann gilt:

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

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

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

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

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

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

Weiter im Beweis:

(ii) (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 $\varepsilon > 0$ existieren $N_1, N_2 \in \mathbb{N}$ mit

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

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

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

Also $\lim\limits_{n \to \infty} (x_n y_n) = xy$. $\square$

Satz

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

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

Beweis. ÜA □

Konvergenz und Ordnungsrelation

Satz

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

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

Beweis. Übung $\square$

Achtung:

Für konvergente Folgen mit $x_n < y_n$ gilt nicht unbedingt $\lim_{n\to\infty} x_n < \lim_{n\to\infty} y_n$.

Beispiel:

Für $x_n := 0 < y_n := \frac{1}{n}$ gilt in der Tat $\lim_{n\to\infty} x_n = 0 \leq \lim_{n\to\infty} y_n = 0$,
aber $0 = \lim_{n\to\infty} x_n \not< \lim_{n\to\infty} y_n = 0$.

Cauchy-Folgen

Definition

Eine Folge (x_n) heißt Cauchy-Folge, wenn es für alle $\varepsilon > 0$ ein $N \in \mathbb{N}$ gibt, so daß

$$|x_n - x_m| < \varepsilon \quad \text{für alle} \quad n, m \geq N.$$

Satz

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

Beweis. Sei $\ell = \lim_{n\to\infty} x_n$. Für alle $n, m \geq N(\frac{\varepsilon}{2})$ gilt:

$$|x_n - x_m| = |x_n - \ell + \ell - x_m| \leq \underbrace{|x_n - \ell|}_{<\frac{\varepsilon}{2}} + \underbrace{|\ell - x_m|}_{<\frac{\varepsilon}{2}} < \varepsilon.$$

$\square$

Das Vollständigkeitsaxiom

Die Umkehrung des letzten Satzes ist das:

Vollständigkeitsaxiom

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

Bemerkungen

(i) Führt man die reellen Zahlen mathematisch exakt ein, so kann man dieses Axiom beweisen (siehe z.B. Königsberger: Analysis 1). Wir setzten voraus, daß es gilt (Axiom).

(ii) 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 $\mathbb{Q}$ nicht vollständig.

Monotone Folgen

Definition

- *Eine Folge (x_n) 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 mit dem Grenzwert ℓ. 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 $\varphi : \mathbb{N} \to \mathbb{N}$ gibt mit $y_n = x_{\varphi(n)}$ für alle $n \in \mathbb{N}$.

Beispiele

- Sei $x_n = \frac{1}{n}$ konvergent. $y_n := x_{n^2} = \frac{1}{n^2}$ ist konvergente Teilfolge.
- Sei $x_n = (-1)^n$ nicht konvergent, aber $y_n := x_{2n} = 1$ und $z_n := x_{2n+1} = -1$ sind konvergente Teilfolgen.

Es gilt:

Ist (x_n) eine konvergente Folge mit dem Grenzwert ℓ. Dann konvergiert jede Teilfolge von (x_n) ebenfalls gegen ℓ.

Beweis. Übung $\square$

Der Satz von Bolzano-Weierstraß

Theorem (Bolzano-Weierstraß)

Jede beschränkte Folge reeller Zahlen besitzt eine konvergente Teilfolge.

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

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

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

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

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

(ii) $[A_k, B_k] \subset [A_{k-1}, B_{k-1}]$, falls $k \geq 2$ und

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

Weiter im Beweis des Satzes von Bolzano-Weierstraß

- Wir setzen $A_1 := A$, $B_1 := B$ und nehmen an, die Intervalle $[A_k, B_k]$ mit den obigen Eigenschaften seien für $k \in \{1, \ldots \ell\}$ bereits konstruiert.
- Wir definieren dann $[A_{\ell+1}, B_{\ell+1}]$ wie folgt:
  - Sei $M := \frac{1}{2}(A_\ell + B_\ell)$ der Mittelpunkt des ℓ-ten Intervalls.
  - Wir setzen $[A_{\ell+1}, B_{\ell+1}] := [A_\ell, M]$, falls $[A_\ell, M]$ unendlich viele Glieder der Folge (x_n) enthält und $[A_{\ell+1}, B_{\ell+1}] := [M, B_\ell]$ sonst.

$\Longrightarrow$ Die Intervalle $[A_k, B_k]$, $k = 1, 2, \ldots, \ell + 1$, die Bedingungen (i)-(iii).

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

- Wir setzen $y_1 := x_1$ und nehmen an, $y_1, \ldots, y_k$ seien 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 Vollständigkeitsaxiom ist (y_k) konvergent, somit folgt aus der Behauptung den Beweis des Satzes von B-W. $\square$

Beweis der Behauptung.

- Zum Beweis der Behauptung benutzen wir folgende Tatsache:

$$\lim_{n\to\infty} \frac{1}{2^n} = 0 \quad (\text{wegen } 0 < \tfrac{1}{2^n} < \tfrac{1}{n} \to_{n\to\infty} 0).$$

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

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

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 (x_n) gegen $\ell := \lim_{k\to\infty} x_{n_k}$ konvergiert.

Weiter im Beweis:

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

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

- Wir setzen $N := n_{k_0}$.
- Für alle $n \geq N = n_{k_0}$ existiert $k \geq k_0$, so dass

$$n_k \leq n < n_{k+1}.$$

- Da (x_n) monoton wachsend ist, folgt daraus

$$x_{n_k} \leq x_n \leq x_{n_{k+1}} \leq \ell$$

- und somit $|x_n - \ell| \leq |x_{n_k} - \ell| < \varepsilon$.

$\square$

Anwendung: Die Quadratwurzel einer positiven Zahl

Theorem

Sei $a > 0$.

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

(ii) *Sei $b > 0$ und (x_n) die durch $x_1 := b$ und*

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

rekursiv definierte 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 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 Lösung geben.

Beweis von (ii): Wir zeigen zuerst, dass die Folge (x_n) gegen eine positive Zahl konvergiert.

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 $n \geq 2$:

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

Weiter im Beweis:

3) Die Folge $(x_n)_{n\geq 2}$ ist monoton fallend:

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

- Die Folge $(x_n)_{n\geq 2}$ ist also monoton fallend und durch 0 nach unten beschränkt.
- Somit konvergiert (x_n) gegen eine Zahl $\ell \geq 0$.

Weiter im Beweis:

4) Behauptung.
$\ell = \lim_{n\to\infty} x_n > 0.$

Beweis der Behauptung.

- Wegen $a \leq x_n^2$ und $x_n \geq 0$ monoton fallend gilt

$$0 < \frac{a^2}{x_2^2} \leq \frac{a^2}{x_n^2} \leq \frac{a^2}{a} = a \leq x_n^2.$$

- Somit $\frac{a}{x_2} \leq x_n$ für alle $n \geq 2$.
- Daraus folgt wie behauptet $\ell = \lim_{n\to\infty} x_n \geq \frac{a}{x_2} > 0$. □

Weiter im Beweis:

- Wir haben gezeigt, dass die Folge (x_n) gegen eine positive Zahl ℓ konvergiert.
- 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 ℓ. Rechenregeln für konvergente Folgen ergeben:

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

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

Unvollständigkeit von $\mathbb{Q}$

Folgerung

$\mathbb{Q}$ *ist nicht vollständig.*

Beweis.

- Wie wir gesehen haben, konvergiert die durch $x_1 := b > 0$ 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 Cauchyfolge.
- Wenn wir b rational wählen, so sind alle Folgenglieder rational.
- Dann ist (x_n) eine rationale Cauchyfolge von rationalen Zahlen, die gegen die irrationale Zahl $\sqrt{2}$ konvergiert. (Wir hatten in der zweiten Vorlesung gesehen daß $\sqrt{2} \notin \mathbb{Q}$).

$\square$

Approximation reeller Zahlen durch rationale Zahlen

Satz

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

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

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

ist x der Grenzwert. □

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

Konstruktion der reellen Zahlen

Eine mathematisch korrekte Beschreibung der reellen Zahlen erhält man dadurch, dass man reelle Zahlen mit Cauchy-Folgen rationaler Zahlen identifiziert. Wobei man zwei solche Cauchy-Folgen als gleich ansieht, falls ihre Differenzfolge gegen 0 konvergiert. Unschwer sieht man, dass die Multiplikation von Folgen mit der der reellen Zahlen verträglich ist. So erhält man per Konstruktion die Vollständigkeit der reellen Zahlen.
Die Details dazu sind jedoch nicht Stoff der Vorlesung.

Vollständigkeit von $\mathbb{C}$

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

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

$|\,.\,|$ erfüllt dieselben Eigenschaften wie der Absolutbetrag. Daher:

Vollständigkeit von $\mathbb{C}$

Die Begriffe 'konvergente Folge', 'Grenzwert' und 'Cauchyfolge' übertragen sich von reellen auf komplexe Zahlenfolgen, indem man den Absolutbetrag reeller Zahlen durch den Betrag komplexer Zahlen ersetzt.

Es gilt dann:

Satz

Der Körper $\mathbb{C}$ ist vollständig, d.h. jede Cauchyfolge komplexer Zahlen konvergiert gegen eine komplexe Zahl.

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

Beweis.

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

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

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

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

□

Bemerkung:

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

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

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

Geometrische Folge

Eine Folge von Potenzen (z^n) einer gegebenen komplexe (oder reelle) 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 $\lim\limits_{n \to \infty} z^n = 0$.*
2. *Ist $|z| > 1$, dann ist (z^n) nicht beschränkt und damit nicht konvergent.*

Schritt 1 des Beweises

Lemma (Bernoullische Ungleichung)

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

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

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

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

□

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 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 $t > 1$. D.h. es gibt ein $x > 0$, so daß $t = 1 + x$.
Sei $K \in \mathbb{R}$, o.B.d.A. $K > 1$.
Wegen (O3) finden wir ein $n \in \mathbb{N}$, so daß $n > \frac{K-1}{x}$.
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 $\varepsilon > 0$ ein N mit $|z^n| = |z|^n < \varepsilon$ für alle $n \geq N$.
Für $|z| > 1$ findet man zu jedem K ein n mit $|z^n| > K$, d.h. (z^n) ist nicht beschränkt. □

Konvergenz von Reihen

Definition

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

$$s_n = \sum_{k=1}^{n} z_k ,$$

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

$$\sum_{k=1}^{\infty} z_k := \lim_{n\to\infty} s_n \quad \text{bezeichnet.}$$

Definition

Eine Reihe

$$\left(\sum_{k=1}^{n} z_k\right)_{n\in\mathbb{N}}$$

heißt absolut konvergent, wenn die Reihe

$$\left(\sum_{k=1}^{n} |z_k|\right)_{n\in\mathbb{N}}$$

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 $-\ln 2$ (werden wir später sehen).

Satz (Cauchy'sches Konvergenzkriterium)

Eine Reihe $(\sum_{k=1}^{n} z_k)_{n\in\mathbb{N}}$ *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 Cauchyfolge ist, denn

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

□

Satz ("Absolute Konvergenz impliziert Konvergenz")

Jede absolut konvergente Reihe ist konvergent.

Beweis. Es genügt zu überprüfen, dass die Reihe $(\sum_{k=1}^{n} z_k)_{n\in\mathbb{N}}$ das Cauchysche Konvergenzkriterium erfüllt.

Dies ist aber erfüllt für die Reihe $(\sum_{k=1}^{n} |z_k|)_{n\in\mathbb{N}}$.

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

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

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

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

Aus dem Cauchy-Kriterium folgt die Konvergenz der Reihe $(\sum_{k=1}^{n} z_k)_{n\in\mathbb{N}}$. $\square$

Satz (Majorantenkriterium)

Sei $(\sum_{k=1}^{n} z_k)_{n\in\mathbb{N}}$ *eine Reihe, und* $a_k \geq 0$ *eine Folge reeller Zahlen mit*

- $|z_k| \leq a_k$ *für alle* $k \in \mathbb{N}$ *und*
- $(\sum_{k=1}^{n} a_k)_{n\in\mathbb{N}}$ *ist eine konvergente Reihe.*

Dann ist die Reihe $(\sum_{k=1}^{n} z_k)_{n\in\mathbb{N}}$ *absolut konvergent und*

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

Beweis. Die absolute Konvergenz folgt aus dem Cauchy'schen Kriterium, der Konvergenz von $(\sum_{k=1}^{n} a_k)$ und der Abschätzung $\sum_{k=m}^{n} |z_k| \leq \sum_{k=m}^{n} a_k$.
Die Ungleichungskette folgt dann aus $|\sum_{k=1}^{n} z_k| \leq \sum_{k=1}^{n} |z_k| \leq \sum_{k=1}^{n} a_k$ durch Grenzübergang. $\square$

Definition (geometrische Reihe)

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

Satz

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

Beweis. Für alle $z \neq 1$ gilt (beweist man mit Induktion über n)

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

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

Wegen (1) konvergiert dann die geometrische Reihe:

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

□

Bemerkung.
Die Konvergenz der geometrischen Reihe ist absolut:

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

Beispiele.

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

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

Satz (Quotientenkriterium)

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

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

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

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

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

Definition (Exponentialreihe)

Sei $z \in \mathbb{C}$. 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

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

Satz

Die Exponentialreihe ist absolut konvergent.

Beweis.

- Finde zu jedem $z \in \mathbb{C}$ 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 $(\sum\limits_{k=N}^{n} \frac{z^k}{k!})_{n \geq N}$ 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*

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

Definition

Man definiert die Eulersche Zahl

$$e := \exp(1).$$

$e \simeq 2,7182818284590452353602874713526624977572470936999595749669$

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

$$
\begin{aligned}
p_n \cdot q_n &= \Big(\sum_{k=0}^{n} a_k\Big) \cdot \Big(\sum_{k=0}^{n} b_k\Big) = \sum_{1 \le i,j \le n} a_i \cdot b_j \\
\text{und} \quad & \Big(\sum_{k=0}^{n} |a_k|\Big) \cdot \Big(\sum_{k=0}^{n} |b_k|\Big) = \sum_{1 \le i,j \le n} |a_i| \cdot |b_j|
\end{aligned}
$$

sind konvergent, aber formal keine Reihe. Daher definiert man:

Definition (Cauchy Produkt)

Das Cauchy-Produkt zweier Reihen $\Big(\sum_{k=0}^{n} a_k\Big)$ *und* $\Big(\sum_{k=0}^{n} b_k\Big)$ *ist die Reihe* $\Big(\sum_{k=0}^{n} c_k\Big)$ *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*

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

Beweis. Die absolute Konvergenz der Reihe $(\sum_{k=0}^{n} c_k)$ folgt aus:

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

Weiter im Beweis

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

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

$$= \left|\sum_{0\le i,j\le n,\ i+j>n} a_i b_j\right| \le \sum_{0\le i,j\le n,\ i+j>n} |a_i||b_j|$$

$$\le \sum_{k=n+1}^{\infty} |c_k| = \underbrace{\sum_{k=0}^{\infty} |c_k| - \sum_{k=0}^{n} |c_k|}_{\lim\limits_{n\to\infty}(\qquad)=0}$$

Eine Folgerung ist die Funktionalgleichung der Exponentialfunktion:

Satz

Für alle $z, w \in \mathbb{C}$ gilt

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

Beweis. Wir benutzen die binomische Formel:

$$(z+w)^k = \sum_{j=0}^{k} \binom{k}{j} z^{k-j} w^j \text{ , wobei } \binom{k}{j} = \frac{k!}{j!(k-j)!}.$$

Daraus ergibt sich:

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

□

Weitere Folgerungen

Satz

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

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

Bemerkungen:

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

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

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

Definition (Sinus und Cosinus)

Für $x \in \mathbb{R}$ definiert man

$$\begin{aligned} \cos x &:= \operatorname{Re} \exp(ix) \quad \textit{und} \\ \sin x &:= \operatorname{Im} \exp(ix), \quad x \in \mathbb{R}. \end{aligned}$$

Satz

Für alle $x \in \mathbb{R}$ gilt $(\cos x)^2 + (\sin x)^2 = 1$, d.h.

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

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

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

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

- (absolut konvergente Potenzreihenentwicklung)

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

- (Symmetrie, resp. Antisymmetrie)

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

- (Additionstheoreme)

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

Ausblick: Ebene Polarkoordinaten

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

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

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

Multiplikation in Polarkoordinaten

- Die Polarkoordinatendarstellung erleichtert die *Multiplikation* komplexer Zahlen:

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

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 geometrischen Interpretation am Kreis und wird später noch bewiesen.

Weitere Konvergenzkriterien für Reihen

Satz

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

Beweis. Nach dem Cauchyschen Konvergenzkriterium gibt es zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$, so daß $|z_n| = |\sum_{k=n}^{n} z_k| < \varepsilon$ für alle $n \geq N$. $\square$

Bemerkung

Das in der Proposition formulierte Konvergenzkriterium ist notwendig aber nicht hinreichend. Beispielsweise ist die sogenannte harmonische Reihe

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

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

Beispiel.
Die alternierende harmonische Reihe

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

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

Satz (Leibnizkriterium für alternierende Reihen)

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

Beachte: der Satz gilt nicht ohne die Voraussetzung 'monoton fallend', wie man an der harmonischen Reihe sieht ($a_k = \frac{(-1)^k}{k}$ für die harmonische Reihe).

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 $y_{n+1} - y_n = -a_{2n+3} + a_{2n+2} \geq 0$.
- Desweiteren gilt $y_0 \leq y_n \leq x_n \leq x_0$, denn $y_n - x_n = -a_{2n+1} \leq 0$.
- Also existieren $x = \lim_{n\to\infty} x_n$ und $y = \lim_{n\to\infty} y_n$.
- Wegen $\lim_{n\to\infty} (y_n - x_n) = \lim_{n\to\infty} (-a_{2n+1}) = 0$, gilt $x = y$.
- Daraus folgt $\lim_{n\to\infty} s_n = \lim_{n\to\infty} s_{2n} = \lim_{n\to\infty} s_{2n+1}$. □

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 setzen:

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

Weiter im Beweis.

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

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

Andererseits gilt für $n > 2^k$:

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

Somit sind die Folgen $(s_n)_{n \in bN}$ und $(t_k)_{n \in bN}$ 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:

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

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

Umordnungssatz

Definition (Umordnung)

Sei $(k_n)_{n\in\mathbb{N}}$ eine Folge in der jede natürliche Zahl genau einmal vorkommt (i.a.W, wir habe eine Bijektion auf $\mathbb{N}$ gegeben durch $n \longrightarrow k_n$). Weiter sei $\sum\limits_{n=1}^{\infty} a_n$ eine Reihe. Die Reihe $\sum\limits_{n=1}^{\infty} a_{k_n}$ wird eine Umordnung von $\sum\limits_{n=1}^{\infty} a_n$ genannt.

Beispiel.

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

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

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 S_1.

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

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

D.h. a_k ist monoton fallend, denn

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

Daher konvergiert auch 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 $-\frac{1}{2j} + \frac{1}{2j+1} < 0$, und

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

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

Wichtige Bemerkung

In Reihen kommt es auf die Reihenfolge der Summationen an.
In anderen Worten:
In unendlichen Summen gilt das Assoziativitätsgesetz in der Regel nicht.

Für absolut konvergente Reihen gilt jedoch:

Satz

Ist $\sum\limits_{n=1}^{\infty} a_n$ eine Reihe komplexer Zahlen, die absolut konvergiert, dann konvergiert jede Umordnung von $\sum\limits_{n=1}^{\infty} a_n$, und alle Umordnungen konvergieren gegen denselben Wert.

Beweis. Da $\sum a_n$ absolut konvergent ist, findet man zu jedem $\epsilon > 0$ existiert aufgrund der Annahme ein $N \in \mathbb{N}$ mit

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

Es sei $\sum a_{k_n}$ eine Umordnung mit Partialsummen s'_n.
Sei nun $p \in \mathbb{N}$ so groß, daß alle Zahlen $1, 2, ..., N$ in der Menge $k_1, k_2, ..., k_p$ enthalten sind. Es folgt für $n \geq p$, daß in der Differenz der Partialsummen $s_n - s'_n$ sich die Zahlen $a_1, ..., a_N$ gegenseitig aufheben, also

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

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

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

$$\sum_{n=1}^{\infty} a_{k_n} = s.$$

Beweis. H. Heuser: Lehrbuch der Analysis 1, S. 199. Die Beweisidee finden Sie auf
`http://de.wikipedia.org/wiki/Riemannscher_Umordnungssatz` □

Abzählbare Mengen

Definition

Eine Menge A heißt abzählbar, wenn es eine surjektive Abbildung $\varphi : \mathbb{N} \longrightarrow A$ gibt.

Beispiele.

- Jede endliche Menge $A = \{a_1, a_2, \ldots, a_N\}$ ist abzählbar. Eine Surjektion $\varphi : \mathbb{N} \to A$ ist $(\varphi(n))_{n\in\mathbb{N}} = (a_1, a_2, \ldots, a_N, a_N, a_N, \ldots)$.
- Die Menge $\mathbb{Z}$ ist abzählbar. Eine Bijektion, und damit eine Surjektion, $\varphi : \mathbb{N} \to \mathbb{Z}$ ist gegeben durch $\varphi(0) = 0$, $\varphi(2k-1) = k$ und $\varphi(2k) = -k$ für $k = 1, 2, \ldots$ D.h. $(\varphi(n))_{n\in\mathbb{N}} = (0, 1, -1, 2, -2, \ldots)$.

Satz

Die Vereinigung von abzählbar vielen abzählbaren Mengen ist abzählbar.

Beweis. Das beweist man mit dem Cantorschen Diagonalverfahren. $\square$

Abzählbarkeit von $\mathbb{Q}$

Folgerung

$\mathbb{Q}$ *ist abzählbar.*

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

Folgerung

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

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

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

□

Überabzählbarkeit von $\mathbb{R}$

Satz

Die reelle Zahlengerade $\mathbb{R}$ ist nicht abzählbar.

Beweis. Wir nehmen an, die reellen Zahlen seien abzählbar und zeigen, daß das ihrer Vollständigkeit widerspricht. Sei also $\mathbb{R} = \{x_0, x_1, x_2, \dots\}$. Dazu konstruieren wir wieder eine Intervallschachtelung von kleiner werdenden Intervallen $I_k \subset I_{k-1}$ der Länge $\frac{1}{3^k}$.

- I_1 sei das Intervall $I_0 := [x_0 + 1, x_0 + 2]$, d.h. $x_0 \notin I_0$.
- Wir unterteilen I_0 in drei gleich große Intervalle und definieren I_1 durch die Bedingung $x_1 \notin I_1$.
- Durch Fortsetzung dieses rekursiven Verfahrens erhalten wir eine Intervallschachtelung durch Intervalle I_n der Länge $\frac{1}{3^n}$ mit $x_n \notin I_n$.

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

Supremumseigenschaft

Definition (Obere und untere Schranken)

Eine Menge $M \subset \mathbb{R}$ heisst nach oben bzw. nach unten beschränkt, wenn es ein $s \in \mathbb{R}$ gibt so, dass für jedes $x \in M$

$$x \leq s \qquad \textit{bzw. } s \leq x$$

gilt. Man nennt s dann eine obere bzw. untere Schranke. Ferner heisst M beschränkt, wenn M sowohl nach unten als nach oben beschränkt ist.

Bemerkung.
Eine beschränkte Menge enthält nicht immer ein kleinstes bzw. grösstes Element. Betrachte z.B. in $\mathbb{R}$ das offene Intervall

$$(0,1) = \{x \in \mathbb{R} \,|\, 0 < x < 1\}.$$

Definition

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. M hat keine obere Schranke s' mit $s' < s$.*

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 (unten) beschränkte, nicht leere Menge $M \subset \mathbb{R}$ besitzt ein Supremum (Infimum).

Die rationalen Zahlen erfüllen die Supremumseigenschaft nicht.

Betrachte $M = [0, \sqrt{2}] \cap \mathbb{Q}$. M is beschränkt hat aber kein Supremum in $\mathbb{Q}$

Beweis der Supremumseigenschaft für $\mathbb{R}$

Wir beweisen hier nur die Existenz des Supremums (die des Infimum beweist man ähnlich). Dazu konstruieren wir 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) $a_n \leq a_{n+1}$,
(iv) $b_{n+1} \leq b_n$ und
(v) $b_n - a_n \leq \frac{1}{2^n}(b_0 - a_0)$.

- Wir beginnen mit $a_0 \in M$ und einer oberen Schranke b_0 von M.
- Ausgehend von $a_0 \leq b_0, \ldots, a_n \leq b_n$ mit (i-v) konstruieren wir $a_{n+1} \leq b_{n+1}$: Sei dazu $m = \frac{1}{2}(a_n + b_n)$.
  - Setzen $[a_{n+1}, b_{n+1}] := [a_n, m]$, falls m eine obere Schranke von M ist.
  - 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 konvergieren.

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

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

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

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

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

Maximum und Minimum

Definition

- *Sei $M \subset \mathbb{R}$ eine nach unten (bzw. oben) beschränkte Menge.*
- *Falls* $\inf M \in M$ *(bzw.* $\sup M \in M$*) so heißt* $\inf M$ *(bzw.* $\sup M$*) das Minimum (bzw. Maximum) der Menge M.*
- *Wir schreiben dann auch* $\min M$ *bzw.* $\max M$ *statt* $\inf M$ *bzw.* $\sup M$.

Notation

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

Stetigkeit

Definition (Folgenkriterium der Stetigkeit)

Sei $D \subset \mathbb{C}$ und $p \in D$. Eine Funktion $f : D \to \mathbb{C}$ heißt stetig in $p \in D$, wenn für jede gegen p konvergierende Folge $(z_n)_{n\in\mathbb{N}}$ aus D gilt:

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

Eine Funktion $f : D \to \mathbb{C}$ heißt stetig, wenn f in allen Punkten $p \in D$ stetig ist.

Oft benutzt man folgende äquivalente Definition der Stetigkeit:

ϵ-δ-Definition der Stetigkeit

Sei $D \subset \mathbb{C}$ und $p \in D$. Eine Funktion $f : D \to \mathbb{C}$ heißt stetig in p, wenn es zu jedem $\epsilon > 0$ ein $\delta > 0$ gibt, derart, daß gilt

$$|f(x) - f(p)| < \epsilon \quad \text{für alle } x \in D \quad \text{mit } |x - p| < \delta.$$

Beispiele

(i) Jede konstante Funktion ist stetig.

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

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

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

(v) Jede rationale Funktion

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

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

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

Satz (Verkettung stetiger Funktionen)

Sei $f : D \to \mathbb{C}$ stetig in $p \in D \subset \mathbb{C}$, $f(D) \subset E \subset \mathbb{C}$ und $g : E \to \mathbb{C}$ stetig in $f(p)$. Dann ist $g \circ f : D \to \mathbb{C}$ stetig in p.

Beweis.

- Sei $(z_n)_{n \in \mathbb{N}} \subset D$ eine Folge mit
$$\lim_{n \to \infty} z_n = p.$$
- Wir erhalten eine Folge $(f(z_n))_{n \in \mathbb{N}} \subset E$. Wegen der Stetigkeit von f in p folgt
$$\lim_{n \to \infty} f(z_n) = f(p).$$
- Aus der Stetigkeit von g in $f(p)$ folgt schließlich
$$\lim_{n \to \infty} g(f(z_n)) = g(f(p)) = (g \circ f)(p),$$
d.h. $g \circ f$ ist stetig.

□

Satz

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

Beweis.

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

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

- Also erfüllt jede Nullfolge z_n ab einem gewissen Folgenglied die Ungleichung $|\exp(z_n) - 1| \leq |z_n| \cdot (e - 1)$, d.h. $\lim_{n\to\infty} \exp(z_n) = 1$.

□

Folgerung

Die Funktionen $\sin, \cos : \mathbb{R} \to \mathbb{R}$ *sind stetig.*

Beweis.

- Aus der Stetigkeit der Exponentialfunktion folgt die Stetigkeit der Funktion $\mathbb{R} \to \mathbb{C}$ geben durch $x \mapsto ix \mapsto \exp(ix)$ (Verkettung).
- Die Stetigkeit von Sinus und Cosinus folgt nun aus der Darstellung

$$\begin{aligned} \cos(x) &= \operatorname{Re}\exp(ix) \\ \sin(x) &= \operatorname{Im}\exp(ix). \end{aligned}$$

$\square$

Definition (Hyperbolische Funktionen)

Sei $x \in \mathbb{R}$. *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 $\sinh : \mathbb{R} \to \mathbb{R}$ *und* $\cosh : \mathbb{R} \to \mathbb{R}$ *heißen Sinus hyperbolicus bzw. Cosinus hyperbolicus.*

Satz

Die Funktionen $\sinh, \cosh : \mathbb{R} \to \mathbb{R}$ *sind stetig.*

Beweis.

Das folgt aus der Stetigkeit der Exponentialfunktion. $\square$

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

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

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

- (Antisymmetrie, bzw. Symmetrie)

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

- (Additionstheoreme)

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

Zwischenwerteigenschaft stetiger Funktionen

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

Satz (Nullstellensatz von Bolzano)

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

Folgerung (Zwischenwertsatz)

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

Beweis. (Zwischenwertsatz)
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 $g(c) = d$. □

Beweis des Nullstellensatzes.

- Durch Ersetzen von f durch $-f$, falls notwendig, können wir ohne Einschränkung annehmen, daß $f(a) < f(b)$.
- 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 $[a_0, b_0], [a_1, b_1], \ldots, [a_n, b_n]$ mit (i)-(iii) bereits konstruiert. Wir konstruieren $[a_{n+1}, b_{n+1}]$.
- Sei $m = \frac{1}{2}(a_n + b_n)$. Wir setzen:

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

- Die Eigenschaften (i-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

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

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

$$f(c) = \lim_{n\to\infty} f(a_n) = \lim_{n\to\infty} f(b_n) = 0.$$

□

Beispiel

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

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

ist stetig,

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

Für jede monoton wachsende rationale Folge (x_n) mit Grenzwert $\sqrt{2}$ gilt $\lim\limits_{n\to\infty} f(x_n) = 0$ und somit müsste $\tilde{f}(\sqrt{2}) = 0$ gelten.

Andererseits gilt für jede monoton fallende rationale Folge (x_n) mit Grenzwert $\sqrt{2}$ hingegen $\lim\limits_{n\to\infty} f(x_n) = 1$ und somit müsste $\tilde{f}(\sqrt{2}) = 1$ gelten.

Infimum und Supremum einer reellwertigen Funktion

Definition

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

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

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

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

Minimum-Maximum-Eigenschaft

Theorem

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

Beweis. Wir zeigen z.B. daß f nach oben beschränkt ist und das Maximum annimmt.

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

- Sei also $f([a, b])$ nach oben unbeschränkt, d.h. es existiert 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 $(x_{n_k})_{k \in \mathbb{N}}$; $\ell := \lim_{k \to \infty} x_{n_k} \in [a, b]$.
- Stetigkeit liefert nun $f(\ell) = \lim_{k \to \infty} f(x_{n_k})$. Das ist unmöglich, denn jede Teilfolge einer monoton wachsenden unbeschränkten Folge ist monoton wachsend und unbeschränkt, und somit nicht konvergent.

Weiter im Beweis:

(2) Sei also 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 $\ell := \lim\limits_{k\to\infty} x_{n_k} \in [a, b]$.
- Stetigkeit liefert $f(\ell) = \lim\limits_{k\to\infty} f(x_{n_k}) = \lim\limits_{n\to\infty} f(x_n) = \sup f$, d.h. f nimmt an der Stelle ℓ ihr Maximum an. □

Folgerung

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

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

Definition

Sei $D \subset \mathbb{R}$. Eine Funktion $f : D \to \mathbb{R}$ 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 π und trigonometrische Funktionen

Wiederholung: Die Cosinusfunktion ist definiert als

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

Theorem

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

Definition

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

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

1. $\cos(2) < 0$. Die Existenz der 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 $|x| \leq 7$:

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

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

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

- Wir haben zu zeigen $0 \leq x < y \leq 2 \Longrightarrow \cos x - \cos y > 0$.
- Man betrachtet $\alpha = \frac{x+y}{2} \in (0, 2)$ und $\beta = \frac{y-x}{2} \in (0, 2]$.
  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:

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

$\square$

Satz (Euler)

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

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

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

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

Folgerung

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

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

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.

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

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 π zu kennen, z.B. auf $[-\frac{\pi}{2}, \frac{\pi}{2}]$.
  - Wegen der Symmetrie $\cos(x) = \cos(-x)$, genügt $[0, \frac{\pi}{2}]$.

$\square$

Folgerung

- *Die Funktionen* sin, cos *und* $\mathbb{R} \ni x \mapsto \exp(ix)$ *sind periodisch mit Periode* 2π*. Es gilt* $\sin, \cos : \mathbb{R} \to [0, 1]$.
- $\cos x = 0 \iff x = \frac{\pi}{2} + n\pi,\ n \in \mathbb{Z}$.
- $\sin x = 0 \iff x = n\pi,\ n \in \mathbb{Z}$.
- $\exp(ix) = 1 \iff x = n2\pi,\ n \in \mathbb{Z}$.
- *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 $f : D \to \mathbb{C}$ eine Funktion. Eine Funktion $g : f(D) \to D$ heißt Umkehrfunktion von f, falls $g \circ f = Id_D$, d.h. $g(f(z)) = z \; \forall z \in D$.

Bemerkungen:

(i) Eine Umkehrfunktion existiert genau dann, wenn $f : D \to \mathbb{C}$ injektiv ist.

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

(iii) Wenn $f : D \to \mathbb{R}$ eine Umkehrfunktion $g : f(D) \to D \subset \mathbb{R}$ besitzt, dann schreiben wir diese als $g = f^{-1}$.
Vorsicht: Verwechseln sie nicht $f^{-1}(x)$, $x \in f(D)$, mit dem multiplikativen Inversen $f(x)^{-1} = \frac{1}{f(x)}$.

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

Umkehrfunktionen streng monotoner Funktionen

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

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

für alle $x, y \in D$ mit $x < y$.

Satz

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

Beweis.

- Sei $y \in f(D)$, d.h. es existiert ein $x \in D$ mit $f(x) = y$.
  Da f streng monoton ist, ist dieses x eindeutig durch y bestimmt:
  Gäbe es ein weiteres x' mit $f(x') = y$, dann wäre $x < x'$ oder $x > x'$ und damit $f(x) < f(x')$ oder $f(x) < f(x')$ 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, d.h.
  $x < x' \implies f(x) < f(x')$.
- 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 $f : [a, b] \to \mathbb{R}$ eine stetige streng monoton wachsende (bzw. streng monoton fallende) Funktion. Dann ist die streng monotone Umkehrfunktion $f^{-1} : [\min f, \max f] \to [a, b]$ stetig.

Beweis.

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

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

- Sei $y_n \in [c, d]$ eine konvergente Folge, $y = \lim_{n\to\infty} y_n$.
- Wir beweisen durch Widerspruch, dass die Folge $x_n := f^{-1}(y_n) \in [a, b]$ gegen $x := f^{-1}(y)$ konvergiert.
- Wenn (x_n) nicht gegen x konvergiert, so gibt es $\varepsilon > 0$, so dass für alle $N \in \mathbb{N}$ eine natürliche Zahl $n \geq N$ existiert mit $|x_n - x| \geq \varepsilon$.
- Daher kann man eine Teilfolge $(x_{n_k})_{k\in\mathbb{N}}$ konstruieren mit $|x_{n_k} - x| \geq \varepsilon$ für alle k.
- Da $x_{n_k} \in [a, b]$, können wir durch Übergang zu einer Teilfolge annehmen, dass (x_{n_k}) gegen $x' \in [a, b] \setminus \{x\}$ konvergiert (Bolzano-Weierstraß).
- Aus der Stetigkeit von f erhalten wir nun

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

- Im Widerspruch zu $\lim_{n\to\infty} y_n = y$. $\square$

Beispiel 1: Potenz- und Wurzelfunktionen

Satz

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

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

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

Beispiel 2: Exponential- und Logarithmusfunktion

Satz

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

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

erfüllt die Gleichung

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

für alle $x, y \in \mathbb{R}_+$.

Bemerkung

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

Beweis. Die Exponentialfunktion ist stetig. Wir beweisen nun

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

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

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

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

Schluß des Beweises

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

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

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

- Anwendung des Logarithmus ergibt:

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

$\square$

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

$$\begin{array}{rccl} \exp_a : & \mathbb{R} & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & \exp(x \ln(a)). \end{array}$$

heißt Exponentialfunktion zur Basis a.

Es gilt:

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

Satz

Die Funktion $\exp_a : \mathbb{R} \to \mathbb{R}$ *ist stetig und erfüllt:*

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

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

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

$\square$

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

□

Satz

(i) $\exp_1(x) \equiv 1$.

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

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

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

Beweis.

(i-iii) Es ist $\ln 1 = \ln \exp(0) = 0$ und damit $\exp_1(x) = \exp(x \cdot 0) = 1$.
Außerdem ist $\ln : \mathbb{R}_+ \to \mathbb{R}$ ist 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 $\exp_a(\mathbb{R}) = \exp(\mathbb{R}) = \mathbb{R}_+$. $\square$

Definition

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

Satz

Es gilt

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

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

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

Charakterisierung der Exponentialfunktionen

Theorem

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

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

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

Beweis.

- Wegen der Funktionalgleichung $(*)$ ist f nicht nur in 0, sondern überall stetig:
  $f(x_n) = f(x_n - x + x) \overset{(*)}{=} f(x_n - x)f(x) \xrightarrow[x_n \to 0]{\text{stetig in } 0} f(0)f(x) \overset{(*)}{=} f(x)$.
- $a = f(1) \overset{(*)}{=} f(\frac{1}{2})^2 \geq 0$.
- Es gilt $f(x) \overset{(*)}{=} f(x-1)f(1) = f(x-1)a$ für alle $x \in \mathbb{R}$. D.h. f konstant null oder $a > 0$

Weiter im Beweis:

Es bleibt zu zeigen, dass $f(x) = a^x$ für alle $x \in \mathbb{R}$. 2 Schritte:

1. $f(n) = a^n$ für alle $n \in \mathbb{Z}$:
   - $a = f(1) = f(1+0) \overset{(*)}{=} af(0)$, d.h. $f(0) = 1 = a^0$.
   - Für $n \in \mathbb{N}$ haben wir $f(n) \overset{(*)}{=} f(1) \cdots f(1) = a^n$ und
   - Aus $1 = f(n-n) \overset{(*)}{=} a^n f(-n)$ folgt $f(-n) = a^{-n}$.
2. $f(\frac{p}{q}) = \sqrt[q]{a^p} = a^{\frac{p}{q}}$ für alle $\frac{p}{q} \in \mathbb{Q}$, d.h. $p \in \mathbb{Z}$ und $q \in \mathbb{N}$:
   - $f\left(\frac{p}{q}\right)^q = \underbrace{f\left(\frac{p}{q}\right) \cdots f\left(\frac{p}{q}\right)}_{\text{q Faktoren}} \overset{(*)}{=} f(p) = a^p > 0.$
   - Diese Rechnung zeigt, dass $f(\frac{p}{q}) = \sqrt[q]{a^p} = a^{\frac{p}{q}}$.

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

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

$\square$

Uneigentliche Grenzwerte

Definition (Grenzwert einer Funktion)

Sei $f : D \to \mathbb{C}$ eine Funktion. Wir schreiben $\lim_{z \to p} f(z) = q$, falls es

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

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

*Man nennt q den **Grenzwert** von f in p.*

Definition

- *Sei (x_n) eine Folge reeller Zahlen. Man schreibt $\lim_{n \to \infty} x_n = +\infty$, falls es für jedes $K > 0$ ein $N \in \mathbb{N}$ gibt, so daß $x_n \geq K$ für alle $n > N$.*
- *Man schreibt $\lim_{n \to \infty} x_n = -\infty$, falls $\lim_{n \to \infty} (-x_n) = +\infty$.*
- *In beiden Fällen nennt man x_n **bestimmt divergent**.*

Beispiele:

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

Exponentielles Wachstum

Satz

Für alle $k \in \mathbb{N}$ gilt

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

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

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

$\square$

Bemerkung

Die Exponentialfunktion wächst also schneller als jede Potenzfunktion.

Logarithmisches Wachstum

Satz

Für alle $k \in \mathbb{N}$ gilt

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

Beweis.

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

□

Bemerkung

Der Logarithmus wächst also langsamer als jede Wurzelfunktion.

Beispiel 4: Trigonometrische Funktionen und Arcusfunktionen

Wiederholung (Trigonometrische Funktionen).

- Die Funktionen sin, cos und $\mathbb{R} \ni x \mapsto \exp(ix)$ sind periodisch mit Periode 2π.
- $\cos x = 0 \iff x = \frac{\pi}{2} + n\pi,\ n \in \mathbb{Z}$.
- $\sin x = 0 \iff x = n\pi,\ n \in \mathbb{Z}$.
- $\exp(ix) = 1 \iff x = n2\pi,\ n \in \mathbb{Z}$.
- 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 $\tan\left((-\frac{\pi}{2}, \frac{\pi}{2})\right) = \mathbb{R}$.

Arcusfunktionen

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

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

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

Definition

Die (streng montonen und stetigen) Umkehrfunktionen der trigonometrischen Funktion

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

heißen Arcus-Cosinus, Arcus-Sinus und Arcus-Tangens.

Polarkoordinaten

Satz

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

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

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

Definition (Polarkoordinaten)

Das Paar (r, φ) heisst Polarkoordinaten von $z = r\exp(i\varphi)$ und φ wird das Argument von z genannt.

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

Weiter im Beweis.

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

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

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

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

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

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

Also erhalten wir eine Darstellung

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

Wegen der Periodizität der Exponentialfunktion folgt die Eindeutigkeit dieser Darstellung bis auf ganze Vielfache von 2π. □

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

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

Ebenso leicht kann man Wurzeln aus einer komplexen Zahl ziehen:
Ist nämlich $z = r\exp(i\varphi)$, dann ist

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

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

Satz

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

$$\zeta^k = \exp\left(k\frac{2\pi i}{n}\right) = \cos(k2\pi/n) + i\sin(k2\pi/n),$$

mit $k = 0, 1, \ldots, n-1$, diese werden die n-ten Einheitswurzeln genannt.

Beweis. Offensichtlich erfüllen diese n verschiedenen Zahlen die Gleichung $z^n = 1$. Es gibt keine weiteren Lösungen, denn es gilt:

Lemma

Sei $p(z) = a_n z^n + \ldots a_a z + a_0$ ein Polynom. Dann hat p höchstens n verschiedene Nullstellen.

Beweis. Dies gilt, da jedes Polynom durch $(z - \lambda)$ teilbar ist, falls λ eine Nullstelle des Polynoms ist. □

Folgerung

Für $0 \neq c \in \mathbb{C}$ hat die Gleichung $z^n = c$ mit $n \in \mathbb{N}$ genau die n Lösungen

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

dabei ist $k = 0, \ldots, n-1$ 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

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

hat mindestens eine komplexe Lösung.

Beweis.
Wir schreiben $P(z) = z^n + c_{n-1}z^{n-1} + \ldots + c_1 z + c_0$ und setzen

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

Für $|z| = R$ gilt dann

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

Für große R strebt die rechte Seite von (4) gegen ∞.
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 $|P(z)|$ 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 $Q(0) = 1$, d.h. es gibt ein $1 \leq k \leq n$ so daß

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

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 daß $r^k|b_k| < 1$. Dann gilt

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

$$= 1 - r^k(|b_k| - r|b_{k+1}| - \dots - r^{n-k}|b_n|)$$

Für hinreichend kleines r_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 den Widerspruch zu $|Q(z)| \geq 1$ und somit zur Annahme $\mu \neq 0$. D.h. $P(z_0) = \mu = 0$. □

Differentialrechnung

Definition (Differenzenquotient)

Sei $D \subset \mathbb{R}$, $f : D \to \mathbb{R}$ eine Funktion und $x \in D$.
Die auf $D \setminus \{x\}$ definierte Funktion

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

heißt Differenzenquotient von f an der Stelle x.

Definition (Häufungspunkt)

Ein Punkt $x \in \mathbb{R}$ heißt Häufungspunkt von D, wenn es eine Folge $x_n \in D \setminus \{x\} = \{\xi \in D | \xi \neq x\}$ gibt, die gegen x konvergiert.

- Jedes Element in einem (offenen oder abgeschlossenen) Intervall ist ein Häufungspunkt.
- Schreiben wir im Folgenden "$\lim\limits_{\xi \to x}$" meinen wir "$\lim\limits_{\xi \to x,\ x \neq \xi}$".

Definition (Ableitung)

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

- *Man sagt, dass f in x differenzierbar ist,*

  wenn der Grenzwert

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

  existiert.
- *Der Grenzwert $f'(x)$ heißt Ableitung der Funktion f an der Stelle x.*
- *Die Funktion f heißt differenzierbar, wenn sie in allen Punkten $x \in D$ differenzierbar ist.*
- *Die Funktion $f' : x \mapsto f'(x)$ heißt 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', wenn f eine Funktion der Zeit ist. Die Zeitvariable heißt dann in der Regel t

Die heute benutzte Definition der Ableitung wurde nach Vorarbeiten von Cauchy schliesslich von Weierstraß Ende des 19. Jahrhunderts formuliert. [2]

[2] http://de.wikipedia.org/wiki/Differentialrechnung

Geometrische Interpretation

- 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").
- Wenn f in x differenzierbar ist, dann strebt die Sekante für $\xi \longrightarrow x$ gegen eine Grenzgerade, die sogenannte Tangente:
- 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 $f : I \to \mathbb{R}$ eine Funktion.
- Wir können $t \mapsto f(t)$ als die Bewegung eines Punktes im eindimensionalen Raum $\mathbb{R}$ auffassen.
- Der Differenzenquotient $\frac{f(t_0)-f(t)}{t_0-t}$ ist dann eine Durchschnittsgeschwindigkeit 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 $x_i : \mathbb{R} \to \mathbb{R}$, $i = 1, 2, 3$, beschrieben.
- Die Geschwindigkeit $v(t)$ zum Zeitpunkt t hat dementsprechend drei Komponenten:

$$v(t) := \dot{x}(t) := (\dot{x}_1(t), \dot{x}_2(t), \dot{x}_3(t)).$$

- Nochmaliges Ableiten liefert die Beschleunigung

$$a(t) := \dot{v}(t).$$

Beispiele

(i) Für jede konstante Funktion $f : \mathbb{R} \to \mathbb{R}$, $f(x) = c$, gilt

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

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

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

Dasselbe gilt auch für jede affine Funktion $f(x) := ax + b$ mit $a, b \in \mathbb{R}$, denn hier ist

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

(iii) Für $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^2$, gilt

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

Weitere Beispiele

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

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

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

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

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

$$\left| \frac{\exp h - 1}{h} - 1 \right| = \left| \frac{\frac{h^2}{2!} + \frac{h^3}{3!} + \ldots}{h} \right| \leq \frac{|h|}{2!} + \frac{|h|^2}{3!} + \ldots$$

$$\leq |h| + \frac{|h|^2}{2!} + \ldots = \exp(|h|) - 1 \xrightarrow[h \to 0]{} 0.$$

Noch mehr Beispiele

(vi) Es gilt $\sin' = \cos$, denn

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

wobei $\lim_{h\to 0}\frac{\cos h-1}{h}=0$ und $\lim_{h\to 0}\frac{\sin h}{h}=1$.
Das folgt aus:

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

(vii) $\cos' = -\sin$ (ÜA).

Beispiel (stetig aber nicht differenzierbar)

Die stetige Funktion $f(x) := |x|$, $x \in \mathbb{R}$, ist im Nullpunkt nicht differenzierbar.

Beweis.
Wir betrachten die beiden Nullfolgen $x_n^+ := \frac{1}{n}$ und $x_n^- := -\frac{1}{n}$.
Einerseits gilt

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

und andererseits

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

Also ist f in 0 nicht differenzierbar. $\square$

Satz (Differenzierbare Funktionen sind stetig)

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

Beweis. Es gilt

$$|f(x) - f(a)| = \left|\frac{f(x) - f(a)}{x - a}\right| |x - a| \xrightarrow[x \to a]{} |f'(a)| \lim_{x \to a} \cdot |x - a| = 0$$

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

Satz (Affine Approximation)

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

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

Beweis. $\frac{f(x)-h(x)}{x-a} = \frac{f(x)-f(a)}{x-a} - f'(a) \xrightarrow[x \to a]{} 0.$ □

Beispiel:

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

- Wir nehmen an, f ist differenzierbar in 0 mit $f'(0)$ als Ableitung.
- Dann 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 $x \to 0$, was im Widerspruch zum Satz steht.

Umgekehrt gilt:

Satz

Sei $f : D \to \mathbb{R}$ eine Funktion, $a \in D$ ein Häufungspunkt von D und $h(x) = c(x-a) + f(a)$ eine affine Funktion mit $c \in \mathbb{R}$, so daß

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

Dann ist f in a differenzierbar, und $h(x) = f'(a)(x-a) + f(a)$.

Beweis. Es gilt:

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

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

Definition (Affine Approximation/ lineare Approximation)

Sei $f : D \to \mathbb{R}$ 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.*
- *$R(a, x) := f(x) - h(x) = f(x) - f'(a)(x - a) - f(a)$ heißt Restglied. Für $x \to a$ geht das Restglied schneller gegen 0 als $x - a$.*

Satz (Ableitungsregeln)

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

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

$$(\lambda \cdot f + \mu \cdot g)'(a) = \lambda \cdot f'(a) + \mu \cdot g'(a). \quad \textit{(Linearität)}$$

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

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

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

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

Beweis.

(i)

$$\frac{(\lambda f + \mu g)(a+h) - (\lambda f + \mu g)(a)}{h} = \lambda \frac{f(a+h) - f(a)}{h} + \mu \frac{g(a+h) - g(a)}{h} \xrightarrow[h \to 0]{} \lambda \cdot f'(a) + \mu \cdot g'(a)$$

(ii)

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

Weiter im Beweis.

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

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

d.h. $(\frac{1}{g})'(a) = -\frac{g'(a)}{g(a)^2}$. Mit (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

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

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

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

(iii) Die Quotientenregel liefert

$$\tan' = \frac{\sin' \cos - \sin \cos'}{\cos^2} = \frac{\cos^2 + \sin^2}{\cos^2} = \frac{1}{\cos^2} = 1 + \tan^2 .$$

Satz (Ableitung der Umkehrfunktion)

Sei $f : [a, b] \to \mathbb{R}$ eine stetige streng monotone Funktion und $g = f^{-1} : [c, d] \to \mathbb{R}$ sei die zugehörige Umkehrfunktion.
Wenn f in $x \in [a, b]$ differenzierbar ist und $f'(x) \neq 0$, dann ist g in $y = f(x)$ differenzierbar und

$$g'(y) = \frac{1}{f'(x)} = \frac{1}{f'(g(y))}.$$

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

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

d.h. $g'(y) = 1/f'(x)$. $\square$

Bemerkung

Wenn im obigen Satz f überall differenzierbar ist und $f'(x) \neq 0$ für alle $x \in [a, b]$, so erhalten wir für die Ableitung der Umkehrfunktion $g = f^{-1} : [c, d] \to \mathbb{R}$ die Regel

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

Beispiele

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

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

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

denn $\cos = \sqrt{1-\sin^2}$ auf $[-\frac{\pi}{2},\frac{\pi}{2}]$.

(iii) Ebenso hat $\arccos = (\cos|_{[0,\pi]})^{-1} : [-1,1] \to \mathbb{R}$ für $x \in (-1,1)$ die Ableitung

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

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

Satz (Kettenregel)

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

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

Beweis. Sei x fixiert und $y = f(x)$. Wir betrachten die folgende Funktion

$$\begin{array}{rcl} h : E & \to & \mathbb{R} \\ \eta & \mapsto & \begin{cases} \frac{g(\eta)-g(y)}{\eta-y}, & \text{falls} \quad \eta \in E \setminus \{y\} \\ g'(y), & \text{falls} \quad \eta = y. \end{cases} \end{array}$$

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

$$g(\eta) - g(y) = h(\eta)(\eta - y). \tag{5}$$

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

$$\begin{array}{rcl} \frac{(g \circ f)(\xi) - (g \circ f)(x)}{\xi - x} & = & h(f(\xi))\frac{f(\xi) - f(x)}{\xi - x} \\ & \xrightarrow[\xi \to x]{} & h(f(x))f'(x) \ = \ g'(f(x))f'(x) \end{array}$$

□

Beispiele

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

Dann ist g differenzierbar und

$$g'(x) = af'(ax + b).$$

(ii) Sei $f : \mathbb{R}_+ \to \mathbb{R}$, $f(x) = x^\alpha$, wobei $\alpha \in \mathbb{R}$.

Dann ist

$$f'(x) = \alpha x^{\alpha-1},$$

denn aus $f(x) = \exp(\alpha \ln x)$ folgt mit der Kettenregel

$$f'(x) = \exp'(\alpha \ln x)(\alpha \ln x)' = \underbrace{\exp(\alpha \ln x)}_{x^\alpha}\frac{\alpha}{x} = \alpha x^{\alpha-1}.$$

Definition (Lokale Extrema)

Sei $f : D \to \mathbb{R}$ 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 $\varepsilon > 0$ gibt, derart dass*

  $$f(z) \geq f(\zeta) \quad (bzw. \quad f(z) \leq f(\zeta))$$

  für alle $\zeta \in D$ mit $|z - \zeta| < \varepsilon$.
- *Im Gegensatz dazu bezeichnet man $\max f$ und $\min f$ als globales Minimum bz.w. Maximum.*
- *Statt von (lokalen bzw. globalen) Minima und Maxima spricht man auch von (lokalen bzw. globalen) Extrema.*
- *Man spricht von einem isolierten lokalen Extremum, falls zusätzlich $f(\zeta) \neq f(z)$ für alle $\zeta \in D \setminus \{z\}$ mit $|z - \zeta| < \varepsilon$.*

Satz

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

Beweis.

Wir nehmen z.B. an, dass ein lokales Maximum vorliegt.

Dann gibt es $\varepsilon > 0$, so dass $f(x) - f(\xi) \geq 0$ für alle ξ mit $|x - \xi| < \varepsilon$.

Schreibt man "$\lim\limits_{\xi \nearrow x}$" für "$\lim\limits_{\xi \to x, \xi < x}$" und "$\lim\limits_{\xi \searrow x}$" für "$\lim\limits_{\xi \to x, \xi > x}$", dann ist

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

und ebenso

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

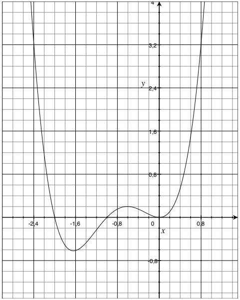
d.h. $f'(x) = 0$. $\square$

Beispiele

(i) Achtung: $f'(x) = 0$ ist nur notwendig, aber nicht hinreichend für die Existenz eines lokalen Extremums.
Die Funktion $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^3$ erfüllt $f'(0) = 0$, hat aber an der Stelle 0 kein lokales Extremum.

(ii) Die Funktion $f : \mathbb{R} \to \mathbb{R}$, $f(x) = (x+2)(x+1)x^2$, ist nach oben unbeschränkt ($\sup f = +\infty$) und hat daher kein (globales) Maximum.
Sie nimmt ihr (globales) Minimum $\min f$ an einer Stelle $a \in (-2, -1)$ an.
Sie hat zwei weitere lokale Extrema: ein lokales Maximum an einer Stelle $b \in (-1, 0)$ und ein lokales Minimum bei 0.
ÜA: Berechnen Sie a, b und $\min f = f(a)$.

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



(iii) Die Funktion $f : [-1, 2] \to \mathbb{R}$, $f(x) = x^2$, nimmt an der Stelle 0 ihr Minimum $\min f = 0$, an der Stelle 2 ihr Maximum $\max f = 4$ und an der Stelle -1 ein lokales Maximum an.

Satz (Satz von Rolle)

Sei $f : [a, b] \longrightarrow \mathbb{R}$ *eine stetige Funktion mit* $f(a) = f(b)$.

Falls f *auf* (a, b) *differenzierbar ist, so existiert* $\xi \in (a, b)$ *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 $\xi \in (a, b)$ mit $f(\xi) = \max f$, im zweiten Fall $\xi \in (a, b)$ mit $f(\xi) = \min f$.

In beiden Fällen ist $f'(\xi) = 0$. $\square$

Satz (Mittelwertsatz)

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

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

Beweis.

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

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

g erfüllt die Voraussetzungen des Satzes von Rolle, insbesondere $g(a) = g(b) = f(a)$.

Somit existiert $\xi \in (a, b)$ mit $0 = g'(\xi) = f'(\xi) - \frac{f(b)-f(a)}{b-a}$. □

Folgerung (Schrankensatz)

Unter den Voraussetzungen des Mittelwertsatz (MWS) gelte zusätzlich

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

Dann gilt

$$m(y - x) \leq f(y) - f(x) \leq M(y - x) \tag{7}$$

für alle $x, y \in [a, b]$ *mit* $x \leq y$.

Beweis.

Für $x = y$ ist nichts zu zeigen. Für $x < y$ gibt es nach dem MWS ein $\xi \in (x, y)$ mit $f'(\xi) = \frac{f(y)-f(x)}{y-x}$.

Multiplikation der Ungleichung (6) mit $y - x > 0$ ergibt dann (7). $\square$

Folgerung

Sei $f : [a,b] \to \mathbb{R}$ stetig, auf (a,b) differenzierbar und $f'(x) = 0$ für alle $x \in (a,b)$. Dann ist f konstant.

Beweis.

- Die Funktion erfüllt die Voraussetzungen des Schrankensatzes mit $m = M = 0$.
- Also $f(y) = f(x)$ für alle $x, y \in [a,b]$ mit $x \leq y$, d.h. $f = const$. □

Bemerkung

- Dieser Satz ist sehr hilfreich beim Studium der Eindeutigkeit von Lösungen von Differentialgleichungen (Dgl.) bei gegebenen Anfangsbedingungen.
- Er besagt, dass die Differentialgleichung $f' = 0$ genau eine Lösung mit der Anfangsbedingung $f(x_0) = c$ hat, nämlich die konstante Funktion $f \equiv c$. (Hierbei ist $x_0 \in [a,b]$ und $c \in \mathbb{R}$.)

Satz (Charakterisierung von exp durch eine Differentialgleichung)

Sei $c \in \mathbb{R}$ und $f : \mathbb{R} \to \mathbb{R}$ eine Lösung der Differentialgleichung

$$f' = c \cdot f. \tag{8}$$

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

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

Ableiten liefert:

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

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

Grenzwertbestimmung nach L'Hospital/Bernoulli

Satz (Verallgemeinerter Mittelwertsatz)

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

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

Beweis. Übungsblatt 8 □

Folgerung (Regel von L'Hospital)

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

Falls $\lim\limits_{x \searrow a} \frac{f'(x)}{g'(x)}$ existiert, so gilt $\lim\limits_{x \searrow a} \dfrac{f(x)}{g(x)} = \lim\limits_{x \searrow a} \dfrac{f'(x)}{g'(x)}$.

Dieselben Aussage gilt für $\lim\limits_{x \nearrow b}$ und damit für beidseitige Grenzwerte.

Beweis der L'Hospital'schen Regel

Da nach Voraussetzung der Grenzwert $\lim\limits_{x \searrow a} \frac{f'(x)}{g'(x)}$ existiert, gilt $g'(x) \neq 0$ für alle $x \in (a, b)$. Sei nun $x \in (a, b)$. Man definiert dann

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

und $\tilde{g}$ genauso. $\tilde{f}$ und $\tilde{g}$ erfüllen dann die Voraussetzungen des verallgemeinerten Mittelwertsatzes, und man findet ein $t \in (a, x)$ 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 Folge mit $x_n \searrow a$, so gilt auch für die Folge der Zwischenwerte $t_n \searrow a$ und damit

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

$\square$

Beispiele

(i) $\lim\limits_{x\to 0} \dfrac{\sin x}{x} = \lim\limits_{x\to 0} \dfrac{\cos x}{1} = 1.$ (ii) $\lim\limits_{x\to 0} \dfrac{1-\cos x}{x^2} = \lim\limits_{x\to 0} \dfrac{\sin x}{2x} = \dfrac{1}{2}.$

Bemerkung

Analog erhält man, daß $\lim\limits_{x\to\infty} \dfrac{f(x)}{g(x)} = \lim\limits_{x\to\infty} \dfrac{f'(x)}{g'(x)}$, falls

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

- Beispiel: $\lim\limits_{x\to\infty} \dfrac{\ln x}{x^\alpha} = \lim\limits_{x\to\infty} \dfrac{1/x}{\alpha x^{\alpha-1}} = \lim\limits_{x\to\infty} \dfrac{1}{\alpha x^\alpha} = 0$ für $\alpha > 0$.
- Achtung: $\lim \frac{f(x)}{g(x)}$ kann existieren, auch wenn $\lim \frac{f'(x)}{g'(x)}$ nicht existiert.
  Bsp.: Für $f(x) = \sin x + 2x$, $g(x) = \cos x + 2x$ ist $\lim\limits_{x\to\infty} \frac{f'(x)}{g'(x)}$ unbestimmt divergent, aber $\lim\limits_{x\to\infty} \frac{f(x)}{g(x)} = \lim\limits_{x\to\infty} (1 + \frac{\sin x - \cos x}{g(x)}) = 1$.

Satz (Monotonie und Ableitung)

Sei $f : [a, b] \to \mathbb{R}$ stetig und auf (a, b) differenzierbar.

(i) *Falls $f'(x) \geq 0$ (bzw. $f'(x) > 0$) für alle $x \in (a, b)$, so ist f auf $[a, b]$ monoton wachsend (bzw. streng monoton wachsend).*

(ii) *Falls $f'(x) \leq 0$ (bzw. $f'(x) < 0$) für alle $x \in (a, b)$, so ist f auf $[a, b]$ monoton fallend (bzw. streng monoton fallend).*

(iii) *Umgekehrt gilt f monoton wachsend (bzw. monoton fallend) impliziert $f'(x) \geq 0$ (bzw. $f'(x) \leq 0$) für alle $x \in (a, b)$.*

Bemerkung

- Aus dem streng monotonen Wachstum von f folgt nur $f' \geq 0$ und nicht die strikte Ungleichung.
- Beispielsweise ist die Funktion $f(x) = x^3$ streng monoton wachsend, aber $f'(0) = 0$.

Beweis des Satzes:

(i-ii) Wir betrachten 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 $f(x) \geq f(y)$.

Wg. des MWS 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 $x \in (a, b)$:

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

$\square$

Folgerung (lokale Extrema)

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

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

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

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

für alle $\xi \in (x, x + \varepsilon)$*.*

Dann hat f *ein lokales Minimum (bzw. Maximum) an der Stelle* x*.*

Ersetzt man die Ungleichungen (9) *und* (10) *durch strikte Ungleichungen,* $f'(\xi) < 0$ *(bzw.* > 0*) usw., so folgt, dass das lokale Extremum isoliert ist.*

Beweis. Aus (9) und (10) folgt, dass f auf $[x - \varepsilon, x]$ monoton fallend (bzw. wachsend) und auf $[x, x + \varepsilon]$ monoton wachsend (bzw. fallend) ist.

Also hat f an der Stelle x ein lokales Minimum (bzw. Maximum). □

Beispiele:

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

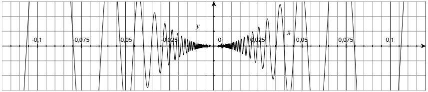
ist differenzierbar, aber $f'(x)$ ist keine stetige Funktion!
Man sagt, f ist nicht stetig-differenzierbar.
Dazu berechnen wir

$$f'(0) = \lim_{x \to 0} \frac{f(x) - f(0)}{x - 0} = \lim_{x \to 0} x \sin(1/x) = 0,$$

und für $x \neq 0$ haben wir

$$f'(x) = 2x \sin(1/x) - 2 \cos(1/x).$$

Weil der Grenzwert $\lim_{x \to 0} f'(x)$ nicht existiert, ist $f'(x)$ nicht stetig!

$graph(f) =$ 

(ii) Es gibt stetige aber nirgendwo differenzierbare Funktionen, z.B. die *Weierstraß-Funktion* $f(x) = \sum_{n=1}^{\infty} \frac{2^n \sin(2^n x)}{3^n}$.

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 $f : (a, b) \to \mathbb{R}$ zweimal differenzierbar und $x \in (a, b)$, derart dass $f'(x) = 0$ und $f''(x) > 0$ (bzw. $f''(x) < 0$).

Dann hat f ein isoliertes lokales Minimum (bzw. Maximum) an der Stelle x.

Beweis.

Wir betrachten z.B. den Fall $f''(x) > 0$.

Wegen

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

gibt es $\varepsilon > 0$, so daß $\frac{f'(\xi)-f'(x)}{\xi-x} > 0$ für alle ξ mit $|x - \xi| < \varepsilon$.

Also gilt daß $f'(\xi) < f'(x) = 0$ für alle $\xi \in (x - \varepsilon, x)$ und $f'(\xi) > f'(x) = 0$ für alle $\xi \in (x, x + \varepsilon)$.

Wir folgern, daß f an der Stelle x ein isoliertes lokales Mininimum hat. $\square$

Beispiele

(i) Für $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^2$, gilt $f'(0) = 0$ und $f''(0) = 2 > 0$. Also hat f an der Stelle 0 ein isoliertes lokales Minimum. Das ist auch das globale Minimum der Funktion.

(ii) Für $f(x) = x^3$ ist auch $f''(0) = 0$ und x ist auch kein lokales Extremum.

(iii) Die Bedingung $f''(x) \neq 0$ ist hinreichend für die Existenz eines lokalen Extremums aber nicht notwendig :
Die Funktion $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^4$, hat ebenfalls an der Stelle 0 ihr (isoliertes) globales Minimum. In diesem Fall gilt jedoch $f''(0) = 0$.

Konvexe Funktionen

- Sei $f : I := [a, b] \to \mathbb{R}$ eine Funktion, $a \leq x_1 < x_2 \leq b$, und

$$s(x) := f(x_1)\frac{x_2 - x}{x_2 - x_1} + f(x_2)\frac{x - x_1}{x_2 - x_1} = \frac{f(x_2) - f(x_1)}{x_2 - x_1}x + \frac{f(x_1)x_2 - f(x_2)x_1}{x_2 - x_1}$$

  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 $f(x) \leq s(x)$.
- Diese Bedingung ist äquivalent zu

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

  (Setze $t := \frac{x - x_1}{x_2 - x_1}$, dann $1 - t = \frac{x_2 - x}{x_2 - x_1}$ und $(1-t)x_1 + tx_2 = x$.)
- Für eine zweimal differenzierbare Funktion $f : I \to \mathbb{R}$ gilt dann:

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

  (Beweis: Forster S. 169)

Theorem (Taylor-Entwicklung)

Es sei $f : [a, b] \to \mathbb{R}$ eine k-mal differenzierbare Funktion.
Zu $x_0 \in [a, b]$ definieren wir das Polynom

$$\begin{aligned} T_{k-1}(x) &= f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 \\ &\quad + \ldots + \frac{f^{(k-1)}(x_0)}{(k-1)!}(x - x_0)^{k-1}. \end{aligned}$$

Dann gibt es zu jedem $x \in [a, b]$ ein ξ zwischen x_0 und x, so daß

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

Bemerkung: Für $k = 1$ entspricht der Satz gerade dem Mittelwertsatz.

Beweis. Es sei $x \in [a, b]$ und M bestimmt durch

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

Definieren nun eine Funktion $g : [a, b] \to \mathbb{R}$ mittels

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

Damit gilt dann für $t \in [a, b]$, daß

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

Wir müssen also die Existenz eines ξ zwischen x_0 und x mit $g^{(k)}(\xi) = 0$ nachweisen, denn dann gilt $M = f^{(k)}(\xi)/k!$.
Zunächst gilt $T_{k-1}^{(n)}(x_0) = f^{(n)}(x_0)$ für $n = 0, \ldots, k-1$, und damit

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

Per Konstruktion 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 ein x_2 zwischen x_0 und x_1 mit $g''(x_2) = 0$.
So fortfahrend erhalten wir $x_3, \ldots, x_{k-1}$ 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

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

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 $x > 1$.
- 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 $n \in \mathbb{N}$.
- Überall konvergente Taylorentwicklung: $\sin(x)\Big|_{x=0} = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} x^{2k+1}$

Integralrechnung

Definition (Treppenfunktion)

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

Bemerkung (ÜA)

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

Definition (Integral einer Treppenfunktion)

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

Man definiert dann

$$\int_a^b \varphi(x)dx := \sum_{i=1}^{n} c_i(x_i - x_{i-1}).$$

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 φ liegende Gebiet.

F ist eine endliche Vereinigung von Rechtecken ($\rightarrow$ Zeichnung). 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)

Es seien $\varphi, \psi : [a, b] \to \mathbb{R}$ *Treppenfunktionen und* $\lambda \in \mathbb{R}$*. Dann gilt:*

(i) $\int_a^b (\varphi + \psi)(x)dx = \int_a^b \varphi(x)dx + \int_a^b \psi(x)dx$.

(ii) $\int_a^b (\lambda\varphi)(x)dx = \lambda \int_a^b \varphi(x)dx$.

(iii) $\varphi \leq \psi \Longrightarrow \int_a^b \varphi(x)dx \leq \int_a^b \psi(x)dx$.

Beweis. Sei $a = x_0 < x_1 < \cdots < x_n = b$ eine Unterteilung, 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 $i \in \{1, 2, \ldots, n\}$. Damit ist dann

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

$\square$

Definition (Ober- und Unterintegral)

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

$$\int_a^{*b} f(x)dx := \inf \left\{ \int_a^b \varphi(x)dx \;\middle|\; \varphi \quad \textit{Treppenfkt. mit} \quad \varphi \geq f \right\}.$$

Das Unterintegral von f ist die Zahl

$$\int_{*a}^{b} f(x)dx := \sup \left\{ \int_a^b \varphi(x)dx \;\middle|\; \varphi \quad \textit{Treppenfkt. mit} \quad \varphi \leq f \right\}.$$

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 Treppenfkt.

Beispiel

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

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

Dann gilt

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

Satz (Eigenschaften des Ober- und Unterintegrals)

Es seien $f, g : [a.b] \to \mathbb{R}$ *beschränkte Funktionen. Dann gilt:*

(i) $\int^*(f+g) \le \int^* f + \int^* g$,

(ii) $\int^*(\lambda f) = \lambda \int^* f$ *für alle* $\lambda \ge 0$,

(iii) $\int_*(f+g) \ge \int_* f + \int_* g$,

(iv) $\int_*(\lambda f) = \lambda \int_* f$ *für alle* $\lambda \ge 0$,

(v) $\int^*(\lambda f) = \lambda \int_* f$ *und* $\int_*(\lambda f) = \lambda \int^* f$ *für alle* $\lambda \le 0$.

(Hierbei ist $\int^* f = \int_a^{*b} f(x)dx$, *usw.)*

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

Insbesondere gilt $\int_* f = -\int^*(-f)$.

(iii-v) folgt damit aus (i-ii).

Weiter im Beweis:

Wir beweisen nun (i) und (ii).

(i) Wir zeigen $\int^*(f+g) \le \int^* f + \int^* g$:
$\int^*(f+g) = \inf\{\int \xi | \xi \ge f+g \text{ Treppenfkt.}\}$
$\le \inf\{\int \xi | \xi = \varphi + \psi, \varphi \ge f, \psi \ge g \text{ Treppenfktn.}\}$
$= \inf\{\int \varphi + \int \psi | \varphi \ge f, \psi \ge g \text{ Treppenfktn.}\}$
$= \inf\{\int \varphi | \varphi \ge f \text{ Treppenfkt.}\} + \inf\{\int \psi | \psi \ge g \text{ Treppenfkt.}\}$
$= \int^* f + \int^* g$.

(ii) Wir zeigen $\int^*(\lambda f) = \lambda \int^* f$ für $\lambda > 0$:
$\int^*(\lambda f) = \inf\{\int \varphi | \varphi \ge \lambda f \text{ Treppenfkt.}\}$
$= \inf\{\int \varphi | \frac{1}{\lambda}\varphi \ge f \text{ Treppenfkt.}\}$
$= \inf\{\lambda \int \psi | \psi \ge f \text{ Treppenfkt.}\}$
$= \lambda \inf\{\int \psi | \psi \ge f \text{ Treppenfkt.}\} = \lambda \int^* f$.

$\square$

Definition (Integrierbare Funktion)

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

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

Man definiert dann das ***Integral*** *von f als*

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

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

Satz (Linearität und Monotonie des Integrals)

Es seien $f, g : [a, b] \to \mathbb{R}$ integrierbar und $\lambda \in \mathbb{R}$. Dann sind $f + g$ und λf integrierbar und es gilt:

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

(ii) $\int_a^b (\lambda f)(x)dx = \lambda \int_a^b f(x)dx$ *und*

(iii) $f \leq g \Longrightarrow \int_a^b f(x)dx \leq \int_a^b g(x)dx.$

Beweis.

(i) Aus

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

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 λf integrierbar und

$$\int(\lambda f) = \lambda \int f.$$

(iii) Aus $f \leq g$ folgern wir

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

□

Bemerkungen (ÜA)

(i) Sei $a < b < c$. Dann ist $f : [a, c] \to \mathbb{R}$ integrierbar, genau dann wenn $f|_{[a,b]}$ und $f|_{[b,c]}$ integrierbar sind.

Es gilt dann

$$\int_a^c f = \int_a^b f + \int_b^c f.$$

(ii) Positiv- und Negativteil $f_\pm : [a, b] \to \mathbb{R}_+ \cup \{0\}$ einer Fkt. $f : [a, b] \to \mathbb{R}$ sind definiert als $f_+(x) := \max\{f(x), 0\}$ bzw. $f_-(x) := -\min\{f(x), 0\}$, $x \in [a, b]$.

Es gilt: f integrierbar $\Longrightarrow f_+$, f_- und $|f|$ integrierbar.

(iii) Aus der Monotonie des Integrals folgt

$$\left|\int_a^b f\right| \leq \int_a^b |f|$$

für jede integrierbare Funktion f.

Theorem (Integrierbarkeit stetiger Funktionen)

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

Beweis.

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

Lemma

Jede stetige Funktion $f : [a, b] \to \mathbb{R}$ ist gleichmäßig stetig, d.h.

Für alle $\varepsilon > 0$ existiert $\delta > 0$, so daß $|f(x) - f(y)| < \varepsilon$ für alle $x, y \in [a, b]$ mit $|x - y| < \delta$.

Beispiel

Die Funktion $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^2$ ist nicht gleichmäßig stetig, ihre Einschränkung auf ein abgeschlossenes 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 $\varepsilon > 0$ und Folgen $x_n, y_n \in [a, b]$ mit

$$|x_n - y_n| < \frac{1}{n} \quad \text{und} \quad |f(x_n) - f(y_n)| \geq \varepsilon.$$

- Nach Bolzano-Weierstraß existiert eine Teilfolge $(x_{n_k})_{k\in\mathbb{N}}$, die gegen $\ell \in [a, b]$ konvergiert.
- Wegen $\lim_{n\to\infty} |x_n - y_n| = 0$ gilt auch $\lim_{k\to\infty} y_{n_k} = \ell$.
- Die Stetigkeit von f liefert dann

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

- im Widerspruch zu $|f(x_n) - f(y_n)| \geq \varepsilon$ für alle $n \in \mathbb{N}$. □

Beweis des Theorems

- Zu $n \in \mathbb{N}$ betrachten wir die äquidistante Unterteilung $x_i = a + i\frac{(b-a)}{n}, \quad i = 0, 1, \dots, n.$
  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 $f : [a, b] \to \mathbb{R}$ ist integrierbar.

Definition (Riemannsche Summe)

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

- $a = x_0 < x_1 < \cdots < x_n = b$ *eine Unterteilung des Intervalls* $[a, b]$ *und*
- $\xi_i \in [x_{i-1}, x_i]$ *beliebige Zahlen in diesen Teilintervallen.*

Die zugehörige Riemannsche Summe ist die Zahl

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

Die ξ_i heißen Stützstellen.

Bemerkung

Die Riemannsche Summe kann aufgefaßt werden als das Integral einer Treppenfunktion.

Satz (Approximation durch Riemannsche Summen)

Für jede stetige Funktion $f : [a, b] \to \mathbb{R}$ *approximiert die Riemannsche Summe in folgendem Sinne das Integral:*

Für alle $\varepsilon > 0$ *gibt es ein* $\delta > 0$*, so daß*

$$\left| \int_a^b f(x)dx - \sum_{i=1}^{n} f(\xi_i)(x_i - x_{i-1}) \right| < \varepsilon$$

für jede Riemannsche Summe mit $\max\limits_{1 \leq i \leq n} (x_i - x_{i-1}) < \delta$*.*

Bemerkung

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

Beweis.

- Wegen der (gleichmäßigen) Stetigkeit von $f : [a, b] \to \mathbb{R}$ gibt es zu jedem $\varepsilon > 0$ ein $\delta > 0$, so daß

$$|f(x) - f(y)| < \frac{\varepsilon}{b-a} \text{ für alle } |x - y| < \delta.$$

- Für Riemannsche Summen mit $\max\limits_{1 \leq i \leq n}(x_i - x_{i-1}) < \delta$ gilt dann: $|f(x) - f(\xi_i)| < \frac{\varepsilon}{b-a}$ für alle $x \in [x_{i-1}, x_i]$.
- Damit erhalten wir

$$\left| \int_a^b f(x)dx - \sum_{i=1}^n f(\xi_i)(x_i - x_{i-1}) \right| = \left| \sum_{i=1}^n \int_{x_{i-1}}^{x_i} (f(x) - f(\xi_i))dx \right|$$

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

$\square$

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

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

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

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

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

Dies ist die Fläche des gleichseitigen Dreiecks unter dem Graphen von $f(x) = x$.

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

Definition (Stammfunktion)

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

$$F' = f.$$

Stammfunktionen F_1 und F_2 zu gegebenem f unterscheiden sich 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 $f : I \longrightarrow \mathbb{R}$ eine stetige Funktion und $x_0 \in I$.

Die durch

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

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

Beweis des Fundamentalsatzes. Wir benutzen folgendes Lemma:

Lemma (Mittelwertsatz der Integralrechnung)

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

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

Beweis.

- Aus $\min f \leq f \leq \max f$ folgt wg. der Monotonie des Integrals
$$(b-a)\min f \leq \int_a^b f(x)dx \leq (b-a)\max f.$$
- Der Zwischenwertsatz 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:

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

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

Folgerung

Sei $f : I \to \mathbb{R}$ stetig und F eine Stammfunktion von f.

Dann gilt für alle $a, b \in I$:

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

Beweis. 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

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

Beispiele

- $\int\limits_0^\pi \sin(x)dx = -\cos(x)|_0^\pi = -\cos(\pi) + \cos(0) = 2.$
- Ist $\alpha \neq -1$, so gilt $\int\limits_a^b x^\alpha dx = \frac{1}{\alpha+1} x^{\alpha+1}|_a^b$, wobei
  - $a, b \in \mathbb{R}$ beliebig, falls $\alpha \in \mathbb{N}$,
  - $0 \notin [a, b]$, falls $\alpha \in \mathbb{Z}$ und $\alpha \leq -2$, und
  - $0 < a < b$, falls $\alpha \notin \mathbb{Z}$.
- Für $0 < a < b$ erhält man $\int\limits_a^b \frac{1}{x} dx = \ln(x)|_a^b$

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

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

Theorem (Substitutionsregel)

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

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

Beweis.

- Sei $F : I \to \mathbb{R}$ eine Stammfunktion von f.
- Nach der Kettenregel gilt für alle $t \in [a, b]$

$$(F \circ \varphi)'(t) = F'(\varphi(t))\varphi'(t) = f(\varphi(t))\varphi'(t).$$

- Somit

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

Schreibweise: Definiert man $d\varphi(t) := \varphi'(t)dt$, 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

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

-

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

Hier ist $f(x) = \frac{1}{x+1}$, $\varphi(t) = t^2$ und damit $\varphi'(t) = 2t$.

Beispiel: Fläche des Kreises vom Radius 1

Wir stellen den Halbkreis als Graphen der Funktion $f : [-1, 1] \to \mathbb{R}$, $f(x) = \sqrt{1 - x^2}$ dar und müssen $\int_{-1}^{1} \sqrt{1 - x^2} dx$ berechnen.
Die Substitution $x = \varphi(t) = \sin(t)$ ergibt

$$\int_{-1}^{1} \sqrt{1 - x^2} dx = \int_{-\pi/2}^{\pi/2} \sqrt{1 - \sin^2 t}\; d\sin(t) = \int_{-\pi/2}^{\pi/2} \cos^2(t) dt.$$

Es ist aber $\cos^2 t \; = \; \left(\frac{e^{it} + e^{-it}}{2}\right)^2 \; = \; \frac{e^{2it} + e^{-2it}}{4} + \frac{1}{2} \; = \; \frac{1}{2}(\cos 2t + 1)$
Damit ist die Fläche des Kreises vom Radius 1 gegeben durch

$$2 \int_{-1}^{1} \sqrt{1 - x^2} dx \; = \; 2 \int_{-\pi/2}^{\pi/2} \cos^2(t) dt \; = \; \int_{-\pi/2}^{\pi/2} \cos 2t \; dt + \int_{-\pi/2}^{\pi/2} dt$$

$$= \; \frac{1}{2} \int_{-\pi}^{\pi} \cos s \; ds + \int_{-\pi/2}^{\pi/2} dt \; = \; \sin s|_{-\pi}^{\pi} + t|_{-\pi/2}^{\pi/2} \; = \; 0 + \pi \; = \; \pi.$$

Theorem (Partielle Integration)

Es seien $f, g : [a, b] \to \mathbb{R}$ *stetig und 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:

$$(fg)|_a^b = \int_a^b f(fg)'(x)dx = \int_a^b f'(x)g(x)dx + \int_a^b f(x)g'(x)dx.$$

$\square$

Kurzschreibweise für die partielle Integration:

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

Beispiele

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

  Hierbei ist $f(x) = x$ und $g(x) = \ln x$.
  D.h. $x(\ln x - 1)$ ist eine Stammfunktion des Logarithmus.
- $x\arctan x - \frac{1}{2}\ln(x^2+1)$ ist eine Stammfunktion des arctan, denn:

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

  Mittels der Substitution $t = x^2$ hatten wir gesehen, daß
  $\int_a^b \frac{x}{x^2+1}dx = \frac{1}{2}\int_{a^2}^{b^2}\frac{1}{t+1}dx = \frac{1}{2}\ln(t+1)|_{a^2}^{b^2} = \frac{1}{2}\ln(x^2+1)|_a^b$.
  Damit ist $\displaystyle\int_a^b \arctan x\;dx = x\arctan x - \frac{1}{2}\ln(x^2+1)\Big|_a^b$.

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 auf der zwei Verknüpfungen gegeben sind,

die Addition

$$+ : V \times V \to V, \quad (v, w) \mapsto v + w$$

und die skalare Multiplikation

$$\mathbb{K} \times V \to V, \quad (\lambda, v) \mapsto \lambda \cdot v \; (=: \lambda v),$$

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 $0 \in V$, 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, daß 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) $(\lambda \cdot \mu) \cdot v = \lambda \cdot (\mu \cdot v)$,
   (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.

Beispiele von Vektorräumen

1) Der kartesische Raum

$\mathbb{K}^n = \{(x_1, \dots, x_n) | x_1, \dots, x_n \in \mathbb{K}\}$

mit der Addition

$$(x_1, \dots, x_n) + (y_1, \dots, y_n) := (x_1 + y_1, \dots, x_n + y_n)$$

und der skalaren Multiplikation

$$\lambda \cdot (x_1, \dots, x_n) := (\lambda x_1, \dots, \lambda x_n)$$

ist ein Vektorraum über $\mathbb{K}$.

Das neutrale Element 0 der kommutativen Gruppe $(\mathbb{K}^n, +)$ ist der Vektor $(0, \dots, 0)$ und das additive Inverse $-(x_1, \dots, x_n)$ des Vektors $(x_1, \dots, x_n)$ ist $(-x_1, \dots, -x_n)$.

2) Funktionenräume

Sei X eine Menge und $F(X, \mathbb{K})$ die Menge der Funktionen $f : X \to \mathbb{K}$. $F(X, \mathbb{K})$ ist mit der üblichen, punktweise definierten, Addition und skalaren Multiplikation

$$\begin{aligned}(f+g)(x) &= f(x)+g(x) \\ (\lambda f)(x) &= \lambda f(x) \quad (f, g \in F(X, \mathbb{K}), \quad \lambda \in \mathbb{K}, \quad x \in X)\end{aligned}$$

ein Vektorraum über $\mathbb{K}$.

Übungsaufgabe

Sei V ein Vektorraum über $\mathbb{K}$. Für alle $v \in V$, $\lambda \in \mathbb{K}$ gilt

(i) $\lambda v = 0 \iff \lambda = 0$ oder $v = 0$,

(ii) $-v = (-1) \cdot v$.

Definition

Ein Unterraum (genauer: ein Untervektorraum) eines Vektorraumes V ist eine nicht leere Teilmenge $U \subset V$, 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

$$+ : U \times U \to U$$

und eine skalare Multiplikation

$$\cdot : \mathbb{K} \times U \to U,$$

die U zu einem Vektorraum machen.

Beispiele von Unterräumen

Sei V ein Vektorraum.

1) $\{0\}$ und V sind Untervektorräume von V.
2) Für jeden Vektor $v \in V$ ist $\mathbb{K}v = \{\lambda v | \lambda \in \mathbb{K}\} \subset V$ der kleinste Unterraum, der v enthält.
   - Wenn $v \neq 0$, so ist $\mathbb{K}v \neq \{0\}$ und heißt die von v erzeugte (lineare) Gerade.
3) Für jedes Paar von Vektoren v, w ist

$$\mathbb{K}v + \mathbb{K}w = \{\lambda v + \mu w | \lambda, \mu \in \mathbb{K}\} \subset V$$

der kleinste Unterraum, der v und w enthält.

Wenn $v \neq 0$ und $w \notin \mathbb{K}v$, so heißt $\mathbb{K}v + \mathbb{K}w \underset{\neq}{\supset} \mathbb{K}v \underset{\neq}{\supset} \{0\}$ die von v und w erzeugte (lineare) Ebene.

Beachte, die (affine) Gerade $w + \mathbb{K}v$ ist im allgemeinen *kein* Unterraum des Vektorraums V.

4) Sei $I \subset \mathbb{R}$ ein Intervall und $k \in \mathbb{N}$. Die folgenden Mengen sind jeweils Unterräume von $F(I, \mathbb{R})$:

$$\begin{aligned} C^k(I,\mathbb{R}) &:= \{f : I \to \mathbb{R} | f \quad \text{k-mal stetig differenzierbar}\} \\ &\subset Diff(I,\mathbb{R}) := \{f : I \to \mathbb{R} | f \quad \text{differenzierbar}\} \\ &\subset C(I,\mathbb{R}) := \{f : I \to \mathbb{R} | f \quad \text{stetig}\} \end{aligned}$$

5) Sei V ein Vektorraum und $U_j \subset V$ durch $j \in J$ indizierte Unterräume, wobei J eine beliebige Menge ist.

Der Durchschnitt $U = \cap_{j \in J} U_j \subset V$ ist ein Unterraum (ÜA).

Beispielsweise ist

$$C^\infty(I,\mathbb{R}) := \cap_{k \in \mathbb{N}} C^k(I,\mathbb{R})$$

ein Unterraum. Die Funktionen in $C^\infty(I,\mathbb{R})$ heißen unendlich oft differenzierbar.

6) Ein Polynom in der Variablen x mit Koeffizienten in einem Körper $\mathbb{K}$ ist ein Ausdruck der Form

$$a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n,$$

wobei $a_0, \ldots, a_n \in \mathbb{K}$ und $n \in \mathbb{N} \cup \{0\}$.

Die Menge aller solchen Polynome wird mit $\mathbb{K}[x]$ bezeichnet und bildet in offensichtlicher Weise (ÜA) einen Vektorraum.

Jedes Polynom $a_0 + a_1 x + \cdots + a_n x^n \in \mathbb{K}[x]$ definiert eine Funktion

$$P : \mathbb{K} \to \mathbb{K}, \quad \lambda \mapsto P(\lambda) = a_0 + a_1 \lambda + \cdots + a_n \lambda^n.$$

Solche Funktionen P heißen polynomial.

Wenn $\mathbb{K}$ unendlich ist (z.B. $\mathbb{K} = \mathbb{Q}, \mathbb{R}, \mathbb{C}$), so ist diese Zuordnung $a_0 + \cdots + a_n x^n \mapsto P$ injektiv (ÜA) und $\mathbb{K}[x]$ wird auf diese Weise zu einem Unterraum $\mathbb{K}[x] \subset F(\mathbb{K}, \mathbb{K})$. Z.B. haben wir die Inklusion $\mathbb{R}[x] \subset C^\infty(\mathbb{R}, \mathbb{R}) \subset F(\mathbb{R}, \mathbb{R})$.

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 $\lambda_i \in \mathbb{K}$ gibt, so daß*

$$v = \sum_{i=1}^{r} \lambda_i v_i.$$

Die Vektoren $v_1, \dots, v_r$ heißen linear unabhängig*, wenn*

$$\sum_{i=1}^{r} \lambda_i v_i = 0 \implies \lambda_1 = \dots = \lambda_r = 0,$$

d.h. sind $\lambda_1, \dots, \lambda_r$ beliebige Skalare, dann gilt entweder $\sum_{i=1}^{r} \lambda_i v_i \neq 0$ oder $\lambda_1 = \dots = \lambda_r = 0$.

(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 Vektoren $(v_j)_{j \in J}$ heißen linear unabhängig, wenn die Vektoren $(v_j)_{j \in J_0}$ für jede endliche Teilmenge $J_0 \subset J$ linear unabhängig sind.

Beispiele

1) Wir setzen
$$e_i := \underset{\text{(i-te Stelle)}}{(0,\dots,0,1,0,\dots,0)} \in \mathbb{K}^n.$$
Die Vektoren $(e_1,\dots,e_n)$ sind linear unabhängig. In der Tat:
$$0 = \sum_{i=1}^{n} \lambda_i e_i = (\lambda_1,\dots,\lambda_n) \Longrightarrow \lambda_1 = \dots = \lambda_n = 0.$$

2) Die Monome $(1, x, x^2, x^3, \dots)$ sind linear unabhängig in $\mathbb{K}[x]$.

Definition (Lineare Hülle)

Sei V ein Vektorraum und $(v_j)_{j\in J}$ eine Familie von Vektoren.

Die Lineare Hülle der $(v_j)_{j\in J}$ ist der Unterraum

$$span\{v_j | j \in J\} := \{v | v \text{ ist eine Linearkombination der } (v_j)_{j\in J}\}.$$

Beispiel: Sei $v_1 = (1,1,0)$ und $v_2 = (1,-1,0) \in \mathbb{R}^3$. Dann ist

$$\begin{aligned} span(v_1, v_2) &= \{\lambda(1,1,0) + \mu(1,-1,0) \mid \lambda, \mu \in \mathbb{R}\} \\ &= \{(\lambda, \lambda, 0) + (\mu, -\mu, 0) \mid \lambda, \mu \in \mathbb{R}\} \\ &= \{(x, y, 0) \mid x, y \in \mathbb{R}\} \\ &= span(e_1, e_2) \end{aligned}$$

für $e_1 = (1,0,0)$ und $e_2 = (0,1,0) \in \mathbb{R}^3$.

Satz

Die Vektoren $v_j \in V$, $j \in J$, *seien linear unabhängig.*

Dann hat jeder Vektor $v \in \mathrm{span}\{v_j | j \in J\}$ *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.

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

$$\begin{aligned} \lambda_j - \lambda_j' &= 0, \quad \text{wenn} \quad j \in J_0 \cap J_0' \\ \lambda_j &= 0, \quad \text{wenn} \quad j \in J_0 \setminus J_0 \cap J_0' \\ \lambda_j' &= 0, \quad \text{wenn} \quad j \in J_0' \setminus J_0 \cap J_0'. \end{aligned}$$

Also stimmen die beiden Darstellungen
$v = \sum_{j\in J_0} \lambda_j v_j = \sum_{j\in J_0\cap J_0'} \lambda_j v_j$ und $v = \sum_{j\in J_0'} \lambda_j' v_j = \sum_{j\in J_0\cap J_0'} \lambda_j' v_j$
überein. $\square$

Definition (Erzeugendensysteme, Basen)

*Sei V ein Vektorraum. Eine Familie von Vektoren $(v_j)_{j\in J}$ heißt **Erzeugendensystem** von V, falls $V = \mathrm{span}\{v_j | j \in J\}$.*

*Eine **Basis** von V ist ein linear unabhängiges Erzeugendensystem.*

Beispiele

1) $(e_1, \ldots, e_n)$ ist eine Basis von $\mathbb{K}^n$, die so genannte kanonische Basis.
2) $(1, x, x^2, \ldots)$ ist eine Basis von $\mathbb{K}[x]$.
3) $(e_1, e_2, e_1 + e_2, 2e_1)$ ist ein Erzeugendensytem von $\mathbb{K}^2$, aus dem wir folgende Basen auswählen können: (e_1, e_2), $(e_1, e_1 + e_2)$ und $(e_2, e_1 + e_2)$. Falls in $\mathbb{K}$ $0 \neq 2$ $(:= 1 + 1)$ gilt, so sind $(e_2, 2e_1)$ und $(e_1 + e_2, 2e_1)$ ebenfalls Basen.

Satz (Charakterisierung von Basen)

Sei $V \neq 0$ ein VR. 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 $(v_j)_{j \in J_0}$ 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 $(v_j)_{j \in J_0}$ linear unabhängig ist, so ist $J_0 = J$.*

(iii) *Jeder Vektor $v \in V$ besitzt genau eine Darstellung*

$$v = \sum_{j \in J} \lambda_j v_j,$$

wobei $\lambda_j = 0$ für fast alle $j \in J$.

Beweis.

Sei (0) die Aussage, daß $(v_j)_{j\in J}$ eine Basis bildet. Wir zeigen $(0) \Rightarrow (iii) \Rightarrow (i) \Rightarrow (ii) \Rightarrow (0)$.

$(0) \Rightarrow (iii)$ Wir wissen bereits, daß jeder Vektor aus $\mathrm{span}\{v_j | j \in J\}$ eine eindeutige Darstellung als (endliche) Linearkombination der linear unabhängigen Vektoren $(v_j)_{j\in J}$ hat. Da $(v_j)_{j\in J}$ ein Erzeugendensystem ist, gilt das für jeden Vektor aus $V = \mathrm{span}\{v_j | j \in J\}$.

$(iii) \Rightarrow (i)$ Aus (iii) folgt offensichtlich, daß $(v_j)_{j\in J}$ ein Erzeugendensystem ist.

Um die Minimalität zu zeigen, nehmen wir an, daß $(v_j)_{j\in J_0}$, $J_0 \subset J$, ein Erzeugendensystem ist.

Dann hat jeder der Vektoren v_i, $i \in J$, eine Darstellung (*) $v_i = \sum_{j\in J_0} \lambda_j v_j$. Nach (iii) ist das die eindeutige Darstellung als Linearkombination der $(v_j)_{j\in J}$.

Daraus folgt $i \in J_0$, denn sonst wären $v_i = v_i$ und (*) zwei verschiedene Darstellungen. 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 $j_0 \in J_0 \quad \lambda_{j_0} \neq 0$ 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 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 $(v_j)_{j\in J_0}$, $J \neq J_0 \supset J$, 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 Darst. $v_{j_0} = \sum_{j\in J} \lambda_j v_j$.

Das zeigt, daß $(v_j)_{j\in J_0}$ linear abhängig ist. Somit ist die lin. unabh. Familie $(v_j)_{j\in J}$ maximal.

Ende des Beweises

$(ii) \Rightarrow (0)$ Sei $(v_j)_{j \in J}$ eine maximale lin. unabh. Familie.

Wir zeigen, daß $(v_j)_{j \in J}$ ein Erzeugendensystem und somit eine Basis ist.

Sonst gäbe es $v \in V \setminus \mathrm{span}\{v_j | j \in J\}$ und $(v, (v_j)_{j \in J})$ wäre eine linear unabhängige Familie, im Widerspruch zur Maximalität von $(v_j)_{j \in J}$. □

Folgerung

Jedes endliche Erzeugendensystem eines Vektorraums V enthält eine Basis.

Beweis. Jedes endliche Erzeugendensystem enthält ein minimales Erzeugendensystem. □

Lemma (Austauschlemma)

Sei $(v_1, \ldots, v_n)$ *eine Basis von* V *und*

$$w = \sum_{i=1}^{n} \lambda_i v_i$$

eine Linearkombination.

Wenn $\lambda_k \neq 0$*, so ist* $(v_1, \ldots, v_{k-1}, w, v_{k+1}, \ldots, v_n)$ *eine Basis von* V*.*

Beweis.

Übungsaufgabe. $\square$

Satz (Steinitzscher Austauschsatz)

Sei V ein VR, $(v_1, \dots, v_n)$ eine Basis und $(w_1, \dots, w_r)$ linear unabhängige Vektoren.

Dann gilt $n \geq r$ und es gibt eine Permutation $\varphi : \{1, \dots, n\} \to \{1, \dots, n\}$, so daß $(w_1, \dots, w_r, v_{\varphi(r+1)}, \dots, v_{\varphi(n)})$ eine Basis ist.

Beweis.

Beweis durch Induktion nach r:

Für $r = 0$ ist nichts zu zeigen. Wir führen den Induktionsschritt $r \to r + 1$ aus.

Seien also $(w_1, \dots, w_{r+1})$ lin. unabh. und $(v_1, \dots, v_n)$ eine Basis.

Nach Induktionsvor. gilt $r \leq n$ und es gibt eine Permutation $\varphi : \{1, \dots, n\} \to \{1, \dots, n\}$, so daß $(w_1, \dots, w_r, v_{\varphi(r+1)}, \dots, v_{\varphi(n)})$ eine Basis ist.

Weiter im Beweis:

Wir können daher schreiben

$$w_{r+1} = \sum_{i=1}^{r} \lambda_i w_i + \sum_{j=r+1}^{n} \lambda_j v_{\varphi(j)}.$$

Da die $(w_1, \ldots, w_{r+1})$ lin. unabh. sind, existiert $j_0 \in \{r+1, \ldots, n\}$ mit $\lambda_{j_0} \neq 0$. Insbesondere ist $r + 1 \leq n$.

Wir können (ggf. nach Abänderung von φ) annehmen, daß $j_0 = r + 1$.

Nach dem Austauschlemma können wir dann in der Basis $(w_1, \ldots, w_r, v_{\varphi(r+1)}, \ldots, v_{\varphi(n)})$ den Vektor $v_{\varphi(j_0)} = v_{\varphi(r+1)}$ durch w_{r+1} ersetzen und erhalten die Basis $(w_1, \ldots, w_{r+1}, v_{\varphi(r+2)}, \ldots, v_{\varphi(n)})$. $\square$

Theorem

Sei V ein Vektorraum.

(i) *Wenn V eine endliche Basis besitzt, dann ist jede Basis von V endlich.*

(ii) *Alle endlichen Basen von V haben die gleiche Anzahl von Elementen.*

Beweis.

(i) Sei $(v_1, \dots, v_n)$ 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, $r \leq n$ und ebenso $n \leq r$.

□

Definition (Dimension)

Sei V ein Vektorraum über $\mathbb{K}$.

Die Dimension von V ist die Zahl

$$\dim V \; (= \dim_{\mathbb{K}} V) := \begin{cases} 0, & \textit{falls} \quad V = \{0\} \\ n, & \textit{falls} \quad V \quad \textit{eine endliche Basis} \\ & \qquad (v_1, \ldots, v_n) \quad \textit{hat} \\ \infty & \textit{sonst} \end{cases}$$

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 $\dim_{\mathbb{R}} V = 2 \dim_{\mathbb{C}} V$. 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) $\dim_{\mathbb{K}} F(X, \mathbb{K}) = \mathrm{card}(X)$, wobei

$$\mathrm{card}(X) := \begin{cases} n, & \text{falls} \quad X \quad \text{aus n Elementen besteht} \\ \infty, & \text{falls} \quad X \quad \text{unendlich ist.} \end{cases}$$

5) $\dim_{\mathbb{K}} \mathbb{K}[x] = \infty$.
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

Sei $V = \mathbb{K}^n$ und $v_1, \dots, v_m \in V$ und

$$U = span\{v_1, \dots, v_m\} \subset V$$

der von $(v_1, \dots, v_m)$ aufgespannte Unterraum.
Wir wollen nun eine Basis des Unterraumes U bestimmen. Dazu benutzt man den Gaußschen Algorithmus (Gaußsches Eliminierungsverfahren). Dieser beruht auf den folgenden Fakten, die sich aus dem Austauschsatz ergeben:

(i) $span\{v_1, \dots, v_m\} = span\{v_{\varphi(1)}, \dots, v_{\varphi(m)}\}$ für jede Permutation $\varphi : \{1, \dots, m\} \to \{1, \dots, m\}$,

(ii) $span\{\lambda v_1, v_2, \dots, v_m\} = span\{v_1, \dots, v_m\}$ für alle $\lambda \in \mathbb{K}^* := \mathbb{K} \setminus \{0\}$ und

(iii) $span\{v_1, \dots, v_{i-1}, v_i + \lambda v_1, v_{i+1}, \dots, v_m\} = span\{v_1, \dots, v_m\}$ 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 ersten $k \le m$ Zeilen $w_i = (b_{i1}, \dots, b_{in})$, $i = 1, \dots, k$, eine Basis von U bilden und $b_{ij} = 0$ für alle $k < i \le m$, $j = 1, \dots, n$.

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 einer Zahl $\lambda \in \mathbb{K}^* := \mathbb{K} \setminus \{0\}$,

(iii) Addieren einer Zeile zu einer anderen.

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 $0 \le k \le m$ und k Indizes $1 \le j_1 < j_2 < \ldots < j_k \le n$ so daß die Matrixeinträge b_{ij} die folgenden Eigenschaften haben:

(1) $b_{1j_1} = b_{2j_2} = \ldots = b_{kj_k} = 1$,

(2) $b_{ij} = 0$ für alle $i = 1, \ldots, k$ und $1 \le j < j_i$,

(3) $b_{ij} = 0$ für alle $i = k+1, \ldots, m$ und $j = 1, \ldots n$.

Eine Matrix $B = (b_{ij})_{\substack{i=1\dots m \\ j=1,\dots n}}$ in Zeilenstufenform sieht so aus:

$$\begin{pmatrix}
0 & \overset{j_1-1 \text{ viele}}{\dots} & 0 & \color{red}{1} & b_{1j_1+1} & & & \dots & & & b_{1n} \\
0 & & \overset{j_2-1 \text{ viele}}{\dots} & & 0 & \color{red}{1} & b_{2j_2+1} & & \dots & & b_{2n} \\
\vdots & & & & & & & & & & \vdots \\
0 & & & \overset{j_i-1 \text{ viele}}{\dots} & & & 0 & \color{red}{1} & b_{ij_i+1} & \dots & b_{in} \\
\vdots & & & & & & & & & & \vdots \\
0 & & & & \overset{j_k-1 \text{ viele}}{\dots} & & & 0 & \color{red}{1} & b_{kj_k+1} \dots & b_{kn} \\
0 & & & & & \dots & & & & & 0 \\
\vdots & & & & & & & & & & \vdots \\
0 & & & & & \dots & & & & & 0
\end{pmatrix}$$

Konklusion

Die Vektoren, die aus den ersten k Zeilen von B bestehen,

$$
\begin{array}{rcl}
w_1 & = & (0 \overset{j_1-1 \text{ viele}}{\dots} 0, 1, b_{1j_1+1}, \dots\dots\dots, b_{1n}) \\
\vdots & & \vdots \\
w_k & = & (0 \qquad \overset{j_k-1 \text{ viele}}{\dots} \qquad 0, 1, b_{kj_k+1}, \dots, b_{kn})
\end{array}
$$

bilden eine Basis von $U = span(v_1, \dots v_m)$, 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 $(w_1, \dots w_k)$ sind linear unabhängig, da

$$0 = \sum_{i=1}^{k} \lambda_i w_i = (0 \dots 0, \lambda_1, *, \lambda_2 + \lambda_1(\ *\ ), *, \lambda_3 + \lambda_2(\ *\ ) + \lambda_1(\ *\ ), etc.)$$

impliziert $\lambda_i = 0$ für $i = 1, \dots k$.

Zahlenbeispiel.
Sei z.B. $\mathbb{K} = \mathbb{Q}$, $= \mathbb{R}$ oder $= \mathbb{C}$ und

$$\begin{aligned} v_1 &= (0,0,2,1,0) \\ v_2 &= (0,1,0,2,1) \\ v_3 &= (0,2,1,1,1) \\ v_4 &= (0,4,4,3,2) \end{aligned}$$

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

$$\begin{pmatrix} \color{red}{0} & \color{red}{1} & 0 & 2 & 1 \\ \color{red}{0} & 0 & 2 & 1 & 0 \\ \color{red}{0} & 2 & 1 & 1 & 1 \\ \color{red}{0} & 4 & 4 & 3 & 2 \end{pmatrix}$$

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:

$$\begin{pmatrix} 0 & 1 & 0 & 2 & 1 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 2 & 1 & 1 & 1 \\ 0 & 4 & 4 & 3 & 2 \end{pmatrix} \overset{-2\times I+III}{\longrightarrow} \begin{pmatrix} 0 & 1 & 0 & 2 & 1 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 1 & -3 & -1 \\ 0 & 4 & 4 & 3 & 2 \end{pmatrix}$$

$$\overset{-4\times I+IV}{\longrightarrow} \begin{pmatrix} 0 & 1 & 0 & 2 & 1 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 1 & -3 & -1 \\ 0 & 0 & 4 & -5 & -2 \end{pmatrix}$$

Als Nächstes produzieren wir Nullen unterhalb von $w_{2j_2} = w_{23} = 2 \neq 0$ durch Addition von Vielfachen der zweiten Zeile:

Weiter im Zahlenbeispiel:

$$\begin{pmatrix} 0 & 1 & 0 & 2 & 1 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 1 & -3 & -1 \\ 0 & 0 & 4 & -5 & -2 \end{pmatrix} \overset{-\frac{1}{2}\times II + III}{\underset{-2\times II + IV}{\longrightarrow}} \begin{pmatrix} 0 & 1 & 0 & 2 & 1 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & -\frac{7}{2} & -1 \\ 0 & 0 & 0 & -7 & -2 \end{pmatrix}$$

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

$$\begin{aligned} w_1 &= (0, 1, 0, 2, 1) \\ w_2 &= (0, 0, 1, \frac{1}{2}, 0) \\ w_3 &= (0, 0, 0, 1, \frac{2}{7}) \end{aligned}$$

Somit ist (w_1, w_2, w_3) eine Basis von $U = span\{v_1, v_2, v_3, v_4\} \subset \mathbb{K}^5$.

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

$$\begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & = & b_1 \\ & \vdots & \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n & = & b_m \end{array}$$

verwendet werden. Hierbei sind die $a_{ij}, b_j \in \mathbb{K}$ gegeben. Gesucht sind die n Unbekannten $x_1, x_2, \ldots, x_n \in \mathbb{K}$.
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).

Lineare Abbildungen

Definition

V, W seien Vektorräume über $\mathbb{K}$.

Eine Abbildung $F : V \to W$ heißt linear (genauer: $\mathbb{K}$-linear), wenn

$$\begin{aligned} F(v+w) &= F(v)+F(w) \quad \text{und} \\ F(\lambda v) &= \lambda F(v) \end{aligned}$$

für alle $v, w \in V$, $\lambda \in \mathbb{K}$.

Eine bijektive lineare Abbildung $F : V \to W$ heißt auch Isomorphismus (von Vektorräumen).

Die Vektorräume V und W heißen isomorph, wenn es einen Isomorphismus $F : V \to W$ gibt.

Grundlegende Eigenschaften:

1. Für jede Menge X und jeden VR V bilden die Abbildungen $X \to V$ einen Vektorraum, der mit $Abb(X, V)$ bezeichnet wird. Die Operationen sind $(\lambda \cdot f)(x) := \lambda \cdot f(x)$ und $(f+g)(x) := f(x) + g(x)$. Das Nullelement von $L(V, W)$ ist die Abbildung $f(v) = 0$ für alle $v \in V$.
2. Seien V, W Vektorräume. Die Menge aller linearen Abbildungen $V \to W$ ist ein Unterraum von $Abb(V, W)$, der mit $L(V, W)$ bezeichnet wird. D.h. λf und $f + g$ sind linear, falls f und g linear sind.
3. Lineare Abbildung sind bestimmt durch ihre Werte auf Basisvektoren. D.h. sind $f, g \in L(V, W)$ und $(v_1, \ldots, v_n)$ ist eine Basis von V mit $f(v_i) = g(v_i)$ für $i = 1, \ldots, n$, dann ist $f = g$.
   Dies gilt da jedes $v \in V$ eine eindeutige Darstellung $v = \sum_{i=1}^n \lambda_i v_i$ hat und f und g linear sind, d.h.
   $$f(v) = f\left(\sum_{i=1}^n \lambda_i v_i\right) = \sum_{i=1}^n \lambda_i f(v_i) = \sum_{i=1}^n \lambda_i g(v_i) = g(v).$$

Weitere wichtige Eigenschaften

1. Sei $f : U \to V$ linear und $g : V \to W$ linear. Dann ist auch $g \circ f : U \to W$ linear.
2. Sei $F : V \to W$ linear. Für jeden Unterraum $V' \subset V$ ist das Bild $F(V') \subset W$ ein Unterraum.
3. Ebenso ist das Urbild $F^{-1}(W') \subset V$ ein Unterraum für jeden Unterraum $W' \subset W$.
4. Insbesondere sind das Bild, $Im(f) := F(V) \subset W$, und der Kern, $\ker F := F^{-1}(0) = \{v \in V | F(v) = 0\} \subset V$, einer linearen Abbildung F Unterräume.
5. F ist genau dann injektiv, wenn $\ker F = \{0\}$.
6. Sei $F : V \to W$ ein Isomorphismus. Dann ist $F^{-1} : W \to V$ linear.
7. Eine lineare Abbildung $F : V \to W$ ist ein Isomorphismus genau dann, wenn $Im(F) = W$ und $Ker(F) = \{0\}$.

Satz

Es seien V, W Vektorräume, $F : V \to W$ linear und $(v_1, \ldots, v_n)$ 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 $(v_1, \ldots, v_n)$ eine Basis von V.

'$\Rightarrow$' wegen (iii).

'$\Leftarrow$' Sei umgekehrt $\dim W = \dim V = n$. Dann existiert eine Basis $(w_1, \ldots, w_n)$ von W mit n Elementen. Die lineare Abbildung $F : V \to W$ definiert durch $F(v_i) = w_i$ ist ein Isomorphismus. □

Satz

Sei V ein Vektorraum über $\mathbb{K}$ und $(v_1, \dots, v_n)$ eine Basis von V. Dann definiert die Zuordnung

$$\begin{aligned} \varphi \ : \ L(V, \mathbb{K}) & \rightarrow \mathbb{K}^n \\ F & \mapsto (F(v_i))_{i=1}^n := (F(v_1), F(v_2), \dots, F(v_n)) \end{aligned}$$

einen Isomorphismus. Insbesondere gilt $\dim(V) = dim(L(V, \mathbb{K}))$.

Beweis.

- φ ist offensichtlich linear.
- $\varphi : F \rightarrow (F(v_i))_i$ ist injektiv, denn aus $0 = \varphi(F)$ folgt $F(v_1) = F(v_2) = \cdots = F(v_n) = 0$ und damit wegen der Linearität $F(v) = 0$ für jeden Vektor $v = \sum_{i=1}^n \lambda_i v_i \in V$. D.h. $F = 0$.
- φ ist auch surjektiv: Sei $(a_1, \dots, a_n) \in \mathbb{K}^n$ beliebig. Dann ist die Abbildung $F(\sum_{i=1}^n \lambda_i v_i) = (a_1\lambda_1, \dots a_n\lambda_n)$ linear und erfüllt $F(v_i) = a_i$, d.h. $\varphi(F) = (a_1, \dots, a_n)$. □

Der Dualraum

Definition

Sei V ein Vektorraum der Dimension n. Der n dimensionale Vektorraum

$$V^* := L(V, \mathbb{K})$$

heißt Dualraum zu V oder auch Raum der Linearformen.

Duale Basen

- Ist $(v_1, \ldots, v_n)$ eine Basis von V so liefert der Isomorphismus φ aus dem Satz eine Basis von V^* mittels $\sigma^i := \varphi^{-1}(e_i)$, d.h.

  $$\sigma^i(v_j) = \begin{cases} 1 & \text{falls } i = j \\ 0 & \text{falls } i \neq j \end{cases}$$

  Diese Basis heißt die zu $(v_1, \ldots, v_n)$ duale Basis.
- Ist $V = \mathbb{K}^n$ so heißt die zur kanonischen Basis $(e_1, \ldots, e_n)$ von $\mathbb{K}^n$ duale Basis die kanonische Basis von $(\mathbb{K}^n)^*$

Der Vektorraum der $(m \times n)$-Matrizen

Die Menge der Matrizen mit m Zeilen und n Spalten, definiert durch

$$Mat_n^m(\mathbb{K}) := \left\{ A = \left(a_j^i\right)_{\substack{i=1,\dots m \\ j=1,\dots n}} := \begin{pmatrix} a_1^1 & \dots & a_n^1 \\ a_1^2 & \dots & a_n^2 \\ \vdots & & \vdots \\ a_1^m & \dots & a_n^m \end{pmatrix} \mid a_j^i \in \mathbb{K} \right\}$$

bildet mit den Operationen

$$\begin{aligned} \lambda\left(a_j^i\right)_{\substack{i=1,\dots m \\ j=1,\dots n}} &:= \left(\lambda a_j^i\right)_{\substack{i=1,\dots m \\ j=1,\dots n}} \text{ und} \\ \left(a_j^i\right)_{\substack{i=1,\dots m \\ j=1,\dots n}} + \left(b_j^i\right)_{\substack{i=1,\dots m \\ j=1,\dots n}} &:= \left(a_j^i + b_j^i\right)_{\substack{i=1,\dots m \\ j=1,\dots n}} \end{aligned}$$

einen Vektorraum, der in offensichtlicher Weise isomorph ist zum Vektorraum $\mathbb{K}^{nm}$.

Notation: Ab jetzt schreiben wir Vektoren in $\mathbb{K}^r$ als Spaltenvektoren mit oberen Indizes

$$\begin{pmatrix} a^1 \\ \vdots \\ a^r \end{pmatrix} \in \mathbb{K}^r.$$

Satz

Der VR $L(\mathbb{K}^n, \mathbb{K}^m)$ ist in kanonischer Weise isomorph zu $Mat_n^m(\mathbb{K})$.

Beweis. Wir fixieren in $\mathbb{K}^n$ die kanonische Basis $(e_i)_{i=1}^n$.
Einem $F \in L(\mathbb{K}^n, \mathbb{K}^m)$ ordnen wir dann die $m \times n$-Matrix

$$\varphi \ : \ F \mapsto (F(e_1) \ \dots \ F(e_n))$$

zu wobei die $F(e_i) \in \mathbb{K}^m$ als Spaltenvektoren zu verstehen sind.
Diese Abbildung ist linear und injektiv, denn $F(e_i) = 0$ für alle $i = 1, \dots, n$ impliziert $F = 0$.

Weiter im Beweis:
Die lineare Abbildung φ ist aber auch surjektiv:
Sei $A = \left(a^i_j\right)_{\substack{i=1,\dots m\\ j=1,\dots n}}$ eine beliebige Matrix.
Dazu definiert man ein $F \in L(\mathbb{K}^n, \mathbb{K}^m)$ mittels

$$F(e_j) := \sum_{i=1}^{m} a^i_j e_i.$$

Damit ist die j-te Spalte von $\varphi(F)$,

$$a_j = \begin{pmatrix} a^1_j \\ \vdots \\ a^m_j \end{pmatrix} = \sum_{i=1}^{m} a^i_j e_i \in \mathbb{K}^m,$$

das Bild $F(e_j)$ des j-ten Basisvektors e_j der kanonischen Basis von $\mathbb{K}^n$.
Somit ist $\varphi(F) = A$. $\square$

Produkt einer Matrix mit einem Vektor

Definition

Das Produkt einer $(m \times n)$-Matrix $A = \left(a^i_j\right)_{\substack{i=1,\dots m\\ j=1,\dots n}}$ 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

$$Ax = \begin{pmatrix} a^1_1 & \dots & a^1_n \\ \vdots & & \vdots \\ a^m_1 & \dots & a^m_n \end{pmatrix} \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix} := \begin{pmatrix} y^1 := \sum_{j=1}^n a^1_j x_j \\ \vdots \\ y^m := \sum_{j=1}^n a^m_j x_j \end{pmatrix} \in \mathbb{K}^m$$

Sein nun $F \in L(\mathbb{K}^n, \mathbb{K}^m)$ eine lineare Abbildung und $A = (a^i_j)_{\substack{i=1,\dots m \\ j=1,\dots n}}$ die kanonisch zugeordnete Matrix, d.h. die j-te Spalte

$$\begin{pmatrix} a^1_j \\ \vdots \\ a^m_j \end{pmatrix} = \sum_{i=1}^{m} a^i_j e_i \in \mathbb{K}^m$$

der Matrix A ist dabei genau das Bild $F(e_j)$ des j-ten Basisvektors e_j der kanonischen Basis von $\mathbb{K}^n$. Daher ist für $x = \sum\limits_{j=1}^{n} x^j e_j$

$$\begin{aligned} F(x) &= F(\sum_{j=1}^{n} x^j e_j) = \sum_{j=1}^{n} x^j F(e_j) \\ &= \sum_{j=1}^{n} x^j \sum_{i=1}^{m} a^i_j e_i = \sum_{j=1}^{n} \left(\sum_{i=1}^{m} x^j a^i_j \right) e_i = Ax \end{aligned}$$

D.h. das Bild von x unter F ist gleich dem Produkt Ax.

Matrixprodukt zweier Matrizen

Definition

Seien nun $A = (a^i_j)_{i,j} \in Mat^m_n(K)$ *und* $B = (b^j_k)_{j,k} \in Mat^n_p(\mathbb{K})$
Das ***Produkt*** $C = AB \in \mathrm{Mat}^m_p(\mathbb{K})$ *von* A *und* B *ist die Matrix*
$C = (c^i_k)_{\substack{i=1\dots m \\ k=1\dots p}}$ *mit* $c^i_k = \sum_{j=1}^n a^i_j b^j_k$.

Diese Matrix entspricht der Matrix der Verknüpfung der linearen Abbildungen $A \circ B \in L(\mathbb{K}^p, \mathbb{K}^m) = \mathrm{Mat}^m_p(\mathbb{K})$, denn für $x = \sum_{k=1}^p x^k e_k \in \mathbb{K}^p$ ist

$$\begin{aligned}(AB)x &= A(Bx) = A(\sum_{j=1}^n \sum_{k=1}^p b^j_k x^k e_j) = \sum_{k=1}^p \sum_{j=1}^n b^j_k x^k A(e_j) \\ &= \sum_{k=1}^p \sum_{j=1}^n b^j_k x^k \sum_{i=1}^m a^i_j e_i = \sum_{i=1}^m \sum_{k=1}^p (\sum_{j=1}^n a^i_j b^j_k) x^k e_i\end{aligned}$$

Also $AB = C = (c^i_k)_{i,k}$ mit $c_{ik} = \sum_{j=1}^n a^i_j b^j_k$.

Darstellende Matrix einer linearen Abbildung

$B = (v_1, \ldots, v_n)$ bzw. $B' = (w_1, \ldots, w_m)$ seien Basen der $\mathbb{K}$-Vektorräume V bzw. W. $\phi_B : \mathbb{K}^n \to V$ und $\phi_{B'} : \mathbb{K}^m \to W$ seien die zugehörigen Isomorphismen definiert durch $\phi_B(e_i) = v_i$, $\phi_{B'}(e_j) = w_j$, wobei $i = 1, \ldots, n$, $j = 1, \ldots, m$ und $(e_i)_i$ die entsprechende kanonische Basis ist.

Definition

Sei $F \in L(V, W)$. Die Matrix

$$M_B^{B'}(F) := \phi_{B'}^{-1} \circ F \circ \phi_B \in L(\mathbb{K}^n, \mathbb{K}^m) = \mathrm{Mat}_n^m(\mathbb{K})$$

heißt darstellende Matrix der linearen Abbildung F bez. der Basen B, B'.

Beispiel

Seien $V = \mathbb{K}^n$, $W = \mathbb{K}^m$ und B bzw. B' die zugehörigen *kanonischen* Basen. Dann gilt $\phi_B = \mathrm{Id}_V$, $\phi_{B'} = \mathrm{Id}_W$ und $M_B^{B'}(F) = F$ für jede lineare Abbildung $F \in L(\mathbb{K}^n, \mathbb{K}^m) = \mathrm{Mat}_n^m(\mathbb{K})$.

Satz

Mit den obigen Bezeichnungen ist die darstellende Matrix $A = (a^i_j)_{i,j} = M^{B'}_B(F)$ *durch folgende Gleichung bestimmt:*

$$F(v_j) = \sum_{i=1}^{m} a^i_j w_i, \quad j = 1, \ldots, n. \tag{11}$$

Beweis. Die $A = M^{B'}_B(F)$ definierende Gleichung

$$(\phi^{-1}_{B'} \circ F \circ \phi_B)(e_j) = A(e_j) = \sum_{i=1}^{m} a^i_j e_i$$

geht durch Anwenden von $\phi_{B'}$ in die äquivalente Gleichung (11) über, da $\phi_B(e_j) = v_j$ und $\phi_{B'}(e_i) = w_i$. $\square$

Satz

Seien V, V', V'' endlichdimensionale 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 leicht aus der von F und G.
Die darstellende Matrix $M_B^{B''}(G \circ F)$ ergibt sich aus

$$\phi_{B''}^{-1} \circ (G \circ F) \circ \phi_B = (\phi_{B''}^{-1} \circ G \circ \phi_{B'}) \circ (\phi_{B'}^{-1} \circ F \circ \phi_B).$$

□

Endomorphismen

Definition

Sei V ein Vektorraum. Lineare Abbildungen von V nach V heißen auch ***Endomorphismen*** *von V, $\mathrm{End}(V) := L(V, V)$.*

Quadratische $(n \times n)$-Matrizen bezeichnet man mit $\mathrm{Mat}(n, \mathbb{K}) := \mathrm{Mat}_n^n(\mathbb{K})$.

Für die darstellende Matrix eines Endomorphismus $F \in \mathrm{End}(V)$ bez. einer Basis B von V schreibt man $M_B(F) := M_B^B(F)$.

Beispiel

Sei $F = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \in \mathrm{Mat}(2, \mathbb{R}) = \mathrm{End}(\mathbb{R}^2)$. Die darstellende Matrix bez. der Basis $B = (b_1, b_2) = (e_1 + e_2, e_1 - e_2)$ ist $M_B(F) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

In der Tat: $F(b_1) = e_2 + e_1 = b_1$ und $F(b_2) = e_2 - e_1 = -b_2$.

Rang von linearen Abbildungen und Matrizen

Definition (Rang einer linearen Abbildung)

Der Rang einer linearen Abbildung $F : V \to W$ ist

$$rg(F) := \dim F(V) \leq \dim W.$$

Bemerkung

Sei $A = (a^i_j)_{i,j} \in \mathrm{Mat}^m_n(\mathbb{K}) = L(\mathbb{K}^n, \mathbb{K}^m)$ eine Matrix. Deren n Spalten sind gegeben durch $A(e_1), \ldots, A(e_n) \in \mathbb{K}^m$.

- D.h. $rg(A)$ ist gleich der Dimension des von den Spalten von A erzeugten Unterraumes von $\mathbb{K}^m$. Diese nennt man Spaltenrang von A.
- Die Dimension des von den m Zeilen $(a^i_1, \ldots, a^i_n) \in \mathbb{K}^n$ erzeugten Unterraumes von $\mathbb{K}^n$ heißt Zeilenrang von A.
- Wir werden später sehen, dass Zeilenrang=Spaltenrang. Damit kann man $rg(A)$ mit dem Gaußschen Algorithmus bestimmen.

Satz

Seien V, W endlichdimensionale Vektorräume und B, B' Basen von V bzw. W. Dann gilt für jede lineare Abbildung $F : V \to W$

$$\mathrm{rg}(F) = \mathrm{rg}(M_B^{B'}(F)).$$

Beweis. Das Bild von $F \circ \phi_B$ stimmt mit dem von F überein und wird durch $\phi_{B'}^{-1}$ isomorph auf das Bild von $M_B^{B'}(F)$ abgebildet. Also $\mathrm{rg}(F) = \mathrm{rg}(F \circ \phi_B) = \mathrm{rg}(M_B^{B'}(F))$. □

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 zweier Basen bestimmt.

Rang und Dimension

Satz (Dimensionsformel)

Seien V, W Vektorräume, dim V *endlich und* $F \in L(V, W)$. *Dann gilt*

$$\mathrm{rg}(F) + \dim \ker(F) \ = \ \dim V.$$

Beweis.

F bildet jede Basis von V auf ein Erzeugendensystem von $F(V)$ ab.
Also ist $\dim F(V) \leq \dim V$.

Sei $(u_1, \ldots, u_k)$ eine Basis von $\ker(F)$,
$(w_1, \ldots, w_r)$ eine Basis von $F(V)$ und $v_1, \ldots, v_r \in V$, so daß $F(v_i) = w_i$.

Behauptung:

$(u_1, \ldots, u_k, v_1, \ldots, v_r)$ ist eine Basis von V.

Aus der Behauptung folgt die Aussage des 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

$$0 \;=\; F(\sum_{i=1}^k \lambda_i u_i + \sum_{j=1}^r \mu_j v_j) \overset{u_i \in Ker(F)}{=} \sum_{j=1}^r \mu_j w_j$$

und somit $\mu_1 = \ldots = \mu_r = 0$, da die w_j linearen unabhängig sind.

Also $0 = \sum_{i=1}^k \lambda_i u_i$, woraus $\lambda_1 = \cdots = \lambda_k = 0$ folgt, wegen der linearen Unabhängigkeit der u_i.

2) Nun zeigen wir, daß $span\{u_1, \dots, u_k, v_1, \dots, v_r\} = V$.

Sei $v \in V$. Wir schreiben $F(v) = \sum_{j=1}^r \mu_j w_j$.

Dann gilt $v - \sum_{j=1}^r \mu_j v_j \in \ker(F) = span\{u_1, \dots, u_k\}$

und somit $v \in span\{u_1, \dots, u_k, v_1, \dots, v_r\}$. $\square$

Folgerung

Seien V, W *Vektorräume,* $\dim V < \infty$ *und* $F \in L(V, W)$*. Dann gilt:*

(i) F *ist genau dann surjektiv, wenn* $\mathrm{rg}(F) = \dim W$,

(ii) F *ist genau dann injektiv, wenn* $\mathrm{rg}(F) = \dim V$,

(iii) F *ist genau dann bijektiv, wenn* $\mathrm{rg}(F) = \dim V = \dim W$.

Beweis.

(i) F surjektiv $\iff F(V) = W \iff \dim F(V) = \dim W$.

(ii) F injektiv $\iff \ker(F) = \{0\} \overset{\text{(Dim.formel)}}{\iff} \mathrm{rg}(F) = \dim V$.

(iii) folgt aus (i-ii). $\square$

Folgerung

Seien V, W Vektorräume mit $\dim V = \dim W < \infty$.

Dann gilt: Eine lineare Abbildung $F : V \to W$ ist genau dann ein Isomorphismus, wenn eine der folgenden äquivalenten Bedingungen erfüllt ist:

(i) *F ist injektiv,*

(ii) *F ist surjektiv,*

(iii) *F ist bijektiv.*

Beweis. Wegen der Dimensionsformel gilt

$$dim(V) \;=\; rg(F) + dim(Ker(V)) \;=\; dim(W).$$

Daraus folgen die Äquivalenzen. $\square$

Lineare Gleichungssysteme

Definition

Ein System von Gleichungen der Form

$$\begin{aligned} a^1_1x^1 + a^1_2x^2 + \cdots + a^1_nx^n &= b^1 \\ &\vdots \\ a^m_1x^1 + a^m_2x^2 + \cdots + a^m_nx^n &= b^m \end{aligned}$$

wobei $a_{ij}, b_j \in \mathbb{K}$ *gegeben und* $x_1, x_2, \ldots, x_n \in \mathbb{K}$ *gesucht sind, bezeichnet man als Lineares Gleichungssystem.*

Ein solches lässt sich in der Form $Ax = b$ *schreiben, mit* $b = \begin{pmatrix} b^1 \\ \vdots \\ b^m \end{pmatrix} \in \mathbb{K}^m$

und $A = \left(a^i_j\right) \in L(\mathbb{K}^n, \mathbb{K}^m)$ *gegeben und* $x \in \mathbb{K}^n$ *gesucht.*

m *ist die Anzahl der Gleichungen und* n *die Anzahl der Unbekannten.*

Das System heißt homogen, falls $b = 0$ *und inhomogen falls* $b \neq 0$.

Satz

(i) *Die Lösungsmenge eines homogenen linearen Gleichungssytems* $Ax = 0$, $A \in L(\mathbb{K}^n, \mathbb{K}^m)$, *ist genau der Unterraum* $U = \ker(A) \subset \mathbb{K}^n$.

(ii) *Die Dimension von* U *beträgt* $n - r$, *wobei* $r = \mathrm{rg}(A)$.

Beweis. (i) folgt aus der Definition und (ii) aus der Dimensionsformel. □

Satz

Sei $A \in L(\mathbb{K}^n, \mathbb{K}^m)$ *eine Matrix und* B *die aus* A *mittels Gaußschem Algorithmus hervorgegangene Matrix in Zeilenstufenform.*
Dann gilt $Ker(A) = Ker(B)$, *d.h. die linearen homogenen Gleichungssysteme* $Ax = 0$ *und* $Bx = 0$ *haben denselben Lösungsraum.*

Beweis. Die Spalten von A sind gegeben durch die Bilder der kanonischen Basis, $A(e_i)$. Der Beweis beruht nun darauf, daß wir die Zeilen von A als Elemente von $(\mathbb{K}^n)^* = L(\mathbb{K}^n, \mathbb{K})$ interpretieren.
Für $i = 1, \ldots m$ definierien wir $\alpha^i \in L(\mathbb{K}^n, \mathbb{K})$ mittels

$$\alpha^i(e_j) := a^i_j, \quad \text{d.h.} \qquad A(x) = \begin{pmatrix} \alpha_1(x) \\ \vdots \\ \alpha_m(x) \end{pmatrix}.$$

Die a^i_j sind also die Komponenten der α^i in der kanonischen dualen Basis von $(\mathbb{K}^n)^*$. Also ist $V^* := span(\alpha_1, \ldots, \alpha_m)$ der durch die Zeilen von A aufgespannte Unterraum. Mittels des Gaußschen Algorithmus erhält man die Matrix B in Zeilenstufenform, die eine Basis $(\beta_1, \ldots, \beta_k)$ von V^* liefert. Also gilt

$$ker(A) = \{x \in \mathbb{K}^n \mid L(x) = 0 \; \forall \; L \in V^*\} = ker(B)$$

D.h. beide Lösungsräume sind gleich. $\square$

Lösen von homogenen linearen Gleichungssystemen mittels Gauß-Algorithmus

- Wir wollen das Gleichungssystem $Ax = 0$ lösen. Dazu bringen wir A auf Zeilenstufenform und erhalten B vom Zeilenrang k.
- D.h. wir erhalten ein Gleichungssystem der Form

$$\begin{array}{ccccc} x^{j_1} & + & \sum\limits_{l=j_1+1}^{n} b_l^1 x^l & = & 0 \\ & & & \vdots & \\ x^{j_i} & + & \sum\limits_{l=j_i+1}^{n} b_l^i x^l & = & 0 \\ & & & \vdots & \\ x^{j_k} & + & \sum\limits_{l=j_k+1}^{n} b_l^k x^l & = & 0 \end{array}$$

- Dieses löst man, indem man schrittweise die Unbekannten x^j mit $j \notin \{j_1, \ldots, j_k\}$ frei wählt und die Lösungen $x^{j_1}, \ldots, x^{j_k}$ bestimmt.
- D.h. der Raum der Lösungen ist $(n-k)$-dimensional. Damit gilt auch $k = \mathit{Zeilenrang} = \mathit{Spaltenrang}$ von A.

Satz

Wir betrachten ein inhomogenes lineares Gleichungssytem

$$Ax = b, \qquad \text{mit } A \in L(\mathbb{K}^n, \mathbb{K}^m). \tag{12}$$

(i) *Falls* $b \notin A(\mathbb{K}^n)$*, so hat* (12) *keine Lösung.*

(ii) *Falls* $b \in A(\mathbb{K}^n)$*, so ist die Lösungsmenge* U' *von* (12) *von der Form*

$$U' = x_0 + U = \{x_0 + u | u \in U\},$$

wobei $x_0 \in A^{-1}(b)$ *eine spezielle Lösung von* (12) *ist und* $U \subset \mathbb{K}^n$ *der Lösungsraum des zugehörigen homogenen Systems* $Ax = 0$ *ist.*

Beweis. (i) ist klar. Im Fall (ii) ist klar, daß $x_0 + U \subset U'$. Wir zeigen $U' \subset x_0 + U$. Sei also v eine Lösung von (*). Dann gilt $A(v - x_0) = b - b = 0$ und somit $v = x_0 + (v - x_0) \in x_0 + U$. □

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 $Ax = b$, $A \in L(\mathbb{K}^n, \mathbb{K}^m)$, ein affiner Unterraum von $\mathbb{K}^n$.

Der Lösungsraum ist genau dann ein Untervektorraum von $\mathbb{K}^n$, wenn das Gleichungssytem 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

$$
\begin{array}{ccccccc}
x^{j_1} & & & + & \displaystyle\sum_{l=j_1+1}^{n} b^1_l x^l & = & c^1 \\
 & & & & & \vdots & \\
 & x^{j_i} & & + & \displaystyle\sum_{l=j_i+1}^{n} b^i_l x^l & = & c^i \\
 & & & & & \vdots & \\
 & & x^{j_k} & + & \displaystyle\sum_{l=j_k+1}^{n} b^k_l x^l & = & c^k \\
 & & & & 0 & = & c^{k+1}.
\end{array}
$$

- $Ax = b \iff (A|b)\begin{pmatrix} x \\ -1 \end{pmatrix} = 0 \iff (B|c)\begin{pmatrix} x \\ -1 \end{pmatrix} = 0$
  $\iff Bx = c \iff 0 = c^{k+1}$.
- Die restlichen k Gleichungen löst man, indem man wieder schrittweise 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.

Zahlenbeispiel: Wir wollen das inhomogene lineare Gleichungssystem

$$\begin{aligned} x^1 + 2x^2 + 3x^3 &= 0 \\ 4x^1 + 5x^2 + 6x^3 &= 3 \\ 7x^1 + 8x^2 + 9x^3 &= 6 \end{aligned}$$

lösen. D.h. $Ax = b$ mit $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$ und $b = \begin{pmatrix} 0 \\ 3 \\ 6 \end{pmatrix}$. Also

$$(A|b) = \begin{pmatrix} 1 & 2 & 3 & 0 \\ 4 & 5 & 6 & 3 \\ 7 & 8 & 9 & 6 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & 3 & 0 \\ 0 & -3 & -6 & 3 \\ 0 & -6 & -12 & 6 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & 3 & 0 \\ 0 & -3 & -6 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\to \begin{pmatrix} 1 & 2 & 3 & 0 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix} = (B|c).$$

Da $c^3 = 0$ ist $Ax = b$ lösbar. Das zugehörige Gleichungssystem ist:

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 0 \\ x_2 + 2x_3 &= -1. \end{aligned}$$

D.h. $x_2 = -2x_3 - 1$, $x_1 = -2x_2 - 3x_3 = -2(-2x_3 - 1) - 3x_3 = x_3 + 2$. Die allgemeine Lösung ist also $x_1 = \lambda + 2$, $x_2 = -2\lambda - 1$, $x_3 = \lambda \in \mathbb{K}$ beliebig.

Direktes Produkt von Vektorräumen

Definition

U, V seien $\mathbb{K}$-Vektorräume.

Das direkte Produkt von U und V ist das kartesische Produkt $U \times V = \{(u, v) | u \in U, v \in V\}$ ausgestattet mit der durch die komponentenweise definierten Addition und skalaren Multiplikation gegebenen Vektorraumstruktur, d.h.

$$\begin{aligned} \lambda(u, v) &:= (\lambda u, \lambda v) \\ (u, v) + (u', v') &:= (u + u', v + v') \end{aligned}$$

Es gilt:

$$\dim(U \times V) = \dim U + \dim V,$$

denn: Ist $(u_1, \dots, u_m)$ eine Basis von U und $(v_1, \dots, v_n)$ eine Basis von V, dann ist $((u_i, 0), (0, v_j) \mid i = 1, \dots m, j = 1, \dots n)$ eine Basis von $U \times V$.

Summe von Unterräumen

Definition

W, W' seien Unterräume eines Vektorraums V.

Die Summe

$$W + W' := \{w + w' | w \in W, w' \in W'\}$$

ist der kleinste Unterraum von V, der W und W' enthält.

Satz

Falls W und W' endlichdimensional sind, so gilt folgende Dimensionsformel

$$\dim(W + W') = \dim W + \dim W' - \dim(W \cap W').$$

Beweis der Dimensionsformel für die Summe.
Wir betrachten die Abbildung $F : W \times W' \to W + W'$,

$$(w, w') \mapsto F(w, w') := w - w'.$$

Offenbar ist F surjektiv und

$$\ker(F) = \{(w, w) | w \in W \cap W'\} \cong W \cap W'.$$

Daher folgt aus der Dimensionsformel für lineare Abbildungen

$$\begin{aligned} \operatorname{rg}(F) &= \dim(W + W') = \dim(W \times W') - \dim(W \cap W') \\ &= \dim W + \dim W' - \dim(W \cap W'). \end{aligned}$$

$\square$

Direkte Summe von Unterräumen

Definition (Direkte Summe von Unterräumen)

W, W' seien Unterräume eines Vektorraums V.
Falls $W \cap W' = \{0\}$ bezeichnet man die Summe $W + W'$ als direkte Summe und schreibt

$$W \oplus W'$$

Die Unterräume W, W' heißen komplementär, wenn $V = W \oplus W'$.
Wir sagen dann, dass W' ein Komplement zu W (in V) ist.

Beispiel

- U, V seien $\mathbb{K}$-Vektorräume und $U \times V$ ihr direktes Produkt. Die Unterräume

$$\begin{aligned} U \times \{0\} &= \{(u,0) \in U \times V | u \in U\} \subset U \times V, \\ \{0\} \times V &= \{(0,v) \in U \times V | v \in V\} \subset U \times V \end{aligned}$$

  sind komplementär: $(U \times \{0\}) \cap (\{0\} \times V = \{0\})$ und

$$U \times V = (U \times \{0\}) \oplus (\{0\} \times V). \tag{13}$$

  Durch die injektiven linearen Abbildungen $u \mapsto (u,0)$, $U \to U \times V$, und $v \mapsto (0,v)$, $V \to U \times V$, können wir U und V mit den Unterräumen $U \times \{0\}$ bzw. $\{0\} \times V$ von $U \times V$ identifizieren und (13) auch in der Form $U \times V = U \oplus V$ schreiben.
- Sei $V = \mathbb{R}^3$, $W = span(e_1, e_2)$ und $W' = span(e_1 + e_2, e_3)$. Die Summe $W + W' = V$ ist nicht direkt, denn $W \cap W' = \mathbb{R} \cdot (e_1 + e_2)$.

Satz

W, W' seien Unterräume eines Vektorraums V. Es gilt:

$V = W \oplus W' \Longleftrightarrow$ Jeder Vektor $v \in V$ hat eine eindeutige Darstellung

$$v = w + w', \quad w \in W, \quad w' \in W'. \tag{14}$$

Beweis.

'$\Rightarrow$' Aus $V = W + W'$ folgt, dass jeder Vektor $v \in V$ eine Darst. (14) hat. Sei $v = w_1 + w_1'$ eine weitere solche Darst.
Es folgt $0 = v - v = (w - w_1) + (w' - w_1')$ und daraus $W \ni w - w_1 = -(w' - w_1') \in W'$.
Wg. $W \cap W' = \{0\}$ folgt nun $w - w_1 = w' - w_1' = 0$ und somit die Eindeutigkeit der Darstellung (14).

'$\Leftarrow$' Wenn umgekehrt jeder Vektor $v \in V$ eine eindeutige Darstellung (14) hat, so gilt $V = W + W'$ und aus $0 = w + (-w)$, $w \in W \cap W'$, folgt wg. der Eindeutigkeit von (14) $w = 0$, d.h. $W \cap W' = \{0\}$.

□

Direkte Summe von Unterräumen

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 Basis von 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.
   $W' := \mathrm{span}\{w_1', \dots, w_l'\}$ ist dann ein Komplement zu W.

b) Die Äquivalenz der Eigenschaften (i-iii) folgt aus der Dimensionsformel.

Gruppen

Definition (Gruppe)

Eine Gruppe $(G, \cdot)$ *ist eine Menge* G *zusammen mit einer assoziativen Verknüpfung* $\cdot : G \times G \to G$, *so daß*

(i) *es gibt (genau) ein* neutrales Element $e \in G$ *mit*

$$e \cdot a = a \cdot e = a \quad \text{für alle} \quad a \in G \quad \text{und}$$

(ii) *zu jedem* $a \in G$ *gibt es (genau) ein* Inverses a^{-1} *mit* $a \cdot a^{-1} = a^{-1} \cdot a = e$.

Eine Gruppe G *heißt* kommutativ *falls* $a \cdot b = b \cdot a$ *für alle* $a, b \in G$.

Bemerkung

Das neutrale Element und das Inverse a^{-1} zu a sind eindeutig bestimmt. Dies muß man nicht in der Definition fordern, es kann aus der Existenz gefolgert werden (vgl. Fischer, S. 44).

Beispiele von Gruppen

(i) Die Bijektionen $\varphi : X \to X$ einer Menge X in sich bilden eine Gruppe, die mit $\mathrm{Bij}(X)$ bezeichnet wird. Die Vernüpfung ist die Verkettung, das neutrale Element ist Id_X und das Inverse einer Bijektion ist ihre Umkehrabbildung.

(ii) Im Fall einer endlichen Menge X nennt man die Bijektionen $\sigma \in \mathrm{Bij}(X)$ auch Permutationen von X. Die Permutationsgruppe $\mathrm{Bij}(X)$ ist dann eine endliche Gruppe mit $n!$ Elementen, wobei $n = \mathrm{card}(X)$.

(iii) Die Permutationsgruppe $S_n := \mathrm{Bij}(X_n)$ der n-elementigen Menge $X_n = \{1, 2, \ldots, n\}$ heißt die symmetrische Gruppe.

(iv) Sei V ein VR. Die Isomorphismen $V \to V$ heißen auch Automorphismen des Vektorraums V und bilden die sogenannte Automorphismengruppe $\mathrm{Aut}(V)$ von V.

weitere Beispiele von Gruppen

(v) Im Fall eines endlichdimensionalen Vektorraums V spricht man auch von der allgemeinen linearen Gruppe und schreibt dafür $\mathrm{GL}(V) := \mathrm{Aut}(V)$.

Man setzt $\mathrm{GL}(n, \mathbb{K}) := \mathrm{GL}(\mathbb{K}^n)$. Die Elemente von $\mathrm{GL}(n, \mathbb{K})$ heißen invertierbare Matrizen.

Es gilt

$$\begin{aligned} \mathrm{GL}(n, \mathbb{K}) &= \{A \in \mathrm{Mat}(n, \mathbb{K}) | \mathrm{rg}(A) = n\} \\ &= \{A \in \mathrm{Mat}(n, \mathbb{K}) | \ker(A) = 0\}, \end{aligned}$$

wobei wir der Übersichtlichkeit halber ab jetzt den Nullvektorraum $\{0\}$ einfach mit 0 bezeichnen.

(vi) $\mathrm{GL}(1, \mathbb{K}) = (\mathbb{K}^* := \mathbb{K} \setminus \{0\}, \cdot)$ ist eine kommutative Gruppe.

(vii) Aufgrund der Definition, ist für jeden Vektorraum V durch $(V, +)$ eine kommutative Gruppe gegeben.

noch mehr Beispiele von Gruppen

(viii) 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.

(ix) Die Kreislinie $S^1 := \{z \in \mathbb{C} | |z| = 1\}$ ist mit der Multiplikation komplexer Zahlen eine kommutative Gruppe.

Definition (Gruppenmorphismen)

Eine Abbildung $\varphi : G \to H$ zwischen Gruppen G und H heißt ein Gruppenmorphismus, falls

$$\varphi(a \cdot b) = \varphi(a) \cdot \varphi(b)$$

für alle $a, b \in G$.

Ein bijektiver Gruppenmorphismus 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 $\varphi : G \to H$ gibt.

Eine Untergruppe einer Gruppe G ist eine Teilmenge $H \subset G$, die mit $a, b \in H$ auch $a \cdot b$ und a^{-1} enthält.

Übungsaufgaben/Beispiele I

- Jeder Gruppenmorphismus $\varphi : G \to H$ bildet das neutrale Element in G auf das neutrale Element in H ab und erfüllt

$$\varphi(a^{-1}) = \varphi(a)^{-1}$$

  für alle $a \in G$.
- Das Inverse eines Isomorphismus $\varphi : G \to H$ ist wieder ein Isomorphismus.
- Jede Untergruppe $H \subset G$ einer Gruppe G ist mit der induzierten Verknüpfung wieder eine Gruppe.
- Das Bild $\varphi(K) \subset H$ einer Untergruppe $K \subset G$ unter einem Gruppenmorphismus $\varphi : G \to H$ ist wieder eine Untergruppe.
- Das Urbild $\varphi^{-1}(K) \subset G$ einer Untergruppe $K \subset H$ unter einem Gruppenmorphismus $\varphi : G \to H$ ist wieder eine Untergruppe.

Übungsaufgaben/Beispiele II

- Die Automorphismen einer Gruppe G bilden eine Untergruppe $\mathrm{Aut}(G) \subset \mathrm{Bij}(G)$.
- Die Gruppen $(\mathbb{R}, +)$ und $(\mathbb{R}_+, \cdot)$ sind isomorph. Die Exponentialfunktion $\exp : \mathbb{R} \to \mathbb{R}_+$ ist ein Isomorphismus, ebenso wie ihre Umkehrfunktion $\ln : \mathbb{R}_+ \to \mathbb{R}$.
- Die Abbildung $t \mapsto \exp(it)$ definiert einen surjektiven aber nicht injektiven Gruppenmorphismus $\varphi : \mathbb{R} \to S^1$.

Die symmetrische Gruppe

Jede Permutation $\sigma \in S_n$ wird durch ein Diagramm

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix}$$

beschrieben. 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:

$$\begin{aligned} \tau_{12} &= \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \\ \tau_{13} &= \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \\ \tau_{23} &= \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \end{aligned}$$

Ist τ eine Transpositon, so gilt offensichtlich $\tau^{-1} = \tau$

Satz

Die symmetrische Gruppe S_n, $n \in \mathbb{N}$, ist nur für $n \leq 2$ kommutativ.

Beweis.

Es ist klar, dass S_1 und S_2 kommutativ sind. S_3 ist nicht kommutativ, denn z.B. ist $\tau_{12} \circ \tau_{13} \neq \tau_{13} \circ \tau_{12}$. Letzteres folgt bereits aus:

$$\tau_{12} \circ \tau_{13}(1) = \tau_{12}(3) = 3 \neq \tau_{13} \circ \tau_{12}(1) = \tau_{13}(2) = 2$$

Die Abbildung

$$\sigma \mapsto \begin{pmatrix} 1 & 2 & 3 & 4 & \cdots & n \\ \sigma(1) & \sigma(2) & \sigma(3) & 4 & \cdots & n \end{pmatrix}$$

ist ein injektiver Gruppenmorphismus $S_3 \to S_n$ $(n \geq 3)$. Das Bild ist eine nicht kommutative Untergruppe von S_n. Insbesondere ist 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 $\sigma = \tau_1 \circ \cdots \circ \tau_k$.

Beweis.
Für jede Transposition τ gilt $\tau^{-1} = \tau$ und somit $\mathrm{Id} = \tau \circ \tau$.
Sei $\mathrm{Id} \neq \sigma \in S_n$. Dann existiert ein $1 \leq i_1 \leq n$ mit

- $\sigma(i) = i$ für alle $1 \leq i \leq i_1 - 1$ und
- $\sigma(i_1) > i_1$.

Wir setzen $\tau_1 := \tau_{i_1 \sigma(i_1)}$ und $\sigma_1 := \tau_1 \circ \sigma$. Dann gilt $\sigma_1(i) = i$ für alle $1 \leq i \leq i_1$.
Falls $\sigma_1 \neq \mathrm{Id}$, so gibt es wieder ein i_2, $i_1 < i_2 \leq n$, so daß $\sigma_1(i) = i$ für alle $1 \leq i \leq i_2 - 1$ und $\sigma_1(i_2) > i_2$. Wir setzen $\tau_2 := \tau_{i_2 \sigma(i+2)}$ und $\sigma_2 := \tau_2 \circ \sigma_2$.
Durch Fortsetzen dieses Iterationsverfahren erhalten wir nach endlich vielen (genauer: nach $k \leq n-1$) Schritten $\sigma_k = \tau_k \circ \cdots \circ \tau_1 \circ \sigma = \mathrm{Id}$ und somit $\sigma = \tau_1 \circ \cdots \circ \tau_k$. $\square$

Definition (Vorzeichen einer Permutation)

Das Vorzeichen $\varepsilon(\sigma)$ *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 $\sigma(i) > \sigma(j)$. Die Permutation σ ist also genau dann gerade (bzw. ungerade), wenn diese Anzahl gerade (bzw. ungerade) ist.

Satz

$\varepsilon : S_n \to \{-1, 1\}$ ist ein Gruppenmorphismus 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

$$\begin{aligned}\prod_{i<j} f_{ij} &= \prod_{\substack{i<j \\ \tau(i)<\tau(j)}} f_{ij} \prod_{\substack{i<j \\ \tau(i)>\tau(j)}} f_{ij} \\ &\overset{(*)}{=} \prod_{\substack{i<j \\ \tau(i)<\tau(j)}} f_{ij} \prod_{\substack{i>j \\ \tau(i)<\tau(j)}} f_{ij} = \prod_{\tau(i)<\tau(j)} f_{ij} = \varepsilon(\sigma).\end{aligned}$$

Damit folgt $\prod_{i<j} f_{ij} = \varepsilon(\sigma)$ und die Behauptung. □

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
Die Permutation

$$\zeta := (23 \cdots n1) := \begin{pmatrix} 1 & 2 & \cdots & n-1 & n \\ 2 & 3 & \cdots & n & 1 \end{pmatrix}$$

schreibt sich als $\zeta = \tau_{12} \circ \tau_{23} \circ \cdots \circ \tau_{n-1n}$.
Also ist $\varepsilon(\zeta) = (-1)^{n-1}$.
ζ erzeugt eine n-elementige Untergruppe

$$\langle\zeta\rangle := \{\zeta, \zeta^2, \ldots, \zeta^n = \mathrm{Id}\} \subset S_n.$$

Die ζ^k, $k = 1, 2, \ldots, n$, heißen zyklische Permutationen.
Für ungerades n sind also alle zyklischen Permutationen gerade.
(ÜA: Für $n = 3$ sind alle ungeraden Permutationen Transpositionen und alle geraden Permutationen zyklisch.)

Determinante

Definition (Determinante)

Eine Abbildung $\det : \mathrm{Mat}(n, \mathbb{K}) \to \mathbb{K}$ *heißt Determinante, wenn sie folgende Eigenschaften hat:*

D1) $\det$ *ist linear in jeder Spalte, d.h.*

$$\begin{aligned} & \det(a_1 \cdots a_{i-1}\ \lambda a_i + \mu b_i\ a_{i+1} \cdots a_n) \\ = \quad & \lambda \det(a_1 \cdots\ a_i\ \cdots a_n) + \mu \det(a_1 \cdots a_{i-1}\ b_i\ a_{i+1} \cdots a_n) \end{aligned}$$

für alle Spaltenvektoren $a_1, \ldots, a_n, b_i \in \mathbb{K}^n$ *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 $\det : \mathrm{Mat}(n, \mathbb{K}) \to \mathbb{K}$ mit den Eigenschaften D1-D3 gibt.

Zunächst zeigen wir, dass D1-D3 eine Reihe weiterer Rechenregeln nach sich ziehen.

Satz

Sei $A = (a_1 \cdots a_n) \in \mathrm{Mat}(n, \mathbb{K})$*. 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 $i < j$

$$0 = \det(\cdots a_i + a_j \cdots a_i + a_j \cdots) = \det(\cdots a_i \cdots a_j \cdots) + \det(\cdots a_j \cdots a_i \cdots).$$

Hierbei ist die i-te und j-te Spalte angegeben. Die Auslassungspunkte stehen für $a_1 \cdots a_{i-1}$, $a_{i+1} \cdots a_{j-1}$ 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 $A = (a_1 \cdots a_n) \in \mathrm{Mat}(n, \mathbb{K})$ und $\lambda \in \mathbb{K}$. Aus D1-D2 folgt, dass Addition des λ-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:

$$\begin{aligned} & \det(a_1 \cdots a_{i-1}\; a_i + \lambda a_j\; a_{i+1} \cdots a_n) \\ = \; & \det A + \lambda \det(\cdots a_j \cdots a_j \cdots) \overset{D2}{=} \det A. \end{aligned}$$

□

Folgerung

Sei $A = (a_1 \cdots a_n) \in \mathrm{Mat}(n, \mathbb{K})$*. Aus D1-D2 folgt: Falls* $\mathrm{rg}(A) \neq n$*, so ist* $\det(A) = 0$.

Beweis. $\mathrm{rg}(A) \neq n$ bedeutet, daß die Dimension des Bildes von A kleiner als n ist. D.h. aber, daß sich eine der Spalten a_i als Linearkombination der anderen schreiben läßt:

$$a_i = \sum_{j \neq i} \lambda_j a_j.$$

Daraus ergibt sich

$$\det A = \det(a_1 \cdots a_{i-1}\ \underbrace{a_i - \sum_{j \neq i} \lambda_j a_j}_{=0}\ a_{i+1} \cdots a_n) = 0.$$

□

Berechnung der Determinante mittels Spaltenumformungen

Satz

Jede invertierbare Matrix $A \in \mathrm{GL}(n, \mathbb{K})$ *läßt sich durch wiederholtes Anwenden folgender zwei Spaltenumformungen in eine Diagonalmatrix* $A' = \mathrm{diag}(\lambda_1, \ldots, \lambda_n) := (\lambda_1 e_1 \cdots \lambda_n e_n)$ *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:
$\det A = \det A' \overset{D1}{=} \lambda_1 \cdots \lambda_n \det(\mathbf{1}_n) \overset{D3}{=} \lambda_1 \cdots \lambda_n$. □

Beweis des Satzes: Da A invertierbar ist, ist $rg(A) = n$. 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 Umformumgen S1-S2 den Rang nicht ändern.

Also ist das Produkt $\lambda_1 \cdots \lambda_n \neq 0$.

Durch Umformungen S2 können wir deswegen alle Matrixeinträge unterhalb der Diagonalen eliminieren und erhalten die gewünsche Diagonalgestalt $A' = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$. □

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 $\det : \mathrm{Mat}(n, \mathbb{K}) \to \mathbb{K}$ *mit den Eigenschaften D1-D3.*

Unter der Voraussetzung, dass eine solche Abbildung existiert, gilt:

$A \in \mathrm{Mat}(n, \mathbb{K})$ *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 $\det : \mathrm{Mat}(n, \mathbb{K}) \to \mathbb{K}$ *mit den Eigenschaften D1-D3.*

Die Determinante einer $(n \times n)$*-Matrix* $A = (a^i_j)_{i,j}$ *ist durch folgende Formel gegeben:*

$$\det A = Det\, A := \sum_{\sigma \in S_n} \varepsilon(\sigma) a_1^{\sigma(1)} \cdots a_n^{\sigma(n)}.$$

Beweis. Da wir bereits wissen, dass es *höchstens* eine Abbildung mit den Eigenschaften D1-D3 gibt, genügt es D1-D3 für die Abbildung *Det* nachzuweisen.

Beweis der expliziten Formel für die Determinante:

D1:

$$
\begin{aligned}
& \sum_{\sigma\in S_n} \varepsilon(\sigma) a_1^{\sigma(1)} \cdots (\lambda a_i^{\sigma(i)} + \mu b_i^{\sigma(i)}) \cdots a_n^{\sigma(n)} \\
= \quad & \lambda Det\, A + \mu \sum_{\sigma\in S_n} \varepsilon(\sigma) a_1^{\sigma(1)} \cdots b_i^{\sigma(i)} \cdots a_n^{\sigma(n)}.
\end{aligned}
$$

Weiter Beweis der expliziten Formel für die Determinante:

D2: Es genügt zu zeigen, dass beim Vertauschen der Spalten

$$\underbrace{Det\left(a_{\tau(1)} \cdots a_{\tau(n)}\right)}_{A_\tau :=} = -DetA$$

für jede Transposition τ. Aus $A = A_\tau$ folgt dann $Det\,A = Det\,A_\tau = -Det\,A$ und somit $Det\,A = 0$. Mit der Substitution $\sigma' := \sigma\tau = \sigma \circ \tau$ haben wir $\sigma = \sigma'\tau$, $\varepsilon(\sigma') = -\varepsilon(\sigma)$ und somit

$$\begin{aligned}
Det(A_\tau) &= \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{\tau(1)}^{\sigma(1)} \cdots a_{\tau(n)}^{\sigma(n)} \\
&= -\sum_{\sigma' \in S_n} \varepsilon(\sigma') a_{\tau(1)}^{\sigma'(\tau(1))} \cdots a_{\tau(n)}^{\sigma'(\tau(n))} \\
&= -\sum_{\sigma' \in S_n} \varepsilon(\sigma') a_1^{\sigma'(1)} \cdots a_n^{\sigma'(n)} = -Det\,A
\end{aligned}$$

Weiter im Beweis der expliziten Formel für die Determinante:

D3: Sei nun $A = \mathbf{1}_n$, d.h.

$$a^i_j = \delta^i_j := \begin{cases} 1, & \text{falls} \quad i = j \\ 0 & \text{sonst} \end{cases} \qquad (\delta^i_j \quad \text{heißt Kroneckersymbol}).$$

Daraus folgt

$$\varepsilon(\sigma)\delta^{\sigma(1)}_1 \cdots \delta^{\sigma(n)}_n = \begin{cases} 1, & \text{falls} \quad \sigma = \mathrm{Id} \\ 0 & \text{sonst} \end{cases}$$

und somit

$$Det(A) = Det(\mathbf{1}_n) = \sum_{\sigma \in S_n} \varepsilon(\sigma)\delta^{\sigma(1)}_1 \cdots \delta^{\sigma(n)}_n = 1.$$

□

Berechnung von Determinanten

Die Berechnung der Determinante einer Matrix $A \in \mathrm{Mat}(n, \mathbb{K})$ mittels der Formel

$$\det A = \sum_{\sigma \in S_n} \varepsilon(\sigma) a_1^{\sigma(1)} \cdots a_n^{\sigma(n)}$$

ist wegen des schnellen Wachstums von $\mathrm{card}(S_n) = n!$ für große n sehr aufwendig.

Beachte, daß $n!$ schneller wächst als jede Exponentialfunktion a^n, $a > 1$.

Wir werden später (im nächsten Semester) weitere Formeln zur Berechnung der Determinante kennenlernen.

Formeln für Determinanten

Für $n \leq 3$ erhalten wir:

- $n = 1$: $a = a_1^1 \in \mathrm{Mat}(1, \mathbb{K}) = \mathbb{K}, \quad \det a = a_1^1$,
- $n = 2$:
$$\det \begin{pmatrix} a_1^1 & a_2^1 \\ a_1^2 & a_2^2 \end{pmatrix} = a_1^1 a_2^2 - a_2^1 a_1^2,$$
- $n = 3$:
$$\begin{aligned} \det \begin{pmatrix} a_1^1 & a_2^1 & a_3^1 \\ a_1^2 & a_2^2 & a_3^2 \\ a_1^3 & a_2^3 & a_3^3 \end{pmatrix} &= a_1^1 a_2^2 a_3^3 + a_2^1 a_3^2 a_1^3 + a_3^1 a_1^2 a_2^3 \\ &\quad - a_3^1 a_2^2 a_1^3 - a_2^1 a_1^2 a_3^3 - a_1^1 a_3^2 a_2^3, \end{aligned}$$

  dabei entsprechen die ersten drei Terme den zyklischen Permutationen (123), (231) und (312), die letzten drei den Transpositionen τ_{13}, τ_{12} und τ_{23}. Das erste Tripel beginnt mit dem Produkt der Hauptdiagonalelemente, das zweite mit dem Produkt der Nebendiagonalelemente.

Satz (Determinantenmultiplikationssatz)

Aus D1-D3 folgt für alle $A, B \in \mathrm{Mat}(n, \mathbb{K})$*:*

$$\det(AB) = \det(A)\det(B).$$

Beweis. Falls $\mathrm{rg}(A) < n$, so ist $\mathrm{rg}(AB) < n$ und somit $\det(AB) = 0 = \det(A)\det(B) = 0 \cdot \det(B)$.

Falls $\mathrm{rg}(B) < n$, so ist $\ker(B) \neq 0$ und somit $\ker(AB) \neq 0$. Also ebenfalls $\det(AB) = 0 = \det(A)\det(B) = \det(A) \cdot 0$. Seien also A, B invertierbar.

Wir benutzen folgendes Lemma:

Lemma

Jede invertierbare Matrix $A \in \mathrm{GL}(n, \mathbb{K})$ schreibt sich als Produkt

$$A = DE_1 \cdots E_k,$$

wobei $D \in \mathrm{GL}(n, \mathbb{K})$ eine Diagonalmatrix ist und die $E_1, \ldots, E_k$ von folgendem Typ sind:

$$\begin{array}{rrcl} (1) & E(i,j) & := & (e_{\tau(1)} \cdots e_{\tau(n)}), \\ (2) & E(i,j,\lambda) & := & (e_1 \cdots e_{i-1}\; e_i + \lambda e_j\; e_{i+1} \cdots e_n). \end{array}$$

Hierbei ist $\lambda \in \mathbb{K}$, $i \neq j$ und $\tau \in S_n$ die Vertauschung von i und j. Matrizen der Form (1) oder (2) heißen Elementarmatrizen.

Beweis des Lemmas. Wir bemerken zunächst, dass $A \mapsto AE(i,j)$ die i-te und j-te Spalte von A vertauscht und

$A \mapsto AE(i,j,\lambda)$ das λ-fache der j-ten Spalte von A zur i-ten Spalte von A addiert .

Desweiteren gilt

$$\begin{aligned} E(i,j)^{-1} &= E(i,j), \\ E(i,j,\lambda)^{-1} &= E(i,j,-\lambda). \end{aligned}$$

Wegen des Gaußalgorithmus gibt es also Matrizen $F_1,\dots,F_k$ vom Typ (1-2), so dass $D = AF_1 \cdots F_k$ eine Diagonalmatrix ist. Daraus folgt mit $E_1 := F_k^{-1},\dots,E_k := F_1^{-1}$:

$$A = DF_k^{-1}\cdots F_1^{-1} = DE_1 \cdots E_k.$$

□

Beweis des Determinantenmultiplikationssatzes
Wegen des vorherigen Lemmas genügt es $\det(AB) = \det(A)\det(B)$ für den Fall zu beweisen, dass B eine Diagonalmatrix oder vom Typ (1-2) ist.

Für eine Diagonalmatrix $B = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ gilt $\det B = \lambda_1 \cdots \lambda_n$ und somit

$$\det(AB) = \det(\lambda_1 a_1 \cdots \lambda_n a_n) \overset{D1}{=} \lambda_1 \cdots \lambda_n \det A = \det A \det B.$$

Für $B = E(i,j)$ gilt $\det B = \det(e_{\tau(1)} \cdots e_{\tau(n)}) = \varepsilon(\tau)\det(\mathbf{1}_n) = -1$ und somit

$$\begin{aligned}\det(AB) &= \det(a_{\tau(1)} \cdots a_{\tau(n)}) = \varepsilon(\tau)\det A = -\det A \\ &= \det A \det B.\end{aligned}$$

Für $B = E(i,j,\lambda)$ gilt $\det B = 1$ und somit
$\det(AB) = \det(a_1 \cdots a_{i-1}\ a_i + \lambda a_j\ a_{i+1} \cdots a_n) = \det A = \det A \det B.$ $\square$

Folgerung

(i)
$$\det : \mathrm{GL}(n, \mathbb{K}) \longrightarrow (\mathbb{K}^*, \cdot)$$
ist ein Gruppenmorphismus.

(ii) *Insbesondere gilt für alle* $A \in \mathrm{GL}(n, \mathbb{K})$*:*
$$\det(A^{-1}) = (\det A)^{-1}$$

(iii)
$$\mathrm{SL}(n, \mathbb{K}) := \{A \in \mathrm{GL}(n, \mathbb{K}) | \det A = 1\} \subset \mathrm{GL}(n, \mathbb{K}).$$
ist eine Untergruppe (die sogenannte spezielle lineare Gruppe).

Beweis. (i-ii) sind klar. (iii) folgt aus $\mathrm{SL}(n, \mathbb{K}) = \det^{-1}(1)$. □

Achtung: Für $n \geq 2$ gilt nicht $det(A + B) = det(A) + det(B)$!
(z.B. $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ und $A = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$.)

Definition (Determinante eines Endomorphismus)

Sei V ein $\mathbb{K}$-Vektorraum der Dimension n, $F \in \mathrm{End}(V)$ ein Endorphismus und $A = M_B(F) \in \mathrm{Mat}(n, \mathbb{K})$ seine darstellende Matrix bez. einer 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 dann gilt:

$$A = \phi_B^{-1} \circ F \circ \phi_B, \quad A' := M_{B'}(F) = \phi_{B'}^{-1} \circ F \circ \phi_{B'}$$

und somit

$$A' = \phi_{B'}^{-1} \circ \phi_B \circ \underbrace{\phi_B^{-1} \circ F \circ \phi_B}_{=A} \circ \phi_B^{-1} \circ \phi_{B'}$$

d.h. $A' = SAS^{-1}$ mit $S = \phi_{B'}^{-1} \circ \phi_B$.
Also in der Tat, $\det A' = \det(SAS^{-1}) = \det S \det A (\det S)^{-1} = \det A$. □

Orientierung

Definition

Sei V ein endlichdimensionaler reeller VR. Zwei Basen $B = (b_1, \ldots, b_n)$ und $B' = (b'_1, \ldots, b'_n)$ heißen gleich orientiert, wenn der Basiswechsel $F = \phi_{B'} \circ \phi_B^{-1} : V \to V$ 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 $M_B(F) = \phi_B^{-1} \circ (\phi_{B'} \circ \phi_B^{-1}) \circ \phi_B = \phi_B^{-1} \circ \phi_{B'} : \mathbb{R}^n \to \mathbb{R}^n$.
  Daher gilt

$$\det(\phi_{B'} \circ \phi_B^{-1}) = \det F = \det M_B(F) = \det(\phi_B^{-1} \circ \phi_{B'}).$$

Äquivalenzrelation

Definition

- *Eine Relation auf einer Menge X ist eine Teilmenge $R \subset X \times X$. Wir schreiben $x \sim y$ für $(x, y) \in R$.*
- *Eine Relation $R \subset X \times X$ heißt Äquivalenzrelation, wenn sie folgende Eigenschaften hat:*
  (i) $x \sim x$ *für alle* $x \in X$,
  (ii) $x \sim y \Longrightarrow y \sim x$ *und*
  (iii) $x \sim y \sim z \Longrightarrow x \sim z$.

  Jedes Element $x \in X$ definiert eine Äquivalenzklasse

$$[x] := \{y \in X | y \sim x\}.$$

Orientierung

Satz

Sei V ein reeller VR der Dimension $n \in \mathbb{N}$. Wir schreiben $B \sim B'$, falls B und B' gleich orientierte Basen von V sind.

(a) *Die Relation "$\sim$" ist eine Äquivalenzrelation auf der Menge der Basen von V.*

(b) *Jede Basis B definiert eine Äquivalenzklasse*

$$[B] := \{B' \quad \textit{Basis von} \quad V | B' \sim B\}$$

und es gibt genau zwei Äquivalenzklassen.

Beweis.

a) ist eine einfache Übungsaufgabe.

b) Sei $B = (b_1, \ldots, b_n)$ eine Basis. Jede Basis ist entweder gleich orientiert zu B oder zu $(-b_1, b_2, \ldots, b_n)$. 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 Basen von V.

Bemerkungen

- Jede Basis $B = (b_1, \ldots, b_n)$ definiert eine Orientierung $[B]$. Wie wir gesehen haben gibt es genau zwei Orientierungen $[B]$ und $[B]^{op} = [(-b_1, b_2, \ldots, b_n)]$. Letztere heißt die zu $[B]$ entgegengesetzte (oder umgekehrte) Orientierung.
- Automorphismen $F \in \mathrm{GL}(V)$ mit $\det F > 0$ heißen orientierungserhaltend, denn $FB = (Fb_1, \ldots, Fb_n) \in [B]$ für jede Basis $B = (b_1, \ldots, b_n)$. $F \in \mathrm{GL}(V)$ mit $\det F < 0$ heißt orientierungsumkehrend, denn $FB = (Fb_1, \ldots, Fb_n) \in [B]^{op}$ für jede Basis B.

Orientierung

Beispiele I

- Die kanonische Orientierung des $\mathbb{R}^n$ ist die durch die kanonische Basis definierte Orientierung $[(e_1, \ldots, e_n)]$. Die Basen (e_2, e_3, e_1) und $(-e_2, e_1, e_3)$ definieren z.B. die kanonische Orientierung des $\mathbb{R}^3$.
- Die Basis (e_2, e_1, e_3) definiert die entgegengesetzte Orientierung $[(-e_1, e_2, e_3)]$.
- Die Drehung (um die z-Achse, entgegen dem Uhrzeigersinn um den Winkel φ)

$$D = \begin{pmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

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.