Mathematik I-II
für Informatiker und Softwaretechniker

Wolf-Patrick Düll

Inhaltsverzeichnis

Teil I

Grundlagen

1 Logik

Aussagen

Die Mathematik ist eine Ansammlung von Aussagen über bestimmte Objekte (z.B. Mengen, Zahlen). Diese Aussagen (Formeln) sind dadurch charakterisiert, dass sie entweder wahr oder falsch sind ("tertium non datur").

Beispiel
1+1=2 ist eine wahre Aussage.
1+1=3 ist eine falsche Aussage.
1+1 ist keine Aussage.

Um Aussagen unmissverständlich und platzsparend aufschreiben zu können, führen wir eine formale mathematische Sprache ein.

Die formale mathematische Sprache

Aufbau der formalen mathematischen Sprache:

1. Alphabet

   Vorrat an Zeichen für

   - Konstanten, z.B. 1, 0, π
   - Variablen, z.B. x, y, ε

     Variablen können mit Konstanten belegt werden, z.B. beim Lösen von Gleichungen wie z.B. $x^2 = 4$
   - Relationen, z.B. $<$, $>$, $\subset$
   - Funktionen und Verknüpfungen, z.B. sin, cos, $+$, $\cup$
   - Relationsvariablen, z.B. R
   - Funktions- und Verknüpfungsvariablen, z.B. $f, g, *$
   - Logische Zeichen
     - Junktoren, z.B. $\wedge, \vee, \neg$
     - Quantoren, z.B. $\exists, \forall$
     - Gleichheitszeichen $=$
   - Klammern, z.B. (,)

   Mit Hilfe dieser Zeichen werden alle mathematischen Objekte definiert und bezeichnet.

2. Aneinanderreihung der Zeichen zu Termen, z.B. $1 + 1$, und zu Formeln, z.B. $1 + 1 = 2$.

Axiome und logisches Schließen

Frage: Welche Aussagen (Formeln) sind wahr?

Einige Formeln werden als wahr vorausgesetzt (Axiome).
Der Wahrheitswert der anderen Aussagen soll durch logisches Schließen unter Voraussetzung der Axiome ermittelt werden (mathematischer Beweis).
Das logische Schließen folgt nach festen Regeln.
Einige dieser Regeln werden ebenfalls als logische Axiome vorausgesetzt, die anderen Regeln lassen sich durch Kombinationen der logischen Axiome gewinnen.

Festlegung der logischen Schlussregeln:

1. Möglichkeit: durch Wahrheitstafeln,

2. Möglichkeit: durch einen Logikkalkül.

Wahrheitstafeln

Jede wahre Aussage erhält den Wahrheitswert 1, jede falsche den Wahrheitswert 0. Seien A, B, ... Aussagen. In Abhängigkeit ihrer Wahrheitswerte $\mathrm{w}(A)$, $\mathrm{w}(B)$, ... werden die Wahrheitswerte von Aussagen, die aus A, B, ... und logischen Zeichen zusammengesetzt sind, festgelegt.

Wichtige Beispiele:

1) $\neg A$: nicht A, Negation von A

A	$\neg A$
1	0
0	1

2) $A \wedge B$: A und B

3) $A \vee B$: A oder B

4) A XOR B: entweder A oder B

5) A NAND B: nicht (A und B)

6) A NOR B: nicht (A oder B)

A	B	$A \wedge B$	$A \vee B$	A XOR B	A NAND B	A NOR B
1	1	1	1	0	0	0
1	0	0	1	1	1	0
0	1	0	1	1	1	0
0	0	0	0	0	1	1

7) $A \to B$: "aus A folgt B", "A impliziert B", "wenn A, dann B"

A	B	$A \to B$
1	1	1
1	0	0
0	1	1
0	0	1

Die Zuweisung w($A \to B$)=1 in der dritten und vierten Zeile nennt man auch "ex falso quodlibet".
Ist $A \to B$ wahr, also w($A \to B$)=1, dann schreiben wir auch $A \Rightarrow B$.
Man nennt in diesem Fall A eine hinreichende Bedingung für B und B eine notwendige Bedingung für A.

Beispiel
$x = 1 \Rightarrow x^2 = 1$

Kennt man den Wahrheitswert einer aus mehreren Aussagen und logischen Zeichen zusammengesetzten Aussage, dann kann man manchmal daraus Wahrheitswerte von einzelnen Bestandteilen ermitteln.

Beispiel
Ist w($A \to B$)=0, dann ist w(A)=1 und w(B)=0. Ist w($A \to B$)=1 und w(A)=1, dann ist w(B)=1. Ist w($A \to B$)=1 und w(A)=0, dann kann w(B) daraus nicht ermittelt werden.

8) $A \leftrightarrow B$: "A ist äquivalent zu B", "genau dann A, wenn B"

A	B	$A \leftrightarrow B$
1	1	1
1	0	0
0	1	0
0	0	1

Ist w($A \leftrightarrow B$)=1, dann schreiben wir auch $A \Leftrightarrow B$.

Beispiel
$2x - 1 = 1 \Leftrightarrow x = 1$

Tautologien

Eine aus den Aussagen A, B, ... und logischen Zeichen zusammengesetzte Aussage, die unabhängig von w(A), w(B), ... immer wahr ist, heißt Tautologie. Tautologien können durch Wahrheitstafeln nachgewiesen werden.

Beispiele

1) $A \vee \neg A$

2) $\neg(A \wedge \neg A)$ (Satz vom Widerspruch)

3) $A \wedge (A \to B) \Rightarrow B$

4) $A \to B \Leftrightarrow \neg B \to \neg A$

A	B	$\neg A$	$\neg B$	$A \to B$	$\neg B \to \neg A$	$A \to B \Leftrightarrow \neg B \to \neg A$
1	1	0	0	1	1	1
1	0	0	1	0	0	1
0	1	1	0	1	1	1
0	0	1	1	1	1	1

5) $A \to B \Leftrightarrow \neg(A \wedge \neg B)$

6) $A \leftrightarrow B \Leftrightarrow (A \to B) \wedge (B \to A)$

7) $(A \to B) \wedge (\neg A \to B) \Leftrightarrow B$ (Fallunterscheidungsregel)

8) De Morgan'sche Regeln

a) $\neg(A \wedge B) \Leftrightarrow \neg A \vee \neg B$

b) $\neg(A \vee B) \Leftrightarrow \neg A \wedge \neg B$

9) $(A \to B) \wedge (B \to C) \Rightarrow A \to C$ (Transitivität der Implikation)

10) Distributivgesetze

a) $A \wedge (B \vee C) \Leftrightarrow (A \wedge B) \vee (A \wedge C)$

b) $A \vee (B \wedge C) \Leftrightarrow (A \vee B) \wedge (A \vee C)$

Anwendungen von Tautologien:

1) Tautologien liefern wichtige Beweistechniken.

Beispiel
Es gibt 3 Alternativen, um $A \Rightarrow B$ zu beweisen:

a) Direkter Beweis
Angenommen, A sei wahr (anderenfalls gilt ohnehin $A \Rightarrow B$). Dann zeigt man durch eine Kette von logischen Schlüssen die Gültigkeit von B.

b) Indirekter Beweis
Zeige $\neg B \Rightarrow \neg A$ durch einen direkten Beweis.

c) Widerspruchsbeweis
Zeige, dass aus $A \wedge \neg B$ ein Widerspruch folgt, d.h. $A \wedge \neg B \Rightarrow C \wedge \neg C$ für irgendeine Aussage C. Dann muss $A \wedge \neg B$ falsch sein und somit $A \Rightarrow B$ gelten.

Beispiel
Sei $A :\Leftrightarrow |x-4| < 1$ und $B :\Leftrightarrow x < 5$.
Zeige $A \Rightarrow B$.

a) direkt
Sei $|x-4| < 1$.

1.Fall: $x \geq 4 \Rightarrow x - 4 = |x-4| < 1 \Rightarrow x < 5$
2.Fall: $x < 4 \Rightarrow x < 5$

b) indirekt
$\neg B \Leftrightarrow x \geq 5 \Rightarrow x - 4 \geq 1 > 0$
$\Rightarrow |x-4| = x - 4 \geq 1$
$\Leftrightarrow \neg A$

c) Widerspruchsbeweis
Angenommen, es gelte $|x-4| < 1 \wedge x \geq 5$.
$\Rightarrow 1 \leq x - 4 = |x-4| < 1 \Rightarrow 1 < 1$ Widerspruch!

2) Aufgrund von Tautologien ist jede logische Verknüpfung von 2 Aussagen A, B (es gibt 16 verschiedene solche Verknüpfungen) jeweils äquivalent zu einer Aussage, in der nur A, B, $\neg$ und $\vee$ auftreten, sowie äquivalent zu einer Aussage, in der nur A, B, $\neg$ und $\wedge$ auftreten.

Beispiel
$A \to B \Leftrightarrow (A \wedge \neg B) \Leftrightarrow \neg A \vee B$

Außerdem ist jede dieser 16 Aussagen jeweils äquivalent zu einer Aussage, in der nur A, B und NAND auftreten, sowie äquivalent zu einer Aussage, in der nur A, B und NOR auftreten.

Beispiel
$A \wedge B \Leftrightarrow (A \text{ NAND } B) \text{ NAND } (A \text{ NAND } B)$,
$\neg(A \wedge B) \Leftrightarrow ((A \text{ NAND } B) \text{ NAND } (A \text{ NAND } B)) \text{ NAND } ((A \text{ NAND } B) \text{ NAND } (A \text{ NAND } B))$

3) Tautologien helfen bei der Konstruktion und bei der Vereinfachung von elektrischen Schaltkreisen (Grundlage eines jeden Computers).

Quantoren

Sei $A(x)$ eine Aussage, in der die Variable x vorkommt.

Bezeichnungen:
$\forall x$: $A(x)$ "für alle x gilt $A(x)$"
$\exists x$: $A(x)$ "es existiert ein x, so dass $A(x)$ gilt"
$\exists! x$: $A(x)$ "es existiert genau ein x, so dass $A(x)$ gilt"
$\forall$ heißt Allquantor, $\exists$ Existenzquantor.

Beispiel
Die Aussagen
$$\forall x : x > 1 \to x > 0$$
sowie
$$\exists! x : 1 + x = 2$$
sind beide wahr, wenn x jeweils mit reellen Zahlen belegt ist.

Verneinung von Quantoren:

Es gilt:
$\neg(\forall x : A(x)) \Leftrightarrow \exists x : \neg A(x)$
$\neg(\exists x : A(x)) \Leftrightarrow \forall x : \neg A(x)$

Beispiel

$\forall x : x^2 \geq 0 \rightarrow x \geq 0$ ist falsch für reelle x.

Beweis durch Gegenbeispiel:
Es ist $(-1)^2 = 1 \geq 0$, aber $-1 < 0 \Rightarrow \exists x : \neg(x^2 \geq 0 \rightarrow x \geq 0)$.

Logikkalküle

Für Details bezüglich der Einführung eines Logikkalküls (es gibt mehrere äquivalente Möglichkeiten) und der Beziehung zwischen Logikkalkülen und Wahrheitstafeln (Vollständigkeitssätze, 1. und 2. Gödelscher Unvollständigkeitssatz, Turingmaschinen) siehe zum Beispiel Ebbinghaus, Plum, Thomas: Einführung in die mathematische Logik.

2 Mengen

Die grundliegenden Objekte der Mathematik sind Mengen. Dabei setzt man voraus, dass eine Art "Urmenge" existiert, welche eine Sammlung von unendlich vielen Elementen ist.
In den Lehrveranstaltungen zur Mathematik I/II genügt es, von den natürlichen Zahlen $\mathbb{N} = \{1, 2, 3, ...\}$ als "Urmenge" auszugehen. Eine mathematisch exakte Definition von $\mathbb{N}$ folgt demnächst.
Ausgehend von dieser "Urmenge" können dann durch festgelegte Mengenbildungsregeln weitere Mengen gebildet werden.

Bezeichnungen

$m \in M$ " m ist Element der Menge M "

Bsp.: $1 \in \mathbb{N}$

$m \notin M$ " m ist nicht Element der Menge M "

Bsp.: $-1 \notin \mathbb{N}$

Es gilt also $m \notin M \Leftrightarrow \neg(m \in M)$.

Mengenbildungsregeln

1. Aussonderungsregel
Ist M_1 eine Menge und $A(x)$ eine Aussage, dann ist $M_2 = \{x \in M_1 : A(x)\}$ eine Menge, d.h., M_2 enthält alle Elemente von M_1, für die $A(x)$ wahr ist.

Beispiel
$M_1 = \mathbb{N}$, $M_2 = \{x \in \mathbb{N} : x > 10\}$

2. Vereinigungsregel
Sind M_1 und M_2 Mengen, dann ist $M_3 = M_1 \cup M_2 = \{x : x \in M_1 \vee x \in M_2\}$ eine Menge. M_3 heißt dann Vereinigung von M_1 und M_2.

Beispiel
$M_1 = \{1, 2, 3\}$, $M_2 = \{3, 4\}$, $M_3 = \{1, 2, 3, 4\}$

3. Paarmengenregel
Sind M_1 und M_2 Mengen, dann ist $M_3 = \{M_1, M_2\}$ eine Menge. M_3 heißt dann Paarmenge von M_1 und M_2. M_3 enthält die Mengen M_1 und M_2 als Elemente.

Beispiel
$M_1 = \{1, 2, 3\}$, $M_2 = \{3, 4\}$, $M_3 = \{\{1, 2, 3\}, \{3, 4\}\}$

4. Gleichheitsregel
Zwei Mengen M und N sind gleich, wenn sie die gleichen Elemente besitzen.
In der formalen mathematischen Sprache schreibt sich diese Regel folgendermaßen:

$$\forall M \forall N : (\forall x : ((x \in M \rightarrow x \in N) \wedge (x \in N \rightarrow x \in M)) \rightarrow M = N)$$

Beispiel
Sei $M = \{x \in \mathbb{N} : x > 1\}$ und $N = \{x \in \mathbb{N} : x + 1 > 2\}$. Zeige: $M = N$.

Beweis

$$\begin{array}{lcl} x \in M & \Rightarrow & x \in \mathbb{N} \;\wedge\; x > 1 \\ & \Rightarrow & x \in \mathbb{N} \;\wedge\; x + 1 > 2 \\ & \Rightarrow & x \in N \\ x \in N & \Rightarrow & x \in \mathbb{N} \;\wedge\; x + 1 > 2 \\ & \Rightarrow & x \in \mathbb{N} \;\wedge\; x > 1 \\ & \Rightarrow & x \in M \end{array}$$

Also ist $M = N$. $\square$

Eine Menge N heißt Teilmenge von M, in Zeichen $N \subseteq M$, falls gilt:

$$\forall x : x \in N \Rightarrow x \in M.$$

M heißt dann Obermenge von N.
Ist $N \subseteq M$ und $N \neq M$, dann schreiben wir auch $N \subset M$.
In mancher Literatur wird "$\subset$" statt "$\subseteq$" und "$\subsetneq$" statt "$\subset$"verwendet.

Beispiel
$\{1, 2, 3\} \subseteq \mathbb{N}$, $\{1, 2, 3\} \subset \mathbb{N}$

5. Potenzmengenregel
Sei M eine Menge, dann ist $\mathcal{P} = \{N : N \subseteq M\}$ eine Menge, die sogenannte Potenzmenge von M.

Beispiel
$M = \{1, 2, 3\}$, $\mathcal{P}(M) = \{\{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}, \emptyset\}$
Hierbei ist $\emptyset$ die leere Menge. Sie enthält keine Elemente.

Seien $M, N \neq \emptyset$. Eine Zuordnungsvorschrift f, die jedem $x \in M$ genau ein $y = f(x) \in N$ zuordnet, heißt Abbildung oder Funktion.
Kurzschreibweise: $f : M \to N$, $x \mapsto f(x)$.
M heißt Definitionsbereich von f.
Ist $y = f(x)$, dann heißt y das Bild von x (unter f) und x das Urbild von y.

Beispiel
$f : \mathbb{N} \to \mathbb{N}$, $x \mapsto x^2$ ist eine Funktion.

6. Ersetzungsregel
Sei f eine Funktion und M Teilmenge des Definitionsbereichs von f. Dann ist $f(M) = \{f(x) : x \in M\}$ eine Menge, die sogenannte Bildmenge von M (unter der Funktion f).

Beispiel
$f : \mathbb{N} \to \mathbb{N}$, $x \mapsto x + 1$, $M = \{1, 2, 3\}$, $f(M) = \{2, 3, 4\}$

Aufgrund der Aussonderungsregel können wir auch für jede Teilmenge Y des Wertebereichs einer Funktion $f : M \to N$ die Menge $f^{-1}(Y) = \{x \in M : f(x) \in Y\}$, die sogenannte Urbildmenge von Y, bilden.

Beispiel
$f : \mathbb{N} \to \mathbb{N}$, $x \mapsto x + 1$, $Y = \{3, 4, 5\}$, $f^{-1}(Y) = \{2, 3, 4\}$

Die Mengenbildungsregeln 1–6 erlauben weitere Mengenbildungen. Für zwei Mengen M_1, M_2 ist zum Beispiel:

a) $M_1 \cap M_2 = \{x : x \in M_1 \wedge x \in M_2\} = \{x \in M_1 : x \in M_2\}$

   die Schnittmenge von M_1 und M_2 (bei dieser Mengenbildung wird die Aussonderungsregel angewandt). Ist $M_1 \cap M_2 = \emptyset$, dann heißen M_1 und M_2 disjunkt.

b) $M_1 \setminus M_2 = \{x : x \in M_1 \wedge x \notin M_2\} = \{x \in M_1 : x \notin M_2\}$

   die Differenzmenge von M_1 und M_2. Ist $M_2 \subseteq M_1$, dann heißt $M_1 \setminus M_2$ auch das Komplement von M_2 in M_1, in Zeichen: M_2^c.

c) $M_1 \times M_2 = \{(m_1, m_2) : m_1 \in M_1 \wedge m_2 \in M_2\}$

   das kartesische Produkt von M_1 und M_2. Hierbei ist $(m_1, m_2) = \{m_1, \{m_2\}\}$ das sogenannte geordnete Paar von m_1 und m_2. Im Unterschied zur Menge $\{m_1, m_2\} = \{m_2, m_1\}$ ist bei geordneten Paaren die Reihenfolge von m_1 und m_2 entscheidend, d.h., es ist $(m_1, m_2) \neq (m_2, m_1)$, falls $m_1 \neq m_2$ ist. Insbesondere gilt $(m_1, m_2) = (m_1', m_2') \Leftrightarrow m_1 = m_1' \wedge m_2 = m_2'$.

Beispiel
$M_1 = \{x \in \mathbb{R} : 1 \leq x \leq 3\}$, $M_2 = \{x \in \mathbb{R} : 2 \leq x \leq 4\}$,
$M_1 \cap M_2 = \{x \in \mathbb{R} : 2 \leq x \leq 3\}$,
$M_1 \setminus M_2 = \{x \in \mathbb{R} : 1 \leq x < 2\}$,
$M_1 \times M_2 = \{(x, y) \in \mathbb{R} \times \mathbb{R} : 1 \leq x \leq 3 \ \wedge \ 2 \leq y \leq 4\}$.

Allgemeiner erhält man für Mengen $M_1, \ldots, M_n$:

a) $\displaystyle\bigcup_{k=1}^{n} M_k \ = \ M_1 \cup M_2 \cup \cdots \cup M_n \ = \ \{x : \exists j \in \{1, \ldots, n\} : x \in M_j\}$

   die Vereinigung der Mengen $M_1, \ldots, M_n$.

b) $\displaystyle\bigcap_{k=1}^{n} M_k \ = \ M_1 \cap M_2 \cap \cdots \cap M_n \ = \ \{x : \forall j \in \{1, \ldots, n\} : x \in M_j\}$

   die Schnittmenge von $M_1, \ldots, M_n$.

c) $\displaystyle\prod_{k=1}^{n} M_k \ = \ M_1 \times M_2 \times \cdots \times M_n \ = \ \{ \underbrace{(x_1, \ldots, x_n)}_{\text{geordnetes n-Tupel}} : \forall j \in \{1, \ldots, n\} : x_j \in M_j\}$

   das kartesische Produkt von $M_1, \ldots, M_n$. Ist $M_1 = M_2 = \cdots = M_n = M$, dann schreibt man auch M^n statt $\underbrace{M \times \cdots \times M}_{\text{n-fach}}$.

Beispiel
$\mathbb{R}^2 = \mathbb{R} \times \mathbb{R} = \{(x, y) : x, y \in \mathbb{R}\}$,
$\mathbb{R}^3 = \mathbb{R} \times \mathbb{R} \times \mathbb{R} = \{(x, y, z) : x, y, z \in \mathbb{R}\}$.

Rechengesetze für Mengen

Seien L, M, N Mengen, dann gilt:

1. Assoziativgesetz
   (a) $(L \cup M) \cup N = L \cup (M \cup N)$
   (b) $(L \cap M) \cap N = L \cap (M \cap N)$
2. Kommutativgesetz
   (a) $M \cup N = N \cup M$
   (b) $M \cap N = N \cap M$
3. Distributivgesetz
   (a) $L \cup (M \cap N) = (L \cup M) \cap (L \cup N)$
   (b) $L \cap (M \cup N) = (L \cap M) \cup (L \cap N)$
4. De Morgan'sche Regeln
   (a) $L \setminus (M \cup N) = (L \setminus M) \cap (L \setminus N)$
   (b) $L \setminus (M \cap N) = (L \setminus M) \cup (L \setminus N)$

Beweis
dieser Regeln durch logisches Schließen unter Ausnutzung von Tautologien, zum Beispiel für 3.(a):
Betrachte die Aussagen:

$$A(x) :\Leftrightarrow x \in L, \quad B(x) :\Leftrightarrow x \in M, \quad C(x) :\Leftrightarrow x \in N$$

Dann gilt:

$$\begin{aligned} x \in L \cup (M \cap N) \quad &\Leftrightarrow \quad A(x) \vee (B(x) \wedge C(x)) \\ &\overset{1.5\ 10)b)}{\Leftrightarrow} \quad (A(x) \vee B(x)) \wedge (A(x) \vee C(x)) \\ &\Leftrightarrow \quad x \in (L \cup M) \cap (L \cup N) \end{aligned}$$

Also gilt 3.(a). □

Wir erkennen: $\cup$ bei Mengen entspricht $\vee$ bei Aussagen, $\cap$ bei Mengen entspricht $\wedge$ bei Aussagen und $\setminus$ bei Mengen entspricht $\neg$ bei Aussagen.

Relationen

Seien M_1, M_2 Mengen. Jede Teilmenge $R \subseteq M_1 \times M_2$ heißt (zweistellige, binäre) Relation (zwischen M_1 und M_2). Für $(x,y) \in R$ schreibt man auch xRy.

Beispiele

a) $M_1 = M_2 = \mathbb{R}$,
$R = \{(x,y) \in \mathbb{R}^2 : x < y\}$ ist eine Relation auf $\mathbb{R}$.
Bei dieser Relation schreibt man auch $(x,y) \in <$ statt $(x,y) \in R$ sowie $x < y$ statt xRy.

b) $K = \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$ ist ebenfalls eine Relation auf $\mathbb{R}$.

Sei $f : M_1 \to M_2$ eine Funktion. Dann ist der sogenannte Graph

$$G(f) \; = \; \{(x,y) \in M_1 \times M_2 : y = f(x)\} \; = \; \{(x, f(x)) \in M_1 \times M_2\}$$

von f eine Relation. Ist umgekehrt R eine Relation zwischen M_1 und M_2 mit der Eigenschaft

$$\forall x \in M_1 \exists! y \in M_2 : (x,y) \in R,$$

dann ist $R = G(f)$ mit $f : M_1 \to M_2$, $x \mapsto y$.

Beispiel

$M = \{x \in \mathbb{R} : -1 \le x \le 1\}$,
$H = \{(x,y) \in M \times M : x^2 + y^2 = 1 \wedge y \ge 0\}$
$\Rightarrow H = G(f)$ mit $f : M \to M$, $f(x) = \sqrt{1 - x^2}$.
Beachte: Hier gilt $f(M) \subset M$.

Eine Relation $R \subseteq M \times M$ heißt Äquivalenzrelation (auf M), falls

(i) $\forall a \in M : \; aRa \quad$ (R ist reflexiv)

(ii) $\forall a, b \in M : \; aRb \; \Leftrightarrow \; bRa \quad$ (R ist symmetrisch)

(iii) $\forall a, b, c \in M : \; aRb \wedge bRc \; \Rightarrow \; aRc \quad$ (R ist transitiv)

Für Äquivalenzrelationen verwendet man häufig das Zeichen $\sim$ statt R.
Sei $\sim$ eine Äquivalenzrelation auf M und $a \in M$. Dann heißt die Menge
$[a] := \{b \in M : b \sim a\}$ die Äquivalenzklasse von a (bzgl. $\sim$).

Beispiel

Sei $M = \mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2 \ldots\}$ $n \in \mathbb{N} \backslash \{1\}$. Dann definiert

$$a \sim b :\Leftrightarrow \exists k \in \mathbb{Z} : kn = a - b$$

eine Äquivalenzrelation (Nachweis der Gültigkeit von (i), (ii), (iii) durch Nachrechnen).
Bei dieser Äquivalenzrelation schreibt man für $a \sim b$ auch $a = b \mod n$ (b modulo n) Die Menge aller Äquivalenzklassen bzgl. dieser Äquivalenzrelation bezeichnet man mit $\mathbb{Z}_n$ oder mit $\mathbb{Z}/n\mathbb{Z}$.
Im Falle $n = 2$ gibt es genau 2 verschiedene Äquivalenzklassen, nämlich $[0] = \{0, \pm 2, \pm 4, \ldots\}$ und $[1] = \{\pm 1, \pm 3, \pm 5, \ldots\}$. Also ist $\mathbb{Z}_2 = \{[0], [1]\}$.

Axiome der Mengenlehre

Formuliert man die Mengenbildungsregeln 1–6 aus 2.1 sowie das Postulat der Existenz einer unendlichen "Urmenge" mathematisch exakt in der formalen mathematischen Sprache, dann erhält man das Zermelo-Fraenkelsche Axiomensystem der Mengenlehre.
Dieses Axiomensystem schließt Mengenbildungen aus, die in sich widersprüchlich sind, wie z.B. $\{M : M \notin M\}$.
Ein zusätzliches Axiom der Mengenlehre ist das sogenannte Auswahlaxiom. Es erleichtert viele mathematische Argumentationen, impliziert aber auch überraschende Aussagen wie das Banach-Tarski Paradoxon.
Für Details siehe zum Beispiel Ebbinghaus.

3 Die reellen Zahlen

3.1 Einführung der reellen Zahlen

Möglichkeit 1:
Konstruktion der reellen Zahlen ausgehend von den natürlichen Zahlen:

1. Schritt:
Definition der natürlichen Zahlen durch ein Axiomensystem (Peano-Axiome):
Die natürlichen Zahlen bilden eine Menge $\mathbb{N}$, auf der eine Abbildung $\nu : \mathbb{N} \to \mathbb{N}$ (Nachfolgerfunktion) definiert ist mit folgenden Eigenschaften:

($\mathbb{N}_1$) $\exists!\ x \in \mathbb{N} : x \notin \nu(\mathbb{N})$. Bezeichnung: $1 := x$

($\mathbb{N}_2$) $\forall n_1, n_2 \in \mathbb{N} : \nu(n_1) = \nu(n_2) \Rightarrow n_1 = n_2$.

($\mathbb{N}_3$) Induktionsaxiom:
Ist $M \subseteq \mathbb{N}$ und gilt

(i) $1 \in M$ und

(ii) $n \in M \Rightarrow \nu(n) \in M$,

dann ist $M = \mathbb{N}$.

Man bezeichnet nun: $2 := \nu(1)$, $3 := \nu(2), \ldots$.
Für $m, n \in \mathbb{N}$ sei

$$m + 1 := \nu(m), \quad m + (n + 1) := \nu(m + n),$$

$$m \cdot 1 := m, \quad m \cdot (n + 1) := m \cdot n + m.$$

Aufgrund von ($\mathbb{N}_3$) sind auf diese Weise $m + n$ und $m \cdot n$ für alle $n, m \in \mathbb{N}$ definiert.
Außerdem definiert man

$$m < n :\Leftrightarrow \exists d \in \mathbb{N} : m + d = n.$$

Ausgehend von diesen Definitionen von $+, \cdot, <$ können mit Hilfe von ($\mathbb{N}_3$) alle bekannten Rechengesetze von $\mathbb{N}$ (z.B. Assoziativgesetz, Kommutativgesetz, Distributivgesetz, Umformungsregeln für Ungleichungen) bewiesen werden.
Alternativ kann in ($\mathbb{N}_1$) und ($\mathbb{N}_3$) die 1 durch die 0 ersetzt werden und $1 := \nu(0)$ gesetzt werden. Dann ist $0 \in \mathbb{N}$.
Anderenfalls führt man $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$ ein, um die 0 einzubeziehen.

2. Schritt:
Konstruktion der ganzen Zahlen $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$ aus $\mathbb{N}$ durch Bildung geordneter Paare, z.B. $(1, 2)$ für $1 - 2 =: -1$, und Äquivalenzklassenbildung zur Gleichsetzung gleichwertiger Differenzen, z.B. $(1, 2) \sim (2, 3)$ wegen $1 - 2 = 2 - 3$.
Erweiterung der Definitionen von $+, \cdot, <$ auf $\mathbb{Z}$.

3. Schritt:
Konstruktion der rationalen Zahlen $\mathbb{Q} = \frac{m}{n} : m \in \mathbb{Z}, n \in \mathbb{N}\}$ aus $\mathbb{Z}$ durch Bildung geordneter Paare, z.B. $(1,2)$ für $1 : 2 =: \frac{1}{2}$, und Äquivalenzklassenbildung zur Gleichsetzung erweiterter oder gekürzter Brüche, z.B. $(2,4) \sim (1,2) \sim (3,6)$ wegen $\frac{2}{4} = \frac{1}{2} = \frac{3}{6}$.
Erweiterung der Definitionen von $+, \cdot, <$ auf $\mathbb{Q}$.

4. Schritt:
Konstruktion der reellen Zahlen $\mathbb{R}$ aus $\mathbb{Q}$ zum Beispiel mit Hilfe von Intervallschachtelungen zur Definition der irrationalen Zahlen (nicht abbrechende, nicht periodische Dezimalbrüche).
Erweiterung der Definitionen von $+, \cdot, <$ auf $\mathbb{R}$.
Auf der Basis der oben skizzierten Konstruktionen und Definitionen lassen sich alle bekannten Rechengesetze von $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$ und $\mathbb{R}$ beweisen.

Für mehr Details siehe Lehrbücher für MathematikstudentInnen, teilweise aber auch Lehrbücher für InformatikstudentInnen.

Möglichkeit 2:
Beschreibung von $\mathbb{R}$ durch Axiome:

Die reellen Zahlen bilden eine Menge $\mathbb{R}$, in der folgende Axiome gelten:

1) die Körperaxiome, siehe 3.2,

2) das Induktionsaxiom, siehe 3.3,

3) die Anordnungsaxiome, siehe 3.4,

4) das Supremumsaxiom, siehe 3.5.

Mit Hilfe der Konstruktionen aus Möglichkeit 1 kann unter Voraussetzung der Existenz von $\mathbb{N}$ bewiesen werden, dass eine (bis auf strukturerhaltende Kopien) eindeutige Menge $\mathbb{R}$ existiert, welche die Axiome 1) - 4) erfüllt. Aus den Axiomen 1) - 4) folgen alle bekannten Rechengesetze von $\mathbb{R}, \mathbb{Q}, \mathbb{Z}$ und $\mathbb{N}$.

3.2 Die Körperaxiome und Folgerungen

Axiom 1 *Auf der Menge der reellen Zahlen $\mathbb{R}$ sind zwei Verknüpfungen, d.h. Abbildungen von $\mathbb{R} \times \mathbb{R}$ nach $\mathbb{R}$, nämlich $+$ (Addition) und $\cdot$ (Multiplikation) definiert mit folgenden Eigenschaften:*

(A_1) $\forall a, b, c \in \mathbb{R} : \ (a + b) + c = a + (b + c)$ *(Assoziativität von +)*

(A_2) $\forall a, b \in \mathbb{R} : \ a + b = b + a$ *(Kommutativität von +)*

(A_3) $\exists!\ x \in \mathbb{R} \forall a \in \mathbb{R} : \ a + x = a$ *(Neutrales Element bzgl. +)*
Bezeichnung: $0 := x$.

(A_4) $\forall a \in \mathbb{R} \exists!\ y \in \mathbb{R}:\ a + y = 0$ *(Inverses Element bzgl. +)*
Bezeichnung: $-a := y$

(M_1) $\forall a, b, c \in \mathbb{R}:\ (a \cdot b) \cdot c = a \cdot (b \cdot c)$ *(Assoziativität von ·)*

(M_2) $\forall a, b \in \mathbb{R}:\ a \cdot b = b \cdot a$ *(Kommutativität von ·)*

(M_3) $\exists!\ p \in \mathbb{R} \forall a \in \mathbb{R}:\ a \cdot p = a$ *(Neutrales Element bzgl. ·)*
Bezeichnung: $1 := p$.

(M_4) $\forall a \in \mathbb{R} \backslash \{0\} \exists!\ q \in \mathbb{R}:\ a \cdot q = 1$ *(Inverses Element bzgl. ·)*
Bezeichnung: $a^{-1} := \frac{1}{a} := q$.

(D) $\forall a, b, c \in \mathbb{R}:\ a \cdot (b + c) = a \cdot b + a \cdot c$ *(Distributivität)*
Hierbei sei Punktrechnung vor Strichrechnung vereinbart, d.h. $a \cdot b + a \cdot c = (a \cdot b) + (a \cdot c)$.

Definition

a) *Eine Menge M mit einer Verknüpfung* $*:\ H * H \to H$ *heißt Halbgruppe, falls gilt*
$$\forall x, y, z \in M:\ x * (y * z) = (x * y) * z,$$
d.h., falls $*$ *assoziativ ist.*

b) *Eine Halbgruppe* $(M, *)$ *heißt Monoid, falls gilt*
$$\exists e \in M \forall x \in M:\ e * x = x = x * e,$$
d.h., falls ein neutrales Element existiert.

c) *Ein Monoid* $(G, *)$ *heißt Gruppe, falls gilt*
$$\forall x \in G \exists y \in G.\ y * x = e,$$
d.h., falls ein linksinverses Element zu x *existiert, nämlich* y

d) *Eine Gruppe* $(A, *)$ *heißt abelsch oder kommutativ, falls gilt*
$$\forall x, y \in A:\ x * y = y * x,$$
d.h., falls $*$ *kommutativ ist.*

Satz 3.1 *Sei* $(M, *)$ *ein Monoid. Dann gibt es genau ein neutrales Element in M.*

Beweis
Seien e, e' neutral. Dann gilt $e = e' * e = e' \Rightarrow e = e'$. $\square$

Satz 3.2 *In jeder Gruppe* $(G, *)$ *gilt:*

(i) $\forall x, y \in G: \; y * x = e \Rightarrow x * y = e$, *d.h., jedes linksinverse Element ist auch rechtsinvers und somit invers (d.h. links- und rechtsinvers).*

(ii) $\forall x \in G \exists!\, y \in G: \; y * x = e = x * y$
Bezeichnung: $x^{-1} := y$

(iii) $\forall a, b \in G \exists!\, x \in G: \; a * x = b$

(iv) $\forall a, b \in G \exists!\, y \in G: \; y * a = b$

(v) $\forall x \in G: \; (x^{-1})^{-1} = x$

(vi) $\forall x, y \in G: \; (x * y)^{-1} = y^{-1} * x^{-1}$

Beweis

(i) Seien $x, y \in G$ und es gelte $y * x = e$.
$\Rightarrow \exists z \in G: \; z * y = e$
$\Rightarrow x * y = e * (x * y) = (z * y) * (x * y) \overset{Ass.}{=} z * (y * (x * y)) = z * ((y * x) * y) = z * (e * y) = z * y = e$

(ii) Seien $x, y, z \in G$ und es gelte $y * x = e \overset{(i)}{=} x * y \;\wedge\; z * x = e \overset{(i)}{=} x * z$.
$\Rightarrow z = e * z = (y * x) * z = y * (x * z) = y * e = y$

(iii) Eindeutigkeit:
Sei $a * x = b$.
$\Rightarrow a^{-1} * a * x = a^{-1} * b$
$\Rightarrow x = a^{-1} * b$
Existenz:
Sei $x := a^{-1} * b$.
$\Rightarrow a * x = a * (a^{-1} * b) = (a * a^{-1}) * b = e * b = b$
Also gilt $a * b = x \Leftrightarrow x = a^{-1} * b$. $\square$

Definition

a) Eine Menge R *mit zwei Verknüpfungen* $+, \cdot$ *heißt Ring, falls gilt:*

(i) $(R, +)$ *ist kommutative Gruppe.*

(ii) $(R, \cdot)$ *ist Monoid.*

(iii) $\forall a, b, c \in R: \; (a + b) \cdot c = a \cdot c + b \cdot c \;\wedge\; a \cdot (b + c) = a \cdot b + a \cdot c$ *(Distributivgesetz)*
Hierbei gilt Punkt vor Strich, d.h. $a \cdot c + b \cdot c = (a \cdot c) + (b \cdot c)$

(iv) $0 \neq 1$, *wobei* 0 *das neutrale Element bzgl.* $+$ *und* 1 *das neutrale Element bzgl.* $\cdot$ *bezeichnet.*

b) Ein Ring $(R, +, \cdot)$ *heißt kommutativ, falls* $\cdot$ *kommutativ ist.*

c) Ein Ring $(R, +, \cdot)$ *heißt Körper, falls* $(R\backslash\{0\}, \cdot)$ *eine kommutative Gruppe ist.*

Bezeichnungen
$-\, x :=$ Inverses von x bzgl. $+$,
$x - y := x + (-y)$,
$\frac{1}{x} := x^{-1} :=$ Inverses bzgl. $\cdot$,
$\frac{x}{y} := x : y := x \cdot y^{-1}$,
$x^2 := x \cdot x$,
$xy := x \cdot y$.

Mit Hilfe der eben eingeführten Fachbegriffe und wegen der Sätze 3.1 und 3.2 können wir das Axiom 1 erheblich kürzer formulieren:

Axiom 1 *Auf* $\mathbb{R}$ *sind zwei Verknüpfungen* $+$ *und* $\cdot$ *definiert und* $(\mathbb{R}, +, \cdot)$ *ein Körper.*

Satz 3.3 *In jedem Ring* $(R, +, \cdot)$ *(und somit auch in jedem Körper, insbesondere in* $(\mathbb{R}, +, \cdot)$*) gelten:*

(i) $\forall x \in R:\ 0 \cdot x = 0 = x \cdot 0$

(ii) $\forall x, y \in R:\ (-x) \cdot y = -(x \cdot y)$ *und damit* $(-1) \cdot x = -x$

(iii) $\forall x, y \in R:\ (-x) \cdot (-y) = x \cdot y$

(iv) $\forall x, y, z \in R:\ -x(y - z) = -xy + xz$

Beweis

(*i*) $0 \cdot x + 0 \cdot x \overset{Distr.}{=} (0 + 0) \cdot x = 0 \cdot x$
$\overset{3.1}{\Rightarrow} 0 \cdot x = 0$
$x \cdot 0 + x \cdot 0 = x \cdot (0 + 0) = x \cdot 0$
$\Rightarrow x \cdot 0 = 0$

(*ii*) $x \cdot y + (-x) \cdot y = (x + (-x)) \cdot y = 0 \cdot y \overset{(i)}{=} 0$
$\overset{3.2}{\Rightarrow} (-x) \cdot y = -(x \cdot y)$ □

Satz 3.4 *In jedem kommutativen Ring* $(R, +, \cdot)$ *(und somit auch in jedem Körper, insbesondere in* $(\mathbb{R}, +, \cdot)$*) gelten die binomischen Formeln*

(*i*) $\forall a, b \in R:\ (a + b)^2 = a^2 + 2ab + b^2$,

(*ii*) $\forall a, b \in R:\ (a - b)^2 = a^2 - 2ab + b^2$,

(*iii*) $\forall a, b \in R:\ (a + b)(a - b) = (a^2 - b^2)$.

Beweis

(*i*) $(a+b)^2 = (a+b)(a+b) = a(a+b) + b(a+b) = a^2 + ab + ba + b^2 \overset{Ring\ komm.}{=}$
$a^2 + ab + ab + b^2 = a^2 + 2ab + b^2$

(*ii*), (*iii*) ähnlich $\square$

Satz 3.5 *In jedem Körper* $(K, +, \cdot)$ *(und somit auch in* $(\mathbb{R}, +, \cdot)$*) gelten:*

(i) $\forall a, b \in K: \ ab = 0 \ \Leftrightarrow \ a = 0 \lor b = 0$

(ii) $\forall a \in K \ \forall b, c, d \in K \backslash \{0\}: \ \dfrac{a}{b} : \dfrac{c}{d} = \dfrac{a}{b} \cdot \dfrac{d}{c} = \dfrac{ad}{bc}$

(iii) $\forall a, c \in K \ \forall b, d \in K \backslash \{0\}: \ \dfrac{a}{b} \pm \dfrac{c}{d} = \dfrac{ad \pm bc}{bd}$

Beweis

(*i*) " $\Leftarrow$ ": folgt aus 3.3(i)

" $\Rightarrow$ ": Sei $ab = 0$.

1. Fall: $a \neq 0$
   $b = 1 \cdot b = (\frac{1}{a} \cdot a) \cdot b = \frac{1}{a}(a \cdot b) = \frac{1}{a} \cdot 0 = 0 \Rightarrow \ a = 0 \ \lor \ b = 0$
2. Fall: $a = 0 \Rightarrow \ a = 0 \ \lor \ b = 0$ $\square$

3.3 Das Induktionsaxiom und Folgerungen

Axiom 2 $(\mathbb{R}, +, \cdot)$ *enthält eine Teilmenge* $\mathbb{N}$*, welche das folgende Induktionsaxiom erfüllt: Ist* $M \subseteq \mathbb{N}$ *und gilt*

(i) $1 \in M$ *(wobei* 1 *das neutrale Element der Multiplikation in* $\mathbb{R}$ *ist),*

(ii) $n \in M \Rightarrow n + 1 \in M$,

dann ist $M = \mathbb{N}$. $\mathbb{N}$ *ist die Menge der natürlichen Zahlen.*

Satz 3.6 (Vollständige Induktion) *Für* $n \in \mathbb{N}$ *sei* $A(n)$ *eine Aussage. Es gelte*

(i) $A(1)$ *ist wahr. (Induktionsanfang)*

(ii) Für alle $n \in \mathbb{N}$ *gilt: Ist* $A(n)$ *wahr (Induktionsannahme, Induktionsvoraussetzung), dann ist auch* $A(n+1)$ *wahr (Induktionsschritt).*

Dann ist $A(n)$ *für alle* $n \in \mathbb{N}$ *wahr (Induktionsschluss).*

Beweis
Sei

$$M = \{n \in \mathbb{N} : A(n) \text{ ist wahr}\}$$

Wegen (i) gilt:

$$1 \in M$$

und wegen (ii) gilt:

$$n \in M \Rightarrow n + 1 \in M$$

Nach dem Induktionsaxiom folgt

$$M = \mathbb{N}.$$

□

Anwendung:
Sei

$$A(n) : 1 + 2 + \ldots + n = \frac{n \cdot (n+1)}{2}$$

Zeige, dass $A(n)$ für alle $n \in \mathbb{N}$ wahr ist.

- Induktionsanfang:

$$A(1) : 1 = \frac{2}{2} = \frac{1 \cdot (1+1)}{2} \text{ ist wahr.}$$

- Induktionsvoraussetzung (IV):

$$\begin{aligned} & A(n) \text{ ist wahr.} & & \\ \Rightarrow \quad & 1 + 2 + \ldots + n + (n+1) & \overset{\text{(IV)}}{=} & \frac{n \cdot (n+1)}{2} + n + 1 \\ & & = & \frac{n \cdot (n+1) + 2 \cdot (n+1)}{2} \\ & & = & \frac{(n+1) \cdot (n+2)}{2} \\ \Rightarrow \quad & A(n+1) \text{ ist wahr.} & & \end{aligned}$$

- Nach Satz 3.6 gilt der Induktionsschluss:

$$A(n) \text{ ist für alle } n \in \mathbb{N} \text{ wahr.}$$

□

Definition *(Summenzeichen, rekursive Definition)*

$$\begin{aligned} \sum_{k=1}^{1} a_k &:= a_1 \\ \sum_{k=1}^{n+1} a_k &:= \left(\sum_{k=1}^{n} a_k \right) + a_{n+1} \end{aligned}$$

Nach dem Induktionsprinzip gilt dann

$$\sum_{k=1}^{n} a_k = a_1 + a_2 + \ldots + a_n \quad \text{für alle } n \in \mathbb{N},$$

$$\sum_{k=0}^{n} a_k := \sum_{k=1}^{n+1} a_{k-1} \quad \text{für } n \in \mathbb{N}_0.$$

Damit können wir die eben bewiesene Formel kürzer schreiben:

Satz 3.7 (arithmetische Summenformel)

$$\forall n \in \mathbb{N} : \sum_{k=1}^{n} k = \frac{n \cdot (n+1)}{2}$$

Definition *($\mathbb{N}_0$, $\mathbb{Z}$, $\mathbb{Q}$ als Teilmengen von $\mathbb{R}$)*

$$\begin{array}{lcll} \mathbb{N}_0 & := & \mathbb{N} \cup \{0\} & \\ \mathbb{Z} & := & \mathbb{N}_0 \cup \{-n : n \in \mathbb{N}\} & \textit{(ganze Zahlen)} \\ \mathbb{Q} & := & \left\{\frac{n}{m} : n \in \mathbb{Z}, m \in \mathbb{N}\right\} & \textit{(rationale Zahlen)} \end{array}$$

Satz 3.8

(i) $(\mathbb{N}, +)$ *ist eine Halbgruppe.*

(ii) $(\mathbb{N}_0, +)$, $(\mathbb{N}, \cdot)$ *sind Monoide.*

(iii) $(\mathbb{Z}, +, \cdot)$ *ist ein kommutativer Ring.*

(iv) $(\mathbb{Q}, +, \cdot)$ *ist ein Körper.*

Definition *Für $x \in \mathbb{R}$ und $n \in \mathbb{N}$ sei*

$$\begin{array}{lcl} x^0 & := & 1, \\ x^1 & := & x, \\ x^{n+1} & := & x \cdot x^n. \end{array}$$

Ist $x \neq 0$, dann sei außerdem

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

Satz 3.9

(i) $\forall x \in \mathbb{R} \setminus \{0\}\ \forall m, n \in \mathbb{Z} :\ x^{m+n} = x^m \cdot x^n$

(ii) $\forall x \in \mathbb{R} \setminus \{0\}\ \forall m, n \in \mathbb{Z} :\ x^{m \cdot n} = (x^m)^n$

(iii) $\forall x, y \in \mathbb{R} \setminus \{0\}\ \forall n \in \mathbb{Z} :\ (x \cdot y)^n = x^n \cdot y^n$

Beweis
durch vollständige Induktion. □

Satz 3.10

(i) $\forall q \in \mathbb{R} \setminus \{1\} : \sum_{k=0}^{n} q^k = \frac{1 - q^{n+1}}{1 - q}$ ***(geometrische Summenformel)***

(ii) $\forall a, b \in \mathbb{R} : (a - b) \cdot \sum_{k=0}^{n-1} \left(a^{n-1-k} \cdot b^k\right) = a^n - b^n$

Beweis
durch vollständige Induktion. □

3.4 Die Anordnungsaxiome und Folgerungen

Axiom 3 $(\mathbb{R}, +, \cdot, <)$ *ist ein angeordneter Körper, das heißt, auf dem Körper* $\mathbb{R}$ *ist eine Relation* $<$ *(Kleiner-Relation) definiert, welche die folgenden Anordnungsaxiome erfüllt:*

(O1) $\forall a, b \in \mathbb{R} : a < b\ XOR\ a = b\ XOR\ b < a$ *(Trichotomie)*

(O2) $\forall a, b, c \in \mathbb{R} : a < b \wedge b < c \Rightarrow a < c$ *(Transitivität)*

(O3) $\forall a, b, c \in \mathbb{R} : a < b \Rightarrow a + c < b + c$ *(Verträglichkeit mit +)*

(O4) $\forall a, b, c \in \mathbb{R} : a < b \wedge 0 < c \Rightarrow a \cdot c < b \cdot c$ *(Verträglichkeit mit Multiplikation mit positiver Zahl)*

Bezeichnungen

$a > b$	$:\Leftrightarrow$	$b < a$	(Größer-Zeichen)
$a \leq b$	$:\Leftrightarrow$	$a < b \vee a = b$	(Kleiner-Gleich-Zeichen)
$a \geq b$	$:\Leftrightarrow$	$a > b \vee a = b$	(Größer-Gleich-Zeichen)
$\mathbb{R}^+$	$:=$	$\{a \in \mathbb{R} : a > 0\}$	(positive reelle Zahlen)
$\mathbb{R}_0^+$	$:=$	$\{a \in \mathbb{R} : a \geq 0\}$	(nicht negative reelle Zahlen)

Definition *Eine Relation* $R \subseteq M \times M$ *heißt Ordnungsrelation (auf* M*), falls gilt:*

(i) $\forall a \in M : aRa$ *(R ist reflexiv)*

(ii) $\forall a, b \in M : aRb \wedge bRa \Rightarrow a = b$ *(R ist antisymmetrsich)*

(iii) $\forall a, b, c \in M : aRb \wedge bRc \Rightarrow aRc$ *(R ist transitiv)*

Beispiel
$\leq$ ist Ordnungsrelation auf $\mathbb{R}$ (folgt aus $(O1), (O2)$)

Aus $(O1) - (O4)$ folgt außerdem

Satz 3.11 *Für alle $a, b, c \in \mathbb{R}$ gelten:*

(1) $a > 0 \Leftrightarrow -a < 0$

$a < 0 \Leftrightarrow -a > 0$

(2) $a \neq 0 \Rightarrow a^2 > 0$ *(insbesondere $1 = 1^2 > 0$)*

(3) $a > 0 \Rightarrow \dfrac{1}{a} > 0$

$a < 0 \Rightarrow \dfrac{1}{a} < 0$

(4) $(a < b \wedge c < 0) \Rightarrow a \cdot c > b \cdot c$

(5) $0 < a < b \Rightarrow \dfrac{1}{b} < \dfrac{1}{a}$

Anwendung finden diese Regeln beim Lösen von Ungleichungen.

Satz 3.12 (Bernoulli-Ungleichung) *Für $x \in \mathbb{R}$, $x \geq 1$ und $n \in \mathbb{N}_0$ gilt*

$$(1 + x)^n \geq 1 + n \cdot x$$

Beweis
durch vollständige Induktion unter Verwendung von $(O4)$. $\square$

Definition *(Intervalle)*

$$\begin{aligned}
[a, b] &:= \{x \in \mathbb{R} : a \leq x \leq b\} && \textit{(abgeschlossenes Intervall)} \\
]a, b[&:= \{x \in \mathbb{R} : a < x < b\} && \textit{(offenes Intervall)} \\
[a, b[&:= \{x \in \mathbb{R} : a \leq x < b\} && \textit{(rechts halboffenes Intervall)} \\
[a, \infty[&:= \{x \in \mathbb{R} : a \leq x\}
\end{aligned}$$

Analog definiert man $]a, b]$, $]-\infty, b]$, $]a, \infty[$, $]-\infty, \infty[= \mathbb{R}$

Definition *(Betrag von x)*

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

Satz 3.13 *Für alle $a, b \in \mathbb{R}$ gelten:*

(i) $|a| \geq 0 \wedge (|a| = 0 \Leftrightarrow a = 0)$

(ii) $|a| \geq a \wedge (|a| = a \Leftrightarrow a \geq 0)$

(iii) $|a \cdot b| = |a| \cdot |b|$

(iv) $|a+b| \leq |a| + |b|$ ***(Dreiecksungleichung)***

(v) $||a| - |b|| \leq |a-b|$ ***(Dreiecksungleichung nach unten)***

(vi) $|a \cdot b| \leq \frac{1}{2} \cdot (a^2 + b^2)$

Beweis
von (iv) (unter Voraussetzung, dass (ii) schon bewiesen ist):
1. Fall: $a+b \geq 0$
$\overset{(ii)}{\Rightarrow} |a+b| = a+b \leq |a| + |b|$
2. Fall: $a+b < 0$
$\Rightarrow |a+b| = -(a+b) = (-a) + (-b) \leq |-a| + |-b| = |a| + |b|$ □

Satz 3.14 $\forall a, b \in \mathbb{Q} : a < b \Rightarrow \exists c \in \mathbb{Q} : a < c < b$

Beweis
Wähle zum Beispiel $c = \frac{a+b}{2}$. □

3.5 Das Supremumsaxiom und Folgerungen

Definition *Sei $M \subseteq \mathbb{R}$.*

a) $a \in \mathbb{R}$ *heißt obere Schranke, falls* $\forall x \in M : x \leq a$.
$a \in \mathbb{R}$ *heißt untere Schranke, falls* $\forall x \in M : x \geq a$.

b) *M heißt nach oben (bzw. nach unten) beschränkt, falls M eine (und damit unendlich viele) obere (bzw. untere) Schranken besitzt. M heißt beschränkt, falls M nach oben und nach unten beschränkt ist.*

c) $a \in \mathbb{R}$ *heißt Maximum (bzw. Minimum) von M, falls* $a \in M \wedge \forall x \in M : x \leq a$ *(bzw. $x \geq a$).*
Bezeichnung: $a = \max M$ *(bzw.* $a = \min M$*).*

d) $a \in \mathbb{R}$ *heißt Supremum von M, falls a die kleinste obere Schranke von M ist, d.h.*

$$(\forall x \in M : x \leq a) \wedge (\forall \varepsilon > 0 \; \exists m \in M : a - \varepsilon < m \leq a) .$$

$a \in \mathbb{R}$ *heißt Infimum von M, falls a die größte untere Schranke von M ist.*
Bezeichnung: $a = \sup M$ *(bzw.* $a = \inf M$*).*

Bemerkung
Ein Maximum ist gleichzeitig Supremum, ein Minimum gleichzeitig Infimum, aber nicht umgekehrt.

Beispiel
Für $M = [0, 2[$ gilt: $0 = \min M = \inf M$, $2 = \sup M$. M hat kein Maximum, da $2 \notin M$.

Axiom 4 *Jede nichtleere, nach oben beschränkte Teilmenge von $\mathbb{R}$ besitzt ein Supremum.*

Satz 3.15 *Jede nichtleere, nach unten beschränkte Menge $M \subseteq \mathbb{R}$ besitzt ein Infimum.*

Beweisidee
Folgt aus Axiom 4, da $-M := \{-m : m \in M\}$ nichtleer und nach oben beschränkt ist. $\square$

Satz 3.16 $\mathbb{R}$ *ist ein archimedisch geordneter Körper, d.h., es gilt*

$$\forall x \in \mathbb{R}^+ \ \exists n \in \mathbb{N} : n > x. \tag{A}$$

Beweis
Angenommen, (A) gilt nicht. Dann ist $\mathbb{N}$ nach oben beschränkt und besitzt nach Axiom 4 ein Supremum $s = \sup \mathbb{N}$. $\Rightarrow \exists n \in \mathbb{N} : s - 1 < n \leq s \ \Rightarrow s < n + 1 \in \mathbb{N}$, ein Widerspruch zu $s = \sup \mathbb{N}$. $\square$

Mit $x = \frac{1}{\varepsilon}$, $\varepsilon > 0$, folgt aus (A):

Korollar 3.17

$$\forall \varepsilon > 0 \ \exists n_0 \in \mathbb{N} \ \forall n \geq n_0 : \frac{1}{n} < \varepsilon.$$

Satz 3.18 (Existenz der n-ten Wurzel)

$$\forall a \in \mathbb{R}_0^+ \ \forall n \in \mathbb{N} \ \exists! x \in \mathbb{R}_0^+ : x^n = a.$$

Bezeichnung: $x = \sqrt[n]{a} = a^{\frac{1}{n}}$, $\sqrt[2]{a} = \sqrt{a}$.

Der Satz folgt aus der Tatsache, dass $\{x \in \mathbb{R}_0^+ : x^n \leq a\}$ nichtleer und nach oben beschränkt ist und aus Axiom 4. Er kann aber auch mit fortgeschritteneren Methoden (konvergente Folgen) einfacher bewiesen werden.

Definition *Für $a \in \mathbb{R}^+$ und $m, n \in \mathbb{N}$ sei*

$$a^{\frac{m}{n}} := \sqrt[n]{a^m}, \quad a^{-\frac{m}{n}} := \frac{1}{a^{\frac{m}{n}}}.$$

Satz 3.19 *i)* $\forall x \in \mathbb{R}^+ \; \forall r, s \in \mathbb{Q} : x^{r+s} = x^r \cdot x^s$

ii) $\forall x \in \mathbb{R}^+ \; \forall r, s \in \mathbb{Q} : x^{r \cdot s} = (x^r)^s$

iii) $\forall x, y \in \mathbb{R}^+ \; \forall r \in \mathbb{Q} : (x \cdot y)^r = x^r \cdot y^r$

Satz 3.20 *$\sqrt{2}$ ist irrational, d.h. $\sqrt{2} \in \mathbb{R} \backslash \mathbb{Q}$.*

Zum Beweis benötigen wir folgende Vorbereitungen.

Definition

a) Sei $n \in \mathbb{Z}$. Die Zahl $m \in \mathbb{N}$ heißt Teiler von n (Bezeichnung: $m|n$), falls

$$\exists k \in \mathbb{Z} : n = km.$$

n heißt dann teilbar durch m und m teilt n.

b) Die Zahlen $p, q \in \mathbb{Z}$ heißen teilerfremd, falls kein $k \in \mathbb{Z} \setminus \{1\}$ existiert mit $k|p$ und $k|q$.

c) Die Zahl $n \in \mathbb{Z}$ heißt gerade, falls $2|n$. Anderenfalls heißt n ungerade.

Satz 3.21 *Für alle $n \in \mathbb{N}$ gilt:*

$$n \text{ gerade} \iff n^2 \text{ gerade.}$$

Beweis
" $\Rightarrow$ ": (direkter Beweis)
Sei $n \in \mathbb{N}$. Dann gilt:

$$\begin{aligned} n \text{ gerade} &\Rightarrow \exists k \in \mathbb{N} : n = 2k \\ &\Rightarrow n^2 = 4k^2 = 2(2k^2) \\ &\Rightarrow n^2 \text{ gerade.} \end{aligned}$$

" $\Leftarrow$ ": (indirekter Beweis)
Sei $n \in \mathbb{N}$. Dann gilt:

$$\begin{aligned} n \text{ ungerade} &\Rightarrow \exists k \in \mathbb{N}_0 : n = 2k + 1 \\ &\Rightarrow n^2 = (2k+1)^2 = 4k^2 + 4k + 1 = 2(2k^2 + 2k) + 1 \\ &\Rightarrow n^2 \text{ ungerade.} \end{aligned}$$

Hierbei gilt $k = \min\{p \in \mathbb{N} : 2p > n\} - 1$ (Beweis durch vollständige Induktion). □

Beweis von Satz 3.20 (Widerspruchsbeweis)
Angenommen, $\sqrt{2} \in \mathbb{Q}$.
$\Rightarrow \; \exists n, m \in \mathbb{N} : \sqrt{2} = \frac{n}{m}$ und n, m sind teilerfremd

$\Rightarrow\ 2m^2 = n^2$

$\Rightarrow\ n^2$ gerade

$\overset{3.21}{\Rightarrow}\ n$ gerade

$\Rightarrow\ \exists k \in \mathbb{N} : n = 2k$

$\Rightarrow\ 2m^2 = n^2 = 4k^2$

$\Rightarrow\ m^2 = 2k^2$

$\Rightarrow\ m^2$ gerade

$\overset{3.21}{\Rightarrow}\ m$ gerade.

Also gilt $2|n$ und $2|m$, ein Widerspruch, denn n, m sind teilerfremd. $\square$

Satz 3.22 *Es seien $a \in \mathbb{R}\backslash\{0\},\ b, c \in \mathbb{R}$ und $b^2 \geq 4ac$. Dann gilt:*

$$ax^2 + bx + c = 0 \iff x_{1/2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.$$

Beweis

Unter den Voraussetzungen des Satzes gilt:

$$\begin{aligned}
& ax^2 + bx + c = 0 \quad \mid : a \neq 0 \\
\iff\ & x^2 + \frac{b}{a}x + \frac{c}{a} = 0 \\
\iff\ & x^2 + 2\frac{b}{2a}x + \left(\frac{b}{2a}\right)^2 = \left(\frac{b}{2a}\right)^2 - \frac{c}{a} \\
\iff\ & \left(x + \frac{b}{2a}\right)^2 = \frac{1}{4a^2}(b^2 - 4ac) \geq 0 \\
\iff\ & \left(x + \frac{b}{2a} = \frac{1}{2a}\sqrt{b^2 - 4ac}\right) \vee \left(x + \frac{b}{2a} = -\frac{1}{2a}\sqrt{b^2 - 4ac}\right) \\
\iff\ & x_{1/2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}\ .
\end{aligned}$$

$\square$

4 Zahlentheorie

4.1 Kombinatorik

Definition *Es seien $M \subseteq \mathbb{N}$ und $r \in \mathbb{N}$.*

a) *Jedes r-Tupel $(a_1, \dots, a_r) \in M^r$ heißt eine r-Permutation (aus M) mit Wiederholung.*

   Anwendungen:

   1. *Ziehen von r Kugeln aus einer Urne mit Zurücklegen. Die Reihenfolge der Ziehung ist von Bedeutung.*
   2. *Verteilen von r unterscheidbaren Kugeln auf so viele Zellen, wie M Elemente hat, mit Mehrfachbesetzung von Zellen.*

b) *Jedes r-Tupel $(a_1, \dots, a_r) \in M^r$ mit $a_i \neq a_j$ für $i \neq j$ heißt eine r-Permutation ohne Wiederholung.*

   Anwendungen:

   1. *Ziehen ohne Zurücklegen, Reihenfolge ist von Bedeutung.*
   2. *Verteilen ohne Mehrfachbesetzung, Kugeln sind unterscheidbar.*

c) *Jede Teilmenge von M mit r Elementen $\{a_1, \dots, a_r\} \in \mathcal{P}(M)$ heißt r-Kombination (aus M) ohne Wiederholung.*

   Anwendungen:

   1. *Ziehen ohne Zurücklegen, Reihenfolge ist egal.*
   2. *Verteilen ohne Mehrfachbesetzung, Kugeln sind ununterscheidbar.*

d) *Jedes r-Tupel $(a_1, \dots, a_r) \in M^r$ mit $a_1 \leq a_2 \leq \dots \leq a_r$ heißt eine r-Kombination mit Wiederholung.*

   Anwendungen:

   1. *Ziehen mit Zurücklegen, die Reihenfolge ist egal.*
   2. *Verteilen mit Mehrfachbesetzung, Kugeln sind ununterscheidbar.*

Definition

a) *(Produktzeichen)*

$$\prod_{k=1}^{1} a_k := a_1, \quad \prod_{k=1}^{n+1} a_k := \left(\prod_{k=1}^{n} a_k\right) \cdot a_{n+1}, \; n \in \mathbb{N}. \text{ Also } \prod_{k=1}^{n} a_k = a_1 \cdot a_2 \cdots a_n,$$

$$\prod_{k=0}^{n} a_k := \prod_{k=1}^{n+1} a_{k-1}.$$

b) (Fakultät)

$$0! := 1, \quad n! := \prod_{k=1}^{n} k,\ n \in \mathbb{N}.$$

Also $n! = 1 \cdot 2 \cdot ... \cdot n$.

c) (Binomialkoeffizient)

$$\binom{\alpha}{0} := 1, \qquad \binom{\alpha}{k+1} := \frac{\alpha - k}{k+1}\binom{\alpha}{k},\ k \in \mathbb{N}_0$$

Also

$$\binom{\alpha}{k} := \frac{\alpha(\alpha-1)\cdot ... \cdot(\alpha-k+1)}{k!}\ ,\ \textit{falls } k \in \mathbb{N},$$

$$\binom{n}{k} := \frac{n!}{k!(n-k)!}\ ,\ \textit{falls } n, k \in \mathbb{N}_0 \textit{ und } k \leq n.$$

Satz 4.1 *Es sei* $M \subseteq \mathbb{N}$ *mit* n *Elementen und* $r \in \mathbb{N}$ *mit* $r \leq n$. *Dann gibt es genau*

a) n^r *mögliche* r*-Permutationen mit Wiederholung,*

b) $\binom{n}{r} r! = n\cdot(n-1)\cdots(n-r+1)$ *mögliche* r*-Permutationen ohne Wiederholung, insbesondere* $n!$ *mögliche* n*-Permutationen ohne Wiederholung.*

c) $\binom{n}{r}$ *mögliche* r*-Kombinationen ohne Wiederholung.*

d) $\binom{n+r-1}{r}$ *mögliche* r*-Kombinationen mit Wiederholung.*

Beweisskizze
$a), b)$ zeigt man durch vollständige Induktion nach n und r.
$c)$ folgt aus $b)$.
$d)$ Sei o. B. d. A. (ohne Beschränkung der Allgemeinheit) $M = \{1, \ldots, n\}$. Es sei außerdem $N := \{1, \ldots, n+r-1\}$ und

$$f\left((a_1, \ldots, a_r)\right) := \{a_1, a_2+1, a_3+2, \ldots, a_r+r-1\},$$

für alle $(a_1, \ldots, a_r) \in M^r$ mit $a_1 \leq a_2 \leq \cdots \leq a_r$.
Dann gibt es zu jeder r-Kombination $\{b_1, \ldots, b_r\}$ aus N ohne Wiederholung genau eine r-Kombination $(a_1, \ldots, a_r)$ aus M mit Wiederholung, so dass

$$f\left((a_1, \ldots, a_r)\right) = \{b_1, \ldots, b_r\}.$$

Daraus folgt die Behauptung unter Verwendung von $c)$.

Satz 4.2 *Es seien* $n, k \in \mathbb{N}_0,\ k \leq n$ *und* $\alpha, a, b \in \mathbb{R}$. *Dann gilt:*

a) $\binom{n}{0} = 1$

b) $\binom{n}{k} = \binom{n}{n-k}$

c) $\binom{\alpha}{k} + \binom{\alpha}{k+1} = \binom{\alpha+1}{k+1}$

d) (Binomischer Satz)

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

Bemerkung

Gemäß Satz 4.2 ergeben sich die Werte von $\binom{n}{k}$ aus dem Pascalschen Dreieck:

$$\begin{array}{ccccccccc}
 & & & & \binom{0}{0} & & & & \\
 & & & \binom{1}{0} & & \binom{1}{1} & & & \\
 & & \binom{2}{0} & & \binom{2}{1} & & \binom{2}{2} & & \\
 & \binom{3}{0} & & \binom{3}{1} & & \binom{3}{2} & & \binom{3}{3} & \\
\binom{4}{0} & & \binom{4}{1} & & \binom{4}{2} & & \binom{4}{3} & & \binom{4}{4} \\
 & & & & \vdots & & & &
\end{array}
\qquad
\begin{array}{ccccccccc}
 & & & & 1 & & & & \\
 & & & 1 & & 1 & & & \\
 & & 1 & & 2 & & 1 & & \\
 & 1 & & 3 & & 3 & & 1 & \\
1 & & 4 & & 6 & & 4 & & 1 \\
 & & & & \vdots & & & &
\end{array}$$

Somit gilt z. B. $(a+b)^4 = a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + b^4$.

Beweis von Satz 4.2
durch vollständige Induktion, z. B. bei d):
Für $n = 0$:

$$\sum_{k=0}^{0} \binom{0}{k} a^{0-k}\, b^k = \binom{0}{0} a^0\, b^0 = 1 = (a+b)^0.$$

Für $n = 1$:

$$\sum_{k=0}^{1} \binom{1}{k} a^{1-k}\, b^k = \binom{1}{0} ab^0 + \binom{1}{1} a^0\, b = a + b = (a+b)^1.$$

$n \to n+1$:

$$\begin{aligned}
(a+b)^{n+1} &= (a+b)(a+b)^n \\
&= (a+b)\sum_{k=0}^{n}\binom{n}{k}a^{n-k}\,b^k \qquad \text{(nach Induktionsannahme)} \\
&= \sum_{k=0}^{n}\binom{n}{k}a^{n+1-k}\,b^k + \sum_{k=0}^{n}\binom{n}{k}a^{n-k}\,b^{k+1} \\
&= \sum_{k=0}^{n}\binom{n}{k}a^{n+1-k}\,b^k + \sum_{k=1}^{n+1}\binom{n}{k-1}a^{n-(k-1)}\,b^k \\
&= \binom{n}{0}a^{n+1}\,b^0 + \sum_{k=1}^{n}\left(\binom{n}{k}+\binom{n}{k-1}\right)a^{n+1-k}\,b^k + \binom{n}{n}a^0\,b^{n+1} \\
&\overset{a),b),c)}{=} \binom{n+1}{0}a^{n+1}\,b^0 + \sum_{k=1}^{n}\binom{n+1}{k}a^{n+1-k}\,b^k + \binom{n+1}{n+1}a^0\,b^{n+1} \\
&= \sum_{k=0}^{n+1}\binom{n+1}{k}a^{n+1-k}\,b^k.
\end{aligned}$$

□

Korollar 4.3

$$\sum_{k=0}^{n}\binom{n}{k} = 2^n$$

Bemerkung

Wegen Satz 4.1 c) folgt aus Korollar 4.3:
Die Potenzmenge einer n-elementigen Menge hat genau 2^n Elemente.
Diese Aussage kann allerdings auch direkt mit vollständiger Induktion bewiesen werden.

4.2 Teilbarkeit und Primzahlen

Definition

a) Für $a, b \in \mathbb{Z}$, $(a,b) \neq (0,0)$ heißt

$$\operatorname{ggT}(a,b) := \max\{d \in \mathbb{N} : d|a \wedge d|b\}$$

der größte gemeinsame Teiler von a und b.

b) Für $m, n \in \mathbb{N}$ heißt

$$\operatorname{kgV}(m,n) := \min\{q \in \mathbb{N} : m|q \wedge n|q\}$$

das kleinste gemeinsame Vielfache von m und n.

Satz 4.4 (Teilen mit Rest) *Seien $n, m \in \mathbb{N}$ mit $n > m$. Dann gibt es eindeutig bestimmte Zahlen $q \in \mathbb{N},\ r \in \mathbb{N}_0$ mit*

$$n = qm + r \quad \wedge \quad r \leq m - 1.$$

Dabei gilt $\mathrm{ggT}(n, m) = \mathrm{ggT}(m, r)$.

Beispiel
$30 = 1 \cdot 24 + 6$ und $\mathrm{ggT}(30, 24) = \mathrm{ggT}(24, 6) = 6$.

Hilfssatz 4.5

$$\forall n, m, \alpha, \beta \in \mathbb{Z}\ \forall d \in \mathbb{N} : d|n \wedge d|m \;\Rightarrow\; d|\alpha n + \beta m.$$

Beweis

$$\begin{aligned} d|n \wedge d|m \;&\Rightarrow\; \exists k_1, k_2 \in \mathbb{Z} : n = k_1 d \wedge m = k_2 d \\ &\Rightarrow\; \alpha n + \beta m = (\alpha k_1 + \beta k_2) d \\ &\Rightarrow\; d|\alpha n + \beta m. \end{aligned}$$

□

Beweisskizze für Satz 4.4
Existenz: Sei $q := \min\{k \in \mathbb{N} : km > n\} - 1 = \max\{k \in \mathbb{N} : km \leq n\}$ und $r := n - qm$. Dann liefern q und r die gewünschte Formel.
Eindeutigkeit:

$$\begin{aligned} & qm + r = q'm + r' \;\wedge\; r, r' \leq m - 1 \\ \Rightarrow\; & m|q - q'| = |r' - r| \;\wedge\; |r' - r| \leq m - 1 \\ \Rightarrow\; & r = r' \;\wedge\; q = q'. \end{aligned}$$

ggT-Formel:

$$d|n \wedge d|m \;\overset{4.5}{\Rightarrow}\; d|n - qm \;\Rightarrow\; d|r$$
$$d|m \wedge d|r \;\overset{4.5}{\Rightarrow}\; d|qm + r \;\Rightarrow\; d|n.$$

Satz 4.6 (Euklidischer Algorithmus zur Berechnung des ggTs) *Seien $n, m \in \mathbb{N}$ mit $n > m$. Dann existieren eindeutig bestimmte Zahlen N, $q_1, ..., q_{N+1}, r_1, ..., r_N \in \mathbb{N}$ mit*

$n = q_1 m + r_1 \wedge r_1 \leq m - 1,$
$m = q_2 r_1 + r_2 \wedge r_2 \leq r_1 - 1 \leq m - 2,$
$r_1 = q_3 r_2 + r_3 \wedge r_3 \leq r_2 - 1,$
$\vdots$
$r_{N-2} = q_N r_{N-1} + r_N \wedge r_N \leq r_{N-1} - 1,$
$r_{N-1} = q_{N+1} r_N + 0,$
und es gilt $r_N = ggT(m, n)$.

Beweisskizze
Beweis durch vollständige Induktion.
Wegen $0 < r_j < r_{j-1} < r_{j-2} < ...$ bricht das Verfahren nach endlich vielen Schritten ab.
Wegen Satz 4.4 gilt:
$\mathrm{ggT}(n, m) = \mathrm{ggT}(n, r_1) = \mathrm{ggT}(r_1, r_2) = ... = \mathrm{ggT}(r_{N-1}, r_N) = r_N$

Beispiel
$30 = 1 \cdot 24 + 6$, $24 = 4 \cdot 6 + 0$, also $6 = \mathrm{ggT}(30, 24)$.

Definition *$p \in \mathbb{N} \setminus \{1\}$ heißt Primzahl, falls 1 und p die einzigen Teiler von p sind.*

Hilfssatz 4.7 *Seien $m, n \in \mathbb{N}$ und p eine Primzahl. Dann gilt: $p|mn \Rightarrow p|m \vee p|n$.*

Beweis
Angenommen: $p|mn \ \wedge \ \neg(p|m)$
$\Rightarrow p|pn \ \wedge \ p|mn = kp$ für ein $k \in \mathbb{Z}$
$\overset{4.6}{\Rightarrow} \mathrm{ggT}(mn, pn) = \mathrm{ggT}(kp, np) = p\,\mathrm{ggT}(k, n)$
$\Rightarrow p|\mathrm{ggT}(mn, pn) = n\,\mathrm{ggT}(m, p) = n$ (da p Primzahl)
$\Rightarrow p|n$. □

Satz 4.8 (Fundamentalsatz der Arithmetik) *Jede natürliche Zahl $n \geq 2$ lässt sich als Produkt von Primzahlen darstellen, wobei die Darstellung bis auf die Reihenfolge der Primfaktoren eindeutig ist.*

Beweisskizze
Existenz: durch vollständige Induktion

$n = 2: \ 2 = 2 \ \wedge$ 2 ist Primzahl

$n \to n+1$: 1.Fall: $n+1$ ist Primzahl
$\Rightarrow n+1 = n+1$
2.Fall: $n+1$ ist keine Primzahl
$\Rightarrow \exists k, l \in \mathbb{N} : n+1 = kl \ \wedge \ k, l < n+1$

Hier machen wir die Induktionsannahme, dass die Behauptung bereits für alle natürlichen Zahlen $\leq n$ gezeigt ist.

Aus dieser Annahme folgt:

$n + 1 = \underbrace{p_1 \cdot \ldots \cdot p_j}_{=k} \cdot \underbrace{q_1 \cdot \ldots \cdot q_i}_{=l}$, wobei $p_1, \ldots, p_j,\ q_1, \ldots, q_i$ Primzahlen sind.

$\Rightarrow$ Behauptung für $n+1$.

Eindeutigkeit:
Angenommen $n = p_1 \cdot \ldots \cdot p_s = q_1 \cdot \ldots \cdot q_t$ (mit $p_1, \ldots, p_s, q_1, \ldots, q_t$ Primzahlen).
Zeige die behauptete Eindeutigkeit durch Induktion nach s unter Verwendung von Hilfssatz 4.7.

Korollar 4.9 *Sei* $n = p_1^{r_1} \cdot \ldots \cdot p_k^{r_k}$ *und* $m = p_1^{s_1} \cdot \ldots \cdot p_k^{s_k}$*, wobei* $p_1, \ldots, p_k$ *Primzahlen und* $r_1, \ldots, r_k \in \mathbb{N}_0$ *seien. Dann gilt:*

$$ggT(n,m) = p_1^{\min\{r_1,s_1\}} \cdot \ldots \cdot p_k^{\min\{r_k,s_k\}},$$

$$kgV(n,m) = p_1^{\max\{r_1,s_1\}} \cdot \ldots \cdot p_k^{\max\{r_k,s_k\}}.$$

Satz 4.10 (Satz von Euklid) *Es gibt unendlich viele Primzahlen.*

Beweis
Angenommen, es gäbe es nur endlich viele Primzahlen $p_1, \ldots, p_n$. Dann gilt nach Satz 4.8:
$p_1 \cdot \ldots \cdot p_n + 1 = q_1 \cdot \ldots \cdot q_n$, wobei $q_1, \ldots, q_n$ Primzahlen sind.
Da alle $p_1, \ldots, p_n$ Primzahlen sind, gilt:
$\exists\, j \in \{1, \ldots, n\} : \ q_1 = p_j$
$\Rightarrow q_1 | p_1 \cdot \ldots \cdot p_n \ \wedge \ q_1 | q_1 \cdot \ldots \cdot q_k$
$\overset{HS4.5}{\Rightarrow} \ q_1 | q_1 \cdot \ldots \cdot q_k - p_1 \cdot \ldots \cdot p_n$
$\Rightarrow q_1 | 1$ Widerspruch zu q_1 ist Primzahl. $\square$

Bemerkung
Anwendungen von Teilbarkeit und Primzahlen gilt es zum Beispiel beim modularen Rechnen und in der Kryptographie.

4.3 Stellenwertsysteme

Üblicherweise stellen wir natürliche Zahlen mit Hilfe der zehn Ziffern $0, 1, \ldots, 9$ dar, zum Beispiel $108 = 1 \cdot 10^2 + 0 \cdot 10^1 + 8 \cdot 10^0$. Das geht auch mit weniger oder mehr Ziffern.

Definition *Sei* $n \in \mathbb{N}$ *und* $g \in \mathbb{N} \setminus \{1\}$*. Die Darstellung*

$$n = (a_N a_{N-1} a_{N-2} \ldots a_0)_g := \sum_{j=0}^{N} a_j g^j$$

mit $N \in \mathbb{N}$ *und* $a_0, \ldots, a_N \in Z_g$*, wobei* Z_g*eine Menge mit* g *Elementen ist, heißt g-adische Entwicklung von n. g heißt Ziffernbasis und die Elemente aus* Z_g *heißen Ziffern.*

Beispiele
Dezimalsystem: $g = 10, \ Z_g = \{0, 1, ..., 9\}$
$108 = (108)_{10}$

Dualsystem: $g = 2, \ Z_g = \{0, 1\}$ (Zahlendarstellung im Computer)
$1011_2 = (1 \cdot 2^3 + 0 \cdot 2^2 + 1 \cdot 2^1 + 1 \cdot 2^0)_{10} = (8 + 0 + 2 + 1)_{10} = 11_{10} = 11$

Hexadezimalsystem: $g = 16, \ Z_g = \{0, 1, \ldots, 9, A, B, C, D, E, F\}$
$D2A_{16} = (D \cdot (16)^2 + 2 \cdot (16)^1 + A \cdot (16)^0)_{10} = (13 \cdot 256 + 2 \cdot 16 + 10 \cdot 1)_{10} = 3370_{10} = 3370$

Satz 4.11 *Sei $g \in \mathbb{N} \setminus \{1\}$ vorgegeben. Dann besitzt jedes $n \in \mathbb{N}$ eine eindeutige g-adische Entwicklung, d.h., es existieren eindeutig bestimmte $N \in \mathbb{N}$ und $a_1, \ldots, a_N \in Z_g$, so dass $n = (a_N, a_{N-1}, \ldots, a_0)_g \ \wedge \ a \neq 0$ gilt.*

Die g-adische Entwicklung kann zum Beispiel folgendermaßen berechnet werden:

1. Schritt: Bestimme $N \in \mathbb{N}$ mit $g^N \leq n < g^{N+1}$.

2. Schritt: Teile n durch g^n mit Rest, d.h., bestimme $n = a_N - g^N + r_N$ mit $0 \leq r_N < g^N$ und $0 < a_N \leq g - 1$.

3. Schritt: Teile r_N durch g^{N-1} mit Rest, d.h., bestimme

$$\begin{array}{rcl} r_N &=& a_{N-1}g^{N-1} + r_{N-1} \quad \text{mit } 0 \leq r_{N-1} < g^{N-1},\ 0 \leq a_{N-1} \leq g-1, \\ & \vdots & \\ r_2 &=& a_1 g^1 + r_1 \quad \text{mit } 0 \leq r_1 < g,\ 0 \leq a_1 \leq g-1, \\ r_1 &=& a_0 g^0 \ = \ a_0. \end{array}$$

Dann ist $n = (a_{N,} a_{N-1}, \ldots, a_0)_g$.

5 Reelle Funktionen

5.1 Allgemeine Grundbegriffe

Definition

a) Zwei Funktionen $f_1 : M_1 \to N_1$ und $f_2 : M_2 \to N_2$ heißen gleich, falls $M_1 = M_2$ und $f_1(x) = f_2(x)$ für alle $x \in M_1$ gilt.

b) Seien $f_1 : M_1 \to N_1$ und $f_2 : M_2 \to N_2$ zwei Funktionen, wobei $M_1 \subseteq M_2$ und $f_1(x) = f_2(x)$ für alle $x \in M_1$ gilt. Dann heißt f_2 die Fortsetzung von f_1 auf M_2 und f_1 die Einschränkung von f_2 auf M_1, in Zeichen : $f_1 = f_2 |_{M_1}$.

Definition *Eine Funktion $f : M \to N$ heißt*

a) injektiv, falls $\forall x, y \in M : \ f(x) = f(y) \ \Rightarrow \ x = y$

b) surjektiv, falls $\forall y \in N \ \exists x \in M : \ y = f(x)$

c) bijektiv, falls f injektiv und surjektiv ist.

Beispiel
$f_1 : \mathbb{R} \to \mathbb{R}, x \mapsto x^2$ ist weder injektiv noch surjektiv.
$f_2 : \mathbb{R}_0^+ \to \mathbb{R}, x \mapsto x^2$ ist injektiv, aber nicht surjektiv.
$f_3 : \mathbb{R} \to \mathbb{R}_0^+, x \mapsto x^2$ ist surjektiv, aber nicht injektiv.
$f_4 : \mathbb{R}_0^+ \to \mathbb{R}_0^+, x \mapsto x^2$ ist bijektiv.
Außerdem gilt:
$f_2 \neq f_1, \ f_2 = f_1 |_{\mathbb{R}_0^+}, \ f_2 = f_4$

Definition *Seien $f : A \to B$ und $g : B' \to C$ Funktionen, wobei $B \subseteq B'$ sei. Dann heißt*

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

Verkettung oder Hintereinanderausführung von f und g.

Bemerkung
Die Verkettung von Funktionen ist assoziativ, denn es gilt

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

aber im Allgemeinen nicht kommutativ, denn für $f(x) = x^2 + 1, \ g(x) = x - 1$ ist $f(g(x)) = (x-1)^2 + 1 = x^2 - 2x + 2$, aber $g(f(x)) = (x^2 + 1) - 1 = x^2$.

Satz 5.1 *Ist eine Funktion $f : M \to N$ bijektiv, dann existiert genau eine Funktion $f^{-1} : N \to M$, die sogenannte Umkehrfunktion oder inverse Abbildung von f, mit*

$$f^{-1} \circ f = Id_M \quad \wedge \quad f \circ f^{-1} = Id_N,$$

wobei $Id_M : M \to M, x \mapsto x,\ Id_N : N \to N, x \mapsto x$ die sogenannten identischen Abbildungen oder Identitäten auf M bzw. N sind.

Definition *Unter einer Menge $\{f, g, \ldots\}$ von Funktionen versteht man die Menge ihrer Funktionsgraphen, wobei jede Funktion mit ihrem Graph identifiziert wird, d.h. $\{f, g, \ldots\} := \{G(f), G(g), \ldots\}$.*

Satz 5.2 *Sei S_M die Menge aller bijektiven Abbildungen (Funktionen) auf einer nichtleeren Menge M. Dann definiert die Verkettung $\circ$ eine Verknüpfung auf S_M und $(S_M, \circ)$ ist eine Gruppe.*

Definition *Seien $M \neq \emptyset,\ \alpha \in \mathbb{R}$ und $f, g : M \to \mathbb{R}$ reellwertige Funktionen. Dann definiert man $f + g, \alpha g$ und fg durch*

$$\begin{aligned} f + g : M \to \mathbb{R},\ x &\mapsto f(x) + g(x), \\ \alpha f : M \to \mathbb{R},\ x &\mapsto \alpha f(x), \\ fg : M \to \mathbb{R},\ x &\mapsto f(x) \cdot g(x). \end{aligned}$$

Definition *Sei I ein Intervall und $f : I \to \mathbb{R}$ eine reelle Funktion. Dann heißt f*

a) monoton steigend, falls $\forall x, y \in I : x < y \;\Rightarrow\; f(x) \leq f(y)$,

b) streng monoton steigend, falls $\forall x, y \in I : x < y \;\Rightarrow\; f(x) < f(y)$,

c) monoton fallend, falls $\forall x, y \in I : x < y \;\Rightarrow\; f(x) \geq f(y)$,

d) streng monoton fallend, falls $\forall x, y \in I : x < y \;\Rightarrow\; f(x) > f(y)$,

Bemerkung
Streng monoton steigende und streng monoton fallende Funktionen sind injektiv.

5.2 Polynomfunktionen

Definition

a) Sei $(R, +, \cdot)$ ein kommutativer Ring. Ein Term der Form $\sum\limits_{k=0}^{n} a_k x^k$ mit $a_k \in R$ für $k = 0, \ldots, n$ heißt Polynom und eine Funktion $p : R \to R,\ x \mapsto \sum\limits_{k=0}^{n} a_k x^k$ heißt Polynomfunktion.

b) Ist $a_n \neq 0$, dann heißt n der Grad des Polynoms. Sind alle Koeffizienten $a_k = 0$, dann heißt das Polynom Nullpolynom und sein Grad wird gleich -1 gesetzt. Polynome vom Grad < 1 heißen konstant, vom Grad 1 linear und vom Grad 2 quadratisch.

c) Gilt $a_0, \ldots, a_n \in \mathbb{R}$, dann heißt das Polynom reell und die zugehörige Polynomfunktion ganzrationale reelle Funktion oder reelle Polynomfunktion.

Satz 5.3 (Koeffizientenvergleich) *Seien $a_0, \ldots a_n, b_0, \ldots, b_m \in \mathbb{R}$ mit $a_n, b_m \neq 0$,*

$$p(x) = \sum_{k=0}^{n} a_k x^k \text{ und } q(x) = \sum_{k=0}^{m} b_k x^k.$$

Gilt $p(x) = q(x)$ für alle $x \in \mathbb{R}$, dann folgt $m = n$ und $a_k = b_k$ für alle $k = 0, \ldots, n$.

Aus diesem Satz folgt, dass die zugehörigen Polynomfunktionen zweier verschiedener reeller Polynome, d.h. zweier reeller Polynome mit verschiedenen Koeffizienten, nicht gleich sein können. Daher werden wir reelle Polynome mit reellen Polynomfunktionen identifizieren und zu reellen Polynomfunktionen auch reelle Polynome sagen. Sind dagegen die Koeffizienten $a_0, \ldots, a_n, b_0, \ldots, b_m$ Elemente eines endlichen Körpers, dann gilt die Behauptung des Satzes 5.3 nicht und wir müssen zwischen Polynomen und Polynomfunktionen unterscheiden.

Satz 5.4 *Seien p, q reelle Polynome mit $p(x) = \sum\limits_{k=0}^{n} a_k x^k$, $q(x) = \sum\limits_{k=0}^{m} b_k x^k$, $n > m$ und seien $\alpha, \beta \in \mathbb{R}$. Dann gilt*

$$\begin{aligned}
(\alpha p + \beta q)(x) &= \alpha a_0 + \beta b_0 + (\alpha a_1 + \beta b_1)x + \ldots + (\alpha a_m + \beta b_m)x^m + \ldots + \alpha a_n x^n \;, \\
(p \cdot q)(x) &= a_0 b_0 + (a_1 b_0 + a_0 b_1)x + (a_2 b_0 + a_1 b_1 + a_0 b_2)x^2 + \ldots + a_n b_m x^{n+m} \\
&= \sum_{k=0}^{m+n} c_k x^k \\
\text{mit } c_k &= \sum_{j=0}^{k} a_j b_{k-j} \text{ , falls man } \; a_{n+1} = \ldots = a_{n+m} := 0 \text{ und} \\
& \qquad\qquad b_{m+1} = \ldots = b_{m+n} := 0 \text{ setzt.}
\end{aligned}$$

Die Menge $\mathbb{R}[x]$ aller reellen Polynome bildet mit den Verknüpfungen $+$ und $\cdot$ einen kommutativen Ring.
Dabei gilt für alle $p \neq 0$ und $q \neq 0$:
$Grad(p + q) \leq max\{Grad(p), Grad(q)\}$ mit Gleichheit für $Grad(p) \neq Grad(q)$ und $Grad(p \cdot q) = Grad(p) + Grad(q)$.

Definition

a) *Ein Polynom* p_2 *heißt Teiler eines Polynoms* p_1 *(Bezeichnung* $p_2 \mid p_1$*), falls ein Polynom* q *existiert mit* $p_1 = q \cdot p_2$.

b) *Zwei Polynome* p_1, p_2 *heißen teilerfremd, falls kein Polynom* p *mit* $Grad(p) \geq 1$ *existiert mit* $p \mid p_1$ *und* $p \mid p_2$.

c) *Für zwei Polynome* a, b*, wobei* $(a, b) \neq (0, 0)$*, heißt ein Polynom* d *ein größter gemeinsamer Teiler von* a *und* b *(Bezeichnung* $d \in ggT(a, b)$*), falls*

$$\begin{array}{ll} (i) & d \mid a \;\wedge\; d \mid b \\ (ii) & q \mid a \;\wedge\; q \mid b \;\Rightarrow\; Grad(q) \leq Grad(d). \end{array}$$

d) *Ein Polynom* p *mit* $Grad(p) \geq 1$ *heißt Primpolynom, falls gilt*

$$p = p_1 \cdot p_2 \quad \Rightarrow \quad Grad(p_1) = 0 \;\vee\; Grad(p_2) = 0.$$

Satz 5.5 a) *Seien* p_1, p_2 *reelle Polynome mit* $Grad(p_1) \geq Grad(p_2) \geq 1$*. Dann gibt es eindeutig bestimmte reelle Polynome* q, r *mit*

$$p_1 = qp_2 + r \quad \wedge \quad Grad(r) < Grad(p_1)$$

Dabei gilt $ggT(p_1, p_2) = ggT(p_2, r)$.

b) *Zu je zwei reellen Polynomen* a, b *mit* $(a, b) \neq (0, 0)$ *gibt es größte gemeinsame Teiler. Aus* $d_1, d_2 \in ggT(a, b)$ *folgt* $d_1 = cd_2$ *für ein* $c \in \mathbb{R} \setminus \{0\}$*. Die größten gemeinsamen Teiler* $ggT(a, b)$ *können mit Hilfe des euklidischen Algorithmus (mit reellen Polynomen statt mit natürlichen Zahlen) berechnet werden.*

Bemerkung
Das Teilen von reellen Polynomen kann man mit Hilfe des Verfahrens der Polynomdivision durchführen, z.B.

$$\begin{array}{l} (x^2+1) : (x+1) = x - 1 + \frac{2}{x+1} \\ \underline{-(x^2+x)} \\ \quad -x+1 \\ \underline{\;-(-x-1)} \\ \qquad\quad 2 \end{array}$$

Satz 5.6 *Jedes reelle Polynom vom Grad* ≥ 1 *lässt sich als Produkt von Primpolynomen darstellen, wobei die Darstellung bis auf die Reihenfolge der Primfaktoren eindeutig ist.*

Definition *Sei p ein Polynom. λ heißt Nullstelle von p, falls $p(\lambda) = 0$ ist.*

Satz 5.7 *Sei p ein reelles Polynom vom Grad ≥ 1. Dann gilt :*

$$\lambda \textit{ ist Nullstelle von } p \;\Leftrightarrow\; x - \lambda \;\textit{ ist Teiler von } p$$

Beweis
"$\Rightarrow$" : Nach Satz 5.5 a) existieren reelle Polynome q, r mit

$$p(x) = (x - \lambda)q(x) + r(x) \quad \text{und} \quad Grad(r) < 1,$$

also r ist konstant. Somit gilt:

$$\begin{aligned} 0 &= p(\lambda) = 0 \cdot q(\lambda) + r(\lambda) \;\text{ und damit} \\ r(x) &= r(\lambda) = 0 \\ \Rightarrow p(x) &= (x - \lambda)q(x) \end{aligned}$$

"$\Leftarrow$" : Sei $p(x) = (x - \lambda)q(x)$.
$\Rightarrow\; p(\lambda) = 0 \cdot q(\lambda) = 0$ □

Korollar 5.8 *Jedes reelle Polynom vom Grad $n \geq 1$ hat höchstens n verschiedene Nullstellen.*

Beweis
Angenommen $p(\lambda_j) = 0$ für $j = 1, \ldots, n+1$ und $\lambda_1 \neq \lambda_2 \neq \ldots \neq \lambda_{n+1}$.
$\Rightarrow\; p(x) = (x - \lambda_1) \cdot \ldots \cdot (x - \lambda_{n+1}) \cdot q(x)$ mit $q(x) \neq 0$
$\Rightarrow\; Grad(p) > n \quad ↯$ □

Definition *λ heißt k-fache Nullstelle eines Polynoms p oder Nullstelle mit Vielfachheit k, falls*

$$p(x) = (x - \lambda)^k p_1(x) \quad \textit{mit } p_1(\lambda) \neq 0$$

Satz 5.9 *Sei $p(x) = \sum\limits_{k=0}^{n} a_k x^k$ mit $a_0, \ldots, a_n \in \mathbb{Z}$. Ist $x = \dfrac{a}{b}$ Nullstelle von p, sind $a, b \in \mathbb{Z}$ und sind a, b teilerfremd, dann gilt $a \mid a_0$ und $b \mid a_n$.*

Beweis

$$\begin{aligned} & 0 = a_n \left(\frac{a}{b}\right)^n + a_{n-1} \left(\frac{a}{b}\right)^{n-1} + \ldots + a_0 \\ \Leftrightarrow \quad & 0 = a_n a^n + a_{n-1} a^{n-1} b + \ldots + a_0 b^n \\ \Leftrightarrow \quad & a_n a^n = -b(\ldots) \\ \Rightarrow \quad & b \mid a_n a^n \;\Rightarrow\; b \mid a_n \;, \quad \text{da } a, b \text{ teilerfremd.} \end{aligned}$$

Analog zeigt man

$$b^n a_0 = -a(\ldots) \;\Rightarrow\; a \mid b^n a_0 \;\Rightarrow\; a \mid a_0$$

$\square$

Beispiele

1) $p(x) = x^3 - 2x^2 - 6x + 4$, also $a_n = a_3 = 1,\ a_0 = 4$
mögliche rationale Nullstellen: $x = \pm 1,\ \pm 2,\ \pm 4$
Einsetzen liefert $p(-2) = 0$.

2) $p(x) = 12x^4 - 4x^3 + 6x^2 + x - 1$, also $a_n = a_4 = 12,\ a_0 = -1$
mögliche rationale Nullstellen: $x = \pm 1,\ \pm\frac{1}{2},\ \pm\frac{1}{3},\ \pm\frac{1}{4},\ \pm\frac{1}{6},\ \pm\frac{1}{12}$
Einsetzen liefert $p\left(\frac{1}{3}\right) = 0$.

Definition *Seien p,q reelle Polynome. Eine Funktion $f : D \to \mathbb{R},\ x \mapsto \dfrac{p(x)}{q(x)}$ mit $D = \{x \in \mathbb{R} : q(x) \neq 0\}$ heißt (gebrochen-)rationale reelle Funktion.*

Bemerkung
Ist λ eine k-fache Nullstelle von q und eine m-fache Nullstelle von p mit $m \geq k$, dann existieren reelle Polynome p_1, q_1, so dass $\tilde{f} : \tilde{D} \to \mathbb{R},\ x \mapsto \dfrac{p_1(x)}{q_1(x)}$ mit $\tilde{D} = D \cup \{\lambda\}$ eine Fortsetzung von f ist.

Bemerkung
Die Berechnung von Funktionswerten von $p(x) = \sum\limits_{k=0}^{n} a_k x^k$ an einer Stelle $x = x_0$ lässt sich in vielen Fällen effizienter durchführen, wenn man $p(x)$ in der Form

$$p(x) \quad = \quad (((a_n x + a_{n-1})x + a_{n-2})x + \ldots + a_1)x + a_0$$

schreibt und dann x_0 einsetzt. Dieses Rechenverfahren nennt man Horner-Schema. Beim Horner-Schema benötigt man zur Berechnung von $p(x_0)$ nur n Multiplikationen anstatt $\frac{1}{2}n(n+1)$.

Beispiel
$p(x) = x^4 - 5x^2 - 2x = 1x^4 + 0x^3 - 5x^2 - 2x + 0$, und $x_0 = -2$
Dann ist das Horner-Schema:

a_k	1	0	-5	-2	0
$x_0 = -2$		-2	4	2	0
	1	-2	-1	0	$0 = p(-2)$

Die Werte in der untersten Zeile liefern auch das Ergebnis der Polynomdivision $p(x) : (x - x_0)$, nämlich

$$\begin{aligned}(x^4 - 5x^2 - 2x) : (x+2) &= 1x^3 - 2x^2 - 1x + 0 + \frac{0}{x+2} \\ &= x^3 - 2x^2 - x\end{aligned}$$

Satz 5.10 *Zu $n+1$ verschiedenen Stützstellen $x_0, \ldots, x_n \in \mathbb{R}$ mit zugehörigen Funktionswerten $y_0, \ldots, y_n \in \mathbb{R}$ gibt es genau ein reelles Polynom p mit $Grad(p) \leq n$ und $p(x_k) = y_k$ für $k = 0, \ldots, n$. Dabei gilt:*

$$\begin{aligned}p(x) &= \sum_{k=0}^{n} y_k L_k(x) \quad \textit{mit} \\ L_k(x) &= \frac{(x - x_0) \cdot \ldots \cdot (x - x_{k-1}) \cdot (x - x_{k+1}) \cdot \ldots \cdot (x - x_n)}{(x_k - x_0) \cdot \ldots \cdot (x_k - x_{k-1}) \cdot (x_k - x_{k+1}) \cdot \ldots \cdot (x_k - x_n)}\end{aligned}$$

Beweisstrategie
Die Darstellung von $p(x)$ zeigt man durch vollständige Induktion. Die Eindeutigkeit folgt aus Korollar 5.8.

5.3 Trigonometrische Funktionen

Trigonometrische Funktionen setzen geometrische Größen wie Winkelmaße und Streckenlängen zueinander in Beziehung. Für die Einführung von trigonometrischen Funktionen benötigen wir einige Vorbereitungen.

Um geometrische Objekte (Punkte, Strecken, Geraden, Ebenen, Kreise, ...) mathematisch exakt zu definieren, stellen wir uns die geometrischen Objekte in einem Koordinatensystem vor und definieren sie mathematisch durch Angabe ihrer Koordinaten. Als Erstes definieren wir:

Definition *Ein Punkt auf einer Geraden ist eine Zahl $x \in \mathbb{R}$. Ein Punkt in einer Ebene ist ein geordnetes Paar $(x, y) \in \mathbb{R}^2$. Ein Punkt im (dreidimensionalen) Raum ist ein geordnetes 3-Tupel $(x, y, z) \in \mathbb{R}^3$.*

Alle anderen geometrischen Objekte definieren wir als Punktmengen, also als Teilmengen von $\mathbb{R}$ bzw. $\mathbb{R}^2$ bzw. $\mathbb{R}^3$.
Da man ein und dieselbe Punktmenge in Abhängigkeit von den verwendeten mathematischen Mitteln (z.B. Funktionen, Vektoren, ...) auf verschiedene Weise beschreiben kann, gibt es mehrere äquivalente Definitionen für geometrische Objekte.
Beispielsweise können wir definieren:

Definition *Geraden in einer Ebene sind entweder Graphen von reellen Polynomfunktionen der Form $p : \mathbb{R} \to \mathbb{R},\ x \mapsto ax + b$ mit $a, b \in \mathbb{R}$ oder Relationen der Form*

$R = \{(x,y) \in \mathbb{R}^2 : x = c\}$ *mit* $c \in \mathbb{R}$.

Auf der Basis dieser Definition kann man dann geometrische Eigenschaften von Punkten und Geraden in einer Ebene beweisen, zum Beispiel:

Zu je zwei Punkten $P = (x_0, y_0)$ *und* $Q = (x_1, y_1)$ *mit* $P \neq Q$ *gibt es genau eine Gerade* g *mit* $P, Q \in g$. *Die Gerade* g *kann explizit angegeben werden.*

Im Falle von $x_0 \neq x_1$ folgt dies unmittelbar aus Satz 5.10. Gilt $x_0 = x_1$, dann ist $g = \{(x,y) \in \mathbb{R}^2 : x = x_0\}$.

Die *Strecke* PQ zwischen zwei verschiedenen Punkten P und Q können wir als Teilmenge der Geraden g, auf der P und Q liegen, definieren. Ist beispielsweise $P = (0,0)$ und $Q = (1,2)$, dann ist $PQ := \{(x, f(x)) \in \mathbb{R}^2 : x \in [0,1]\}$ mit $f(x) = 2x$.

Die *Streckenlänge* $|PQ|$ einer Strecke zwischen den Punkten $P = (x_0, y_0)$ und $Q = (x_1, y_1)$ können wir definieren durch

$$|PQ| := \sqrt{(x_1 - x_0)^2 + (y_1 - y_0)^2}.$$

Hierbei gehen wir implizit von einem kartesischen Koordinatensystem aus. Alternativ kann man Streckenlängen auch auf eine abstraktere Weise definieren und dann den Satz des Pythagoras beweisen, aus dem schließlich die obige Formel für die Streckenlänge folgt.

Definition *Der Kreis* $K_r(M)$ *mit Mittelpunkt* $M = (x_0, y_0)$ *und Radius* r *ist definiert durch*

$$K_r(M) := \{X \in \mathbb{R}^2 : |MX| = r\} = \{(x,y) \in \mathbb{R}^2 : (x - x_0)^2 + (y - y_0)^2 = r^2\}.$$

Der Kreis $K_1(O)$ *um den Ursprung* $O = (0,0)$ *mit Radius* 1 *wird auch Einheitskreis oder Einheitssphäre genannt und mit* S^1 *bezeichnet.*

Man kann Kreise auch mit Hilfe von Vereinigungen von Funktionsgraphen beschreiben, zum Beispiel $S^1 = \{(x, \sqrt{1 - x^2}) : x \in [-1,1]\} \cup \{(x, -\sqrt{1 - x^2}) : x \in [-1,1]\}$.

Unter Verwendung dieser Darstellung kann man dann beispielsweise den *Kreisbogen* $\overset{\frown}{PQ}$, der den Punkt $P = (1,0)$ mit dem Punkt $Q = (-\frac{1}{2}, -\frac{1}{2}\sqrt{3})$ durch Durchlaufen des Einheitskreises gegen den Uhrzeigersinn verbindet, definieren durch $\{(x, \sqrt{1 - x^2}) : x \in [-1,1]\} \cup \{(x, -\sqrt{1 - x^2}) : x \in [-1, -\frac{1}{2}]\}$.

Die *Länge eines Kreisbogens* $\overset{\frown}{AB}$ kann man definieren durch das Supremum der Längen aller möglichen Polygonzüge (aneinandergesetzte Strecken) mit Anfangspunkt A, Endpunkt B und Eckpunkten auf $\overset{\frown}{AB}$. Dieses Supremum kann dann mit Hilfe der Integralrechnung berechnet werden. Für jeden Halbkreis mit Radius 1 kommt als Länge die sogenannte Kreiszahl π heraus. π ist eine irrationale Zahl und es gilt $\pi = 3,141592653589793\ldots$.

Den *Winkel* zwischen zwei Punkten $A, B \in S^1$ bzw. zwischen den Strecken OA und OB können wir mit einem Kreisbogen auf dem Einheitskreis mit Anfangspunkt A und Endpunkt B identifizieren. Wird dabei A mit B gegen Uhrzeigersinn verbunden, nennen wir den Winkel positiv orientiert, wird A mit B im Uhrzeigersinn verbunden, heißt der Winkel negativ orientiert.
Das *Bogenmaß* eines positiv orientierten Winkels definieren wir durch die Länge des zugehörigen Kreisbogens, das Bogenmaß eines negativ orientierten Winkels durch das (-1)-fache der Länge des zugehörigen Kreisbogens. Das *Gradmaß* eines Winkels in der Maßeinheit ° (Grad) ist dann definiert durch (Bogenmaß : π) $\cdot$ 180°.

Mit Hilfe der obigen Vorbereitungen können wir nun die trigonometrischen Funktionen definieren.

Definition *Sei* $\varphi \in]-2\pi, 2\pi[$ *das Bogenmaß eines Winkels zwischen* $(1,0)$ *und* $(x,y) \in S^1$. *Dann sei*

$$\begin{aligned} \sin\varphi &:= y \quad (\text{Sinus von } \varphi), \\ \cos\varphi &:= x \quad (\text{Kosinus von } \varphi) \end{aligned}$$

sowie

$$\sin(\varphi + 2k\pi) := \sin\varphi, \qquad \cos(\varphi + 2k\pi) := \cos\varphi$$

für alle $k \in \mathbb{Z}$. *Außerdem sei*

$$\tan\varphi := \frac{\sin\varphi}{\cos\varphi} \quad (\text{Tangens von } \varphi), \qquad \textit{falls } \cos\varphi \neq 0.$$

Eigenschaften von Sinus, Kosinus und Tangens:

a) Es gilt

φ	0	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$	$\frac{\pi}{2}$
$\sin\varphi$	0	$\frac{1}{2}$	$\frac{1}{2}\sqrt{2}$	$\frac{1}{2}\sqrt{3}$	1
$\cos\varphi$	1	$\frac{1}{2}\sqrt{3}$	$\frac{1}{2}\sqrt{2}$	$\frac{1}{2}$	0

b) Für alle $\varphi \in \mathbb{R}$ gilt

$$\begin{aligned} \sin(-\varphi) &= -\sin\varphi, & \cos(-\varphi) &= \cos\varphi \\ \sin(\tfrac{\pi}{2} - \varphi) &= \sin(\tfrac{\pi}{2} + \varphi), & \cos(\tfrac{\pi}{2} - \varphi) &= -\cos(\tfrac{\pi}{2} + \varphi). \end{aligned}$$

c) Für alle $\varphi \in \mathbb{R}$ gilt

$$\begin{aligned} &\sin^2\varphi + \cos^2\varphi := (\sin\varphi)^2 + (\cos\varphi)^2 = 1, \\ &\cos\varphi = \sin(\tfrac{\pi}{2} + \varphi). \end{aligned}$$

d) Für alle $\alpha, \beta \in \mathbb{R}$ gelten die Additionstheoreme

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

e) Für $|\varphi| < \frac{\pi}{2}$ gelten die Abschätzungen

$$1 - \tfrac{1}{2}\varphi^2 \;\le\; \cos\varphi \;\le\; 1,$$
$$\cos\varphi \cdot |\varphi| \;\le\; |\sin\varphi| \;\le\; |\varphi|.$$

f) Für alle $\varphi \in \mathbb{R}$ mit $\cos\varphi \neq 0$ und $k \in \mathbb{Z}$ gilt

$$\tan(\varphi + k\pi) \;=\; \tan\varphi.$$

g) $\sin|_{[-\frac{\pi}{2},\frac{\pi}{2}]} : [-\frac{\pi}{2}, \frac{\pi}{2}] \to [-1,1]$, $x \mapsto \sin x$ ist bijektiv. Die zugehörige Umkehrfunktion heißt arcsin (Arkussinus).

h) $\cos|_{[0,\pi]} : [0,\pi] \to [-1,1]$, $x \mapsto \cos x$ ist bijektiv. Die zugehörige Umkehrfunktion heißt arccos (Arkuscosinus).

i) $\tan|_{]-\frac{\pi}{2},\frac{\pi}{2}[} :]-\frac{\pi}{2}, \frac{\pi}{2}[\to \mathbb{R}$, $x \mapsto \tan x$ ist bijektiv. Die zugehörige Umkehrfunktion heißt arctan (Arkustangens).

Die obigen Eigenschaften kann man unter Ausnutzung elementargeometrischer Argumente, welche man mit Hilfe von Koordinatenmengen mathematisch exakt formulieren kann, beweisen. Diese Vorgehensweise ist allerdings ziemlich aufwendig. Daher findet man in der Literatur meistens eine alternative, äquivalente Definition der trigonometrischen Funktionen, welche mit Begriffen der Analysis (unendliche Reihen) auskommt und keinen Rückgriff auf die Geometrie benötigt. Diese Definition werden wir später behandeln. Auf der Basis dieser Definition können die obigen Eigenschaften der trigonometrischen Funktionen dann eleganter und nur mit Mitteln der Analysis bewiesen werden.

Bemerkung:
Geometrische Objekte kann man alternativ zu den in diesem Abschnitt diskutierten Definitionen mit Hilfe von Koordinatenmengen auch durch ein Axiomensystem definieren (siehe Hilbert: Grundlagen der Geometrie) und dann zeigen, dass dieses Axiomensystem (abgesehen von strukturerhaltenden Kopien) alleine von den oben definierten Koordinatenmengen erfüllt wird.

6 Komplexe Zahlen

Gesucht ist ein Körper $(\mathbb{C}, +, \cdot)$ mit $\mathbb{R} \subseteq \mathbb{C}$ und einer Zahl $i \in \mathbb{C}$ mit $i^2 = -1$.

Ein solcher Körper müsste auch Elemente der Form $x + iy = x + i \cdot y$ für beliebige $x, y \in \mathbb{R}$ haben und es müsste gelten:

(1) $\forall\, x_1, x_2, y_1, y_2 \in \mathbb{R}:$

$$(x_1 + iy_1) + (x_2 + iy_2) = (x_1 + x_2) + i(y_1 + y_2)$$

(2) $\forall\, x_1, x_2, y_1, y_2 \in \mathbb{R}:$

$$(x_1+iy_1)\cdot(x_2+iy_2) = x_1\cdot x_2+i(x_1y_2+x_2y_1)+i^2y_1y_2 = (x_1x_2-y_1y_2)+i(x_1y_2+x_2y_1)$$

Frage: Lässt sich ein solcher Körper widerspruchsfrei konstruieren?

Definition *Sei $\mathbb{C} := \mathbb{R}^2$ mit den Verknüpfungen*

$$(x_1, y_1) + (x_2, y_2) := (x_1 + x_2, y_1 + y_2)$$

$$(x_1, y_1) \cdot (x_2, y_2) := (x_1x_2 - y_1y_2, x_1y_2 + x_2y_1)$$

Die Elemente von $\mathbb{C}$ heißen komplexe Zahlen.

Satz 6.1 *$(\mathbb{C}, +, \cdot)$ ist ein Körper und es gilt $(0, 1)^2 = (0, 1) \cdot (0, 1) = (-1, 0)$.*

Beweis
Durch direktes Nachrechnen zeigt man, dass + und $\cdot$ assoziativ und kommutativ sind, dass das Distributivgesetz gilt, dass $(0, 0)$ das neutrale Element bzgl. + und das $(1, 0)$ das neutrale Element bzgl. $\cdot$ sowie das $(-x, -y)$ das inverse Element von (x, y) bzgl. + ist.
Sei $(x, y) \neq (0, 0)$. Dann gilt:

$$(x, y) \cdot (a, b) = (1, 0)$$

$$\Leftrightarrow (xa - yb,\ xb + ya) = (1, 0)$$

$$\Leftrightarrow \left\{ \begin{array}{l} xa - yb = 1 \\ xb + ya = 0 \end{array} \right. \Leftrightarrow: (*)$$

1. Fall: $x \neq 0$

$$(*) \Leftrightarrow \left\{ \begin{array}{c} xa - yb = 1 \\ b = -\dfrac{y}{x}\, a \end{array} \right.$$

$$\Leftrightarrow \left\{ \begin{array}{c} \left(x + \dfrac{y^2}{x}\right) a = 1 \\ b = -\dfrac{y}{x}\, a \end{array} \right.$$

$$\Leftrightarrow \begin{cases} a = \dfrac{x}{x^2+y^2} \\ b = -\dfrac{y}{x^2+y^2} \end{cases}$$

2. Fall: $x = 0$

$$(*) \Leftrightarrow \begin{cases} xa - yb = 1 \\ a = 0 \end{cases}$$

$$\Leftrightarrow \begin{cases} b = -\dfrac{1}{y} \\ a = 0 \end{cases}$$

$$\Leftrightarrow \begin{cases} a = \dfrac{x}{x^2+y^2} \\ b = -\dfrac{y}{x^2+y^2} \end{cases}$$

Also ist $\left(\dfrac{x}{x^2+y^2}, -\dfrac{y}{x^2+y^2}\right)$ das inverse Element von $(x,y) \neq (0,0)$ bzgl. $\cdot$.

Außerdem gilt:

$$(0,1)^2 = (0,1) \cdot (0,1) = (0 \cdot 0 - 1 \cdot 1,\, 0 \cdot 1 + 0 \cdot 1) = (-1,0)$$

$\square$

Satz 6.2 *Sei $E : \mathbb{R} \to \mathbb{C}$, $x \mapsto (x,0)$. Dann ist E injektiv und es gilt:*

$$E(x) + E(y) = E(x+y)\,, \quad E(0) = E(0,0)\,,$$

$$E(x) \cdot E(y) = E(x \cdot y)\,, \quad E(1) = E(1,0)\,.$$

Beweis
durch direktes Nachrechnen. $\square$

$\mathbb{R}$ kann also mittels E in $\mathbb{C}$ eingebettet werden, wobei die Verknüpfungen $+, \cdot$ auf $\mathbb{R}$ den Verknüpfungen $+, \cdot$ auf $E(\mathbb{R}) \subseteq \mathbb{C}$ entsprechen.
Somit macht es Sinn, die folgende vereinfachende Notation einzuführen. Sei

$$x := (x,0),$$

$$i := (0,1).$$

Dann gilt:

$$\forall\, x,y \in \mathbb{R} :\ (x,y) = (x,0) + (0,y) = (x,0) + (0,1) \cdot (y,0) = x + iy$$

Daher gilt $\mathbb{C} = \{x + iy : \ x, y \in \mathbb{R}\}$, und $+, \cdot$ sind durch (1), (2) eindeutig als Verknüpfungen auf $\mathbb{C}$ definiert. Außerdem gilt $i^2 = -1$.

Schreibt man die komplexen Zahlen in der Form $x + iy$ und beachtet, dass $(\mathbb{C}, +, \cdot)$ ein Körper ist und $i^2 = -1$ gilt, dann kann man mit komplexen Zahlen nach denselben algebraischen Regeln rechnen wie man es von den reellen Zahlen gewohnt ist. Insbesondere gelten die Sätze 3.3 bis 3.10 sowie der binomische Satz 4.2 d) auch in $\mathbb{C}$, denn sie folgen aus den Körperaxiomen.

Beispiel

$$\frac{1+i}{1-i} = \frac{(1+i)^2}{(1-i)(1+i)} = \frac{1+2i+i^2}{1-i^2} = \frac{2i}{2} = i \ ,$$

$$\begin{aligned} \frac{a+ib}{c+id} &= \frac{(a+ib)(c-id)}{c^2+d^2} = \frac{ac+bd+i(bc-ad)}{c^2+d^2} \\ &= \frac{ac+bd}{c^2+d^2} + i\,\frac{bc-ad}{c^2+d^2} \ . \end{aligned}$$

Bemerkung
Die kleiner-Relation $<$ kann nicht von $\mathbb{R}$ auf $\mathbb{C}$ fortgesetzt werden, so dass die Anordnungsaxiome (O1) bis (O4) auf $\mathbb{C}$ gelten. Denn dann würde wie in 3.11(2) folgen, dass $i^2 > 0$ ist, was ein Widerspruch zu $i^2 = -1 < 0$ wäre.

Definition *Sei $z = x + iy \in \mathbb{C}$ mit $x, y \in \mathbb{R}$. Dann heißt:*

(i) $Re\, z := x$ *der Realteil von* z,

(ii) $Im\, z := y$ *der Imaginärteil von* z,

(iii) $\overline{z} : x - iy$ *die zu* z *konjugiert komplexe Zahl,*

(iv) $|z| = \sqrt{x^2 + y^2}$ *der Betrag von* z.

Satz 6.3 *Für alle $z, w \in \mathbb{C}$ gelten:*

(i) $\overline{\overline{z}} = z$

(ii) $z = \overline{z} \Leftrightarrow z \in \mathbb{R}$

(iii) $Re\, z = \dfrac{z + \overline{z}}{2}, \ Im\, z = \dfrac{z - \overline{z}}{2i}$

(iv) $\overline{z + w} = \overline{z} + \overline{w}$

(v) $\overline{z \cdot w} = \overline{z} \cdot \overline{w}$

(vi) $|Re\, z| \leq |z|, \ |Im\, z| \leq |z|, \ |\overline{z}| = |z|$

(vii) $z \cdot \overline{z} = |z|^2$, $\frac{1}{z} = \frac{\overline{z}}{|z|^2}$ *(falls* $z \neq 0$*)*

(viii) $|z| \geq 0$, $|z| = 0 \Leftrightarrow z = 0$

(ix) $|z \cdot w| = |z| \cdot |w|$

(x) Dreiecksungleichung:
$|z + w| \leq |z| + |w|$
$|z + w| = |z| + |w| \Leftrightarrow (w = 0 \vee \exists \lambda \in \mathbb{R}_0^+ : z = \lambda w)$

(xi) Dreiecksungleichung nach unten:
$||z| - |w|| \leq |z - w|$

Beweis
durch direktes Nachrechnen. Z.B. Nachweis von (x) (nachdem (i)-(ix) schon gezeigt sind):

$$\begin{aligned}
|z+w|^2 &\overset{(vii)}{=} (z+w) \cdot (\overline{z+w}) \\
&\overset{(iv)}{=} (z+w)(\overline{z} + \overline{w}) \\
&= z\overline{z} + z\overline{w} + w\overline{z} + w\overline{w} \\
&\overset{(vii)}{=} |z|^2 + z\overline{w} + w\overline{z} + |w|^2 \\
&\overset{(i),(v)}{=} |z|^2 + z\overline{w} + \overline{z\overline{w}} + |w|^2 \\
&\overset{(iii)}{=} |z|^2 + 2\,Re(z\overline{w}) + |w|^2 \\
&\overset{(vi)}{\leq} |z|^2 + 2\,|z\overline{w}| + |w|^2 \\
&\overset{(v),(ix)}{=} |z|^2 + 2\,|z|\,|w| + |w|^2 \\
&= (|z| + |w|)^2
\end{aligned}$$

$\Longrightarrow \; |z + w| \leq |z| + |w|$

Aus der obigen Rechnung folgt:

$$\begin{aligned}
|z+w| = |z| + |w| \quad &\Longrightarrow \quad |z\overline{w}| = Re(z\overline{w}) \\
&\Longrightarrow \quad \sqrt{(Re(z\overline{w}))^2 + (Im(z\overline{w}))^2} = Re(z\overline{w}) \\
&\Longrightarrow \quad Re(z\overline{w}) \geq 0 \wedge (Re(z\overline{w}))^2 + (Im(z\overline{w}))^2 = (Re(z\overline{w}))^2 \\
&\Longrightarrow \quad Re(z\overline{w}) \geq 0 \wedge (Im(z\overline{w})) = 0 \\
&\Longrightarrow \quad z\overline{w} \in \mathbb{R}_0^+ \\
&\overset{(vii),(viii)}{\Longrightarrow} \quad w = 0 \vee z = z\frac{w}{w} = \frac{z\overline{w}}{|w|^2} w, \text{ wobei } \frac{z\overline{w}}{|w|^2} \in \mathbb{R}_0^+.
\end{aligned}$$

Sei umgekehrt $w = 0 \vee \exists \lambda \in \mathbb{R}_0^+ : z = \lambda w$.

1. Fall: $w = 0$

$$|z+w| = |z| \overset{(viii)}{=} |z| + |w|$$

2. Fall: $\exists \lambda \in \mathbb{R}_0^+ : z = \lambda w$

$$\begin{aligned} |z+w| &= |(\lambda+1)w| \\ &\overset{(ix)}{=} |\lambda+1| \cdot |w| \\ &= (\lambda+1) \cdot |w| \\ &= \lambda|w| + |w| \\ &= |\lambda| \cdot |w| + |w| \\ &\overset{(ix)}{=} |\lambda w| + |w| = |z| + |w| \end{aligned}$$

□

Bemerkung
Die Abbildung $z \mapsto \overline{z}$ kann geometrisch interpretiert werden als Spiegelung an der reellen Achse.

Satz 6.4 $\forall z \in \mathbb{C} \setminus \{0\} \, \exists! r \in \mathbb{R}_0^+ \exists! \varphi \in [0, 2\pi[: \; z = r(\cos\varphi + i \cdot \sin\varphi)$. *Dabei gilt* $r = |z|$.

Bezeichnung: arg(z) $:= \varphi$ (Argument von z).

Beispiel
$1 + i = \sqrt{2}\left(\cos\left(\frac{\pi}{4}\right) + i \cdot \sin\left(\frac{\pi}{4}\right)\right)$

Bemerkung
Sei $z \in \mathbb{C}$ mit $z = x + iy = r(\cos\varphi + i \cdot \sin\varphi)$. Dann ist $r = \sqrt{x^2+y^2}$ und

$$\varphi = \begin{cases} \arctan\left(\frac{\mathrm{y}}{\mathrm{x}}\right), & x > 0,\, y \geq 0 \\ \frac{\pi}{2}, & x = 0,\, y > 0 \\ \pi + \arctan\left(\frac{\mathrm{y}}{\mathrm{x}}\right), & x < 0 \\ \frac{3\pi}{2}, & x = 0,\, y < 0 \\ 2\pi + \arctan\left(\frac{\mathrm{y}}{\mathrm{x}}\right), & x > 0,\, y < 0 \end{cases}$$

Satz 6.5 *Seien* $z_1, z_2 \in \mathbb{C}$ *mit* $z_i := r_i(\cos\varphi_i + i\sin\varphi_i)$ *für* $i = 1, 2$. *Dann gilt*

$$z_1 \cdot z_2 = r_1 r_2 \left(\cos(\varphi_1 + \varphi_2) + i\sin(\varphi_1 + \varphi_2)\right).$$

Beweis
Mit dem Additionstheorem folgt

$$\begin{aligned} z_1 \cdot z_2 &= r_1 (\cos\varphi_1 + i \sin\varphi_2) \cdot r_2 (\cos\varphi_2 + i\sin\varphi_2) \\ &= r_1 r_2 ((\cos\varphi_1 \cos\varphi_2 - \sin\varphi_1 \sin\varphi_2) + i(\sin\varphi_1 \cos\varphi_2 + \cos\varphi_1 \sin\varphi_2)) \\ &= r_1 r_2 (\cos(\varphi_1 + \varphi_2) + i \sin(\varphi_1 + \varphi_2)) \,. \end{aligned}$$

□

Bemerkung
Für festes $w \in \mathbb{C} \setminus \{0\}$ kann daher die Multiplikationsabbildung $z \mapsto wz$ geometrisch interpretiert werden als Drehstreckung (Verkettung von Drehung und zentrischer Streckung) mit Streckzentrum im Ursprung, Streckfaktor $|w|$ und Drehwinkel $\arg(w)$.

Korollar 6.6 *Sei $z = r(\cos\varphi + i\sin\varphi)$. Dann gilt für jedes $n \in \mathbb{N}$:*

$$z^n = r^n (\cos(n\varphi) + i \sin(n\varphi)) \,.$$

Beweisskizze
Folgt aus Satz 6.5 durch vollständige Induktion.

Wir führen folgende vereinfachende Notation ein. Sei

$$e^{i\varphi} := \cos\varphi + i\sin\varphi.$$

Dann gilt:

(i) $e^{i\cdot 0} = 1$ und $e^{i(\varphi_1+\varphi_2)} = e^{i\varphi_1} \cdot e^{i\varphi_2}$ (wegen Satz 6.5). Für jedes $n \in \mathbb{N}$ gilt wegen Satz 6.6 $\left(e^{i\varphi}\right)^n = e^{in\varphi}$.

(ii) $|e^{i\varphi}| = 1$.

(iii) Für jedes $z \in \mathbb{C} \setminus \{0\}$ existiert genau ein $\varphi \in [0, 2\pi[$, so dass $z = |z|\, e^{i\varphi}$.

(iv) Sind $z_1 = |z_1|\, e^{i\varphi_1}$ und $z_2 = |z_2|\, e^{i\varphi_2}$, dann gilt

$$z_1 z_2 = |z_1|\,|z_2|\, e^{i(\varphi_1+\varphi_2)}.$$

Satz 6.7 *Sei $a \in \mathbb{C} \setminus \{0\}$ mit $a = |a|\, e^{i\varphi}$ und $n \in \mathbb{N}$. Dann hat die Gleichung*

$$z^n = a$$

genau n verschiedene komplexe Lösungen, nämlich

$$z_k = |a|^{\frac{1}{n}} \cdot e^{i\frac{\varphi+2k\pi}{n}} = |a|^{\frac{1}{n}} \left(\cos\left(\frac{\varphi + 2k\pi}{n}\right) + i \sin\left(\frac{\varphi+2k\pi}{n}\right)\right)$$

für $k = 0, 1, \ldots, n-1$.

Beweisskizze
Folgt aus Korollar 6.6.

Definition *Ein Polynom*

$$\sum_{k=0}^{n} a_k z^k \quad mit \quad a_k \in \mathbb{C} \text{ für } k = 0, \dots, n$$

heißt komplexes Polynom.

Bemerkung
Alle Definitionen, Rechenverfahren, Sätze und Korollare aus Abschnitt 5.2 gelten auch für komplexe Polynome.

Zusätzlich gelten:

Satz 6.8 *Seien $a, b, c \in \mathbb{C}$ mit $a \neq 0$. Dann gilt*

$$az^2 + bz + c = 0 \iff z_{1/2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.$$

Hierbei bezeichnet $\pm\sqrt{b^2 - 4ac}$ die Lösungen von $w^2 = b^2 - 4ac$.

Beweisskizze
Analog zum Beweis von Satz 3.22 mit dem einzigen Unterschied, dass hier keine Einschränkung an $b^2 - 4ac$ gemacht werden muss.

Satz 6.9 (Fundamentalsatz der Algebra) *Jedes Polynom $\sum\limits_{k=0}^{n} a_k z^k$ mit $a_0, \dots, a_n \in \mathbb{C}$ vom Grad ≥ 1 besitzt mindestens eine Nullstelle in $\mathbb{C}$.*

Der Beweis benötigt fortgeschrittene Methoden der Analysis (Theorie komplex differenzierbarer Funktionen) und kann an dieser Stelle der Vorlesung noch nicht geführt werden.

Korollar 6.10 *Jedes komplexe Polynom*

$$p(z) = \sum_{k=0}^{n} a_k z^k \quad \text{mit } a_0, \dots, a_n \in \mathbb{C}$$

vom Grad ≥ 1 lässt sich über $\mathbb{C}$ in Linearfaktoren zerlegen, das heißt es existieren nicht notwendigerweise verschiedene $\lambda_j \in \mathbb{C}$, so dass

$$p(z) = a_n(z - \lambda_1) \cdot \ldots \cdot (z - \lambda_n).$$

Die Linearfaktoren $(z - \lambda_j)$ sind bis auf die Reihenfolge eindeutig.

Beweisskizze
Vollständige Induktion unter Ausnutzung von Satz 6.9 und Polynomdivision.

Korollar 6.11 (Satz von Vieta) *Sei* $p(z) = \sum_{k=0}^{n} a_k z^k$ *und* $a_n = 1$. *Dann gilt*

$$a_{n-1} = -\sum_{k=1}^{n} \lambda_k \quad \textit{und} \quad a_0 = (-1)^n \prod_{k=1}^{n} \lambda_n.$$

Beweisskizze
Vollständige Induktion und Ausmultiplizieren der Linearfaktorzerlegung.

Satz 6.12 *Sei* p *ein reelles Polynom. Dann gilt:*

a) Ist $\lambda \in \mathbb{C}$ *eine Nullstelle von* p, *dann ist auch* $\overline{\lambda}$ *eine Nullstelle von* p.

b) p *besitzt eine der folgenden Darstellungen:*

1. p *ist Produkt von linearen rellen Polynomen.*
2. p *ist Produkt von quadratischen reellen Polynomen, welche keine reellen Nullstellen besitzen.*
3. p *ist Produkt von linearen reellen und quadratischen reellen Polynomen, wobei die quadratischen Polynome keine reellen Nullstellen haben.*

Beweisskizze
a) Zeige und verwende, dass für ein reelles Polynom p gilt $p(\overline{z}) = \overline{p(z)}$.
b) Verwende 6.10, 6.12 1. und

$$(z-\lambda)(z-\overline{\lambda}) = z^2 - 2(Re(\lambda))z + |\lambda|^2 .$$

Teil II
Lineare Algebra

7 Lineare Gleichungssysteme

Beispiel
Gesucht sind alle $(x_1, x_2, x_3) \in \mathbb{R}^3$, die alle drei Gleichungen

$$\begin{array}{rcrcrcrl}
x_1 & + & x_2 & + & x_3 & = & 6 & \qquad (1) \\
x_1 & + & 2x_2 & - & x_3 & = & 2 & \qquad (2) \\
-2x_1 & + & 3x_2 & - & 3x_3 & = & -5 & \qquad (3)
\end{array}$$

lösen.
Die Lösungsmenge ändert sich nicht, wenn wir die Gleichung (2) durch die Differenz von (2) und (1) und die Gleichung (2) durch die Summe von (3) und dem Zweifachen von (1) ersetzen.
Dies liefert die Gleichungen

$$\begin{array}{rcrcrcrlcl}
x_1 & + & x_2 & + & x_3 & = & 6 & \qquad (1) & & \\
 & & x_2 & - & 2x_3 & = & -4 & \qquad (4) & = & (2)-(1) \\
 & & 5x_2 & - & x_3 & = & 7 & \qquad (5) & = & 2\cdot(1)+(3)
\end{array}$$

Multiplizieren wir jetzt (4) mit -5 auf beiden Seiten und ersetzen (5) durch die Summe der erhaltenen Gleichung und (5), dann ergibt sich

$$\begin{array}{rcrcrcrlcl}
x_1 & + & x_2 & + & x_3 & = & 6 & \qquad (1) & & \\
 & & x_2 & - & 2x_3 & = & -4 & \qquad (4) & & \\
 & & & & 9x_3 & = & 27 & \qquad (6) & = & -5\cdot(4)+(5)
\end{array}$$

Dieses Gleichungssystem kann nun von unten nach oben aufgelöst werden:

1. Aus (6) folgt $x_3 = 3$.

2. Einsetzen in (4) und Auflösen nach x_2 liefert $x_2 = -4 + 6 = 2$.

3. Einsetzen von $x_2 = 2$ und $x_3 = 3$ in (1) und Auflösen nach x_1 liefert $x_1 = 1$.

Also ist $(1, 2, 3)$ die einzige Lösung der Gleichungen $(1), (2), (3)$.

Das gerade durchgeführte Lösungsverfahren lässt sich auch kürzer aufschreiben:

$$\begin{array}{rrr|rll}
1 & 1 & 1 & 6 & |\cdot(-1) & |\cdot 2 \\
1 & 2 & -1 & 2 & \longleftarrow^{|+} & |+ \\
-2 & 3 & -3 & -5 & \longleftarrow & \swarrow \\
\hline
1 & 1 & 1 & 6 & & \\
0 & 1 & -2 & -4 & |\cdot(-5) & \\
0 & 5 & -1 & 7 & \longleftarrow^{|+} & \\
\hline
1 & 1 & 1 & 6 & & \\
0 & 1 & -2 & -4 & & \\
0 & 0 & 9 & 27 & & \\
\hline
\end{array}$$

$$\begin{array}{rrcrcl} \Longrightarrow & x_3 & = & 3 & & \\ & x_2 & = & -4+6 & = & 2 \\ & x_1 & = & 6-3-2 & = & 1 \end{array}$$

Allgemeine Problemstellung:
Seien $m, n \in \mathbb{N}$, K ein Körper und $a_{ij}, b_i \in K$ für $i = 1, ..., m$ und $j = 1, ...n$. Gleichungen der Form

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & \cdots & & \cdots & + & a_{2n}x_n & = & b_2 \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ a_{m1}x_1 & + & \cdots & & \cdots & + & a_{mn}x_n & = & b_m \end{array}$$

heißen ein lineares Gleichungssystem (LGS) mit m Gleichungen, n Unbekannten $x_1, ..., x_n$, Koeffizienten $a_{ij} \in K$ und rechter Seite $(b_1, ..., b_n) \in K^m$.
Gesucht ist die Lösungsmenge

$$L = \left\{ (x_1, ..., x_n) \in K^n \,:\, \forall i \in \{1, ..., m\} : \sum_{j=1}^{n} a_{ij}x_j = b_i \right\}$$

Eine Möglichkeit, L zu bestimmen, ist der Algorithmus von Gauß:

1. Bringe - gegebenenfalls durch Vertauschen von Zeilen - das LGS auf eine Form, bei der keine Zeile weiter links beginnt als die erste, d.h., es soll gelten

   $$\forall i \geq 2 : \min \{j : a_{ij} \neq 0\} \geq \min \{j : a_{1j} \neq 0\}$$

2. Erreiche - gegebenenfalls durch Addition von Vielfachen der 1. Zeile zu den anderen Zeilen, dass gilt

   $$\forall i \geq 2 : \min \{j : a_{ij} \neq 0\} > \min \{j : a_{1j} \neq 0\}$$

3. Wiederhole den bisherigen Vorgang, wobei die nächste Zeile (von oben nach unten gezählt) die Rolle der ersten Zeile in 1. und 2. übernimmt.
   Führe das Verfahren so lange durch, bis das entstandene LGS Zeilenstufenform hat, d.h., es ist von der Form

$$\begin{array}{cccccccccccc} c_{1j_1}x_{j_1} & + & c_{1(j_1+1)}x_{j_1+1} & + & \cdots & & & & \cdots & + & c_{1n}x_n & = & d_1 \\ & & & & c_{2j_2}x_{j_2} & + & \cdots & & & & \cdots & = & d_2 \\ & & & & & \ddots & & & & & \vdots & \vdots & \vdots \\ & & & & & & c_{lj_l}x_l & + & \cdots & & & = & d_l \\ & & & & & & & & & & 0 & = & d_{l+1} \\ & & & & & & & & & & & \vdots & \vdots \\ & & & & & & & & & & 0 & = & d_m \,, \end{array}$$

wobei

(i) $l \leq \min \{m, n\}$,

(ii) $j_1 < j_2 < ... < j_l$,

(iii) $\forall i \in \{1, ..., l\} \, \forall j \in \{j_i, ..., n\} : c_{ij} \in K$,

(iv) $\forall i \in \{1, ..., n\} : d_i \in K$,

(v) $\forall i \in \{1, ..., l\} : c_{ij_i} \neq 0$.

Die Lösungsmenge dieses LGS ist gleich der Lösungsmenge des ursprünglichen LGS.

1. Fall: $l < m \, \wedge \, \exists i \in \{l+1, ..., m\} : d_i \neq 0$.
Das LGS hat keine Lösung.

2. Fall: $(l = m \, \vee \, (l < m \, \wedge \, \forall i \in \{l+1, ..., m\} : d_i = 0)) \wedge \, l = n$
Das LGS kann von unten nach oben eindeutig nach $x_m, ..., x_1$ aufgelöst werden, es besitzt somit genau eine Lösung.

3. Fall: $(l = m \, \vee \, (l < m \, \wedge \, \forall i \in \{l+1, ..., m\} : d_i = 0)) \wedge \, l < n$
Alle Variablen x_j mit $j \notin \{j_1, j_2, ..., j_l\}$ können mit beliebigen Werten aus K belegt werden. Für jede dieser Belegungen kann dann jeweils das LGS von unten nach oben eindeutig nach $x_{j_l}, ..., x_{j_1}$ aufgelöst werden. Das LGS hat somit unendlich viele Lösungen.

Der Beweis, dass der Gauß-Algorithmus in endlich vielen Schritten die Lösungsmenge des LGS liefert, wird durch vollständige Induktion geführt.

8 Vektoren

8.1 Der Vektorraum $\mathbb{R}^n$

Definition *Auf $\mathbb{R}^n$ ist eine Addition $+\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ definiert durch:*

$$\forall x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{R}^n\colon\ x + y = \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix},$$

sowie eine Skalarmultiplikation $\cdot\colon \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ definiert durch

$$\forall \lambda \in \mathbb{R}\, \forall x \in \mathbb{R}^n\colon\ \lambda x = \lambda \cdot x = \begin{pmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \end{pmatrix}.$$

Definition *Ein Vektorraum V über einem Körper K (ein K-Vektorraum V) ist eine abelsche Gruppe $(V, +)$ versehen mit einer Skalarmultiplikation, d.h. mit einer Abbildung $\cdot\colon K \times V \to V$, $(\lambda, x) \mapsto \lambda x = \lambda \cdot x$, so dass gilt*

$$\begin{aligned}
&(S1) \quad \forall \lambda, \mu \in K\, \forall x \in V\colon\ (\lambda + \mu)x = \lambda x + \mu x, \\
&(S2) \quad \forall \lambda \in K\, \forall x, y \in V\colon\ \lambda(x + y) = \lambda x + \lambda y, \\
&(S3) \quad \forall \lambda, \mu \in K\, \forall x \in V\colon\ \lambda(\mu x) = (\lambda\mu)x, \\
&(S4) \quad \forall x \in V\colon\ 1x = x.
\end{aligned}$$

Die Elemente von V heißen Vektoren.

Satz 8.1 *$\mathbb{R}^n$ versehen mit den oben definierten Abbildungen $+$ und $\cdot$ ist ein $\mathbb{R}$-Vektorraum.*

Beweis
durch direktes Nachrechnen. $\square$

Satz 8.2 *Sei V ein K-Vektorraum. Dann gilt:*

(i) $\forall \lambda \in K\, \forall x \in V\colon\ \lambda x = 0 \Leftrightarrow \lambda = 0 \lor x = 0$ *(x ist der Nullvektor),*

(ii) $\forall x \in V\colon\ (-1)x = -x$,

(iii) $\forall u, v \in V\, \exists!\, x \in V\colon\ u + x = v$, *und zwar* $x = v - u$.

Beweis
durch direktes Nachrechnen. $\square$

Definition

a) Seien $a, b \in \mathbb{R}^n$ *und* $b \neq 0$. *Dann heißt*

$$g := \{x \in \mathbb{R}^n \colon \exists \lambda \in \mathbb{R} \colon x = a + \lambda b\}$$

Gerade (durch die Punkte a *und* $b + a$*).*

b) Seien $p, q, r \in \mathbb{R}^n$ *und* $q, r \neq 0$ *und* $r \neq \lambda q$ *für alle* $\lambda \in \mathbb{R}$. *Dann heißt*

$$E := \{x \in \mathbb{R}^n \colon \exists \lambda, \mu \in \mathbb{R} \colon x = p + \lambda q + \mu r\}$$

Ebene (durch die Punkte $p, q + p$ *und* $r + p$*).*

Die obigen Darstellungen der Mengen g *bzw.* E *heißen eine Parameterdarstellung der Geraden* g *bzw. der Ebene* E. *Der Punkt* a *heißt ein Aufpunkt,* b *ein Richtungsvektor von* g, p *ein Aufpunkt von* E, q *und* r *Richtungsvektoren von* E.

Definition *Sei* $(V, +, \cdot)$ *ein* K*-Vektorraum.*

a) $U \subseteq V$ *heißt Untervektorraum (von* V*), falls* U *mit den Abbildungen* $+|_{K \times U}$ *und* $\cdot|_{K \times U}$ *ein* K*-Vektorraum ist.*

b) $A \subseteq V$ *heißt affiner Raum, falls ein Vektor* $a \in V$ *und ein Unterraum* $U \subseteq V$ *existieren mit* $A = a + U$, *also*

$$A = \{x \in V : \exists\, u \in U \colon x = a + u\}\,.$$

Beispiele

1) Geraden und Ebenen sind affine Räume und genau dann Untervektorräume, wenn sie den Ursprung (den Nullvektor) enthalten.
2) Jeder Vektorraum ist ein Untervektorraum von sich selbst.

Satz 8.3 *Sei* V *ein* K*-Vektorraum und* $U \subseteq V$. *Dann sind äquivalent:*

(i) U *ist ein Untervektorraum.*

(ii) $U \neq \emptyset \ \wedge\ \forall x, y \in U\, \forall \lambda, \mu \in K : \lambda x + \mu y \in U.$

Definition

a) Die Länge bzw. der Betrag bzw. die euklidische Norm von Vektoren im $\mathbb{R}^n$ *ist definiert durch die Abbildung*

$$\|\,.\,\| \colon \mathbb{R}^n \to \mathbb{R}, \qquad x \mapsto \|x\|,$$

wobei

$$\|x\| = \left\| \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \right\| = \left(\sum_{i=1}^{n} x_i^2 \right)^{\frac{1}{2}}.$$

b) Der Abstand zwischen zwei Vektoren im $\mathbb{R}^n$ ist definiert durch die Abbildung

$$d\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}, \qquad (x,y) \mapsto d(x,y),$$

wobei

$$d(x,y) = \|x-y\|.$$

Satz 8.4 *Für alle $x, y \in \mathbb{R}^n$ und alle $\lambda \in \mathbb{R}$ gilt*

$$\begin{array}{lll} (N1) & \|x\| \geq 0 \,\wedge\, (\|x\| = 0 \Leftrightarrow x = 0) & \textit{(Positive Definitheit)} \\ (N2) & \|\lambda x\| = |\lambda|\,\|x\| & \textit{(Homogenität)} \\ (N3) & \|x+y\| \leq \|x\| + \|y\| & \textit{(Dreiecksungleichung)} \end{array}$$

Satz 8.5 *Für alle $x, y, z \in \mathbb{R}^n$ gilt*

$$\begin{array}{lll} (M1) & d(x,y) \geq 0 \,\wedge\, (d(x,y) = 0 \Leftrightarrow x = y) & \textit{(Positive Definitheit)} \\ (M2) & d(x,y) = d(y,x) & \textit{(Symmetrie)} \\ (M3) & d(x,y) \leq d(x,z) + d(z,y) & \textit{(Dreiecksungleichung)} \end{array}$$

Definition *Auf $\mathbb{R}^n$ ist das euklidische Skalarprodukt definiert durch die Abbildung*

$$\langle\,.\,,\,.\,\rangle\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R},$$

wobei

$$\langle x,y\rangle = \left\langle \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right\rangle = \sum_{i=1}^n x_i y_i.$$

Satz 8.6 *Für alle $x, y, z \in \mathbb{R}^n$ und alle $\lambda \in \mathbb{R}$ gilt*

$$\begin{array}{lll} (SP1) & \langle x,x\rangle \geq 0 \,\wedge\, (\langle x,x\rangle = 0 \Leftrightarrow x = 0) & \textit{(Positive Definitheit)} \\ (SP2) & \langle x,y\rangle = \langle y,x\rangle & \textit{(Symmetrie)} \\ (SP3) & \langle \lambda x + y, z\rangle = \lambda\langle x,z\rangle + \langle y,z\rangle & \textit{(Linearität in der 1. Komponente)} \end{array}$$

Außerdem gilt

$$\|x\| = \sqrt{\langle x,x\rangle}\ .$$

Korollar 8.7 *Für alle $x, y, z \in \mathbb{R}^n$ und alle $\lambda \in \mathbb{R}$ gilt*

$$\langle z, \lambda x + y\rangle = \lambda\langle z,x\rangle + \langle z,y\rangle,$$

also die Linearität in der 2. Komponente.

Satz 8.8 (Parallelogrammgleichung) *Für alle $x, y \in \mathbb{R}^n$ gilt*

$$\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2).$$

Beweis
Es gilt

$$\begin{aligned}\|x+y\|^2+\|x-y\|^2 &= \langle x+y,x+y\rangle+\langle x-y,x-y\rangle \\ &= \langle x,x\rangle+\langle x,y\rangle+\langle y,x\rangle+\langle y,y\rangle \\ &\quad +\langle x,x\rangle-\langle x,y\rangle-\langle y,x\rangle+\langle y,y\rangle \\ &= 2(\|x\|^2+\|y\|^2).\end{aligned}$$

□

Korollar 8.9 *Für $x, y \in \mathbb{R}^n$ gilt*

$$\langle x,y\rangle = \frac{1}{4}(\|x+y\|^2-\|x-y\|^2).$$

Beweis
durch Nachrechnen unter Verwendung der Parallelogrammgleichung. □

Definition *Für $x, y \in \mathbb{R}^n \setminus \{0\}$ heißt*

$$\varphi = \measuredangle(x,y) = \arccos\left(\frac{\langle x,y\rangle}{\|x\|\|y\|}\right)$$

der Winkel (bzw. das Bogenmaß des nicht orientierten Winkels) zwischen x und y.

Beispiel
Seien $x = \begin{pmatrix}1\\0\end{pmatrix}$, $y = \begin{pmatrix}0\\1\end{pmatrix}$ und $z = \begin{pmatrix}1\\1\end{pmatrix}$. Dann ist

$$\measuredangle(x,y) = \arccos\frac{\left\langle\begin{pmatrix}1\\0\end{pmatrix},\begin{pmatrix}0\\1\end{pmatrix}\right\rangle}{1\cdot 1} = \arccos(0) = \frac{\pi}{2}\ ,$$

$$\measuredangle(x,z) = \arccos\frac{\left\langle\begin{pmatrix}1\\0\end{pmatrix},\begin{pmatrix}1\\1\end{pmatrix}\right\rangle}{1\cdot \sqrt{2}} = \arccos\left(\frac{1}{2}\cdot\sqrt{2}\right) = \frac{\pi}{4}\ .$$

Bemerkungen

1) Es gilt $\varphi \in [0,\pi]$ und $\langle x,y\rangle = \|x\|\,\|y\|\,\cos\varphi$.
2) Seien $x, y \in \mathbb{R}^2$ mit $\|x\| = \|y\| = 1$. Dann kann man zeigen, dass $\measuredangle(x,y)$ mit der Länge des kürzeren Kreisbogens zwischen x und y übereinstimmt.

Definition *Zwei Vektoren $x, y \in \mathbb{R}^n$ heißen orthogonal zueinander, in Zeichen $x \perp y$, falls $\langle x,y\rangle = 0$ ist.*

Satz 8.10 (Satz des Pythagoras) *Seien $x, y \in \mathbb{R}^n$ mit $x \perp y$. Dann gilt*

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2.$$

Beweis
Da $x \perp y$ ist, gilt $\langle x, y\rangle = 0$. Damit folgt

$$\begin{aligned} \|x + y\|^2 &= \langle x + y, x + y\rangle \\ &= \langle x, y\rangle + \langle x, y\rangle + \langle y, x\rangle + \langle y, y\rangle \\ &= \langle x, x\rangle + 2\langle x, y\rangle + \langle y, y\rangle \\ &= \|x\|^2 + \|y\|^2. \end{aligned}$$

□

Definition *Auf $\mathbb{R}^3$ ist das Vektorprodukt bzw. Kreuzprodukt definiert durch die Abbildung*

$$\times \colon \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$$

mit

$$a \times b = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \times \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} a_2b_3 - b_2a_3 \\ a_3b_1 - b_3a_1 \\ a_1b_2 - b_1a_2 \end{pmatrix}.$$

Satz 8.11 *Für $e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$, $e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ und für alle $a, b, c \in \mathbb{R}^3$ und alle $\lambda \in \mathbb{R}$ gilt*

(i) $e_i \times e_i = 0$ *für* $i = 1, 2, 3$,

(ii) $e_1 \times e_2 = e_3$ *und* $e_2 \times e_3 = e_1$ *und* $e_3 \times e_1 = e_2$,

(iii) $a \times b = -(b \times a)$,

(iv) $(\lambda a + b) \times c = \lambda(a \times c) + b \times c$,

(v) $\langle a \times b, a\rangle = \langle a \times b, b\rangle = 0$,

(vi) $\|a \times b\| = \|a\| \cdot \|b\| \cdot \sin(\measuredangle(a, b))$.

Definition

a) Der Flächeninhalt des von $a, b \in \mathbb{R}^3$ aufgespannten Parallelogramms P, d.h.

$$P = \left\{x \in \mathbb{R}^3 \colon \exists \lambda, \mu \in [0, 1] \colon x = \lambda a + \mu b\right\},$$

ist definiert durch $\|a \times b\|$.

b) Der Flächeninhalt des von a, b aufgespannten Dreiecks D, d.h.

$$D = \left\{ x \in \mathbb{R}^3 \colon \exists \lambda \in [0,1] \, \exists \mu \in [0, \lambda] \colon x = \lambda a + \mu(b-a) \right\},$$

ist definiert durch $\frac{1}{2}\|a \times b\|$.

Definition

a) Auf $\mathbb{R}^3$ ist das Spatprodukt definiert durch die Abbildung

$$[\,.\,,\,.\,,\,.\,] \colon \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}, \qquad [a,b,c] = \langle a \times b, c \rangle.$$

b) Das Volumen des von $a, b, c \in \mathbb{R}^3$ aufgespannten Spats S, d.h.

$$S = \left\{ x \in \mathbb{R}^3 \colon \exists \lambda, \mu, \nu \in [0,1] \colon x = \lambda a + \mu b + \nu c \right\},$$

ist definiert durch $|[a,b,c]|$.

Satz 8.12 *Sei $E \subseteq \mathbb{R}^3$ eine Ebene. Dann gibt es einen Vektor $n = \begin{pmatrix} n_1 \\ n_2 \\ n_3 \end{pmatrix} \in \mathbb{R}^3$ mit $\|n\| = 1$ und genau eine Zahl $d \in \mathbb{R}_0^+$, so dass*

$$E = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3 : n_1 x_1 + n_2 x_2 + n_3 x_3 - d = 0 \right\} = \{ x \in \mathbb{R}^3 : \langle n, x \rangle = d \} \qquad (1)$$

gilt. Außerdem gilt:

$$\forall a, b \in E : \langle n, b - a \rangle = 0 \qquad (2)$$

Umgekehrt gibt es für jeden Vektor $n \in \mathbb{R}^3$ mit $\|n\| = 1$ und jede Zahl $d \in \mathbb{R}_0^+$ genau eine Ebene $E \subseteq \mathbb{R}^3$, für die (1) und (2) gilt.

Bemerkungen

1) Eine Gleichung der Form $n_1 x_1 + n_2 x_2 + n_3 x_3 - d = 0$ mit $n_1, n_2, n_3 \in \mathbb{R}$, $\sqrt{n_1^2 + n_2^2 + n_3^2} = 1$ und $d \in \mathbb{R}_0^+$ nennt man Hessesche Normalenform (HNF) der Ebene, die durch die Lösungsmenge dieser Gleichung beschrieben wird.

2) Eine Ebene besitzt genau dann eine eindeutig bestimmte HNF, falls $d > 0$ ist. Im Fall $d = 0$ sind sowohl $n_1 x_1 + n_2 x_2 + n_3 x_3 = 0$ als auch $-n_1 x_1 - n_2 x_2 - n_3 x_3 = 0$ Hessesche Normalenform derselben Ebene.

3) Ein Vektor $n \in \mathbb{R}^3$, der (2) erfüllt, heißt ein Normalenvektor von E. Ist zusätzlich $\|n\| = 1$, dann heißt n ein Normaleneinheitsvektor von E.

4) Seien $q, r \in \mathbb{R}^3 \setminus \{0\}$ zwei Richtungsvektoren einer Ebene $E \subseteq \mathbb{R}^3$ mit $r \neq \lambda q$ für alle $\lambda \in \mathbb{R}$. Dann ist $\{\alpha(q \times r) : \alpha \in \mathbb{R} \setminus \{0\}\}$ die Menge aller Normalenvektoren von E und $\left\{\pm \dfrac{q \times r}{\|q \times r\|}\right\}$ die Menge aller Normaleneinheitsvektoren von E.

Definition *Seien $v_i \in \mathbb{R}^3$, $i = 1{,}2$, Richtungsvektoren der Geraden g_i und $n_i \in \mathbb{R}^3$ Normalenvektoren der Ebenen E_i. Dann heißt*

a) g_1 parallel zu g_2, in Zeichen: $g_1 \| g_2$, falls gilt: $\exists \lambda \in \mathbb{R} \setminus \{0\} : v_1 = \lambda v_2$,

b) $E_1 \| E_2$, falls gilt $\exists \lambda \in \mathbb{R} \setminus \{0\} : n_1 = \lambda n_2$,

c) $g_1 \| E_1$, falls gilt $n_1 \perp v_1$.

8.2 Weitere Beispiele für Vektorräume

Weitere Beispiele für Vektorräume sind:

1) $\mathbb{C}^n = \left\{ \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} : z_1, \dots, z_n \in \mathbb{C} \right\}$ mit

$$\begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} + \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} = \begin{pmatrix} z_1 + w_1 \\ \vdots \\ z_n + w_n \end{pmatrix}, \qquad \lambda \cdot \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} = \begin{pmatrix} \lambda z_1 \\ \vdots \\ \lambda z_n \end{pmatrix}$$

ist ein $\mathbb{C}$-Vektorraum.

2) Sei K ein beliebiger Körper. Dann ist K^n, wobei $+$ und $\cdot$ wie in 1) definiert sind, ein K-Vektorraum.

3) Die Menge aller reellen Folgen, d.h. Abbildungen von $\mathbb{N}$ nach $\mathbb{R}$, $(a_n)_{n\in\mathbb{N}}$ mit

$$(a_n)_{n\in\mathbb{N}} + (b_n)_{n\in\mathbb{N}} = (a_n + b_n)_{n\in\mathbb{N}}, \qquad \lambda(a_n)_{n\in\mathbb{N}} = (\lambda a_n)_{n\in\mathbb{N}}$$

ist ein $\mathbb{R}$-Vektorraum.

4) Die Menge aller reellen Funktionen $f : \mathbb{R} \to \mathbb{R}$ mit

$$f + g : x \mapsto f(x) + g(x), \qquad \lambda f : x \mapsto \lambda f(x)$$

ist ein $\mathbb{R}$-Vektorraum.
In diesem Vektorraum ist die Menge aller reellen Polynome $p : \mathbb{R} \to \mathbb{R}$ ein Untervektorraum, der wiederum die Mengen aller reellen Polynome vom Grade $\leq n$ für jedes $n \in \mathbb{N} \cup \{0, -1\}$ als Untervektorräume enthält.

5) In 3) und 4) kann $\mathbb{R}$ auch durch $\mathbb{C}$ ersetzt werden.

8.3 Basis und Dimension

Definition *Sei V ein K-Vektorraum und $M \subseteq V$.*

a) *Ein Vektor $x \in V$ heißt eine Linearkombination von Vektoren aus M, falls es endlich viele Vektoren $v_1, \ldots, v_n \in M$ gibt, so dass $x = \sum_{i=1}^{n} \lambda_i v_i$ mit $\lambda_i \in K$ für $i = 1, \ldots, n$ gilt. x heißt dann auch eine Linearkombination von $v_1, \ldots, v_n$.*

b) *Ist $M \neq \emptyset$, dann heißt die Menge aller Linearkombinationen von Vektoren aus M der Aufspann von M oder die lineare Hülle von M, in Zeichen: $Span(M)$. Für $\emptyset$ definiert man $Span(\emptyset) = \{0\}$.*

c) *M heißt ein Erzeugendensystem von V, falls $Span(M) = V$ ist.*

d) *M heißt linear unabhängig (l.u.), falls für jede endliche Teilmenge $\{v_1, \ldots, v_n\}$ von M gilt:*

$$\forall \lambda_1, \ldots, \lambda_n \in K : \sum_{i=1}^{n} \lambda_i v_i = 0 \Rightarrow \lambda_1 = \cdots = \lambda_n = 0.$$

In diesem Fall heißen auch die Vektoren $v_1, \ldots, v_n$ linear unabhängig. M heißt linear abhängig (l.a.), falls M nicht linear unabhängig ist.

e) *$B \subseteq V$ heißt Basis von V, falls B linear unabhängig und ein Erzeugendensystem von V ist.*

Beispiele

1) $V = K^n$,

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots, e_n = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

$\{e_1, \ldots, e_n\}$ ist eine Basis von K^n, die sogenannte kanonische Basis oder Standardbasis.

2) $V = \mathbb{R}^2$,

$v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $v_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $v_3 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$.

Dann ist:

- v_3 eine Linearkombination von v_1 und v_2, denn es ist $v_3 = v_1 + v_2$,
- $\{v_1, v_2, v_3\}$ linear abhängig,

- $\{v_1, v_2\}$ linear unabhängig,
- $Span(\{v_1\})$ eine Ursprungsgerade mit Richtungsvektor v_1,
- $\{v_1, v_2, v_3\}$ ein Erzeugendensystem von V,
- $\{v_1, v_2\}$ eine Basis von V.

3) Sei $p_j : \mathbb{R} \to \mathbb{R}, x \mapsto x^j$.
Dann ist $\{p_0, p_1, \ldots, p_n\}$ eine Basis des Vektorraums der reellen Polynome von Grad $\leq n$.

4) Sei 0 das neutrale Element eines Körpers K.
Dann ist $\emptyset$ eine Basis des K-Vektorraums $\{0\}$.

Satz 8.13 *Sei V ein K-Vektorraum und $M \subseteq V$ mit $M \neq 0$. Dann sind äquivalent:*

(i) *M ist linear unabhängig.*

(ii) *Für jedes Element aus $Span(M)$ gibt es eine eindeutige Darstellung als Linearkombination von Vektoren aus M, d.h.*
$\forall x \in Span(M) \setminus \{0\}\, \exists!\, n \in \mathbb{N}\, \exists!\, (v_1, \ldots, v_n) \in M^n$ *mit paarweise verschiedenen Vektoren* v_i $\exists!\, (\lambda_1, \ldots, \lambda_n) \in K^n : x = \sum\limits_{i=1}^{n} \lambda_i v_i \wedge \forall i \in \{1, \ldots, n\} : \lambda_i \neq 0.$

Satz 8.14 *Sei V ein K-Vektorraum und $B \subseteq V$. Dann sind äquivalent:*

(i) *B ist eine Basis von V.*

(ii) *B ist eine maximale linear unabhängige Menge von V, d.h.*
B ist l.u. $\wedge\ \forall B' \in \mathcal{P}(V) : B \subseteq B' \wedge B'$ *l.u.* $\Rightarrow B' = B$.

(iii) *B ist ein minimales Erzeugendensystem von V, d.h.*
$Span(B) = V \ \wedge\ \forall B' \in \mathcal{P}(V) : B' \subseteq B \ \wedge\ Span(B') = V \Rightarrow B' = B$.

Ist $B \neq \emptyset$, dann sind $(i), (ii), (iii)$ auch äquivalent zu

(iv) $\forall x \in V \setminus \{0\}\ \exists! n \in \mathbb{N}\ \exists!\, (b_1, \ldots, b_n) \in B^n$ *mit paarweise verschiedenen* $b_i \wedge$
$\exists!\, (\lambda_1, \ldots, \lambda_n) \in K^n : x = \sum\limits_{i=1}^{n} \lambda_i b_i \ \wedge\ \forall i \in \{1, \ldots, n\} : \lambda_i \neq 0.$
Jedes λ_i heißt dann Koordinate von x bezüglich des Basisvektors b_i.

Satz 8.15 (Basisergänzungssatz) *Sei V ein K-Vektorraum mit einer endlichen Basis $B = \{b_1, \ldots, b_n\}$ und $M = \{v_1, \ldots, v_m\}$ sei eine linear unabhängige Teilmenge von V. Dann ist $m \leq n$ und es existieren $n - m$ Elemente aus B, durch die M zu einer Basis von V ergänzt werden kann.*

Beweisskizze
Vollständige Induktion nach m:
$m = 0$, d.h. $M = \emptyset$: M kann durch $b_1, \ldots, b_n$ zu einer Basis von V ergänzt werden.
$m > 0$, $m - 1 \to m$:
Induktionsvoraussetzung: Die Behauptung gelte für $m - 1$.
Dann ist $m - 1 \leq n$ und $\exists b_{j_1}, \ldots, b_{j_{n-m+1}} \in B$, so dass
$B' = \{v_1, \ldots, v_{m-1}, b_{j_1}, \ldots, b_{j_{n-m+1}}\}$ eine Basis von V ist.
Da $\{v_1, \ldots, v_m\}$ l.u. ist, folgt:

$$v_m = \sum_{j=1}^{m-1} \lambda_j v_j + \sum_{i=1}^{n-m+1} \lambda_{j_i} b_{j_i} \ \wedge \exists k \in \{1, \ldots, n-m+1\} : \lambda_{j_k} \neq 0 \wedge m \leq n.$$

Dann kann in B' das Element b_{j_k} durch v_m ersetzt werden und
$\{v_1, \ldots, v_m, b_{j_1}, \ldots, b_{j_{k-1}}, b_{j_{k+1}}, \ldots, b_{j_{n-m+1}}\}$ ist eine Basis von V.

Korollar 8.16 *Sei V ein K-Vektorraum mit einer endlichen Basis $\{b_1, \ldots, b_n\}$. Dann hat jede Basis von V genau n Elemente.*

Definition *Besitzt ein K-Vektorraum V eine Basis mit genau n Elementen, dann heißt n die Dimension von V und V heißt n-dimensional.*
Bezeichnung: $n = \dim V$.
Besitzt V keine endliche Basis, dann heißt V unendlich-dimensional.
Bezeichnung: $\dim V = \infty$.

Beispiele

1) $\dim K^n = n$

2) $\dim \{0\} = 0$

3) Der Vektorraum der reellen Polynome vom Grad $\leq n$ ist $n + 1$ - dimensional, der Vektorraum aller reellen Polynome unendlich-dimensional.

Satz 8.17 *Jeder Vektorraum besitzt eine Basis.*

Beweisstrategie
Besitzt ein Vektorraum ein endliches Erzeugendensystem, dann kann daraus mit Hilfe der Aussonderungsregel eine maximale l.u. Teilmenge ausgesondert werden, die nach 8.14 Basis des Vektorraums ist.
Andernfalls kann man mit Hilfe des Auswahlaxioms eine Basis des Vektorraums bilden (Bildung eines minimalen Erzeugendensystems).

9 Lineare Abbildungen

9.1 Grundlegende Eigenschaften

Definition *Seien V, W K-Vektorräume.*

a) *Eine Abbildung $L : V \to W$ heißt linear oder Vektorraum-Homomorphismus, falls gilt:*

$$\forall x, y \in V \, \forall \lambda, \mu \in K : L(\lambda x + \mu y) = \lambda L(x) + \mu L(y)$$

Kurzschreibweise: Lx statt $L(x)$ für alle $x \in V$.

b) *Eine lineare Abbildung $L : V \to K$ heißt Linearform.*

c) *Eine bijektive lineare Abbildung $L : V \to W$ heißt (Vektorraum-) Isomorpismus.*
Existiert ein Isomorphismus $L : V \to W$, dann heißen V und W isomorph.

d) *Eine Abbildung $A : V \to W$ heißt affin, falls ein Vektor $a \in W$ und eine lineare Abbildung $L : V \to W$ existieren, so dass gilt:*

$$\forall x \in V : A(x) = L(x) + a.$$

Beispiele

1) Sei V ein beliebiger Vektorraum. Dann ist die identische Abbildung $Id : V \to V, x \mapsto x$ linear und bijektiv.

2) Jede lineare Abbildung von $\mathbb{R}$ nach $\mathbb{R}$ ist von der Form $x \mapsto ax$ für ein $a \in \mathbb{R}$.

3) Sei $\varphi \in [-2\pi, 2\pi]$. Dann ist $D_\varphi : \mathbb{R}^2 \to \mathbb{R}^2$,
$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} (\cos\varphi)\, x_1 - (\sin\varphi)\, x_2 \\ (\sin\varphi)\, x_1 + (\cos\varphi)\, x_2 \end{pmatrix}$$
linear und bijektiv. D_φ beschreibt eine Drehung um den Winkel φ.

4) $S_{x_1} : \mathbb{R}^2 \to \mathbb{R}^2$,
$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} x_1 \\ -x_2 \end{pmatrix}$$
ist linear und bijektiv. S_{x_1} beschreibt eine Spiegelung an der x_1-Achse.

5) Sei $a \in \mathbb{R}^2$. Dann ist $T_a : \mathbb{R}^2 \to \mathbb{R}^2$, $x \mapsto x + a$ eine affine Abbildung, aber keine lineare Abbildung. T_a beschreibt eine Verschiebung um den Vektor a.

6) Seien $i, j \in \{1, 2, 3\}$ und $a_{ij} \in \mathbb{R}$. Dann sind alle linearen Abbildungen von $\mathbb{R}^3$ nach $\mathbb{R}^3$ von der Form
$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mapsto \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 \end{pmatrix}.$$

7) $f : \mathbb{R}^3 \to \mathbb{R}^3$,
$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mapsto \begin{pmatrix} x_1^2 \\ x_1 + x_2 \\ sin\, x_3 \end{pmatrix}$$
ist nicht linear.

8) Sei P_n der Vektorraum der reellen Polynome vom Grad $\leq n$. Dann ist
$$L : \mathbb{R}^{n+1} \to P_n, \begin{pmatrix} a_1 \\ \vdots \\ a_{n+1} \end{pmatrix} \mapsto a_{n+1}x^n + \cdots + a_2 x + a_1$$
ein Isomorphismus.

9) Sei V ein n-dimensionaler K-Vektorraum und $B = \{b_1, \ldots, b_n\}$ eine Basis von V. Dann ist
$$L : V \to K^n : x = \sum_{i=1}^{n} x_i b_i \mapsto \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$
ein Isomorphismus, die sogenannte Koordinatendarstellung von x bzgl. der Basis B.

10) Sei V der Vektorraum aller reellen Funktionen. Dann ist $L : V \to \mathbb{R}$, $f \mapsto f(0)$ eine Linearform.

Definition *Seien V, W Vektorräume und sei $L : V \to W$ linear. Dann heißt*

a) *$Bild(L) = L(V) = \{L(x) : x \in V\}$ das Bild von L,*

b) *$Kern(L) = L^{-1}(\{0\}) = \{x \in V : L(x) = 0\}$ der Kern von L.*

Satz 9.1 *Seien V, W Vektorräume und sei $L : V \to W$ linear. Dann gilt:*

(i) $L(0) = 0$

(ii) *Alle Bildmengen von Untervektorräumen von V sind Untervektorräume von W, insbesondere auch $Bild(L)$.*

(iii) *Alle Urbildmengen von Untervektorräumen von W sind Untervektorräume von V, insbesondere auch $Kern(L)$.*

(iv) *L ist injektiv $\Leftrightarrow$ $Kern(L) = \{0\}$.*

Beweisskizze
(i) Sei $x \in V$. Mit Hilfe von Satz 8.2(i) folgt $L(0) = L(0x) = 0 \cdot L(x) = 0$.
$(ii), (iii)$ folgt aus der Definition einer linearen Abbildung und Satz 8.3.
(iv) "$\Rightarrow$": Nach (i) ist $L(0) = 0$. Da L injektiv ist, muss $L^{-1}(\{0\}) = \{0\}$, also

$Kern(L) = \{0\}$ sein.
"$\Leftarrow$": Sei $Kern(L) = \{0\}$. Dann gilt für $x, y \in V$:
$Lx = Ly \Rightarrow Lx - Ly = 0 \Rightarrow L(x-y) = 0 \Rightarrow x - y \in Kern(L) \Rightarrow x - y = 0 \Rightarrow x = y$
Also ist L injektiv.

Satz 9.2 *Seien U, V, W K-Vektorräume. Dann gilt:*

a) Sind $J : U \to V$ und $L : U \to V$ linear, dann sind für alle $\alpha, \beta \in K$ auch $\alpha J + \beta L : U \to V$ linear.

b) Sind $L : U \to V$ und $S : V \to W$ linear, dann ist auch $S \circ L : U \to W$ linear.

c) Ist $T : V \to V$ Isomorphismus, dann ist auch $T^{-1} : V \to V$ Isomorphismus.

d) Die Menge aller linearen Abbildungen von U nach V ist mit der Addition und der Skalarmultiplikation aus a) ein K-Vektorraum.

e) Die Menge aller linearen Abbildungen von V nach V ist mit den Verknüpfungen $+$ und $\circ$ ein Ring.

f) Die Menge aller Isomorphismen von V nach V ist bezüglich $\circ$ eine Gruppe.

Satz 9.3 (Dimensionsformel) *Seien V, W K-Vektorräume, wobei V endlich-dimensional sei, und $L : V \to W$ sei linear. Dann gilt:*

$$dim\, V = dim\, Kern(L) + dim\, Bild(L)$$

Beispiel
$L : \mathbb{R}^3 \to \mathbb{R}^3, \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mapsto \begin{pmatrix} x_1 \\ x_2 \\ 0 \end{pmatrix}$ ist linear (orthogonale Projektion auf die x_1-x_2-Ebene).
$Kern(L) = \left\{ \begin{pmatrix} 0 \\ 0 \\ \alpha \end{pmatrix} : \alpha \in \mathbb{R} \right\}$,
$dim(Kern(L)) = 1$,
$Bild(L) = \left\{ \begin{pmatrix} \lambda \\ \mu \\ 0 \end{pmatrix} : \lambda, \mu \in \mathbb{R} \right\}, dim(Bild(L)) = 2 = 3 - 1.$

Beweis Satz 9.3
Sei $dim\, V = n, dim\,(Kern(L)) = k$ und B_k eine Basis von $Kern(L)$.
Zu zeigen: $dim(Bild(L)) = n - k$.
Aufgrund des Basisergänzungssatzes 8.15 existiert eine Basis $B = \{b_1, \ldots, b_n\}$ von V, so dass $B_k = B \setminus \{b_{k+1}, \ldots, b_n\}$ ist. Da L linear ist, gilt:
$Bild(L) = L(V) = L(Span\{b_1, \ldots, b_n\}) = Span\{Lb_1, \ldots, Lb_n\} = Span\{Lb_{k+1}, \ldots, Lb_n\}$.

$\{Lb_{k+1}, \ldots, Lb_n\}$ ist l.u., denn sei

$$0 = \sum_{j=k+1}^{n} \lambda_j Lb_j = L\left(\sum_{j=k+1}^{n} \lambda_j b_j\right)$$

$$\Rightarrow \sum_{j=k+1}^{n} \lambda_j b_j \in Kern(L)$$

$$\Rightarrow \sum_{j=k+1}^{n} \lambda_j b_j = \sum_{j=1}^{k} (-\lambda_j) b_j \text{ mit } (-\lambda_j) \in K$$

(im Falle $k = 0$ sei hierbei $\sum_{j=1}^{0} \cdots := 0$)

$$\Rightarrow \sum_{j=1}^{n} \lambda_j b_j = 0$$

$\Rightarrow \lambda_1 = \cdots = \lambda_n = 0$ (da B l.u. ist)
$\Rightarrow \lambda_{k+1} = \cdots = \lambda_n = 0$
$\{Lb_{k+1}, \ldots, Lb_n\}$ ist somit eine Basis von $Bild(L)$.
Also ist $dim\, Bild(L) = n - k$. □

Definition *Seien V, W Vektorräume und sei $L : V \to W$ linear. Dann heißt $Rang(L) := dim\, Bild(L)$ der Rang von L.*

Satz 9.4 *Seien V, W K-Vektorräume, B eine Basis von V und $f : B \to W$ eine beliebige Abbildung. Dann gibt es genau eine lineare Abb. $L : V \to W$ mit $L(y) = f(y)$ für alle $y \in B$.*

Beweis
Existenz:
Sei $x \in V \setminus \{0\}$. Nach Satz 8.14 gilt:
$\exists! n \in \mathbb{N}\, \exists! \,(b_1, \ldots, b_n) \in B^n$ mit $b_1 \neq \cdots \neq b_n$ $\exists! \,(\lambda_1 \ldots, \lambda_n) \in K^n : x = \sum_{i=1}^{n} \lambda_i b_i \wedge \forall i \in \{1, \ldots, n\} : \lambda_i \neq 0$.
Definiere $L(0) = 0$ und $\forall x \neq 0 : L(x) := \sum_{i=1}^{n} \lambda_i f(b_i)$.
Dann ist $L : V \to W, x \mapsto L(x)$ linear mit $L(y) = f(y)$ für alle $y \in B$ (Nachweis durch direktes Nachrechnen).
Eindeutigkeit:
Sei $\tilde{L} : V \to W$ linear mit $\tilde{L}(y) = f(y)$ für alle $y \in B$ und sei $x = \sum_{i=1}^{n} \lambda_i b_i$ mit $b_i \in B$.
Dann gilt:

$$\tilde{L}(x) = \tilde{L}\left(\sum_{i=1}^{n} \lambda_i b_i\right) = \sum_{i=1}^{n} \lambda_i \tilde{L}(b_i) = \sum_{i=1}^{n} \lambda_i f(b_i) = L(x).$$

Also ist $\tilde{L} = L$. □

9.2 Lineare Abbildungen und Matrizen

Defintion *Sei K ein Körper und $m, n \in \mathbb{N}$. Eine Abbildung $A : \{1, \ldots, m\} \times \{1, \ldots, n\} \to K, (i, j) \mapsto a_{ij} \in K$ heißt $m \times n$-Matrix über K.*

Eine $m \times n$-Matrix A wird dargestellt als rechteckiges Schema der Form:

$$A = (a_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}.$$

Die Vektoren $\begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix} \in K^m$ *heißen Spaltenvektoren, die Vektoren* $(a_{i1}, \dots, a_{in}) \in K^n$ *heißen Zeilenvektoren von A. Die Menge aller $m \times n$-Matrizen über K wird mit $K^{m \times n}$ bezeichnet.*

Definition *Das Produkt einer Matrix $A \in K^{m \times n}$ mit einem Vektor $x \in K^n$ ist definiert durch den Vektor $Ax \in K^m$ mit*

$$Ax = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \sum\limits_{j=1}^{n} a_{1j} x_j \\ \vdots \\ \sum\limits_{j=1}^{n} a_{mj} x_j \end{pmatrix}.$$

Beispiel

$A = \begin{pmatrix} 1 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix}$, $x = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, $Ax = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$

Satz 9.5 *Sei V ein K-Vektorraum mit endlicher Basis $B = \{b_1, \dots, b_n\}$, W ein K-Vektorraum mit endlicher Basis $C = \{c_1, \dots, c_m\}$.*

a) Sei $L : V \to W$ linear. Dann existiert genau eine Matrix $M = M_L^{C,B} \in K^{m \times n}$, so dass gilt:

Ist $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in K^n$ *die Koordinatendarstellung von $x \in V$ bezüglich B, dann ist*

$M \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in K^m$ *die Koordinatendarstellung von Lx bezüglich C.*

Insbesondere ist der j-te Spaltenvektor $\begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}$ *von M die Koordinatendarstellung von Lb_j bezüglich C.*

b) Sei $M \in K^{m\times n}$. Dann existiert genau eine lineare Abbildung $L : V \to W$, so dass $M = M_L^{C,B}$ ist.

Bemerkung
Wir verwenden hier und im Folgenden die Schreibweise $B = \{b_1, \dots, b_n\}$ zur Bezeichnung einer geordneten Basis, d.h. eines n-Tupels $(b_1, \dots, b_n) \in V^n$ mit der Eigenschaft, dass $\{b_1, \dots, b_n\}$ eine Basis von V ist. Die Indizes legen also die Reihenfolge der Basisvektoren fest. Entsprechendes gilt für $C = \{c_1, \dots, c_m\}$.

Bezeichnung
Sei $M \in K^{m\times n}$. Dann sei $L_M : K^n \to K^m$ die eindeutig bestimmte lineare Abbildung, für die $M = M_{L_M}^{\mathcal{E}_m,\mathcal{E}_n}$, wobei $\mathcal{E}_j = \{e_1, \dots, e_j\}$ die geordnete kanonische Basis von K^j mit $j \in \{m, n\}$ sei.

Beispiel

Sei $V = \mathbb{R}^3$, $W = \mathbb{R}^2$, $B = \mathcal{E}_3$ und $C = \mathcal{E}_2$.

$$L : \mathbb{R}^3 \to \mathbb{R}^2,\ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mapsto \begin{pmatrix} x_1 + 2x_2 \\ x_1 + x_3 \end{pmatrix} \ \Rightarrow\ L\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix},$$

$$M = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \end{pmatrix} \ \Rightarrow\ L_M : \mathbb{R}^3 \to \mathbb{R}^2,\ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mapsto \begin{pmatrix} x_1 + 2x_2 \\ x_2 \end{pmatrix}.$$

Beweisskizze zu Satz 9.5

Sei $x = \sum_{j=1}^{n} x_j b_j$. Dann gilt:

$$\begin{aligned}
Lb_j &= \sum_{i=1}^{m} a_{ij} c_i \ \text{ mit } \ a_{ij} \in K \\
\Leftrightarrow\ L(x) &= L\left(\sum_{j=1}^{n} x_j b_j\right) = \sum_{j=1}^{n} x_j L(b_j) \\
&= \sum_{j=1}^{n} \left(\sum_{i=1}^{m} a_{ij} c_i\right) x_j \\
&= \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} x_j\right) c_i \\
\Leftrightarrow\ M_L^{C,B} &= (a_{ij}) \in K^{m\times n}
\end{aligned}$$

Definition

a) Auf $K^{m\times n}$ sind eine Addition und eine Skalarmultiplikation definiert durch:

$$\forall A = (a_{ij}) \in K^{m\times n}\ \forall B = (b_{ij}) \in K^{m\times n}:\ A + B := (a_{ij} + b_{ij}),$$

$$\forall \lambda \in K\ \forall A = (a_{ij}) \in K^{m\times n}:\ \lambda A := (\lambda a_{ij}).$$

b) Die Matrizenmultiplikation ist eine Abbildung von $K^{l\times m} \times K^{m\times n}$ nach $K^{l\times n}$, definiert durch

$$\forall A \in K^{l\times m}\ \forall B \in K^{m\times n}:\ AB = C := (c_{ij}) \in K^{l\times n} \text{ mit } c_{ij} = \sum_{k=1}^{m} a_{ik}b_{kj}.$$

c) Eine Matrix $A \in K^{n\times n}$ heißt invertierbar, falls eine Matrix $A^{-1} \in K^{n\times n}$ existiert mit $A^{-1}A = E_n = AA^{-1}$, wobei

$$E_n = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix} = (\delta_{ij}) \text{ mit } \delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}.$$

A^{-1} heißt inverse Matrix zu A und E_n die Einheitsmatrix.

Beispiel

$A = \begin{pmatrix} 1 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix}$, $M = \begin{pmatrix} 1 & 0 & 1 \\ 2 & 2 & 2 \end{pmatrix}$, $B = \begin{pmatrix} 2 & 3 \\ 1 & -1 \\ 1 & 0 \end{pmatrix}$, $N = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$

$\Rightarrow A + M = \begin{pmatrix} 2 & 2 & 1 \\ 3 & 2 & 3 \end{pmatrix}$, $2A = \begin{pmatrix} 2 & 4 & 0 \\ 2 & 0 & 2 \end{pmatrix}$, $AB = \begin{pmatrix} 1 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 3 \\ 1 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 4 & 1 \\ 3 & 3 \end{pmatrix}$,

$N^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{pmatrix}$.

Satz 9.6 *Seien U, V, W K-Vektorräume mit $\dim U = n \geq 1$, $\dim V = m \geq 1$, $\dim W = l \geq 1$ und geordneten Basen B_U, B_V, B_W. Dann gilt:*

a) Sind $J : U \to V$ und $L : U \to V$ linear und haben die Matrixdarstellungen $M_J^{B_V,B_U} \in K^{m\times n}$ bzw. $M_L^{B_V,B_U} \in K^{m\times n}$, dann folgt:

$$\forall \alpha, \beta \in K : M_{\alpha J+\beta L}^{B_V,B_U} = \alpha M_J^{B_V,B_U} + \beta M_L^{B_V,B_U}.$$

b) Sind $L : U \to V$ und $S : V \to W$ linear mit den Matrixdarstellungen $M_L^{B_V,B_U} \in K^{m\times n}$ bzw. $M_S^{B_W,B_V} \in K^{l\times m}$, dann folgt:

$$M_{S\circ L}^{B_W,B_U} = M_S^{B_W,B_V} M_L^{B_V,B_U}.$$

c) Ist $T : V \to V$ Isomorphismus mit Matrixdarstellung $M_T^{B_V,B_V} \in K^{m\times m}$, dann folgt:

$$M_{T^{-1}}^{B_V,B_V} = (M_T^{B_V,B_V})^{-1}.$$

Beweisstrategie:
Durch Nachrechnen, z.B. bei b):
Sei $B_U = \{a_1,\ldots,a_n\}, B_V = \{b_1,\ldots,b_m\}, B_W = \{c_1,\ldots,c_l\}$, $M_S^{B_W,B_V} = (\alpha_{ik})$ und $M_L^{B_V,B_U} = (\beta_{kj})$. Dann gilt:
$S \circ L(a_j) = S(L(a_j)) = S\left(\sum\limits_{k=1}^{m} \beta_{kj} b_k\right) = \sum\limits_{k=1}^{m} \beta_{kj} S(b_k) = \sum\limits_{k=1}^{m} \beta_{kj} \left(\sum\limits_{i=1}^{l} \alpha_{ik} c_i\right)$
$= \sum\limits_{i=1}^{l} \left(\sum\limits_{k=1}^{m} \alpha_{ik}\beta_{kj}\right) c_i.$

Satz 9.7 *a) $K^{m\times n}$ ist mit der oben def. Addition und Skalarmultiplikation ein K-Vektorraum, welcher isomorph zu $K^{m\cdot n}$ sowie zum Vektorraum aller linearen Abbildungen von V nach W ist, wobei V ein beliebiger n-dimensionaler und W ein beliebiger m-dimensionaler K-Vektorraum ist, z.B. K^n und K^m.*

b) $K^{n\times n}$ ist bzgl. der Addition und der Matrixmultiplikation ein Ring.

c) Die Menge aller invertierbaren Matrizen aus $K^{n\times n}$ ist bezüglich der Matrizenmultiplikation eine Gruppe, die sogenannte allgemeine lineare Gruppe $GL(n, K)$.

Beweisstrategie
Folgt durch Kombination von Satz 9.2, Satz 9.5 und Satz 9.6 oder durch direktes Nachrechnen der Vektorraum-, Ring- und Gruppenaxiome und Anwendung der Sätze 9.5 und 9.6 zum Nachweis der Isomorphieeigenschaft.

Bemerkung
Für $n \geq 2$ sind weder der Ring $K^{n\times n}$ noch die Gruppe $GL(n, K)$ kommutativ, da die Matrixmultiplikation im Allgemeinen nicht kommutativ ist. Denn es gilt z.B.
$A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \Rightarrow AB = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}, BA = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \Rightarrow AB \neq BA$

Korollar 9.8 *a) $A \in K^{n\times n} \wedge \exists B \in K^{n\times n}$: $BA = E_n \Rightarrow A$ ist invertierbar und A^{-1} ist eindeutig bestimmt.*

b) $A \in K^{n\times n}$ ist invertierbar $\Rightarrow A^{-1}$ ist invertierbar mit $(A^{-1})^{-1} = A$.

c) $A, B \in K^{n\times n}$ sind invertierbar $\Rightarrow AB \in K^{n\times n}$ ist invertierbar mit $(AB)^{-1} = B^{-1}A^{-1}$.

Beweisstrategie
Folgt aus Satz 9.7 c) und Satz 3.2 $(i), (ii), (v), (vi)$.

Korollar 9.9 *Sei V ein n-dimensionaler K-Vektorraum mit einer geordneten Basis B_V und sei $L: V \to V$ linear. Dann gilt :*

$$L \text{ ist bijektiv.} \Leftrightarrow M_L^{B_V,B_V} \text{ ist invertierbar.}$$

Definition *Sei V ein endlich-dimensionaler K-Vektorraum mit geordneten Basen B und B'. Dann heißt die Matrix $M_{Id_V}^{B',B}$ die Basiswechselmatrix (oder Übergangsmatrix) von B zu B'. Dabei ist der j-te Spaltenvektor von $M_{Id_V}^{B',B}$ die Darstellung des j-ten Basisvektors der alten Basis B in den Koordinaten der neuen Basis B'.*

Satz 9.10 (Koordinatentransformation) *Seien V, W endlich-dimensionale K-Vektorräume, B, B' geordnete Basen von V und C, C' geordnete Basen von W. Sei $L: V \to W$ linear, $M = M_L^{C,B}, R = M_{Id_V}^{B,B'}$ und $S = M_{Id_W}^{C,C'}$. Dann gilt:*

a) $M_L^{C',B} = M_{Id_W}^{C',C} M_L^{C,B} = \left(M_{Id_W}^{C,C'}\right)^{-1} M_L^{C,B} = S^{-1}M,$

b) $M_L^{C,B'} = M_L^{C,B} M_{Id_V}^{B,B'} = MR,$

c) $M_L^{C',B'} = M_{Id_W}^{C',C} M_L^{C,B} M_{Id_V}^{B,B'} = S^{-1}MR.$

d) Ist $V = W$, $B = C$ und $B' = C'$, dann folgt $M_L^{B',B'} = S^{-1}M_L^{B,B}S$.

Beweisstrategie
Durch Nachrechnen.

Beispiel
$V = W = \mathbb{R}^2$, $B = \{b_1, b_2\}$ mit $b_1 = \begin{pmatrix}1\\0\end{pmatrix}$, $b_2 = \begin{pmatrix}0\\1\end{pmatrix}$, $B' = \{b_1', b_2'\}$ mit $b_1' = \begin{pmatrix}2\\0\end{pmatrix}$, $b_2' = b_2$

$M_L^{B,B} = \begin{pmatrix}1 & 2\\1 & 0\end{pmatrix}$

$\Rightarrow M_{Id_{\mathbb{R}^2}}^{B,B'} = \begin{pmatrix}2 & 0\\0 & 1\end{pmatrix} =: S$

$\Rightarrow M_{Id_{\mathbb{R}^2}}^{B',B} = \left(M_{Id_{\mathbb{R}^2}}^{B,B'}\right)^{-1} = \begin{pmatrix}\frac{1}{2} & 0\\0 & 1\end{pmatrix} = S^{-1}$

$\Rightarrow M_L^{B',B'} = S^{-1}M_L^{B,B}S = \begin{pmatrix}\frac{1}{2} & 0\\0 & 1\end{pmatrix}\begin{pmatrix}1 & 2\\1 & 0\end{pmatrix}\begin{pmatrix}2 & 0\\0 & 1\end{pmatrix} = \begin{pmatrix}1 & 1\\2 & 0\end{pmatrix}$

Definition

a) $A, B \in K^{m\times n}$ heißen äquivalent, falls $S \in GL(m, K)$ und $R \in GL(n, K)$ existieren mit $B = S^{-1}AR$.

b) $A, B \in K^{n\times n}$ heißen ähnlich, falls $S \in GL(n, K)$ existiert mit $B = S^{-1}AS$.

9.3 Der Zusammenhang zwischen linearen Abbildungen, Matrizen und linearen Gleichungssystemen

Satz 9.11 *Sei $A = (a_{ij}) \in K^{n\times n}$ invertierbar. Dann kann A^{-1} folgendermaßen berechnet werden. Bringe das LGS mit n rechten Seiten $A \mid E_n$, d.h. $Ax = E_n$,*

$$\begin{array}{rcccccccc} a_{11}x_1 + \cdots + a_{1n}x_n & = & 1 & bzw. & 0 & \cdots & bzw. & 0 \\ a_{21}x_1 + \cdots + a_{2n}x_n & = & 0 & " & 1 & \cdots & " & 0 \\ & \vdots & & & & & & \\ a_{n1}x_1 + \cdots + a_{nn}x_n & = & 0 & " & 0 & \cdots & " & 1 \end{array}$$

durch Anwendung der Zeilenumformungen des Gauß-Algorithmus (gleichzeitig für alle n rechten Seiten) auf die Form

$$\left.\begin{array}{ccccc|ccc} 1 & 0 & \cdots & 0 & 0 & b_{11} & \cdots & b_{1n} \\ 0 & 1 & 0 & \cdots & \vdots & \vdots & \cdots & \vdots \\ \vdots & 0 & \ddots & \ddots & \vdots & \vdots & \cdots & \vdots \\ \vdots & \cdots & 0 & 1 & 0 & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 & 1 & b_{n1} & \cdots & b_{nn} \end{array}\right. .$$

Dann ist $A^{-1} = (b_{ij})$.

Definition *Sei $A \in K^{m\times n}$.*

a) *Die maximale Anzahl linear unabhängiger Spaltenvektoren von A heißt Spaltenrang von A.*

b) *Die maximale Anzahl linear unabhängiger Zeilenvektoren von A heißt Zeilenrang von A.*

Satz 9.12

a) *Der Zeilenrang von $A \in K^{m\times n}$ ändert sich nicht durch elementare Zeilenumformungen, d.h. Vertauschen von Zeilen, Multiplikation einer Zeile mit $a \in K\backslash\{0\}$ oder Addition eines Vielfachen einer Zeile zu einer anderen Zeile.*

b) *Der Spaltenrang von $A \in K^{m\times n}$ ändert sich nicht durch elementare Spaltenumformungen, d.h. Vertauschen von Spalten, Multiplikation einer Spalte mit $a \in K\backslash\{0\}$ oder Addition eines Vielfachen einer Spalte zu einer anderen Spalte.*

Satz 9.13 *Sei $A \in K^{m\times n}$. Dann gilt:*

$$\textit{Zeilenrang von } A \;=\; \textit{Spaltenrang von } A$$

Beweisstrategie
Strategie 1: Sei der Zeilenrang von $A = r$ und seien die Zeilen $(a_{i_k 1}, \ldots, a_{i_k n})$, $1 \le k \le r$, linear unabhängig. Dann lässt sich der j-te Spaltenvektor, $j \in \{1, \ldots, n\}$, von A darstellen in der Form

$$\begin{pmatrix} a_{1j} \\ \vdots \\ a_{ij} \\ \vdots \\ a_{mj} \end{pmatrix} = a_{i_1 j} \begin{pmatrix} c_{1i_1} \\ \vdots \\ c_{ii_1} \\ \vdots \\ c_{mi_1} \end{pmatrix} + \cdots + a_{i_r j} \begin{pmatrix} c_{1i_r} \\ \vdots \\ c_{ii_r} \\ \vdots \\ c_{mi_r} \end{pmatrix}$$

$\Rightarrow$ Spaltenrang von $A \le$ Zeilenrang von A
Ersetzt man in der obigen Argumentation Zeilen durch Spalten und Spalten durch Zeilen, dann erhält man:
Zeilenrang von $A \le$ Spaltenrang von A
Also folgt die Behauptung.

Strategie 2: Man zeigt die Behauptung zunächst für Matrizen in Zeilenstufenform (für diese Matrizen ist der Beweis einfacher). Dann bringt man beliebige Matrizen durch elementare Zeilen- oder Spaltenumformungen auf Zeilenstufenform und nutzt Satz 9.12 aus.

Definition *Für $A \in K^{m \times n}$ heißt Rang(A) := Spaltenrang von A = Zeilenrang von A der Rang von A.*

Satz 9.14 *Sei $A \in K^{m \times n}$ und L_A wie in Abschnitt 9.2. Dann gilt:*

$$Rang(A) = Rang(L_A).$$

Definition *Für $A \in K^{m \times n}$ heißt $Kern(A) := Kern(L_A)$ der Kern von A.*

Satz 9.15 *Sei $A \in K^{m \times n}$ und $0 \neq b \in K^m$. Dann gilt:*

a) Die Lösungsmenge des homogenen LGS $Ax = 0$ stimmt mit $Kern(A)$ überein.

b) Das inhomogene LGS $Ax = b$ ist genau dann lösbar, wenn $b \in Bild(L_A)$. In diesem Fall ist die Lösungsmenge gegeben durch

$$x_0 + Kern(A) = \{x_0 + y : y \in Kern(A)\},$$

wobei x_0 eine beliebige Lösung von $Ax = b$ ist.

Korollar 9.16 *Sei $A \in K^{m \times n}$ und $b \in K^m$. Dann gilt: Die Lösungsmenge des LGS $Ax = b$ ist die leere Menge oder ein affiner Raum. Ist $b = 0$, dann ist sie ein Untervektorraum von K^n.*

Korollar 9.17 *Sei $A \in K^{m\times n}$. Dann gilt:*

a) Das LGS $Ax = b$ besitzt für jedes $b \in K^m$ mindestens eine Lösung.
$\Leftrightarrow$ *Rang*$(A) = m$

b) Das LGS $Ax = b$ besitzt für jedes $b \in K^m$ höchstens eine Lösung.
$\Leftrightarrow$ *Rang*$(A) = n$

Korollar 9.18 *Sei $A \in K^{n\times n}$. Dann ist äquivalent:*

(i) Das LGS $Ax = b$ besitzt für jedes $b \in K^n$ genau eine Lösung.

(ii) Das homogene LGS $Ax = 0$ besitzt nur die triviale Lösung $x = 0$.

(iii) Rang$(A) = n$

(iv) A ist invertierbar.

(v) L_A ist bijektiv.

10 Skalarprodukträume

10.1 Skalarprodukte, Normen und Metriken

Definition

a) Sei V ein $\mathbb{R}$-Vektorraum. Eine Abbildung $\langle\cdot,\cdot\rangle : V \times V \to \mathbb{R}$ heißt (reelles) Skalarprodukt, wenn für alle $x, y, z \in V$ und alle $\lambda \in \mathbb{R}$ gilt:

$$\begin{array}{ll} (SP1) & \langle x,x\rangle \;\geq\; 0 \quad \wedge \quad (\,\langle x,x\rangle = 0 \;\Leftrightarrow\; x = 0\,), \\ (SP2) & \langle x,y\rangle \;=\; \langle y,x\rangle, \\ (SP3) & \langle \lambda x + y, z\rangle \;=\; \lambda\langle x,z\rangle + \langle y,z\rangle. \end{array}$$

V zusammen mit einem reellen Skalarprodukt heißt (reeller) Skalarproduktraum oder euklidischer Vektorraum.

b) Sei V ein $\mathbb{C}$-Vektorraum. Eine Abbildung $\langle\cdot,\cdot\rangle : V \times V \to \mathbb{C}$ heißt (komplexes) Skalarprodukt, wenn für alle $x, y, z \in V$ und alle $\lambda \in \mathbb{C}$ gilt:

$$\begin{array}{ll} (SP1) & \langle x,x\rangle \;\geq\; 0 \quad \wedge \quad (\,\langle x,x\rangle = 0 \;\Leftrightarrow\; x = 0\,), \\ (SP2) & \langle x,y\rangle \;=\; \overline{\langle y,x\rangle}, \\ (SP3) & \langle \lambda x + y, z\rangle \;=\; \lambda\langle x,z\rangle + \langle y,z\rangle. \end{array}$$

V zusammen mit einem komplexen Skalarprodukt heißt (komplexer) Skalarproduktraum oder unitärer Vektorraum.

Beispiele

Folgende $\mathbb{R}$-Vektorräume bilden mit den zugehörigen Abbildungen (reelle) Skalarprodukträume:

1) $\mathbb{R}^n$, $\langle x,y\rangle_{\mathbb{R}^n} := \sum\limits_{i=1}^{n} x_i y_i$ (Standardskalarprod. auf $\mathbb{R}^n$, euklidisches Skalarprod.),

2) $\mathbb{R}^n$, $\langle x,y\rangle := \sum\limits_{i=1}^{n} \lambda_i x_i y_i$ mit $\lambda_1, ..., \lambda_n > 0$.

Folgende $\mathbb{C}$-Vektorräume bilden mit den zugehörigen Abbildungen (komplexe) Skalarprodukträume:

3) $\mathbb{C}^n$, $\langle x,y\rangle_{\mathbb{C}^n} := \sum\limits_{i=1}^{n} x_i \overline{y_i}$ (Standardskalarprodukt auf $\mathbb{C}^n$),

4) $\mathbb{C}^n$, $\langle x,y\rangle := \sum\limits_{i=1}^{n} \lambda_i x_i \overline{y_i}$ mit $\lambda_1, ..., \lambda_n > 0$.

Satz 10.1 *a) Sei $(V, \langle\cdot,\cdot\rangle)$ ein reeller Skalarproduktraum. Dann gilt für alle $x, y, z \in V$ und alle $\lambda \in \mathbb{R}$:*

$$\langle z, \lambda x + y\rangle \;=\; \lambda\langle z,x\rangle + \langle z,y\rangle.$$

b) Sei $(V, \langle\cdot,\cdot\rangle)$ ein komplexer Skalarproduktraum. Dann gilt für alle $x, y, z \in V$ und alle $\lambda \in \mathbb{C}$:

$$\langle z, \lambda x + y\rangle \;=\; \overline{\lambda}\langle z,x\rangle + \langle z,y\rangle.$$

Sei von nun an $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$.

Satz 10.2 (Cauchy-Schwarzsche Ungleichung) *Sei V ein Skalarproduktraum und $x, y \in V$. Dann ist*

$$|\langle x, y\rangle| \leq \sqrt{\langle x, x\rangle} \cdot \sqrt{\langle y, y\rangle}.$$

Außerdem gilt für $x, y \in V \backslash \{0\}$ die Gleichheit genau dann, wenn es ein $\alpha \in \mathbb{K}$ gibt, so dass $x + \alpha y = 0$, d.h., wenn x und y linear abhängig sind.

Beweis
Für $y = 0$ gilt $|\langle x, 0\rangle| = 0 = \sqrt{\langle x, x\rangle} \cdot \sqrt{\langle 0, 0\rangle}$.
Sei $y \neq 0$ und $\alpha \in \mathbb{K}$ beliebig. Dann gilt

$$0 \leq \langle x + \alpha y, x + \alpha y\rangle = \langle x, x\rangle + \alpha\langle y, x\rangle + \overline{\alpha}\langle x, y\rangle + |\alpha|^2\langle y, y\rangle \qquad (1)$$

Mit $\alpha := -\frac{\langle x,y\rangle}{\langle y,y\rangle}$ ergibt sich nach Multiplikation beider Seiten der Ungleichung mit $\langle y, y\rangle > 0$:

$$\langle x, x\rangle\langle y, y\rangle - |\langle x, y\rangle|^2 - |\langle x, y\rangle|^2 + |\langle x, y\rangle|^2 = \langle x, x\rangle\langle y, y\rangle - |\langle x, y\rangle|^2 \geq 0.$$

Der letzte Teil der Aussage folgt daraus, dass in (1) Gleichheit genau dann gilt, wenn $\langle x + \alpha y, x + \alpha y\rangle = 0$. $\square$

Skalarprodukte können zur Abstandsmessung verwendet werden.

Definition *Sei V ein $\mathbb{K}$-Vektorraum. Eine Abbildung $\|\cdot\| : V \to \mathbb{R}$ heißt Norm, wenn für alle $x, y, z \in V$ und alle $\lambda \in \mathbb{K}$ gilt:*

$$\begin{array}{lrcl} (N1) & \|x\| & \geq & 0 \quad \wedge \quad (\,\|x\| = 0 \Leftrightarrow x = 0\,), \\ (N2) & \|\lambda x\| & = & |\lambda|\,\|x\|, \\ (N3) & \|x + y\| & \leq & \|x\| + \|y\| \quad (\textit{Dreiecksungleichung}). \end{array}$$

V zusammen mit einer Norm heißt normierter Raum.

Satz 10.3 *In jedem Skalarproduktraum V lässt sich durch*

$$\|x\| := \sqrt{\langle x, x\rangle}$$

eine Norm einführen. Man nennt sie die durch das Skalarprodukt $\langle\cdot,\cdot\rangle$ induzierte Norm.

Beweis
Die Eigenschaften $(N1) - (N3)$ einer Norm sind zu zeigen. $(N1)$ folgt aus der Eigenschaft $(SP1)$ des Skalarprodukts. $(N2)$ folgt aus

$$\|\lambda x\| = \sqrt{\langle \lambda x, \lambda x\rangle} = \sqrt{\lambda\overline{\lambda}\langle x, x\rangle} = |\lambda|\sqrt{\langle x, x\rangle}.$$

$(N3)$ ergibt sich folgendermaßen:

$$\begin{aligned} \|x+y\|^2 &= \langle x+y, x+y\rangle \\ &= \langle x,x\rangle + \langle x,y\rangle + \langle y,x\rangle + \langle y,y\rangle \\ &= \|x\|^2 + \|y\|^2 + \langle x,y\rangle + \langle y,x\rangle. \qquad (1) \end{aligned}$$

Außerdem ist

$$\begin{aligned} \langle x,y\rangle + \langle y,x\rangle &= \langle x,y\rangle + \overline{\langle x,y\rangle} \;=\; 2\mathrm{Re}\langle x,y\rangle \\ &\leq 2|\langle x,y\rangle| \\ &\overset{Cauchy-Schwarz}{\leq} 2\sqrt{\langle x,x\rangle}\sqrt{\langle y,y\rangle} \;=\; 2\|x\|\,\|y\| \qquad (2) \end{aligned}$$

(1) und (2) ergeben zusammen

$$\begin{aligned} \|x+y\|^2 &\leq \|x\|^2 + \|y\|^2 + 2\|x\|\,\|y\| \\ &\leq (\|x\| + \|y\|)^2, \end{aligned}$$

woraus die Eigenschaft $(N3)$ folgt. □

Beispiel
Das euklidische Skalarprodukt induziert auf $\mathbb{R}^n$ die euklidische Norm

$$\|x\|_2 := \left(\sum_{i=1}^n x_i^2\right)^{1/2}.$$

Satz 10.4 (Parallelogrammgleichung) *Sei $(V, \langle\cdot,\cdot\rangle)$ ein Skalarproduktraum und $\|\cdot\|$ die von $\langle\cdot,\cdot\rangle$ induzierte Norm. Dann gilt:*

$$\forall x,y \in V: \quad \|x+y\|^2 + \|x-y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right).$$

Beweis
durch analoge Argumentation wie im Beweis von Satz 8.8. □

Satz 10.5 *Genau diejenigen normierten Räume V, in denen die Parallelogrammgleichung gilt, sind Skalarprodukträume. Im reellen Fall lässt sich dann durch*

$$\langle x,y\rangle := \frac{1}{4}\left(||x+y||^2 - ||x-y||^2\right)$$

und im komplexen Fall durch

$$\langle x,y\rangle := \frac{1}{4}\left(||x+y||^2 - ||x-y||^2 + i\cdot(||x+iy||^2 - ||x-iy||^2)\right)$$

ein Skalarprodukt auf V erklären, welches die Norm $||\cdot||$ induziert.

Beispiel
Sei $V = \mathbb{R}^2$, $||x|| := \max_{i=1,2} |x_i|$. Dann ist $(V, \|\cdot\|)$ ein normierter Raum. Allerdings gilt die Parallelogrammgleichung nicht. Denn sei $x = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, $y = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$. Dann ist $\|x\| = \|y\| = 2$, $\|x+y\| = 3$ und $\|x-y\| = 2$.
Also ist $\|x+y\|^2 + \|x-y\|^2 = 3^2 + 2^2 = 13$,
aber $2\,(\|x\|^2 + \|y\|^2) = 2(2^2+2^2) = 16 \neq 13$,
d.h., $\|\cdot\|$ kann nicht von einem Skalarprodukt induziert sein.

Definition *Sei V ein reeller Skalarproduktraum und seien $x, y \in V\backslash\{0\}$. Dann nennt man $\alpha \in [0,\pi]$ mit*

$$\cos(\alpha) := \frac{\langle x, y\rangle}{\|x\|\,\|y\|}$$

den Winkel zwischen x und y.

Definition *Sei V ein Skalarproduktraum. $x, y \in V$ heißen orthogonal zueinander, in Zeichen $x \perp y$, wenn $\langle x, y\rangle = 0$ ist.*

Satz 10.6 (Satz des Pythagoras) *Sei V ein Skalarproduktraum und $x, y \in V$ orthogonal zueinander. Dann gilt*

$$\|x+y\|^2 = \|x\|^2 + \|y\|^2.$$

Beweis
durch analoge Argumentation wie im Beweis von Satz 8.10. □

Definition *(X, d) heißt metrischer Raum, falls X eine nichtleere Menge ist und $d : X \times X \longrightarrow \mathbb{R}$ eine Abbildung mit folgenden Eigenschaften: Für alle $x, y, z \in X$ gilt*

$$\begin{array}{lrcll} (M1) & d(x,y) & \geq & 0 \quad \wedge \quad (\,d(x,y) = 0 \;\Leftrightarrow\; x = y\,), & \\ (M2) & d(y,x) & = & d(x,y) \quad (Symmetrie), & \\ (M3) & d(x,y) & \leq & d(x,z) + d(z,y) \quad (Dreiecksungleichung). & \end{array}$$

d heißt dann eine Metrik (auf X).

Satz 10.7 *Sei $(V, \|\cdot\|)$ ein normierter Raum. Dann ist durch $d(x,y) := \|x-y\|$ eine Metrik auf V definiert.*

Beispiel
Sei $X \neq \emptyset$ und $d(x,y) \;:=\; \begin{cases} 1 & , \quad x \neq y \\ 0 & , \quad x = y \end{cases}$.
Dann ist d eine Metrik, die sogenannte diskrete Metrik, welche allerdings, sollte X ein $\mathbb{K}$-Vektorraum sein, keine Norm induziert.

Definition

a) *Sei V ein K-Vektorraum. Eine Abbildung $b : V \times V \to K$ heißt Bilinearform (auf V), falls sie bezüglich jeder Komponente linear ist, also, falls gilt:*

$$\forall x, y, z \in V \ \forall \lambda \in K : \ b(\lambda x + y, z) = \lambda b(x, z) + b(y, z),$$

$$\forall u, v, w \in V \ \forall \mu \in K : \ b(u, \mu v + w) = \mu b(u, v) + b(u, w).$$

b) *Sei V ein $\mathbb{C}$-Vektorraum. Eine Abbildung $b : V \times V \to \mathbb{C}$ heißt Sesquilinearform (auf V), falls sie linear bezüglich der ersten Komponente und konjugiert linear bzgl. der zweiten Komponente ist, d.h., falls gilt:*

$$\forall x, y, z \in V \ \forall \lambda \in \mathbb{C} : \ b(\lambda x + y, z) = \lambda b(x, z) + b(y, z),$$

$$\forall u, v, w \in V \ \forall \mu \in \mathbb{C} : \ b(u, \mu v + w) = \overline{\mu} b(u, v) + b(u, w).$$

c) *Eine Bilinearform b auf V heißt symmetrisch, falls gilt:*

$$\forall x, y \in V : b(x, y) = b(y, x).$$

d) *Eine Sesquilinearform b auf V heißt hermitesch, falls gilt:*

$$\forall x, y \in V : \ b(x, y) = \overline{b(y, x)}.$$

e) *Sei b eine Bilinearform oder eine Sesquilinearform. Dann heißt die Abbildung $q : x \mapsto b(x, x)$ die zu b gehörende quadratische Form q.*

f) *Eine Bilinearform bzw. eine Sesquilinearform b auf V bzw. die zugehörige quadratische Form q heißt*

$$\left\{\begin{matrix} positiv\ definit \\ positiv\ semidefinit \\ negativ\ definit \\ negativ\ semidefinit \end{matrix}\right\}, falls\ \forall x \in V \backslash \{0\} : \left\{\begin{matrix} q(x) > 0 \\ q(x) \geq 0 \\ q(x) < 0 \\ q(x) \leq 0 \end{matrix}\right\}.$$

Anderenfalls heißen b bzw. q indefinit.

Satz 10.8 a) *Sei V ein $\mathbb{R}$-Vektorraum. Eine Abbildung $f : V \times V \to \mathbb{R}$ ist genau dann ein reelles Skalarprodukt, wenn sie eine positiv definite, symmetrische Bilinearform ist.*

b) *Sei V ein $\mathbb{C}$-Vektorraum. Eine Abbildung $f : V \times V \to \mathbb{C}$ ist genau dann ein komplexes Skalarprodukt, wenn sie eine positiv definite, hermitesche Sesquilinearform ist.*

Definition *Sei* $A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \in K^{m \times n}.$

a) $A^T \in K^{n\times m}$ *mit* $A^T = \begin{pmatrix} a_{11} & \cdots & a_{m1} \\ \vdots & \vdots & \vdots \\ a_{1n} & \cdots & a_{mn} \end{pmatrix}$ *heißt transponierte Matrix von A.*

b) Ist $K = \mathbb{C}$*, dann heißt* $\overline{A} \in \mathbb{C}^{m\times n}$ *mit* $\overline{A} = \begin{pmatrix} \overline{a_{11}} & \cdots & \overline{a_{1n}} \\ \vdots & \vdots & \vdots \\ \overline{a_{m1}} & \cdots & \overline{a_{mn}} \end{pmatrix}$ *komplex konjugierte Matrix zu A.*

c) Ist $K = \mathbb{C}$*, dann heißt* $A^* := \overline{A}^T$ *die adjungierte Matrix zu A. Ist* $K = \mathbb{R}$*, dann heißt* $A^* := A^T$ *die adjungierte Matrix zu A.*

d) Ist $A = A^T$*, dann heißt A symmetrisch.*

e) Ist $A = \overline{A}^T$*, dann heißt A hermitesch.*

f) Ist $A = A^*$*, dann heißt A selbstadjungiert.*

g) Sei $K = \mathbb{R}$ *oder* $= \mathbb{C}$ *und A symmetrisch oder hermitesch. Dann heißt A positiv definit / positiv semidefinit / negativ definit / negativ semidefinit / indefinit, falls die quadratische Form* $q_A : K^n \to K,\ x \mapsto x^T Ax$ *positiv definit / positiv semidefinit / negativ definit / negativ semidefinit / indefinit ist. (Hierbei wird* x^T *als* $1 \times n$*-Matrix aufgefasst.)*

Satz 10.9 *(i)* $(A^T)^T = A$

(ii) $(A + B)^T = A^T + B^T$

(iii) $(\lambda A)^T = \lambda A^T$

(iv) $(AC)^T = C^T A^T$

Satz 10.10 *Sei* $A \in \mathbb{K}^{m\times n}$*. Dann gilt:*

$$\forall x \in \mathbb{K}^n\ \forall y \in \mathbb{K}^m :\ \langle Ax, y\rangle_{\mathbb{K}^m} = \langle x, A^*y\rangle_{\mathbb{K}^n}.$$

Korollar 10.11 *Sei* $A \in \mathbb{K}^{m\times n}$*. Dann gilt:*

$$x \in Kern(L_A) \Leftrightarrow \forall y \in Bild(L_{A^*}) :\ \langle x, y\rangle_{\mathbb{K}^n} = 0.$$

Korollar 10.12 *Sei* $A \in \mathbb{K}^{n\times n}$ *selbstadjungiert. Dann gilt:*

$$\forall x, y \in \mathbb{K}^n :\ \langle Ax, y\rangle_{\mathbb{K}^n} = \langle x, Ay\rangle_{\mathbb{K}^n}.$$

Satz 10.13 *Sei V ein n-dim. K-Vektorraum mit geordneter Basis* $B = \{b_1, \cdots, b_n\}$*.*

a) *Sei $b : V \times V \to K$ bilinear. Dann existiert genau eine Matrix $M = M_b^B \in K^{n\times n}$, so dass gilt: Sind $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in K^n$ die Koordinatendarstellungen von $x, y \in V$ bzgl. B, dann ist $b(x,y) = \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} M_b^B \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$.*

   Insbesondere ist $M_b^B = \begin{pmatrix} b(b_1,b_1) & \cdots & b(b_1,b_n) \\ \vdots & \vdots & \vdots \\ b(b_n,b_1) & \cdots & b(b_n,b_n) \end{pmatrix}$.

b) *Sei $K = \mathbb{C}$ und $\sigma : V \times V \to \mathbb{C}$ sesquilinear. Dann existiert genau eine Matrix $M = M_\sigma^B \in \mathbb{C}^{n\times n}$, so dass gilt: Sind $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{C}^n$ die Koordinatendarstellungen von $x, y \in V$ bzgl. B, dann ist $\sigma(x,y) = \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} M_\sigma^B \begin{pmatrix} \overline{y_1} \\ \vdots \\ \overline{y_n} \end{pmatrix}$.*

   Insbesondere ist $M_\sigma^B = \begin{pmatrix} \sigma(b_1,b_1) & \cdots & \sigma(b_1,b_n) \\ \vdots & \vdots & \vdots \\ \sigma(b_n,b_1) & \cdots & \sigma(b_n,b_n) \end{pmatrix}$.

c) *Sei $M \in K^{n\times n}$. Dann existiert genau eine Bilinearform b und im Falle $K = \mathbb{C}$ zusätzlich genau eine Sesquilinearform σ, so dass $M = M_b^B$ bzw. $M = M_\sigma^B$ gilt.*

Korollar 10.14 *Seien V, B, b und σ wie in Satz 10.13. Sei $B' = \{b'_1, \ldots, b'_n\}$ eine weitere Basis von V und $S = M_{Id_V}^{B,B'}$. Dann gilt:*

$$M_b^{B'} = S^T M_b^B S,$$
$$M_\sigma^{B'} = S^T M_\sigma^B \overline{S}.$$

Korollar 10.15 *Seien V, B, b und σ wie in Satz 10.13. Dann gilt:*

(i) *b ist symmetrisch. $\Leftrightarrow$ M_b^B ist symmetrisch.*

(ii) *σ ist hermitesch. $\Leftrightarrow$ M_σ^B ist hermitesch.*

(iii) *b bzw. σ ist positiv definit / positiv semidefinit / negativ definit / negativ semidefinit / indefinit.*
$\Leftrightarrow$ M_b^B bzw. M_σ^B ist positiv definit / positiv semidefinit / negativ definit / negativ semidefinit / indefinit.

10.2 Orthonormalsysteme und orthogonale Projektionen

Definition *Sei $(V, \langle\cdot,\cdot\rangle)$ ein Skalarproduktraum und $I \neq \emptyset$.*

a) Eine Menge $(v_i)_{i\in I} = \{v_i : i \in I\} \subseteq V$ mit $v_i \neq 0$ für alle $i \in I$ heißt Orthogonalsystem, falls $\langle v_i, v_j\rangle = 0$ für alle $i \neq j$ ist.

b) Ein Orthogonalsystem heißt Orthonormalsystem (ONS), falls $\langle v_i, v_i\rangle = 1$ für alle $i \in I$ ist.

c) Ein ONS heißt Orthonormalbasis (ONB) von V, falls es eine Basis von V ist.

Beispiele

1) Die kanonische Basis $\mathcal{E}_n$ ist ONB von $(\mathbb{R}^n, \langle\cdot,\cdot\rangle_{\mathbb{R}^n})$.

2) Sei $\varphi \in \mathbb{R}$. Dann ist $\left\{\begin{pmatrix}\cos\varphi\\ \sin\varphi\end{pmatrix}, \begin{pmatrix}-\sin\varphi\\ \cos\varphi\end{pmatrix}\right\}$ ONB von $(\mathbb{R}^2, \langle\cdot,\cdot\rangle_{\mathbb{R}^2})$.

Satz 10.16 *Jedes Orthogonalsystem $(v_i)_{i\in I}$ ist linear unabhängig.*

Beweis
Sei $J \subseteq I$, J endlich und $\sum\limits_{j\in J} \lambda_j v_j = 0$. Für alle $k \in J$ folgt:

$$0 = \left\langle \sum_{j\in J} \lambda_j v_j,\, v_k \right\rangle = \sum_{j\in J} \lambda_j \langle v_j, v_k\rangle = \lambda_k \langle v_k, v_k\rangle \Rightarrow \;\lambda_k = 0$$

□

Satz 10.17 *Sei $(V, \langle\cdot,\cdot\rangle)$ ein Skalarproduktraum mit induzierter Norm $\|\cdot\|$, $I = \{i \in \mathbb{N} : i \leq m\}$ für ein $m \in \mathbb{N}$ oder $I = \mathbb{N}$ und $(w_i)_{i\in I} \subseteq V$ linear unabhängig. Dann gibt es ein ONS $(v_i)_{i\in I}$ mit*

$$\forall n \in I: \;\; Span\{v_1, \dots, v_n\} = Span\{w_1, \dots, w_n\}.$$

$(v_i)_{i\in I}$ kann durch das Gram-Schmidtsche Orthogonalisierungsverfahren konstruiert werden:

1. Schritt: Sei $v_1 := \dfrac{1}{\|w_1\|} w_1$.

n-ter Schritt: Sei $\{v_1, \dots, v_{n-1}\}$ bereits konstruiert. Dann sei $v_n := \dfrac{\widetilde{v}_n}{\|\widetilde{v}_n\|}$ mit

$$\widetilde{v}_n := w_n - \sum_{i=1}^{n-1} \langle w_n, v_i\rangle v_i.$$

Insbesondere gilt: Ist V endlich-dimensional, dann besitzt V eine ONB und jedes ONS in V kann zu einer ONB von V ergänzt werden.

Satz 10.18 *Sei $(V, \langle\cdot,\cdot\rangle)$ ein Skalarproduktraum mit induzierter Norm $\|\cdot\|$ und endlicher ONB $\{v_1, \dots, v_n\}$. Dann gilt:*

(i) $\forall x \in V: \; x = \sum\limits_{i=1}^{n} \langle x, v_i\rangle v_i$

(ii) $u = \sum\limits_{i=1}^{n} \lambda_i v_i \wedge v = \sum\limits_{i=1}^{n} \mu_i v_i \;\Rightarrow\; \langle u, v\rangle = \sum\limits_{i=1}^{n} \lambda_i \overline{\mu_i}$

(iii) $\forall x \in V: \|x\|^2 = \sum\limits_{i=1}^{n} |\langle x, v_i\rangle|^2$ *(Parsevalsche Gleichung)*

Beweis
(i): Da $\{v_1, \dots v_n\}$ ist ONB von V ist, gilt:
$\forall x \in V \exists! (\lambda_1, \dots, \lambda_n) \in \mathbb{K} : x = \sum\limits_{i=1}^{n} \lambda_i v_i \wedge \forall k \in \{1, \dots, n\} : \langle x, v_k\rangle = \sum\limits_{i=1}^{n} \lambda_i \langle v_i, v_k\rangle = \lambda_k$
$(ii), (iii)$: durch direktes Nachrechnen □

Satz 10.19 *Es seien $(V, \langle\cdot,\cdot\rangle)$ ein Skalarproduktraum mit induzierter Norm $\|\cdot\|$, U ein endlich-dimensionaler Unterraum von V, $\{u_1, \dots, u_n\}$ eine ONB von U, $x \in V$ und $y \in U$. Dann sind äquivalent:*

(i) $y = \sum\limits_{i=1}^{n} \langle x, u_i\rangle u_i$

(ii) $\exists z \in V : x = y + z \wedge z \perp U$ *(d. h. $\forall w \in U : \langle z, w\rangle = 0$)*

(iii) $\|x - y\| = \min\{\|x - w\| : w \in U\}$

Die Abbildung $P_U : V \to U, \; x \mapsto \sum\limits_{i=1}^{n} \langle x, u_i\rangle u_i$ ist linear und heißt orthogonale Projektion von V auf U.

Beweisstrategie
$(i) \Rightarrow (ii)$: durch direktes Nachrechnen
$(ii) \Rightarrow (iii)$: durch Nachrechnen unter Verwendung des Satzes von Pythagoras
$(iii) \Rightarrow (i)$: Nach Satz 10.17 existiert eine endliche ONB von $\mathrm{Span}\{u_1, \dots, u_n, x\}$, welche $\{u_1, \dots, u_n\}$ enthält. Stelle x, y, w mit Hilfe von Satz 10.18 (i) bzgl. dieser ONB dar und berechne $\min\{\|x - w\| : w \in U\}$ mit Hilfe von Satz 10.18 (iii).

Bemerkung
In der Analysis werden wir sehen, dass die Menge aller stetigen Funktionen von $[0, 2\pi]$ nach $\mathbb{R}$ mit der Abbildung $\langle f, g\rangle := \int_0^{2\pi} f(x)g(x)\,dx$ ein Skalarproduktraum ist und $(v_i)_{i \in \mathbb{N}_0}$ mit

$$v_0 : x \mapsto \frac{1}{\sqrt{2\pi}}, \quad v_{2n-1} : x \mapsto \frac{1}{\sqrt{\pi}} \sin(nx), \quad v_{2n} : x \mapsto \frac{1}{\sqrt{\pi}} \cos(nx)$$

für $n \in \mathbb{N}$, ein ONS in diesem Raum ist.
Dies ist der Anfangspunkt der Theorie der Fourieranalyse.

10.3 Orthogonale und unitäre Abbildungen

Satz 10.20 *Sei $(V, \langle\cdot,\cdot\rangle)$ ein Skalarproduktraum mit induzierter Norm $\|\cdot\|$ und $L : V \to V$ linear. Dann sind äquivalent:*

(i) *Für jedes ONS $\{v_1, \ldots, v_n\} \subseteq V$ ist $\{Lv_1, \ldots, Lv_n\}$ ein ONS.*

(ii) *L ist isometrisch (längenerhaltend), d. h. $\forall x \in V : \|Lx\| = \|x\|$.*

(iii) *$\forall x, y \in V : \langle Lx, Ly\rangle = \langle x, y\rangle$. Insbesondere ist L kongruent (längen- und winkelerhaltend).*

Definition *Sei V wie in Satz 10.20 und $L : V \to V$ eine lineare Abbildung, welche die äquivalenten Eigenschaften (i)–(iii) aus Satz 10.20 erfüllt.*

a) *Ist $(V, \langle\cdot,\cdot\rangle)$ ein reeller Skalarproduktraum, dann heißt L eine orthogonale Abbildung.*

b) *Ist $(V, \langle\cdot,\cdot\rangle)$ ein komplexer Skalarproduktraum, dann heißt L eine unitäre Abbildung.*

Satz 10.21 *Orthogonale und unitäre Abbildungen sind injektiv.*

Satz 10.22 *Sei $(V, \langle\cdot,\cdot\rangle)$ ein n-dimensionaler reeller bzw. komplexer Skalarproduktraum, $B = \{b_1, \ldots, b_n\}$ eine beliebige ONB von V und $L : V \to V$ linear. Dann sind äquivalent:*

(i) *L ist orthogonal bzw. unitär.*

(ii) *$\{Lb_1, \ldots, Lb_n\}$ ist eine ONB von V.*

(iii) *Die Spaltenvektoren von $M_L^{B,B}$ bilden eine ONB von $\mathbb{R}^n$ bzw. $\mathbb{C}^n$.*

(iv) *$M_L^{B,B}$ ist invertierbar mit $\left(M_L^{B,B}\right)^{-1} = \left(M_L^{B,B}\right)^*$.*

(v) *Die Zeilenvektoren von $M_L^{B,B}$ bilden eine ONB von $\mathbb{R}^n$ bzw. $\mathbb{C}^n$.*

Beweisstrategie
Zeige $(i) \Rightarrow (ii) \Rightarrow (iii) \Rightarrow (iv) \Leftrightarrow (v)$ durch Nachrechnen und $(iv) \Rightarrow (i)$ unter Ausnutzung, dass nach Satz 10.10 gilt:

$$\forall x \in K^n : \langle M_L^{B,B} x, M_L^{B,B} x\rangle_{\mathbb{K}^n} = \langle x, \left(M_L^{B,B}\right)^* M_L^{B,B} x\rangle_{\mathbb{K}^n}.$$

Definition

a) *$A \in \mathbb{R}^{n\times n}$ heißt orthogonal, falls A invertierbar ist mit $A^{-1} = A^{\mathsf{T}}$.*

b) *$A \in \mathbb{C}^{n\times n}$ heißt unitär, falls A invertierbar ist mit $A^{-1} = \overline{A}^{\mathsf{T}}$.*

Beispiel

$A = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$ für $\varphi \in \mathbb{R}$ ist orthogonal und L_A eine orthogonale Abbildung auf $\mathbb{R}^2$.

Satz 10.23

(i) $O(n) := \{A \in GL(n,\mathbb{R}) : A^{-1} = A^{\mathsf{T}}\}$ *ist bzgl. der Matrizenmultiplikation eine Gruppe, d. h.* $O(n)$ *ist eine Untergruppe von* $GL(n,\mathbb{R})$ *– die sogenannte orthogonale Gruppe.*

(ii) $U(n) := \{A \in GL(n,\mathbb{C}) : A^{-1} = \overline{A}^{\mathsf{T}}\}$ *ist eine Untergruppe von* $GL(n,\mathbb{C})$ *– die sogenannte unitäre Gruppe.*

11 Determinanten

Definition *Sei K ein Körper, $n \in \mathbb{N}$ und $V = K^n$.*

a) Eine Abbildung $\det : V^n \to K$ *heißt Determinante, wenn sie die folgenden Eigenschaften hat:*

(D1) $\det$ *ist eine Multilinearform, d.h. eine Abbildung von V^n nach K, die linear bzgl. jedes Arguments ist, also für jedes $k \in \{1, \dots, n\}$ gilt:*

$$\begin{aligned}\det(a_1, \dots, a_{k-1}, \alpha\, a_k + b, a_{k+1}, \dots, a_n) = {} & \alpha \det(a_1, \dots, a_{k-1}, a_k, a_{k+1}, \dots, a_n) \\ & + \det(a_1, \dots, a_{k-1}, b, a_{k+1}, \dots, a_n).\end{aligned}$$

(D2) $\det$ *ist alternierend, d.h., für alle $j, k \in \{1, \dots, n\}$ mit $j \neq k$ gilt:*

$$\det(a_1, \dots, a_j, \dots, a_k, \dots, a_n) = -\det(a_1, \dots, a_k, \dots, a_j, \dots, a_n).$$

Vertauschen zweier Argumente ändert also das Vorzeichen.

(D3) $\det$ *ist normiert, d.h.*

$$\det(e_1, \dots, e_n) = 1,$$

wobei $\{e_1, \dots, e_n\}$ die (geordnete) kanonische Basis von K^n ist.

b) Sei $(a_1, \dots, a_n) \in V^n$. Dann heißt $\det(a_1, \dots, a_n)$ *das orientierte (n-dimensionale) Volumen des von den Vektoren $a_1, \dots, a_n$ aufgespannten Spats und* $|\det(a_1, \dots, a_n)|$ *das (n-dimensionale) Volumen des von $a_1, \dots, a_n$ aufgespannten Spats. Im Fall $n = 2$ sagt man auch Flächeninhalt statt 2-dimensionales Volumen und Parallelogramm statt Spat.*

c) Sei $A \in K^{n \times n}$ mit den Zeilenvektoren $a_1, \dots, a_n$. Dann ist $\det(A) := \det(a_1, \dots, a_n)$.

Kurzschreibweise: Statt $\det\left(\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}\right)$ *schreibt man auch* $\begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix}$.

Bemerkung
Mit Hilfe der im Folgenden aufgestellten Resultate kann man zeigen, dass es genau eine Determinante mit den Eigenschaften $(D1)$–$(D3)$ gibt (bei vorgegebenen V, K und n).

Satz 11.1 *Sei $A \in K^{n \times n}$ mit den Zeilenvektoren $a_1, \dots, a_n$. Dann gilt:*

(i) $\exists i \in \{1, \dots, n\} : a_i = 0 \;\Rightarrow\; \det(A) = 0.$

(ii) Sind zwei Zeilenvektoren von A gleich, dann ist $\det(A) = 0$.

(iii) Addition eines Vielfachen einer Zeile von A zu einer anderen Zeile von A ändert den Wert der Determinante nicht.

Beweis
durch direktes Nachrechnen unter Anwendung von $(D1)$ und $(D2)$. $\square$

Satz 11.2 *Sei $A \in K^{n \times n}$ eine obere Dreiecksmatrix, d. h. A ist von der Form*

$$A = \begin{pmatrix} a_{11} & \dots & \dots & a_{1n} \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & a_{nn} \end{pmatrix}.$$

Dann gilt $\det(A) = \prod_{i=1}^{n} a_{ii}$.

Beweis
1. Fall: $\forall i \in \{1, \dots, n\} : a_{ii} \neq 0$.
Dann kann man, wenn nötig, durch Addition von Vielfachen von Zeilen zu anderen Zeilen erreichen, dass aus A eine Diagonalmatrix, d. h. eine Matrix der Form

$$\begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & a_{nn} \end{pmatrix}$$

entsteht. Aufgrund $(D3)$ und $(D1)$ ist

$$\begin{vmatrix} a_{11} & 0 & \dots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & a_{nn} \end{vmatrix} = \prod_{i=1}^{n} a_{ii}$$

und wegen 11.1 (iii) folgt die Behauptung von Satz 11.2 für den 1. Fall.
2. Fall: $\exists i \in \{1, \dots, n\} : a_{ii} = 0$.
Sei $k := \max\{j \in \{1, \dots, n\} : a_{jj} = 0\}$. Dann kann man, wenn nötig, durch Addition von Vielfachen von Zeilen unter der k-ten Zeile zur k-ten Zeile, erreichen, dass die k-te Zeile nur aus Nullen besteht. Aufgrund von 11.1 (iii), (i) folgt dann die Behauptung auch für den 2. Fall. $\square$

Bemerkung
Aufgrund der bisherigen Resultate ist eine Möglichkeit, die Determinante einer beliebigen Matrix zu berechnen, die Folgende. Man formt die Matrix A mit Hilfe des Gauß-Algorithmus zu einer oberen Dreiecksmatrix D um und notiert dabei die Anzahl p der durchgeführten Zeilenvertauschungen. Dann berechnet man mit Satz 11.2 $\det(D)$ und daraus $\det(A) = (-1)^p \det(D)$.

Satz 11.3
Sei $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \in K^{2\times 2}$. *Dann gilt*

$$\det(A) = a_{11}a_{22} - a_{12}a_{21}.$$

Beweis
1. Fall: $a_{11} = 0$. Dann gilt:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \overset{(D2)}{=} - \begin{vmatrix} a_{21} & a_{22} \\ 0 & a_{12} \end{vmatrix} \overset{11.2}{=} -a_{21}a_{12} = a_{11}a_{22} - a_{12}a_{21}.$$

2. Fall: $a_{11} \neq 0$. Dann gilt:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \overset{11.1}{=} \begin{vmatrix} a_{11} & a_{12} \\ 0 & \left(a_{22} - \frac{a_{21}}{a_{11}}a_{12}\right) \end{vmatrix} \overset{11.2}{=} a_{11}\left(a_{22} - \frac{a_{21}}{a_{11}}a_{12}\right) = a_{11}a_{22} - a_{12}a_{21}.$$

□

Satz 11.4 *Sei* $A \in K^{n\times n}$ *Dann ist äquivalent:*

(i) *Die Zeilenvektoren von* A *sind linear unabhängig.*

(ii) $\det(A) \neq 0$.

Beweis
$(i) \Rightarrow (ii)$:
Sind die Zeilenvektoren von A linear unabhängig, dann kann man A mit Hilfe des Gauß-Algorithmus zu einer oberen Dreiecksmatrix D mit $d_{ii} \neq 0$ für alle $i \in \{1, \dots, n\}$ umformen. Aufgrund von Satz 11.2 und der anschließenden Bemerkung folgt $\det(A) \neq 0$.
$(ii) \Rightarrow (i)$: indirekter Beweis.
Sind die Zeilenvektoren von A linear abhängig, dann existiert ein $i \in \{1, \dots, n\}$, so dass gilt: Man kann, wenn nötig, durch Addition von Vielfachen von anderen Zeilenvektoren zum i-ten Zeilenvektor, erreichen, dass der i-te Zeilenvektor der Nullvektor ist. Aufgrund von 11.1 folgt $\det(A) = 0$. □

Satz 11.4 und Korollar 9.18 liefern die folgende Anwendung von Determinanten bei linearen Geichungssystemen:

Korollar 11.5 *Sei* $A \in K^{n\times n}$. *Dann sind äquivalent:*

(i) $\det(A) \neq 0$.

(ii) *Das homogene LGS* $Ax = 0$ *besitzt nur die triviale Lösung* $x = 0$.

(iii) *Das LGS* $Ax = b$ *besitzt für jedes* $b \in K^n$ *genau eine Lösung.*

Satz 11.6 (Determinantenmultiplikationssatz) *Es seien $A, B \in K^{n\times n}$. Dann gilt:*

$$\det(AB) = \det(A)\det(B).$$

Beweis
1. Fall: $\det(A) = 0$.
$\Rightarrow A$ ist nicht invertierbar.
$\Rightarrow L_{AB}$ ist nicht surjektiv.
$\Rightarrow AB$ ist nicht invertierbar.
$\overset{11.4}{\Rightarrow} \det(AB) = 0 = \det(A)\det(B)$.
2. Fall: $\det(A) \neq 0$.
Für $k, l \in \{1, \ldots, n\},\ k \neq l$ und $\alpha \in K$ sei $S(\alpha, k, l) = (s_{ij}) \in K^{n\times n}$ mit

$$s_{ij} = \begin{cases} \alpha, & i = k \wedge j = l \\ 1, & i = j \\ 0, & \text{sonst} \end{cases},$$

also

$$S(\alpha, k, l) = \begin{pmatrix} 1 & 0 & \ldots & \ldots & 0 \\ 0 & 1 & & & \vdots \\ \vdots & & \ddots & & \vdots \\ \vdots & & \alpha & \ddots & 0 \\ 0 & \ldots & \ldots & 0 & 1 \end{pmatrix} \quad (\alpha \text{ in } k\text{-ter Zeile und } l\text{-ter Spalte}).$$

Dann kann A durch Addition des α-fachen der l-ten Zeile zur k-ten Zeile zur Matrix $S(\alpha, k, l)A$ umgeformt werden. Daher und wegen $\det(A) \neq 0$ kann A durch Multiplikation mit einer Matrix S, die das Produkt von endlich vielen Matrizen der Form $S(\alpha, k, l)$ ist, zu einer Diagonalmatrix D mit $d_{ii} \neq 0$ für alle $i \in \{1, \ldots, n\}$ umgeformt werden.
Mit Hilfe der Bemerkung nach 11.2 und vollständiger Induktion kann man berechnen, dass $\det(S(\alpha_m, k_m, l_m)) = 1$ für alle $k_m, l_m \in \{1, \ldots, n\}$ und alle $\alpha_m \in K$, sowie $\det(S) = 1$ ist. Also gilt

$$\det(SA) = \det(D) = \det(A) = \det(S)\det(A).$$

Des Weiteren gilt:

$$\begin{aligned} \det(AB) &= \det(SAB) = \det(DB) = \begin{vmatrix} d_{11}b_{11} & \ldots & d_{11}b_{1n} \\ \vdots & & \vdots \\ d_{nn}b_{n1} & \ldots & d_{nn}b_{nn} \end{vmatrix} \\ &\overset{(D1)}{=} \left(\prod_{i=1}^{n} d_{ii}\right)\det(B) \overset{11.2}{=} \det(D)\det(B) \\ &= \det(A)\det(B). \end{aligned}$$

□

Bemerkung
Im Allgemeinen gilt aber nicht $\det(A+B) = \det(A) + \det(B)$.
Gegenbeispiel: $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix}$

$$\Rightarrow \ \det(A) = \det(B) = 0, \text{ aber } \det(A+B) = \begin{vmatrix} 1 & 1 \\ 0 & 1 \end{vmatrix} = 1.$$

Korollar 11.7

(i) Sei $A \in K^{n\times n}$ invertierbar. Dann gilt:

$$\det(A^{-1}) = (\det(A))^{-1}.$$

(ii) Seien $A, B \in K^{n\times n}$ ähnlich. Dann gilt:

$$\det(A) = \det(B).$$

Aufgrund von Korollar 11.7 (ii) und Satz 9.10 $d)$ können wir definieren:

Definition *Sei V ein endlich-dimensionaler Vektorraum und $L : V \to V$ linear. Dann sei*

$$\det(L) := \det(M_L^{B,B}),$$

wobei B eine beliebige (geordnete) Basis von V ist.

Satz 11.8 *Sei $A \in K^{n\times n}$. Dann gilt*

$$\det(A^\top) = \det(A).$$

Bemerkung
Aufgrund von Satz 11.8 kann man in den Sätzen 11.1 und 11.4 Zeilen durch Spalten ersetzen, sowie in Satz 11.2 obere Dreiecksmatrizen durch untere Dreiecksmatrizen ersetzen und erhält die selben Schlussfolgerungen.

Korollar 11.9 *Sei $A \in \mathbb{C}^{n\times n}$. Dann gilt:*

$$\det(A^*) = \overline{\det(A)}$$

Korollar 11.10 *Sei A eine orthogonale oder unitäre Matrix. Dann ist*

$$|\det(A)| = 1.$$

Satz 11.11 (Entwicklungssatz von Laplace) *Sei $A = (a_{ij}) \in K^{n\times n}$ und $A_{ij} \in K^{(n-1)\times(n-1)}$ die Matrix, die aus A durch Streichen der i-ten Zeile und j-ten Spalte entsteht. Dann gilt für jedes $i \in \{1, \ldots, n\}$:*

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij}) \quad \textit{(Entwicklung nach der i-ten Zeile).}$$

Korollar 11.12 *Seien A, A_{ij} wie in Satz 11.11. Dann gilt für jedes $j \in \{1, \ldots, n\}$:*

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij}) \quad \textit{(Entwicklung nach } j\textit{-ter Spalte).}$$

Beispiel

$$\begin{vmatrix} 2 & 3 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 2 \end{vmatrix} = -\begin{vmatrix} 3 & 1 \\ 1 & 2 \end{vmatrix} + \begin{vmatrix} 2 & 1 \\ 1 & 2 \end{vmatrix} = -5 + 3 = -2 \quad \text{(Entw. nach 2. Zeile)}$$

$$= \begin{vmatrix} 1 & 1 \\ 1 & 1 \end{vmatrix} + 2\begin{vmatrix} 2 & 3 \\ 1 & 1 \end{vmatrix} = 0 - 2 = -2 \quad \text{(Entw. nach 3. Spalte).}$$

Korollar 11.13
Sei $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \in K^{2\times 2}$. Dann gilt:

$$A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}.$$

Zur geometrischen Bedeutung der Determinante:

Satz 11.14 *Sei $L : \mathbb{R}^n \to \mathbb{R}^n$ linear und bijektiv. Dann wird jeder Spat mit n-dimensionalem Volumen V durch L auf einen Spat mit n-dimensionalem Volumen $|detL|\, V$ abgebildet.*

Beweisstrategie
Folgt mit Hilfe des Determinantenmultiplikationssatzes.

Bemerkung
Sei $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \in \mathbb{R}^{2\times 2}$. Aufgrund von Satz 11.3 und der Definition des Vektorproduktes gilt

$$\begin{pmatrix} a_{11} \\ a_{21} \\ 0 \end{pmatrix} \times \begin{pmatrix} a_{12} \\ a_{22} \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \det(A) \end{pmatrix}$$

und somit

$$\left\| \begin{pmatrix} a_{11} \\ a_{21} \\ 0 \end{pmatrix} \times \begin{pmatrix} a_{12} \\ a_{22} \\ 0 \end{pmatrix} \right\|_2 = |\det(A)|$$

Sei $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \in \mathbb{R}^{3\times 3}$. Aufgrund von Korollar 11.12, Satz 11.3 und der Definition des Spatproduktes gilt

$$\left[\begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix}, \begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \end{pmatrix}, \begin{pmatrix} a_{13} \\ a_{23} \\ a_{33} \end{pmatrix}\right] = \det(A),$$

also

$$\left|\left[\begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix}, \begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \end{pmatrix}, \begin{pmatrix} a_{13} \\ a_{23} \\ a_{33} \end{pmatrix}\right]\right| = |\det(A)|$$

Die Definition des Flächeninhalts von Parallelogrammen bzw. des Volumens von Spaten mit Hilfe von Determinanten stimmen also im $\mathbb{R}^2$ bzw. im $\mathbb{R}^3$ mit den entsprechenden Definitionen aus Abschnitt 8.1 überein.

Definition *Sei V ein endlich-dimensionaler $\mathbb{R}$-Vektorraum.*

a) Zwei geordnete Basen B und B' von V heißen gleich orientiert, falls $det\left(M_{Id}^{B',B}\right) > 0$ ist.

b) Ein Isomorphismus $L : V \to V$ heißt orientierungserhaltend, falls $det(L) > 0$ ist.

Satz 11.15

(i) $SL(n,K) := \{A \in GL(n,K) : \ det(A) = 1\}$ ist eine Untergruppe von $GL(n,K)-$ die sogenannte spezielle lineare Gruppe.

(ii) $SO(n) := O(n) \cap SL(n,\mathbb{R})$ ist sowohl eine Untergruppe von $O(n)$ als auch eine Untergruppe von $SL(n,\mathbb{R})-$ die sogenannte spezielle orthogonale Gruppe.

(iii) $SU(n) := U(n) \cap SL(n,\mathbb{C})$ ist sowohl eine Untergruppe von $U(n)$ als auch eine Untergruppe von $SL(n,\mathbb{C})-$die sogenannte spezielle unitäre Gruppe.

Satz 11.16

(i) $SO(2) = \left\{A \in \mathbb{R}^{2\times 2} : \exists \varphi \in \mathbb{R} : \ A = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}\right\},$

(ii) $O(2) \setminus SO(2) = \left\{A \in \mathbb{R}^{2\times 2} : \exists \varphi \in \mathbb{R} : \ A = \begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix}\right\}.$

Bemerkungen

1) Matrizen aus $SL(n,\mathbb{R})$ bzw., genauer gesagt, die durch diese Matrizen dargestellten linearen Abbildungen beschreiben volumen- und orientierungserhaltende lineare Abbildungen auf $\mathbb{R}^n$.

2) Matrizen aus $O(n)$ beschreiben längen- und winkelerhaltende (und somit auch volumenerhaltende) lineare Abbildungen auf $\mathbb{R}^n$.

3) Matrizen aus $SO(n)$ beschreiben orientierungserhaltende und längen- und winkelerhaltende (und somit auch volumenerhaltende) lineare Abbildungen auf $\mathbb{R}^n$.

4) $\begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$ beschreibt eine Drehung im $\mathbb{R}^2$ mit Drehwinkel φ.

5) $\begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix}$ beschreibt eine Achsenspiegelung im $\mathbb{R}^2$, wobei $\frac{1}{2}\varphi$ der Winkel zwischen der x_1-Achse und der Spiegelachse ist.

Bemerkung
Alternativ zu der in diesem Abschnitt vorgestellten Vorgehensweise gibt es auch die folgende Möglichkeit, Determinanten einzuführen. Für $A = (a_{ij}) \in K^{n\times n}$ definiert man

$$\det(A) = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} := \sum_{\sigma\in S_n} sign(\sigma) \prod_{i=1} a_{i\sigma(i)}$$

Dabei ist S_n die Gruppe aller bijektiven Abbildungen auf $\{i \in \mathbb{N} :\ i \leq n\}$ und $sign(\sigma)$ (Signum von σ) ist $\left\{ \begin{array}{c} 1 \\ -1 \end{array} \right\}$, falls sich σ als Verkettung einer $\left\{ \begin{array}{c} \text{geraden} \\ \text{ungeraden} \end{array} \right\}$ Anzahl von Transpositionen (Vertauschung von jeweils zwei Zahlen) darstellen lässt. (Man kann zeigen, dass solche Darstellungen immer möglich sind und $sign(\sigma)$ für jede mögliche Darstellung von σ den gleichen Wert annimmt).
Man kann dann die Resultate aus diesem Abschnitt auch mit Hilfe dieser alternativen Determinantendefinition herleiten sowie die Äquivalenz der beiden Determinantendefinitionen beweisen.

12 Eigenwerte und Eigenvektoren

Sei weiterhin $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$.

Definition

a) Sei V ein $\mathbb{K}$-Vektorraum und $L : V \to V$ linear. $\lambda \in \mathbb{K}$ heißt Eigenwert von L, falls ein $v \in V \setminus \{0\}$ existiert mit

$$Lv = \lambda v\,.$$

Der Vektor v heißt dann Eigenvektor von L zum Eigenwert λ. Die Menge $\sigma(L) := \{\lambda \in \mathbb{K} :\ \lambda \text{ ist Eigenwert von } L\}$ heißt Spektrum von L.

b) Sei $A \in \mathbb{K}^{n\times n}$. $\lambda \in \mathbb{K}$ heißt Eigenwert von A, falls ein $x \in \mathbb{K}^n \setminus \{0\}$ existiert mit

$$Ax = \lambda x\,.$$

Der Vektor x heißt dann Eigenvektor von A zum Eigenwert λ. Die Menge $\sigma(A) := \{\lambda \in \mathbb{K} :\ \lambda \text{ ist Eigenwert von } A\}$ heißt Spektrum von A.

Beispiel

Sei $A = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$. Dann gilt:

$\sigma(A) = \{1, 2\}$,

$x \in \mathbb{K}^2$ ist Eigenvektor von A zum Eigenwert 1

$\Leftrightarrow x \in \left\{ \begin{pmatrix} 0 \\ \alpha \end{pmatrix} :\ \alpha \in \mathbb{K} \setminus \{0\} \right\}$,

$x \in \mathbb{K}^2$ ist Eigenvektor von A zum Eigenwert 2

$\Leftrightarrow x \in \left\{ \begin{pmatrix} \alpha \\ 0 \end{pmatrix} :\ \alpha \in \mathbb{K} \setminus \{0\} \right\}$.

Satz 12.1

a) Sei V ein $\mathbb{K}$-Vektorraum und $L : V \to V$ linear. Dann gilt:
$\lambda \in \mathbb{K}$ ist Eigenwert von L und $v \in V \setminus \{0\}$ ist Eigenvektor von L zum Eigenwert $\lambda \Leftrightarrow v \in Kern(L - \lambda Id) \setminus \{0\}$

b) Sei $A \in \mathbb{K}^{n\times n}$. Dann gilt:
$\lambda \in \mathbb{K}$ ist Eigenwert von A und $x \in \mathbb{K}^n \setminus \{0\}$ ist Eigenvektor von A zum Eigenwert $\lambda \Leftrightarrow x \in Kern(A - \lambda E_n) \setminus \{0\}$

Korollar 12.2 *Seien V, L und A wie in 12.1, $\lambda \in \mathbb{K}$ ein Eigenwert von L und $\mu \in \mathbb{K}$ ein Eigenwert von A. Dann gilt:*

$$Eig_\lambda(L) := \{v \in V : \lambda v\}$$

und

$$Eig_\mu(A) := \{x \in \mathbb{K}^n : Ax = \mu x\}$$

sind Untervektorräume von V bzw. $\mathbb{K}^n$ mit Dimension ≥ 1, genannt der Eigenraum von L zum Eigenwert λ bzw. der Eigenraum von A zum Eigenwert μ.

Bemerkung
Da lineare Abbildungen auf endlich-dimensionalen Vektorräumen durch Matrizen dargestellt werden können, genügt es in diesem Fall, Eigenwerte, Eigenvektoren und Eigenräume von Matrizen zu untersuchen.

Satz 12.3 *Sei $A \in \mathbb{K}^{n \times n}$. Dann gilt:*

a) λ ist Eigenwert von $A \Leftrightarrow det(A - \lambda E_n) = 0$

b) Die Funktion $\chi_A : \mathbb{K} \to \mathbb{K}$, $\lambda \mapsto det(A - \lambda E_n)$ ist ein Polynom n-ten Grades, das sogenannte charakteristische Polynom von A, und es gilt:
λ ist Eigenwert von $A \Leftrightarrow \lambda$ ist Nullstelle von χ_A.

Beweisskizze
$a)$ folgt aus Satz 12.1 und Korollar 11.5.
$b)$ folgt aus $a)$ und den Eigenschaften von Determinanten und Polynomen.

Korollar 12.4 *Sei $A \in \mathbb{K}^{n \times n}$. Dann gilt:*

a) A hat höchstens n verschiedene Eigenwerte.

b) A mindestens einen komplexen Eigenwert. Dieser Eigenwert kann reell sein, muss aber nicht, selbst wenn $A \in \mathbb{R}^{n \times n}$ ist.

c) Ist $A \in \mathbb{R}^{n \times n} \subseteq \mathbb{C}^{n \times n}$ und $\lambda \in \mathbb{C} \setminus \mathbb{R}$ Eigenwert von A, dann ist auch $\overline{\lambda}$ Eigenwert von A.

Beweisskizze
$a)$ folgt aus Satz 12.3 $b)$ und Korollar 5.8.
$b), c)$ folgen aus Satz 12.3 $b)$, Satz 6.9 (Fundamentalsatz derAlgebra) und Satz 6.12.

Definition *Sei $A \in \mathbb{K}^{n \times n}$ und λ Eigenwert von A. Dann heißt:*

- *die Vielfachheit der Nullstelle λ von χ_A die algebraische Vielfachheit von λ, bezeichnet mit $n_a(\lambda)$,*
- *$n_g(\lambda) := dim\, Eig_\lambda(A)$ die geometrische Vielfachheit von λ.*

Beispiele

1) Sei $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, d.h., L_A beschreibt die Achsenspiegelung an der Geraden $x_2 = x_1$ im $\mathbb{R}^2$. Dann gilt:

$$\chi_A(\lambda) = \begin{vmatrix} -\lambda & 1 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 - 1 = (\lambda + 1)(\lambda - 1)$$

Also

$$\chi_A(\lambda) = 0 \Leftrightarrow \lambda = \pm 1,$$

d.h., A besitzt die Eigenwerte $\lambda_1 = 1,\ \lambda_2 = -1$.
Berechnung der Eigenvektoren:

$$(A - \lambda_1 E_2)x = 0$$

$$\Leftrightarrow \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\Leftrightarrow -x_1 + x_2 = 0$$

d.h.

$$Eig_1(A) = \left\{ \begin{pmatrix} \alpha \\ \alpha \end{pmatrix} : \alpha \in \mathbb{K} \right\},$$

wobei $\mathbb{K} = \mathbb{R}$, falls A als Matrix im $\mathbb{R}^{2\times 2}$ betrachtet wird und $\mathbb{K} = \mathbb{C}$, falls A als Matrix im $\mathbb{C}^{2\times 2}$ betrachtet wird.
Die Menge aller Eigenvektoren von A zum Eigenwert 1 ist $Eig_1(A) \setminus \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\}$.
Außerdem gilt: $n_a(1) = 1 = n_g(1)$.

$$(A - \lambda_2 E_2)x = 0$$

$$\Leftrightarrow \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\Leftrightarrow x_1 + x_2 = 0,$$

d.h.

$$Eig_{-1}(A) = \left\{ \begin{pmatrix} \alpha \\ -\alpha \end{pmatrix} : \alpha \in K \right\},$$

$$n_a(-1) = 1 = n_g(-1).$$

2) $B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, d.h., L_B beschreibt die Verkettung von L_A mit der orthogonalen Projektion auf die x_1-Achse im $\mathbb{R}^2$ (zuerst spiegeln, dann projizieren).

$$\chi_B(\lambda) = \begin{vmatrix} -\lambda & 1 \\ 0 & -\lambda \end{vmatrix} = \lambda^2$$

Also ist $\lambda = 0$ einziger Eigenwert und es gilt $n_a(0) = 2$.

$$Eig_0(B) = \left\{ \begin{pmatrix} \alpha \\ 0 \end{pmatrix} : \alpha \in \mathbb{K} \right\}$$

$$\Rightarrow n_g(0) = 1.$$

Satz 12.5 *Sei $A \in \mathbb{R}^{n\times n} \subseteq \mathbb{C}^{n\times n}$ und $x \in \mathbb{C}^n$ ein Eigenvektor von A zum Eigenwert $\lambda \in \mathbb{C}$. Dann ist $\overline{x}$ Eigenvektor von A zum Eigenwert $\overline{\lambda}$.*

Beweis

$$A\bar{x} = \bar{A}\bar{x} = \overline{Ax} = \overline{\lambda x} = \bar{\lambda}\bar{x}.$$

□

Satz 12.6 *Sei $A = (a_{ij}) \in \mathbb{K}^{n\times n}$. Dann gilt:*

$$\chi_A(\lambda) = (-1)^n\lambda^n + (-1)^{n-1}Sp(A)\lambda^{n-1} + \sum_{k=1}^{n-2} c_k\lambda^k + det(A)$$

mit $c_k \in \mathbb{K}$ und $Sp(A) = \sum\limits_{i=1}^{n} a_{ii}$ (die sogenannte Spur von A).

Beweisskizze
Wegen $\chi_A(\lambda) = \det(A - \lambda E_n)$ ist $\chi_A(0) = \det(A)$ und durch vollständige Induktion nach n mit Entwicklung von $\det(A - \lambda E_n)$ nach der 1. Spalte erhält man:

$$\chi_A(\lambda) = \prod_{i=1}^{n}(a_{ii} - \lambda) + q(\lambda)$$

mit $Grad\,(q) \leq n - 2$. Hieraus folgt die Behauptung des Satzes.

Korollar 12.7 *Seien $\lambda_1, \ldots, \lambda_k$ die (paarweise verschiedenen) Eigenwerte von $A \in \mathbb{K}^{n\times n}$. Dann gilt:*

(i) $\prod\limits_{i=1}^{k} \lambda_i^{n_a(\lambda_i)} = det(A)$

(ii) $\sum\limits_{i=1}^{k} n_a(\lambda_i)\,\lambda_i = Sp(A)$

Beweis
Folgt aus 12.6 und dem Satz von Vieta (Korollar 6.11). □

Satz 12.8 *Seien $A,\ B \in \mathbb{K}^{n\times n}$ ähnlich. Dann gilt: $\chi_A = \chi_B$.*

Beweis
Da A, B ähnlich sind, sind auch $A - \lambda E_n$ und $B - \lambda E_n$ ähnlich und wegen Korollar 11.7 (ii) folgt die Behauptung. $\square$

Korollar 12.9 *Seien A, $B \in \mathbb{K}^{n\times n}$ ähnlich. Dann gilt:*

(i) $Sp(A) = Sp(B)$

(ii) *λ ist Eigenwert von A mit $n_a(\lambda) = k \Leftrightarrow \lambda$ ist Eigenwert von B mit $n_a(\lambda) = k$*

(iii) *Sei $S \in GL\,(n, \mathbb{K})$ mit $B = S^{-1}AS$ und x Eigenvektor von A zum Eigenwert λ mit $n_g(\lambda) = m$. Dann ist $S^{-1}x$ Eigenvektor von B zum Eigenwert λ und es ist $n_g(\lambda) = m$ (als Eigenwert von B).*

Beweisskizze
(i) folgt aus 12.8 und 12.6. Alternativ kann man auch direkt nachrechnen, dass $Sp(M\,N) = Sp(N\,M)$ für alle M, $N \in \mathbb{K}^{n\times n}$ ist daher $Sp(S^{-1}AS) = Sp(SS^{-1}A) = Sp(E_nA) = Sp(A)$.
(ii) folgt aus 12.8 und 12.3.
(iii) folgt aus $Ax = \lambda x \Leftrightarrow S^{-1}ASS^{-1}x = S^{-1}(\lambda x) = \lambda S^{-1}x \Leftrightarrow BS^{-1}x = \lambda S^{-1}x$.

Korollar 12.10 *Sei λ_i Eigenwert von $A \in \mathbb{K}^{n\times n}$. Dann gilt $n_g(\lambda_i) \leq n_a(\lambda_i)$.*

Beweis
Sei $m = n_g(\lambda_i)$ und $\{b_1, \ldots, b_m\}$ eine Basis von $Eig_{\lambda_i}(A)$. Ergänze diese Basis zu einer Basis $B = \{b_1, \ldots, b_m, \ldots, b_n\}$ von $\mathbb{K}^n$. Dann gilt: A ist ähnlich zu $M_{L_A}^{B,B}$ und es ist

$$M_{L_A}^{B,B} = \left(\begin{array}{ccc|c} & \lambda_i \; E_m & & F \\ \hline 0 & \cdots & 0 & \\ \vdots & \ddots & \vdots & C \\ 0 & \cdots & 0 & \end{array}\right)$$

mit $F \in \mathbb{K}^{m\times(n-m)}$ und $C \in \mathbb{K}^{(n-m)\times(n-m)}$.

$$\Rightarrow \chi_A(\lambda) = \chi_{M_{L_A}^{B,B}}(\lambda) = \det((\lambda_i - \lambda)E_m) \cdot \det(C - \lambda E_{n-m}) = (\lambda_i - \lambda)^m \det(C - \lambda E_{n-m})$$

$$\Rightarrow n_a(\lambda_i) \geq m = n_g(\lambda_i)$$

$\square$

Satz 12.11 *Sei $A \in \mathbb{K}^{n\times n}$ und $S \in GL(n, \mathbb{K})$ mit den Spaltenvektoren s_j, $j = 1, \ldots, n$. Dann ist äquivalent:*

(i) *$\{s_j : j = 1, \ldots n\}$ ist eine Basis von $\mathbb{K}^n$ aus Eigenvektoren von A.*

(ii) *$D = (d_{ij})$ mit $D = S^{-1}AS$ ist eine Diagonalmatrix.*

Gelten die äquivalenten Aussagen (i) und (ii), dann ist $A_{s_j} = d_{jj}s_j$, $j = 1, \ldots, n$, *und* $\{d_{jj}: \ j = 1, \ldots, n\} = \sigma(A)$.

Beweisskizze
$(i) \Rightarrow (ii)$: durch direktes Nachrechnen.
$(ii) \Rightarrow (i)$: Aus (ii) folgt $A = SDS^{-1}$ und aus 12.9 (iii) erhält man (i).
Die restliche Behauptungen ergeben sich durch direktes Nachrechen und wegen 12.9 (ii).

Korollar 12.12 *Sei* $A \in \mathbb{K}^{n \times n}$. *Dann ist äquivalent:*

(i) $\mathbb{K}^n$ *besitzt eine Basis aus Eigenvektoren von* A.

(ii) A *ist ähnlich zu einer Diagonalmatrix* $D = (d_{ij})$.

Definition $A \in \mathbb{K}^{n \times n}$ *heißt diagonalisierbar, falls die äquivalenten Eigenschaften* $(i), (ii)$ *aus 12.11 erfüllt sind.*

Satz 12.13 *Sei* $A \in \mathbb{K}^{n \times n}$. *Dann ist äquivalent:*

(i) A *ist diagonalisierbar*

(ii) χ_A *zerfällt über* $\mathbb{K}$ *in Linearfaktoren und für jede Nullstelle* λ_i *von* χ_A *gilt:*

$$n_g(\lambda_i) = n_a(\lambda_i).$$

Satz 12.14 *Sei* V *ein* $\mathbb{K}$*-Vektorraum,* $L: V \to V$ *linear und* $\lambda_i, \ldots, \lambda_k$ *paarweise veschiedene Eigenwerte von* L *mit zugehörigen Eigenvektoren* $v_1, \ldots v_k$. *Dann ist* $\{v_1, \ldots v_k\}$ *linear unabhängig. Also sind Eigenvektoren zu verschiedenen Eigenwerten linear unabhängig.*

Beweis
Vollständige Induktion:
$k = 1: \quad \{v_1\}$ ist linear unabhängig, da $v_1 \neq 0$.
$k \to k+1: \quad \sum_{i=1}^{k+1} \alpha_i v_i = 0 = (L - \lambda_{k+1} Id)(\sum_{i=1}^{k+1} \alpha_i v_i) = \sum_{i=1}^{k} \alpha_i(\lambda_i - \lambda_{k+1})v_i$
$\overset{IV}{\Rightarrow} \alpha_1 = \ldots = \alpha_k = 0 \Rightarrow \alpha_{k+1}v_{k+1} = 0 \Rightarrow \alpha_{k+1} = 0,$
da $v_{k+1} \neq 0$. □

Beweis Satz 12.13
$(i) \Rightarrow (ii):$
Ist A diagonalisierbar, dann ist A ähnlich zu einer Diagonalmatrix D. Da χ_D über $\mathbb{K}$ in Linearfaktoren zerfällt und für jede Nullstelle λ_i von χ_D gilt

$$n_g(\lambda_i) = n_a(\lambda_i),$$

folgt wegen 12.8 und 12.9 $(ii), (iii)$ die Aussage (ii).

$(ii) \Rightarrow (i)$:
Seien $\lambda_1, \ldots, \lambda_k$ die paarweise verschiedenen Nullstellen von χ_A. Da χ_A in Linearfaktoren zerfällt, gilt

$$\sum_{i=1}^{k} n_a(\lambda_i) = n.$$

Wegen $n_g(\lambda_i) = n_a(\lambda_i)$ folgt

$$\sum_{i=1}^{k} n_g(\lambda_i) = n.$$

Sei B_i eine Basis von $Eig_{\lambda_i}(A)$. Dann folgt aus Satz 12.14, dass $B := \bigcup_{i=1}^{k} B_i$ linear unabhängig ist. Wegen $dim(Span(B)) = \sum_{i=1}^{k} n_g(\lambda_i) = n$ muss B eine Basis von $\mathbb{K}^n$ sein. $\square$

Beispiele

1) $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ ist diagonalisierbar. Wählt man zum Beispiel $S = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$, dann ist $S^{-1}AS = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

2) $B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ ist nicht diagonalisierbar.

Satz 12.15 *Sei $A \in \mathbb{K}^{n \times n}$ selbstadjungiert. Dann gilt:*

(i) $\sigma(A) \subseteq \mathbb{R}$

(ii) Eigenvektoren zu verschiedenen Eigenwerten von A sind orthogonal zueinander.

Beweis
(i) Sei $\lambda \in \sigma(A)$ und $x \in \mathbb{K}^n$ ein EV von A zum EW λ. Dann gilt:

$$\begin{aligned} \lambda\langle x, x\rangle_{\mathbb{K}^n} &= \langle \lambda x, x\rangle_{\mathbb{K}^n} = \langle Ax, x\rangle_{\mathbb{K}^n} = \langle x, Ax\rangle_{\mathbb{K}^n} \\ &= \langle x, \lambda x\rangle_{\mathbb{K}^n} = \bar{\lambda}\langle x, x\rangle_{\mathbb{K}^n} \end{aligned}$$

Wegen $x \neq 0$ folgt daraus $\lambda = \bar{\lambda}$ und somit $\lambda \in \mathbb{R}$.

(ii) Seien $x, y \in \mathbb{K}^n \setminus \{0\}$, $\lambda, \mu \in \mathbb{K}$, $\lambda \neq \mu$, $Ax = \lambda x$ und $Ay = \mu y$.
Wegen (i) gilt $\lambda, \mu \in \mathbb{R}$ und damit folgt

$$\begin{aligned}
(\lambda - \mu)\langle x, y\rangle_{\mathbb{K}^n} &= \langle \lambda x, y\rangle_{\mathbb{K}^n} - \langle x, \bar{\mu} y\rangle_{\mathbb{K}^n} \\
&= \langle \lambda x, y\rangle_{\mathbb{K}^n} - \langle x, \mu y\rangle_{\mathbb{K}^n} \\
&= \langle Ax, y\rangle_{\mathbb{K}^n} - \langle x, Ay\rangle_{\mathbb{K}^n} = 0 \quad |: (\lambda - \mu) \neq 0 \\
\Rightarrow \langle x, y\rangle_{\mathbb{K}^n} = 0.
\end{aligned}$$

□

Satz 12.16 *Sei $A \in \mathbb{K}^{n\times n}$ selbstadjungiert. Dann existiert eine Orthonormalbasis von $\mathbb{K}^n$ aus Eigenvektoren von A.*

Beweis Vollständige Induktion nach n :
$n = 1$:
$\{1\}$ ist ONB von $\mathbb{K}$ und 1 ist EV von jeder Matrix aus $\mathbb{K}^{1\times 1}$
$n \to n+1$:
Nach 12.4 b) und 12.15 (i) besitzt A mindestens einen reellen EW λ_1. Sei v_1 ein EV zu λ_1 und $W = \{v_1\}^\perp = \{x \in \mathbb{K}^{n+1} : \langle x, v_1\rangle_{\mathbb{K}^{n+1}} = 0\}$
Dann ist W ein n-dimensionaler Unterraum von $\mathbb{K}^{n+1}$ und es gilt:

$$\begin{aligned}
\forall x \in W: \ \langle Ax, v_1\rangle_{\mathbb{K}^{n+1}} &= \langle x, Av_1\rangle_{\mathbb{K}^{n+1}} = \langle x, \lambda_1 v_1\rangle_{\mathbb{K}^{n+1}} \\
&= \lambda_1 \langle x, v_1\rangle_{\mathbb{K}^{n+1}} = 0
\end{aligned}$$

$\Rightarrow\ \forall x \in W: \ Ax \in W$
$\Rightarrow\ \widetilde{L} := L_A\,|_W$ ist eine lineare Abbildung von W nach W.
Sei $B = \{b_1, \dots, b_n\}$ eine ONB von W. Dann ist $M_{\widetilde{L}}^{B,B} \in \mathbb{K}^{n\times n}$ selbstadjungiert und nach Induktionsvoraussetzung existiert eine ONB von $\mathbb{K}^n$ aus EV von $M_{\widetilde{L}}^{B,B}$.
Folglich existiert auch eine ONB $\{\tilde{b}_1, \dots, \tilde{b}_n\}$ von W aus EV von A.
Daher ist $\{\frac{v_1}{\|v_1\|_{\mathbb{K}^{n+1}}}, \tilde{b}_1, \dots, \tilde{b}_n\}$ eine ONB von $\mathbb{K}^{n+1}$ aus EV von A. □

Korollar 12.17 *a) Sei $A \in \mathbb{R}^{n\times n}$ symmetrisch. Dann existiert eine orthogonale Matrix S, so dass $S^T AS$ eine Diagonalmatrix ist.*

b) Sei $A \in \mathbb{C}^{n\times n}$ hermitesch. Dann existiert eine Matrix S, so dass $\bar{S}^T AS$ eine Diagonalmatrix ist, die in $\mathbb{R}^{n\times n}$ enthalten ist.

Bemerkung
Koordinatentransformationen, die symmetrische Matrizen mit Hilfe von orthogonalen Matrizen zu Diagonalmatrizen transformieren, nennt man auch Hauptachsentransformationen.

Korollar 12.18 *Sei $A \in \mathbb{K}^{n\times n}$ selbstadjungiert. Dann gilt :*

$$A \text{ ist } \begin{cases} \textit{positiv definit} \\ \textit{positiv semidefinit} \\ \textit{negativ definit} \\ \textit{negativ semidefinit} \end{cases} \Leftrightarrow \; \sigma(A) \subseteq \begin{cases} \mathbb{R}^+ \\ \mathbb{R}_0^+ \\ \mathbb{R}^- \\ \mathbb{R}_0^- \end{cases}$$

Bemerkung
Sei $A \in \mathbb{R}^{n\times n}$. Dann ist $\frac{1}{2}(A + A^T) \in \mathbb{R}^{n\times n}$ symmetrisch und für die quadratische Form $q_A : \mathbb{R}^n \to \mathbb{R}\,,\; x \longmapsto x^T Ax$ gilt

$$q_A(x) = \langle x, Ax\rangle_{\mathbb{R}^n} = \langle x, \frac{1}{2}(A + A^T)x\rangle_{\mathbb{R}^n}.$$

Ist $\{b_1, \dots b_n\}$ eine ONB von $\mathbb{R}^n$ aus EV von $\frac{1}{2}(A + A^T)$, $\frac{1}{2}(A + A^T)b_i = \lambda_i b_i$ und $x = \sum\limits_{i=1}^{n} \widetilde{x}_i b_i$, dann gilt:

$$q_A(x) = \sum_{i=1}^{n} \lambda_i \widetilde{x}_i^{\,2}.$$

Bemerkung
Es gibt verschiedene Verallgemeinerungen von Satz 12.16. Für Details sei auf die Literatur verwiesen.

Beispiel
Sei $C \in SO(2)$, also

$$C = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$$

mit $\varphi \in \mathbb{R}$, d.h., L_C beschreibt eine Drehung im $\mathbb{R}^2$ um φ.

$$\chi_C(\lambda) = \begin{vmatrix} \cos\varphi - \lambda & -\sin\varphi \\ \sin\varphi & \cos\varphi - \lambda \end{vmatrix} = \lambda^2 - 2(\cos\varphi)\lambda + 1 = 0$$

$$\Leftrightarrow \quad \lambda_{1,2} = \cos\varphi \pm i\sin\varphi$$

1. Fall: $\varphi = 2k\pi,\; k \in \mathbb{Z}$
(Drehung um $2k\pi$ entspricht der identischen Abbildung.)
$\sin\varphi = 0 \,\wedge\, \cos\varphi = 1$
$\Rightarrow\; \lambda = 1$ ist einziger EW mit $n_a(1) = 2$,
$Eig_1(C) = \mathbb{R}^2$ bzw. $\mathbb{C}^2$, $\;n_g(1) = 2$.

2. Fall: $\varphi = (2k + 1)\pi,\; k \in \mathbb{Z}$
(Drehung um $(2k + 1)\pi$ entspricht Punktspiegelung am Ursprung.)
$\sin\varphi = 0 \,\wedge\, \cos\varphi = -1$

$\Rightarrow\ \lambda = -1$ ist einziger EW mit $n_a(-1) = 2$,
$Eig_{-1}(C) = \mathbb{R}^2$ bzw. $\mathbb{C}^2,\ n_g(-1) = 2$.

3.Fall: $\varphi \in \mathbb{R} \setminus \{k\pi : k \in \mathbb{Z}\}$
$\sin\varphi \neq 0$
$\Rightarrow\ \lambda_1 = \cos\varphi + i\sin\varphi \in \mathbb{C} \setminus \mathbb{R},\ n_a(\lambda_1) = 1$,
$Eig_{\lambda_1}(C) = Kern \begin{pmatrix} -i\sin\varphi & -\sin\varphi \\ \sin\varphi & i\sin\varphi \end{pmatrix} = \left\{\alpha \begin{pmatrix} 1 \\ -i \end{pmatrix} : \alpha \in \mathbb{C}\right\},\ n_g(\lambda_1) = 1$,
$\lambda_2 = \overline{\lambda_1} = \cos\varphi - i\sin\varphi,\ n_a(\lambda_2) = 1$
$Eig_{\lambda_2}(C) = \left\{\alpha \begin{pmatrix} 1 \\ i \end{pmatrix} : \alpha \in \mathbb{C}\right\},\ n_g(\lambda_2) = 1.$

Satz 12.19 *Sei $A \in SO(3)$. Dann gilt:*

(i) $1 \in \sigma(A)$

(ii) *Für jede ONB $B = \{b_1, b_2, b_3\}$ von $\mathbb{R}^3$ mit $Ab_1 = b_1$ existiert genau ein $\varphi \in\ ]-\pi, \pi]$ mit $|\varphi| = \arccos(\frac{1}{2}(Sp(A) - 1))$, so dass*
$$M_{L_A}^{B,B} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & -\sin\varphi \\ 0 & \sin\varphi & \cos\varphi \end{pmatrix} \quad \textit{ist.}$$

Bemerkung
Für weitere Resultate bzgl. Eigenwerten und Eigenvektoren von Matrizen aus $SO(n), O(n), SU(n), U(n)$ sei auf die Literatur verwiesen.

Definition *Sei $\lambda \in \mathbb{K}$. Die Matrix*
$$J(k,\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ & & & \ddots & 1 \\ 0 & \dots & & 0 & \lambda \end{pmatrix} \in \mathbb{K}^{k\times k}$$
heißt Jordan-Matrix oder Jordan-Block.

Satz 12.20 *Für eine Jordan-Matrix $J(k,\lambda)$ gelten:*

(i) *λ ist einziger EW von $J(k,\lambda)$.*

(ii) *$n_a(\lambda) = k$ und $n_g(\lambda) = 1$.*
Insbesondere ist $J(k,\lambda)$ für $k \geq 2$ nicht diagonalisierbar.

(iii) *Die Matrix $J(k,\lambda) - \lambda E_k$ ist nilpotent, d.h.*
$\exists m \in \mathbb{N}\ \forall n \in \mathbb{N} : n \geq m\ \Rightarrow\ (J(k,\lambda) - \lambda E_k)^n = 0$,
und es gilt $m = k$.

Satz 12.21 *Sei $k \geq 2$, B eine Basis von $\mathbb{K}^k$ und $A \in \mathbb{K}^{k\times k}$ eine Matrix, die nur einen EW λ besitzt und für den gilt: $n_a(\lambda) = k$ und $n_g(\lambda) = 1$.*
Dann ist äquivalent:

(i) $M_{L_A}^{B,B} = J(k, \lambda)$.

(ii) $B = \{b_1, \ldots, b_k\}$ *mit* $(A - \lambda E_k)b_1 = 0$ *und*
$\forall j \in \{1, 2, \ldots, k-1\} : (A - \lambda E_k)b_{j+1} = b_j$.

Definition *Sei $A \in \mathbb{K}^{n\times n}$ und λ ein EW von A mit $n_a(\lambda) = m$. Dann heißt*

$$H_\lambda(A) := Kern((A - \lambda E_n)^m)$$

der Hauptraum von A zum EW λ und die Elemente aus $H_\lambda(A)$ heißen Hauptvektoren von A zum EW λ.

Satz 12.22 (Jordansche Normalform) *Sei $A \in \mathbb{K}^{n\times n}$ und $\chi_A(\lambda) = \prod_{i=1}^{m} (\lambda_i - \lambda)^{k_i}$ mit $\lambda_i \neq \lambda_j$ für $i \neq j$. Dann existiert eine Basis B von $\mathbb{K}^n$ aus Hauptvektoren von A, so dass*

$$M_{L_A}^{B,B} = \begin{pmatrix} J(k_{11}, \lambda_1) & & & \ldots & & & 0 \\ & \ddots & & & & & \\ & & J(k_{1l_1}, \lambda_1) & & & & \\ & & & \ddots & & & \\ & & & & J(k_{m1}, \lambda_m) & & \\ & & & & & \ddots & \\ 0 & \ldots & & & & & J(k_{ml_m}, \lambda_m) \end{pmatrix}$$

ist, wobei $\sum_{j=1}^{l_i} k_{ij} = k_i = n_a(\lambda_i)$ und $l_i = n_g(\lambda_i)$ ist.
$M_{L_A}^{B,B}$ heißt dann Jordansche Normalform von A.

Beweis
siehe Literatur

Beispiel

Sei $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 4 & 1 \\ 0 & -1 & 2 \end{pmatrix}$. Dann gilt:

$\chi_A(\lambda) = (\lambda - 1)(\lambda - 3)^2 \;\Rightarrow\; \sigma(A) = \{1, 3\},\; n_a(1) = 1\;, n_a(3) = 2$

$Eig_1(A) = Span\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\} \;\Rightarrow\; n_g(1) = 1$

$$Eig_3(A) = Span\left\{ \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \right\} \ \Rightarrow \ n_g(3) = 1$$

$$(A - 3E)x = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \ \Rightarrow \ x = \tfrac{1}{2} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \ \Rightarrow \ H_3(A) = Span\left\{ \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right\}$$

Für $B = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \tfrac{1}{2} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right\}$ ist $M_{L_A}^{B,B} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{pmatrix}$.

Teil III
Eindimensionale Analysis

13 Konvergenz

Definition

a) *Eine reellwertige Folge $(a_n)_{n\in\mathbb{N}}$ heißt konvergent gegen $a \in \mathbb{R}$, falls gilt*

$$\forall \varepsilon > 0 \; \exists n_\varepsilon \in \mathbb{N} \; \forall n > n_\varepsilon : \mid a_n - a \mid < \varepsilon \qquad (K)$$

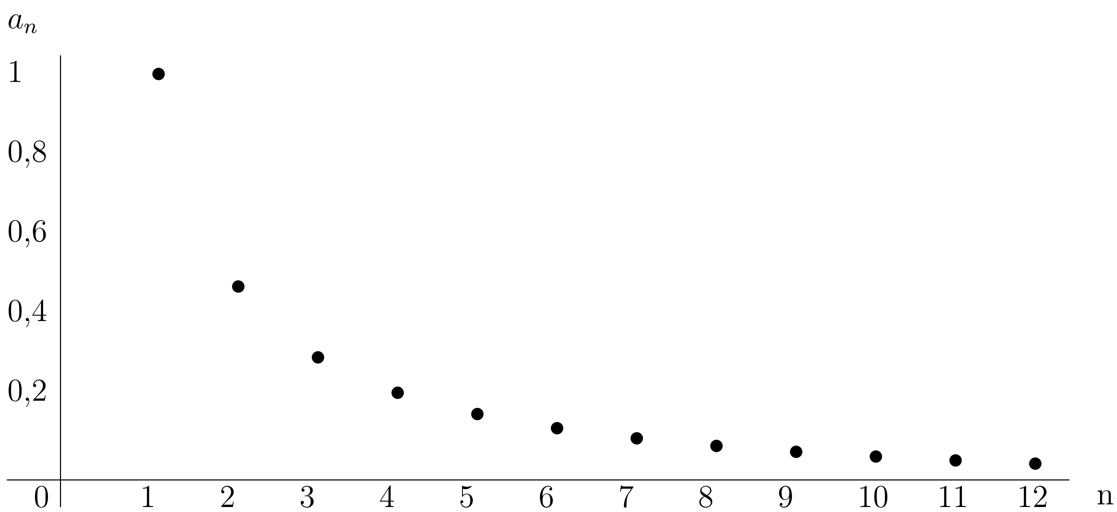
Man sagt dann auch "$(a_n)_{n\in\mathbb{N}}$ konvergiert gegen a" oder "a ist Grenzwert (Limes) von $(a_n)_{n\in\mathbb{N}}$".
Kurzschreibweisen : $a_n \to a$ für $n \to \infty$ oder $\lim\limits_{n\to\infty} a_n = a$.

b) *Eine reellwertige Folge, die gegen 0 konvergiert, heißt Nullfolge.*

c) *Eine reellwertige Folge, die keinen Grenzwert besitzt, heißt divergent.*

Beispiele

1) Sei $(a_n)_{n\in\mathbb{N}}$ mit $a_n = \frac{1}{n}$. Dann ist $(a_n)_{n\in\mathbb{N}}$ eine Nullfolge (folgt aus Korollar 3.17)



2) Sei $N \in \mathbb{N}$ und $(b_n)_{n\in\mathbb{N}}$ mit $b_n = 0$ für alle $n \geq N$. Dann gilt $b_n \to 0$ für $n \to \infty$.

3) Sei $(c_n)_{n\in\mathbb{N}}$ mit $c_n = 1 - \frac{1}{n}$. Dann gilt $\lim\limits_{n\to\infty} c_n = 1$

Satz 13.1 a) *Jede konvergente Folge hat genau einen Grenzwert.*

b) *Jede konvergente Folge $(a_n)_{n\in\mathbb{N}}$ ist beschränkt, d.h.*

$$\exists C > 0 \; \forall n \in \mathbb{N} : \; |a_n| < C$$

c) Verändert man endlich viele Folgeglieder einer konvergenten Folge, dann hat die dadurch entstandene Folge denselben Grenzwert wie die ursprüngliche.

Beweis

$a)$ Angenommen $\exists a, b \in \mathbb{R}: \; a \neq b \; \wedge \; \lim\limits_{n\to\infty} a_n = a \; \wedge \; \lim\limits_{n\to\infty} a_n = b.$

$\Rightarrow \; \exists N, \widetilde{N} \in \mathbb{N} : (\; \forall n > N : \; |a_n - a| < \frac{|a-b|}{2} \;) \wedge (\; \forall n > \widetilde{N} : \; |a_n - b| < \frac{|a-b|}{2} \;)$

$\Rightarrow \; \forall n > max\{N, \widetilde{N}\} : \quad |a-b| = |a - a_n + a_n - b| \leq |a - a_n| + |a_n - b| < |a-b| \; \lightning$

$b)$ Sei $(a_n)_{n\in\mathbb{N}}$ konvergent.

$\Rightarrow \; \exists a \in \mathbb{R} \; \exists n_1 \in \mathbb{N} \; \forall n > n_1 : \; |a_n - a| < 1$

$\Rightarrow \; \forall n \in \mathbb{N} : \; |a_n| \leq max\{|a_1|, \ldots, |a_{n_1}|, |a| + 1\}.$

$c)$ Seien $(a_n)_{n\in\mathbb{N}}, \; (b_n)_{n\in\mathbb{N}}$ mit $\lim\limits_{n\to\infty} a_n = a$ und $\exists N \in \mathbb{N} \; \forall n > N : \; b_n = a_n.$

$\Rightarrow \; \forall \varepsilon > 0 \; \exists n_\varepsilon \in \mathbb{N} \; \forall n > n_\varepsilon : | \; a_n - a \; | < \varepsilon$

$\Rightarrow \; \forall \varepsilon > 0 \; \forall n > max\{N, n_\varepsilon\} : | \; b_n - a \; | < \varepsilon$

$\Rightarrow \; \lim\limits_{n\to\infty} b_n = a.$ $\square$

Beispiele

1) $(a_n)_{n\in\mathbb{N}}$ mit $a_n = n$ ist nicht beschränkt, kann also nicht konvergent sein.

2) Sei $b_n = \begin{cases} 1, & \text{falls } n \in \{2k : 1 \leq k \leq 10\} \\ \frac{1}{n}, & \text{sonst} \end{cases}. \Rightarrow \; \lim\limits_{n\to\infty} b_n = 0$

Satz 13.2 *Seien* $(a_n)_{n\in\mathbb{N}}, \; (b_n)_{n\in\mathbb{N}}$ *konvergente Folgen mit* $\lim\limits_{n\to\infty} a_n = a$ *und* $\lim\limits_{n\to\infty} b_n = b$. *Dann gilt:*

(i) $\forall \alpha, \beta \in \mathbb{R} : \; \lim\limits_{n\to\infty} (\alpha a_n + \beta b_n) = \alpha a + \beta b$

(ii) $\lim\limits_{n\to\infty} (a_n b_n) = ab$

(iii) $(b \neq 0 \; \wedge \; \forall n \in \mathbb{N} : \; b_n \neq 0) \; \Rightarrow \; \lim\limits_{n\to\infty} \dfrac{a_n}{b_n} = \dfrac{a}{b}$

(iv) $\lim\limits_{n\to\infty} |a_n| = |a|$

(v) $\forall m \in \mathbb{N} : \; \lim\limits_{n\to\infty} a_n^m = a^m$

(vi) $(a \geq 0 \; \wedge \; \forall n \in \mathbb{N} : \; a_n \geq 0) \; \Rightarrow \; \forall m \in \mathbb{N} : \; \lim\limits_{n\to\infty} \sqrt[m]{a_n} = \sqrt[m]{a}$

Beweis

$\forall \varepsilon > 0 \; \exists n_\varepsilon, \widetilde{n}_\varepsilon : \; (\forall n > n_\varepsilon : \; |a_n - a| < \varepsilon) \; \wedge \; (\forall n > \widetilde{n}_\varepsilon : \; |b_n - b| < \varepsilon)$

(i) $\forall \alpha, \beta \in \mathbb{R}\ \forall n > max\{n_\varepsilon, \widetilde{n}_\varepsilon\}$:
$|\alpha a_n - \beta b_n - (\alpha a - \beta b)| = |\alpha(a_n - a) - \beta(b_n - b)| \leq |\alpha|\,|a_n - a| + |\beta|\,|b_n - b|$
$\leq (|\alpha| + |\beta|)\,\varepsilon$,
woraus die Behauptung (i) folgt.

(ii) $\forall n > max\{n_\varepsilon, \widetilde{n}_\varepsilon\}$:
$|a_n b_n - ab| = |a_n b_n - a_n b + a_n b - ab| \leq |a_n|\,|b_n - b| + |b|\,|a_n - a|$
$\overset{13.1b)}{\leq} \big(max\{|a_n| :\ n \in \mathbb{N}\} + |b|\big)\,\varepsilon$,
woraus die Behauptung (ii) folgt.

(iii) Sei $b \neq 0\ \wedge\ \forall n \in \mathbb{N} :\ b_n \neq 0$.

$$\Rightarrow\ \exists N \in \mathbb{N}\ \forall n \in \mathbb{N} :\ |b_n| > min\{|b_1|, \ldots, |b_N|, \frac{|b|}{2}\} > 0$$

$$\Rightarrow\ \forall n > \widetilde{n}_\varepsilon : \left|\frac{1}{b_n} - \frac{1}{b}\right| = \left|\frac{b - b_n}{b_n b}\right| \leq \frac{|b - b_n|}{|b| min\left\{|b_1|, \ldots, |b_n|, \frac{|b|}{2}\right\}}$$

$$< \left(\frac{1}{|b| min\{|b_1|, \ldots, |b_n|, \frac{|b|}{2}\}}\right)\varepsilon$$

$$\Rightarrow\ \lim_{n\to\infty} \frac{1}{b_n} = \frac{1}{b}$$

$\overset{(ii)}{\Rightarrow}$ Behauptung (iii)

(iv) Aufgrund der Dreiecksungleichung nach unten (3.13 (v)) gilt :
$\forall n \in \mathbb{N} :\ |\,|a_n| - |a|\,| \leq |a_n - a|$
$\Rightarrow$ Behauptung (iv)

(v) folgt durch vollständige Induktion nach m unter Ausnutzung von (ii).

(vi) Sei $\forall n \in \mathbb{N} :\ a_n \geq 0$.
1. Fall: $a = 0$
$\Rightarrow\ \forall n > n_{\varepsilon^m} : |\ a_n\ | < \varepsilon^m$
$\Rightarrow\ \forall n > n_{\varepsilon^m} :\ \sqrt[m]{a_n} < \varepsilon$
$\Rightarrow\ \lim_{n\to\infty} \sqrt[m]{a_n} = 0 = \sqrt[m]{a}$

2. Fall: $a > 0$
Nach 3.10 (ii) gilt:

$$\forall x, y \in \mathbb{R} :\ (x - y) \sum_{k=0}^{m-1} x^{m-1-k} y^k = x^m - y^m$$

Für $x = \sqrt[m]{a_n}$ und $y = \sqrt[m]{a}$ folgt daher

$$|\sqrt[m]{a_n} - \sqrt[m]{a}| = \frac{|a_n - a|}{\left|\sum\limits_{k=0}^{m-1} a_n^{\frac{m-1-k}{m}} a^{\frac{k}{m}}\right|} \leq a^{-\frac{m-1}{m}} |a_n - a|$$

$$\Rightarrow \ \lim_{n\to\infty} \sqrt[m]{a_n} = \sqrt[m]{a}$$

$\square$

Beispiele

1) $$\frac{2n^4 + n^2 + 5}{3n^5 + 2n} = \underbrace{\frac{1}{n}}_{\to 0} \underbrace{\frac{2 + n^{-2} + 5n^{-4}}{3 + 2n^{-4}}}_{\to \frac{2}{3}} \overset{n\to\infty}{\longrightarrow} 0$$

2) $$\frac{2n^5 + n^2 + 5}{3n^5 + 2n} = \frac{2 + n^{-3} + 5n^{-5}}{3 + 2n^{-4}} \overset{n\to\infty}{\longrightarrow} \frac{2}{3}$$

3) $$\frac{2n^6 + n^2 + 5}{3n^5 + 2n} = n \underbrace{\frac{2 + n^{-2} + 5n^{-4}}{3 + 2n^{-4}}}_{\overset{n\to\infty}{\longrightarrow} \frac{2}{3}}$$

$$\Rightarrow \ \forall N \in \mathbb{N} \ \exists n_N : \ \forall n > n_N : \ \frac{2n^6 + n^2 + 5}{3n^5 + 2n} > N$$

Man schreibt dafür auch:

$$\frac{2n^6 + n^2 + 5}{3n^5 + 2n} \longrightarrow \infty \text{ für } n \to \infty.$$

Wegen 13.1 b) ist die Folge $(a_n)_{n\in\mathbb{N}}$ mit $a_n = \dfrac{2n^6 + n^2 + 5}{3n^5 + 2n}$ divergent.

Analog gilt

$$\frac{-2n^6 + n^2 + 5}{3n^5 + 2n} = n \cdot \underbrace{\frac{-2 + n^{-2} + 5n^{-4}}{3 + 2n^{-4}}}_{\overset{n\to\infty}{\longrightarrow} -\frac{2}{3}}$$

$$\Rightarrow \ \forall N \in \mathbb{N} \ \exists n_N \in \mathbb{N} \ \forall n > n_N : \frac{-2n^6 + n^2 + 5}{3n^5 + 2n} < -N$$

Kurzschreibweise: $\dfrac{-2n^6 + n^2 + 5}{3n^5 + 2n} \longrightarrow -\infty$ für $n \to \infty$.

Wegen 13.1 b) ist auch diese Folge divergent.

4) $$\sqrt{n^2 + n} - n = \frac{(n^2 + n) - n^2}{\sqrt{n^2 + n} + n} = \frac{n}{n(\sqrt{1 + n^{-1}} + 1)} = \frac{1}{\sqrt{1 + n^{-1}} + 1} \overset{n\to\infty}{\longrightarrow} \frac{1}{2}$$

Satz 13.3 *Seien $(a_n)_{n\in\mathbb{N}}$, $(b_n)_{n\in\mathbb{N}}$ konvergente Folgen und*

$$\exists N \in \mathbb{N} \, \forall n > N : a_n \leq b_n \,,$$

dann folgt

$$\lim_{n\to\infty} a_n \leq \lim_{n\to\infty} b_n\,.$$

Beweis
Sei $a = \lim\limits_{n\to\infty} a_n$ und $b = \lim\limits_{n\to\infty} b_n$.

$$\Longrightarrow \exists \widetilde{N}\, \forall n > \widetilde{N} : |a_n - b_n - (a-b)| < \frac{|a-b|}{2}$$

Angenommen, es sei $a > b$.

$$\Longrightarrow \forall n > \widetilde{N} : a_n - b_n > \frac{a-b}{2} > 0,$$

ein Widerspruch zu $a_n \leq b_n$ für hinreichend große n. □

Bemerkung
Ersetzt man in den Voraussetzungen von Satz 13.3 $a_n \leq b_n$ durch $a_n < b_n$, dann kann aber im Allgemeinen nicht

$$\lim_{n\to\infty} a_n < \lim_{n\to\infty} b_n$$

gefolgert werden, sondern auch nur

$$\lim_{n\to\infty} a_n \leq \lim_{n\to\infty} b_n\,,$$

wie das Beispiel $a_n = 0$ und $b_n = \frac{1}{n}$ zeigt.

Korollar 13.4 *(“Sandwich-Satz”, “Dreifolgensatz”, “Prinzip der Polizisten”) Seien $(a_n)_{n\in\mathbb{N}}$, $(b_n)_{n\in\mathbb{N}}$ und $(c_n)_{n\in\mathbb{N}}$ reellwertige Folgen. Gilt*

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

und

$$\exists N \in \mathbb{N}\, \forall n > N : a_n \leq b_n \leq c_n,$$

dann ist $(b_n)_{n\in\mathbb{N}}$ konvergent und

$$\lim_{n\to\infty} b_n = a.$$

Beweis
Folgt aus

$$\begin{aligned}\forall n > N:\ 0 &\leq |b_n - a| \\ &\leq |b_n - a_n| + |a_n - a| \\ &= b_n - a_n + |a_n - a| \\ &\leq c_n - a_n + |a_n - a| \\ &= |c_n - a_n| + |a_n - a| \xrightarrow{n\to\infty} 0\end{aligned}$$

und Satz 13.3. □

Beispiel
Sei $m \in \mathbb{N}_0$, $|q| < 1$ und $a_n = n^m q^n$. Dann gilt $\lim\limits_{n\to\infty} a_n = 0$.

1. Fall: $q = 0$
$\implies \forall n \in \mathbb{N} : a_n = 0 \implies \lim\limits_{n\to\infty} a_n = 0$

2. Fall: $0 < |q| < 1$
Nachweis durch vollständige Induktion nach m.
Dazu benötigen wir die folgende Abschätzung:
Sei $x \in \mathbb{R}$ mit $0 < |x| < 1$. Dann $\exists C > 0$, so dass gilt:

$$\forall n \in \mathbb{N} : |x|^n \leq \frac{C}{n} \quad (*)$$

Nachweis von $(*)$:
$0 < |x| < 1 \implies \dfrac{1}{|x|} > 1$
$\implies \dfrac{1}{|x|} = 1 + h$ mit $h = \dfrac{1}{|x|} - 1 > 0$.
Nach der Bernoulli-Ungleichung (Satz 3.12) folgt:
$\forall n \in \mathbb{N} : \dfrac{1}{|x|^n} = (1+h)^n \geq 1 + nh > nh$
$\implies |x|^n < \dfrac{1}{hn}$.

Nun zur vollständigen Induktion:

$m = 0$:
Sei $0 < |q| < 1 \overset{(*)}{\implies} \exists C > 0\, \forall n \in \mathbb{N} : 0 < |q^n| = |q|^n \leq \dfrac{C}{n} \xrightarrow{n\to\infty} 0$
$\overset{13.4}{\implies} \lim\limits_{n\to\infty} q^n = 0$.

$m \to m+1$:
Sei $0 < |q| < 1 \implies 0 < \sqrt{|q|} < 1$

$$\begin{aligned} \overset{(*)}{\implies} \exists C > 0 \, \forall n \in \mathbb{N}:\ 0 &\leq |n^{m+1} q^n| \\ &= n \left(\sqrt{|q|}\right)^n n^m \left(\sqrt{|q|}\right)^n \\ &\leq C \cdot n^m \left(\sqrt{|q|}\right)^n \overset{n\to\infty}{\longrightarrow} 0 \text{ (nach IV)} \end{aligned}$$

$\overset{13.4}{\implies} \lim\limits_{n\to\infty} n^{m+1} q^n = 0.$

Definition *Eine reellwertige Folge* $(a_n)_{n\in\mathbb{N}}$ *heißt*

a) monoton wachsend, falls gilt $\forall n \in \mathbb{N} : a_{n+1} \geq a_n$,

b) streng monoton wachsend, falls gilt $\forall n \in \mathbb{N} : a_{n+1} > a_n$,

c) monoton fallend, falls gilt $\forall n \in \mathbb{N} : a_{n+1} \leq a_n$,

d) streng monoton fallend, falls gilt $\forall n \in \mathbb{N} : a_{n+1} < a_n$,

e) nach oben beschränkt, falls gilt $\exists C \in \mathbb{R} \, \forall n \in \mathbb{N} : a_n \leq C$,

f) nach unten beschränkt, falls gilt $\exists C \in \mathbb{R} \, \forall n \in \mathbb{N} : a_n \geq C$.

Satz 13.5 *Jede monoton wachsende, nach oben beschränkte reelle Folge* $(a_n)_{n\in\mathbb{N}}$ *ist konvergent und es gilt*

$$\lim_{n\to\infty} a_n = \sup \{a_n : n \in \mathbb{N}\} .$$

Beweis
Nach Voraussetzung ist $\{a_n : n \in \mathbb{N}\}$ nach oben beschränkt. Aufgrund des Supremumsaxioms $\exists a \in \mathbb{R} : a = \sup\{a_n : n \in \mathbb{N}\}$

$$\implies \forall \varepsilon \, \exists n_\varepsilon : a - \varepsilon < a_{n_\varepsilon} < a.$$

Da $(a_n)_{n\in\mathbb{N}}$ monoton wachsend ist, folgt

$$\forall \varepsilon \, \exists n_\varepsilon \forall n \geq n_\varepsilon : |a_n - a| = a - a_n < \varepsilon$$

$\implies a_n \overset{n\to\infty}{\longrightarrow} a.$ $\square$

Korollar 13.6 *Jede monoton fallende, nach unten beschränkte reelle Folge $(a_n)_{n\in\mathbb{N}}$ ist konvergent und es gilt*

$$\lim_{n\to\infty} a_n = \inf\{a_n : n \in \mathbb{N}\}\ .$$

Beispiel (Babylonisches Wurzelziehen, Heron-Verfahren)
Seien $a, x_0 \in \mathbb{R}^+$. Wegen $x > 0 \implies \left(x + \frac{a}{x}\right) > 0$ ist durch

$$\forall n \in \mathbb{N}_0 : x_{n+1} := \frac{1}{2}\left(x_n + \frac{a}{x_n}\right)$$

eine Folge $(x_n)_{n\in\mathbb{N}}$ rekursiv definiert und es gilt $x_n > 0$ für alle $n \in \mathbb{N}_0$. Außerdem gilt:

- $\forall n \in \mathbb{N} : x_n \geq \sqrt{a}\,,$

  denn

$$\begin{aligned}
\forall n \in \mathbb{N}_0 :\ x_{n+1} - \sqrt{a} &= \frac{1}{2}\left(x_n + \frac{a}{x_n}\right) - \sqrt{a} \\
&= \frac{1}{2x_n}\left(x_n^2 + a - 2x_n\sqrt{a}\right) \\
&= \frac{1}{2x_n}\left(x_n - \sqrt{a}\right)^2 \\
&> 0 \qquad (*)
\end{aligned}$$

- $(x_n)_{n\in\mathbb{N}}$ ist monoton fallend,

  denn

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

Daher ist nach Korollar 13.6 die Folge $(x_n)_{n\in\mathbb{N}}$ konvergent mit $\lim\limits_{n\to\infty} x_n \geq \sqrt{a} > 0$. Es gilt sogar

$$\lim_{n\to\infty} x_n = \sqrt{a}\,,$$

denn es ist
$\lim\limits_{n\to\infty} x_n = \lim\limits_{n\to\infty} x_{n+1}$ (folgt direkt aus der Definition eines Grenzwerts)

$$\Longrightarrow \lim_{n\to\infty} x_n = \lim_{n\to\infty} \frac{1}{2}\left(x_n + \frac{a}{x_n}\right) \overset{13.2}{=} \frac{1}{2}\left(\lim_{n\to\infty} x_n\right) + \frac{a}{2}\frac{1}{\underbrace{\lim_{n\to\infty} x_n}_{>0}}$$

$$\Longrightarrow \left(\underbrace{\lim_{n\to\infty} x_n}_{>0}\right)^2 = a$$

$\Longrightarrow \lim\limits_{n\to\infty} x_n = \sqrt{a}.$

Aus $(*)$ folgt außerdem:

$$x_{n+1} - \sqrt{a} \leq \frac{1}{2\sqrt{a}}\left(x_n - \sqrt{a}\right)^2$$

und daher

$$|x_n - \sqrt{a}| = x_n - \sqrt{a} \leq 10^{-k} \Longrightarrow |x_{n+1} - \sqrt{a}| = x_{n+1} - \sqrt{a} \leq \frac{1}{2\sqrt{a}} 10^{-2k},$$

also eine hohe Konvergenzgeschwindigkeit für genügend große n.

Definition *Seien $(a_n)_{n\in\mathbb{N}}$, $(b_n)_{n\in\mathbb{N}}$ reellwertige Folgen. Die Menge $\{[a_n, b_n] : n \in \mathbb{N}\}$ heißt Intervallschachtelung, falls gilt:*

$$(i)\ \forall n \in \mathbb{N} : [a_{n+1}, b_{n+1}] \subseteq [a_n, b_n],$$
$$(ii)\ \lim_{n\to\infty} (b_n - a_n) = 0.$$

Korollar 13.7 *Sei $\{[a_n, b_n] : n \in \mathbb{N}\}$ eine Intervallschachtelung. Dann*

$$\exists! x \in \mathbb{R} : \{x\} = \bigcap_{n\in\mathbb{N}} [a_n, b_n]$$

und es gilt

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

Definition

a) Sei $(x_n)_{n\in\mathbb{N}}$ eine Folge und $k \to n_k$ eine streng monoton wachsende Abbildung von $\mathbb{N}$ nach $\mathbb{N}$. Dann heißt $(x_{n_k})_{k\in\mathbb{N}}$ eine Teilfolge von $(x_n)_{n\in\mathbb{N}}$.

b) Der Grenzwert einer jeden konvergenten Teilfolge von $(x_n)_{n\in\mathbb{N}}$ heißt ein Häufungspunkt von $(x_n)_{n\in\mathbb{N}}$.

*c) Besitzt eine Folge $(x_n)_{n\in\mathbb{N}}$ einen größten Häufungspunkt x, dann heißt er Limes superior von $(x_n)_{n\in\mathbb{N}}$.
Bezeichnung: $x = \limsup\limits_{n\to\infty} x_n = \overline{\lim\limits_{n\to\infty}} x_n$.*

d) *Besitzt eine Folge* $(x_n)_{n\in\mathbb{N}}$ *einen kleinsten Häufungspunkt* x, *dann heißt er Limes inferior von* $(x_n)_{n\in\mathbb{N}}$.
Bezeichnung: $x = \liminf_{n\to\infty} x_n = \varliminf_{n\to\infty} x_n$.

Beispiel
Sei $x_n = (-1)^n$. Dann ist $(x_n)_{n\in\mathbb{N}}$ nicht konvergent, aber es existieren $\liminf_{n\to\infty} x_n = -1$ und $\limsup_{n\to\infty} x_n = 1$.

Satz 13.8 *Sei* x *der Grenzwert einer Folge* $(x_n)_{n\in\mathbb{N}}$. *Dann konvergiert auch jede Teilfolge von* $(x_n)_{n\in\mathbb{N}}$ *gegen* x.

Satz 13.9 (Satz von Bolzano-Weierstraß) *Jede beschränkte Folge* $(x_n)_{n\in\mathbb{N}}$ *in* $\mathbb{R}$ *besitzt eine konvergente Teilfolge.*

Beweis
Sei $(x_n)_{n\in\mathbb{N}}$ eine beschränkte reellwertige Folge.
$\implies \exists a_1, b_1 \in \mathbb{R} : \forall n \in \mathbb{N} : \ x_n \in [a_1, b_1]$.
Sei $\forall m \in \mathbb{N}$:

$$[a_{m+1}, b_{m+1}] := \begin{cases} \left[a_m, \frac{a_m+b_m}{2}\right], & \text{falls eine Teifolge } (x_{n_k})_{k\in\mathbb{N}} \text{ existiert mit} \\ & x_{n_k} \in \left[a_m, \frac{a_m+b_m}{2}\right] \text{ für alle } k \in \mathbb{N} \\ \left[\frac{a_m+b_m}{2}, b_m\right], & \text{sonst} \end{cases}$$

Dann enthält jedes Intervall $[a_m, b_m]$, $m \in \mathbb{N}$, unendlich viele Folgenglieder von $(x_n)_{n\in\mathbb{N}}$. Wähle für jedes $m \in \mathbb{N}$ ein Folgenglied x_{n_m} aus mit $x_{n_m} \in [a_m, b_m]$ und $n_m > n_j$ für alle $j \in \mathbb{N}$ mit $1 \leq j < m$.
Da $\{[a_m, b_m] : m \in \mathbb{N}\}$ eine Intervallschachtelung ist, ist wegen Korollar 13.7 die Folge $(x_{n_m})_{m\in\mathbb{N}}$ konvergent mit $\lim_{m\to\infty} x_{n_m} = \lim_{m\to\infty} a_m = \lim_{m\to\infty} b_m$. □

Definition *Eine reellwertige Folge* $(x_n)_{n\in\mathbb{N}}$ *heißt Cauchy-Folge, falls gilt:*

$$\forall \varepsilon > 0 \, \exists n_\varepsilon \in \mathbb{N} \, \forall m, n > n_\varepsilon : \ |a_m - a_n| < \varepsilon$$

Satz 13.10 a) *Jede konvergente Folge ist eine Cauchy-Folge.*

b) *Jede Cauchy-Folge ist beschränkt.*

c) *Enthält eine Cauchy-Folge* $(x_n)_{n\in\mathbb{N}}$ *eine Teifolge* $(x_{n_k})_{k\in\mathbb{N}}$ *mit* $\lim_{n\to\infty} x_{n_k} = x$, *dann gilt auch* $\lim_{n\to\infty} x_n = x$.

Satz 13.11 *$\mathbb{R}$ ist vollständig, d.h., jede Cauchy-Folge in $\mathbb{R}$ hat eine Grenzwert in $\mathbb{R}$.*

Beweis
Sei $(x_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge in $\mathbb{R}$. Dann ist $(x_n)_{n\in\mathbb{N}}$ nach 13.10 b) beschränkt. Daher enthält sie nach 13.9 eine konvergente Teilfolge, und somit ist $(x_n)_{n\in\mathbb{N}}$ nach 13.10 c) ebenfalls konvergent. $\square$

Bemerkungen

1) Die Vollständigkeit von $\mathbb{R}$ ist letztendlich eine Folgerung von Satz 13.5 und damit eine Konsequenz des Supremumsaxioms, siehe den Beweis von 13.5. Alternativ kann man auch die Vollständigkeit von $\mathbb{R}$ als Axiom postulieren und daraus die Gültigkeit des Supremumsaxioms beweisen.

2) $\mathbb{Q}$ ist nicht vollständig, denn es gibt Cauchy-Folgen in $\mathbb{Q}$, deren jeweilige Grenzwerte in $\mathbb{R} \setminus \mathbb{Q}$ liegen.

3) Die Vollständigkeit von $\mathbb{R}$ ist ein grenzwertfreies Konvergenzkriterium, denn damit kann man für diejenigen Cauchy-Folgen, bei denen man den Grenzwert nicht explizit berechnen kann, immerhin deren Konvergenz nachweisen.

Satz 13.12 *Sei $x \in \mathbb{R}$ und $m = \min\{n \in \mathbb{N} : n > -x\}$.*

(i) Die Folge $\left(\left(1+\frac{x}{n}\right)^n\right)_{n\in\mathbb{N}}$ ist ab dem m-ten Folgenglied monoton wachsend.

(ii) Die Folge $\left(\left(1-\frac{x}{n}\right)^{-n}\right)_{n\in\mathbb{N}}$, definiert für $n \geq m$, ist monoton fallend.

(iii) $\left\{\left[\left(1+\frac{x}{n}\right)^n, \left(1-\frac{x}{n}\right)^{-n}\right] : n \in \mathbb{N} \wedge n > |x|\right\}$ ist eine Intervallschachtelung und somit existiert

$$e^x := \lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n = \lim_{n\to\infty}\left(1-\frac{x}{n}\right)^{-n}.$$

Beweisstrategie
Durch Nachrechnen, z.B. gilt

$$\begin{aligned}\frac{\left(1+\frac{1}{n+1}\right)^{n+1}}{\left(1+\frac{1}{n}\right)^n} &= \left(1-\frac{1}{(n+1)^2}\right)^{n+1}\frac{n+1}{n} \\ &\overset{Bernoulli}{\geq} \left(1-\frac{n+1}{(n+1)^2}\right)\frac{n+1}{n} = 1\end{aligned}$$

Satz 13.13 *(i)* $e^0 = 1$

(ii) $\forall x \in \mathbb{R}:\ e^x > 0$

(iii) $\forall x, y \in \mathbb{R}:\ e^{x+y} = e^x e^y$

(iv) $\forall x \in \mathbb{R}:\ e^{-x} = \dfrac{1}{e^x}$

(v) $\forall x \in \mathbb{Q}:\ e^x = (e^1)^x$, *wobei* $(e^1)^x$ *die* x*-te Potenz von* (e^1) *ist. Daher definiert man* $e := e^1$ *(Eulersche Zahl).*

(vi) $e = 2{,}718281828459045\ldots$

(vii) $\forall x \in \mathbb{R}:\ e^x \geq 1 + x \land (e^x = 1 + x \Leftrightarrow x = 0)$

(viii) $\forall x, y \in \mathbb{R}:\ x < y \implies e^x < e^y$

(ix) $\forall x \in \mathbb{R}:\ |x| < 1 \implies |e^x - 1| \leq \dfrac{|x|}{1 - |x|}$

(x) $\lim\limits_{n\to\infty} e^{\frac{1}{n}} = 1 = \lim\limits_{n\to\infty} e^{-\frac{1}{n}}$.

(xi) $\forall m \in \mathbb{N}_0:\ \lim\limits_{n\to\infty} n^m e^{-n} = 0$.

Beweisstrategie
Durch Nachrechnen.

Satz 13.14 *(i) Die Exponentialfunktion* $\exp : \mathbb{R} \to \mathbb{R}^+,\ x \mapsto e^x$ *ist bijektiv. Die Umkehrfunktion heißt natürlicher Logarithmus* $\ln : \mathbb{R}^+ \to \mathbb{R},\ x \mapsto \ln x$.

(ii) $\ln 1 = 0$

(iii) $\forall x, y \in \mathbb{R}^+ : x < y \implies \ln x < \ln y$

(iv) $\forall a, b \in \mathbb{R}^+ : \ln(ab) = \ln a + \ln b$.

(v) $\forall a, b \in \mathbb{R}^+ : \ln \dfrac{a}{b} = \ln a - \ln b$

(vi) $\forall a \in \mathbb{R}^+ \forall n \in \mathbb{N}:\ \ln a^n = n \cdot \ln a$

(vii) $\forall x \in \mathbb{R}^+:\ \ln x \leq x - 1 \land (\ln x = x - 1 \Leftrightarrow x = 1)$

(viii) $\forall m \in \mathbb{N}:\ \lim\limits_{n\to\infty} n^{-\frac{1}{m}} \ln n = 0$

(ix) $\forall a \in \mathbb{R}^+ \forall x \in \mathbb{Q}:\ a^x = e^{x \cdot \ln a}$

Beweisstrategie
(i): Injektivität folgt aus 13.13 $(viii)$.
Surjektivität: Zeige, dass $\{x \in \mathbb{R} : e^x < y\}$ nichtleer und nach oben beschränkt ist und somit ein Supremum s besitzt.
$\overset{13.13\,(viii)}{\Longrightarrow} \forall n \in \mathbb{N} : e^{s-\frac{1}{n}} < y < e^{s+\frac{1}{n}}$
$\overset{13.13\,(x),13.4}{\Longrightarrow} e^s = y$

$(ii) - (ix)$: durch Nachrechnen unter Ausnutzung von 13.13.

Definition (Reelle Potenzen) *Für $a \in \mathbb{R}^+$ und $x \in \mathbb{R}$ sei $a^x := e^{x \cdot \ln a}$.*

Satz 13.15 *(i)* $\forall a \in \mathbb{R}^+ \forall x, y \in \mathbb{R} : a^{x+y} = a^x a^y$

(ii) $\forall a \in \mathbb{R}^+ \forall x, y \in \mathbb{R} : (a^x)^y = a^{xy}$

(iii) $\forall a, b \in \mathbb{R}^+ \forall x \in \mathbb{R} : (ab)^x = a^x b^x$

(iv) Sei $a \in \mathbb{R}^+ \setminus \{1\}$. Dann ist $f : \mathbb{R} \to \mathbb{R}^+$, $x \mapsto a^x$ streng monoton wachsend und bijektiv. Die Umkehrfunktion heißt Logarithmus zur Basis a, in Zeichen: $\log_a : \mathbb{R}^+ \to \mathbb{R}$, $x \mapsto \log_a x$.

(v) $\forall a \in \mathbb{R}^+ \setminus \{1\} \, \forall x \in \mathbb{R}^+ : \log_a x = \dfrac{\ln x}{\ln a}$.

Beweisstrategie
Durch Nachrechnen.

Definition

(i) Eine komplexwertige Folge $(z_n)_{n\in\mathbb{N}}$ heißt konvergent gegen $z \in \mathbb{C}$, in Zeichen: $z_n \to z$ für $n \to \infty$ oder $\lim\limits_{n\to\infty} z_n = z$, falls $\lim\limits_{n\to\infty} |z_n - z| = 0$ gilt.

(ii) Eine komplexwertige Folge $(z_n)_{n\in\mathbb{N}}$ heißt Cauchy-Folge, falls gilt

$$\forall \varepsilon > 0 \, \exists n_\varepsilon \in \mathbb{N} \, \forall m, n > n_\varepsilon : \, |z_m - z_n| < \varepsilon$$

Satz 13.16 *Seien $z_n = x_n + iy_n$, $n \in \mathbb{N}$ und $z = x + iy$ mit $x, x_n, y, y_n \in \mathbb{R}$. Dann gilt:*

(i) $\lim\limits_{n\to\infty} z_n = z \Longleftrightarrow \lim\limits_{n\to\infty} x_n = x \wedge \lim\limits_{n\to\infty} y_n = y$

(ii) $(z_n)_{n\in\mathbb{N}}$ ist Cauchy-Folge $\Longleftrightarrow$ $(x_n)_{n\in\mathbb{N}}$ und $(y_n)_{n\in\mathbb{N}}$ sind Cauchy-Folgen.

Beweisstrategie
Folgt aus

$$\forall z \in \mathbb{C} : \left\{ \begin{array}{c} |\mathrm{Re}\, z| \\ |\mathrm{Im}\, z| \end{array} \right\} \leq |z| \leq |\mathrm{Re}\, z| + |\mathrm{Im}\, z|.$$

Satz 13.17 *Für komplexwertige Folgen gelten:*

(i) *die Aussagen aus Satz 13.1.*

(ii) *die Aussagen aus Satz 13.2 (i)-(v), wobei* $\alpha, \beta \in \mathbb{C}$ *sein darf.*

(iii) $\lim\limits_{n\to\infty} z_n = z \implies \lim\limits_{n\to\infty} \overline{z_n} = \overline{z}$.

(iv) $\left((\exists C > 0\, \forall n \in \mathbb{N} :\, |z_n| \leq C) \wedge \lim\limits_{n\to\infty} z_n = z\right) \implies |z| \leq C$.

(v) *die Aussage aus 13.8.*

(vi) *die Aussage aus 13.9 für* $\mathbb{C}$ *statt* $\mathbb{R}$.

(vii) *die Aussagen aus 13.10.*

(viii) *die Aussage aus 13.11 für* $\mathbb{C}$ *statt* $\mathbb{R}$.

Beweisstrategie
Analog zum reellen Fall.

Beispiel

$$\lim_{n\to\infty} \frac{\frac{1}{n} + i\left(1 - \dfrac{1}{n^2}\right)}{2 + i\left(1 + \dfrac{1}{\sqrt{n}}\right)} = \frac{i}{2+i} = \frac{1}{5} + \frac{2}{5}i$$

14 Reihen

Definition *Sei $(a_n)_{n\in\mathbb{N}_0}$ eine Folge in $\mathbb{R}$. Dann heißt die Folge $(s_n)_{n\in\mathbb{N}_0}$ mit $s_n = \sum\limits_{k=0}^{n} a_k$ eine reelle (unendliche) Reihe.*

Kurzschreibweise: $\sum\limits_{k=0}^{\infty} a_k$ statt $\Big(\sum\limits_{k=0}^{n} a_k\Big)_{n\in\mathbb{N}_0}$

Die Folgenglieder s_n heißen Partialsummen der Reihe.

Konvergiert die Reihe $\sum\limits_{k=0}^{\infty} a_k$ gegen $a \in \mathbb{R}$, dann schreibt man $\sum\limits_{k=0}^{\infty} a_k = a$ als Abkürzung für $\lim\limits_{n\to\infty} \sum\limits_{k=0}^{n} a_k = a$.

Allgemeiner definiert man für beliebige $k_0 \in \mathbb{Z}$ auch Reihen der Form

$\sum\limits_{k=k_0}^{\infty} a_k := \sum\limits_{m=0}^{\infty} a_{m+k_0}$.

Beispiele

1) Die geometrische Reihe $\sum\limits_{k=0}^{\infty} x^k$ für $x \in \mathbb{R}$:

   1. Fall: $|x| < 1$

   Mit Hilfe der geometrischen Summenformel (Satz 3.10(i)) folgt:

$$\sum_{k=0}^{n} x^k = \frac{1-x^{n+1}}{1-x} = \frac{1}{1-x} - \frac{x}{1-x} \cdot \underbrace{x^n}_{\xrightarrow{n\to\infty} 0} \xrightarrow{n\to\infty} \frac{1}{1-x}$$

   $\Rightarrow$ Für $|x| < 1$ ist $\sum\limits_{k=0}^{\infty} x^k$ konvergent und es gilt

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

   2. Fall: $x = 1$

$$\sum_{k=0}^{n} x^k = n+1 \xrightarrow{n\to\infty} \infty$$

   3. Fall: $x > 1$

$$\sum_{k=0}^{n} x^k = \frac{1-x^{n+1}}{1-x} \xrightarrow{n\to\infty} \infty$$

   4. Fall: $x = -1$

$$\sum_{k=0}^{2m} x^k = 1$$

$$\sum_{k=0}^{2m+1} x^k = 0$$

5. Fall: $x < -1$

$$\sum_{k=0}^{2m} x^k = \frac{1 - x^{2m+1}}{1 - x} \xrightarrow{m\to\infty} \infty$$

$$\sum_{k=0}^{2m+1} x^k = \frac{1 - x^{2m+2}}{1 - x} \xrightarrow{m\to\infty} -\infty$$

$\Rightarrow$ Für $|x| \geq 1$ ist $\sum\limits_{k=0}^{\infty} x^k$ divergent.

Folgerung:

$$\forall m \in \mathbb{N}\, \forall\, |x| < 1 : \sum_{k=m}^{\infty} x^k = x^m \sum_{k=0}^{\infty} x^k = \frac{x^m}{1-x}\,.$$

2) Die harmonische Reihe $\sum\limits_{k=1}^{\infty} \frac{1}{k}$ ist divergent, denn es gilt

$$\begin{aligned} S_n &= 1 + \frac{1}{2} + \underbrace{\left(\frac{1}{3} + \frac{1}{4}\right)}_{>2\cdot\frac{1}{4}=\frac{1}{2}=2\cdot\frac{1}{2^2}} + \underbrace{\left(\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}\right)}_{>4\cdot\frac{1}{8}=\frac{1}{2}=2^2\cdot\frac{1}{2^3}} + ... + \underbrace{\left(\frac{1}{2^{n-1}+1} + ... + \frac{1}{2^n}\right)}_{>2^{n-1}\cdot\frac{1}{2^n}=\frac{1}{2}} \\ &> 1 + \frac{n}{2} \xrightarrow{n\to\infty} \infty. \end{aligned}$$

3) $\sum\limits_{k=1}^{\infty} \frac{1}{k(k+1)} = 1$, denn es gilt

$$\begin{aligned} \sum_{k=1}^{n} \frac{1}{k(k+1)} &= \sum_{k=1}^{n} \left(\frac{1}{k} - \frac{1}{k+1}\right) \\ &= \left(1 - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + ... + \left(\frac{1}{n-1} - \frac{1}{n}\right) + \left(\frac{1}{n} - \frac{1}{n+1}\right) \\ &= 1 - \frac{1}{n+1} \xrightarrow{n\to\infty} 1. \end{aligned}$$

4) Jede reelle Folge $(x_n)_{n\in\mathbb{N}}$ kann als reelle Reihe dargestellt werden, denn es gilt $x_n = \sum\limits_{k=1}^{n} a_k$ mit $a_1 = x_1$ und $a_k = x_k - x_{k-1}$ für alle $k > 1$.

Satz 14.1 *Seien* $\sum\limits_{k=0}^{\infty} a_k$*,* $\sum\limits_{k=0}^{\infty} b_k$ *konvergente Reihen mit* $\sum\limits_{k=0}^{\infty} a_k = a$ *und* $\sum\limits_{k=0}^{\infty} b_k = b$*. Dann gilt:*

$$\forall \alpha, \beta \in \mathbb{R} : \sum_{k=0}^{\infty} \alpha a_k + \beta b_k = \alpha a + \beta b$$

Beweis
Folgt aus Satz 13.2. □

Beispiel

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

Satz 14.2 *Sind* $(a_k)_{k\in\mathbb{N}_0}$ *und* $(b_k)_{k\in\mathbb{N}_0}$ *Folgen, die sich nur durch endlich viele Folgenglieder unterscheiden, dann ist* $\sum\limits_{k=0}^{\infty} a_k$ *genau dann konvergent wenn* $\sum\limits_{k=0}^{\infty} b_k$ *konvergiert. Die Grenzwerte der beiden Reihen können allerdings verschieden sein.*

Beispiel

Sei $x \in \mathbb{R}$ und $a_k = \begin{cases} (10x)^k, & k \leq 100, \\ (2x)^k, & \text{sonst.} \end{cases}$

Dann konvergiert $\sum\limits_{k=0}^{\infty} a_k$ genau dann, wenn $|x| < \frac{1}{2}$ ist.

Satz 14.3 (Leibniz-Kriterium) *Sei* $(a_k)_{k\in\mathbb{N}_0}$ *eine monoton fallende Nullfolge mit* $a_k \in \mathbb{R}^+$ *für alle* $k \in \mathbb{N}_0$*. Dann konvergiert die alternierende Reihe* $\sum\limits_{k=0}^{\infty} (-1)^k a_k$ *und für deren Grenzwert a gilt:*

$$\forall n \in \mathbb{N}_0 : \sum_{k=0}^{2n+1} (-1)^k a_k \leq a \leq \sum_{k=0}^{2n} (-1)^k a_k \tag{1}$$

und

$$\forall n \in \mathbb{N}_0 : |a - \sum_{k=0}^{n} (-1)^k a_k| \leq a_{n+1}. \tag{2}$$

Beweis
Betrachte die Partialsummen $s_n = \sum\limits_{k=0}^{n} (-1)^k a_k$.
Aufgrund der vorausgesetzten Eigenschaften von $(a_k)_{k\in\mathbb{N}_0}$ folgt:

- $\forall m \in \mathbb{N} : s_{2m+1} = s_{2m-1} + a_{2m} - a_{2m+1} \geq s_{2m-1}$
- $\forall m \in \mathbb{N} : s_{2m} = s_{2m-2} - a_{2m-1} + a_{2m} \leq s_{2m-2}$

- $\forall m \in \mathbb{N} : s_{2m+1} - s_{2m} = a_{2m+1} \xrightarrow{m\to\infty} 0$

Somit ist $\{[s_{2m+1}, s_{2m}] : m \in \mathbb{N}_0\}$ eine Intervallschachtelung.
Nach Korollar 13.7 existiert ein $a \in \mathbb{R}$ mit

$a = \lim\limits_{m\to\infty} s_{2m+1} = \lim\limits_{m\to\infty} s_{2m}.$

$\Rightarrow a = \lim\limits_{n\to\infty} s_n = \sum\limits_{k=0}^{\infty} (-1)^k a_k$

Mit Hilfe von Satz 13.5 und Korollar 13.6 folgt außerdem

- $\forall m \in \mathbb{N} : s_{2m-1} \leq a \leq s_{2m}$

  $\Rightarrow |a - s_{2m-1}| = a - s_{2m-1} \leq s_{2m} - s_{2m-1} = a_{2m}$

- $\forall m \in \mathbb{N}_0 : s_{2m+1} \leq a \leq s_{2m}$

  $\Rightarrow |a - s_{2m}| = s_{2m} - a \leq s_{2m} - s_{2m+1} = a_{2m+1}$

und daher $(1), (2)$. □

Beispiel
Die Leibniz-Reihe $\displaystyle\sum_{n=1}^{\infty} (-1)^{n-1} \frac{1}{n} = \sum_{k=0}^{\infty} (-1)^k \frac{1}{k+1}$ konvergiert nach dem Leibniz-Kriterium.
Ihren Grenzwert ($\ln 2$) können wir mit den bisher vorgestellten Mitteln noch nicht berechnen.

Bemerkung
Ist $(a_k)_{k\in\mathbb{N}_0}$ mit $a_k \in \mathbb{R}^+$ für alle $k \in \mathbb{N}_0$ eine Nullfolge, die nicht monoton fallend ist, dann muss $\sum\limits_{k=0}^{\infty} (-1)^k a_k$ nicht notwendigerweise konvergieren, wie das Beispiel

$a_k = \begin{cases} \frac{1}{k}, & k \in \{2m : m \in \mathbb{N}\}, \\ (\frac{1}{2})^k, & \text{sonst}, \end{cases}$

zeigt.

Satz 14.4 (Cauchy-Kriterium)

$$\sum_{k=0}^{\infty} a_k \textit{ ist konvergent.} \quad \Leftrightarrow \quad \forall \varepsilon > 0 \exists n_\varepsilon \forall m > n > n_\varepsilon : |\sum_{k=n+1}^{m} a_k| < \varepsilon$$

Beweis
Folgt direkt aus der Vollständigkeit von $\mathbb{R}$. □

Korollar 14.5

$$\sum_{k=0}^{\infty} a_k \textit{ ist konvergent.} \quad \Rightarrow \quad (a_k)_{k\in\mathbb{N}_0} \textit{ ist eine Nullfolge.}$$

Beispiel
Sei $x \in \mathbb{R}$ und $a_k = x^k$. Da $(a_k)_{k\in\mathbb{N}_0}$ für $|x| \geq 1$ keine Nullfolge ist, erhält man mit Korollar 14.5 die Divergenz der geometrischen Reihe für $|x| \geq 1$ einfacher mit Hilfe der zu Beginn dieses Kapitels diskutierten Fallunterscheidung.

Bemerkung
Die Umkehrung von Korollar 14.5 gilt aber im Allgemeinen nicht, d.h., ist $(a_k)_{k\in\mathbb{N}_0}$ eine Nullfolge, dann muss $\sum\limits_{k=0}^{\infty} a_k$ nicht konvergent sein, wie das Beispiel $a_k = \frac{1}{k+1}$ zeigt. Dass $(a_k)_{k\in\mathbb{N}_0}$ eine Nullfolge ist, ist also eine notwendige Bedingung, aber keine hinreichende Bedingung für die Konvergenz von $\sum\limits_{k=0}^{\infty} a_k$.

Definition *Eine Reihe $\sum\limits_{k=0}^{\infty} a_k$ heißt absolut konvergent, falls die Reihe $\sum\limits_{k=0}^{\infty} |a_k|$ konvergiert.*

Satz 14.6 *Jede absolut konvergente Reihe ist konvergent.*

Beweis
Folgt aus $|\sum\limits_{k=n}^{m} a_k| \leq \sum\limits_{k=n}^{m} |a_k|$ und Satz 14.4. □

Bemerkung
Die Umkehrung von Satz 14.6 gilt aber im Allgemeinen nicht, denn die Leibniz-Reihe ist konvergent, aber nicht absolut konvergent.

Satz 14.7 (Monotoniekriterium)

$$\sum_{k=0}^{\infty} a_k \textit{ ist absolut konvergent.} \quad \Leftrightarrow \quad \left(\sum_{k=0}^{n} |a_k|\right)_{n\in\mathbb{N}_0} \textit{ist eine beschränkte Folge.}$$

Beweis
"$\Leftarrow$": mit Hilfe von Satz 13.5, da $\left(\sum\limits_{k=0}^{n} |a_k|\right)_{n\in\mathbb{N}_0}$ monoton wachsend ist.

"$\Rightarrow$": folgt mit Hilfe von Satz 13.1 b). □

Beispiel
$\displaystyle\sum_{k=1}^{\infty} \frac{1}{k^2}$ ist absolut konvergent und somit konvergent, denn

$$\forall k \geq 2 : \frac{1}{k^2} < \frac{1}{(k-1)k}$$

$$\Rightarrow \forall n \in \mathbb{N} : \sum_{k=1}^{n} \left|\frac{1}{k^2}\right| = \sum_{k=1}^{n} \frac{1}{k^2} < 1 + \sum_{k=1}^{n-1} \frac{1}{k(k+1)} = 2 - \frac{1}{n} < 2$$

Den Grenzwert von $\sum\limits_{k=1}^{\infty} \frac{1}{k^2}$, nämlich $\frac{\pi^2}{6}$, können wir mit den bisher vorgestellten Mitteln noch nicht berechnen.

Satz 14.8 (Majorantenkriterium) *Ist $N \in \mathbb{N}_0$, $|a_k| \leq b_k$ für alle $k \geq N$ und $\sum\limits_{k=0}^{\infty} b_k$ konvergent, dann konvergiert $\sum\limits_{k=0}^{\infty} a_k$ absolut. Die Reihe $\sum\limits_{k=0}^{\infty} b_k$ heißt dann konvergente Majorante für $\sum\limits_{k=0}^{\infty} a_k$.*

Beweis
Seien die Annahmen aus Satz 14.8 an a_k und b_k erfüllt. Dann gilt nach Satz 14.4

$\forall \varepsilon > 0\ \exists n_\varepsilon\ \forall m > n > n_\varepsilon$: $|\sum\limits_{k=n+1}^{m} b_k| < \varepsilon$

$\overset{b_k \geq 0}{\Rightarrow}\ \forall \varepsilon \exists n_\varepsilon \forall m > n > n_\varepsilon : \sum\limits_{k=n+1}^{m} b_k < \varepsilon$

$\Rightarrow \forall \varepsilon \exists n_\varepsilon \forall m > n > max\{N, n_\varepsilon\}$:

$|\sum\limits_{k=n+1}^{m} a_k| \leq \sum\limits_{k=n+1}^{m} |a_k| \leq \sum\limits_{k=n+1}^{m} b_k < \varepsilon$

$\overset{14.4}{\Rightarrow} \sum\limits_{k=0}^{\infty} a_k$ konvergiert absolut. □

Korollar 14.9 (Minorantenkriterium) *Ist $N \in \mathbb{N}_0$, $0 \leq b_k \leq a_k$ für alle $k \geq N$ und $\sum\limits_{k=0}^{\infty} b_k$ divergent, dann divergiert $\sum\limits_{k=0}^{\infty} a_k$. Die Reihe $\sum\limits_{k=0}^{\infty} b_k$ heißt dann divergente Minorante für $\sum\limits_{k=0}^{\infty} a_k$.*

Beispiele

1) $\sum\limits_{k=0}^{\infty} \frac{(sin(k))^2}{5^k} \cdot 4^{k+1}$ konvergiert absolut.

2) $\sum\limits_{k=1}^{\infty} \frac{1}{k^\alpha}$ divergiert für alle $\alpha \leq 1$ und konvergiert absolut für alle $\alpha \geq 2$.

   Später werden wir sehen, dass diese Reihe sogar für alle $\alpha > 1$ absolut konvergiert.

Korollar 14.10 *Ist $|a_k| < b_k$ für alle $k \in \mathbb{N}_0$ und $\sum\limits_{k=0}^{\infty} b_k$ konvergent, dann gilt:*

$$|\sum_{k=0}^{\infty} a_k| \leq \sum_{k=0}^{\infty} |a_k| \leq \sum_{k=0}^{\infty} b_k$$

Beweis
Folgt aus Satz 14.8 und Satz 13.3. □

Korollar 14.11 *Sind* $\sum\limits_{k=0}^{\infty} |a_k|^2$ *und* $\sum\limits_{k=0}^{\infty} |b_k|^2$ *konvergent, dann ist* $\sum\limits_{k=0}^{\infty} a_k b_k$ *absolut konvergent.*

Beweisstrategie
Folgt wegen $|a_k b_k| \leq \frac{1}{2}|a_k|^2 + \frac{1}{2}|b_k|^2$.

Satz 14.12 (Quotientenkriterium) *Sei* $n \in \mathbb{N}_0$ *und* $a_k \neq 0$ *für alle* $k \geq n_0$.

a) *Falls ein* $q \in [0, 1[$ *und ein* $N > n_0$ *existiert, so dass*

$$\forall k \geq N : \left|\frac{a_{k+1}}{a_k}\right| \leq q$$

gilt, dann ist $\sum\limits_{k=0}^{\infty} a_k$ *absolut konvergent.*

b) *Falls ein* $N > n_0$ *existiert, so dass*

$$\forall k \geq N : \left|\frac{a_{k+1}}{a_k}\right| \geq 1$$

gilt, dann ist $\sum\limits_{k=0}^{\infty} a_k$ *divergent.*

Beweis

a) Aus $\left|\frac{a_{k+1}}{a_k}\right| \leq q$ für alle $k \geq N$ folgt:

$$\forall k > N : |a_{k+1}| \leq q|a_k| \leq ... \leq q^{k+1-N}|a_N|$$

$$\Rightarrow \sum_{k=N}^{\infty} |a_k| \leq \frac{|a_N|}{q^N} \sum_{k=N}^{\infty} q^k$$

Wegen $q < 1$ folgt aufgrund des Majorantenkriteriums und des Satzes 14.2 die absolute Konvergenz von $\sum\limits_{k=0}^{\infty} a_k$.

b) folgt mit Hilfe des Minorantenkriteriums. □

Korollar 14.13 *a)* $\limsup\limits_{k\to\infty} \left|\frac{a_{k+1}}{a_k}\right| \leq q < 1 \Rightarrow \sum\limits_{k=0}^{\infty} a_k$ *konvergiert absolut.*

b) $\liminf\limits_{k\to\infty} \left|\frac{a_{k+1}}{a_k}\right| > 1 \Rightarrow \sum\limits_{k=0}^{\infty} a_k$ *divergiert.*

Beweis
Die Voraussetzung von 14.13 a) impliziert die Voraussetzung von 14.12 a) und dasselbe gilt für die Voraussetzungen von 14.13 b) und 14.12 b). $\square$

Bemerkung
Gilt weder die Voraussetzung von 14.12 a) noch die Voraussetzung von 14.12 b), z.B. im Falle $\liminf\limits_{k\to\infty} |\frac{a_{k+1}}{a_k}| = 1$, dann kann $\sum\limits_{k=0}^{\infty} a_k$ konvergent sein wie im Falle $a_k = \frac{1}{(k+1)^2}$ oder divergent sein wie im Falle $a_k = \frac{1}{k+1}$.

Beispiel
Die Reihe $\displaystyle\sum_{n=0}^{\infty} \frac{x^n}{n!}$ konvergiert nach dem Quotientenkriterium für alle $x \in \mathbb{R}$ absolut, denn es gilt

$$\forall x \in \mathbb{R} : \left| \frac{\frac{x^{n+1}}{(n+1)!}}{\frac{x^n}{n!}} \right| = \frac{|x|}{n+1} \xrightarrow{n\to\infty} 0 < 1.$$

Später werden wir sehen, dass $\displaystyle\sum_{n=0}^{\infty} \frac{x^n}{n!} = e^x$ ist.

Satz 14.14 (Wurzelkriterium) *a) Falls ein* $q \in [0, 1[$ *und ein* $N \in \mathbb{N}_0$ *existiert, so dass*

$\forall k \geq N : \sqrt[k]{|a_k|} \leq q$

gilt, dann ist $\sum\limits_{k=0}^{\infty} a_k$ *absolut konvergent.*

b) Falls ein $N \in \mathbb{N}_0$ *existiert, so dass*

$\forall k \geq N : \sqrt[k]{|a_k|} \geq 1$

gilt, dann ist $\sum\limits_{k=0}^{\infty} a_k$ *divergent.*

Beweis

a) Unter der in a) gegebenen Voraussetzung ist $|a_k| \leq q^k$ für alle $k \geq N$ und wegen $q < 1$ konvergiert $\sum\limits_{k=0}^{\infty} a_k$ aufgrund des Majorantenkriteriums und des Satzes 14.2 absolut.

b) folgt mit Hilfe des Minorantenkriteriums. $\square$

Korollar 14.15 *a)* $\limsup\limits_{k\to\infty} \sqrt[k]{|a_k|} < 1 \Rightarrow \displaystyle\sum_{k=0}^{\infty} a_k$ *konvergiert absolut.*

b) $\limsup\limits_{k\to\infty} \sqrt[k]{|a_k|} > 1 \Rightarrow \sum\limits_{k=0}^{\infty} a_k$ *divergiert.*

Bemerkung
Gilt weder die Voraussetzung von 14.14 *a)* noch die Voraussetzung von 14.14 *b)*, z.B. im Falle $\limsup\limits_{k\to\infty} \sqrt[k]{|a_k|} = 1$, dann kann $\sum\limits_{k=0}^{\infty} a_k$ konvergent sein wie im Falle $a_k = (-1)^k \frac{1}{k+1}$ oder divergent sein, wie im Falle $a_k = \frac{1}{k+1}$.

Beispiel
$\sum\limits_{k=0}^{\infty} \dfrac{(2+(-1)^k)^k}{4^k}$ ist nach dem Wurzelkriterium absolut konvergent,

denn es gilt $\sqrt[k]{|a_k|} = \begin{cases} \sqrt[k]{\dfrac{3^k}{4^k}} = \dfrac{3}{4}, & \text{falls } k \text{ gerade,} \\ \sqrt[k]{\dfrac{1^k}{4^k}} = \dfrac{1}{4}, & \text{falls } k \text{ ungerade,} \end{cases}$

also $\limsup\limits_{k\to\infty} \sqrt[k]{|a_k]} = \dfrac{3}{4} < 1$.

Das Quotientenkriterium dagegen wäre bei dieser Reihe nicht anwendbar, denn es ist

$$\left|\frac{a_{k+1}}{a_k}\right| = \begin{cases} \dfrac{(\frac{1}{4})^{k+1}}{(\frac{3}{4})^k} = \dfrac{1}{4\cdot 3^{k+1}}, & \text{falls } k \text{ gerade,} \\ \dfrac{(\frac{3}{4})^{k+1}}{(\frac{1}{4})^k} = \dfrac{3^{k+1}}{4}, & \text{falls } k \text{ ungerade.} \end{cases}$$

Satz 14.16 (Umordnungssatz) *Sei* $\sum\limits_{k=0}^{\infty} a_k$ *absolut konvergent und* $\sigma : \mathbb{N}_0 \to \mathbb{N}_0$ *eine bijektive Abbildung. Dann ist die umgeordnete Reihe* $\sum\limits_{k=0}^{\infty} a_{\sigma(k)}$ *ebenfalls absolut konvergent und es gilt* $\sum\limits_{k=0}^{\infty} a_{\sigma(k)} = \sum\limits_{k=0}^{\infty} a_k$.

Beweis
Sei $\sum\limits_{k=0}^{\infty} a_k$ absolut konvergent.

$\Rightarrow \forall \varepsilon > 0\ \exists n_\varepsilon\ \forall m > n > n_\varepsilon : \sum\limits_{k=n}^{m} |a_k| < \varepsilon$

$\Rightarrow \forall \varepsilon > 0\ \forall m > n > max\{n_\varepsilon, \sigma(1), .., \sigma(n_\varepsilon)\} : \sum\limits_{k=n}^{m} |a_{\sigma(k)}| < \varepsilon$

$\overset{14,4}{\Rightarrow}$ $\sum\limits_{k=0}^{\infty} a_{\sigma(k)}$ ist absolut konvergent.

Außerdem gilt

$$|\sum_{k=0}^{n_\varepsilon} a_{\sigma(k)} - \sum_{k=0}^{\infty} a_k| = |\sum_{k=n_\varepsilon+1}^{\infty} a_{\sigma(k)}| \leq \sum_{k=n_\varepsilon+1}^{\infty} |a_{\sigma(k)}| < \varepsilon$$

und somit

$$\sum_{k=0}^{\infty} a_{\sigma(k)} = \sum_{k=0}^{\infty} a_k. \qquad \square$$

Bemerkung

Ist $\sum_{k=0}^{\infty} a_k$ konvergent, aber nicht absolut konvergent, dann kann man zeigen, dass es für jede Zahl $s \in \mathbb{R}$ eine Umordnung $\sigma_s : \mathbb{N}_0 \to \mathbb{N}_0$ gibt mit $\sum_{k=0}^{\infty} a_{\sigma_s(k)} = s$ sowie eine Umordnung $\psi : \mathbb{N}_0 \to \mathbb{N}_0$, so dass $\sum_{k=0}^{\infty} a_{\psi(k)}$ divergiert.

Es gilt also nur für absolut konvergente Reihen eine Verallgemeinerung des Kommutativgesetzes von endlich vielen auf unendlich viele Summanden.

Beispiel

Die Leibniz-Reihe $\displaystyle\sum_{k=0}^{\infty}(-1)^k \frac{1}{k+1}$ ist konvergent, aber nicht absolut konvergent. Wir betrachten die folgende Umordnung:

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \left(\frac{1}{5} + \frac{1}{7}\right) - \frac{1}{6} + \left(\frac{1}{9} + \frac{1}{11} + \frac{1}{13} + \frac{1}{15}\right) - \frac{1}{8} + ... + \left(\frac{1}{2^n+1} + \frac{1}{2^n+3} + \right.$$
$$\left. ... + \frac{1}{2 \cdot 2^n - 1}\right) - \frac{1}{2n+2} + ...$$

Wegen

$$\forall n \in \mathbb{N}_0 : \left(\frac{1}{2^n+1} + \frac{1}{2^n+3} + ... + \frac{1}{2 \cdot 2^n - 1}\right) \overset{Vollst.Ind.}{\geq} \frac{2^n}{2} \cdot \frac{1}{2 \cdot 2^n} = \frac{1}{4}$$

divergiert die umgeordnete Reihe.

Bemerkung

Im Allgemeinen gilt auch keine Verallgemeinerung des Assoziativgesetzes von endlich vielen auf unendlich viele Summanden. Denn es gilt

$$\sum_{n=0}^{\infty}(-1)^k = 1 - 1 + 1 - 1 + ... \text{ ist divergent,}$$

$$\sum_{n=0}^{\infty}((-1)^{2n} + (-1)^{2n+1}) = (1-1) + (1-1) + ... = 0,$$

$$1+\sum_{n=1}^{\infty}((-1)^{2n-1}+(-1)^{2n}) = 1+(-1+1)+(-1+1)+... = 1$$

Satz 14.17 (Cauchy-Produkt von Reihen) *Seien* $\sum\limits_{k=0}^{\infty} a_k$ *und* $\sum\limits_{l=0}^{\infty} b_l$ *absolut konvergent und* $c_n = \sum\limits_{k=0}^{n} a_k b_{n-k}$*, dann konvergiert auch* $\sum\limits_{n=0}^{\infty} c_n$ *absolut und es gilt*
$$\Big(\sum_{k=0}^{\infty} a_k\Big)\Big(\sum_{k=0}^{\infty} b_k\Big) = \sum_{n=0}^{\infty} c_n = a_0b_0 + (a_0b_1 + a_1b_0) + (a_0b_2 + a_1b_1 + a_2b_0) + \dots .$$
Die Reihe $\sum\limits_{n=0}^{\infty} c_n$ *heißt das Cauchy-Produkt von* $\sum\limits_{k=0}^{\infty} a_k$ *und* $\sum\limits_{k=0}^{\infty} b_k$*.*

Beispiele

1) Seien $x, y \in \mathbb{R}$, $a_k = \dfrac{x^k}{k!}$, $b_l = \dfrac{y^l}{l!}$ und $c_n = \sum\limits_{k=0}^{n} a_k b_{n-k}$.

   $$\Rightarrow c_n = \sum_{k=0}^{n} \frac{x^k}{k!}\frac{y^{n-k}}{(n-k)!} = \frac{1}{n!}\sum_{k=0}^{n}\binom{n}{k} x^k y^{n-k} = \frac{1}{n!}(x+y)^n$$

   Da $\sum\limits_{k=0}^{\infty} \dfrac{x^k}{k!}$ und $\sum\limits_{l=0}^{\infty} \dfrac{y^l}{l!}$ absolut konvergieren, folgt

   $$\Big(\sum_{n=0}^{\infty} \frac{(x+y)^n}{n!}\Big) = \Big(\sum_{k=0}^{\infty} \frac{x^k}{k!}\Big)\Big(\sum_{l=0}^{\infty} \frac{y^l}{l!}\Big).$$

2) Sei $|q| < 1$, $a_k = b_k = q^k$ und $c_n = \sum\limits_{k=0}^{n} a_k b_{n-k}$.

   $$\Rightarrow c_n = \sum_{k=0}^{n} q^k q^{n-k} = \sum_{k=0}^{n} q^n = (n+1)q^n$$

   Da $\sum\limits_{k=0}^{\infty} q^k$ für $|q| < 1$ absolut konvergiert, folgt

   $$\sum_{n=1}^{\infty} nq^{n-1} = \sum_{n=0}^{\infty}(n+1)q^n = \Big(\sum_{k=0}^{\infty} q^k\Big)\Big(\sum_{k=0}^{\infty} q^k\Big) = \frac{1}{(1-q)^2}.$$

3) Sei $a_k = (-1)^k \dfrac{1}{\sqrt{k+1}}$. Dann ist $\sum\limits_{k=0}^{\infty} a_k$ konvergent, aber nicht absolut konvergent. Für das Cauchy-Produkt $\sum\limits_{n=0}^{\infty} c_n$ von $\sum\limits_{k=0}^{\infty} a_k$ mit sich selber gilt

   $$c_n = (-1)^n \sum_{k=0}^{n} \frac{1}{\sqrt{k+1}\sqrt{n-k-1}}$$

$$\Rightarrow |c_n| = \frac{1}{\sqrt{n-1}} + \frac{1}{\sqrt{2}}\frac{1}{\sqrt{n-2}} + ... + \frac{1}{\sqrt{n-2}}\frac{1}{\sqrt{2}} + \frac{1}{\sqrt{n-1}} > (n+1)\frac{1}{n}$$

$\Rightarrow (c_n)_{n\in\mathbb{N}}$ ist keine Nullfolge.

$\Rightarrow \sum_{n=0}^{\infty} c_n$ divergiert.

Definition

a) *Reihen der Form* $\sum_{n=0}^{\infty} \frac{d_n}{10^n}$ *oder* $\sum_{n=0}^{\infty} \frac{-d_n}{10^n}$ *mit* $d_0 \in \mathbb{N}_0$ *und* $d_n \in \{0, 1, .., 9\}$ *für alle* $n \in \mathbb{N}$ *heißen Dezimalbrüche.*
Ist $a_N a_{N-1}...a_0$ *die Darstellung von* d_0 *im Dezimalsystem, dann schreibt man auch abkürzend* $\pm a_N a_{N-1}...a_0, d_1 d_2...$ *statt* $\pm \sum_{n=0}^{\infty} \frac{d_n}{10^n}$.

b) *Ein Dezimalbruch heißt abbrechend, falls gilt:*
$\exists m \in \mathbb{N}\ \forall n \geq m : d_n = 0.$
Kurzschreibweise: $\pm a_N a_{N-1}...a_0, d_1...d_{m-1}$.

c) *Ein Dezimalbruch heißt periodisch, falls er nicht abbrechend ist und falls gilt:*
$\exists j, k \in \mathbb{N}\ \forall n \geq j : d_{n+k} = d_n.$
Kurzschreibweise: $\pm a_N a_{N-1}...a_0, d_1...\overline{d_j d_{j+1}...d_{j+k-1}}$.
Ein Dezimalbruch hat eine Neunerperiode, falls gilt:
$\exists l \in \mathbb{N}\ \forall n \geq l : d_n = 9.$

Satz 14.18 a) *Jeder Dezimalbruch konvergiert gegen eine reelle Zahl.*

b) *Für jede reelle Zahl* x *existiert genau ein Dezimalbruch, der keine Neunerperiode hat und gegen* x *konvergiert. Dieser Dezimalbruch heißt Dezimalbruchentwicklung von* x.

c) *Eine reelle Zahl ist genau dann rational wenn ihre Dezimalbruchentwicklung abbrechend oder periodisch ist.*

Beweisskizze

a) Sei $\sum_{n=0}^{\infty} \frac{d_n}{10^n}$ ein Dezimalbruch, $s_m = \sum_{n=0}^{m} \frac{d_n}{10^n}$ und $S_m = s_m + \frac{1}{10^m}$.
Dann ist $\{[s_m, S_m] : m \in \mathbb{N}_0\}$ eine Intervallschachtelung und somit besitzen $\pm \sum_{n=0}^{\infty} \frac{d_n}{10^n}$ jeweils einen reellen Granzwert.

b) Sei $x \in \mathbb{R}_0^+$.
Setze $d_0 = \lfloor x \rfloor = max\{n \in \mathbb{N}_0 : n \leq x\}$
(die sogenannte untere Gaußklammer von x)

und $\forall n \in \mathbb{N}$:

$$d_n = \lfloor 10^n (x - \sum_{m=0}^{n-1} \frac{d_m}{10^m}) \rfloor$$

Dann sind $\pm \sum_{n=0}^{\infty} \frac{d_n}{10^n}$ Dezimalbrüche mit Grenzwert $\pm x$.

Unter Ausnutzung von $0,\overline{9} = 1$ zeigt man dann, dass die oben konstruierten Dezimalbrüche die einzigen Dezimalbrüche ohne Nennerperiode sind, die gegen x bzw. $-x$ konvergieren.

c) zeigt man unter Ausnutzung der Bruchrechenregeln und der Formel für den Grenzwert der geometrischen Reihe.

Beispiel

$$\begin{array}{llllll} \frac{1}{6} & = & 1 & : & 6 & = 0,1\bar{6} \\ & & 10 & & & \\ & & \underline{\ 6} & & & \\ & & \ 40 & & & \\ & & \ \underline{36} & & & \\ & & \quad 4 & & & \end{array}$$

Bemerkung

Entsprechend kann man für $x \in \mathbb{R}$ auch eine g-adische Entwicklung

$$x = (x_N x_{N-1} ... x_0, x_{-1} x_{-2} ...)_g := \sum_{n=0}^{\infty} a_{N-n} g^{N-n}$$

mit $N \in \mathbb{N}$ und $x_j \in \{0, 1, ..., g-1\}$ konstruieren.

Definition

a) Zwei Mengen A, B heißen gleich mächtig, falls es eine bijektive Abbildung $f : A \to B$ gibt.

b) Eine Menge A heißt endlich, falls ein $m \in \mathbb{N}$ existiert, so dass A und $\{n \in \mathbb{N} : n < m\}$ gleich mächtig sind.

c) Eine Menge A heißt unendlich, falls sie nicht endlich ist.

d) Eine Menge A heißt abzählbar unendlich, falls sie gleich mächtig wie $\mathbb{N}$ ist.

e) Eine Menge A heißt überabzählbar, falls sie unendlich und nicht abzählbar unendlich ist.

Beispiele

1) $\{1, 2, 3\}$ und $\{3, 4, 5\}$ sind gleich mächtig und endlich.

2) $\emptyset$ ist endlich.

3) $\mathbb{N}$, $\{n+1 : n \in \mathbb{N}\}$ und $\{2n : n \in \mathbb{N}\}$ sind gleich mächtig und abzählbar unendlich.

4) Eine Menge A ist genau dann endlich, wenn jede injektive Abbildung $f : A \to A$ auch surjektiv ist.

Satz 14.19 $\mathbb{Q}$ *ist abzählbar unendlich.*

Beweisidee
Cantorsches Diagonalverfahren:

$$
\begin{array}{cccccccccccc}
\frac{1}{1} & & \frac{2}{1} & \to & \frac{3}{1} & & \frac{4}{1} & \to & \frac{5}{1} & & \frac{6}{1} & \cdots \\
\downarrow & \nearrow & & \diagup & & \nearrow & & \diagup & & \nearrow & & \\
\frac{1}{2} & & \left(\frac{2}{2}=\frac{1}{1}\right) & & \frac{3}{2} & & \left(\frac{4}{2}=\frac{2}{1}\right) & & \frac{5}{2} & & \cdots & \\
 & \swarrow & & \nearrow & & \diagup & & \nearrow & & & & \\
\frac{1}{3} & & \frac{2}{3} & & \left(\frac{3}{3}=\frac{1}{1}\right) & & \frac{4}{3} & & \cdots & & & \\
\downarrow & \nearrow & & \diagup & & \nearrow & & & & & & \\
\frac{1}{4} & & \left(\frac{2}{4}=\frac{1}{2}\right) & & \frac{3}{4} & & \cdots & & & & & \\
 & \swarrow & & \nearrow & & & & & & & & \\
\frac{1}{5} & & \frac{2}{5} & & \cdots & & & & & & & \\
\downarrow & \nearrow & & & & & & & & & & \\
\frac{1}{6} & & \cdots & & & & & & & & &
\end{array}
$$

Es liefert $1 \mapsto \frac{1}{1}$, $2 \mapsto -\frac{1}{1}$, $3 \mapsto \frac{1}{2}$, $4 \mapsto -\frac{1}{2}$, $5 \mapsto \frac{2}{1}$, $6 \mapsto -\frac{2}{1}$, $7 \mapsto \frac{3}{1}$, $8 \mapsto -\frac{3}{1}$, $9 \mapsto \frac{1}{3}$, $10 \mapsto -\frac{1}{3}$, $\cdots$ eine bijektive Abbildung $\mathbb{N} \to \mathbb{Q}$.

Satz 14.20 $\mathbb{R}$ *ist überabzählbar.*

Beweis
Sei eine beliebige Abbildung $f : \mathbb{N} \to \mathbb{R}$ gegeben. Zeige, dass f nicht surjektiv und

somit nicht bijektiv sein kann.
Sei $x = 0, d_1 d_2 ...$, wobei

$$d_j = \begin{cases} 0, & \text{falls } f(j) = \tilde{d}_0, \tilde{d}_1 \tilde{d}_2 ... \wedge \tilde{d}_j \neq 0, \\ 1, & \text{falls } f(j) = \tilde{d}_0, \tilde{d}_1 \tilde{d}_2 ... \wedge \tilde{d}_j = 0. \end{cases}$$

$\Rightarrow x \in \mathbb{R} \wedge \forall j \in \mathbb{N} : x \neq f(j)$
$\Rightarrow x \in \mathbb{R} \setminus f(\mathbb{N})$ $\square$

Definition *Sei $(a_n)_{n \in \mathbb{N}_0}$ eine Folge in $\mathbb{C}$. Dann heißt $\sum\limits_{k=0}^{\infty} a_k$ eine komplexe (unendliche) Reihe.*

Satz 14.21 *Für komplexe Reihen gelten die Aussagen aus den Sätzen bzw. Korollaren* 14.1 *(mit $\alpha, \beta \in \mathbb{C}$),* 14.2*,* 14.4-14.8 *und* 14.10-14.17.

Beweisstrategie
Analog zum reellen Fall.

Beispiele

1) Die komplexe geometrische Reihe $\sum\limits_{k=0}^{\infty} z^k$ konvergiert für $|z| < 1$ aufgrund der geometrischen Summenformel und divergiert für $|z| \geq 1$ wegen Korollar 14.5.

2) Die Reihe $\sum\limits_{n=0}^{\infty} \frac{z^n}{n!}$ konvergiert nach dem Quotientenkriterium für alle $z \in \mathbb{C}$ absolut und wegen Satz 14.17 gilt für alle $z, w \in \mathbb{C}$:

$$\Big(\sum_{n=0}^{\infty} \frac{(z+w)^n}{n!} \Big) = \Big(\sum_{k=0}^{\infty} \frac{z^k}{k!} \Big) \Big(\sum_{l=0}^{\infty} \frac{w^l}{l!} \Big)$$

15 Stetigkeit

Definition *Sei $M \subseteq \mathbb{R}$. $x \in \mathbb{R}$ heißt Häufungspunkt von M, falls eine Folge $(x_n)_{n\in\mathbb{N}}$ existiert mit $x_n \in M\backslash\{x\}$ für alle $n \in \mathbb{N}$ und $\lim\limits_{n\to\infty} x_n = x$.*

Beispiele

1) Sei $M = [0,1[$. Dann gilt: x ist Häufungspunkt von M $\Leftrightarrow x \in [0,1]$.

2) Sei $M = \{\frac{1}{n} : n \in \mathbb{N}\}$. Dann ist 0 der einzige Häufungspunkt von M.

Definition *Sei $D \subseteq \mathbb{R}$ und $f : D \to \mathbb{R}$ eine Funktion.*

a) Sei $x_0 \in \mathbb{R}$ ein Häufungspunkt von D. Die Zahl $a \in \mathbb{R}$ heißt Grenzwert der Funktion f in x_0, in Zeichen: $\lim\limits_{x\to x_0} f(x) = a$, falls $\lim\limits_{n\to\infty} f(x_n) = a$ für jede Folge $(x_n)_{n\in\mathbb{N}}$ mit $x_n \in D\backslash\{x_0\}$ für alle $n \in \mathbb{N}$ und $\lim\limits_{n\to\infty} x_n = x_0$.

b) Sei $x_0 \in \mathbb{R}$ ein Häufungspunkt von $D \cap]-\infty, x_0[$. Die Zahl $a \in \mathbb{R}$ heißt linksseitiger Grenzwert von f in x_0, in Zeichen: $\lim\limits_{x\to x_0^-} f(x) = a$, falls $\lim\limits_{n\to\infty} f(x_n) = a$ für jede Folge $(x_n)_{n\in\mathbb{N}}$ in $D \cap]-\infty, x_0[$ mit $\lim\limits_{n\to\infty} x_n = x_0$.

c) Sei $x_0 \in \mathbb{R}$ ein Häufungspunkt von $D \cap]x_0, \infty[$. Die Zahl $a \in \mathbb{R}$ heißt rechtsseitiger Grenzwert von f in x_0, in Zeichen: $\lim\limits_{x\to x_0^+} f(x) = a$, falls $\lim\limits_{n\to\infty} f(x_n) = a$ für jede Folge $(x_n)_{n\in\mathbb{N}}$ in $D \cap]x_0, \infty[$ mit $\lim\limits_{n\to\infty} x_n = x_0$.

d) Es existiere eine Folge $(x_n)_{n\in\mathbb{N}}$ in D mit $x_n \to \infty$ für $n \to \infty$. Dann heißt die Zahl $a \in \mathbb{R}$ Grenzwert von f für $x \to \infty$, in Zeichen: $\lim\limits_{x\to\infty} f(x) = a$, falls $\lim\limits_{n\to\infty} f(x_n) = a$ für jede Folge $(x_n)_{n\in\mathbb{N}}$ in D mit $x_n \to \infty$ für $n \to \infty$.

e) Analog definiert man $\lim\limits_{x\to-\infty} f(x) = a$.

Beispiele

1) $f : \mathbb{R} \to \mathbb{R},\ f(x) = x,\ g : \mathbb{R} \to \mathbb{R},\ g(x) = x^2$
$\lim\limits_{x\to 2} f(x) = 2,\ \lim\limits_{x\to 2} g(x) = 4$

2) $h : \mathbb{R}\backslash\{0\} \to \mathbb{R},\ h(x) = \frac{1}{x}$
$\lim\limits_{x\to\infty} h(x) = \lim\limits_{x\to-\infty} h(x) = 0$

3) $s : \mathbb{R} \to \mathbb{R},\ s(x) = \begin{cases} -1, & x < 0 \\ 0, & x = 0 \\ 1, & x > 0 \end{cases}$

$\lim\limits_{x\to 0} s(x)$ existiert nicht,

$\lim\limits_{x\to 0^-} s(x) = -1,\ \lim\limits_{x\to 0^+} s(x) = 1$

4) $q : \mathbb{R} \to \mathbb{R}$, $q(x) = \begin{cases} 0, & x \neq 0 \\ 1, & x = 0 \end{cases}$

$\lim\limits_{x \to 0} q(x) = 0$

Bemerkung
In manchen Büchern findet man eine leicht modifizierte Definition von $\lim\limits_{x \to x_0} f(x) = a$, bei der $\lim\limits_{n \to \infty} f(x_n) = a$ für jede Folge $(x_n)_{n \in \mathbb{N}}$ in D (und nicht nur in $D \backslash \{x_0\}$) gefordert wird. Verwendet man diese Definition, dann würde im vorigen Beispiel $\lim\limits_{x \to 0} q(x)$ nicht existieren.

Bemerkung
Die Gesetze für konvergente Folgen aus Kapitel 13 übertragen sich sinngemäß auf Grenzwerte von Funktionen. Beispielsweise gilt:
Sei $f : D \to \mathbb{R}$, $g : D \to \mathbb{R}$, $\lim\limits_{x \to x_0} f(x) = a$ und $\lim\limits_{x \to x_0} g(x) = b$. Dann folgt

$$\forall \alpha, \beta \in \mathbb{R} : \lim_{x \to x_0} (\alpha f(x) + \beta g(x)) = \alpha a + \beta b.$$

Satz 15.1 (ε-δ-Kriterium) *Sei $D \subseteq \mathbb{R}$, $f : D \to \mathbb{R}$ eine Funktion und $x_0 \in \mathbb{R}$ ein Häufungspunkt von D. Dann ist äquivalent:*

(i) $\lim\limits_{x \to x_0} f(x) = a$

(ii) $\forall \varepsilon > 0 \exists \delta > 0 \forall x \in D \backslash \{x_0\} : |x - x_0| < \delta \Rightarrow |f(x) - a| < \varepsilon$

Beweis
(ii) $\Rightarrow$ (i):
Sei $(x_n)_{n \in \mathbb{N}}$ eine beliebige Folge in $D \backslash \{x_0\}$ mit $\lim\limits_{n \to \infty} x_n = x_0$ und $\varepsilon > 0$.
Wegen (ii) können wir ein $\delta > 0$ wählen, so dass gilt:
$\forall x \in D \backslash \{x_0\} : |x - x_0| < \delta \Rightarrow |f(x) - a| < \varepsilon$.
Wegen $\lim\limits_{n \to \infty} x_n = x_0$ folgt:
$\exists n_\delta \in \mathbb{N}\, \forall n > n_\delta : |x_n - x_0| < \delta \Rightarrow \forall n > n_\delta : |f(x_n) - a| < \varepsilon$. Also gilt $\lim\limits_{n \to \infty} f(x_n) = a$.
(i) $\Rightarrow$ (ii): indirekter Beweis
Angenommen, (ii) gelte nicht.
$\Rightarrow\ \exists \varepsilon > 0\ \forall \delta > 0\ \exists x \in D \backslash \{x_0\} : |x - x_0| < \delta \wedge |f(x) - a| \geq \varepsilon$
$\Rightarrow\ \exists \varepsilon > 0\ \forall n > 0\ \exists x_n \in D \backslash \{x_0\} : |x_n - x_0| < \frac{1}{n} \wedge |f(x_n) - a| \geq \varepsilon$
$\Rightarrow \lim\limits_{n \to \infty} x_n = x_0 \wedge \neg(\lim\limits_{n \to \infty} f(x_n) = a)$. Somit kann (i) nicht gelten. $\square$

Definition *Sei $D \subseteq \mathbb{R}$ und $f : D \to \mathbb{R}$ eine Funktion.*

a) f heißt stetig in $x_0 \in D$, falls $\lim\limits_{n \to \infty} f(x_n) = f(x_0)$ für jede Folge $(x_n)_{n \in \mathbb{N}}$ in D mit $\lim\limits_{n \to \infty} x_n = x_0$.

b) f heißt stetig auf der Menge $M \subseteq D$, falls f in jedem $x \in M$ stetig ist.

Satz 15.2 *Sei $D \subseteq \mathbb{R}$, $f : D \to \mathbb{R}$ eine Funktion und $x_0 \in D$. Dann ist äquivalent:*

(i) f ist stetig in x_0.

(ii) $\forall \varepsilon > 0 \exists \delta > 0 \forall x \in D : |x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \varepsilon$

(iii) Ist x_0 Häufungspunkt von D, dann folgt $\lim\limits_{x \to x_0} f(x) = f(x_0)$.

Korollar 15.3 *Sei $D \subseteq \mathbb{R}$, $f : D \to \mathbb{R}$ stetig in $x_0 \in D$ und $f(x_0) \neq 0$. Dann gilt: $\exists \delta > 0 \forall x \in D : |x - x_0| < \delta \Rightarrow f(x) \neq 0$.*

Beispiele

1) Konstante Funktionen sind auf ihrem gesamten Definitionsbereich stetig.

2) Die Identität $Id : \mathbb{R} \to \mathbb{R}$, $x \mapsto x$ ist auf ganz $\mathbb{R}$ stetig.

3) Die Betragsfunktion $|\cdot| : \mathbb{R} \to \mathbb{R}$, $x \mapsto |x|$ ist auf ganz $\mathbb{R}$ stetig.

4) Die Exponentialfunktion $\exp : \mathbb{R} \to \mathbb{R}^+$, $x \mapsto e^x$ ist auf ganz $\mathbb{R}$ stetig.
Denn aufgrund von $\lim\limits_{n \to \infty} e^{\frac{1}{n}} = \lim\limits_{n \to \infty} e^{-\frac{1}{n}} = 1 = e^0$ (siehe 13.13 *a)j)*) und da exp streng monoton steigend ist (13.13 *h)*) gilt $\lim\limits_{x \to 0} e^x = e^0$. Wegen $e^{x+y} = e^x e^y$ (13.13 *c)*) folgt: $\forall x_0 \in \mathbb{R} : \lim\limits_{x \to x_0} e^x = \lim\limits_{x \to x_0} e^{x_0} e^{x - x_0} = e^{x_0} \lim\limits_{y \to 0} e^y = e^{x_0}$.

5) Der Sinus und der Kosinus sind auf ganz $\mathbb{R}$ stetig.
Denn aufgrund von $|\sin(x)| \leq |x|$ und $1 - \frac{1}{2}x^2 \leq \cos(x) \leq 1$ für $|x| < \frac{\pi}{2}$ (siehe Kapitel 5.3) ist $\lim\limits_{x \to 0} \sin(x) = 0 = \sin(0)$ und $\lim\limits_{x \to 0} \cos(x) = 1 = \cos(0)$. Aufgrund der Additionstheoreme ergibt sich aus diesen Grenzwerten auch $\lim\limits_{x \to x_0} \sin(x) = \sin(x_0)$ und $\lim\limits_{x \to x_0} \cos(x) = \cos(x_0)$ für alle $x_0 \in \mathbb{R}$.

6) Die Heavisidefunktion $H : \mathbb{R} \to \mathbb{R}$, $H(x) = \begin{cases} 0, & x < 0, \\ 1, & x \geq 0, \end{cases}$ ist für alle $x \neq 0$ stetig und in 0 nicht stetig.

7) Sei $f : \mathbb{R} \to \mathbb{R}$, $f(x) = \begin{cases} x, & x \in \mathbb{R} \backslash \mathbb{Q}, \\ 0, & x \in \mathbb{Q}. \end{cases}$
Dann ist f in 0 stetig und für alle $x \in \mathbb{R} \backslash \{0\}$ nicht stetig.

Satz 15.4 *Sei $D_f \subseteq \mathbb{R}$, $D_g \subseteq \mathbb{R}$ und $f : D_f \to \mathbb{R}$, $g : D_g \to \mathbb{R}$ Funktionen. Dann gilt:*

a) Ist f stetig in x_0, dann ist auch $|f|$ stetig in x_0.

b) Sind f und g stetig in x_0, dann sind auch $\alpha f + \beta g$ für alle $\alpha, \beta \in \mathbb{R}$ stetig in x_0.
Folglich ist $C^0(M) := \{f : M \to \mathbb{R} : f$ stetig auf $M\}$ ein $\mathbb{R}$-Vektorraum.

c) Sind f und g stetig in x_0, dann ist auch fg stetig in x_0.

d) Ist f stetig in x_0 und $f(x_0) \neq 0$, dann ist auch $\frac{1}{f}$ stetig in x_0.

e) Ist g stetig in x_0 und f stetig in $g(x_0)$, dann ist $f \circ g$ stetig in x_0.

Beweis
a) - *d)* folgen aus der Definition der Stetigkeit und den Rechenregeln für Grenzwerte.
e) Sei $(x_n)_{n\in\mathbb{N}}$ eine Folge im Definitionsbereich von $f \circ g$ mit $\lim\limits_{n\to\infty} x_n = x_0$. Da g stetig in x_0 ist, folgt $\lim\limits_{n\to\infty} g(x_n) = g(x_0)$ und da f stetig in $g(x_0)$ ist, folgt $\lim\limits_{n\to\infty} f \circ g(x_n) = \lim\limits_{n\to\infty} f(g(x_n)) = f(g(x_0)) = f \circ g(x_0)$. $\square$

Beispiele

1) Polynome sind auf ganz $\mathbb{R}$ stetig.
2) Gebrochen-rationale Funktionen sind auf ihrem gesamten Definitionsbereich stetig.
3) Der Tangens ist stetig auf $]-\frac{\pi}{2}, \frac{\pi}{2}[$.
4) Sei $n \in \mathbb{N}_0$ und $f : \mathbb{R} \to \mathbb{R}$ mit $f(x) = \begin{cases} |x|^n \sin\left(\frac{1}{|x|}\right), & x \neq 0, \\ 0, & x = 0. \end{cases}$

   Dann ist f auf $\mathbb{R}\backslash\{0\}$ stetig und, falls $n > 0$ ist, auch in 0.

Definition *Eine Funktion $f : D \to \mathbb{R}$ heißt stetig fortsetzbar in $x_0 \notin D$, falls $\lim\limits_{x\to x_0} f(x)$ existiert. In diesem Fall heißt $\tilde{f} : D \cup \{x_0\} \to \mathbb{R}$,*

$$\tilde{f}(x) = \begin{cases} f(x), & x \in D\,, \\ \lim\limits_{x\to x_0} f(x), & x = x_0\,, \end{cases}$$

die stetige Fortsetzung von f auf $D \cup \{x_0\}$.

Beispiel
Sei $n \in \mathbb{N}$ und $f : \mathbb{R}\backslash\{1\} \to \mathbb{R}$, $f(x) = \dfrac{x^2-1}{(x-1)^n}$.
Wegen $\dfrac{x^2-1}{(x-1)^n} = \dfrac{x+1}{(x-1)^{n-1}}$ für alle $x \in \mathbb{R}\backslash\{1\}$ folgt:
$n = 1 : \lim\limits_{x\to 1} f(x) = 2 \;\Rightarrow\; f$ ist in 1 stetig fortsetzbar durch 2.

$n > 1 \wedge n$ ungerade: $f(x) \to \infty$ für $x \to 1$, d.h. $f(x_n) \to \infty$ für jede Folge $(x_n)_{n\in\mathbb{N}}$ in $\mathbb{R}\backslash\{1\}$ mit $\lim\limits_{n\to\infty} x_n = 1$. Man sagt, f hat in $x = 1$ eine Polstelle ohne Vorzeichenwechsel.
n gerade: $f(x) \to -\infty$ für $x \to 1^-$ und $f(x) \to \infty$ für $x \to 1^+$. Man sagt, f hat in $x = 1$ eine Polstelle mit Vorzeichenwechsel.

Satz 15.5 (Zwischenwertsatz) *Sei $f : \mathbb{R} \supseteq D \to \mathbb{R}$ stetig und $[a, b] \subseteq D$. Dann nimmt f jeden Wert y zwischen f(a) und f(b) mindestens einmal an.*
Insbesondere gilt: Ist $(f(a) \leq 0 \wedge f(b) \geq 0)$ oder $(f(a) \geq 0 \wedge f(b) \leq 0)$, dann hat f mindestens eine Nullstelle in $[a, b]$.

Beweis
Intervallhalbierungsverfahren:
Sei o.B.d.A. $y = 0$ und $f(a) \leq 0 \leq f(b)$. (Denn für $y \neq 0$ kann man $f - y$ und für $f(a) \geq y \geq f(b)$ kann man $-(f - y)$ betrachten.)
Zeige: $\exists x \in\,]a, b[\colon\; f(x) = 0$.
Ist $f(a) = 0 \vee f(b) = 0$, dann sind wir bereits fertig. Anderenfalls definiere rekursiv:

$a_1 := a \wedge b_1 := b$,

$a_{n+1} := a_n \wedge b_{n+1} := \frac{a_n+b_n}{2}$, falls $f(\frac{a_n+b_n}{2}) \geq 0$,

$a_{n+1} := \frac{a_n+b_n}{2} \wedge b_{n+1} := b_n$, falls $f(\frac{a_n+b_n}{2}) < 0$.

Ist $f(\frac{a_m+b_m}{2}) = 0$ für ein $m \in \mathbb{N}$, dann liefert der obige Algorithmus nach endlich vielen Schritten eine Nullstelle von f in $[a, b]$.
Anderenfalls erhält man eine Intervallschachtelung $\{[a_n, b_n] : n \in \mathbb{N}\}$. Für $x = \lim\limits_{n\to\infty} a_n = \lim\limits_{n\to\infty} b_n$ gilt aufgrund der Stetigkeit von f : $f(x) = \lim\limits_{n\to\infty} f(a_n) = \lim\limits_{n\to\infty} f(b_n)$.
Da nach Konstruktion der Intervallschachtelung $f(a_n) \leq 0$ und $f(b_n) \geq 0$ für alle $n \in \mathbb{N}$ ist, folgt aufgrund von Satz 13.3, dass $f(x) = 0$ ist. $\square$

Beispiel

Jedes reelle Polynom von ungeradem Grad hat mindestens eine Nullstelle.

Satz 15.6 (Satz vom Maximum und Minimum) *Sei I ein beschränktes, abgeschlossenes Intervall und $f : I \to \mathbb{R}$ stetig. Dann nimmt f auf I ein Maximum und ein Minimum an, d.h. $\exists x_1, x_2 \in I\, \forall x \in I : f(x_1) \leq f(x) \leq f(x_2)$.*

Beweis
Wir zeigen zunächst, dass f beschränkt ist, d.h. $\exists M\; \forall x \in I : |f(x)| \leq M$. Angenommen, f sei unbeschränkt.

$\Rightarrow\; \forall n \in \mathbb{N}\; \exists x_n \in I : |f(x_n)| > n \qquad (*)$

Da I beschränkt ist, ist $(x_n)_{n\in\mathbb{N}}$ eine beschränkte Folge. Nach dem Satz von Bolzano-Weierstraß (Satz 13.9) besitzt daher $(x_n)_{n\in\mathbb{N}}$ eine konvergente Teilfolge $(x_{n_k})_{k\in\mathbb{N}}$.

Da I beschränkt und abgeschlossen ist, $\exists a, b \in \mathbb{R} : I = [a, b]$.
Also gilt $a \leq x_{n_k} \leq b$ für alle n_k und wegen Satz 13.3 auch $a \leq \lim\limits_{k\to\infty} x_{n_k} \leq b$. Da f stetig auf I ist, folgt $\lim\limits_{k\to\infty} f(x_{n_k}) = f\left(\lim\limits_{k\to\infty} x_{n_k}\right)$.
Nach Satz 13.1 muss dann aber $(f(x_{n_k}))_{k\in\mathbb{N}}$ eine beschränkte Folge sein, was ein Widerspruch zu $(*)$ ist.
Also muss f beschränkt sein.
Daher existiert $s = \sup\{f(x) : x \in I\}$

$\Rightarrow\ \forall n \in \mathbb{N}\ \exists x_n \in I :\ s - \frac{1}{n} < f(x_n) \leq s$

$\Rightarrow\ \lim\limits_{n\to\infty} f(x_n) = s$

Da $(x_n)_{n\in\mathbb{N}}$ beschränkt ist, besitzt $(x_n)_{n\in\mathbb{N}}$ eine konvergente Teilfolge $(x_{n_k})_{k\in\mathbb{N}}$ mit Grenzwert in I. Nach Satz 13.8 folgt $\lim\limits_{k\to\infty} f(x_{n_k}) = s$ und, da f stetig ist, auch $f(\lim\limits_{k\to\infty} x_{n_k}) = s$.

$\Rightarrow\ s = \max\{f(x) : x \in I\}$

Da $-f$ ebenfalls stetig auf I ist, nimmt auch $-f$ ein Maximum auf I an, das zugleich das Minimum von f auf I ist. □

Beispiele

1) $x \mapsto \sin(x)$ ist stetig auf $[0, 2\pi]$ und nimmt bei $\frac{\pi}{2}$ ein Maximum und bei $\frac{3}{2}\pi$ ein Minimum an.

2) $x \mapsto \frac{1}{x}$ ist stetig auf dem beschränkten, aber nicht abgeschlossenen Intervall $]0, 1[$ und nimmt auf $]0, 1[$ weder ein Maximum noch ein Minimum an.

3) $x \mapsto x$ ist auf dem abgeschlossenen, aber nicht beschränkten Intervall $[0, \infty[$ stetig, aber nicht beschränkt.

Satz 15.7 (Umkehrsatz für stetige Funktionen) *Sei $f : [a, b] \to \mathbb{R}$ stetig und streng monoton wachsend. Dann bildet f das Intervall $[a, b]$ bijektiv auf das Intervall $[f(a), f(b)]$ ab. Die Umkehrfunktion $f^{-1} : [f(a), f(b)] \to [a, b]$ ist ebenfalls stetig und streng monoton wachsend.*

Beweis:

- f ist bijektiv als Funktion von $[a, b]$ nach $[f(a), f(b)]$.
  Da f streng monoton wachsend ist, ist f injektiv und $f([a, b]) \subseteq [f(a), f(b)]$.
  Nach dem Zwischenwertsatz muss dann $f([a, b]) = [f(a), f(b)]$ sein.

- f^{-1} ist streng monoton wachsend.
  Sei $y_1 < y_2$ mit $y_j = f(x_j)$. Angenommen, $f^{-1}(y_1) \geq f^{-1}(y_2)$
  $\Rightarrow\ x_1 = f^{-1}(y_1) \geq f^{-1}(y_2) = x_2$
  $\Rightarrow\ y_1 = f(x_1) \geq f(x_2) = y_2$ ↯

- f^{-1} ist stetig
  Sei $y_0 = f(x_0)$ und $\varepsilon > 0$.
  Da f streng monoton wachsend ist, gilt $f(x_0 - \varepsilon) < f(x_0) < f(x_0 + \varepsilon)$
  $\Rightarrow\ \exists \delta > 0:\ f(x_0 - \varepsilon) < f(x_0) - \delta \wedge f(x_0) + \delta < f(x_0 + \varepsilon)$
  $\Rightarrow\ \forall y \in]f(x_0) - \delta, f(x_0) + \delta[:\ y \in [f(x_0 - \varepsilon), f(x_0 + \varepsilon)] = f([x_0 - \varepsilon, x_0 + \varepsilon])$
  $\Rightarrow\ f^{-1}(y) \in [x_0 - \varepsilon, x_0 + \varepsilon] \Rightarrow |f^{-1}(y) - x_0| < \varepsilon$. Nach Satz 15.2 folgt die Stetigkeit von f^{-1}. $\square$

Bemerkungen

1) Man kann in Satz 15.7 auch streng monoton wachsend durch streng monoton fallend oder $[a, b]$ und $[f(a), f(b)]$ durch $]a, b[$ und $]f(a), f(b)[$ ersetzen, und die dadurch entstehenden Aussagen sind auch gültig.

2) Man kann zeigen, dass jede auf einem beliebigen Intervall stetige und injektive Funktion entweder streng monoton wachsend oder streng monoton fallend sein muss.

Beispiele

1) Nach Satz 15.7 ist $x \mapsto \ln(x)$ stetig auf $\mathbb{R}^+$. Infolgedessen sind auch die Funktionen

   $x \mapsto \log_b(x) = \dfrac{\ln(x)}{\ln(b)}$ für jedes $b \in \mathbb{R}^+ \setminus \{1\}$,

   $x \mapsto x^r = e^{r \ln(x)}$ für jedes $r \in \mathbb{R}$,

   $x \mapsto a^x = e^{x \ln(a)}$ für jedes $a \in \mathbb{R}^+$,

   $x \mapsto x^x = e^{x \ln(x)}$

   stetig auf $\mathbb{R}^+$.
   Folglich erhalten wir auch: $\forall a \in \mathbb{R}^+:\ \lim\limits_{n\to\infty} \sqrt[n]{a} = 1,\ \lim\limits_{n\to\infty} \sqrt[n]{n} = 1$.

2) Der Arkussinus und der Arkuskosinus sind stetig auf $[-1, 1]$. Der Arkustangens ist stetig auf $\mathbb{R}$. Insbesondere erhalten wir auch $\lim\limits_{x\to-\infty} \arctan(x) = -\dfrac{\pi}{2}$, $\lim\limits_{x\to\infty} \arctan(x) = \dfrac{\pi}{2}$.

3) $x \mapsto \ln\left(\sqrt{1 + |\arctan(x)|}\right)$ ist stetig auf ganz $\mathbb{R}$.

Definition

a) $z \in \mathbb{C}$ heißt Häufungspunkt von $M \subseteq \mathbb{C}$, falls eine Folge $(z_n)_{n\in\mathbb{N}}$ in $M \setminus \{z\}$ existiert mit $\lim\limits_{n\to\infty} z_n = z$.

b) *Sei $D \subseteq \mathbb{R}$ oder $D \subseteq \mathbb{C}$, $f : D \to \mathbb{C}$ eine Funktion und z_0 ein Häufungspunkt von D. Die Zahl $a \in \mathbb{C}$ heißt Grenzwert der Funktion f in z_0, in Zeichen: $\lim\limits_{z \to z_0} f(z) = a$, falls $\lim\limits_{n \to \infty} f(z_n) = a$ für jede Folge $(z_n)_{n \in \mathbb{N}}$ in $D \backslash \{z_0\}$ mit $\lim\limits_{n \to \infty} z_n = z_0$.*

c) *Sei $D \subseteq \mathbb{R}$ oder $D \subseteq \mathbb{C}$ und $f : D \to \mathbb{C}$ eine Funktion. f heißt stetig in $z_0 \in D$, falls $\lim\limits_{n \to \infty} f(z_n) = f(z_0)$ für jede Folge $(z_n)_{n \in \mathbb{N}}$ in D mit $\lim\limits_{n \to \infty} z_n = z_0$. f heißt stetig auf $M \subseteq D$, falls f in jedem $z \in M$ stetig ist.*

Bemerkungen

1) Die Gesetze für konvergente komplexwertige Folgen aus Kapitel 13 übertragen sich sinngemäß auf Grenzwerte von komplexwertigen Funktionen.

2) Die Sätze 15.1, 15.2 (samt Korollar 15.3) und 15.4 übertragen sich sinngemäß auf den komplexen Fall.

Beispiel

Komplexe Polynome sind stetig auf ganz $\mathbb{C}$.

16 Differenzierbarkeit

Motivation

1) Sei $I \subseteq \mathbb{R}$ ein Intervall und $f : I \to \mathbb{R}$ eine Funktion.
Problem: Existiert im Punkt $(x_0, f(x_0)) \in I \times f(I)$ eine Tangente an den Graphen von f ? Wenn ja, welche Steigung hat sie?
Lösungsstrategie: Betrachte für alle $x \in I$, die nahe bei x_0 liegen, die jeweilige Sekante durch die Punkte $(x_0, f(x_0))$ und $(x, f(x))$. Ist der Graph von f in der Nähe von $(x_0, f(x_0))$ genügend glatt, dann geht für $x \to x_0$ die Sekantensteigung $\dfrac{f(x) - f(x_0)}{x - x_0}$ in die Tangentensteigung über.

2) Sei $y(t)$ der Ort eines längs der y-Achse bewegten Massenpunktes zur Zeit t. Bewegt sich der Massenpunkt mit konstanter Geschwindigkeit, so beträgt diese $\dfrac{y(t) - y(t_0)}{t - t_0}$ unabhängig von $t \neq t_0$.
Beim freien Fall dagegen gilt $y(t) = y(t_0) - \frac{1}{2}\, g(t - t_0)^2$ mit $g = 9,81[\frac{m}{s^2}]$, so dass die Momentangeschwindigkeit von der Zeit abhängt, die Bewegung also ungleichförmig ist. Bei ungleichförmigen Bewegungen geht die Durchschnittsgeschwindigkeit während der Zeitspanne $[t_0, t]$, welche $\dfrac{y(t) - y(t_0)}{t - t_0}$ beträgt, für $t \to t_0$ in die Momentangeschwindigkeit zum Zeitpunkt t_0 über.

Definition

a) Sei $x \in \mathbb{R}$ und $\varepsilon > 0$. Dann heißt das offene Intervall $]x - \varepsilon,\ x + \varepsilon[$ eine $\varepsilon-$Umgebung von x.

b) Sei $x \in \mathbb{R}$ und $\varepsilon > 0$. Dann heißt $]x - \varepsilon,\ x]$ eine linksseitige $\varepsilon-$Umgebung von x und $[x,\ x + \varepsilon[$ eine rechtsseitige $\varepsilon-$Umgebung von x.

c) Sei $x \in D \subseteq \mathbb{R}$. x heißt innerer Punkt von D, falls x eine $\varepsilon-$Umgebung besitzt, die in D enthalten ist, d.h. $\exists\, \varepsilon > 0 :]x - \varepsilon,\ x + \varepsilon[\subseteq D$.

Definition *Sei $D \subseteq \mathbb{R}$ und $f : D \to \mathbb{R}$ eine Funktion.*

a) Sei x_0 ein innerer Punkt von D. f heißt differenzierbar in x_0, falls

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

existiert. $f'(x_0)$ heißt dann die Ableitung von f in x_0. Statt f' kann man auch das Symbol $\dfrac{df}{dx}$ schreiben.

b) f heißt differenzierbar in $M \subseteq D$, falls f in jedem $x \in M$ differenzierbar ist. Die Funktion $f' : M \to \mathbb{R},\ x \mapsto f'(x)$ heißt dann Ableitung von f.

c) *Sei $x_0 \in D$ ein Punkt, der eine linksseitige $\varepsilon-$Umgebung, die in D enthalten ist, besitzt. f heißt linksseitig differenzierbar in x_0, falls*

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

existiert.

d) *Sei $x_0 \in D$ ein Punkt, der eine rechtsseitige $\varepsilon-$Umgebung, die in D enthalten ist, besitzt. f heißt rechtsseitig differenzierbar in x_0, falls*

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

existiert.

Bemerkung

$\dfrac{f(x) - f(x_0)}{x - x_0}$ heißt auch ein Differenzenquotient von f in x_0. Er beschreibt die Steigung der Sekante durch $(x_0, f(x_0))$ und $(x, f(x))$.

$\lim\limits_{x \to x_0} \dfrac{f(x) - f(x_0)}{x - x_0}$ heißt auch der Differentialquotient von f in x_0. Wenn er existiert, dann beschreibt er die Steigung der Tangente an der Graphen von f in $(x_0, f(x_0))$. Die Gleichung der Tagente ist dann

$$y = f(x_0) + f'(x_0)(x - x_0).$$

Beispiele

1) Sei $c \in \mathbb{R}$ und $f : \mathbb{R} \to \mathbb{R}$ mit $f(x) = c$. Dann ist f differenzierbar in $\mathbb{R}$ mit $f'(x) = 0$ für alle $x \in \mathbb{R}$.

2) Sei $f : \mathbb{R} \to \mathbb{R}$ mit $f(x) = x$. Dann ist f differerenzierbar in $\mathbb{R}$ mit $f'(x) = 1$ für alle $x \in \mathbb{R}$.

3) Sei $f : \mathbb{R} \to \mathbb{R}_0^+$ mit $f(x) = |x|$. Dann ist f differenzierbar in $\mathbb{R} \setminus \{0\}$ mit $f'(x) = -1$ für alle $x < 0$ und $f'(x) = 1$ für alle $x > 0$. f ist in 0 linksseitig differenzierbar mit $f'(0^-) = -1$ und rechtsseitig differenzierbar mit $f'(0^+) = 1$, aber nicht differenzierbar wegen $f'(0^-) \neq f'(0^+)$.

4) $f : \mathbb{R} \to \mathbb{R}_0^+$ mit $f(x) = x^2$ ist differenzierbar in $\mathbb{R}$ mit $f'(x) = 2x$ für alle $x \in \mathbb{R}$, denn es gilt

$$\forall\, x \in \mathbb{R} : \lim_{y \to x} \frac{f(y) - f(x)}{y - x} = \lim_{y \to x} \frac{y^2 - x^2}{y - x} = \lim_{y \to x} y + x = 2x$$

5) $f : \mathbb{R} \to \mathbb{R}^+$ mit $f(x) = e^x$ ist differenzierbar in $\mathbb{R}$ mit $f' = f$.

Denn aufgrund von $e^x \geq 1 + x$ für alle $x \in \mathbb{R}$ und $|e^x - 1| \leq \frac{|x|}{1-|x|}$ für alle $x \in \mathbb{R}$ mit $|x| < 1$ (Satz 13.13 (vii), (ix)) folgt

$$e^{-|x|} \leq \left| \frac{e^x - 1}{x} \right| \leq \frac{1}{1 - |x|}$$

für alle $x \in \mathbb{R}$ mit $|x| < 1$.

$$\Rightarrow f'(0) = \lim_{x \to 0} \frac{e^x - e^0}{x} = \lim_{x \to 0} \frac{e^x - 1}{x} = 1 = e^0.$$

Wegen $e^{x+h} = e^x e^h$ folgt daraus

$$\forall\, x \in \mathbb{R} \,:\, f'(x) = \lim_{h \to 0} \frac{e^{x+h} - e^x}{h} = e^x \lim_{h \to 0} \frac{e^h - 1}{h} = e^x.$$

6) $\dfrac{d}{dx} \sin x = \cos x$ und $\dfrac{d}{dx} \cos x = -\sin x$ für alle $x \in \mathbb{R}$.

Denn aufgrund von $|\sin x| \leq |x| \leq \frac{1}{\cos x} |\sin x|$ und $1 - \frac{1}{2}x^2 \leq \cos x \leq 1$ für $|x| < \frac{\pi}{2}$ (siehe Kapitel 5.3) folgt

$$\frac{1}{1 - \frac{x^2}{2}} \leq \frac{\sin x}{x} \leq 1$$

für alle $|x| < 1$ und

$$0 \leq \left| \frac{1 - \cos x}{x} \right| \leq \frac{1}{2}|x|$$

für alle $|x| < 1$. Daher folgt

$$\frac{d}{dx} \sin x|_{x=0} = \lim_{x \to 0} \frac{\sin x - \sin 0}{x} = \lim_{x \to 0} \frac{\sin x}{x} = 1 = \cos 0.$$

und

$$\frac{d}{dx} \cos x|_{x=0} = \lim_{x \to 0} \frac{\cos x - 1}{x} = 0 = \sin 0.$$

Aufgrund der Additionstheoreme ergibt sich daraus für alle $x \in \mathbb{R}$:

$$\begin{aligned}
\frac{d}{dx} \sin x &= \lim_{h \to 0} \frac{\sin(x+h) - \sin x}{h} \\
&= \lim_{h \to 0} \frac{\sin x \cos h + \cos x \sin h - \sin x}{h} \\
&= \sin x \lim_{h \to 0} \frac{\cos h - 1}{h} + \cos x \lim_{h \to 0} \frac{\sin h}{h} \\
&= \cos x
\end{aligned}$$

und auf ähnliche Weise erhält man

$$\frac{d}{dx}\cos x = -\sin x$$

für alle $x \in \mathbb{R}$.

Satz 16.1 *Ist f differenzierbar in x_0, dann ist f auch stetig in x_0.*

Beweis

$$\lim_{x\to x_0}(f(x) - f(x_0)) = \lim_{x\to x_0}\frac{f(x)-f(x_0)}{x-x_0}\lim_{x\to x_0}(x-x_0) = f'(x_0)\cdot 0 = 0$$

□

Bemerkung
Aus der Stetigkeit einer Funktion f in x_0 folgt aber im Allgemeinen nicht die Differenzierbarkeit von f in x_0, denn $x \mapsto |x|$ ist stetig in $\mathbb{R}$, aber nur in $\mathbb{R}\setminus\{0\}$ differenzierbar.

Satz 16.2 *Sei $D \subseteq \mathbb{R}$, $f : D \to \mathbb{R}$ eine Funktion und $x_0 \in D$. Dann ist äquivalent:*

(i) *f ist differenzierbar in x_0.*

(ii) $\exists c \in \mathbb{R}\,\forall x \in D:\ f(x) = f(x_0) + c(x-x_0) + R(x,x_0)$ *mit* $\displaystyle\lim_{x\to x_0}\frac{R(x,x_0)}{x-x_0} = 0$

Sind (i) bzw. (ii) erfüllt, dann ist $c = f'(x_0)$.

Beweis
$(i) \Rightarrow (ii)$:
Sei f differenzierbar in x_0 und $R(x,x_0) := f(x) - f(x_0) - f'(x_0)(x-x_0)$.

$$\Rightarrow \lim_{x\to x_0}\frac{R(x,x_0)}{x-x_0} = \lim_{x\to x_0}\left(\frac{f(x)-f(x_0)}{x-x_0} - f'(x_0)\right) = 0$$

$(ii) \Rightarrow (i)$:
Aus (ii) folgt

$$\lim_{x\to x_0}\frac{f(x)-f(x_0)}{x-x_0} = \lim_{x\to x_0}\left(c + \frac{R(x,x_0)}{x-x_0}\right) = c$$

□

Satz 16.3 *Sind f und g differenzierbar in x_0 und $\alpha, \beta \in \mathbb{R}$, dann ist auch $\alpha f + \beta g$ differenzierbar in x_0 und es*

$$(\alpha f + \beta g)'(x_0) = \alpha f'(x_0) + \beta g'(x_0)$$

Beweis
Folgt aus der Definition der Differenzierbarkeit und den Rechnenregeln für Grenzwerte. □

Satz 16.4 (Produktregel) *Sind f und g in x_0 differenzierbar, dann auch fg mit*

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

Beweis

$$\lim_{x\to x_0} \frac{f(x)g(x) - f(x_0)g(x_0)}{x - x_0} = \lim_{x\to x_0} \left(\frac{f(x) - f(x_0)}{x - x_0} g(x) + f(x) \frac{g(x) - g(x_0)}{x - x_0} \right)$$
$$= \left(\lim_{x\to x_0} \frac{f(x) - f(x_0)}{x - x_0} \right) \left(\lim_{x\to x_0} g(x) \right) + \left(\lim_{x\to x_0} f(x) \right) \left(\lim_{x\to x_0} \frac{g(x) - g(x_0)}{x - x_0} \right)$$
$$\stackrel{16.1}{=} f'(x_0)g(x_0) + f(x_0)g'(x_0)$$ □

Beispiel
Sei $p(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0$. Dann gilt:

$$p'(x) = n a_n x^{n-1} + (n-1)a_{n-1}x^{n-2} + \ldots + a_1$$

Zeige zunächst durch vollständige Induktion:

$$\forall\, n \in \mathbb{N} \,:\, \frac{d}{dx} x^n = n x^{n-1} \qquad (*)$$

$n = 1 : \dfrac{d}{dx} x = 1 = 1 \cdot x^0$
$n \to n+1 :$
$$\frac{d}{dx} x^{n+1} = \frac{d}{dx}(x^n x) \stackrel{16.4}{=} \left(\frac{d}{dx} x^n \right) x + x^n \left(\frac{d}{dx} x \right) \stackrel{IV}{=} n x^{n-1} x + x^n = (n+1)x^n$$

Also gilt (*).

Wegen (*) und Satz 16.3 folgt die Aussage über p'.

Satz 16.5 (Quotientenregel) *Sind f, g differenzierbar in $x_0$und ist $g(x_0) \neq 0$, dann ist $\frac{f}{g}$ differenzierbar in x_0 mit*

$$\left(\frac{f}{g} \right)' (x_0) = \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{(g(x_0))^2}$$

Insbesondere gilt

$$\left(\frac{1}{g} \right)' (x_0) = -\frac{g'(x_0)}{(g(x_0))^2}$$

Beweis
Sei g differenzierbar in x_0 und $g(x_0) \neq 0$. Nach Satz 16.1 ist g stetig in x_0 und wegen Korollar 15.3 $\exists \delta > 0 \forall\, x \in]x_0 - \delta,\ x_0 + \delta[:\ g(x) \neq 0$.
Sei $h = \frac{1}{g}|_{]x_0-\delta, x_0+\delta[}$. Dann gilt:

$$\lim_{x\to x_0} \frac{h(x) - h(x_0)}{x - x_0} = \lim_{x\to x_0} \frac{1}{x - x_0} \left(\frac{1}{g(x)} - \frac{1}{g(x_0)} \right)$$

$$= \lim_{x\to x_0} \left(\frac{-(g(x) - g(x_0))}{x - x_0} \cdot \frac{1}{g(x)g(x_0)} \right) = \frac{-g'(x_0)}{(g(x_0))^2}$$

Mit Hilfe der Produktregel folgt daraus die Behauptung für $\left(\frac{f}{g}\right)'$. □

Beispiele

1) $\forall n \in \mathbb{N} :\ \frac{d}{dx}\frac{1}{x^n} = \frac{d}{dx}x^{-n} = -nx^{-n-1} = \frac{-n}{x^{n+1}}$ in $\mathbb{R} \setminus \{0\}$.
   Nach dem Quotientenregel gilt

$$\frac{d}{dx}\frac{1}{x} = -\frac{1}{x^2}$$

   für $x \neq 0$. Die restliche Behauptung ergibt sich durch vollständige Induktion.

2) $\frac{d}{dx}\tan x = \frac{1}{\cos^2 x}$ in $\mathbb{R} \setminus \{(\frac{1}{2} + k)\pi : k \in \mathbb{Z}\}$. Denn es gilt:
   $\frac{d}{dx}\tan x = \frac{d}{dx}\frac{\sin x}{\cos x} \overset{16.5}{=} \frac{\cos^2 x + \sin^2 x}{\cos^2 x} = \frac{1}{\cos^2 x}$

Satz 16.6 (Kettenregel) *Ist g differenzierbar in x_0 und f differenzierbar in $g(x_0)$, dann ist $f \circ g$ differenzierbar in x_0 mit*

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

Beweis
Aufgrund der Annahmen in Satz 16.6 ist g nach Satz 16.1 stetig in x_0. Also gilt

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

Sei D der Definitionsbereich von g. Für jede Folge $(x_n)_{n\in\mathbb{N}}$ mit $x_n \in D \setminus \{x_0\}$ für alle $n \in \mathbb{N}$, $\lim\limits_{n\to\infty} x_n = x_0$ und $g(x_n) \neq g(x_0)$ für alle $n \in \mathbb{N}$ ergibt sich

$$\lim_{n\to\infty} \frac{f(g(x_n)) - f(g(x_0))}{x_n - x_0} = \lim_{n\to\infty} \left(\frac{f(g(x_n)) - f(g(x_0))}{g(x_n) - g(x_0)} \cdot \frac{g(x_n) - g(x_0)}{x_n - x_0} \right)$$

$$= \left(\lim_{n\to\infty} \frac{f(g(x_n)) - f(g(x_0))}{g(x_n) - g(x_0)} \right) \left(\lim_{n\to\infty} \frac{g(x_n) - g(x_0)}{x_n - x_0} \right) \overset{(*)}{=} f'\left(g(x_0)\right) g'(x_0).$$

Existiert eine Folge $(x_n)_{n\in\mathbb{N}}$ mit $x_n \in D \setminus \{x_0\}$ für alle $n \in \mathbb{N}$, $\lim_{n\to\infty} x_n = x_0$ und $g(x_n) = g(x_0)$ für alle $n \in \mathbb{N}$, dann gilt $g'(x_0) = 0$ sowie $f \circ g(x_n) = f \circ g(x_0)$ für alle $n \in \mathbb{N}$. Daher folgt

$$\lim_{n\to\infty} \frac{f(g(x_n)) - f(g(x_0))}{x_n - x_0} = 0 = f'(g(x_0)) \cdot 0 = f'(g(x_0))g'(x_0).$$

Unter Ausnutzung dieser beiden Fälle erhalten wir die Behauptung. □

Beispiele

1) $\dfrac{d}{dx} \sin\left(e^{2x}\right) = 2e^{2x} \cos\left(e^{2x}\right)$

2) Sei $n \in \mathbb{N}_0$ und $f : \mathbb{R} \to \mathbb{R}$ mit

$$f(x) = \begin{cases} x^n \sin\left(\frac{1}{x}\right), & x \neq 0, \\ 0, & x = 0. \end{cases}$$

Dann ist f in $\mathbb{R} \setminus \{0\}$ differenzierbar mit $f'(x) = nx^{n-1} \sin\left(\frac{1}{x}\right) - x^{n-2} \cos\left(\frac{1}{x}\right)$. Ist $n > 1$, dann ist f auch in 0 differenzierbar mit $f'(x) = 0$, denn wegen $n - 1 > 0$ gilt

$$\left| \frac{x^n \sin\left(\frac{1}{x}\right)}{x} \right| = |x^{n-1} \sin\left(\tfrac{1}{x}\right)| \le |x|^{n-1} \overset{x\to 0}{\to} 0.$$

Satz 16.7 (Umkehrsatz für differenzierbare Funktionen) *Sei $I \subseteq \mathbb{R}$ ein offenes Intervall, $f : I \to \mathbb{R}$ differenzierbar in I mit $f'(x) \neq 0$ für alle $x \in I$ und f entweder streng monoton wachsend oder streng monoton fallend. Dann ist $f : I \to f(I)$ bijektiv und die Umkehrfunktion $f^{-1} : f(I) \to I$ ist differenzierbar in $f(I)$ mit*

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

für alle $y \in f(I)$.

Beweis
Aufgrund des Umkehrsatzes für stetige Funktionen (Satz 15.7 und anschließende Bemerkungen), ist $f : I \to f(I)$ bijektiv, $f(I)$ ein offenes Intervall und $f^{-1} : f(I) \to I$ stetig. Sei $y_0 \in f(I)$ und $(y_n)_{n\in\mathbb{N}}$ eine Folge in $f(I)$ mit $\lim_{n\to\infty} y_n = y_0$. Wegen der Stetigkeit von f^{-1} folgt $\lim_{n\to\infty} f^{-1}(y_n) = f^{-1}(y_0)$ und daher

$$\lim_{n\to\infty} \frac{f^{-1}(y_n) - f^{-1}(y_0)}{y - y_0} = \lim_{n\to\infty} \frac{f^{-1}(y_n) - f^{-1}(y_0)}{f\left(f^{-1}(y_n)\right) - f\left(f^{-1}(y_0)\right)} = \frac{1}{f'\left(f^{-1}(y_0)\right)}. \qquad \square$$

Bemerkung
Da der Graph von f^{-1} aus dem Graphen von f durch Achsenspiegelung an der 1. Winkelhalbierenden $y = x$ entsteht, könnte man die Formel für die Ableitung der Umkehrfunktion mit Hilfe geometrischer Argumente unter Verwendung der Eigenschaften von linearen Abbildungen beweisen. Der oben vorgestellte, rein analytische Beweis ist jedoch weniger aufwendig durchzuführen.

Beispiele

1) Für $x > 0$ und $n \in \mathbb{N}$ ist $\dfrac{d}{dx}x^{\frac{1}{n}} = \dfrac{1}{n}\,x^{\frac{1}{n}-1}$,

   also insbesondere $\dfrac{d}{dx}\sqrt{x} = \dfrac{1}{2\sqrt{x}}$.

   Denn es gilt nach Satz 16.7:
   $$\frac{d}{dx}\,x^{\frac{1}{n}} = \frac{1}{n\left(x^{\frac{1}{n}}\right)^{n-1}} = \frac{1}{n}\,x^{\frac{1}{n}-1} \quad \text{für } x > 0.$$
   Für ungerade n gilt die obige Formel auch für $x < 0$.

   In $x = 0$ ist $x \mapsto x^{\frac{1}{n}}$ für $n > 1$ nicht differenzierbar wegen
   $$\left|\frac{x^{\frac{1}{n}} - 0}{x - 0}\right| = |x^{\frac{1}{n}-1}| \to \infty \text{ für } x \to 0.$$

2) Für $x \in]-r, r[$ ist $\dfrac{d}{dx}\sqrt{r^2 - x^2} = -\dfrac{x}{\sqrt{r^2 - x^2}}$,

   was aus 1) und der Kettenregel folgt.

3) Für $x > 0$ ist $\dfrac{d}{dx}\ln x = \dfrac{1}{e^{\ln x}} = \dfrac{1}{x}$ für $x > 0$.

4) Für $r \in \mathbb{R}$ und $x > 0$ ist $\dfrac{d}{dx}x^r = rx^{r-1}$,

   was aus $x^r = e^{rlnx}$, der Kettenregel und 3) folgt.

5) Für $a > 0$ und $x \in \mathbb{R}$ ist $\dfrac{d}{dx}a^x = a^x \ln a$,

   was aus $a^x = e^{x \ln a}$ und der Kettenregel folgt.

6) Für $x \in]-1, 1[$ ist $\dfrac{d}{dx}\arcsin x = \dfrac{1}{\sqrt{1 - x^2}}$,

   denn es gilt für $x \in]-1, 1[$:
   $$\frac{d}{dx}\arcsin x = \frac{1}{\cos(\arcsin x)} = \frac{1}{\sqrt{1 - \sin^2(\arcsin x)}} = \frac{1}{\sqrt{1 - x^2}}$$
   Durch analoge Argumentation ergibt sich:

Für $x \in]-1,1[$ gilt $\dfrac{d}{dx} \arccos x = -\dfrac{1}{\sqrt{1-x^2}}$.

7) Für $x \in \mathbb{R}$ ist $\dfrac{d}{dx} \arctan x = \dfrac{1}{x^2+1}$,

denn es gilt für $x \in \mathbb{R}$:

$$\begin{aligned} \frac{d}{dx} \arctan x &= \frac{1}{\frac{1}{\cos^2(\arctan x)}} = \frac{1}{\frac{\sin^2(\arctan x)+\cos^2(\arctan x)}{\cos^2(\arctan x)}} = \frac{1}{\tan^2(\arctan x)+1} \\ &= \frac{1}{x^2+1} \end{aligned}$$

Definition *Sei $D \subseteq \mathbb{R}$, $f : D \to \mathbb{R}$ eine Funktion und $k \in \mathbb{N}$.*

a) Sei x_0 ein innerer Punkt von D. f heißt k-mal differenzierbar in x_0, falls es eine $\varepsilon-$Umgebung von x_0 gibt, so dass $f^{(0)} := f$, $f^{(1)} := f'$, $f^{(2)} := (f')', \ldots, f^{(k-1)} := \left(f^{(k-2)}\right)'$ in dieser $\varepsilon-$Umgebung existieren und $f^{(k-1)}$ in x_0 differenzierbar ist.

b) f heißt k-mal differenzierbar in $M \subseteq D$, falls f in jedem $x \in M$ k-mal differenzierbar ist. Die Funktion $f^{(m)}$, $m = 0, 1, \ldots, k$, heißt dann die m-te Ableitung von f. Weitere Schreibweisen sind:
f'' statt $f^{(2)}$ und f''' statt $f^{(3)}$,
$\dfrac{d^m}{dx^m} f$ oder $\dfrac{d^m f}{dx^m}$ statt $f^{(m)}$.

c) f heißt k-mal stetig differenzierbar in $M \subseteq D$, falls f in M differenzierbar ist und $f^{(k)}$ stetig in M ist. Bezeichnung:
$C^k(M) := \{f : M \to \mathbb{R} : f$ ist k-mal stetig differenzierbar in $M\}$

d) f heißt $\infty-$ oft differenzierbar in $M \subseteq D$, falls in jedem $x \in M$ für alle $k \in \mathbb{N}$ die k-te Ableitung $f^{(k)}$ existiert. Bezeichnung:
$C^\infty(M) := \{f : M \to \mathbb{R} : f$ ist $\infty-$oft differenzierbar in $M\}$.

Bemerkung
$C^k(M)$ und $C^\infty(M)$ sind $\mathbb{R}-$Vektorräume und $\frac{d}{dx} : f \mapsto \frac{d}{dx} f$ ist eine lineare Abbildung von $C^k(M)$ nach $C^{k-1}(M)$ für $k \geq 1$. Dies ist eine Konsequenz von Satz 16.3.

Beispiele

1) Polynome, die Exponentialfunktion, der Sinus und Kosinus sind $\infty-$oft differenzierbar auf $\mathbb{R}$.

2) ln ist $\infty-$oft differenzierbar auf $\mathbb{R}^+$

3) Sei $f : \mathbb{R} \to \mathbb{R}_0^+$, $x \mapsto |x|^3$. Dann ist $f \in C^2(\mathbb{R})$, aber $f \notin C^3(\mathbb{R})$.

4) Sei $f : \mathbb{R} \to \mathbb{R}$ mit $f(x) = \begin{cases} x^2 \sin\left(\frac{1}{x}\right), & x \neq 0, \\ 0\,, & x = 0. \end{cases}$

   Dann ist f zwar differenzierbar in $\mathbb{R}$, aber nicht stetig differenzierbar in $\mathbb{R}$, denn es ist

   $f'(x) = \begin{cases} 2x \sin\left(\frac{1}{x}\right) - \cos\left(\frac{1}{x}\right), & x \neq 0, \\ 0\,, & x = 0, \end{cases}$

   und somit ist f' nicht stetig in 0.

Satz 16.8 (Leibniz-Regel) *Sind f, g n-mal differenzierbar in x_0, dann auch fg mit*

$$(f \cdot g)^{(n)}(x_0) = \sum_{k=0}^{n} \binom{n}{k} f^{(k)} g^{(n-k)}(x_0)$$

Beweis
Folgt durch vollständige Induktion unter Verwendung der Produktregel und des Binomischen Satzes (Satz 4.2 d)). □

Definition *Eine Funktion $f : \mathbb{R} \supseteq D \to \mathbb{R}$ hat in $x_0 \in D$ ein lokales Maximum, falls gilt:*

$$\exists \varepsilon > 0 \, \forall x \in D : |x - x_0| < \varepsilon \Rightarrow f(x) \leq f(x_0)$$

f hat in x_0 ein lokales Minimum, falls gilt:

$$\exists \varepsilon > 0 \, \forall x \in D : |x - x_0| < \varepsilon \Rightarrow f(x) \geq f(x_0).$$

Ein lokales Maximum bzw. ein lokales Minimum heißt auch lokales Extremum.
Ein lokales Extremum heißt strikt, falls $f(x) = f(x_0)$ nur für $x = x_0$ gilt.
Gilt $\forall x \in D: \; f(x) \leq f(x_0)$ bzw. $\forall x \in D: \; f(x) \geq f(x_0)$, dann hat f in x_0 ein globales (absolutes) Maximum bzw. ein globales Minimum in D.

Satz 16.9 (Notwendige Bedingung für lokale Extrema) *Sei $f : \mathbb{R} \supseteq D \to \mathbb{R}$ eine Funktion, x_0 ein innerer Punkt von D und f differenzierbar in x_0. Dann gilt:*

$$f \text{ hat in } x_0 \text{ ein lokales Extremum} \Rightarrow f'(x_0) = 0$$

Beweis

f habe in x_0 ein lokales Maximum.

$\Rightarrow \exists \varepsilon > 0 \forall\, h \in]0, \varepsilon[: \; \dfrac{f(x_0 + h) - f(x_0)}{h} \leq 0 \, \wedge \, \dfrac{f(x_0 - h) - f(x_0)}{-h} \geq 0$

Da f in x_0 differenzierbar ist, folgt

$f'(x_0) = f'(x_0^+) = \lim\limits_{h \to 0^+} \dfrac{f(x_0 + h) - f(x_0)}{h} \leq 0 \; \wedge$

$f'(x_0) = f'(x_0^-) = \lim\limits_{h \to 0^-} \dfrac{f(x_0+h) - f(x_0)}{h} \geq 0$

$\Rightarrow f'(x_0) = 0$

Für ein lokales Minimum zeigt man die Implikation analog. □

Bemerkung
Das Verschwinden der 1.Ableitung ist aber keine hinreichende Bedingung für lokale Extrema, denn für $f : x \mapsto x^3$ ist $f'(0) = 0$, f hat aber kein lokales Extremum in 0.

Definition *Ist f differenzierbar in x_0 und $f'(x_0) = 0$, dann heißt x_0 stationärer oder kritischer Punkt von f.*

Satz 16.10 (Satz von Rolle) *Sei $f : [a,b] \to \mathbb{R}$ stetig auf $[a,b]$, differenzierbar in $]a,b[$ und $f(a) = f(b)$. Dann gilt:*

$$\exists \xi \in]a,b[: \; f'(\xi) = 0$$

Beweis
Da f stetig auf $[a,b]$ ist, nimmt f nach Satz 15.6 auf $[a,b]$ ein Maximum und ein Minimum an. Liegt das Maximum oder das Minimum in $x_0 \in]a,b[$, dann handelt es sich um ein lokales Extremum und nach Satz 16.9 folgt $f'(x_0) = 0$. Liegen das Maximum und das Minimum in a oder b, dann ist f konstant auf $[a,b]$ und somit $f'(x) = 0$ für alle $x \in]a,b[$. □

Satz 16.11 (Mittelwertsatz der Differentialrechnung) *Sei $f : [a,b] \to \mathbb{R}$ stetig auf $[a,b]$ und differenzierbar in $]a,b[$. Dann gilt:*

$$\exists \xi \in \,]a,b[: \quad f'(\xi) = \frac{f(b) - f(a)}{b - a}.$$

Beweis
Sei $g(x) := f(x) - \big(f(a) + \frac{f(b)-f(a)}{b-a}(x-a)\big)$. Dann erfüllt g die Voraussetzungen des Satzes von Rolle. Daher gilt:

$$\exists \xi \in \,]a,b[: 0 = g'(\xi) = f'(\xi) - \frac{f(b) - f(a)}{b - a}.$$

□

Korollar 16.12 *Sei $f : [a,b] \to \mathbb{R}$ stetig auf $[a,b]$ und differenzierbar in $]a,b[$. Ist $f'(x) = 0$ für alle $x \in \,]a,b[$, dann ist f auf $[a,b]$ konstant.*

Korollar 16.13 *Sei $f : [a,b] \to \mathbb{R}$ stetig auf $[a,b]$ und differenzierbar in $]a,b[$.*

a) Ist $f'(x) \geq 0$ (bzw. $f'(x) > 0$ / $f'(x) \leq 0$ / $f'(x) < 0$), so ist f auf $[a,b]$ monoton wachsend (bzw. streng monoton wachsend / monoton fallend / streng monoton fallend).

b) Ist f auf $[a,b]$ monoton wachsend / fallend, so ist $f'(x) \geq 0$ / $f'(x) \leq 0$ auf $]a,b[$.

Bemerkung
Ist f auf $[a,b]$ streng monoton wachsend / fallend, so muss nicht notwendigerweise $f'(x) > 0$ / $f'(x) < 0$ auf $]a,b[$ gelten, wie die Beispiele $f(x) = x^3$ / $f(x) = -x^3$ zeigen.

Satz 16.14 (Hinreichende Bedingungen für lokale Extrema) *Sei $f : [a,b] \to \mathbb{R}$ stetig auf $[a,b]$, differenzierbar in $]a,b[$ und $x_0 \in\,]a,b[$. Dann gilt:*

a) $\exists \varepsilon > 0\ \forall x \in\,]x_0 - \varepsilon, x_0[\, : f'(x) \geq 0 \ \wedge\ \forall x \in\,]x_0, x_0 + \varepsilon[\, : f'(x) \leq 0$
$\Rightarrow$ *f hat in x_0 ein lokales Maximum.*

Gilt in obiger Implikation sogar $f'(x) > 0$ statt $f'(x) \geq 0$ und $f'(x) < 0$ statt $f'(x) \leq 0$, dann ist das lokale Maximum strikt.

$\exists \varepsilon > 0\ \forall x \in\,]x_0 - \varepsilon, x_0[\, : f'(x) \leq 0 \ \wedge\ \forall x \in\,]x_0, x_0 + \varepsilon[\, : f'(x) \geq 0$
$\Rightarrow$ *f in x_0 ein lokales Minimum.*

Gilt in obiger Implikation sogar $f'(x) < 0$ statt $f'(x) \leq 0$ und $f'(x) > 0$ statt $f'(x) \geq 0$, dann ist das lokale Minimum strikt.

b) Sei f in x_0 zweimal differenzierbar. Dann gilt:
$f'(x_0) = 0 \ \wedge\ f''(x_0) < 0 \ \Rightarrow$ *f hat in x_0 ein striktes lokales Maximum.*
$f'(x_0) = 0 \ \wedge\ f''(x_0) > 0 \ \Rightarrow$ *f hat in x_0 ein striktes lokales Minimum.*

Beweis
$a)$ folgt aus Korollar 16.13 $a)$.

$b)$ Sei $f'(x_0) = 0 \wedge f''(x_0) = \lim\limits_{x \to x_0} \dfrac{f'(x) - 0}{x - x_0} < 0.$

$\Rightarrow \exists \varepsilon > 0\ \forall x \in\,]x_0 - \varepsilon, x_0[\, \cup\,]x_0, x_0 + \varepsilon[\, : \dfrac{f'(x_0)}{x - x_0} < 0$

$\Rightarrow \forall x \in\,]x_0 - \varepsilon, x_0[\, : f'(x) > 0 \ \wedge\ \forall x \in\,]x_0, x_0 + \varepsilon[\, : f'(x) < 0$

$\Rightarrow$ f hat in x_0 ein striktes lokales Maximum.

Die Implikation für das strikte lokale Minimum zeigt man analog. □

Bemerkungen

1) Satz 16.14 b) gibt hinreichende, aber keine notwendige Bedingungen für strikte lokale Extrema an, wie die Beispiele $f(x) = -x^4$ bzw. $f(x) = x^4$ in $x_0 = 0$ zeigen.

2) Aus $f'(x_0) = 0 \ \wedge\ f''(x_0) = 0$ lässt sich keine Aussage über lokale Extrema ableiten, wie die Beispiele $f(x) = x^3$, $g(x) = x^4$ und $h(x) = -x^4$ in $x_0 = 0$ zeigen.

Definition *Ist f differenzierbar in einer $\varepsilon-$Umgebung von x_0 und hat f' in x_0 ein striktes lokales Extremum, dann heißt x_0 Wendepunkt von f.*

Beispiel
$f(x) = xe^x$
$\Rightarrow\ f'(x) = (1+x)e^x, \quad f''(x) = (2+x)e^x, \quad f'''(x) = (3+x)e^x$
$f'(-1) = 0 \wedge f''(-1) > 0 \ \Rightarrow\ f$ hat in $x_0 = -1$ ein striktes lokales Minimum. Das Minimum ist wegen $f(-1) = -e^{-1}$, $\lim\limits_{x\to-\infty} f(x) = 0$ und $f(x) \to \infty$ für $x \to \infty$ sogar global.
$f''(-2) = 0 \wedge f'''(-2) > 0 \ \Rightarrow\ f$ hat in $x_0 = -2$ einen Wendepunkt.

Satz 16.15 (Verallgemeinerter Mittelwertsatz der Differentialrechnung) *Seien $f, g : [a,b] \to \mathbb{R}$ stetig auf $[a,b]$ und differenzierbar in $]a,b[$, wobei $g'(x) \neq 0$ für alle $x \in\ ]a,b[$ sei. Dann gilt:*

$$\exists \xi \in\ ]a,b[: \quad \frac{f(b)-f(a)}{g(b)-g(a)} = \frac{f'(\xi)}{g'(\xi)}.$$

Beweis
Wegen $g'(x) \neq 0$ für alle $x \in\ ]a,b[$ folgt $g(a) \neq g(b)$. Somit ist durch

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

eine Funktion h auf $[a,b]$ definiert, die die Voraussetzungen des Satzes von Rolle erfüllt. Daher gilt

$$\exists \xi \in\ ]a,b[: 0 = h'(\xi) = f'(\xi) - \frac{f(b)-f(a)}{g(b)-g(a)} g'(\xi).$$

$\square$

Satz 16.16 (Regel von de l'Hospital) *Seien $f, g :]a,b[\to \mathbb{R}$, wobei $-\infty \le a < b \le \infty$, differenzierbar, es gelte $\lim\limits_{x\to a^+} f(x) = \lim\limits_{x\to a^+} g(x) = 0$, $g(x) \neq 0 \wedge g'(x) \neq 0$ für alle $x \in\ ]a,b[$ und es existiere $\lim\limits_{x\to a^+} \frac{f'(x)}{g'(x)}$. Dann gilt:*

$$\lim_{x\to a^+} \frac{f(x)}{g(x)} = \lim_{x\to a^+} \frac{f'(x)}{g'(x)}.$$

Die analoge Aussage gilt, wenn man $\lim\limits_{x\to a^+}$ durch $\lim\limits_{x\to b^-}$ ersetzt.

Beweis
Setzt man $f(a) := 0$ und $g(a) := 0$, dann erfüllen f und g auf $[a,x]$ mit $a < x < b$ die Voraussetzungen des verallgemeinerten Mittelwertsatzes. Daher gilt:

$$\forall x \in\ ]a,b[\ \ \exists \xi(x) \in\ ]a,x[: \frac{f(x)}{g(x)} = \frac{f(x)-f(a)}{g(x)-g(a)} = \frac{f'(\xi(x))}{g'(\xi(x))}$$

$$\Rightarrow\ \lim_{x\to a^+} \frac{f(x)}{g(x)} = \lim_{x\to a^+} \frac{f'(x)}{g'(x)}.$$

$\square$

Beispiele

1)

$$\lim_{x\to 0}\frac{1-\cos x}{x^2}=\lim_{x\to 0}\frac{\sin x}{2x}=\lim_{x\to 0}\frac{\cos x}{2}=\frac{1}{2}.$$

2)

$$\begin{aligned}\lim_{x\to 0}\left(\frac{1}{\sin x}-\frac{1}{x}\right)&=\lim_{x\to 0}\frac{x-\sin x}{x\sin x}=\lim_{x\to 0}\frac{1-\cos x}{\sin x+x\cos x}\\&=\lim_{x\to 0}\frac{\sin x}{2\cos x-x\sin x}=0.\end{aligned}$$

3)

$$\begin{aligned}\lim_{x\to\infty}x\ln\left(\frac{x+1}{x-1}\right)&=\lim_{x\to\infty}\frac{\ln\left(\frac{x+1}{x-1}\right)}{\frac{1}{x}}=\lim_{x\to\infty}\frac{\ln(x+1)-\ln(x-1)}{\frac{1}{x}}\\&=\lim_{x\to\infty}\frac{\frac{1}{x+1}-\frac{1}{x-1}}{-\frac{1}{x^2}}=\lim_{x\to\infty}\frac{2x^2}{x^2-1}=2.\end{aligned}$$

Definition

a) $z_0\in\mathbb{C}$ *heißt innerer Punkt von* $M\subseteq\mathbb{C}$*, falls* z_0 *eine* ε*-Umgebung besitzt, die in* M *enthalten ist, d. h.*

$$\exists\varepsilon>0:\{z\in\mathbb{C}:|z-z_0|<\varepsilon\}\subseteq M.$$

b) Sei $D\subseteq\mathbb{R}$ *oder* $D\subseteq\mathbb{C}$*,* $f:D\to\mathbb{C}$ *eine Funktion und* z_0 *ein innerer Punkt von* D*.*
f *heißt komplex differenzierbar in* z_0*, falls*

$$f'(z_0):=\lim_{z\to z_0}\frac{f(z)-f(z_0)}{z-z_0}$$

existiert. f *heißt komplex differenzierbar in* D*, falls* f *in jedem* $z\in D$ *komplex differenzierbar ist.*

Bemerkung
Die Sätze 16.1–16.6 und 16.8 übertragen sich sinngemäß auf den komplexen Fall.

Beispiel
Sei $f:\mathbb{C}\to\mathbb{C}$ mit $f(z)=2iz^2+3z+1-5i\Rightarrow f'(z)=4iz+3$ für alle $z\in\mathbb{C}$.

17 Der Satz von Taylor, Taylorreihen und Potenzreihen

Satz 17.1 (Satz von Taylor) *Sei $I \subseteq \mathbb{R}$ ein offenes Intervall, $n \in \mathbb{N}_0$, $f \in C^{n+1}(I)$ und $x_0 \in I$. Dann gilt:*

$$\forall x \in I \; \exists \vartheta \in \,]0,1[\, : f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k + \frac{f^{(n+1)}(x_0 + \vartheta(x-x_0))}{(n+1)!}(x-x_0)^{n+1}.$$

Dabei heißt

$$T_n(f,x,x_0) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k$$

das n-te Taylorpolynom von f zum Entwicklungspunkt x_0 und

$$R_n(f,x,x_0) = f(x) - T_n(f,x,x_0)$$

das n-te Restglied. Die Darstellungsformel

$$R_n(f,x,x_0) = \frac{f^{(n+1)}(x_0 + \vartheta(x-x_0))}{(n+1)!}(x-x_0)^{n+1}$$

heißt Lagrangesche Form des Restglieds.

Beweis
Sei $F(x) := f(x) - T_n(f,x,x_0)$ und $G(x) := (x-x_0)^{n+1} \Rightarrow \forall k \in \{0,1,\ldots,n\} : F^{(k)}(x_0) = G^{(k)}(x_0) = 0$. Daher erhält man durch wiederholte Anwendung des verallgemeinerten Mittelwertsatzes (Satz 16.15):

$$\forall x \in I \; \forall k \in \{1,\ldots,n+1\} \; \exists \vartheta_k \in \,]0,1[\, : \frac{F(x)}{G(x)} = \frac{F^{(k)}(x_0 + \vartheta_k(x-x_0))}{G^{(k)}(x_0 + \vartheta_k(x-x_0))}.$$

Wegen $F^{(n+1)}(x_0 + \vartheta_{n+1}(x-x_0)) = f^{(n+1)}(x_0 + \vartheta_{n+1}(x-x_0))$ und
$G^{(n+1)}(x_0 + \vartheta_{n+1}(x-x_0)) = (n+1)!$ folgt

$$\begin{aligned} f(x) - T_n(f,x,x_0) &= G(x)\frac{F^{(n+1)}(x_0 + \vartheta_{n+1}(x-x_0))}{G^{(n+1)}(x_0 + \vartheta_{n+1}(x-x_0))} \\ &= \frac{f^{(n+1)}(x_0 + \vartheta_{n+1}(x-x_0))}{(n+1)!}(x-x_0)^{n+1}. \end{aligned}$$

□

Definition *Sei $D \subseteq \mathbb{R}$, x_0 ein Häufungspunkt von D und seien $f,g : D \to \mathbb{R}$ zwei Funktionen. Man schreibt:*

a) $f(x) = O(g(x))$ für $x \to x_0$ („f ist Groß O von g in x_0"), genau dann wenn gilt

$$\exists C, \delta > 0 \; \forall x \in D \cap \,]x_0 - \delta, x_0 + \delta[\, : \quad |f(x)| \le C|g(x)|.$$

b) $f(x) = o(g(x))$ für $x \to x_0$ („f ist Klein O von g in x_0”), genau dann wenn gilt

$$\lim_{x \to x_0} \frac{f(x)}{g(x)} = 0.$$

Korollar 17.2 *Unter den Voraussetzungen des Satzes von Taylor gilt:*

$$f(x) - T_n(f,x,x_0) = o((x-x_0)^n) \qquad \text{für } x \to x_0,$$
$$f(x) - T_n(f,x,x_0) = O((x-x_0)^{n+1}) \quad \text{für } x \to x_0.$$

Beispiele

1) $f(x) = \cos x$ ist C^∞-differenzierbar in $\mathbb{R}$, $f(0) = 1, f'(0) = -\sin 0 = 0, f''(0) = -\cos 0 = -1$. Daher gilt:

$$\begin{aligned} T_2(f,x,0) &= 1 - \frac{1}{2}x^2, \\ \cos x &= 1 - \frac{1}{2}x^2 + o(x^2) \quad \text{für } x \to 0 \\ \Rightarrow \lim_{x \to 0} \frac{1-\cos x}{x^2} &= \frac{1}{2}. \end{aligned}$$

2) $f(x) = \sin x$ ist C^∞-differenzierbar in $\mathbb{R}$, $f(0) = 0, f'(0) = \cos 0 = 1, f''(0) = -\sin 0 = 0$.

$$\Rightarrow \sin x = x + o(x^2) \quad \text{für } x \to 0$$
$$\Rightarrow \frac{1}{\sin x} - \frac{1}{x} = \frac{x - \sin x}{x \sin x} = \frac{o(x^2)}{x^2 + o(x^3)} \quad \text{für } x \to 0$$
$$\Rightarrow \lim_{x \to 0} \left(\frac{1}{\sin x} - \frac{1}{x} \right) = 0.$$

3) $f(x) = \ln x$ ist C^∞-differenzierbar in $\mathbb{R}^+$, $f(1) = 0, f'(1) = 1$.

$$\Rightarrow \ln(1+x) = x + O(x^2) \quad \text{für } x \to 0$$
$$\Rightarrow \ln\left(\frac{x+1}{x-1}\right) = \ln\left(1 + \frac{2}{x-1}\right) = \frac{2}{x-1} + O\left(\left(\frac{2}{x-1}\right)^2\right) \quad \text{für } x \to \infty$$
$$\Rightarrow \lim_{x \to \infty} x \ln\left(\frac{x+1}{x-1}\right) = \lim_{x \to \infty}\left(\frac{2x}{x-1} + O\left(x\left(\frac{2}{x-1}\right)^2\right)\right) = 2.$$

Anwendungen des Satzes von Taylor

1) Das Newton-Verfahren zur näherungsweisen Berechnung von Nullstellen von Funktionen

Sei $f :]a, b[\to \mathbb{R}$ differenzierbar und f besitze eine Nullstelle x_*, die in der Nähe von $x_0 \in]a, b[$ liege. Zur näherungsweisen Berechnung von x_* bestimmt man den Schnittpunkt der Tangente an den Graphen von f in $(x_0, f(x_0))$ mit der x-Achse, nimmt die x-Koordinate dieses Schnittpunktes als neuen Näherungswert x_1 für die Nullstelle x_* und wiederholt das Verfahren.
Man erhält:

$$\begin{aligned} f'(x_0) &= \frac{f(x_0)}{x_0 - x_1} \\ \Rightarrow x_1 &= x_0 - \frac{f(x_0)}{f'(x_0)}, \quad \text{falls } f'(x_0) \neq 0 \\ \Rightarrow x_{n+1} &= x_n - \frac{f(x_n)}{f'(x_n)}, \quad \text{falls } f'(x_n) \neq 0. \end{aligned}$$

Satz 17.3 *Sei $f \in C^3(]a, b[)$, $x_* \in]a, b[$ eine Nullstelle von f und $\exists \delta > 0\ \forall x \in]x_* - \delta, x_* + \delta[\ :\ f'(x) \neq 0$. Dann existiert eine ε-Umgebung von x_*, in der das Newton-Verfahren*

$$\begin{aligned} x_0 &\in]x_* - \varepsilon, x_* + \varepsilon[\\ x_{n+1} &:= x_n - \frac{f(x_n)}{f'(x_n)} \end{aligned}$$

quadratisch (von Ordnung 2) gegen x_* konvergiert, d. h.,

$$\lim_{n\to\infty} x_n = x_* \ \wedge\ \exists C > 0\ \forall n \in \mathbb{N} : |x_{n+1} - x_*| \leq C|x_n - x_*|^2.$$

Beweis
Sei $F :]x_* - \delta, x_* + \delta[\to \mathbb{R}$ mit $F(x) := x - \dfrac{f(x)}{f'(x)}$. Dann gilt $F(x_n) = x_{n+1}$ sowie

$$F(x) = x \iff f(x) = 0$$

und daher $F(x_*) = x_*$. Außerdem gilt

$$F'(x) = 1 - \frac{f'(x)}{f'(x)} + \frac{f(x)f''(x)}{(f'(x))^2} = \frac{f(x)f''(x)}{(f'(x))^2}$$

$$\Rightarrow F'(x_*) = 0.$$

Des Weiteren ist $F \in C^2(]x_* - \delta, x_* + \delta[)$, da $f \in C^3(]a, b[)$ ist. Nach dem Satz von Taylor folgt:

$$\forall x \in \left[x_* - \tfrac{\delta}{2}, x_* + \tfrac{\delta}{2}\right]\ \exists \xi(x) \in \left[x_* - \tfrac{\delta}{2}, x_* + \tfrac{\delta}{2}\right] : F(x) = x_* + \frac{1}{2}F''(\xi(x))(x - x_*)^2.$$

Sei $x_n \in \left[x_* - \frac{\delta}{2}, x_* + \frac{\delta}{2}\right]$. Dann gilt

$$\begin{aligned}
|x_{n+1} - x_*| &= |F(x_n) - x_*| \\
&= \frac{1}{2}|F''(\xi(x_n))||x_n - x_*|^2 \\
&\leq \frac{1}{2}\left(\max_{x\in\left[x_*-\frac{\delta}{2},x_*+\frac{\delta}{2}\right]} |F''(x)|\right)|x_n - x_*|^2 \\
&\leq C|x_n - x_*|^2
\end{aligned}$$

mit $C \in \mathbb{R}^+$, wobei die letzte Abschätzung aus dem Satz vom Maximum und Minimum (Satz 15.6) folgt. Per Induktion folgt:

$$\forall x_0 \in \left[x_* - \tfrac{\delta}{2}, x_* + \tfrac{\delta}{2}\right] \ \forall n \in \mathbb{N} : |x_n - x_*| \leq \left(C|x_0 - x_*|^2\right)^n .$$

Wähle nun $\varepsilon \in \left]0, \frac{\delta}{2}\right]$ so, dass $C|x - x_*|^2 < 1$ für alle $x \in [x_* - \varepsilon, x_* + \varepsilon]$ gilt. Dann folgt die quadratische Konvergenz des Newton-Verfahrens gegen x_* für alle Startwerte $x_0 \in \left]x_* - \varepsilon, x_* + \varepsilon\right[$. □

2) Weitere hinreichende Bedingungen für lokale Extrema

Satz 17.4 *Sei $I \subseteq \mathbb{R}$ ein offenes Intervall, $m \geq 2$, $f \in C^m(I)$, $x_0 \in I$, $f'(x_0) = \cdots = f^{(m-1)}(x_0) = 0$ und $f^{(m)}(x_0) \neq 0$. Dann gilt:*

a) Ist m gerade, dann gilt:
$f^{(m)}(x_0) > 0 \Rightarrow f$ hat in x_0 ein striktes lokales Minimum.
$f^{(m)}(x_0) < 0 \Rightarrow f$ hat in x_0 ein striktes lokales Maximum.

b) Ist m ungerade, dann hat f in x_0 kein lokales Extremum, sondern einen Wendepunkt mit waagerechter Tangente.

Beweis
Unter den Voraussetzungen des Satzes liefert der Satz von Taylor

$$f(x) = f(x_0) + \frac{1}{m!} f^{(m)}(x_0 + \vartheta(x)(x - x_0))(x - x_0)^m$$

mit $\vartheta(x) \in \left]0, 1\right[$, wobei $f^{(m)}$ auf einer hinreichend kleinen ε-Umgebung von x_0 konstantes Vorzeichen hat. Daraus folgen die Behauptungen. □

Beispiel
$x \mapsto x^n$ erfüllt für jedes $n \in \mathbb{N} \setminus \{1\}$ in $x_0 = 0$ die Voraussetzungen von Satz 17.4.

Definition *Sei $I \subseteq \mathbb{R}$ ein offenes Intervall und $f \in C^\infty(I)$. Dann heißt*

$$T(f, x, x_0) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k$$

für $x, x_0 \in I$ die Taylorreihe von f um x_0.

Satz 17.5 *Sei $I \subseteq \mathbb{R}$ ein offenes Intervall, $f \in C^\infty(I)$ und $x_0 \in I$. Dann konvergiert $T(f, x, x_0)$ genau dann gegen $f(x)$ wenn $\lim\limits_{n\to\infty} R_n(f, x, x_0) = 0$ gilt.*

Beispiele

1) Sei $f(x) = e^x$. Dann ist $f \in C^\infty(\mathbb{R})$ und $f^{(n)} = f$ für alle $n \in \mathbb{N}_0$. Nach dem Satz von Taylor folgt für alle $x \in \mathbb{R}$:

$$e^x = \sum_{k=0}^{n} \frac{x^k}{k!} + \frac{e^{\vartheta(x)x}}{(n+1)!} x^{n+1}$$

mit geeignetem $\vartheta(x) \in \,]0, 1[$.

Da die Exponentialfunktion streng monoton wächst, gilt:

$$\forall x < 0 : \ \left| e^{\vartheta(x)x} \frac{x^{n+1}}{(n+1)!} \right| < \frac{|x|^{n+1}}{(n+1)!}\,, \tag{1}$$

$$\forall x > 0 : \ 0 < e^{\vartheta(x)x} \frac{x^{n+1}}{(n+1)!} < e^x \frac{|x|^{n+1}}{(n+1)!}\,. \tag{2}$$

Da die Reihe $\sum\limits_{k=0}^{\infty} \frac{x^k}{k!}$ für alle $x \in \mathbb{R}$ absolut konvergiert, muss $\left(\frac{|x|^n}{n!}\right)_{n\in\mathbb{N}}$ für alle $x \in \mathbb{R}$ eine Nullfolge sein. Daher folgt

$$\forall x \in \mathbb{R} : \lim_{n\to\infty} R_n(f, x, 0) = 0$$
$$\overset{17.5}{\Rightarrow} \ \forall x \in \mathbb{R} : e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}.$$

Die Exponentialreihe (die Taylorreihe von e^x um 0) konvergiert sehr schnell gegen e^x, denn aus (1), (2) folgt zum Beispiel

$$\sup_{x\in[-1,1]} |R_n(f, x, 0)| \le \frac{e}{(n+1)!}$$
$$\Rightarrow \ \sup_{x\in[-1,1]} |R_{10}(f, x, 0)| \le 6{,}81 \cdot 10^{-8}.$$

2) Sei $f(x) = \cos(x)$ und $g(x) = \sin(x)$. Dann sind $f, g \in C^\infty(\mathbb{R})$ mit $f^{(2n)} = (-1)^n f$, $f^{(2n+1)} = (-1)^{n+1} g$, $g^{(2n)} = (-1)^n g$ und $g^{(2n+1)} = (-1)^n f$ für alle $n \in \mathbb{N}_0$. Nach dem Satz von Taylor folgt für alle $x \in \mathbb{R}$:

$$\cos x = \sum_{k=0}^{n} (-1)^k \frac{x^{2k}}{(2k)!} + (-1)^{n+1} \cos(\vartheta_c(x)x) \frac{x^{2n+2}}{(2n+2)!},$$
$$\sin x = \sum_{k=0}^{n} (-1)^k \frac{x^{2k+1}}{(2k+1)!} + (-1)^{n+1} \cos(\vartheta_s(x)x) \frac{x^{2n+3}}{(2n+3)!}$$

mit geeigneten $\vartheta_c(x), \vartheta_s(x) \in \,]0,1[$. Da $\forall y \in \mathbb{R} : |\cos(y)| \leq 1$ und $\frac{y^n}{n!}$ eine Nullfolge ist, folgt

$$\forall x \in \mathbb{R} : \lim_{n\to\infty} R_{2n+1}(f, x, 0) = 0 \;\wedge\; \lim_{n\to\infty} R_{2n+2}(g, x, 0) = 0$$

$$\overset{17.5}{\Rightarrow} \forall x \in \mathbb{R} : \cos x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!} \qquad \text{(Kosinusreihe)}$$
$$\forall x \in \mathbb{R} : \sin x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!} \qquad \text{(Sinusreihe)}$$

Unter Verwendung der Sinus- und der Kosinusreihe kann man alle in Abschnitt 5.3 vorgestellten Eigenschaften des Sinus und des Kosinus allein mit Methoden der Analysis und ohne Rückgriff auf geometrische Argumente herleiten. Deshalb werden in vielen Büchern der Sinus und der Kosinus durch die Sinus- und die Kosinusreihe definiert.

3) Sei $f(x) = \dfrac{1}{1-x}$. Dann ist $f \in C^\infty(]-\infty, 1[)$ mit $f^{(n)}(x) = \dfrac{n!}{(1-x)^{n+1}}$ für alle $n \in \mathbb{N}_0$. Nach dem Satz von Taylor folgt für alle $x \in \,]-\infty, 1[$

$$\frac{1}{1-x} = \sum_{k=0}^{n} x^k + R_n(f, x, 0).$$

Aus Kapitel 14 wissen wir:

$$\frac{1}{1-x} = \sum_{k=0}^{\infty} x^k \iff |x| < 1,$$

also gilt: $\lim_{n\to\infty} R_n(f, x, 0) = 0 \iff |x| < 1$.

4) Sei

$$f(x) = \begin{cases} e^{-x^{-2}}, & x \neq 0 \\ 0, & x = 0 \end{cases}.$$

Dann ist $f \in C^\infty(\mathbb{R})$ mit $f^{(n)}(0) = 0$ für alle $n \in \mathbb{N}_0$. Nach dem Satz von Taylor folgt:

$$\begin{aligned} \forall n \in \mathbb{N}_0 \ \forall x \in \mathbb{R} : T_n(f,x,0) = 0 \\ \Rightarrow \forall x \in \mathbb{R} : T(f,x,0) = \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} x^k = 0 \end{aligned}$$

$\Rightarrow \sum\limits_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} x^k$ konvergiert zwar für alle $x \in \mathbb{R}$, aber für $x \neq 0$ nicht gegen $f(x)$.

Definition *Seien $x_0, x \in \mathbb{R}$ und $(a_n)_{n\in\mathbb{N}_0}$ eine reellwertige Folge. Dann heißt*

$$\sum_{n=0}^{\infty} a_n (x - x_0)^n$$

eine reelle Potenzreihe (um x_0 mit Koeffizienten a_n).

Definition

a) *Sei M eine Menge von Funktionen. Eine Abbildung von $\mathbb{N}$ bzw. $\mathbb{N}_0$ nach M, $n \mapsto f_n$ (kurz: $(f_n)_{n\in\mathbb{N}}$ bzw. $(f_n)_{n\in\mathbb{N}_0}$) heißt Funktionenfolge.*

b) *Eine Folge $(f_n)_{n\in\mathbb{N}_0}$ von Funktionen $f_n : \mathbb{R} \supseteq D \to \mathbb{R}$ heißt punktweise konvergent gegen $f : D \to \mathbb{R}$, falls*

$$\forall x \in D : \lim_{n\to\infty} f_n(x) = f(x).$$

c) *Eine Folge $(f_n)_{n\in\mathbb{N}_0}$ von Funktionen $f_n : \mathbb{R} \supseteq D \to \mathbb{R}$ heißt gleichmäßig konvergent (auf D) gegen $f : D \to \mathbb{R}$, falls*

$$\lim_{n\to\infty} \sup_{x\in D} |f_n(x) - f(x)| = 0.$$

Beispiele

1) Jede Taylorreihe ist eine Potenzreihe.

2) $f_n : \mathbb{R} \to \mathbb{R}$ mit $f_n(x) = \sum\limits_{k=0}^{n} \frac{x^k}{k!}$ für alle $n \in \mathbb{N}_0$ bildet eine Funktionenfolge. Sie ist punktweise konvergent gegen $f : \mathbb{R} \to \mathbb{R}$ mit $f(x) = e^x$. Sie ist gleichmäßig konvergent gegen f auf jedem beschränkten abgeschlossenen Intervall $[a, b]$ mit $a, b \in \mathbb{R}$, konvergiert aber nicht gleichmäßig auf $\mathbb{R}$ gegen f. Denn es gilt

$$\lim_{n\to\infty} \sup_{x\in[a,b]} \left| e^{\vartheta(x)x} \frac{x^{n+1}}{(n+1)!} \right| = 0,$$

aber $\left| e^{\vartheta(x)x} \frac{x^{n+1}}{(n+1)!} \right| \to \infty$ für $x \to \infty$.

Bemerkungen

1) Die Rechenregeln für konvergente Folgen übertragen sich auf punktweise bzw. gleichmäßig konvergente Funktionenfolgen.

2) Aus gleichmäßiger Konvergenz folgt punktweise Konvergenz, aus punktweiser Konvergenz im Allgemeinen aber nicht gleichmäßige Konvergenz.

Satz 17.6 (Kriterien für gleichmäßige Konvergenz) *Seien $f_n : \mathbb{R} \supseteq D \to \mathbb{R}$, $n \in \mathbb{N}_0$, Funktionen. Dann gilt*

a) (Cauchy-Kriterium)

$$\left[\forall \varepsilon > 0 \ \exists n_\varepsilon \in \mathbb{N} \ \forall m, n > n_\varepsilon : \sup_{x \in D} |f_m(x) - f_n(x)| < \varepsilon\right]$$
$$\Rightarrow \exists f : D \to \mathbb{R} : f_n \text{ konvergiert gleichmäßig auf } D \text{ gegen } f.$$

b) (Weierstraß-Kriterium)

$$\left[\left(\forall n \in \mathbb{N}_0 \ \exists c_n \in \mathbb{R} : \sup_{x \in D} |f_n(x)| \leq c_n\right) \ \wedge \ \sum_{n=0}^{\infty} c_n \text{ ist konvergent}\right]$$
$$\Rightarrow \exists f, g : D \to \mathbb{R} : \sum_{n=0}^{\infty} f_n \text{ konvergiert gleichmäßig auf } D \text{ gegen } f \ \wedge$$
$$\sum_{n=0}^{\infty} |f_n| \text{ konvergiert gleichmäßig auf } D \text{ gegen } g.$$

Beweis

a) $\forall \varepsilon > 0 \ \exists n_\varepsilon \in \mathbb{N} \ \forall m, n > n_\varepsilon : \sup_{x \in D} |f_m(x) - f_n(x)| < \varepsilon$

$\Rightarrow \forall x \in D$ ist $(f_n(x))_{n \in \mathbb{N}_0}$ eine Cauchy-Folge und somit konvergent.

Sei $\forall x \in D : f(x) := \lim_{n \to \infty} f_n(x)$. Dann gilt

$$\forall n > n_\varepsilon : \sup_{x \in D} |f_n(x) - f(x)| = \sup_{x \in D} \left(\lim_{m \to \infty} |f_n(x) - f_m(x)|\right) \leq \varepsilon$$

$\Rightarrow$ f_n konvergiert gleichmäßig auf D gegen f.

b) $\left(\forall n \in \mathbb{N}_0 \ \exists c_n \in \mathbb{R} : \sup_{x \in D} |f_n(x)| \leq c_n\right) \wedge \sum_{n=0}^{\infty} c_n$ ist konvergent

$$\begin{aligned}
\Rightarrow \forall \varepsilon > 0 \ \exists n_\varepsilon \in \mathbb{N} \ \forall m, n > n_\varepsilon : &\sup_{x \in D} \left|\sum_{k=0}^{m} f_k(x) - \sum_{k=0}^{n} f_k(x)\right| \\
&= \sup_{x \in D} \left|\sum_{k=n+1}^{m} f_k(x)\right| \\
&\leq \sum_{k=n+1}^{m} |f_k(x)| \ \leq \sum_{k=n+1}^{m} c_k < \varepsilon
\end{aligned}$$

$\overset{a)}{\Rightarrow} \sum\limits_{n=0}^{\infty} f_n$ konvergiert gleichmäßig auf D gegen f mit $f(x) = \lim\limits_{n\to\infty} \sum\limits_{k=0}^{n} f_k(x)$ und $\sum\limits_{n=0}^{\infty} |f_n|$ konvergiert gleichmäßig auf D gegen g mit $g(x) = \lim\limits_{n\to\infty} \sum\limits_{k=0}^{n} |f_k(x)|$. $\square$

Satz 17.7 (Konvergenzverhalten von Potenzreihen) *Sei $\sum\limits_{n=0}^{\infty} a_n(x-x_0)^n$ eine Potenzreihe. Dann gilt:*

a) Es gibt ein eindeutig bestimmtes R mit $R \in \mathbb{R}_0^+$ oder $R = \infty$, so dass gilt:

$$\sum_{n=0}^{\infty} a_n(x-x_0)^n \text{ ist } \begin{cases} \text{absolut konvergent,} & \text{falls } |x-x_0| < R, \\ \text{divergent,} & \text{falls } |x-x_0| > R. \end{cases}$$

R heißt der Konvergenzradius von $\sum\limits_{n=0}^{\infty} a_n(x-x_0)^n$.

b) Für den Konvergenzradius R gilt die Formel von Cauchy-Hadamard, nämlich

$$R = \begin{cases} \dfrac{1}{\limsup\limits_{n\to\infty} \sqrt[n]{|a_n|}}, & \text{falls } \limsup\limits_{n\to\infty} \sqrt[n]{|a_n|} \text{ existiert und} \neq 0 \text{ ist,} \\ \infty, & \text{falls } \limsup\limits_{n\to\infty} \sqrt[n]{|a_n|} = 0 \text{ ist,} \\ 0, & \text{falls } \limsup\limits_{n\to\infty} \sqrt[n]{|a_n|} \text{ nicht existiert.} \end{cases}$$

c) Falls $\lim\limits_{n\to\infty} \frac{1}{\sqrt[n]{|a_n|}}$, $\lim\limits_{n\to\infty} \frac{|a_n|}{|a_{n+1}|}$ oder $\limsup\limits_{n\to\infty} \frac{|a_n|}{|a_{n+1}|}$ existieren, dann gilt $R = \lim\limits_{n\to\infty} \frac{1}{\sqrt[n]{|a_n|}}$, bzw. $R = \lim\limits_{n\to\infty} \frac{|a_n|}{|a_{n+1}|}$, bzw. $R = \limsup\limits_{n\to\infty} \frac{|a_n|}{|a_{n+1}|}$. Falls $\limsup\limits_{n\to\infty} \frac{|a_n|}{|a_{n+1}|}$ nicht existiert, dann gilt $R = \infty$.

d) Falls $R > 0$ und $0 < \rho < R$ ist, dann konvergiert $\sum\limits_{n=0}^{\infty} a_n(x-x_0)^n$ gleichmäßig auf $\{x \in \mathbb{R} : |x-x_0| \leq \rho\}$.

Beweis
Wegen $\sqrt[n]{|a_n(x-x_0)^n|} = \sqrt[n]{|a_n|}\,|x-x_0|$ und $\dfrac{|a_{n+1}(x-x_0)^{n+1}|}{|a_n(x-x_0)^n|} = \dfrac{|a_{n+1}|}{|a_n|}|x-x_0|$ folgen $a)$, $b)$, $c)$ aus dem Wurzelkriterium und dem Quotientenkriterium.
Sei $R > 0$ und $0 < \rho < R$. Wegen $\sup\limits_{|x-x_0|\leq\rho} |a_n(x-x_0)^n| \leq |a_n|\rho^n$ und der Konvergenz von $\sum\limits_{n=0}^{\infty} |a_n|\rho^n$ folgt $d)$ aus dem Weierstraß-Kriterium. $\square$

Bemerkungen

1) Ist der Konvergenzradius R einer Potenzreihe $\sum\limits_{n=0}^{\infty} a_n(x - x_0)^n$ endlich, also $R \in \mathbb{R}_0^+$, und $x \in \mathbb{R}$ mit $|x - x_0| = R$, dann kann die Potenzreihe für dieses x konvergent oder divergent sein, was aus den folgenden Beispielen ersichtlich sein wird.

2) Auf $\{x \in \mathbb{R} : |x - x_0| < R\}$ konvergieren Potenzreihen im Allgemeinen nicht gleichmäßig, was wir bei der Exponentialreihe gesehen haben.

Beispiele

1) Die Exponentialreihe $\sum\limits_{n=0}^{\infty} \frac{1}{n!} x^n$ hat, wie wir bereits gesehen haben, den Konvergenzradius $R = \infty$. Dies ergibt sich auch aus $\dfrac{|a_n|}{|a_{n+1}|} = \dfrac{\frac{1}{n!}}{\frac{1}{(n+1)!}} = n + 1 \to \infty$ für $n \to \infty$.

2) Die Potenzreihe $\sum\limits_{n=0}^{\infty} n!\, x^n$ hat Konvergenzradius $R = 0$, denn es gilt

$$\lim_{n\to\infty} \frac{|a_n|}{|a_{n+1}|} = \lim_{n\to\infty} \frac{n!}{(n+1)!} = \lim_{n\to\infty} \frac{1}{n+1} = 0.$$

Für $x = 0$ konvergiert sie gegen 1.

3) Die geometrische Reihe $\sum\limits_{n=0}^{\infty} x^n$ hat, wie wir bereits wissen, den Konvergenzradius $R = 1$. Dies ergibt sich auch aus $\lim\limits_{n\to\infty} \dfrac{1}{\sqrt[n]{|a_n|}} = \lim\limits_{n\to\infty} \dfrac{|a_n|}{|a_{n+1}|} = \lim\limits_{n\to\infty} 1 = 1$. Für $|x| = 1$ divergiert sie.

4) Die Potenzreihe $\sum\limits_{n=0}^{\infty} 2^n (x-1)^n$ hat den Konvergenzradius $R = \frac{1}{2}$. Dies ergibt sich durch Zurückführung auf die geometrische Reihe ($\sum\limits_{n=0}^{\infty} 2^n (x-1)^n = \sum\limits_{n=0}^{\infty} q^n$ mit $q = 2(x-1)$) oder aus $\lim\limits_{n\to\infty} \dfrac{1}{\sqrt[n]{|a_n|}} = \lim\limits_{n\to\infty} \dfrac{|a_n|}{|a_{n+1}|} = \lim\limits_{n\to\infty} \frac{1}{2} = \frac{1}{2}$. Für $|x - 1| = \frac{1}{2}$ divergiert sie.

5) Die Potenzreihe $\sum\limits_{n=0}^{\infty} n x^n$ hat Konvergenzradius $R = 1$, denn es gilt $\lim\limits_{n\to\infty} \dfrac{|a_n|}{|a_{n+1}|} = \lim\limits_{n\to\infty} \dfrac{n}{n+1} = 1$ oder $\lim\limits_{n\to\infty} \dfrac{1}{\sqrt[n]{|a_n|}} = \lim\limits_{n\to\infty} \dfrac{1}{\sqrt[n]{n}} = 1$. Für $|x| = 1$ divergiert sie.

6) Die Potenzreihe $\sum\limits_{n=1}^{\infty} \frac{1}{n^2} x^n$ hat Konvergenzradius $R = 1$, denn es gilt $\lim\limits_{n\to\infty} \frac{|a_n|}{|a_{n+1}|}$ $= \lim\limits_{n\to\infty} \frac{n^2}{n^2+2n+1} = 1$ oder $\lim\limits_{n\to\infty} \frac{1}{\sqrt[n]{|a_n|}} = \lim\limits_{n\to\infty} \frac{1}{\sqrt[n]{n}\sqrt[n]{n}} = 1$. Für $|x| = 1$ konvergiert sie nach dem Majorantenkriterium ($\sum\limits_{n=1}^{\infty} \frac{1}{n^2}$ ist konvergente Majorante).

7) Die Potenzreihe $\sum\limits_{n=0}^{\infty} \big(3 + (-1)^n\big)^n x^n$ hat Konvergenzradius $R = \frac{1}{4}$, denn es gilt

$$\sqrt[n]{|a_n|} = \begin{cases} 2, & n \text{ ungerade} \\ 4, & n \text{ gerade} \end{cases} \quad \Rightarrow \quad \frac{1}{\limsup\limits_{n\to\infty} \sqrt[n]{|a_n|}} = \frac{1}{4}.$$

Für $|x| = \frac{1}{4}$ divergiert sie, da alle Summanden, die gerades n haben, gleich 1 sind und alle Summanden mit ungeradem n entweder $\big(\frac{1}{2}\big)^n$ oder $-\big(\frac{1}{2}\big)^n$ sind.

8) Die Potenzreihe $\sum\limits_{n=1}^{\infty} (-1)^{n-1} \frac{1}{n} x^n$ hat Konvergenzradius $R = 1$, denn es gilt $\lim\limits_{n\to\infty} \frac{|a_n|}{|a_{n+1}|} = \lim\limits_{n\to\infty} \frac{n+1}{n} = 1$ oder $\lim\limits_{n\to\infty} \frac{1}{\sqrt[n]{|a_n|}} = \lim\limits_{n\to\infty} \frac{1}{\sqrt[n]{1/n}} = \lim\limits_{n\to\infty} \sqrt[n]{n} = 1$.
Für $x = 1$ konvergiert sie nach dem Leibniz-Kriterium, für $x = -1$ divergiert sie wegen $\sum\limits_{n=1}^{\infty} (-1)^{n-1} \frac{1}{n} (-1)^n = \sum\limits_{n=1}^{\infty} \big(-\frac{1}{n}\big)$.

Satz 17.8 (Identitätssatz) *Seien $f(x) = \sum\limits_{n=0}^{\infty} a_n (x-x_0)^n$ und $g(x) = \sum\limits_{n=0}^{\infty} b_n (x-x_0)^n$ Potenzreihen mit positiven Konvergenzradien R_f bzw. R_g. Existiert eine Folge $(x_m)_{m\in\mathbb{N}}$ in $\mathbb{R} \setminus \{x_0\}$ mit $\lim\limits_{m\to\infty} x_m = x_0$ und $f(x_m) = g(x_m)$ für alle $m \in \mathbb{N}$, dann ist $a_n = b_n$ für alle $n \in \mathbb{N}_0$ und somit $R_f = R_g$ und $f = g$.*

Beweis
Sei O. B. d. A. $x_0 = 0$ (anderenfalls ersetze x durch $\tilde{x} = x + x_0$) und sei $R = \min\{R_f, R_g\} \in \mathbb{R}^+$. Zeige $a_n = b_n$ für alle $n \in \mathbb{N}_0$ durch vollständige Induktion.
$n = 0$:

$$\forall m \in \mathbb{N} : 0 = f(x_m) - g(x_m) = \sum_{n=0}^{\infty} (a_n - b_n) x_m^n = (a_0 - b_0) + x_m \sum_{n=1}^{\infty} (a_n - b_n) x_m^{n-1}.$$

Wegen $\lim\limits_{m\to\infty} x_m = x_0 = 0$ und da $\sum\limits_{m=1}^{\infty} (a_n - b_n) x_m^{n-1}$ beschränkt ist, folgt $a_0 = b_0$.
$n \to n+1$:
Sei $a_m = b_m$ für $m = 0, 1, \ldots, n$. Dann gilt

$$\forall m \in \mathbb{N} : 0 = f(x_m) - g(x_m) = \underbrace{x_m^{n+1}}_{\neq 0} \sum_{k=n+1}^{\infty} (a_k - b_k) x_m^{k-(n+1)}$$

$$\Rightarrow \ \forall m \in \mathbb{N} : 0 = (a_{n+1} - b_{n+1}) + x_m \sum_{k=n+2}^{\infty} (a_k - b_k) x_m^{k-(n+2)}.$$

Wie bei $n = 0$ folgt auch hier $a_{n+1} = b_{n+1}$. □

Korollar 17.9 *Sei $x_0 \in \mathbb{R}$, $R > 0$, $B_R(x_0) = \{x \in \mathbb{R} : |x - x_0| < R\}$, $f \in C^\infty(B_R(x_0))$ und $f(x) = \sum_{n=0}^{\infty} a_n (x - x_0)^n$ für alle $x \in B_R(x_0)$. Dann gilt*

$$\forall n \in \mathbb{N}_0 : \quad a_n = \frac{f^{(n)}(x_0)}{n!},$$

d. h., die Potenzreihendarstellung von f ist die Taylorreihe von f um x_0.

Bemerkung
Später werden wir zeigen, dass die Voraussetzung $f \in C^\infty(B_R(x_0))$ bereits aus den anderen Voraussetzungen des Korollars 17.9 folgt.

Beispiel
Sei $f(x) = \dfrac{1}{1 - x^2}$. Dann ist $f \in C^\infty(]-1, 1[)$. Mit Hilfe der Grenzwertformel für die geometrische Reihe folgt $f(x) = \sum_{k=0}^{\infty} x^{2k}$ für $x^2 \in]-1, 1[$ und damit für alle $x \in]-1, 1[$. Nach Korollar 17.9 muss $\sum_{k=0}^{\infty} x^{2k}$ die Taylorreihe von f um 0 sein und daher gilt $f^{(n)} = \begin{cases} 0, & n \text{ ungerade}, \\ n!, & n \text{ gerade}. \end{cases}$

Satz 17.10 (Multiplikation von Potenzreihen) *Seien $f(x) = \sum_{n=0}^{\infty} a_n (x - x_0)^n$ und $g(x) = \sum_{n=0}^{\infty} b_n (x - x_0)^n$ mit positiven Konvergenzradien R_f, bzw. R_g. Dann gilt für $|x - x_0| < \min\{R_f, R_g\}$:*

$$f(x)g(x) = \sum_{n=0}^{\infty} c_n (x - x_0)^n \quad \textit{mit } c_n = \sum_{k=0}^{n} a_k b_{n-k}.$$

Beweis
Da die beiden Potenzreihen nach Satz 17.7 a) für $|x - x_0| < \min\{R_f, R_g\}$ absolut konvergieren, folgt die Behauptung aus der Formel für das Cauchy-Produkt von Reihen (Satz 14.17). □

Bemerkung
Die Sätze 17.6 – 17.8, Korollar 17.9 und Satz 17.10 übertragen sich sinngemäß auf komplexe Funktionenfolgen bzw. komplexe Potenzreihen.

Beispiel
Die Potenzreihe $\sum\limits_{n=0}^{\infty} \frac{1}{n!} z^n$ ist für alle $z \in \mathbb{C}$ konvergent und absolut konvergent. Dies folgt, wie wir bereits wissen, aus dem Quotientenkriterium. Alternativ erhält man dies auch mit Hilfe von Satz 17.7, denn wegen $\dfrac{1/n!}{1/(n+1)!} \to \infty$ für $n \to \infty$ hat diese Reihe den Konvergenzradius $R = \infty$, was außerdem impliziert, dass die Reihe für jedes $\rho \in \mathbb{R}^+$ gleichmäßig auf $\{z \in \mathbb{C} : |z| \leq \rho\}$ konvergiert. Auf ganz $\mathbb{C}$ konvergiert sie nicht gleichmäßig, da sie, wie wir bereits wissen, schon auf $\mathbb{R} \subseteq \mathbb{C}$ nicht gleichmäßig konvergiert.

Definition (Komplexe Exponentialfunktion) *Die Funktion*

$$\exp : z \mapsto \sum_{n=0}^{\infty} \frac{1}{n!} z^n$$

für alle $z \in \mathbb{C}$ heißt komplexe Exponentialfunktion. Man bezeichnet $e^z := \sum\limits_{n=0}^{\infty} \frac{1}{n!} z^n$ für alle $z \in \mathbb{C}$.

Satz 17.11 *(i) $\forall z, w \in \mathbb{C} : e^{z+w} = e^z e^w$.*

(ii) $\forall y \in \mathbb{R} : e^{iy} = \cos y + i \sin y$.

(iii) $\forall y \in \mathbb{R} : \cos y = \dfrac{e^{iy} + e^{-iy}}{2} \ \wedge\ \sin y = \dfrac{e^{iy} - e^{-iy}}{2i}$.

(iv) $\forall x, y \in \mathbb{R} : e^{x+iy} = e^x(\cos y + i \sin y)$.

(v) $\forall z \in \mathbb{C}\ \forall k \in \mathbb{Z} : e^{z+2\pi ki} = e^z$.

(vi) $\exp$ ist bijektiv als Funktion von $\mathbb{R} \times]-\pi, \pi]$ (aufgefasst als Teilmenge von $\mathbb{C}$) nach $\mathbb{C} \setminus \{0\}$. Die Umkehrfunktion heißt komplexer Logarithmus (oder komplexer natürlicher Logarithmus oder Hauptzweig des komplexen Logarithmus)

$$\ln : \mathbb{C} \setminus \{0\} \to \mathbb{R} \times]-\pi, \pi]\,,\ z \mapsto \ln z$$

mit

$$\ln z = \ln |z| + i \cdot \begin{cases} \arctan \frac{Im\, z}{Re\, z}, & Re\, z > 0 \\ \pi + \arctan \frac{Im\, z}{Re\, z}, & Re\, z < 0 \wedge Im\, z \geq 0 \\ -\pi + \arctan \frac{Im\, z}{Re\, z}, & Re\, z < 0 \wedge Im\, z < 0 \\ \frac{\pi}{2}, & Re\, z = 0 \wedge Im\, z > 0 \\ -\frac{\pi}{2}, & Re\, z = 0 \wedge Im\, z < 0 \end{cases}$$

(vii) $\exp$ ist auf ganz $\mathbb{C}$ stetig und unendlich oft komplex differenzierbar, wobei $\dfrac{d}{dz} e^z = e^z$ für alle $z \in \mathbb{C}$ ist.

(viii) $\forall z \in \mathbb{C} : e^z = \lim\limits_{n\to\infty} \left(1 + \frac{z}{n}\right)^n = \lim\limits_{n\to\infty} \left(1 - \frac{z}{n}\right)^{-n}$.

Beweisstrategie
Durch Nachrechnen, z. B. bei
(ii):

$$\begin{aligned}\forall y \in \mathbb{R} : e^{iy} &= \sum_{n=0}^{\infty} \frac{(iy)^n}{n!} = \sum_{k=0}^{\infty} \left((-1)^k \frac{y^{2k}}{(2k)!} + i(-1)^k \frac{y^{2k+1}}{(2k+1)!} \right) \\ &= \sum_{k=0}^{\infty} (-1)^k \frac{y^{2k}}{(2k)!} + i \sum_{k=0}^{\infty} (-1)^k \frac{y^{2k+1}}{(2k+1)!} \\ &= \cos y + i \sin y.\end{aligned}$$

(vii):

$$\begin{aligned}\lim_{z\to z_0} \frac{e^z - e^{z_0}}{z - z_0} &\overset{(i)}{=} e^{z_0} \lim_{z\to z_0} \frac{e^{z-z_0} - 1}{z - z_0} \\ &= e^{z_0} \lim_{z\to z_0} \sum_{n=1}^{\infty} \frac{1}{n!} (z - z_0)^{n-1} \\ &= e^{z_0} \lim_{z\to z_0} \left(1 + (z - z_0) \sum_{n=2}^{\infty} \frac{1}{n!} (z - z_0)^{n-2} \right) \\ &= e^{z_0},\end{aligned}$$

da $\sum\limits_{n=2}^{\infty} \frac{1}{n!} (z - z_0)^{n-2}$ wegen $z \to z_0$ beschränkt ist.

Bemerkung
Die Rechenregel $\ln(ab) = \ln a + \ln b$ des reellen Logarithmus (Satz 13.14 (iv)) lässt sich nicht auf den komplexen Logarithmus ausweiten, denn es gilt $\ln(-1) + \ln(-1) = 2\pi i \neq 0 = \ln 1 = \ln((-1)(-1))$.

Definition (Komplexe Potenzen) *Für $w \in \mathbb{C} \setminus \{0\}$ und $z \in \mathbb{C}$ sei $w^z := e^{z \ln w}$.*

Definition (Komplexe trigonometrische Funktionen, komplexe Hyperbelfunktionen) *Für alle $z \in \mathbb{C}$ ist definiert:*

a) $\sin z := \dfrac{e^{iz} - e^{-iz}}{2i}$ *((komplexer) Sinus)*

b) $\cos z := \dfrac{e^{iz} + e^{-iz}}{2}$ *((komplexer) Kosinus)*

c) $\sinh z := \dfrac{e^z - e^{-z}}{2}$ *((komplexer) Sinus hyperbolicus)* *d)* $\cosh z := \dfrac{e^z + e^{-z}}{2}$ *((komplexer) Kosinus hyperbolicus)*

Außerdem ist definiert:

e) $\tan z := \dfrac{\sin z}{\cos z}$, *falls* $\cos z \neq 0$ *((komplexer) Tangens)*

f) $\cot z := \dfrac{\cos z}{\sin z}$, *falls* $\sin z \neq 0$ *((komplexer) Kotangens)*

g) $\tanh z := \dfrac{\sinh z}{\cosh z}$, *falls* $\cosh z \neq 0$ *((komplexer) Tangens hyperbolicus)*

h) $\coth z := \dfrac{\cosh z}{\sinh z}$, *falls* $\sinh z \neq 0$ *((komplexer) Kotangens hyperbolicus)*

Satz 17.12

(i) $\forall z \in \mathbb{C} : \sinh z = -i \sin(iz)$,
$\forall z \in \mathbb{C} : \cosh z = \cos(iz)$.

(ii) $\forall z \in \mathbb{C} : \sin(-z) = -\sin z$,
$\forall z \in \mathbb{C} : \cos(-z) = \cos z$.

(iii) $\sin z = 0 \iff z \in \{k\pi : k \in \mathbb{Z}\}$,
$\cos z = 0 \iff z \in \{\left(\frac{1}{2} + k\right)\pi : k \in \mathbb{Z}\}$.

(iv) $\forall z \in \mathbb{C} : \sin(z + 2\pi) = \sin z$,
$\forall z \in \mathbb{C} : \cos(z + 2\pi) = \cos z$,
$\forall z \in \mathbb{C} : \sin(z + \pi) = -\sin z$,
$\forall z \in \mathbb{C} : \cos(z + \pi) = -\cos z$,
$\forall z \in \mathbb{C} : \sin(z + \frac{\pi}{2}) = \cos z$,
$\forall z \in \mathbb{C} : \cos(z + \frac{\pi}{2}) = -\sin z$,
$\forall z \in \mathbb{C} \setminus \{\left(\frac{1}{2} + k\right)\pi : k \in \mathbb{Z}\} : \tan(z + \pi) = \tan z$.

(v) (Eulersche Formel)
$\forall z \in \mathbb{C} : e^{iz} = \cos z + i \sin z$.

(vi) (Formel von de Moivre)
$\forall n \in \mathbb{N}\ \forall z \in \mathbb{C} : (\cos z + i \sin z)^n = e^{inz} = \cos(nz) + i \sin(nz)$.

(vii) $\forall z \in \mathbb{C} : \sin^2 z + \cos^2 z = 1$.

(viii) (Additionstheoreme)
$\forall z, w \in \mathbb{C} : \sin(z + w) = \sin z \cdot \cos w + \cos z \cdot \sin w$,
$\forall z, w \in \mathbb{C} : \cos(z + w) = \cos z \cdot \cos w - \sin z \cdot \sin w$.

(ix) $\forall x, y \in \mathbb{R} : \sin(x + iy) = \sin x \cdot \cosh y + i \cos x \cdot \sinh y,$
$\forall x, y \in \mathbb{R} : \cos(x + iy) = \cos x \cdot \cosh y - i \sin x \cdot \sinh y.$

(x) $\forall z \in \mathbb{C} : \dfrac{d}{dz} \sin z = \cos z,$

$\forall z \in \mathbb{C} : \dfrac{d}{dz} \cos z = -\sin z,$

$\forall z \in \mathbb{C} \setminus \{\left(\frac{1}{2} + k\right)\pi : k \in \mathbb{Z}\} : \dfrac{d}{dz} \tan z = \dfrac{1}{\cos^2 z},$

$\forall z \in \mathbb{C} \setminus \{k\pi : k \in \mathbb{Z}\} : \dfrac{d}{dz} \cot z = -\dfrac{1}{\sin^2 z}.$

(xi) $\forall z \in \mathbb{C} : \sin z = \sum\limits_{n=0}^{\infty} \dfrac{(-1)^n}{(2n+1)!} z^{2n+1},$

$\forall z \in \mathbb{C} : \cos z = \sum\limits_{n=0}^{\infty} \dfrac{(-1)^n}{(2n)!} z^{2n}.$

Beweisstrategie
Durch Nachrechnen, z. B. bei
$(viii)$: Für alle $z, w \in \mathbb{C}$ gilt

$$\begin{aligned}
\sin z \cdot \cos w + \cos z \cdot \sin w &= \frac{1}{4i}\left(e^{iz} - e^{-iz}\right)\left(e^{iw} + e^{-iw}\right) + \frac{1}{4i}\left(e^{iw} - e^{-iw}\right)\left(e^{iz} + e^{-iz}\right) \\
&= \frac{1}{4i}\left(2e^{iz}e^{iw} - 2e^{-iz}e^{-iw}\right) \\
&= \frac{1}{2i}\left(e^{i(z+w)} - e^{-i(z+w)}\right) \\
&= \sin(z + w)
\end{aligned}$$

Bemerkungen

1) Wegen 17.12 (i) folgen aus 17.12 $(ii) - (xi)$ auch vergleichbare Formeln für die Hyperbelfunktionen, z. B. gelten

   a) $\forall z \in \mathbb{C} : \sinh(-z) = -\sinh z,$
   $\forall z \in \mathbb{C} : \cosh(-z) = \cosh z.$

   b) $\sinh z = 0 \iff z \in \{ik\pi : k \in \mathbb{Z}\},$
   $\cosh z = 0 \iff z \in \{i\left(\frac{1}{2} + k\right)\pi : k \in \mathbb{Z}\}.$

   c) $\forall z \in \mathbb{C} : \cosh^2 z - \sinh^2 z = 1.$

   d) $\forall z \in \mathbb{C} : \dfrac{d}{dz} \sinh z = \cosh z,$

   $\forall z \in \mathbb{C} : \dfrac{d}{dz} \cosh z = \sinh z.$

2) Der reelle Sinus hyperbolicus und der reelle Kosinus hyperbolicus haben die folgende geometrische Bedeutung: Jede Ursprungsgerade im $\mathbb{R}^2$ der Form $y = ax$ mit $a \geq 0$ schneidet die Hyperbel $x^2 - y^2 = 1$ in der Halbebene $x > 0$ im Punkt $(\cosh A, \sinh A)$, wobei A der Flächeninhalt der von der Geraden, von ihrem Spiegelbild bezüglich der x-Achse und von dem in $x > 0$ gelegenen Hyperbelast begrenzten Fläche ist.

3) $\sinh|_{\mathbb{R}} : \mathbb{R} \to \mathbb{R},\ x \mapsto \sinh x$ ist bijektiv. Die zugehörige Umkehrfunktion heißt arsinh (Areasinus hyperbolicus). Es gilt $\operatorname{arsinh} x = \ln(x + \sqrt{x^2+1})$.

   $\cosh|_{\mathbb{R}_0^+} : \mathbb{R}_0^+ \to [1, \infty[$ ist bijektiv. Die zugehörige Umkehrfunktion heißt arcosh (Areakosinus hyperbolicus). Es gilt $\operatorname{arcosh} x = \ln(x + \sqrt{x^2-1})$.

4) Für eine Diskussion aller existierenden Umkehrfunktionen zu komplexen und reellen trigonometrischen Funktionen und Hyperbelfunktionen sei auf die Literatur verwiesen (Arkusfunktionen und Areafunktionen).

18 Integrierbarkeit

Motivation
Sei $f : [a, b] \to \mathbb{R}$ eine Funktion.
Problem: Welchen Flächeninhalt hat die Fläche zwischen dem Graphen von f, den Geraden $x = a$ und $x = b$ und der x-Achse?
Lösungsstrategie: Approximiert man die Fläche mit immer größer werdender Genauigkeit durch Rechtecke, dann konvergiert der Flächeninhalt der approximierenden Rechtecke gegen den gesuchten Flächeninhalt.

Definition

a) *Für $I \subseteq \mathbb{R}$ mit $I \neq \emptyset$ heißt*

$$\chi_I : \mathbb{R} \to \mathbb{R}, x \to \begin{cases} 1, & x \in I, \\ 0, & x \notin I, \end{cases}$$

die charakteristische Funktion von I.

b) *$f : [a, b] \to \mathbb{R}$ heißt Treppenfunktion auf $[a, b]$, falls es eine endliche Unterteilung (Zerlegung, Partition) $a = x_0 < x_1 < \ldots < x_n = b$ von $[a, b]$ gibt, so dass f auf jedem offenen Intervall $]x_{i-1}, x_i[, 1 \leq i \leq n$, konstant ist. (Die Funktionswerte von f an den Stellen $x_0, x_1, \ldots, x_n$ können beliebig sein.)*

c) *Zwei Treppenfunktionen $f, g : [a, b] \to \mathbb{R}$ heißen fast überall gleich, kurz: $f = g$ f.ü., falls $f(x) \neq g(x)$ für höchstens endlich viele $x \in [a, b]$ gilt.*

Beispiel

$$f : [0, 3] \to \mathbb{R},\ x \to \begin{cases} 2, & 0 \leq x \leq 1, \\ 3, & 1 < x < 2, \\ 4, & x = 2, \\ -1, & 2 < x \leq 3, \end{cases}$$

ist eine Treppenfunktion. Es gilt

$$f = 2\chi_{]0,1[} + 3\chi_{]1,2[} - \chi_{]2,3[}\ f.ü.$$

Satz 18.1 a) *$f : [a, b] \to \mathbb{R}$ ist genau dann eine Treppenfunktion, wenn endlich viele Intervalle $I_k \subseteq [a, b],\ k = 1, \ldots, n$, und Konstanten $c_k \in \mathbb{R}$ existieren, so dass gilt: $f = \sum_{k=1}^{n} c_k \chi_{I_k}$ f.ü..*

b) *Sind $f, g : [a, b] \to \mathbb{R}$ Treppenfunktionen, dann sind für alle $\alpha, \beta \in \mathbb{R}$ die Funktionen $\alpha f + \beta g$ Treppenfunktionen. Die Menge $\mathcal{T}([a, b])$ aller Treppenfunktionen auf $[a, b]$ ist folglich ein $\mathbb{R}$-Vektorraum.*

c) *Sind* $f, g : [a, b] \to \mathbb{R}$ *Treppenfunktionen, dann ist auch* fg *eine Treppenfunktion.*

d) *Ist* $f : [a, b] \to \mathbb{R}$ *eine Treppenfunktion, dann auch* $|f|$*.*

Definition *Ist* $f : [a, b] \to \mathbb{R}$ *eine Treppenfunktion und* $f = \sum_{k=1}^{n} c_k \chi_{]x_{k-1}, x_k[}$ *f.ü., dann heißt*

$$\int_a^b f(x)\,dx := \sum_{k=1}^{n} c_k (x_k - x_{k-1})$$

das (bestimmte) Integral von f *über* $[a, b]$*. Außerdem sei*

$$\begin{aligned} & \int_b^a f(x)\,dx &:=& -\int_a^b f(x)\,dx, \\ \forall a \le c \le d \le b : \quad & \int_c^d f(x)\,dx &:=& \int_a^b (f \cdot \chi_{[c,d]})(x)\,dx. \end{aligned}$$

Die Definition des Integrals ist von der Wahl der Zerlegung von $[a, b]$ *unabhängig, da*

$$c_k(x_k - x_{k-1}) = c_k(t_k - x_{k-1}) + c_k(x_k - t_k)$$

für jeden Punkt $t_k \in]x_{k-1}, x_k[$ *gilt. Außerdem gilt:*

$$f = g \ \text{f.ü.} \ \Rightarrow \ \int_a^b f(x)\,dx = \int_a^b g(x)\,dx.$$

Beispiel

$$\int_0^3 (2\chi_{]0,1[} + 3\chi_{]1,2[} - \chi_{]2,3[})(x)\,dx \ = \ 4$$

$$\int_0^2 (2\chi_{]0,1[} + 3\chi_{]1,2[} - \chi_{]2,3[})(x)\,dx \ = \ 5$$

Satz 18.2 *Seien* f, g *Treppenfunktionen auf* $[a, b]$*. Dann gilt*

a) $\forall \alpha, \beta \in \mathbb{R} : \quad \int_a^b (\alpha f + \beta g)(x)\,dx = \alpha \int_a^b f(x)\,dx + \beta \int_a^b g(x)\,dx$

b) $\forall x \in [a, b] : \quad f(x) \le g(x) \ \Rightarrow \ \int_a^b f(x)\,dx \le \int_a^b g(x)\,dx$

Definition *Sei $f : [a,b] \to \mathbb{R}$ eine beschränkte Funktion.*

a)

$$\underline{\int_a^b} f(x)\,dx := \sup\left\{\int_a^b \varphi(x)\,dx : \ \varphi \in \mathcal{T}([a,b]) \ \wedge \ \forall x \in [a,b] : \varphi(x) \le f(x)\right\}$$

heißt Unterintegral von f über $[a,b]$ und

$$\overline{\int_a^b} f(x)\,dx := \inf\left\{\int_a^b \varphi(x)\,dx : \ \varphi \in \mathcal{T}([a,b]) \ \wedge \ \forall x \in [a,b] : \varphi(x) \ge f(x)\right\}$$

heißt Oberintegral von f über $[a,b]$.

b) f heißt Riemann-integrierbar über $[a,b]$, falls

$$\underline{\int_a^b} f(x)\,dx = \overline{\int_a^b} f(x)\,dx.$$

In diesem Fall heißt

$$\int_a^b f(x)\,dx = \underline{\int_a^b} f(x)\,dx = \overline{\int_a^b} f(x)\,dx$$

das (bestimmte) Riemann-Integral von f über $[a,b]$. Außerdem sei

$$\begin{aligned}\int_b^a f(x)\,dx &:= -\int_a^b f(x)\,dx,\\ \int_a^a f(x)\,dx &:= 0.\end{aligned}$$

Beispiele

1) Jede Treppenfunktion ist Riemann-integrierbar.

2) Die Dirichletsche Sprungfunktion $\chi_{\mathbb{Q}}$ ist nicht Riemann-integrierbar über $[0,1]$, denn es gilt

$$\begin{aligned}\underline{\int_0^1} \chi_{\mathbb{Q}}(x)\,dx &= 0 \qquad \text{und}\\ \overline{\int_0^1} \chi_{\mathbb{Q}}(x)\,dx &= 1.\end{aligned}$$

Definition *Sei* $a = x_0 < x_1 < \ldots < x_n = b$ *eine endliche Unterteilung von* $[a,b]$*,* $\xi_k \in [x_{k-1}, x_k]$ *für* $k = 1, \ldots, n$ *,* $\mathcal{Z} := ((x_0, x_1, \ldots, x_n), (\xi_1, \ldots, \xi_n))$*,*
$\mu(\mathcal{Z}) := \max\limits_{1 \le k \le n} (x_k - x_{k-1})$ *und* $f : [a,b] \to \mathbb{R}$ *eine beschränkte Funktion. Dann heißt*

$$R(f, \mathcal{Z}) := \sum_{k=1}^{n} f(\xi_k)(x_k - x_{k-1})$$

Riemannsche Summe von f *über* $[a,b]$ *bezüglich* $\mathcal{Z}$ *mit Feinheit* $\mu(\mathcal{Z})$*.*

Satz 18.3 *Sei* $f : [a,b] \to \mathbb{R}$ *Riemann-integrierbar. Dann existiert zu jedem* $\varepsilon > 0$ *ein* $\delta > 0$*, so dass für jede Wahl von* $\mathcal{Z}$ *mit* $\mu(\mathcal{Z}) < \delta$ *gilt*

$$\Big| \int_a^b f(x)\, dx - R(f, \mathcal{Z}) \Big| < \varepsilon.$$

Kurzschreibweise:

$$\lim_{\mu(\mathcal{Z}) \to 0} R(f, \mathcal{Z}) = \int_a^b f(x)\, dx.$$

Beweisstrategie
Zeige die Gültigkeit der Behauptung zunächst für Treppenfunktionen. Dies impliziert die Gültigkeit der Behauptung für Riemann-integrierbare Funktionen f, denn für alle Treppenfunktionen φ, ψ mit $\varphi(x) \le f(x) \le \psi(x)$ für alle $x \in [a,b]$ und für jede Wahl von $\mathcal{Z}$ gilt

$$R(\varphi, \mathcal{Z}) \le R(f, \mathcal{Z}) \le R(\psi, \mathcal{Z}).$$

Beispiel

Berechnung von $\int\limits_0^b x\ dx,\ b > 0$, mittels Riemannscher Summen:

Für alle $n \in \mathbb{N}$ und alle $k \in \{0, 1, \ldots, n\}$ sei $x_{nk} := \frac{k}{n} \cdot b,\ \xi_{nk} := x_{nk}$ und $\mathcal{Z}_n := ((x_{n0}, \ldots, x_{nn}), (\xi_{n1}, \ldots, \xi_{nn}))$. Dann gilt

$$\begin{aligned}
\mu(\mathcal{Z}) &= \frac{b}{n} \\
R(x, \mathcal{Z}_n) &= \sum_{k=1}^{n} \frac{kb}{n} \cdot \frac{b}{n} = \frac{b^2}{n^2} \cdot \sum_{k=1}^{n} k = \frac{b^2}{n^2} \frac{n(n+1)}{2} = \frac{b^2}{2}\Big(1 + \frac{1}{n}\Big) \\
\Rightarrow \int_0^b x\ dx &= \lim_{n \to \infty} R(x, \mathcal{Z}_n) = \lim_{n \to \infty} \frac{b^2}{2}\Big(1 + \frac{1}{n}\Big) = \frac{b^2}{2}
\end{aligned}$$

Satz 18.4 *Eine Funktion* $f : [a,b] \to \mathbb{R}$ *ist genau dann Riemann-integrierbar, wenn zu jedem* $\varepsilon > 0$ *Treppenfunktionen* $\varphi, \psi \in \mathcal{T}([a,b])$ *existieren mit* $\varphi \le f \le \psi$*, d.h.* $\varphi(x) \le f(x) \le \psi(x)$ *für alle* $x \in [a,b]$*, so dass gilt:*

$$\int_a^b \psi(x)\ dx - \int_a^b \varphi(x)\ dx \le \varepsilon.$$

Bemerkung
Von nun an schreiben wir integrierbar statt Riemann-integrierbar und Integral statt Riemann-Integral.

Satz 18.5 *Jede monotone Funktion* $f : [a, b] \to \mathbb{R}$ *ist integrierbar.*

Beweis
Sei $\varepsilon > 0$, $n \in \mathbb{N}$ und $x_k := a + k\frac{b-a}{n}$, $k = 0, 1, \ldots, n$. Ist f monoton wachsend, dann definiere $\varphi, \psi \in \mathcal{T}([a, b])$ durch

$$\begin{aligned}
\varphi(x) &:= f(x_{k-1}) = \min_{x\in[x_{k-1},x_k]} f(x) \qquad \text{für } x_{k-1} \le x < x_k, \quad k = 1, \ldots, n, \\
\psi(x) &:= f(x_k) = \max_{x\in[x_{k-1},x_k]} f(x) \qquad \text{für } x_{k-1} \le x < x_k, \quad k = 1, \ldots, n, \\
\varphi(b) &= \psi(b) := f(b).
\end{aligned}$$

Dann gilt $\varphi \le f \le \psi$ und

$$\begin{aligned}
\int_a^b \psi(x)\,dx - \int_a^b \varphi(x)\,dx &= \sum_{k=1}^n f(x_k)(x_k - x_{k-1}) - \sum_{k=1}^n f(x_{k-1})(x_k - x_{k-1}) \\
&= \frac{b-a}{n}\left(\sum_{k=1}^n f(x_k) - \sum_{k=1}^n f(x_{k-1})\right) \\
&= \frac{b-a}{n}\,(f(x_n) - f(x_0)) \\
&< \varepsilon,
\end{aligned}$$

falls n hinreichend groß gewählt ist.
Also folgt nach Satz 18.4 die Behauptung für monoton wachsende Funktionen.
Ist f monoton fallend, dann beweist man die Behauptung durch analoge Argumentation. □

Satz 18.6 *Jede stetige Funktion* $f : [a, b] \to \mathbb{R}$ *ist integrierbar.*

Satz 18.7 *Jede auf einem beschränkten und abgeschlossenen Intervall* $[a, b]$ *stetige Funktion* $f : [a, b] \to \mathbb{R}$ *ist dort gleichmäßig stetig, d.h.*

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x_0, x \in [a, b] : \qquad |x - x_0| < \delta \ \Rightarrow\ |f(x) - f(x_0)| < \varepsilon.$$

Beweis
Sei $\varepsilon > 0$. Angenommen, f sei nicht gleichmäßig stetig auf $[a, b]$.

$$\Rightarrow\ \forall n \in \mathbb{N}\ \exists x_n, x_n' \in [a, b] : \quad |x_n - x_n'| < \frac{1}{n} \ \wedge\ |f(x_n) - f(x_n')| \ge \varepsilon$$

Da $(x_n)_{n\in\mathbb{N}}$ eine beschränkte Folge ist, besitzt sie nach dem Satz von Bolzano-Weierstraß eine konvergente Teilfolge $(x_{n_k})_{k\in\mathbb{N}}$ mit $\lim\limits_{k\to\infty} x_{n_k} \in [a, b]$.

Wegen $|x_{n_k} - x'_{n_k}| < \frac{1}{n_k}$ folgt $\lim\limits_{k\to\infty} x'_{n_k} = \lim\limits_{k\to\infty} x_{n_k}$.
Da f stetig auf $[a,b]$ ist, ergibt sich daraus

$$\lim_{k\to\infty}(f(x_{n_k}) - f(x'_{n_k})) = f(\lim_{n\to\infty} x_{n_k}) - f(\lim_{k\to\infty} x'_{n_k}) = 0,$$

ein Widerspruch zu $|f(x_{n_k}) - f(x'_{n_k})| \geq \varepsilon$ für alle $k \in \mathbb{N}$. □

Bemerkung
Stetige Funktionen auf nicht beschränkten oder nicht abgeschlossenen Intervallen müssen dort nicht gleichmäßig stetig sein, wie die Beispiele $f(x) = x^2$ auf $[0, \infty[$ und $g(x) = \frac{1}{x}$ auf $]0, 1]$ zeigen.

Korollar 18.8 *Sei $f : [a,b] \to \mathbb{R}$ stetig. Dann gibt es zu jedem $\varepsilon > 0$ Treppenfunktionen $\varphi, \psi : [a,b] \to \mathbb{R}$ mit*

(i) $\forall x \in [a,b] : \varphi(x) \leq f(x) \leq \psi(x)$,

(ii) $\forall x \in [a,b] : |\psi(x) - \varphi(x)| < \varepsilon$.

Beweis von Satz 18.6
Sei $f : [a,b] \to \mathbb{R}$ stetig. Nach Korollar 18.8 gilt:

$$\forall \varepsilon > 0\ \exists \varphi, \psi \in \mathcal{T}([a,b]) : \quad \varphi \leq f \leq \psi \quad \wedge \quad \forall x \in [a,b] :\ \psi(x) - \varphi(x) \leq \frac{\varepsilon}{b-a}$$

Daher folgt nach Satz 18.2 :

$$\int_a^b \psi(x)\,dx - \int_a^b \varphi(x)\,dx = \int_a^b (\psi - \varphi)(x)\,dx \leq \int_a^b \frac{\varepsilon}{b-a}\,dx = \varepsilon,$$

woraus sich nach Satz 18.4 die Behauptung ergibt. □

Satz 18.9 *Seien $f, g : [a,b] \to \mathbb{R}$ integrierbar. Dann gilt:*

(i) Für alle $\alpha, \beta \in \mathbb{R}$ sind auch $\alpha f + \beta g$ integrierbar mit

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

(Linearität des Integrals)

(ii) $f \leq g \ \Rightarrow\ \int_a^b f(x)\,dx \leq \int_a^b g(x)\,dx$

(Monotonie des Integrals)

(iii) Die Funktionen $f_+ : [a,b] \to \mathbb{R}$ *mit* $f_+(x) = \begin{cases} f(x), & x > 0, \\ 0, & \text{sonst}, \end{cases}$

$f_- : [a,b] \to \mathbb{R}$ *mit* $f_-(x) = \begin{cases} -f(x), & x < 0, \\ 0, & \text{sonst}, \end{cases}$

und $|f| : [a,b] \to \mathbb{R}$ *sind ebenfalls integrierbar und es gilt*

$$\Big| \int_a^b f(x)\,dx \Big| \;\leq\; \int_a^b |f(x)|\,dx \;\leq\; (b-a) \sup_{x\in[a,b]} |f(x)|$$

(iv) Die Funktion fg ist ebenfalls integrierbar.

Beweisstrategie
Folgt mit Hilfe der Sätze 18.1, 18.2 und 18.4 .

Satz 18.10 *Sei* $a < c < b$ *und* $f : [a,b] \to \mathbb{R}$ *eine Funktion.* f *ist genau dann integrierbar, wenn sowohl* $f_{|_{[a,c]}}$ *als auch* $f_{|_{[c,b]}}$ *integrierbar sind. In diesem Fall gilt*

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

Satz 18.11 (Mittelwertsatz der Integralrechnung) *Seien* $f : [a,b] \to \mathbb{R}$ *und* $g : [a,b] \to \mathbb{R}_0^+$ *stetige Funktionen. Dann gilt:*

$$\exists \xi \in [a,b] : \qquad \int_a^b f(x)g(x)\,dx = f(\xi) \int_a^b g(x)\,dx,$$

insbesondere gilt für $g \equiv 1$*:*

$$\exists \xi \in [a,b] : \qquad \int_a^b f(x)\,dx = f(\xi)(b-a).$$

Beweis
Wegen $g \geq 0$ und Satz 18.9 (ii) gilt

$$\Big(\min_{x\in[a,b]} f(x)\Big) \int_a^b g(x)\,dx \leq \int_a^b f(x)g(x)\,dx \leq \Big(\max_{x\in[a,b]} f(x)\Big) \int_a^b g(x)\,dx$$

und daher

$$\exists \mu \in \big[\min_{x\in[a,b]} f(x), \max_{x\in[a,b]} f(x)\big] : \qquad \int_a^b f(x)g(x)\,dx = \mu \int_a^b g(x)\,dx$$

Nach dem Zwischenwertsatz $\exists \xi \in [a,b] :\; f(\xi) = \mu$. Daraus folgt die Behauptung. □

Bemerkung
Außer dem Riemann-Integral gibt es noch weitere Integralbegriffe:

1) Das Cauchy-Integral oder Regelintegral
Eine Funktion $f : [a,b] \to \mathbb{R}$ heißt Regelfunktion, falls es eine Folge $(t_n)_{n\in\mathbb{N}}$ von Treppenfunktionen auf $[a,b]$ gibt, die gleichmäßig auf $[a,b]$ gegen f konvergiert.
In diesem Fall existiert der Grenzwert $\lim\limits_{n\to\infty} \int_a^b t_n(x)\,dx$ und hat für jede Folge von Treppenfunktionen, die gleichmäßig auf $[a,b]$ gegen f konvergiert, denselben Wert. Es heißt dann
$$\int_a^b f(x)\,dx := \lim_{n\to\infty} \int_a^b t_n(x)\,dx$$
das (bestimmte) Cauchy-Integral oder Regelintegral von f über $[a,b]$ und f heißt Regel-integrierbar oder Cauchy-integrierbar über $[a,b]$.
Jede Cauchy-integrierbare Funktion ist auch Riemann-integrierbar und das Riemann-Integral hat denselben Wert wie das Cauchy-Integral. Es gibt aber Riemann-integrierbare Funktionen, die nicht Cauchy-integrierbar sind. Zum Beispiel ist die Funktion
$$f : [0,1] \to \mathbb{R}, \qquad x \to \begin{cases} 1, & \text{für } x = \frac{1}{n},\ n \in \mathbb{N}, \\ 0, & \text{sonst,} \end{cases}$$
Riemann-integrierbar mit Riemann-Integral $\int_0^1 f(x)\,dx = 0$, aber nicht Cauchy-integrierbar.

2) Das Lebesgue-Integral
Funktionen wie die Dirichletsche Sprungfunktion können mit Hilfe des Lebesgueschen Integralbegriffs integriert werden. Im Gegensatz zum Riemann- und zum Cauchy-Integral wird beim Lebesgue-Integral nicht die Urbildmenge einer Funktion unterteilt, sondern die Bildmenge. Die Einführung des Lebesgue-Integrals im Detail ist allerdings konzeptionell ziemlich aufwändig und würde an dieser Stelle zu weit führen.
Das Lebesgue-Integral von $\chi_{\mathbb{Q}}$ über $[0,1]$ ist übrigens gleich 0.

Satz 18.12 (Hauptsatz der Differential- und Integralrechnung) *Sei $f \in C^0([a,b])$. Dann gilt:*

a) Für jedes $\alpha \in [a,b]$ ist die Funktion $F : [a,b] \to \mathbb{R}$ mit
$$F(x) = \int_\alpha^x f(t)\,dt$$
eine Stammfunktion von f, d.h., F ist differenzierbar in $]a,b[$ mit $F' = f$.

b) Für jede Stammfunktion F von f gilt

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

$$\textit{Kurzschreibweise:} \quad \int_a^b f(x)\,dx = F(x)\Big|_a^b$$

Hilfssatz 18.13 *Sei $F : [a,b] \to \mathbb{R}$ eine Stammfunktion von $f : [a,b] \to \mathbb{R}$. $G : [a,b] \to \mathbb{R}$ ist genau dann eine weitere Stammfunktion von f wenn $F - G$ konstant ist.*

Beweis

"$\Leftarrow$" : Sei $F - G = c$ mit $c \in \mathbb{R}$. Dann gilt $G' = (F - c)' = F' = f$ und somit ist G eine Stammfunktion von f.

"$\Rightarrow$" : Sei G eine Stammfunktion von f. Dann gilt $(F - G)' = F' - G' = f - f = 0$ und nach Korollar 16.12 ist $F - G$ konstant. □

Beweis Satz 18.12

a) Für $h \neq 0$ ist

$$\frac{F(x+h) - F(x)}{h} = \frac{1}{h}\Big(\int_\alpha^{x+h} f(t)\ dt - \int_\alpha^x f(t)\ dt\Big) \overset{18.10}{=} \frac{1}{h}\Big(\int_x^{x+h} f(t)\ dt\Big)$$

Nach dem Mittelwertsatz der Integralrechnung (Satz 18.11) existiert ein ξ_h mit $\xi_h \in [x, x+h]$, falls $h > 0$ und $\xi_h \in [x+h, x]$, falls $h < 0$ ist, und

$$\int_x^{x+h} f(t)\ dt = hf(\xi_h).$$

Da $\lim\limits_{h\to 0} \xi_h = x$ gilt und f stetig ist, folgt

$$F'(x) = \lim_{h\to 0} \frac{1}{h} \int_x^{x+h} f(t)\ dt = \lim_{h\to 0} \frac{1}{h}(hf(\xi_h)) = f(x).$$

b) Sei $F_a(x) = \int_a^x f(t)\ dt$ und F eine beliebige Stammfunktion von f. Aufgrund von a) und Hilfssatz 18.13 existiert ein $c \in \mathbb{R}$ mit $F - F_a = c$. Daher gilt

$$F(b) - F(a) = F_a(b) - F_a(a) = F_a(b) = \int_a^b f(t)\ dt.$$

□

Bemerkung
Für die Menge aller Stammfunktionen einer Funktion f verwendet man auch die Notation

$$\int f(x)\,dx$$

und man nennt diese Menge auch das unbestimmte Integral von f.

Beispiele

1) $\displaystyle\int x\,dx = \frac{1}{2}x^2 + c, \qquad c \in \mathbb{R},$

$\displaystyle\int_1^2 x\,dx = \frac{1}{2}x^2\Big|_1^2 = 2 - \frac{1}{2} = \frac{3}{2}$

2) $\displaystyle\forall r \in \mathbb{R}\setminus\{-1\} : \int x^r\,dx = \frac{1}{r+1}x^{r+1} + c, \qquad c \in \mathbb{R}$

3) $\displaystyle\forall a,b > 0 : \int_a^b \frac{1}{x}\,dx = \ln x\Big|_a^b$

$\displaystyle\forall a,b < 0 : \int_a^b \frac{1}{x}\,dx = \ln(-x)\Big|_a^b$

$\displaystyle\Rightarrow \int \frac{1}{x}\,dx = \ln|x| + c, \qquad c \in \mathbb{R},$

für $x \neq 0$, d.h., 0 liegt nicht im Integrationsintervall.

4) $\displaystyle\int e^x\,dx = e^x + c, \qquad c \in \mathbb{R}$

5) $\displaystyle\int \sin x\;dx = -\cos x + c, \qquad c \in \mathbb{R}$

6) $\displaystyle\int \cos x\;dx = \sin x + c, \qquad c \in \mathbb{R}$

7) $\displaystyle\int \frac{1}{x^2+1}\,dx = \arctan x + c, \qquad c \in \mathbb{R}$

8) $\displaystyle\int \frac{1}{\sqrt{1-x^2}}\,dx = \arcsin x + c, \qquad c \in \mathbb{R} \qquad \text{für } |x| < 1$

9) $\displaystyle\int \frac{1}{\cos^2 x}\,dx = \tan x + c, \qquad c \in \mathbb{R} \qquad \text{für } \cos x \neq 0$

10) $\displaystyle\int \cosh x\;dx = \sinh x + c, \qquad c \in \mathbb{R}$

Bemerkung
Ist $f : [a, b] \to \mathbb{R}$ integrierbar und beschränkt, aber nicht stetig, dann ist für jedes $\alpha \in [a, b]$ die Funktion $F : [a, b] \to \mathbb{R}$ mit

$$F(x) = \int_\alpha^x f(t)\, dt$$

wegen Satz 18.9 (iii) stetig auf $[a, b]$, aber im Allgemeinen nicht differenzierbar in $]a, b[$, wie das Beispiel

$$F : [0, 2] \to \mathbb{R} \qquad f(x) = \begin{cases} 1, & x \in [0, 1[, \\ 2, & x \in [1, 2], \end{cases}$$

zeigt.

Satz 18.14 (Partielle Integration) *Seien $f, g \in C^1([a, b])$, d.h., $f, g \in C^0([a, b]) \cap C^1(]a, b[)$ und f', g' besitzen eine stetige Fortsetzung auf $[a, b]$. Dann gilt*

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

Beweis
Folgt aus dem Hauptsatz und der Produktregel. □

Beispiele

1) Sei $0 < a < b$.
$$\int_a^b \ln x \; dx = \int_a^b 1 \cdot \ln x \; dx = x \ln x \Big|_a^b - \int_a^b x \cdot \frac{1}{x}\, dx = x \ln x \Big|_a^b - (b - a)$$
$$= x(\ln x - 1)\Big|_a^b$$

2) $$\int_a^b x^2 e^x\, dx = x^2 e^x \Big|_a^b - \int_a^b 2x e^x\, dx = (x^2 - 2x)e^x \Big|_a^b + \int_a^b 2e^x\, dx$$
$$= (x^2 - 2x + 2)e^x \Big|_a^b$$

3) $$\int \cos^2 x \; dx = \sin x \cos x + c + \int \sin^2 x \; dx = \sin x \cos x + c + \int 1 - \cos^2 x \; dx$$
$$= \sin x \cos x + x + c - \int \cos^2 x \; dx, \qquad c \in \mathbb{R}$$

$$\Rightarrow \int \cos^2 x \; dx = \frac{1}{2}(\sin x \cos x + x) + c, \qquad c \in \mathbb{R}$$

Satz 18.15 (Substitutionsregel) *Sei* $y \in C^1([a,b])$, $y([a,b]) \subseteq [\alpha,\beta]$ *und* $f \in C^0([\alpha,\beta])$. *Dann gilt*

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

Beweis
Folgt aus dem Hauptsatz und der Kettenregel. □

Korollar 18.16 (Substitutionsregel rückwärts) *Sei* $f \in C^0([a,b])$.

a) Ist $x : [\alpha,\beta] \to [a,b]$, $t \mapsto x(t)$ *bijektiv mit Umkehrfunktion* $t : [a,b] \to [\alpha,\beta]$, $x \mapsto t(x)$ *und* $x \in C^1([\alpha,\beta])$, *dann gilt*

$$\int_a^b f(x)\,dx = \int_{t(a)}^{t(b)} f(x(t))\,\frac{dx}{dt}(t)\,dt = \int_\alpha^\beta f(x(t))\left|\frac{dx}{dt}(t)\right|dt.$$

b) Ist $t \in C^1([a,b])$ *mit* $\frac{dt}{dx}(x) \neq 0$ *für alle* $x \in [a,b]$, *dann ist* $t : [a,b] \to t([a,b])$ *bijektiv mit Umkehrfunktion* $x : t([a,b]) \to [a,b]$, $t \mapsto x(t)$ *und es gilt*

$$\int_a^b f(x)\ dx = \int_{t(a)}^{t(b)} f(x(t))\ \frac{1}{\frac{dt}{dx}(x(t))}\ dt.$$

Beweis
Folgt aus der Substitutionsregel, dem Umkehrsatz und den Monotonieaussagen aus Korollar 16.13 . □

Beispiele

1) $$\int_0^2 \frac{2x}{\sqrt{1+x^2}}\ dx = \int_1^5 \frac{1}{\sqrt{y}}\ dy = 2\sqrt{y}\,\Big|_1^5 = 2\sqrt{5}-2$$

2) $$\int \tan x\ dx = \int \frac{\sin x}{\cos x}\ dx = -\ln(|\cos x|)+c, \qquad c \in \mathbb{R}, \qquad \text{für } \cos x \neq 0$$

3) Sei g stetig differenzierbar. Dann gilt:

$$\int (g(x))^r g'(x)\ dx = \begin{cases} g(x)+c, & r=0 \\ \ln(|g(x)|)+c, & r=-1 \,\wedge\, g(x)\neq 0 \\ \frac{1}{r+1}(g(x))^{r+1}+c, & r>0 \\ \frac{1}{r+1}(g(x))^{r+1}+c, & r \in \mathbb{R}^- \setminus \{-1\} \,\wedge\, g(x)\neq 0 \end{cases}, \quad c \in \mathbb{R}$$

4) $$\int_a^b f(cx+k)\ dx = \frac{1}{c}\int_{ca+k}^{cb+k} f(y)\ dy \text{ für } f \in C^0([a,b]) \text{ und } c \neq 0$$

5) $\displaystyle\int_{-r}^{r} \sqrt{r^2 - x^2}\ dx$

Substitution: $x = r \sin t$ ($t \mapsto x(t)$ ist bijektiv als Abbildung von $[-\frac{\pi}{2}, \frac{\pi}{2}]$ nach $[-r, r]$)

$$\Rightarrow \quad dx = r \cos t\ dt \quad (\text{ d.h. } \frac{dx}{dt}(t) = r \cos t)$$

$$\Rightarrow \quad \int_{-r}^{r} \sqrt{r^2 - x^2}\ dx = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sqrt{r^2 - r^2 \sin^2 t}\ \ r \cos t\ dt$$

$$= \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} r^2 \cos^2 t\ dt = \frac{r^2}{2}(\sin t\ \cos t + t)\Big|_{-\frac{\pi}{2}}^{\frac{\pi}{2}} = \frac{\pi}{2} r^2$$

(Flächeninhalt eines Halbkreises mit Radius r)

Alternativ hätte man auch $x = r \cos t$ substituieren können.

6) $\displaystyle\int \sqrt{1 + x^2}\ dx$

Substitution: $x = \sinh t$ (bijektiv als Abbildung von $\mathbb{R}$ nach $\mathbb{R}$)

$$\Rightarrow \quad dx = \cosh t\ dt$$

$$\Rightarrow \quad \int \sqrt{1 + x^2}\ dx = \int \sqrt{1 + \sinh^2 t}\ \ \cosh t\ dt$$

$$= \int \cosh^2 t\ dt = \sinh t \cosh t + c - \int \sinh^2 t\ dt$$

$$= \sinh t \cosh t + c - \int \cosh^2 t - 1\ dt$$

$$\Rightarrow \quad \int \sqrt{1 + x^2}\ dx = \frac{1}{2}(\sinh t\ \cosh t + t) + c$$

$$= \frac{1}{2}\Big(\sinh t\ \sqrt{1 + \sinh^2 t} + t\Big) + c = \frac{1}{2}\Big(x\sqrt{1 + x^2} + \operatorname{arsinh} x\Big) + c,$$

$c \in \mathbb{R}$

Wegen $\frac{d}{dt} \cosh t = \sinh t = \sqrt{\cosh^2 t - 1} \neq \sqrt{1 + \cosh^2 t}$ wäre hier die Substitution $x = \cosh t$ für $x \geq 0$ und $x = -\cosh t$ für $x < 0$ nicht sinnvoll gewesen. Diese Substitution ist sinnvoll bei der Berechnung von $\displaystyle\int \sqrt{x^2 - 1}\ dx$.

7) $\displaystyle\int x\sqrt{1+x}\; dx\,, \quad x \geq -1$

Substitution $t = \sqrt{1+x} \;\Rightarrow\; x = t^2 - 1\,,\; t \geq 0$

$$\begin{aligned}
&\Rightarrow \quad dx = 2t\; dt \\
&\Rightarrow \quad \int x\sqrt{1+x}\; dx = 2\int (t^2-1)t^2\; dt = 2\int t^4 - t^2\; dt \\
&\quad = \frac{2}{5}t^5 - \frac{2}{3}t^3 + c = \frac{2}{5}(1+x)^{\frac{5}{2}} - \frac{2}{3}(1+x)^{\frac{3}{2}} + c, \quad c \in \mathbb{R}, \quad x \geq -1
\end{aligned}$$

8) $\displaystyle\int x^7\sqrt{1+x^4}\; dx$

Substitution: $t = 1 + x^4$

$$\begin{aligned}
&\Rightarrow \quad dt = 4x^3\; dx \\
&\Rightarrow \quad \int x^7\sqrt{1+x^4}\; dx = \frac{1}{4}\int (t-1)\sqrt{t}\; dt = \frac{1}{4}\int t^{\frac{3}{2}} - t^{\frac{1}{2}}\; dt \\
&\quad = \frac{1}{10}t^{\frac{5}{2}} - \frac{1}{6}t^{\frac{3}{2}} + c = \frac{1}{10}(1+x^4)^{\frac{5}{2}} - \frac{1}{6}(1+x^4)^{\frac{3}{2}} + c, \quad c \in \mathbb{R}
\end{aligned}$$

Bei Integralen der Form $\displaystyle\int x^n\sqrt{1+x^4}\; dx$ hilft die obige Substitution nur bei manchen n weiter.

Integration rationaler Funktionen mit Hilfe der Partialbruchzerlegung

Seien p, q reelle Polynome und sei $Grad(q) \geq 1$. Das Integral

$$\int \frac{p(x)}{q(x)}\; dx\,, \qquad q(x) \neq 0\,,$$

kann explizit angegeben werden, falls die Nullstellen von q bestimmt werden können. Für die Integration von $\frac{p}{q}$ kann man das folgende Verfahren verwenden.

1. Schritt:

Stelle $\frac{p}{q}$ in der Form

$$\frac{p(x)}{q(x)} = h(x) + \frac{r(x)}{q(x)}$$

dar, wobei h, r reelle Polynome sind und $Grad(r) < Grad(q)$ ist. Diese Darstellung ist eindeutig bestimmt und kann mit Hilfe der Polynomdivision berechnet werden (siehe Abschnitt 5.2).

2. Schritt:

Bestimme, zum Beispiel mit Hilfe des euklidischen Algorithmus, den ggT von r und q, kürze $\frac{r}{q}$ mit $ggT(r,q)$ und stelle dann $\frac{p}{q}$ in der Form

$$\frac{p(x)}{q(x)} = h(x) + \frac{\tilde{r}(x)}{\tilde{q}(x)}$$

dar, wobei $\tilde{r}$, $\tilde{q}$ teilerfremde reelle Polynome sind mit $Grad(\tilde{r}) < Grad(\tilde{q})$.

Für die weiteren Schritte benötigen wir folgende Vorüberlegungen.

Seien P, Q teilerfremde reelle oder komplexe Polynome mit $Grad(P) < Grad(Q)$ und sei $\lambda \in \mathbb{C}$ eine n-fache Nullstelle von Q, es existiere also ein eindeutig bestimmtes Polynom Q_1 mit

$$Q(z) = (z-\lambda)^n Q_1(z), \quad \text{wobei } Q_1(\lambda) \neq 0.$$

Sei $a_1 = \frac{P(\lambda)}{Q_1(\lambda)}$. Dann ist λ Nullstelle von $P - a_1 Q_1$, also gibt es ein eindeutig bestimmtes Polynom P_1 mit

$$P(z) = a_1 Q_1(z) + (z-\lambda) P_1(z).$$

Daher folgt

$$\frac{P(z)}{Q(z)} = \frac{a_1}{(z-\lambda)^n} + \frac{P_1(z)}{(z-\lambda)^{n-1} Q_1(z)},$$

wobei $Grad(P_1) < Grad(Q) - 1$ ist.
Durch Wiederholung diese Verfahrens erhält man die komplexe Partialbruchzerlegung von $\frac{P}{Q}$:

Besitzt Q die Linearfaktorzerlegung

$$Q(z) = a \prod_{j=1}^{k} (z - \lambda_j)^{n_j}$$

mit paarweise verschiedenen $\lambda_j \in \mathbb{C}$ (nach Korollar 6.10 existiert für jedes komplexe Polynom eine solche Zerlegung), dann existieren eindeutig bestimmte $a_{ij} \in \mathbb{C}$, so dass

$$\frac{P(z)}{Q(z)} = \sum_{j=1}^{k} \left(\frac{a_{1j}}{z-\lambda_j} + \frac{a_{2j}}{(z-\lambda_j)^2} + \ldots + \frac{a_{n_j j}}{(z-\lambda_j)^{n_j}} \right)$$

gilt.

Sind P, Q reelle Polynome, dann ergibt sich aus der komplexen Partialbruchzerlegung auch die reelle Partialbruchzerlegung von $\frac{P}{Q}$:

Besitzt Q die Zerlegung

$$Q(x) = a \prod_{j=1}^{m} (x - \lambda_j)^{n_j} \prod_{l=1}^{r} (x^2 + b_l x + c_l)^{p_l}$$

mit $a, b_l, c_l \in \mathbb{R}$, paarweise verschiedenen $\lambda_j \in \mathbb{R}$ und quadratischen Polynomen $x^2+b_lx+c_l$, die keine reelle Nullstellen haben (nach Satz 6.12 existiert für jedes reelle Polynom eine solche Zerlegung), dann existieren eindeutig bestimmte $A_{ij}, B_{il}, C_{il} \in \mathbb{R}$, so dass

$$\begin{aligned}\frac{P(x)}{Q(x)} \;&=\; \sum_{j=1}^{m}\left(\frac{A_{1j}}{x-\lambda_j}+\frac{A_{2j}}{(x-\lambda_j)^2}+\ldots+\frac{A_{n_jj}}{(x-\lambda_j)^{n_j}}\right) \qquad (*)\\ &+\sum_{l=1}^{r}\left(\frac{B_{1l}x+C_{1l}}{x^2+b_lx+c_l}+\frac{B_{2l}x+C_{2l}}{(x^2+b_lx+c_l)^2}+\ldots+\frac{B_{p_ll}x+C_{p_ll}}{(x^2+b_lx+c_l)^{p_l}}\right)\end{aligned}$$

gilt.

3. Schritt:

Bestimme die Zerlegung von $\tilde{q}$ in reelle Linearfaktoren und reelle quadratische Polynome ohne reelle Nullstellen.
Dann bestimme die reelle Partialbruchzerlegung von $\frac{\tilde{r}}{\tilde{q}}$.
Die Koeffizienten A_{ij}, B_{il}, C_{il} in (*) für P $=\tilde{r}$ und Q $=\tilde{q}$ kann man zum Beispiel mit folgender Methode berechnen.
Multipliziert man die Gleichung (*) mit $(x-\lambda_s)^{n_s}$, dann erhält man

$$\frac{P(x)}{\left(\frac{Q(x)}{(x-\lambda_s)^{n_s}}\right)} = A_{n_ss} + (x-\lambda_s)(...).$$

Setzt man in diese Gleichung λ_s für x ein, dann erhält man aus der dadurch entstehenden Gleichung den Wert von A_{n_ss}. Nachdem man auf diese Weise die Koeffizienten A_{n_jj} bestimmt hat, multipliziert man (*) mit $Q(x)$ und ermittelt die restlichen Koeffizienten durch Vergleich der Koeffizienten von $P(x)$ mit den Koeffizienten des Polynoms, das auf der rechten Seite von (*) nach der Multiplikation mit $Q(x)$ entstanden ist.

4. Schritt:

Integriere n sowie die einzelnen Summanden der Partialbruchzerlegungen $\frac{\tilde{r}}{\tilde{q}}$. Dabei sind Integrale der folgenden Form zu berechnen:

1) $\displaystyle\int \frac{A}{(x-\lambda)^n}\,dx\,, \quad n \in \mathbb{N}$

2) $\displaystyle\int \frac{Bx+C}{(x^2+bx+c)^n}\,dx\,, \quad n \in \mathbb{N}$

Es gilt:

$$\int \frac{A}{(x-\lambda)^n}\,dx \;=\; \begin{cases} A\,\ln|x-\lambda|+k, & n=1 \\ \dfrac{A}{1-n}\dfrac{1}{(x-\lambda)^{n-1}}+k, & n>1 \end{cases}\,, \qquad k \in \mathbb{R},$$

$$\int \frac{Bx+C}{(x^2+bx+c)^n}\,dx = \frac{B}{2}\int \frac{2x+b}{(x^2+bx+c)^n}\,dx + \Big(C-\frac{Bb}{2}\Big)\int \frac{1}{(x^2+bx+c)^n}\,dx\,,$$

$$\int \frac{2x+b}{(x^2+bx+c)^n}\,dx \;=\; \begin{cases} \ln|x^2+bx+c|+k, & n=1 \\ \dfrac{1}{1-n}\dfrac{1}{(x^2+bx+c)^{n-1}}+k, & n>1 \end{cases}, \quad k\in\mathbb{R}.$$

Es bleibt nun noch die Berechnung von

$$\int \frac{1}{(x^2+bx+c)^n}\,dx\,, \quad n\in\mathbb{N},$$

wobei x^2+bx+c keine reellen Nullstellen hat und somit $D:=b^2-4c<0$ ist. Durch quadratische Ergänzung erhält man

$$x^2+bx+c = \Big(x+\frac{b}{2}\Big)^2 + c - \frac{b^2}{4} = K^{-2}\Big(1+\Big(K\Big(x+\frac{b}{2}\Big)\Big)^2\Big)$$

mit $K=\frac{2}{\sqrt{-D}}>0$. Daher gilt

$$\int \frac{1}{(x^2+bx+c)^n}\,dx = \int K^{2n}\frac{1}{(1+(K(x+\frac{b}{2}))^2)^n}\,dx = K^{2n-1}\int \frac{1}{(1+t^2)^n}\,dt$$

mit $t=K(x+\frac{b}{2})$. Sei

$$I_n := \int \frac{1}{(1+t^2)^n}\,dt\,, \quad n\in\mathbb{N}.$$

Dann gilt

$$I_1 = \arctan t + k, \quad k\in\mathbb{R},$$

und für alle $n>1$:

$$\begin{aligned} I_n &= I_{n-1} - \int \frac{t^2}{(1+t^2)^n}\,dt \\ &= I_{n-1} - \int \frac{t}{2}\cdot\frac{2t}{(1+t^2)^n}\,dt \\ &= I_{n-1} + \frac{1}{2(n-1)}\cdot\frac{t}{(1+t^2)^{n-1}} + k - \frac{1}{2(n-1)}I_{n-1}\,, \quad k\in\mathbb{R}, \\ &= \frac{2n-3}{2n-2}\,I_{n-1} + \frac{1}{2n-2}\cdot\frac{t}{(1+t^2)^{n-1}} + k\,, \quad k\in\mathbb{R}. \end{aligned}$$

Somit ergibt sich für alle $b,c\in\mathbb{R}$ mit $b^2<4c$:

$$\int \frac{1}{x^2+bx+c}\,dx = \frac{1}{\sqrt{c-\frac{b^2}{4}}}\;\arctan\left(\frac{x+\frac{b}{2}}{\sqrt{c-\frac{b^2}{4}}}\right) + k\,, \quad k\in\mathbb{R},$$

sowie

$$\int \frac{1}{(x^2+bx+c)^n}\,dx = \frac{1}{(n-1)(4c-b^2)}\cdot\frac{2x+b}{(x^2+bx+c)^{n-1}} + \frac{2(2n-3)}{(n-1)(4c-b^2)}\int \frac{1}{(x^2+bx+c)^{n-1}}\,dx\,,$$

falls $n \in \mathbb{N}\setminus\{1\}$ ist.

Beispiele

1) $\displaystyle\int \frac{4x^2+9x-4}{(x-1)(x+2)^2}\,dx$

1. Schritt:

Nicht notwendig, da $Grad\,(4x^2+9x-4) < Grad\,((x-1)(x+2)^2)$.

2. Schritt:

Da weder 1 noch -2 eine Nullstelle von $4x^2+9x-4$ ist, sind $4x^2+9x-4$ und $(x-1)(x+2)^2$ bereits teilerfremd, so dass nichts zu kürzen ist.

3. Schritt:

Die für die Partialbruchzerlegung notwendige Zerlegung des Nennerpolynoms liegt bereits vor. Der Ansatz für die Partialbruchzerlegung in der vorliegenden Situation ist:

$$\frac{4x^2+9x-4}{(x-1)(x+2)^2} = \frac{a}{x-1} + \frac{b}{x+2} + \frac{c}{(x+2)^2}$$

Multiplikation mit $(x-1)$ liefert

$$\frac{4x^2+9x-4}{(x+2)^2} = a + (x-1)\Big(\frac{b}{x+2} + \frac{c}{(x+2)^2}\Big)$$

Einsetzen von $x=1$ liefert

$$a = 1$$

Multiplikation mit $(x+2)^2$ liefert

$$\frac{4x^2+9x-4}{(x-1)} = c + (x+2)^2\Big(\frac{a}{x-1} + \frac{b}{x+2}\Big)$$

Einsetzen von $x=-2$ liefert

$$c = 2$$

Multiplikation mit dem Nenner liefert

$$\begin{aligned}4x^2+9x-4 &= a(x+2)^2 + b(x-1)(x+2) + c(x-1)\\ &= (x+2)^2 + b(x-1)(x+2) + 2(x-1)\\ &= (1+b)x^2 + (6+b)x + (2-2b)\end{aligned}$$

Koeffizientenvergleich liefert

$$b = 3$$

4. Schritt:

$$\begin{aligned}
\int \frac{4x^2+9x-4}{(x-1)(x+2)^2}\,dx &= \int \frac{1}{x-1}\,dx + 3\int \frac{1}{x+2}\,dx + 2\int \frac{1}{(x+2)^2}\,dx \\
&= \ln|x-1| + 3\ln|x+2| - \frac{2}{x+2} + C, \quad C \in \mathbb{R}
\end{aligned}$$

2) $\displaystyle\int \frac{x^4+x^3+x}{x^3+1}\,dx$

1. Schritt:

$$\frac{x^4+x^3+x}{x^3+1} = x+1-\frac{1}{x^3+1}$$

2. Schritt:

-1 und x^3+1 sind bereits teilerfremd.

3. Schritt:

-1 ist Nullstelle von x^3+1,

$$(x^3+1):(x+1) = x^2-x+1$$

$$\Rightarrow x^3+1 = (x+1)(x^2-x+1)$$

Ansatz für die Partialbruchzerlegung von $-\dfrac{1}{x^3+1}$:

$$-\frac{1}{x^3+1} = \frac{a}{x+1} + \frac{Bx+C}{x^2-x+1}$$

Multiplikation mit $(x+1)$ liefert

$$-\frac{1}{x^2-x+1} = a + (x+1)\frac{Bx+C}{x^2-x+1}$$

Einsetzen von $x=-1$ liefert

$$a = -\frac{1}{3}$$

Multiplikation mit dem Nenner liefert

$$\begin{aligned}
-1 &= a(x^2-x+1) + (x+1)(Bx+C) \\
&= -\frac{1}{3}(x^2-x+1) + (x+1)(Bx+C) \\
&= \left(B-\frac{1}{3}\right)x^2 + \left(B+C+\frac{1}{3}\right)x + C - \frac{1}{3}
\end{aligned}$$

Koeffizientenvergleich liefert

$$B = \frac{1}{3},\ C = -\frac{2}{3}$$

4. Schritt:

$$\begin{aligned}
\int \frac{x^4+x^3+x}{x^3+1}\,dx &= \int x+1\,dx - \frac{1}{3}\int \frac{1}{x+1}\,dx + \frac{1}{3}\int \frac{x-2}{x^2-x+1}\,dx \\
&= \frac{1}{2}x^2 + x - \frac{1}{3}\ln|x+1| + C + \frac{1}{6}\int \frac{2x-1}{x^2-x+1}\,dx \\
&\quad -\frac{1}{2}\int \frac{1}{x^2-x+1}\,dx\,, \quad C \in \mathbb{R} \\
&= \frac{1}{2}x^2 + x - \frac{1}{3}\ln|x+1| + \frac{1}{6}\ln|x^2-x+1| \\
&\quad -\frac{1}{2}\sqrt{\frac{4}{3}}\,\arctan\left(\sqrt{\frac{4}{3}}\left(x-\frac{1}{2}\right)\right) + C\,, \quad C \in \mathbb{R}
\end{aligned}$$

3) $\displaystyle\int \frac{2x^2-x-1}{(x^2+1)^2(x-1)}\,dx$

1. Schritt:

Nicht notwendig, da $Grad\,(2x^2-x-1) < Grad\,((x^2+1)^2(x-1))$.

2. Schritt:

1 ist Nennernullstelle und Zählernullstelle, somit muss mit $(x-1)$ gekürzt werden. Dies liefert

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

$2x+1$ und $(x^2+1)^2$ sind bereits teilerfremd.

3. Schritt:

Da (x^2+1) keine reellen Nullstellen hat und wegen

$$\frac{2x+1}{(x^2+1)^2} = \frac{0}{(x^2+1)} + \frac{2x+1}{(x^2+1)^2}$$

ist die gewünschte Zerlegung bereits gegeben.

4. Schritt:

$$
\begin{aligned}
\int \frac{2x^2 - x - 1}{(x^2+1)^2(x-1)}\, dx &= \int \frac{2x+1}{(x^2+1)^2}\, dx \\
&= \int \frac{2x}{(x^2+1)^2}\, dx + \int \frac{1}{(x^2+1)^2}\, dx \\
&= -\frac{1}{x^2+1} + \frac{1}{4} \cdot \frac{2x}{x^2+1} + \frac{1}{2} \int \frac{1}{x^2+1}\, dx \\
&= \frac{1}{2}\Big(\frac{x-2}{x^2+1} + \arctan x \Big) + C, \quad C \in \mathbb{R}
\end{aligned}
$$

Bemerkung
Sei $R(\sin x, \cos x)$ eine Funktion, die aus einer rationalen Funktion $\frac{P(x)}{Q(x)}$ entsteht, wenn man x durch $\sin x$ oder $\cos x$ ersetzt, zum Beispiel

$$
\frac{\sin^2 x + 3 \sin x \cos x - \sin x + 1}{(\sin x - 1)^2 (1 + \cos^2 x)}.
$$

Mit Hilfe der Substitution $x = 2 \arctan t$ erhält man

$$
\int R(\sin x, \cos x)\, dx = \int \frac{q(t)}{h(t)}\, dt\,,
$$

wobei q, h Polynome sind. Somit steht die Methode der Partialbruchzerlegung auch als Hilfsmittel zur Berechnung von

$$
\int R(\sin x, \cos x)\, dx
$$

zur Verfügung.

Bemerkung
Für eine Übersicht über weitere Funktionen, deren Stammfunktionen in geschlossener Form, d.h. als Summe, Differenz, Produkt, Quotient, Potenz und Verkettung von endlich vielen elementaren Funktionen wie 1, x, e^x, $\sin x$, $\ln x$ dargestellt werden können, samt den zugehörigen Integrationsmethoden sei auf die Literatur verwiesen, unter anderem auf Integraltafeln. Für einige Funktionen wie z.B.

$$
\frac{\sin x}{x}, \quad \frac{1}{\sqrt{1+x^4}}, \quad e^{-x^2},
$$

kann man allerdings zeigen, dass ihre Stammfunktion nicht in geschlossener Form darstellbar sind.

Uneigentliche Integrale

Definition

a) *Sei I ein offenes oder halboffenes Intervall. $f : I \to \mathbb{R}$ heißt lokalintegrierbar (über I), falls f über jedem beschränkten und abgeschlossenen Intervall $[a,b] \subseteq I$ integrierbar ist.*

b) *Sei $I = [a,\infty[$ bzw $I =]-\infty, b]$ bzw $I = \mathbb{R}$ bzw $I =]a,b]$ bzw $I = [a,b[$ bzw $I =]a,b[$ und f: $I \to \mathbb{R}$ lokalintegrierbar. f heißt uneigentlich integrierbar über I, falls die jeweiligen Grenzwerte*

$$\int_a^\infty f(x)\ dx := \lim_{s\to\infty} \int_a^s f(x)\ dx$$

$$\text{bzw. } \int_{-\infty}^b f(x)\ dx := \lim_{s\to-\infty} \int_s^b f(x)\ dx$$

$$\text{bzw. } \int_{-\infty}^\infty f(x)\ dx := \int_{-\infty}^c f(x)\ dx + \int_c^\infty f(x)\ dx$$

$$\text{bzw. } \int_a^b f(x)\ dx := \lim_{s\to a^+} \int_s^b f(x)\ dx$$

$$\text{bzw. } \int_a^b f(x)\ dx := \lim_{s\to b^-} \int_a^s f(x)\ dx$$

$$\text{bzw. } \int_a^b f(x)\ dx := \int_a^c f(x)\ dx + \int_c^b f(x)\ dx$$

existieren, wobei der dritte und der letzte Grenzwert aufgrund von Satz 18.10 unabhängig von der Wahl von $c \in I$ sind.
Ist f uneigentlich integrierbar über I, dann sagt man auch, das uneigentliche Integral $\int_a^\infty f(x)\ dx$ bzw. $\int_{-\infty}^b f(x)\ dx$ bzw. $\int_{-\infty}^\infty f(x)\ dx$ bzw. $\int_a^b f(x)\ dx$ konvergiert (ist konvergent). Anderenfalls sagt man, das jeweilige uneigentliche Integral divergiert (ist divergent).

Beispiele

1) f mit $f(x) = \dfrac{1}{x^r}$, $r > 0$, ist lokalintegrierbar über $]0,1]$, über $[1,\infty[$ und über $]0,\infty[$.

Sei $0 < r < 1$. Dann gilt

$$\int_s^1 \frac{1}{x^r}\ dx = \frac{1}{1-r}\frac{1}{x^{r-1}}\Big|_s^1 = \frac{1}{1-r} - \frac{1}{1-r}s^{1-r} \overset{s\to 0^+}{\longrightarrow} \frac{1}{1-r},$$

$$\int_1^s \frac{1}{x^r}\ dx = \frac{1}{1-r}s^{1-r} - \frac{1}{1-r} \overset{s\to\infty}{\longrightarrow} \infty.$$

Sei $r = 1$. Dann gilt

$$\int_s^1 \frac{1}{x}\, dx \;=\; \ln 1 - \ln s \;=\; -\ln s \overset{s\to 0^+}{\longrightarrow} \infty\,,$$

$$\int_1^s \frac{1}{x}\, dx \;=\; \ln s \overset{s\to\infty}{\longrightarrow} \infty\,.$$

Sei $r > 1$. Dann gilt

$$\int_s^1 \frac{1}{x^r}\, dx \;=\; -\frac{1}{r-1} + \frac{1}{r-1} s^{1-r} \overset{s\to 0^+}{\longrightarrow} \infty\,,$$

$$\int_1^s \frac{1}{x^r}\, dx \;=\; -\frac{1}{r-1} s^{1-r} + \frac{1}{r-1} \overset{s\to\infty}{\longrightarrow} \frac{1}{r-1}.$$

Also: Für $0 < r < 1$ ist f uneigentlich integrierbar über $]0,1]$ mit

$$\int_0^1 \frac{1}{x^r}\, dx \;=\; \frac{1}{1-r}\,,$$

aber nicht über $[1,\infty[$, für $r > 1$ ist f uneigentlich integrierbar über $[1,\infty[$ mit

$$\int_1^\infty \frac{1}{x^r}\, dx \;=\; \frac{1}{r-1}\,,$$

aber nicht über $]0,1]$ und für $r = 1$ ist f weder über $]0,1]$ noch über $[1,\infty[$ uneigentlich integrierbar. Über $]0,\infty[$ ist f für kein $r > 0$ uneigentlich integrierbar.

2) $$\int_{-1}^1 \frac{1}{\sqrt{1-x^2}}\, dx \;=\; \lim_{s\to -1^+} \int_s^0 \frac{1}{\sqrt{1-x^2}}\, dx + \lim_{s\to 1^-} \int_0^s \frac{1}{\sqrt{1-x^2}}\, dx$$

$$= \lim_{s\to -1^+} \arcsin x \Big|_s^0 + \lim_{s\to 1^-} \arcsin x \Big|_0^s \;=\; \pi$$

3) $$\int_{-\infty}^\infty \frac{1}{1+x^2}\, dx \;=\; 2\int_0^\infty \frac{1}{1+x^2}\, dx \;=\; 2 \lim_{s\to\infty} \arctan x \Big|_0^s \;=\; \pi$$

4) $$\int_0^\infty e^{-x}\, dx \;=\; \lim_{s\to\infty} -e^{-x} \Big|_0^s \;=\; 1$$

Satz 18.17 (Konvergenzkriterien für uneigentliche Integrale) *Sei $I = [a,\infty[$ und $f : I \to \mathbb{R}$ lokalintegrierbar über I. Dann gilt:*

a) (Cauchy-Kriterium)
f ist genau dann uneigentlich integrierbar über I, wenn für jede Folge $(s_n)_{n\in\mathbb{N}} \subseteq I$ mit $s_n \to \infty$ für $n \to \infty$ die Folge $(S_n)_{n\in\mathbb{N}}$ mit $S_n := \int_a^{s_n} f(x)\, dx$ eine Cauchy-Folge ist.

b) *Ist das uneigentliche Integral von f über I absolut konvergent, d.h., das uneigentliche Integral von $|f|$ über I ist konvergent, dann ist das uneigentliche Integral von f über I auch konvergent.*

c) *Das uneigentliche Integral von f über I ist genau dann absolut konvergent, wenn gilt:*

$$\exists C > 0 \forall [\alpha, \beta] \subseteq I : \int_\alpha^\beta |f(x)| \, dx \leq C$$

d) *(Majorantenkriterium)*
Existiert ein $c \in I$ und eine Funktion $g : [c, \infty[\to \mathbb{R}$, die über $[c, \infty[$ uneigentlich integrierbar ist und für die gilt, dass $|f(x)| \leq g(x)$ für alle $x \in [c, \infty[$ ist, dann ist das uneigentliche Integral von f über I absolut konvergent und es gilt

$$\Big| \int_c^\infty f(x) \, dx \Big| \leq \int_c^\infty |f(x)| \, dx \leq \int_c^\infty g(x) \, dx$$

e) *(Minorantenkriterium)*
Existiert ein $c \in I$ und eine lokalintegrierbare Funktion $g : [c, \infty[\to \mathbb{R}$, die über $[c, \infty[$ nicht uneigentlich integrierbar ist und für die gilt, dass $0 \leq g(x) \leq f(x)$ für alle $x \in [c, \infty[$ ist, dann ist das uneigentliche Integral von f über I divergent.

Ersetzt man $I = [a, \infty[$ durch $I = [a, b[\,,\ I = \,]-\infty, b]$ oder $I = \,]a, b]\,$, dann übertragen sich die Kriterien $a) - e)$ sinngemäß.

Beweisstrategie
Mit ähnlichen Argumenten wie in den Beweisen der entsprechenden Konvergenzkriterien für Reihen, z.B. bei $d)$:
Seien die Annahmen von $d)$ an f und g erfüllt. Dann gilt nach $a)$:

Für $(s_n)_{n \in \mathbb{N}} \subseteq [c, \infty[$ mit $s_n \to \infty$ für $n \to \infty$ ist $\displaystyle\int_c^{s_n} g(x) \, dx$ eine Cauchy-Folge

$\Rightarrow \forall \varepsilon > 0 \exists n_\varepsilon \forall m > n > n_\varepsilon : \Big| \displaystyle\int_{s_n}^{s_m} g(x) \, dx \Big| < \varepsilon$

$\overset{g(x) \geq 0}{\Longrightarrow} \forall \varepsilon > 0 \exists n_\varepsilon \forall m > n > n_\varepsilon : \displaystyle\int_{s_n}^{s_m} g(x) \, dx < \varepsilon$

$\Rightarrow \forall \varepsilon > 0 \exists n_\varepsilon \forall m > n > n_\varepsilon : \Big| \displaystyle\int_{s_n}^{s_m} f(x) \, dx \Big| < \int_{s_n}^{s_m} |f(x)| \, dx \leq \int_{s_n}^{s_m} g(x) \, dx < \varepsilon$

$\Rightarrow \displaystyle\int_c^\infty f(x) \, dx$ konvergiert absolut

Da f über $[a, \infty[$ lokalintegrierbar ist, konvergiert auch

$\displaystyle\int_a^\infty f(x) \, dx = \int_a^c f(x) \, dx + \int_c^\infty f(x) \, dx$

absolut.

Beispiele

1) Wegen $\frac{e^x}{x^3} > \frac{1}{x^3} > 0$ für alle $x > 0$ und der Divergenz von $\int_0^1 \frac{1}{x^3}\,dx$ divergiert $\int_0^1 \frac{e^x}{x^3}\,dx$ nach dem Minorantenkriterium.

2) Wegen $\left|\frac{\cos x}{1+x+x^2}\right| \le \frac{1}{x^2}$ für alle $x > 1$ und der Konvergenz von $\int_1^\infty \frac{1}{x^2}\,dx$ und da $x \mapsto \frac{\cos x}{1+x+x^2}$ über $[0,\infty[$ lokalintegrierbar ist, konvergiert $\int_0^\infty \frac{\cos x}{1+x+x^2}\,dx$ absolut nach dem Majorantenkriterium.

3) Nach dem Satz von Taylor gilt
$|\cos x - 1| = O(x^2)$ für $x \to 0$
$\Rightarrow \left|\frac{\cos x - 1}{x^{5/2}}\right| = O\Big(\frac{1}{\sqrt{x}}\Big)$ für $x \to 0$.
Da außerdem gilt, dass $x \mapsto \frac{\cos x - 1}{x^{5/2}}$ über $]0,1]$ lokalintegrierbar ist und $\int_0^1 \frac{1}{\sqrt{x}}\,dx$ konvergiert, konvergiert $\int_0^1 \frac{\cos x - 1}{x^{5/2}}\,dx$ absolut nach dem Majorantenkriterium.

4) $x \mapsto \frac{\sin x}{x}$ ist stetig auf $]0,\infty[$ und es gilt $\lim\limits_{x\to 0^+} \frac{\sin x}{x} = 1$. Daher konvergiert $\int_0^1 \frac{\sin x}{x}\,dx$.
Sei $(s_n)_{n\in\mathbb{N}} \subseteq [1,\infty[$ mit $s_n \to \infty$ für $n \to \infty$. Dann gilt für alle $m \ge n$:
$$\begin{aligned}\Big|\int_{s_n}^{s_m} \frac{\sin x}{x}\,dx\Big| &\le \Big|-\frac{\cos s_m}{s_m} + \frac{\cos s_n}{s_n}\Big| + \int_{s_n}^{s_m} \Big|\frac{\cos x}{x^2}\Big|\,dx \\ &\le \frac{1}{s_m} + \frac{1}{s_n} + \int_{s_n}^{s_m} \frac{1}{x^2}\,dx \\ &= \frac{2}{s_n} \to 0 \quad \text{für } n \to \infty.\end{aligned}$$
Daher folgt nach dem Cauchy-Kriterium die Konvergenz von $\int_1^\infty \frac{\sin x}{x}\,dx$.
Insgesamt erhalten wir, dass $\int_0^\infty \frac{\sin x}{x}\,dx$ konvergiert.

Satz 18.18 (Integralkriterium für Reihen) *Sei $f : [1, \infty[\to \mathbb{R}_0^+$ monoton fallend. Dann gilt:*

$$\sum_{k=1}^{\infty} f(n) \text{ konvergiert} \quad \Longleftrightarrow \quad \int_1^{\infty} f(x)\ dx \text{ konvergiert.}$$

Beweis

"$\Leftarrow$": Sei $\displaystyle\int_1^{\infty} f(x)\ dx$ konvergent.

$\Rightarrow \forall m \geq 2 : \displaystyle\sum_{n=2}^{m} f(n) \leq \int_1^{m} f(x)\ dx \leq \int_1^{\infty} f(x)\ dx$

$\Rightarrow \left(\displaystyle\sum_{n=1}^{m} f(n) \right)_{m \in \mathbb{N}}$ ist eine beschränkte Folge.

Wegen $f(n) \geq 0$ für alle $n \in \mathbb{N}$ und Satz 14.7 folgt die Konvergenz von $\displaystyle\sum_{n=1}^{\infty} f(n)$.

"$\Rightarrow$": Sei $\displaystyle\sum_{n=1}^{\infty} f(n)$ konvergent und sei $\forall n \in \mathbb{N}\, \forall x \in [n, n+1[\, :\ g(x) := f(n)$.

$\Rightarrow 0 \leq f(x) \leq g(x) \ \wedge\ \displaystyle\int_1^{\infty} g(x)\ dx\ =\ \sum_{n=1}^{\infty} f(n)$ ist konvergent.

Nach dem Majorantenkriterium folgt die Konvergenz von $\displaystyle\int_1^{\infty} f(x)\ dx$. $\square$

Beispiel

$\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^{\alpha}}$ ist für alle $\alpha > 1$ konvergent und absolut konvergent und für alle $\alpha \leq 1$ divergent.

Definition

a) Sei $f :]a, c[\cup]c, b[\to \mathbb{R}$ lokalintegrierbar über $]a, c[$ und über $]c, b[$. Existiert der Grenzwert

$$CH \int_a^b f(x)\ dx =: \lim_{\varepsilon \to 0^+} \left(\int_a^{c-\varepsilon} f(x)\ dx + \int_{c+\varepsilon}^{b} f(x)\ dx \right),$$

dann heißt er der Cauchysche Hauptwert von f über $[a, b]$.

b) Sei $f : \mathbb{R} \to \mathbb{R}$ lokalintegrierbar über $\mathbb{R}$. Existiert der Grenzwert

$$CH \int_{-\infty}^{\infty} f(x)\ dx =: \lim_{s \to \infty} \int_{-s}^{s} f(x)\ dx,$$

dann heißt er der Cauchysche Hauptwert von f über $\mathbb{R}$.

Beispiele

1) $CH \int_{-1}^{1} \frac{1}{x}\, dx = \lim_{\varepsilon \to 0^+} (\ln \varepsilon - \ln \varepsilon) = 0$

2) $CH \int_{-1}^{1} \frac{1}{|x|}\, dx$ existiert nicht.

3) $CH \int_{-\infty}^{\infty} \frac{2x}{x^2+1}\, dx = \lim_{s \to \infty} (\ln |s^2+1| - \ln |(-s)^2+1|) = 0$

4) $CH \int_{-\infty}^{\infty} \frac{2|x|}{x^2+1}\, dx$ existiert nicht.

Vertauschbarkeit von Grenzübergängen, Differentation und Integration

Satz 18.19 *Sei $(f_n)_{n\in\mathbb{N}_0}$ eine Folge von Funktionen $f_n : \mathbb{R} \supseteq D \to \mathbb{R}$, die stetig auf D seien. Konvergiert $(f_n)_{n\in\mathbb{N}_0}$ gleichmäßig auf D gegen $f : D \to \mathbb{R}$, dann ist f ebenfalls stetig auf D und folglich gilt für jeden Häufungspunkt x_0 von D:*

$$\lim_{x\to x_0}\left(\lim_{n\to\infty} f_n(x)\right) = \lim_{n\to\infty}\left(\lim_{x\to x_0} f_n(x)\right) = f(x_0).$$

Beweis
Sei $x_0 \in D$ und $\varepsilon > 0$. Wenn $(f_n)_{n\in\mathbb{N}_0}$ gleichmäßig auf D gegen f konvergiert, dann
$\exists n_\varepsilon \in \mathbb{N} : \sup_{x\in D} |f_{n_\varepsilon}(x) - f(x)| \leq \frac{\varepsilon}{3}$.
Da f_{n_ε} stetig auf D ist,
$\exists \delta > 0\, \forall x \in D : |x - x_0| < \delta \;\Rightarrow\; |f_{n_\varepsilon}(x) - f_{n_\varepsilon}(x_0)| \leq \frac{\varepsilon}{3}$.
Folglich gilt dann
$\forall x \in D : |x - x_0| < \delta \;\Rightarrow\; |f(x) - f(x_0)| \leq |f(x) - f_{n_\varepsilon}(x)| + |f_{n_\varepsilon}(x) - f_{n_\varepsilon}(x_0)| +$
$|f_{n_\varepsilon}(x_0) - f(x_0)| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon$,
woraus die Stetigkeit von f auf D folgt. $\square$

Bemerkung
Konvergiert eine Folge von auf D stetigen Funktionen $(f_n)_{n\in\mathbb{N}_0}$ nur punktweise gegen f, dann muss f nicht unbedingt stetig auf D sein. Denn sei $f_n : [0,1] \to \mathbb{R}, x \mapsto x^n$. Dann gilt $f_n \in C^0([0,1])$ und $(f_n)_{n\in\mathbb{N}_0}$ konvergiert punktweise auf $[0,1]$ gegen

$$f : [0,1] \to \mathbb{R}, x \mapsto \begin{cases} 0, & x \neq 1, \\ 1, & x = 1. \end{cases}$$

f ist nicht stetig in 1.

Satz 18.20 *Sei* $(f_n)_{n\in\mathbb{N}_0} \subseteq C^0([a,b])$ *und* $(f_n)_{n\in\mathbb{N}_0}$ *konvergiere gleichmäßig auf* $[a,b]$ *gegen* $f\colon [a,b] \to \mathbb{R}$. *Dann gilt*

$$\int_a^b f(x)\ dx = \lim_{n\to\infty} \int_a^b f_n(x)\ dx.$$

Beweis
Nach Satz 18.19 ist $f \in C^0([a,b])$ und somit integrierbar über $[a,b]$. Wegen Satz 18.9 und dem Satz vom Maximum und Minimum gilt

$$\begin{aligned} \Big| \int_a^b f(x)\ dx - \int_a^b f_n(x)\ dx \Big| &\leq \int_a^b |f(x) - f_n(x)|\ dx \\ &\leq (b-a) \sup_{x\in[a,b]} |f(x) - f_n(x)| \stackrel{n\to\infty}{\longrightarrow} 0. \end{aligned}$$ □

Bemerkung
Konvergiert $(f_n)_{n\in\mathbb{N}_0} \subseteq C^0([a,b])$ nur punktweise auf $[a,b]$ gegen f, dann muss die Aussage aus Satz 18.20 nicht unbedingt gelten. Denn sei

$$f_n : [0,1] \to \mathbb{R}, \quad x \mapsto \begin{cases} n^2 x\,, & 0 \leq x \leq \frac{1}{n}, \\ 2n - n^2 x\,, & \frac{1}{n} < x \leq \frac{2}{n}, \\ 0\,, & \text{sonst.} \end{cases}$$

Dann gilt $f_n \in C^0([0,1])$ und $(f_n)_{n\in\mathbb{N}_0}$ konvergiert punktweise auf $[0,1]$ gegen $f \equiv 0$. f ist in diesem Fall sogar stetig, aber es gilt

$$\lim_{n\to\infty} \int_0^1 f_n(x)\ dx = \lim_{n\to\infty} 1 = 1 \neq 0 = \int_0^1 f(x)\ dx.$$

Satz 18.21 *Sei* $(f_n)_{n\in\mathbb{N}_0} \subseteq C^1([a,b]), (f_n)_{n\in\mathbb{N}_0}$ *konvergiere punktweise auf [a,b] gegen f und* $(f_n')_{n\in\mathbb{N}_0}$ *konvergiere gleichmäßig auf* $[a,b]$ *gegen g. Dann ist* $f \in C^1([a,b])$ *mit*

$$f'(x) = g(x) = \lim_{n\to\infty} f_n'(x)$$

für alle $x \in\ ]a,b[$.

Beweis
Nach dem Hauptsatz gilt

$$\forall x \in [a,b]\, \forall n \in \mathbb{N}_0 : f_n(x) = f_n(a) + \int_a^x f_n'(t)\ dt.$$

Nach Satz 18.19 ist $g \in C^0([a,b])$ und nach Satz 18.20 gilt

$$\forall x \in [a,b] : \lim_{n\to\infty} \int_a^x f_n'(t)\ dt = \int_a^x g(t)\ dt.$$

Somit erhalten wir unter Ausnutzung der punktweisen Konvergenz von $(f_n)_{n\in\mathbb{N}_0}$ gegen f:

$\forall x \in [a,b] : f(x) = f(a) + \displaystyle\int_a^x g(t)\ dt$

Nach dem Hauptsatz folgt daraus $f'(x) = g(x)$ für alle $x \in\,]a,b[$. $\square$

Bemerkungen

1) Konvergiert $(f_n)_{n\in\mathbb{N}_0} \subseteq C^1([a,b])$ gleichmäßig auf $[a,b]$ gegen f, so muss f nicht unbedingt differenzierbar auf $]a,b[$ sein. Denn sei

$$f_n : [-1,1] \to \mathbb{R}, \quad x \mapsto \begin{cases} |x|\,, & x \in [-1, -\frac{1}{n+1}] \cup [\frac{1}{1+n}, 1]\,, \\ \frac{n+1}{2}x^2 + \frac{1}{2(n+1)}\,, & x \in\,]-\frac{1}{n+1}, \frac{1}{n+1}[\,. \end{cases}$$

Dann gilt $f_n \in C^1([-1,1])$ und $(f_n)_{n\in\mathbb{N}_0}$ konvergiert gleichmäßig auf $[-1,1]$ gegen f mit $f(x) = |x|$. f ist in 0 nicht differenzierbar.

2) Konvergiert $(f_n)_{n\in\mathbb{N}_0} \subseteq C^1([a,b])$ gleichmäßig auf $[a,b]$ gegen eine Funktion f aus $C^1([a,b])$ und $(f_n')_{n\in\mathbb{N}_0}$ (auf a und b stetig fortgesetzt) punktweise auf $[a,b]$ gegen g. Dann muss nicht unbedingt $f' = g$ sein. Denn sei $f_n : [0,1] \to \mathbb{R}\,, x \mapsto \frac{1}{n+1}x^{n+1}$. Dann ist $f_n \in C^1([0,1]), (f_n)_{n\in\mathbb{N}_0}$ konvergiert gleichmäßig auf $[0,1]$ gegen die C^1-Funktion $f \equiv 0$, aber es gilt $f_n'(x) = x^n$ für alle $x \in [0,1]$, so dass $(f_n')_{n\in\mathbb{N}_0}$ punktweise gegen $g : [0,1] \to \mathbb{R},\ x \mapsto \begin{cases} 0, & x \neq 1, \\ 1, & x = 1, \end{cases}$ konvergiert.

Satz 18.22 *Sei* $f(x) = \displaystyle\sum_{n=0}^{\infty} a_n(x-x_0)^n$ *eine reelle Potenzreihe mit Konvergenzradius* $R > 0$. *Dann gilt:*

(i) f *ist stetig auf* $B_R(x_0) = \{x \in \mathbb{R} : |x - x_0| < R\}$.

(ii) f *kann in* $B_R(x_0)$ *gliedweise integriert werden, d.h.*

$$\forall x \in B_R(x_0) : \int_{x_0}^{x} \sum_{n=0}^{\infty} a_n(t-x_0)^n\ dt \ = \sum_{n=0}^{\infty} \frac{a_n}{n+1}(x-x_0)^{n+1}.$$

Die gliedweise integrierte Potenzreihe $\displaystyle\sum_{n=0}^{\infty} \frac{a_n}{n+1}(x-x_0)^{n+1}$ *hat ebenfalls den Konvergenzradius R.*

(iii) f *kann in* $B_R(x_0)$ *gliedweise abgeleitet werden, d.h.*

$$\forall x \in B_R(x_0) : \frac{d}{dx}\sum_{n=0}^{\infty} a_n(x-x_0)^n = \sum_{n=1}^{\infty} na_n(x-x_0)^{n-1}.$$

Die gliedweise abgeleitete Potenzreihe $\displaystyle\sum_{n=1}^{\infty} na_n(x-x_0)^{n-1}$ *hat ebenfalls den Konvergenzradius R.*

(iv) $f \in C^\infty(B_R(x_0))$, *wobei*

$$\forall k \in \mathbb{N} \forall x \in B_R(x_0) : \frac{d^k}{dx^k} f(x) = \sum_{n=k}^{\infty} n(n-1) \cdot \ldots \cdot (n-k+1) a_n (x-x_0)^{n-k}.$$

Außerdem gilt:

$$\forall k \in \mathbb{N}_0 : a_k = \frac{1}{k!} f^{(k)}(x_0),$$

d.h., die Potenzreihendarstellung von f ist die Taylorreihe von f um x_0.

Beweis
Sei $x \in B_R(x_0)$. Dann existiert ein $\rho > 0$ mit $x \in [x_0 - \rho, x_0 + \rho]$ und nach Satz 17.7 *d)* konvergiert $\sum\limits_{n=0}^{\infty} a_n (x-x_0)^n$ gleichmäßig auf $[x_0 - \rho, x_0 + \rho]$. Daher folgt nach Satz 18.19 die Stetigkeit von f auf $B_R(x_0)$, nach Satz 18.20 die gliedweise Integrierbarkeit von f auf $B_R(x_0)$ und nach Satz 18.21 die gliedweise Differenzierbarkeit von f auf $B_R(x_0)$.
Die Aussagen über den Konvergenzradius der gliedweise integrierten und der gliedweise abgeleiteten Potenzreihe folgen aus $\sqrt[n]{|na_n|} = \sqrt[n]{n}\sqrt[n]{|a_n|}$, $\lim\limits_{n\to\infty} \sqrt[n]{n} = 1$ und der Formel von Cauchy-Hadamard (Satz 17.7 *b)*).
Die Aussage (iv) folgt aus (iii) durch vollständige Induktion und durch Anwendung von Korollar 17.9 . □

Bemerkung
Für Funktionenreihen, die keine Potenzreihen sind, müssen die Aussagen von Satz 18.22 nicht unbedingt gelten. Denn beispielsweise hat die Funktionenreihe
$\sum\limits_{n=1}^{\infty} \frac{1}{n^2} sin(n^2 x)$ die konvergente Majorante $\sum\limits_{n=1}^{\infty} \frac{1}{n^2}$ und konvergiert somit für alle $x \in \mathbb{R}$. Die gliedweise abgeleitete Reihe $\sum\limits_{n=1}^{\infty} \cos(n^2 x)$ konvergiert aber für kein $x \in \mathbb{R}$, da $(\cos(n^2 x))_{n \in \mathbb{N}}$ für kein $x \in \mathbb{R}$ eine Nullfolge ist, was man folgendermaßen zeigen kann.
Angenommen: $\lim\limits_{n\to\infty} \cos(n^2 x) = 0$. Dann gilt
$\forall k \in \mathbb{N} : \lim\limits_{n\to\infty} \cos((kn)^2 x) = 0$
und daher

$$\begin{aligned}
0 &= \lim_{n\to\infty} \cos((5n)^2 x) \\
&= \lim_{n\to\infty} \cos((3n)^2 x + (4n)^2 x) \\
&= \lim_{n\to\infty} \Big(\cos((3n)^2 x) \cos((4n)^2 x) - \sin((3n)^2 x) \sin((4n)^2 x) \Big) \\
&= \lim_{n\to\infty} \Big(\cos((3n)^2 x) \ \cos((4n)^2 x) - \sqrt{1 - (\cos((3n)^2 x))^2} \sqrt{1 - (\cos((4n)^2 x))^2} \Big) \\
&= -1 \quad \lightning .
\end{aligned}$$

Beispiele

1) Die geometrische Reihe hat Konvergenzradius 1. Daher erhalten wir nach Satz 18.22:
$$\frac{1}{(1-x)^2} = \frac{d}{dx}\left(\frac{1}{1-x}\right) = \frac{d}{dx}\sum_{n=0}^{\infty} x^n = \sum_{n=0}^{\infty}\frac{d}{dx}x^n = \sum_{n=1}^{\infty} nx^{n-1}, \qquad \text{für } \ |x|<1.$$
Für $|x| = 1$ divergiert $\sum\limits_{n=1}^{\infty} n\ x^{n-1}$.

2) Die Sinusreihe hat Konvergenzradius ∞. Folglich hat auch $\sum\limits_{n=0}^{\infty}\frac{(-1)^n}{(2n+1)!}\,x^{2n}$ $= \frac{\sin x}{x}$ Konvergenzradius ∞. Daher erhalten wir nach Satz 18.22 :
$$\begin{aligned}\int_0^x \frac{\sin t}{t}\,dt &= \int_0^x \sum_{n=0}^{\infty}\frac{(-1)^n}{(2n+1)!}t^{2n}\,dt = \sum_{n=0}^{\infty}\int_0^x \frac{(-1)^n}{(2n+1)!}t^{2n}\,dt \\ &= \sum_{n=0}^{\infty}\frac{(-1)^n}{(2n+1)(2n+1)!}\,x^{2n+1}, \qquad \text{für alle } x\in\mathbb{R}.\end{aligned}$$

3) Aufgrund des Konvergenzradius der geometrischen Reihe und Satz 18.22 erhalten wir:
$$\begin{aligned}\ln(1+x) &= \int_0^x \frac{1}{1+t}\,dt = \int_0^x \sum_{n=0}^{\infty}(-t)^n\,dt = \sum_{n=0}^{\infty}\int_0^x (-t)^n\,dt \\ &= \sum_{n=0}^{\infty}(-1)^n\,\frac{x^{n+1}}{n+1}, \qquad \text{für } \ |x|<1.\end{aligned}$$
Aufgrund des Leibniz-Kriteriums gilt außerdem:

$\sum\limits_{n=0}^{\infty}(-1)^n\,\frac{x^{n+1}}{n+1}$ konvergiert auch für $x=1$ und

$\forall x\in[0,1]\,\forall m\in\mathbb{N}: \left|\sum\limits_{n=0}^{\infty}(-1)^n\,\frac{x^{n+1}}{n+1} - \sum\limits_{n=0}^{m}(-1)^n\,\frac{x^{n+1}}{n+1}\right| \le \frac{x^{m+2}}{m+2} \le \frac{1}{m+2}.$

$\sum\limits_{n=0}^{\infty}(-1)^n\,\frac{x^{n+1}}{n+1}$ konvergiert also auch gleichmäßig auf $[0,1]$, so dass nach Satz 18.19 auch
$$\ln(2) = \sum_{n=0}^{\infty}(-1)^n\,\frac{1}{n+1}$$
gilt.

Für $x=-1$ dagegen divergiert $\sum\limits_{n=0}^{\infty}(-1)^n\,\frac{x^{n+1}}{n+1}$.

Teil IV

Mehrdimensionale Analysis

19 Konvergenz und Stetigkeit

Definition

a) Eine Folge $(a_k)_{k\in\mathbb{N}}$ *mit* $a_k \in \mathbb{R}^n$ *für alle* $k \in \mathbb{N}$ *heißt konvergent gegen* $a \in \mathbb{R}^n$*, in Zeichen:* $a_k \to a$ *für* $k \to \infty$ *oder* $\lim\limits_{k\to\infty} a_k = a$*, falls*

$$\lim_{k\to\infty} |a_k - a| = \lim_{k\to\infty} \|a_k - a\|_2 = \lim_{k\to\infty} \sqrt{\sum_{i=1}^{n} (a_k - a)_i^2} = 0$$

gilt.

b) Eine Folge $(a_k)_{k\in\mathbb{N}}$ *in* $\mathbb{R}^n$ *heißt Cauchy-Folge in* $\mathbb{R}^n$*, falls gilt*

$$\forall \varepsilon > 0 \ \exists k_\varepsilon \in \mathbb{N} \ \forall m, k > k_\varepsilon : |a_m - a_k| < \varepsilon$$

Satz 19.1 *Sei* $a \in \mathbb{R}^n$ *und* $a_k \in \mathbb{R}^n$ *für alle* $k \in \mathbb{N}$*. Dann ist äquivalent*

(i) $\lim\limits_{k\to\infty} |a_k - a| = 0$

(ii) $\lim\limits_{k\to\infty} \|a_k - a\|_\infty = \lim\limits_{k\to\infty} \max\limits_{i=1,\ldots,n} |\,(a_k - a)_i\,| = 0$

(iii) $\forall i \in \{1, \ldots, n\} : \lim\limits_{k\to\infty} (a_k)_i = a_i$

Außerdem gilt

$$(a_k)_{k\in\mathbb{N}} \text{ ist Cauchy-Folge in } \mathbb{R}^n \quad \Leftrightarrow \quad ((a_k)_i)_{k\in\mathbb{N}} \text{ sind Cauchy-Folgen in } \mathbb{R} \text{ für alle } i \in \{1, \ldots, n\}$$

Bemerkung
Folgende Aussagen über Folgen in $\mathbb{R}$ übertragen sich sinngemäß auf Folgen in $\mathbb{R}^n$: die Aussagen aus Satz 13.1, Satz 13.2*(i),(iv)*, Satz 13.8, Satz 13.9, Satz 13.10 und Satz 13.11. Außerdem gilt

$$\Big((\exists C > 0 \ \forall k \in \mathbb{N} : \ |a_k| \le C) \wedge \lim_{k\to\infty} a_k = a \Big) \Rightarrow |a| \le C.$$

Beispiel

$$\lim_{k\to\infty} \begin{pmatrix} \dfrac{2k^2}{k^2+1} \\ \dfrac{k^4}{k^6+5k} \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$$

Definition *Sei $D \subseteq \mathbb{R}^m$ und $f : D \to \mathbb{R}^n$ eine Funktion.*

a) f heißt stetig in $a_0 \in D$, falls $\lim\limits_{k\to\infty} f(a_k) = f(a_0)$ für jede Folge $(a_k)_{k\in\mathbb{N}}$ in D mit $\lim\limits_{k\to\infty} a_k = a_0$.

b) f heißt stetig auf der Menge $M \subseteq D$, falls f in jedem $a \in M$ stetig ist.

Bemerkungen

1) Die Aussagen für Funktionen $f : \mathbb{R} \supseteq D \to \mathbb{R}$ aus Satz 15.1, Satz 15.2, Korollar 15.3, Satz 15.4 *a) b) e)* übertragen sich sinngemäß auf Funktionen $f : \mathbb{R}^m \supseteq D \to \mathbb{R}^n$. Falls $n = 1$ ist, übertragen sich außerdem die Aussagen aus Satz 15.4 *c) d)*.

2) $f : \mathbb{R}^m \supseteq D \to \mathbb{R}^n$, $\begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \mapsto \begin{pmatrix} f_1(x_1,\ldots,x_m) \\ \vdots \\ f_n(x_1,\ldots,x_m) \end{pmatrix}$ ist genau dann stetig in $\begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}$, wenn die Funktionen $f_i : \mathbb{R}^m \supseteq D \to \mathbb{R}$ für alle $i \in \{1,\ldots,n\}$ in $\begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}$ stetig sind.

Beispiel

$$f : \mathbb{R}^3 \to \mathbb{R}^2\,, \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} \dfrac{x^2y + z^3}{1 + x^4 + y^6z^2} + 3x - 2y \\ e^{xy} sin\, z + \sqrt{1 + |y|} \end{pmatrix}$$

ist stetig auf ganz $\mathbb{R}^3$.

Bemerkung
Es gibt Funktionen $f : \mathbb{R}^m \supseteq D \to \mathbb{R}^n$, die in einem Punkt $a_0 \in D$ zwar stetig längs jeder Geraden sind, aber dennoch nicht stetig in a_0 sind. Denn sei zum Beispiel

$$f : \mathbb{R}^2 \to \mathbb{R}\,,\ f(x,y) = \begin{cases} \dfrac{xy^2}{x^2 + y^4}\,, & (x,y) \neq (0,0), \\ 0\,, & (x,y) = (0,0). \end{cases}$$

Dann gilt $\forall c \in \mathbb{R} : \lim\limits_{x\to 0} f(x, cx) = \lim\limits_{x\to 0} \dfrac{c^2x^3}{x^2 + c^4x^4} = 0 = f(0,0)$ sowie $\lim\limits_{y\to 0} f(0,y) = \lim\limits_{y\to 0} 0 = 0 = f(0,0)$, so dass f in $(0,0)$ längs jeder Geraden stetig ist.
Aber es gilt $\lim\limits_{x\to 0} f(x, \sqrt{x}) = \lim\limits_{x\to 0} \dfrac{x^2}{2x^2} = \dfrac{1}{2} \neq f(0,0)$, so dass $\lim\limits_{k\to\infty} \left(\dfrac{1}{k}, \dfrac{1}{\sqrt{k}}\right) = (0,0)$ und $\lim\limits_{k\to\infty} f\left(\dfrac{1}{k}, \dfrac{1}{\sqrt{k}}\right) = \dfrac{1}{2} \neq f(0,0)$ ist und daher f nicht stetig in $(0,0)$ ist.

20 Mehrdimensionale Differentialrechnung

Definition *Sei* $f : \mathbb{R} \supseteq D \to \mathbb{R}^n$, $x \mapsto \begin{pmatrix} f_1(x) \\ \vdots \\ f_n(x) \end{pmatrix}$ *eine Funktion.*

a) Sei x_0 ein innerer Punkt von D. f heißt differenzierbar in x_0, falls die folgenden äquivalenten Bedingungen erfüllt sind.

(i) $f'(x_0) := \lim\limits_{x \to x_0} \dfrac{f(x) - f(x_0)}{x - x_0} = \lim\limits_{h \to 0} \dfrac{f(x_0 + h) - f(x_0)}{h}$ *existiert.*

(ii) Für alle $i \in \{1, \ldots, n\}$ sind die Funktionen $f_i : \mathbb{R} \supseteq D \to \mathbb{R}$ differenzierbar in x_0.

In diesem Fall heißt dann $f'(x_0)$ die Ableitung von f in x_0. Statt f' kann man auch das Symbol $\dfrac{df}{dx}$ schreiben.

b) f heißt differenzierbar in $M \subseteq D$, falls f in jedem $x \in M$ differenzierbar ist. Die Funktion $f' : M \to \mathbb{R}^n$, $x \mapsto f'(x)$ heißt dann Ableitung von f.

Bemerkung
Die Aussagen aus dem Sätzen 16.1 und 16.3 übertragen sich sinngemäß auf Funktionen $f : \mathbb{R} \supseteq D \to \mathbb{R}^n$.

Definition *Sei* $f : \mathbb{R} \supseteq D \to \mathbb{R}^n$, $x \mapsto \begin{pmatrix} f_1(x) \\ \vdots \\ f_n(x) \end{pmatrix}$ *eine Funktion und $k \in \mathbb{N}$. f heißt k-mal differenzierbar / k-mal stetig differenzierbar / ∞-oft differenzierbar in einem inneren Punkt $x_0 \in D$ / in $M \subseteq D$, falls für alle $i \in \{1, \ldots, n\}$ die Funktionen f_i in x_0 / in M k-mal differenzierbar / k-mal stetig differenzierbar / ∞-oft differenzierbar sind.*
Die Symbole zur Bezeichnung der höheren Ableitungen von f werden direkt vom eindimensionalen Fall übernommen.
Außerdem seien

$$C^k(M, \mathbb{R}^n) := \left\{ f = \begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix} : M \to \mathbb{R}^n \ : \ \forall i \in \{1, \ldots, n\} : \ f_i \in C^k(M) \right\},$$

$$C^\infty(M, \mathbb{R}^n) := \left\{ f = \begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix} : M \to \mathbb{R}^n \ : \ \forall i \in \{1, \ldots, n\} : \ f_i \in C^\infty(M) \right\}.$$

Bemerkung
Aufgrund von Satz 16.3 sind $C^k(M,\mathbb{R}^n)$ und $C^\infty(M,\mathbb{R}^n)$ $\mathbb{R}$-Vektorräume und $\frac{d}{dx}: f \mapsto \frac{d}{dx}f$ ist eine lineare Abbildung von $C^k(M,\mathbb{R}^n)$ nach $C^{k-1}(M,\mathbb{R}^n)$ für $k \geq 1$.

Definition *Sei $I \subseteq \mathbb{R}$ ein Intervall. $\gamma \in C^1(I,\mathbb{R}^n)$ heißt reguläre parametrisierte Kurve oder reguläre Parametrisierung einer Kurve, falls $\frac{d}{dt}\gamma(t) \neq 0$ für alle $t \in I$ ist. Die Bildmenge $\gamma(I)$ heißt dann Spur der Kurve.*

Bemerkungen

1) Sei $\gamma \in C^1(I,\mathbb{R}^n)$ eine reguläre parametrisierte Kurve und $t_0 \in I$. Dann ist $\frac{d}{dt}\gamma(t_0)$ ein Tangentenvektor an $\gamma(I)$ im Punkt $\gamma(t_0)$, d.h., $\frac{d}{dt}\gamma(t_0)$ ist ein Richtungsvektor der Tangente an $\gamma(I)$ in $\gamma(t_0)$.

2) Sei $\gamma \in C^1(I,\mathbb{R}^n)$ eine reguläre parametrisierte Kurve. Interpretiert man $t \in I$ als Zeit und $\gamma(t)$ als Ort zum Zeitpunkt t, dann beschreibt γ die Bewegung eines Punktes im $\mathbb{R}^n$ und $\frac{d}{dt}\gamma(t)$ den Geschwindigkeitsvektor zum Zeitpunkt t.

Beispiele

Die folgenden Funktionen sind reguläre parametrisierte Kurven.

1) $\gamma : \mathbb{R} \to \mathbb{R}^n$, $t \mapsto a + vt$ mit $a \in \mathbb{R}^n$ und $v \in \mathbb{R}^n \setminus \{0\}$ (Gerade)

2) $\gamma : [0, 2\pi] \to \mathbb{R}^2$, $t \mapsto \begin{pmatrix} r\cos t \\ r\sin t \end{pmatrix}$ mit $r > 0$ (Kreis um 0 mit Radius r)

3) $\gamma : \mathbb{R} \to \mathbb{R}^3$, $t \mapsto \begin{pmatrix} r\cos t \\ r\sin t \\ t \end{pmatrix}$ mit $r > 0$ (Schraubenlinie)

4) $\gamma : [a, b] \to \mathbb{R}^2$, $t \mapsto \begin{pmatrix} t \\ f(t) \end{pmatrix}$ mit $f \in C^1([a, b])$ (Graph von f)

Definition *Sei $\gamma \in C^1([a,b],\mathbb{R}^n)$ eine reguläre parametrisierte Kurve. Dann heißt*

$$L(\gamma) := \int_a^b \left|\frac{d}{dt}\gamma(t)\right| dt$$

Länge der Kurve γ.

Bemerkung
Haben zwei verschiedene reguläre Parametrisierungen einer Kurve dieselbe Spur, dann ergibt sich aufgrund der Substitutionsregel unabhängig von der jeweiligen Parametrisierung dieselbe Kurvenlänge.

Beispiel
Sei $r > 0$,

$$\gamma_1 : [0, 2\pi] \to \mathbb{R}^2\,,\; t \mapsto \begin{pmatrix} r\cos t \\ r\sin t \end{pmatrix}$$

und

$$\gamma_2 : [0, \pi] \to \mathbb{R}^2\,,\; \tilde{t} \mapsto \begin{pmatrix} r\cos 2\tilde{t} \\ r\sin 2\tilde{t} \end{pmatrix}.$$

Dann haben γ_1 und γ_2 dieselbe Spur und es gilt

$$\frac{d}{dt}\gamma_1(t) = \begin{pmatrix} -r\sin t \\ r\cos t \end{pmatrix}\,, \quad \frac{d}{d\tilde{t}}\gamma_2(\tilde{t}) = \begin{pmatrix} -2r\sin 2\tilde{t} \\ 2r\cos 2\tilde{t} \end{pmatrix},$$

$$L(\gamma_1) = \int_0^{2\pi} \sqrt{r^2\sin^2 t + r^2\cos^2 t}\, dt = \int_0^{2\pi} r\, dt = 2\pi r\,,$$

$$L(\gamma_2) = \int_0^{\pi} \sqrt{4r^2\sin^2 2\tilde{t} + 4r^2\cos^2 2\tilde{t}}\, d\tilde{t} = \int_0^{2\pi} \sqrt{r^2\sin^2 t + r^2\cos^2 t}\, dt = L(\gamma_1)\,.$$

Definition

a) *Sei $a \in \mathbb{R}^n$ und $\varepsilon > 0$. Dann heißt*

$$B_\varepsilon(a) := \{x \in \mathbb{R}^n \,:\, |x - a| < \varepsilon\}$$

eine ε-Umgebung von a.

b) *Sei $a \in D \subseteq \mathbb{R}^n$. a heißt innerer Punkt von D, falls a eine ε-Umgebung besitzt, die in D enthalten ist.*

Definition *Sei $f : \mathbb{R}^n \supseteq D \to \mathbb{R}$, $(x_1, \ldots, x_n) \mapsto f(x_1, \ldots, x_n)$ eine Funktion.*

a) *Sei $x = (x_1, \ldots, x_n)$ ein innerer Punkt von D und $i \in \{1, \ldots, n\}$. Existiert der Grenzwert*

$$\partial_{x_i} f(x) := \lim_{h\to 0} \frac{f(x_1, \ldots, x_i + h, \ldots, x_n) - f(x_1, \ldots, x_i, \ldots, x_n)}{h}\,,$$

dann heißt f in x nach der i-ten Variable (bzw. nach der i-ten Koordinate) x_i partiell differenzierbar und $\partial_{x_i} f(x)$ die partielle Ableitung von f nach x_i in x. Statt $\partial_{x_i} f$ kann man auch die Symbole f_{x_i}, $\dfrac{\partial f}{\partial x_i}$, $\dfrac{\partial}{\partial x_i} f$ oder $\partial_i f$ verwenden.

b) *f heißt in $M \subseteq D$ nach x_i partiell differenzierbar, falls f in jedem $x \in M$ partiell nach x_i differenzierbar ist. Die Funktion $\partial_{x_i} f : M \to \mathbb{R}$, $x \mapsto \partial_{x_i} f(x)$ heißt dann partielle Ableitung von f nach x_i.*

Bemerkungen

1) $\partial_{x_i} f(x)$ beschreibt die Steigung der Tangente an den Graphen der Funktion $t \mapsto f(x + te_i)$ in $(0, f(x))$, wobei e_i der i-te Basisvektor der geordneten kanonischen Basis des $\mathbb{R}^n$ ist.

2) Die Rechenregeln für Ableitungen von Funktionen $f : \mathbb{R} \supseteq D \to \mathbb{R}$ übertragen sich sinngemäß auf partielle Ableitungen.

Beispiel
Sei $f : \mathbb{R}^2 \to \mathbb{R}$ mit $f(x_1, x_2) = x_1 x_2^2 + e^{2x_1} - \sin x_2$. Dann gilt:

$$\begin{aligned} \partial_{x_1} f(x_1, x_2) &= x_2^2 + 2e^{2x_1}, & \partial_{x_1} f(1,1) &= 1 + 2e^2, \\ \partial_{x_2} f(x_1, x_2) &= 2x_1 x_2 - \cos x_2, & \partial_{x_2} f(1,0) &= -1. \end{aligned}$$

Definition *Sei $f : \mathbb{R}^n \supseteq D \to \mathbb{R}$ eine Funktion, x ein innerer Punkt von D und $v \in \mathbb{R}^n$. Existiert der Grenzwert*

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

dann heißt er Richtungsableitung von f in x in Richtung v.

Beispiel
Sei $f(x_1, x_2) = x_1 x_2^2 + e^{2x_1} - \sin x_2$ und $v = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$. Dann gilt:

$$\begin{aligned} \partial_v f(x_1, x_2) &= 2x_2^2 - 2x_1 x_2 + 4e^{2x_1} + \cos x_2, \\ \partial_{e_i} f(x_1, x_2) &= \partial_{x_i} f(x_1, x_2). \end{aligned}$$

Definition *Sei $f : \mathbb{R}^n \supseteq D \to \mathbb{R}$ eine Funktion und x ein innerer Punkt von D. Dann heißen*

$$\partial_{x_i}^2 f(x) = f_{x_i x_i}(x) = \frac{\partial^2 f}{\partial x_i^2}(x) = \frac{\partial^2}{\partial x_i^2} f(x) = \partial_i^2 f(x) \quad := \quad \partial_{x_i}(\partial_{x_i} f)(x)$$

für $i \in \{1, \ldots, n\}$ und

$$\partial_{x_i} \partial_{x_j} f(x) = f_{x_i x_j}(x) = \frac{\partial^2 f}{\partial x_i \partial x_j}(x) = \frac{\partial^2}{\partial x_i \partial x_j} f(x) = \partial_i \partial_j f(x) \quad := \quad \partial_{x_i}(\partial_{x_j} f)(x)$$

für $i, j \in \{1, \ldots, n\}$ die zweiten partiellen Ableitungen bzw. die partiellen Ableitungen zweiter Ordnung von f in x, sofern sie existieren.
Induktiv definiert man die partiellen Ableitungen k-ter Ordnung durch

$$\frac{\partial^k}{\partial x_{i_k} \partial x_{i_{k-1}} \ldots \partial x_{i_1}} f := \frac{\partial}{\partial x_{i_k}} \left(\frac{\partial^{k-1} f}{\partial x_{i_{k-1}} \ldots \partial x_{i_1}} \right).$$

Beispiel
Sei $f(x_1, x_2) = x_1 x_2^2 + e^{2x_1} - \sin x_2$. Dann gilt:

$$\begin{aligned}
&\partial_{x_1}^2 f(x_1, x_2) = 4e^{2x_1},\\
&\partial_{x_1}\partial_{x_2} f(x_1, x_2) = 2x_2,\\
&\partial_{x_2}\partial_{x_1} f(x_1, x_2) = 2x_2,\\
&\partial_{x_2}^2 f(x_1, x_2) = 2x_1 + \sin x_2,\\
&\partial_{x_1}^3 f(x_1, x_2) = 8e^{2x_1},\\
&\partial_{x_1}\partial_{x_2}^2 f(x_1, x_2) = 2.
\end{aligned}$$

Definition *Sei $f : \mathbb{R}^n \supseteq D \to \mathbb{R}^m$, $(x_1, \ldots, x_n) \mapsto \begin{pmatrix} f_1(x_1, \ldots, x_n) \\ \vdots \\ f_m(x_1, \ldots, x_n) \end{pmatrix}$. Dann sind die partiellen Ableitungen von f definiert durch*

$$\partial_{x_i} f(x_1, \ldots, x_n) := \begin{pmatrix} \partial_{x_i} f_1(x_1, \ldots, x_n) \\ \vdots \\ \partial_{x_i} f_m(x_1, \ldots, x_n) \end{pmatrix}.$$

Analog definiert man die Richtungsableitung $\partial_v f$ sowie die partiellen Ableitungen k-ter Ordnung von f.

Definition *Sei $f : \mathbb{R}^n \supseteq D \to \mathbb{R}^m$, $(x_1, \ldots, x_n) \mapsto \begin{pmatrix} f_1(x_1, \ldots, x_n) \\ \vdots \\ f_m(x_1, \ldots, x_n) \end{pmatrix}$.*

a) Sei x_0 ein innerer Punkt von D. f heißt (total) differenzierbar in x_0, falls gilt: Es existiert eine Matrix $A \in \mathbb{R}^{m \times n}$, so dass

$$\lim_{x \to x_0} \Big(\frac{f(x) - f(x_0)}{|x - x_0|} - A \frac{x - x_0}{|x - x_0|} \Big) = 0$$

bzw.

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

ist. A heißt dann (totale) Ableitung von f in x_0. Schreibweise: $Df(x_0) = A$ oder $f'(x_0) = A$.

b) f heißt (total) differenzierbar in $M \subseteq D$, falls f in jedem $x \in M$ (total) differenzierbar ist. Die Funktion $f' : M \to \mathbb{R}^{m \times n}$, $x \mapsto f'(x)$ heißt dann (totale) Ableitung von f.

Satz 20.1 *Ist $f : \mathbb{R}^n \supseteq D \to \mathbb{R}^m$ (total) differenzierbar in x_0, dann ist f auch stetig in x_0.*

Beweis
analog zum eindimensionalen Fall. $\square$

Satz 20.2 *Sei $f : \mathbb{R}^n \supseteq D \to \mathbb{R}^m$ und $x_0 \in D$. Dann ist äquivalent:*

(i) f *ist (total) differenzierbar in x_0.*

(ii) $\exists A \in \mathbb{R}^{m\times n}\, \forall x \in D :\ f(x) = f(x_0) + A(x - x_0) + R(x, x_0)$

mit $\lim\limits_{x\to x_0} \dfrac{R(x, x_0)}{|x - x_0|} = 0.$

Sind (i) bzw. (ii) erfüllt, dann ist $A = f'(x_0)$.

Beweis
analog zum eindimensionalen Fall. $\square$

Bemerkung
Ist $f : \mathbb{R}^2 \supseteq D \to \mathbb{R}$ in x_0 (total) differenzierbar, dann beschreibt

$$T_{x_0}G_f := \left\{(x,\ f(x_0) + f'(x_0)(x - x_0))\ :\ x \in \mathbb{R}^2\right\}$$

die Tangentialebene an den Graphen von f im Punkt $(x_0, f(x_0))$.

Satz 20.3 *Ist $f : \mathbb{R}^n \supseteq D \to \mathbb{R}^m$ in x_0 (total) differenzierbar, dann gilt*

a) $f'(x_0) = \begin{pmatrix} \partial_{x_1} f_1(x_0) & \cdots & \partial_{x_n} f_1(x_0) \\ \vdots & & \vdots \\ \partial_{x_1} f_m(x_0) & \cdots & \partial_{x_n} f_m(x_0) \end{pmatrix}$

b) $\forall v \in \mathbb{R}^n :\ \partial_v f(x_0) = f'(x_0)v$

Beweis
durch Nachrechnen. $\square$

Definition *Sei $f : \mathbb{R}^n \supseteq D \to \mathbb{R}^m$ und $x_0 \in D$, dann heißt, sofern existent,*

$$Jf(x_0) := \begin{pmatrix} \partial_{x_1} f_1(x_0) & \cdots & \partial_{x_n} f_1(x_0) \\ \vdots & & \vdots \\ \partial_{x_1} f_m(x_0) & \cdots & \partial_{x_n} f_m(x_0) \end{pmatrix}$$

die Jacobi-Matrix von f in x_0.

Ist $m = 1$, dann heißt außerdem, sofern existent,

$$\nabla f(x_0) := (Jf(x_0))^\top = \begin{pmatrix} \partial_{x_1} f(x_0) \\ \vdots \\ \partial_{x_n} f(x_0) \end{pmatrix}$$

der Gradient von f in x_0. Alternative Schreibweise: $\operatorname{grad} f(x_0)$ *statt* $\nabla f(x_0)$.

Bemerkung
Ist f in x_0 (total) differenzierbar, dann ist

$$f'(x_0) = Jf(x_0) \quad \text{und} \quad \forall v \in \mathbb{R}^n : \partial_v f(x_0) = Jf(x_0)v$$

sowie im Falle $m = 1$ außerdem

$$f'(x_0) = (\nabla f(x_0))^\top \quad \text{und} \quad \forall v \in \mathbb{R}^n : \partial_v f(x_0) = (\nabla f(x_0))^\top v = \langle \nabla f(x_0), v \rangle_{\mathbb{R}^n} .$$

Bemerkung
Aus der Existenz aller partiellen Ableitungen 1. Ordnung und aller Richtungsableitungen einer Funktion f in einem Punkt x_0 folgt nicht notwendigerweise die totale Differenzierbarkeit oder die Stetigkeit von f in x_0.
Denn beispielsweise für

$$f : \mathbb{R}^2 \to \mathbb{R} \quad , \quad f(x,y) = \begin{cases} \dfrac{xy^2}{x^2 + y^4} , & (x,y) \neq (0,0), \\ 0 , & (x,y) = (0,0), \end{cases}$$

gilt

$$\forall v \in \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \in \mathbb{R}^2 : \partial_v f(0,0) = \begin{cases} \dfrac{v_2^2}{v_1} , & v_1 \neq 0 \\ 0 , & v_1 = 0 \end{cases}$$

aber f ist, wie wir in Kapitel 19 gesehen haben, nicht stetig in $(0,0)$.
Und für

$$g : \mathbb{R}^2 \to \mathbb{R} \quad , \quad g(x,y) = \begin{cases} \dfrac{y^3 - x^2 y}{x^2 + y^2} , & (x,y) \neq (0,0), \\ 0 , & (x,y) = (0,0), \end{cases}$$

gilt

$$\forall v \in \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \in \mathbb{R}^2 : \partial_v g(0,0) = \begin{cases} \dfrac{v_2^3 - v_1^2 v_2}{v_1^2 + v_2^2} , & v \neq 0, \\ 0 , & v = 0. \end{cases}$$

g ist wegen $|g(x,y)| \leq |y|$ für alle $(x,y) \in \mathbb{R}^2$ auch stetig in $(0,0)$, aber

$$\partial_v g(0,0) \;=\; (\nabla g(0,0))^\top v \;=\; (0 \quad 1)\, v \;=\; v_2$$

gilt nur, falls $v_1 = 0$ oder $v_2 = 0$ ist. Also kann nach Satz 20.3 *b)* g nicht total differenzierbar in $(0,0)$ sein.

Satz 20.4 *Sei $f : \mathbb{R}^n \supseteq D \to \mathbb{R}^m$ eine Funktion. Existieren alle partiellen Ableitungen erster Ordnung von f in D und sind diese stetig in x, dann ist f (total) differenzierbar in x.*

Beweis
Wir führen nur den Fall $n = 2$ und $m = 1$ explizit vor. Die anderen Fälle funktionieren mit analoger Argumentation.
Seien die Voraussetzungen des Satzes 20.4 erfüllt und sei $h \in \mathbb{R}^2 \setminus \{0\}$ mit $x + h = (x_1 + h_1, x_2 + h_2) \in D$. Dann existieren aufgrund des Mittelwertsatzes der Differentialrechnung $\xi_i = \xi_i(h_i) \in]0, h_i[$, $i \in \{1,2\}$, so dass gilt:

$$\begin{aligned}
f(x_1+h_1, x_2+h_2) - f(x_1,x_2) &= f(x_1+h_1, x_2+h_2) - f(x_1, x_2+h_2) \\
&\quad + f(x_1, x_2+h_2) - f(x_1,x_2) \\
&= \partial_{x_1} f(x_1+\xi_1, x_2+h_2)h_1 + \partial_{x_2} f(x_1, x_2+\xi_2)h_2 \\
&= \partial_{x_1} f(x_1,x_2)h_1 + \partial_{x_2} f(x_1,x_2)h_2 \\
&\quad + \left(\partial_{x_1} f(x_1+\xi_1, x_2+h_2) - \partial_{x_1} f(x_1,x_2)\right) h_1 \\
&\quad + \left(\partial_{x_2} f(x_1, x_2+\xi_2) - \partial_{x_2} f(x_1,x_2)\right) h_2 \\
&= Jf(x_1,x_2)h + R\left((x_1+h_1, x_2+h_2), (x_1,x_2)\right)
\end{aligned}$$

mit

$$\begin{aligned}
R\left((x_1+h_1, x_2+h_2), (x_1,x_2)\right) &= \left(\partial_{x_1} f(x_1+\xi_1, x_2+h_2) - \partial_{x_1} f(x_1,x_2)\right) h_1 \\
&\quad + \left(\partial_{x_2} f(x_1, x_2+\xi_2) - \partial_{x_2} f(x_1,x_2)\right) h_2.
\end{aligned}$$

Da $\partial_{x_1} f$ und $\partial_{x_2} f$ stetig in x sind, gilt

$$\lim_{h\to 0} \frac{R\left((x_1+h_1, x_2+h_2), (x_1,x_2)\right)}{|h|} = 0.$$

Aufgrund von Satz 20.2 folgt daher die totale Differenzierbarkeit von f in x. □

Beispiel
$f : \mathbb{R}^2 \to \mathbb{R},\ f(x_1,x_2) = x_1 x_2^2 + e^{2x_1} - \sin x_2$
$\partial_{x_1} f$ mit $\partial_{x_1} f(x_1,x_2) = x_2^2 + 2e^{2x_1}$ und $\partial_{x_2} f$ mit $\partial_{x_2} f(x_1,x_2) = 2x_1x_2 - \cos x_2$ sind stetig in $\mathbb{R}^2$.
Daher ist f total differenzierbar in $\mathbb{R}^2$ mit

$$f'(x_1,x_2) = Jf(x_1,x_2) = (\nabla f(x_1,x_2))^\top = \begin{pmatrix} x_2^2 + 2e^{2x_1} & 2x_1x_2 - \cos x_2 \end{pmatrix}$$

Infolgedessen gilt

$$\begin{aligned}
\partial_{\binom{2}{-1}} f(x_1,x_2) &= \begin{pmatrix} x_2^2 + 2e^{2x_1} & 2x_1x_2 - \cos x_2 \end{pmatrix} \begin{pmatrix} 2 \\ -1 \end{pmatrix} \\
&= 2x_2^2 + 4e^{2x_1} - 2x_1x_2 + \cos x_2
\end{aligned}$$

$g : \mathbb{R}^2 \to \mathbb{R}^2,\ g(x_1, x_2) = \begin{pmatrix} x_1 x_2^2 \\ e^{2x_1} - \sin x_2 \end{pmatrix}$

$\partial_{x_1} g$ mit $\partial_{x_1} g(x_1, x_2) = \begin{pmatrix} x_2^2 \\ 2e^{2x_1} \end{pmatrix}$ und $\partial_{x_2} g$ mit $\partial_{x_2} g(x_1, x_2) = \begin{pmatrix} 2x_1 x_2 \\ -\cos x_2 \end{pmatrix}$ sind stetig in $\mathbb{R}^2$.

Daher ist g total differenzierbar in $\mathbb{R}^2$ mit

$$g'(x_1, x_2) = Jg(x_1, x_2) = \begin{pmatrix} x_2^2 & 2x_1 x_2 \\ 2e^{2x_1} & -\cos x_2 \end{pmatrix}$$

Infolgedessen gilt

$$\partial_{\binom{2}{-1}} g(x_1, x_2) = \begin{pmatrix} x_2^2 & 2x_1 x_2 \\ 2e^{2x_1} & -\cos x_2 \end{pmatrix} \begin{pmatrix} 2 \\ -1 \end{pmatrix} = \begin{pmatrix} 2x_2^2 - 2x_1 x_2 \\ 4e^{2x_1} + \cos x_2 \end{pmatrix}$$

Satz 20.5 *Sei $f : \mathbb{R}^n \supseteq D \to \mathbb{R}$ (total) differenzierbar in D. Dann gilt*

a) Existiert eine reguläre parametrisierte Kurve $\gamma \in C^1\left(]-\varepsilon, \varepsilon[\,, D\right)$ und ein $c \in \mathbb{R}$ mit $f(\gamma(t)) = c$ für alle $t \in\ ]-\varepsilon, \varepsilon[$, dann gilt

$$\forall t \in\ ]-\varepsilon, \varepsilon[\ :\ \left\langle \nabla f(\gamma(t)), \frac{d}{dt}\gamma(t) \right\rangle_{\mathbb{R}^n} = 0.$$

Der Gradient steht also senkrecht auf jeder Niveaulinie von f.

b) Sei $x_0 \in D$ und $\nabla f(x_0) \neq 0$. Dann gilt

$$\max_{|v|=1} \partial_v f(x_0) = \partial_{\frac{\nabla f(x_0)}{|\nabla f(x_0)|}} f(x_0) = |\nabla f(x_0)|$$

und

$$\forall v \in \mathbb{R}^n \setminus \{0\} : \frac{v}{|v|} \neq \frac{\nabla f(x_0)}{|\nabla f(x_0)|} \implies \nabla_{\frac{v}{|v|}} f(x_0) < \partial_{\frac{\nabla f(x_0)}{|\nabla f(x_0)|}} f(x_0)$$

Der Gradient zeigt also in Richtung des steilsten Anstiegs von f.

Beweis
Da f in D total differenzierbar ist, gilt

$$\forall x \in D\, \forall v \in \mathbb{R}^n : \partial_v f(x) = \langle \nabla f(x), v \rangle_{\mathbb{R}^n}$$

woraus a) folgt und aufgrund der Cauchy-Schwarzschen Ungleichung auch b). □

Beispiel
$f : \mathbb{R}^2 \to \mathbb{R},\ f(x, y) = x^2 + y^2$

$\nabla f(x, y) = \begin{pmatrix} 2x \\ 2y \end{pmatrix}$

Satz 20.6 *a) Seien* $f = \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix}$, $g = \begin{pmatrix} g_1 \\ \vdots \\ g_m \end{pmatrix} : \mathbb{R}^n \supseteq D \to \mathbb{R}^m$ *Funktionen, die in* x_0 *(total) differenzierbar sind. Dann gilt*

(i) Für alle $\alpha, \beta \in \mathbb{R}$ *ist die Funktion* $\alpha f + \beta g$ *in* x_0 *(total) differenzierbar mit*

$$(\alpha f + \beta g)'(x_0) = \alpha f'(x_0) + \beta g'(x_0)$$

(ii) (Produktregel)
Für alle $i \in \{1, \ldots, m\}$ *sind die Funktionen* $f_i g_i$ *in* x_0 *(total) differenzierbar mit*

$$(f_i g_i)'(x_0) = f_i(x_0) g_i'(x_0) + g_i(x_0) f_i'(x_0)\,.$$

(iii) (Quotientenregel)
Ist $g_i(x_0) \neq 0$ *für ein* $i \in \{1, \ldots, m\}$*, dann ist die Funktion* $\dfrac{f_i}{g_i}$ *in* x_0 *(total) differenzierbar mit*

$$\left(\frac{f_i}{g_i}\right)'(x_0) = \frac{1}{\left(g_i(x_0)\right)^2}\left(g_i(x_0) f_i'(x_0) - f_i(x_0) g_i'(x_0)\right)$$

b) (Kettenregel)
Sei $g = \begin{pmatrix} g_1 \\ \vdots \\ g_m \end{pmatrix} : \mathbb{R}^n \supseteq D \to \mathbb{R}^m$ *(total) differenzierbar in* x_0 *und* $h = \begin{pmatrix} h_1 \\ \vdots \\ h_l \end{pmatrix} :$ $\mathbb{R}^m \supseteq g(D) \to \mathbb{R}^l$ *(total) differenzierbar in* $g(x_0)$*, dann ist* $h \circ g$ *(total) differenzierbar in* x_0 *mit*

$$(h \circ g)'(x_0) = h'(g(x_0)) g'(x_0) \quad \textit{(im Sinne der Matrizenmultiplikation)}$$

Beweisstrategie
Durch Nachrechnen.

Beispiel
$g : \mathbb{R}^2 \to \mathbb{R}^2$, $g(x, y) = \begin{pmatrix} x \\ y^2 \end{pmatrix}$

$h : \mathbb{R}^2 \to \mathbb{R}$, $h(x, y) = xy$

$\partial_x g$ mit $\partial_x g(x, y) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ und $\partial_y g$ mit $\partial_y g(x, y) = \begin{pmatrix} 0 \\ 2y \end{pmatrix}$ sind stetig in $\mathbb{R}^2$

$\Longrightarrow$ g ist in $\mathbb{R}^2$ total differenzierbar mit $g'(x, y) = \begin{pmatrix} 1 & 0 \\ 0 & 2y \end{pmatrix}$.

$\partial_x h$ mit $\partial_x h(x,y) = y$ und $\partial_y h$ mit $\partial_y h(x,y) = x$ sind stetig in $\mathbb{R}^2$

$\implies$ h ist in $\mathbb{R}^2$ total differenzierbar mit $h'(x,y) = \begin{pmatrix} y & x \end{pmatrix}$.

Also ist $h \circ g$ total differenzierbar in $\mathbb{R}^2$ mit

$$\begin{aligned}(h \circ g)'(x,y) &= h'(g(x,y))g'(x,y) \\ &= h'(x,y^2)g'(x,y) \\ &= \begin{pmatrix} y^2 & x \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 2y \end{pmatrix} \\ &= \begin{pmatrix} y^2 & 2xy \end{pmatrix}\end{aligned}$$

Satz 20.7 (Mittelwertsatz der mehrdimensionalen Differentialrechnung)
Sei $f : \mathbb{R}^n \supseteq D \to \mathbb{R}$ (total) differenzierbar in D und seien $a, b \in D$ zwei Punkte, deren Verbindungsstrecke $\{a + t(b-a) : t \in [0,1]\}$ in D enthalten ist. Dann gilt

$$\exists \vartheta \in \,]0,1[\, : \; f(b) - f(a) = f'\left(a + \vartheta(b-a)\right)(b-a)$$

Beweis
$h : [0,1] \to \mathbb{R}$, $h(t) = f\left(a + t(b-a)\right)$ erfüllt die Voraussetzungen des Mittelwertsatzes der eindimensionalen Differentialrechnung.
$\implies \exists \vartheta \in \,]0,1[\, : \; h(1) - h(0) = h'(\vartheta)$
$\overset{20.6\,b)}{\implies} f(b) - f(a) = f'\left(a + \vartheta(b-a)\right)(b-a)$ □

Bemerkung
Sei $f = \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix} : \mathbb{R}^n \supseteq D \to \mathbb{R}^m$ (total) differenzierbar in D und
$\{a + t(b-a) : t \in [0,1]\} \subseteq D$. Dann gilt nach Satz 20.7:

$$\forall i \in \{1, \ldots, m\} \; \exists \vartheta_i \in \,]0,1[\, : \; f_i(b) - f_i(a) = f_i'\left(a + \vartheta_i(b-a)\right)(b-a)$$

Im Allgemeinen gilt aber $\vartheta_i \neq \vartheta_j$ für $i \neq j$ und dann existiert kein $\vartheta \in \,]0,1[$ mit $f(b) - f(a) = f'\left(a + \vartheta(b-a)\right)(b-a)$.

Korollar 20.8 *Sei $f : \mathbb{R}^n \supseteq D \to \mathbb{R}^m$ (total) differenzierbar in D. Ist $f'(x) = 0$ für alle $x \in D$, dann ist f längs jeder in D liegenden Strecke konstant.*

Satz 20.9 (Satz von Schwarz) *Sei $f : \mathbb{R}^n \supseteq D \to \mathbb{R}^m$ eine Funktion. Existieren alle partiellen Ableitungen erster und zweiter Ordnung von f in D und sind diese stetig in x, dann gilt*

$$\forall i, j \in \{1, \ldots, n\} : \; \partial_{x_i}\partial_{x_j} f(x) = \partial_{x_j}\partial_{x_i} f(x)\,,$$

d.h., die Reihenfolge der partiellen Differentiation bis zur zweiten Ordnung spielt keine Rolle in x.

Beweisstrategie
Durch Approximation der ersten und zweiten partiellen Ableitungen mit Hilfe von Differenzenquotienten und Ausnutzen des eindimensionalen Mittelwertsatzes und der Stetigkeit der ersten und zweiten partiellen Ableitungen auf ähnliches Weise wie im Beweis von Satz 20.4.

Beispiel
Wie wir bereits explizit berechnet haben, erfüllt die Funktion $f(x_1, x_2) = x_1 x_2^2 + e^{2x_1} - \sin x_2$ auf ganz $\mathbb{R}^2$ die Voraussetzungen und die Behauptung des Satzes von Schwarz.

Bemerkung
Sind nicht alle partiellen Ableitungen erster und zweiter Ordnung einer Funktion stetig, dann ist die Reihenfolge der partiellen Differentiation im Allgemeinen nicht egal. Wie man direkt nachrechnen kann, trifft dies z.B. auf die Funktion f mit

$$f(x,y) = \begin{cases} xy\,\dfrac{x^2-y^2}{x^2+y^2}\,, & (x,y) \neq (0,0), \\ 0\,, & (x,y) = (0,0), \end{cases}$$

in $(0,0)$ zu und man erhält $\partial_x \partial_y f(0,0) = 1 \neq -1 = \partial_y \partial_x f(0,0)$.

Definition *Sei $f : \mathbb{R}^n \supseteq D \to \mathbb{R}$ und $x_0 \in D$. Dann heißt, sofern existent,*

$$Hf(x_0) = \begin{pmatrix} \partial_{x_1}^2 f(x_0) & \cdots & \partial_{x_1}\partial_{x_n} f(x_0) \\ \vdots & & \vdots \\ \partial_{x_n}\partial_{x_1} f(x_0) & \cdots & \partial_{x_n}^2 f(x_0) \end{pmatrix}$$

die Hesse-Matrix von f in x_0.

Bemerkung
Ist $f : \mathbb{R}^n \supseteq D \to \mathbb{R}$ eine Funktion mit der Eigenschaft, dass alle partiellen Ableitungen erster und zweiter Ordnung von f in D existieren und stetig in x sind, dann ist nach Satz 20.4 f (total) differenzierbar in x und ebenso $f' : \mathbb{R}^n \supseteq D \to \mathbb{R}^n$ (total) differenzierbar in x. Außerdem gilt $(Hf)^\top(x) = J(Jf)^\top(x)$ und nach Satz 20.9 ist $Hf(x)$ symmetrisch, also $Hf(x) = (Hf)^\top(x)$. Daher schreibt man in diesem Fall auch $f''(x)$ statt $Hf(x)$.

Definition *$f : \mathbb{R}^n \supseteq D \to \mathbb{R}^m$ heißt k-mal stetig differenzierbar in $M \subseteq D$, falls alle partiellen Ableitungen von f bis zur k-ten Ordnung in M existieren und dort stetig sind.*
Bezeichnung:

$$C^k(M, \mathbb{R}^m) := \{f : M \to \mathbb{R}^m : \ f \textit{ ist } k\textit{-mal stetig differenzierbar in } M\}\,.$$

$f : \mathbb{R}^n \supseteq D \to \mathbb{R}^m$ heißt unendlich oft differenzierbar in $M \subseteq D$, falls für alle $k \in \mathbb{N}$ alle partiellen Ableitungen k-ter Ordnung von f in M existieren.
Bezeichnung:

$$C^\infty(M, \mathbb{R}^m) := \{f : M \to \mathbb{R}^m : \ f \textit{ ist unendlich oft differenzierbar in } M\}.$$

Bemerkung
Aufgrund der Sätze 20.4 und 20.6 sind $C^k(M, \mathbb{R}^m)$ und $C^\infty(M, \mathbb{R}^m)$ $\mathbb{R}$-Vektorräume und $f \mapsto f'$ ist eine lineare Abbildung von $C^k(M, \mathbb{R}^m)$ nach $C^{k-1}(M, \mathbb{R}^m)$ für $k \geq 1$.

Definition *Ein n-Tupel $\alpha = (\alpha_1, \ldots, \alpha_n)$ mit $\alpha_i \in \mathbb{N}_0$ für $i = 1, \ldots, n$ heißt Multiindex. Für Multiindizes $\alpha, \beta \in \mathbb{N}_0^n$ sei*

$$\begin{aligned}
|\alpha| &:= \sum_{i=1}^{n} \alpha_i \,, \\
\alpha! &:= \prod_{i=1}^{n} \alpha_i! \,, \\
\beta = \alpha &:\Leftrightarrow \forall i \in \{1, \ldots, n\} : \ \beta_i = \alpha_i \,, \\
\beta \leq \alpha &:\Leftrightarrow \forall i \in \{1, \ldots, n\} : \ \beta_i \leq \alpha_i \,, \\
\alpha + \beta &:= (\alpha_1 + \beta_1, \ldots, \alpha_n + \beta_n) \,, \\
\alpha - \beta &:= (\alpha_1 - \beta_1, \ldots, \alpha_n - \beta_n) \,, \qquad \textit{falls } \beta \leq \alpha \,, \\
\binom{\alpha}{\beta} &:= \frac{\alpha!}{\beta!(\alpha - \beta)!} = \prod_{i=1}^{n} \binom{\alpha_i}{\beta_i} \,, \qquad \textit{falls } \beta \leq \alpha \,, \\
\forall x \in \mathbb{R}^n : \ x^\alpha &:= \prod_{i=1}^{n} x_i^{\alpha_i} \,, \\
\forall D \subseteq \mathbb{R}^n \, \forall f \in C^{|\alpha|}(D, \mathbb{R}^m) : \ \partial^\alpha f &:= \partial_{x_1}^{\alpha_1} \partial_{x_2}^{\alpha_2} \ldots \partial_{x_n}^{\alpha_n} f \,,
\end{aligned}$$

wobei $\partial_{x_i}^0 g := g$ ist.

Beispiel
$f \in C^3(\mathbb{R}^3, \mathbb{R}^m) \implies \partial^{(1,0,2)} f = \partial_{x_1} \partial_{x_3}^2 f$

Satz 20.10 *a) (Binomischer Satz)*

$$\forall x, y \in \mathbb{R}^n \, \forall \alpha \in \mathbb{N}_0^n : \ (x + y)^\alpha = \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} x^\beta \, y^{\alpha - \beta}$$

b) (Leibniz-Regel)
Sei $D \subseteq \mathbb{R}^n$, $k \in \mathbb{N}$ und seien $f, g \in C^k(D, \mathbb{R})$. Dann ist $fg \in C^k(D, \mathbb{R})$ und für alle $\alpha \in \mathbb{N}_0^n$ mit $|\alpha| \leq k$ gilt

$$\partial^\alpha(fg) = \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} \partial^\beta f \, \partial^{\alpha-\beta} g$$

Satz 20.11 (Satz von Taylor) *Sei $D \subseteq \mathbb{R}^n$, $k \in \mathbb{N}_0$, $f \in C^{k+1}(D, \mathbb{R})$ und seien $x_0, x \in D$ zwei Punkte, deren Verbindungsstrecke $\{x_0 + t(x - x_0) : t \in [0, 1]\}$ in D enthalten ist. Dann gilt*

$$\exists \vartheta \in]0, 1[: \; f(x) = \sum_{|\alpha| \leq k} \frac{\partial^\alpha f(x_0)}{\alpha!}(x - x_0)^\alpha + \sum_{|\alpha| = k+1} \frac{\partial^\alpha f(x_0 + \vartheta(x - x_0))}{\alpha!}(x - x_0)^\alpha.$$

Beweisskizze
Sei $h : [0, 1] \to \mathbb{R}, \; h(t) = f(x_0 + t(x - x_0))$.
Zeige durch Induktion unter Verwendung der Kettenregel 20.6 *b*) und der Formeln der Kombinatorik aus Abschnitt 4.1, dass $h \in C^{k+1}([0, 1])$ ist mit

$$h^{(\tilde{k})}(t) = \sum_{|\alpha| = \tilde{k}} \frac{\tilde{k}!}{\alpha!} \, \partial^\alpha f(x_0 + t(x - x_0)) \, (x - x_0)^\alpha$$

für $\tilde{k} \in \{0, 1, \ldots, k + 1\}$. Daher kann der eindimensionale Satz von Taylor auf h angewandt werden, was auch die Behauptung von Satz 20.11 liefert.

Korollar 20.12 *Sei $D \subseteq \mathbb{R}^n$, $k \in \mathbb{N}_0$, $f \in C^{k+1}(D, \mathbb{R})$ und $x_0 \in D$. Dann gilt*

$$f(x) = \sum_{|\alpha| \leq k} \frac{\partial^\alpha f(x_0)}{\alpha!}(x - x_0)^\alpha + O(|x - x_0|^{k+1}) = \sum_{|\alpha| \leq k} \frac{\partial^\alpha f(x_0)}{\alpha!}(x - x_0)^\alpha + o(|x - x_0|^k)$$

für $x \to x_0$.

Bemerkungen

1) Für $k = 0$ liefert der Satz von Taylor wieder den Mittelwertsatz, für $k = 1$ ergibt sich

   $$f(x) = f(x_0) + \langle \nabla f(x_0), x - x_0 \rangle_{\mathbb{R}^n} + \tfrac{1}{2} \langle x - x_0, Hf(x_0 + \vartheta(x - x_0))(x - x_0) \rangle_{\mathbb{R}^n}$$

   mit $\vartheta \in]0, 1[$ und für $k = 2$ erhalten wir

   $$f(x) = f(x_0) + \langle \nabla f(x_0), x - x_0 \rangle_{\mathbb{R}^n} + \tfrac{1}{2} \langle x - x_0, Hf(x_0)(x - x_0) \rangle_{\mathbb{R}^n} + o(|x - x_0|^2)$$

   für $x \to x_0$.

2) Sei $D \subseteq \mathbb{R}^n$, $k \in \mathbb{N}_0$, $f = \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix} \in C^{k+1}(D, \mathbb{R}^m)$ und $\{x_0+t(x-x_0) : t \in [0,1]\} \subseteq D$. Dann gilt die Aussage des Satzes von Taylor für jede Komponente f_i, aber man erhält im Allgemeinen für jedes f_i ein anderes ϑ.

Beispiel
Sei $f(x,y) = \mathrm{e}^{x+y}$. Dann ist $f \in C^\infty(\mathbb{R}^2, \mathbb{R})$ mit $\partial^\alpha f = f$ für jeden Multiindex α. Daher gilt

$\nabla f(x,y) = \begin{pmatrix} \mathrm{e}^{x+y} \\ \mathrm{e}^{x+y} \end{pmatrix}, \quad \nabla f(0,0) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$

$Hf(x,y) = \begin{pmatrix} \mathrm{e}^{x+y} & \mathrm{e}^{x+y} \\ \mathrm{e}^{x+y} & \mathrm{e}^{x+y} \end{pmatrix}, \quad Hf(0,0) = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$

und somit liefert der Satz von Taylor

$$\begin{aligned} f(x,y) &= 1 + \left\langle \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix} \right\rangle_{\mathbb{R}^2} + \frac{1}{2} \left\langle \begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \right\rangle_{\mathbb{R}^2} + o(x^2+y^2) \\ &= 1 + x + y + \frac{1}{2}(x^2+2xy+y^2) + o(x^2+y^2)\,. \end{aligned}$$

Definition *Eine Funktion $f : \mathbb{R}^n \supseteq D \to \mathbb{R}$ hat in $x_0 \in D$ ein lokales Maximum, falls gilt:*

$$\exists \varepsilon > 0 \forall x \in D : |x - x_0| < \varepsilon \Rightarrow f(x) \leq f(x_0).$$

f hat in x_0 ein lokales Minimum, falls gilt:

$$\exists \varepsilon > 0 \forall x \in D : |x - x_0| < \varepsilon \Rightarrow f(x) \geq f(x_0).$$

Ein lokales Extremum (Maximum oder Minimum) heißt strikt, falls $f(x) = f(x_0)$ nur für $x = x_0$ gilt.
Gilt $\forall x \in D : f(x) \leq f(x_0)$ bzw. $\forall x \in D : f(x) \geq f(x_0)$, dann hat f in x_0 ein globales (absolutes) Maximum bzw. Minimum in D.

Satz 20.13 (Notwendige Bedingung für lokale Extrema) *Sei $f : \mathbb{R}^n \supseteq D \to \mathbb{R}$ eine Funktion, $x_0 \in D$ und f (total) differenzierbar in x_0. Dann gilt:*

$$f \text{ hat in } x_0 \text{ ein lokales Extremum} \;\Rightarrow\; f'(x_0) = 0$$

(also $(\partial_{x_1} f(x_0), \ldots, \partial_{x_n} f(x_0)) = (0, \ldots, 0)$).

Beweis
f habe in x_0 ein lokales Extremum.
Sei $h_i(t) := f(x_0 + te_i)$ für $i = 1, \ldots, n$.

$\Rightarrow \exists \varepsilon > 0 \forall i \in \{1, \ldots, n\} :$ h_i ist in $]-\varepsilon, \varepsilon[$ differenzierbar und hat in 0 ein lokales Extremum

$\overset{16.9}{\Rightarrow} \forall i \in \{1, \ldots, n\} : h_i'(0) = 0$

$\Rightarrow f'(x_0) = 0$ □

Definition *Ist $f : \mathbb{R}^n \supseteq D \to \mathbb{R}$ differenzierbar in x_0 und $f'(x_0) = 0$, dann heißt x_0 stationärer oder kritischer Punkt von f.*

Satz 20.14 (Hinreichende Bedingungen für lokale Extrema) *Sei $D \subseteq \mathbb{R}^n$, $f \in C^2(D, \mathbb{R})$ und $x_0 \in D$. Dann gilt:*

a) $f'(x_0) = 0 \ \wedge\ f''(x_0) > 0$ (d.h., $f''(x_0) = Hf(x_0)$ ist positiv definit)
$\Rightarrow f$ hat in x_0 ein striktes lokales Minimum.

b) $f'(x_0) = 0 \ \wedge\ f''(x_0) < 0$ (d.h., $f''(x_0) = Hf(x_0)$ ist negativ definit)
$\Rightarrow f$ hat in x_0 ein striktes lokales Maximum.

c) $f'(x_0) = 0 \ \wedge\ f''(x_0)$ ist indefinit
$\Rightarrow f$ hat in x_0 einen Sattelpunkt (d.h., einen stationären Punkt, in dem f kein lokales Extremum hat).

Bemerkung
Sei $D \subseteq \mathbb{R}^n$, $f \in C^2(D, \mathbb{R})$ und $x_0 \in D$.
Ist $f'(x_0) = 0$ und $f''(x_0) \geq 0$ (d.h., $f''(x_0) = Hf(x_0)$ ist positiv semidefinit), dann kann daraus nicht geschlossen werden, ob f in x_0 ein lokales oder striktes lokales Minimum hat, wie die Beispiele $f_1(x, y) = x^2 + y^4$ (f_1 hat in $x_0 = (0, 0)$ ein striktes lokales Minimum), $f_2(x, y) = x^2$ (f_2 hat in $x_0 = (0, 0)$ ein lokales Minimum, welches nicht strikt ist) und $f_3(x, y) = x^2 + y^3$ (f_3 hat in $x_0 = (0, 0)$ einen Sattelpunkt) zeigen.
Ist $f'(x_0) = 0$ und gilt nicht $f''(x_0) \geq 0$, dann hat f in x_0 kein lokales Minimum, weil dann entweder die Situation von 20.14 *b*) oder die Situation von 20.14 *c*) vorliegt. Entsprechend ist $f'(x_0) = 0 \wedge f''(x_0) \leq 0$ (d.h., $f''(x_0) = Hf(x_0)$ ist negativ semidefinit) keine hinreichende, aber eine notwendige Bedingung für ein lokales oder striktes lokales Maximum von f in x_0.

Beweis von Satz 20.14:

a) Sei $f \in C^2(D, \mathbb{R})$, $f'(x_0) = 0$ und $f''(x_0) > 0$.
Da $f''(x_0)$ symmetrisch ist, existiert nach 12.15-12.18 eine Orthonormalbasis von

$\mathbb{R}^n$ aus Eigenvektoren von $f''(x_0)$, wobei alle Eigenwerte von $f''(x_0)$ positiv sind. Daher folgt:

$$\forall \xi \in \mathbb{R}^n : \langle \xi, f''(x_0)\xi\rangle_{\mathbb{R}^n} \geq \lambda_{\min}|\xi|^2,$$

wobei $\lambda_{\min} > 0$ der kleinste Eigenwert von $f''(x_0)$ ist. Nach Korollar 20.12 existiert somit ein $\delta > 0$, so dass

$$\begin{aligned} f(x) &= f(x_0) + \frac{1}{2}\langle x - x_0, f''(x_0)(x-x_0)\rangle_{\mathbb{R}^n} + o(|x-x_0|^2) \\ &\geq f(x_0) + \frac{1}{4}\lambda_{\min}|x-x_0|^2 \\ &> f(x_0) \end{aligned}$$

für $0 < |x - x_0| < \delta$ gilt. Folglich hat f in x_0 ein striktes lokales Minimum.

b) folgt durch Anwendung von *a)* auf $-f$.

c) Sei $f \in C^2(D, \mathbb{R})$, $f'(x_0) = 0$ und $f''(x_0)$ indefinit.
$\Rightarrow \exists \xi_1 . \xi_2 \in \mathbb{R}^n \backslash \{0\} \, \exists \alpha_1, \alpha_2 \in \mathbb{R}^+ : \langle \xi_1, f''(x_0)\xi_1\rangle_{\mathbb{R}^n} = -\alpha_1 \wedge \langle \xi_2, f''(x_0)\xi_2\rangle_{\mathbb{R}^n} = \alpha_2$
$\Rightarrow \exists \delta > 0 \, \forall \, 0 < |t| < \delta : \, f(x_0 + t\xi_1) \leq f(x_0) - \frac{\alpha_1}{4}t^2 < f(x_0) \wedge f(x_0 + t\xi_2) \geq f(x_0)$
$+ \frac{\alpha_2}{4}t^2 > f(x_0)$
Also hat f in x_0 einen Sattelpunkt. $\square$

Beispiele

1) $f : \mathbb{R}^2 \to \mathbb{R}, \; f(x,y) = x^2 + y^2$. Dann gilt
$f'(x,y) = 0 \;\Leftrightarrow\; \begin{pmatrix} 2x & 2y \end{pmatrix} = \begin{pmatrix} 0 & 0 \end{pmatrix} \;\Leftrightarrow\; (x,y) = (0,0)$
$f''(x,y) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \;\Rightarrow\; f''(0,0) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} > 0$
Also hat f in $(0,0)$ ein striktes lokales Minimum.

2) $g : \mathbb{R}^2 \to \mathbb{R}, \; g(x,y) = -x^2 - y^2$. Dann gilt
$g'(x,y) = 0 \;\Leftrightarrow\; \begin{pmatrix} -2x & -2y \end{pmatrix} = \begin{pmatrix} 0 & 0 \end{pmatrix} \;\Leftrightarrow\; (x,y) = (0,0)$
$g''(x,y) = \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix} \;\Rightarrow\; g''(0,0) = \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix} < 0$
Also hat g in $(0,0)$ ein striktes lokales Maximum.

3) $h : \mathbb{R}^2 \to \mathbb{R}, \; h(x,y) = x^2 - y^2$ Dann gilt
$h'(x,y) = 0 \;\Leftrightarrow\; \begin{pmatrix} 2x & -2y \end{pmatrix} = \begin{pmatrix} 0 & 0 \end{pmatrix} \;\Leftrightarrow\; (x,y) = (0,0)$
$h''(x,y) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix} \;\Rightarrow\; h''(0,0) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$ ist indefinit.
Also hat h in $(0,0)$ einen Sattelpunkt.

4) $q : \mathbb{R}^2 \to \mathbb{R},\ q(x,y) = x^3 + xy^2 - 3x$. Dann gilt

$$q'(x,y) = 0 \ \Leftrightarrow\ \begin{pmatrix} 3x^2 + y^2 - 3 & 2xy \end{pmatrix} = \begin{pmatrix} 0 & 0 \end{pmatrix} \ \Leftrightarrow\ \begin{cases} 3x^2 + y^2 - 3 = 0 & (I) \\ 2xy = 0 & (II) \end{cases}$$

$(II) \ \Leftrightarrow\ x = 0 \ \vee\ y = 0$

1. Fall: $y = 0$

$\overset{(I)}{\Rightarrow}\ 3x^2 - 3 = 0 \ \Leftrightarrow\ x = \pm 1$

2. Fall: $x = 0$

$\overset{(I)}{\Rightarrow}\ y^2 - 3 = 0 \ \Leftrightarrow\ y = \pm\sqrt{3}$

Also sind $(1,0),\ (-1,0),\ (0,\sqrt{3}),\ (0,-\sqrt{3})$ die stationären Punkte von q.

$$q''(x,y) = \begin{pmatrix} 6x & 2y \\ 2y & 2x \end{pmatrix}$$

$$\Rightarrow\ q''(1,0) = \begin{pmatrix} 6 & 0 \\ 0 & 2 \end{pmatrix} > 0\,,$$

$$q''(-1,0) = \begin{pmatrix} -6 & 0 \\ 0 & -2 \end{pmatrix} < 0\,,$$

$$q''(0,\pm\sqrt{3}) = \begin{pmatrix} 0 & \pm 2\sqrt{3} \\ \pm 2\sqrt{3} & 0 \end{pmatrix} \text{ ist indefinit,}$$

denn weil $\det(q''(0,\pm\sqrt{3})) = -12 < 0$ und die Determinante gleich dem Produkt der Eigenwerte ist, muss $q''(0,\pm\sqrt{3})$ einen positiven und einen negativen Eigenwert haben.

Also hat q in $(1,0)$ ein striktes lokales Minimum, in $(-1,0)$ ein striktes lokales Maximum und in $(0,\sqrt{3})$ sowie in $(0,-\sqrt{3})$ Sattelpunkte.

21 Mehrdimensionale Integralrechnung

Definition *Jede nichtleere Menge B der Form*

$$\begin{aligned} B \;=\; \{(x_1,\, \dots\, , x_n) \in \mathbb{R}^n :\; & a_1 \le x_1 \le b_1,\; a_2(x_1) \le x_2 \le b_2(x_1), \\ & a_3(x_1,x_2) \le x_3 \le b_3(x_1,x_2), \dots , a_n(x_1,\dots,x_{n-1}) \le x_n \le b_n(x_1,\dots,x_{n-1})\} \end{aligned}$$

mit $a_1, b_1 \in \mathbb{R}$ und stetigen Funktionen a_i, b_i für $i = 2,\dots,n$ heißt Normalbereich in $\mathbb{R}^n$ bezüglich der x_1-Achse.
Analog sind Normalbereiche in $\mathbb{R}^n$ bezüglich der anderen Koordinatenachsen definiert.

Satz 21.1 *Sei B_n ein Normalbereich in $\mathbb{R}^n$, B_{n+1} ein Normalbereich in $\mathbb{R}^{n+1}$ mit $B_{n+1} = \{(x_1,\dots,x_n,y) \in \mathbb{R}^{n+1} : (x_1,\dots,x_n) \in B_n \wedge a(x_1,\dots,x_n) \le y \le b(x_1,\dots,x_n)\}$, wobei $a, b \,:\, B_n \to \mathbb{R}$ stetig seien, und $f \,:\, B_{n+1} \to \mathbb{R}$ stetig. Dann ist auch das Parameterintegral $I \,:\, B_n \to \mathbb{R}$ mit*

$$I(x_1,\dots,x_n) = \int_{a(x_1,\dots,x_n)}^{b(x_1,\dots,x_n)} f(x_1,\dots,x_n,y)\; dy$$

stetig.

Beweisskizze
Durch analoge Argumentation wie im Beweis von Satz 18.7 zeigt man, dass jede stetige Funktion auf B_{n+1} und somit auch f gleichmäßig stetig auf B_{n+1} ist.
Nun betrachtet man zunächst die Situation, dass a und b konstant sind.
Dann konvergiert infolge der gleichmäßigen Stetigkeit von f für jede konvergente Folge $((x_1,\dots,x_n)_k)_{k\in\mathbb{N}}$ in B_n die Funktionenfolge $(g_k)_{k\in\mathbb{N}}$ mit $g_k :\; [a,b] \to \mathbb{R}$, $g_k(y) = f\left((x_1,\dots,x_n)_k, y\right)$ gleichmäßig auf $[a,b]$ gegen $g :\; [a,b] \to \mathbb{R}$ mit $g(y) = f\Big(\lim\limits_{k\to\infty}(x_1,\dots,x_n)_k, y\Big)$, so dass nach Satz 18.20 die Behauptung des Satzes für konstante a, b folgt.
Für nicht konstante a, b zeigt man nun die Stetigkeit von I in $(x_1,\dots,x_n)$, indem man für $(\tilde{x}_1,\dots,\tilde{x}_n) \ne (x_1,\dots,x_n)$ das jeweilige Integrationsintervall $[a(\tilde{x}_1,\dots,\tilde{x}_n), b(\tilde{x}_1,\dots,\tilde{x}_n)]$ in $[a(\tilde{x}_1,\dots,\tilde{x}_n), b(\tilde{x}_1,\dots,\tilde{x}_n)] \cap [a(x_1,\dots,x_n), b(x_1,\dots,x_n)]$ und $[a(\tilde{x}_1,\dots,\tilde{x}_n), b(\tilde{x}_1,\dots,\tilde{x}_n)] \setminus [a(x_1,\dots,x_n), b(x_1,\dots,x_n)]$ zerlegt, auf dem Schnittintervall die Behauptung für konstante a, b ausnutzt und zur Abschätzung des Integrals über die Differenzmenge die Stetigkeit von a und b, die Beschränktheit von f auf $[a(x_1,\dots,x_n), b(x_1,\dots,x_n)]$ sowie die gleichmäßige Stetigkeit von f ausnutzt.

Satz 21.2 *Sind die Voraussetzungen aus Satz 21.1 gegeben und existieren zusätzlich $\partial_{x_i} a,\; \partial_{x_i} b \in C^0(B_n,\mathbb{R})$ sowie $\partial_{x_i} f \in C^0(B_{n+1},\mathbb{R})$ für mindestens ein $i \in \{1,\dots,n\}$,* *dann existiert auch* $\partial_{x_i} I \in C^0(B_n, \mathbb{R})$ *und es gilt*

$$\begin{aligned}
\partial_{x_i} I(x_1, \ldots, x_n) &= \partial_{x_i} \int_{a(x_1,\ldots,x_n)}^{b(x_1,\ldots,x_n)} f(x_1, \ldots, x_n, y)\ dy \\
&= \int_{a(x_1,\ldots,x_n)}^{b(x_1,\ldots,x_n)} \partial_{x_i} f(x_1, \ldots, x_n, y)\ dy \\
&\quad + f(x_1, \ldots, x_n, b(x_1, \ldots, x_n))\, \partial_{x_i} b(x_1, \ldots, x_n) \\
&\quad - f(x_1, \ldots, x_n, a(x_1, \ldots, x_n))\, \partial_{x_i} a(x_1, \ldots, x_n).
\end{aligned}$$

Beweisskizze

Sei $\tilde{I}(v_1, \ldots, v_n, w_1, \ldots, w_n) := \int_{a(v_1,\ldots,v_n)}^{b(v_1,\ldots,v_n)} f(w_1, \ldots, w_n, y)\ dy$. Dann gilt nach der mehrdimensionalen Kettenregel:

$$\begin{aligned}
\partial_{x_i} I(x_1, \ldots, x_n) &= \partial_{v_i} \int_{a(v_1,\ldots,v_n)}^{b(v_1,\ldots,v_n)} f(w_1, \ldots, w_n, y)\ dy \\
&\quad + \partial_{w_i} \int_{a(v_1,\ldots,v_n)}^{b(v_1,\ldots,v_n)} f(w_1, \ldots, w_n, y)\ dy\,,
\end{aligned}$$

sofern die partiellen Ableitungen auf der rechten Seite existieren und diese an der Stelle $(v_1, \ldots, v_n, w_1, \ldots, w_n) = (x_1, \ldots, x_n, x_1, \ldots, x_n)$ ausgewertet werden.
Nach dem Hauptsatz der Differential- und Integralrechnung und der Kettenregel gilt

$$\begin{aligned}
\partial_{v_i} \int_{a(v_1,\ldots,v_n)}^{b(v_1,\ldots,v_n)} f(w_1, \ldots, w_n, y)\ dy &= f(w_1, \ldots, w_n, b(v_1, \ldots, v_n)) \partial_{v_i} b(v_1, \ldots, v_n) \\
&\quad - f(w_1, \ldots, w_n, a(v_1, \ldots, v_n)) \partial_{v_i} a(v_1, \ldots, v_n).
\end{aligned}$$

Es bleibt somit zu zeigen:

$$\partial_{w_i} \int_{a(v_1,\ldots,v_n)}^{b(v_1,\ldots,v_n)} f(w_1, \ldots, w_n, y)\ dy = \int_{a(v_1,\ldots,v_n)}^{b(v_1,\ldots,v_n)} \partial_{w_i} f(w_1, \ldots, w_n, y)\ dy.$$

Es gilt

$$\frac{1}{h}\Big(\int_{a(v_1,\ldots,v_n)}^{b(v_1,\ldots,v_n)} f(w_1, \ldots, w_i{+}h, \ldots, w_n, y)\ dy - \int_{a(v_1,\ldots,v_n)}^{b(v_1,\ldots,v_n)} f(w_1, \ldots, w_i, \ldots, w_n, y)\ dy \Big)$$

$$= \int_{a(v_1,\ldots,v_n)}^{b(v_1,\ldots,v_n)} \frac{1}{h}\, f(w_1, \ldots, w_i + h, \ldots, w_n, y) - f(w_1, \ldots, w_i, \ldots, w_n, y)\ dy$$

Aufgrund des Mittelwertsatzes der Differentialrechnung und da $\partial_{w_i} f$ auf B_{n+1} stetig und damit auf B_{n+1} gleichmäßig stetig ist, konvergiert die Funktion
$y \mapsto \frac{1}{h}\left(f(w_1, \ldots, w_i + h, \ldots, w_n, y) - f(w_1, \ldots, w_i, \ldots, w_n, y)\right)$ für $h \to 0$ gleichmäßig auf $[a, b]$ gegen $y \mapsto \partial_{w_i} f(w_1, \ldots, w_i, \ldots, w_n, y)$ so dass nach Satz 18.20 die Behauptung folgt.

Beispiel

$$\begin{aligned}\partial_x \int_{2x}^{x^2} xy\ dy &= \int_{2x}^{x^2} y\ dy + x\cdot x^2 \cdot 2x - x\cdot 2x\cdot 2\\ &= \frac{1}{2}\left(x^2\right)^2 - \frac{1}{2}\left(2x\right)^2 + 2x^4 - 4x^2\\ &= \frac{5}{2}x^4 - 6x^2\end{aligned}$$

Definition

a) *Für $I \subseteq \mathbb{R}^n$ mit $I \neq \emptyset$ heißt*

$$\chi_I : \mathbb{R}^n \to \mathbb{R},\ x \mapsto \begin{cases} 1, & x \in I\\ 0, & x \notin I\end{cases}$$

die charakteristische Funktion von I.

b) *Eine Menge der Form $\prod_{i=1}^{n}[a_i, b_i] \subseteq \mathbb{R}^n$ heißt beschränkter abgeschlossener n-dimensionaler Quader und eine Menge der Form $\prod_{i=1}^{n}]a_i, b_i[\subseteq \mathbb{R}^n$ heißt beschränkter offener n-dimensionaler Quader.*

c) *Sei $Q := \prod_{i=1}^{n}[a_i, b_i]$. Eine Funktion $f : Q \to \mathbb{R}$ heißt Treppenfunktion auf Q, falls es eine endliche Unterteilung (Zerlegung, Partition) $a_i = x_{i0} < x_{i1} < \ldots < x_{im_i} = b_i$ von jedem Intervall $[a_i, b_i]$ gibt, so dass f auf jedem beschränkten offenen Quader $Q_j := \prod_{i=1}^{n}]x_{i(j_i-1)}, x_{ij_i}[$ mit $j = (j_1, \ldots, j_n) \in \prod_{i=1}^{n}\{1, \ldots, m_i\}$ konstant ist. (Die Funktionswerte von f auf den sogenannten Rändern aller Q_j, d.h., auf $Q \setminus \bigcup_{j\in\prod_{i=1}^{n}\{1,\ldots,m_i\}} Q_i$ können beliebig sein).*

Die Menge aller Treppenfunktionen auf Q sei mit $\mathcal{T}(Q)$ bezeichnet.

d) *Zwei Treppenfunktionen $f, g : Q \to \mathbb{R}$ heißen fast überall gleich, kurz : $f = g$ f.ü., falls $f(x) \neq g(x)$ höchstens für Punkte x auf den Rändern von endlich vielen beschränkten offenen Quadern $Q_j \subseteq Q$.*

e) *Ist $Q = \prod_{i=1}^{n}[a_i, b_i]$, $f \in \mathcal{T}(Q)$ und $f = \sum_{j\in\prod_{i=1}^{n}\{1,\ldots,m_i\}} c_j \chi_{Q_j}$ f.ü., wobei Q_j wie in c) ist, dann heißt*

$$\int_Q f(x_1, \ldots, x_n)\ dx_1 \ldots dx_n := \sum_{j\in\prod_{i=1}^{n}\{1,\ldots,m_i\}} c_j \prod_{i=1}^{n}(x_{ij_i} - x_{i(j_i-1)})$$

das (bestimmte) Integral von f über Q. (Die Definition des Integrals ist wie im Eindimensionalen unabhängig von der Wahl der Zerlegung).

Definition *Sei Q ein beschränkter abgeschlossener n-dimensionaler Quader und $f : Q \to \mathbb{R}$ eine beschränkte Funktion.*

a)

$$\underline{\int_Q} f(x_1, \ldots, x_n) dx_1 \ldots dx_n := \sup \left\{ \int_Q \varphi(x_1, \ldots, x_n) dx_1 \ldots dx_n : \varphi \in \mathcal{T}(Q) \wedge \varphi \leq f \right\}$$

heißt Unterintegral von f über Q und

$$\overline{\int_Q} f(x_1, \ldots, x_n) dx_1 \ldots dx_n := \inf \left\{ \int_Q \varphi(x_1, \ldots, x_n) dx_1 \ldots dx_n : \varphi \in \mathcal{T}(Q) \wedge \varphi \geq f \right\}$$

heißt Oberintegral von f über Q.

b) f heißt Riemann-integrierbar über Q, falls

$$\underline{\int_Q} f(x_1, \ldots, x_n)\ dx_1 \ldots dx_n = \overline{\int_Q} f(x_1, \ldots, x_n)\ dx_1 \ldots dx_n.$$

In diesem Fall heißt

$$\begin{aligned} \int_Q f(x_1, \ldots, x_n)\ dx_1 \ldots dx_n &:= \underline{\int_Q} f(x_1, \ldots, x_n)\ dx_1 \ldots dx_n \\ &= \overline{\int_Q} f(x_1, \ldots, x_n)\ dx_1 \ldots dx_n \end{aligned}$$

das (bestimmte) Riemann-Integral von f über Q.

Bemerkungen

1) Die Aussagen aus den Sätzen bzw. Korollaren 18.1, 18.2, 18.3, 18.4, 18.6, 18.7, 18.8, 18.9 und 18.10 übertragen sich sinngemäß auf den mehrdimensionalen Fall.

2) Außer dem Riemann-Integral lassen sich auch das Cauchy-Integral und das Lebesgue-Integral auf den mehrdimensionalen Fall verallgemeinern.

3) Wie im Eindimensionalen schreiben wir von nun an integrierbar für Riemann-intergrierbar und Integral für Riemann-Integral.

Satz 21.3 (Satz von Fubini für beschränkte abgeschlossene Quader)
Sei $\prod_{i=1}^{n}[a_i, b_i]$ *und* $f : Q \to \mathbb{R}$ *stetig. Dann gilt*

$$\begin{aligned}
&\int_Q f(x_1, \ldots, x_n) dx_1 \ldots dx_n \\
= &\int_{a_1}^{b_1} \left(\int_{a_2}^{b_2} \left(\ldots \left(\int_{a_{n-1}}^{b_{n-1}} \left(\int_{a_n}^{b_n} f(x_1, \ldots, x_n) dx_n\right) dx_{n-1}\right) \ldots\right) dx_2\right) dx_1 \\
= &\int_{a_{\sigma(1)}}^{b_{\sigma(1)}} \left(\int_{a_{\sigma(2)}}^{b_{\sigma(2)}} \left(\ldots \left(\int_{a_{\sigma(n-1)}}^{b_{\sigma(n-1)}} \left(\int_{a_{\sigma(n)}}^{b_{\sigma(n)}} f(x_1, \ldots, x_n) dx_{\sigma(n)}\right) dx_{\sigma(n-1)}\right) \ldots\right) dx_{\sigma(2)}\right) dx_{\sigma(1)}
\end{aligned}$$

für jede bijektive Abbildung $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$.
Das Integral kann also durch sukzessive eindimensionale Integration bezüglich der einzelnen Koordinaten berechnet werden, wobei die Reihenfolge der Integration egal ist.

Beweisskizze
Wir diskutieren den Fall $n = 2$, für $n > 2$ folgt die Behauptung durch Induktion.
Sei $f : [a_1, b_1] \times [a_2, b_2] \to \mathbb{R}$ stetig und seien Q_{ij}, $i \in \{1, \ldots, m_i\}$, $j \in \{1, \ldots, m_j\}$ die zu einer endlichen Unterteilung $\{x_i : i \in \{1, \ldots, m_i\}\}$ von $[a_1, b_1]$ und $\{y_j : \{1, \ldots, m_j\}\}$ von $[a_2, b_2]$ gehörenden beschränkten offenen 2-dimensionalen Quader (Rechtecke). Dann gilt

$$\begin{aligned}
&\sum_{i=1}^{m_i}\sum_{j=1}^{m_j} \big(\inf_{(x,y)\in Q_{ij}} f(x, y)\big)(y_j - y_{j-1})(x_i - x_{i-1}) \\
\leq\ &\sum_{i=1}^{m_i}\sum_{j=1}^{m_j} \int_{y_{j-1}}^{y_j} f(x_i, y)\, dy\ (x_i - x_{i-1}) \\
=\ &\sum_{i=1}^{m_i} \int_{a_2}^{b_2} f(x_i, y)\, dy\ (x_i - x_{i-1}) \\
\leq\ &\sum_{i=1}^{m_i}\sum_{j=1}^{m_j} \big(\sup_{(x,y)\in Q_{ij}} f(x, y)\big)(y_j - y_{j-1})(x_i - x_{i-1})
\end{aligned}$$

Nach Satz 21.1 ist $x \mapsto \displaystyle\int_{a_2}^{b_2} f(x, y)\, dy$ stetig auf $[a_1, b_1]$ und somit integrierbar über $[a_1, b_1]$. Lassen wir die Feinheit der Unterteilung von $[a_1, b_1]$ und $[a_2, b_2]$ gegen 0 gehen, dann folgt aufgrund der obigen Ungleichung

$$\underline{\int_Q} f(x, y)\ dxdy\ \leq\ \int_{a_1}^{b_1} \left(\int_{a_2}^{b_2} f(x, y)\ dy\right) dx\ \leq\ \overline{\int_Q} f(x, y)\ dxdy$$

Da f als stetige Funktion integrierbar über Q ist, müssen das Unter-, das Ober- und das bestimmte Integral von f über Q den gleichen Wert haben, so dass

$$\int_Q f(x,y)\ dxdy \ = \ \int_{a_1}^{b_1} \left(\int_{a_2}^{b_2} f(x,y)\ dy \right) dx$$

folgt. Vertauscht man in der obigen Argumentation die Rollen von x und y, dann erhält man

$$\int_Q f(x,y)\ dxdy \ = \ \int_{a_2}^{b_2} \left(\int_{a_1}^{b_1} f(x,y)\ dx \right) dy\,.$$

$\square$

Beispiel
Sei $Q = [0,1] \times [0,2]$. Es gilt

$$\begin{aligned}
\int_Q 1 + xy\ dxdy \ &= \ \int_0^1 \int_0^2 1 + xy\ dydx \\
&= \ \int_0^1 \left[y + \frac{1}{2}xy^2 \right]_{y=0}^{y=2} dx \\
&= \ \int_0^1 2 + 2x\ dx \ = \ \left[2x + x^2\right]_{x=0}^{x=1} \ = \ 3
\end{aligned}$$

bzw.

$$\begin{aligned}
\int_Q 1 + xy\ dxdy \ &= \ \int_0^2 \int_0^1 1 + xy\ dxdy \\
&= \ \int_0^2 \left[x + \frac{1}{2}yx^2 \right]_{x=0}^{x=1} dy \\
&= \ \int_0^2 1 + \frac{1}{2}y\ dy \ = \ \left[y + \frac{1}{4}y^2 \right]_{y=0}^{y=2} \ = \ 3
\end{aligned}$$

Bemerkung
Der Satz von Fubini lässt sich noch auf eine größere Menge von Funktionen als die der auf Q stetigen Funktionen verallgemeinern. Dennoch gibt es Funktionen $f : Q \to \mathbb{R}$, für die mindestens eines der iterierten Integrale

$$\int_{a_{\sigma(1)}}^{b_{\sigma(1)}} \cdots \left(\int_{a_{\sigma(n)}}^{b_{\sigma(n)}} f(x_1, \ldots, x_n)\ dx_{\sigma(n)} \right) \ldots \ dx_{\sigma(1)}$$

existiert, bei Änderung der Integrationsreihenfolge sich aber der Wert des iterierten Integrals ändert und das Integral

$$\int_Q f(x_1, \ldots, x_n)\ dx_1 \ldots dx_n$$

nicht existiert. Ein Beispiel dafür ist die Funktion $f : [0,1] \times [0,1] \to \mathbb{R}$ mit

$$f(x,y) = \begin{cases} \dfrac{x^2 - y^2}{(x^2+y^2)^2}\,, & (x,y) \neq (0,0), \\ 0\,, & (x,y) = (0,0). \end{cases}$$

Unter Ausnutzung von $\dfrac{x^2 - y^2}{(x^2+y^2)^2} = \partial^2_{xy} \arctan\Bigl(\dfrac{x}{y}\Bigr)$ für $y \neq 0$ kann man direkt nachrechnen, dass gilt:

$$\int_0^1 \left(\int_0^1 f(x,y)\ dy \right) dx = \frac{\pi}{4} \neq -\frac{\pi}{4} = \int_0^1 \left(\int_0^1 f(x,y)\ dx \right) dy.$$

Dieses Phänomen ist letztendlich eine Konsequenz der Tatsache, dass für konvergente, aber nicht absolut konvergente Reihen der Umordnungssatz nicht gelten muss.

Definition *Sei $M \subseteq \mathbb{R}^n$ eine beschränkte Menge, $f : M \to \mathbb{R}$ eine beschränkte Funktion, Q ein beschränkter abgeschlossener n-dimensionaler Quader mit $M \subseteq Q$ und $\tilde{f} : Q \to \mathbb{R}$ mit*

$$\tilde{f}(x_1, \dots, x_n) = \begin{cases} f(x_1, \dots, x_n)\,, & (x_1, \dots, x_n) \in M, \\ 0\,, & (x_1, \dots, x_n) \in Q \setminus M. \end{cases}$$

f heißt integrierbar über M, falls $\tilde{f}$ integrierbar über Q ist. In diesem Fall heißt

$$\int_M f(x_1, \dots, x_n)\ dx_1 \dots dx_n := \int_Q \tilde{f}(x_1, \dots, x_n)\ dx_1 \dots dx_n$$

das (bestimmte) Integral von f über M.
Die Definition des Integrals von f über M ist unabhängig von der konkreten Wahl von $Q \supseteq M$.

Satz 21.4 *Sei B ein Normalbereich in $\mathbb{R}^n$ bezüglich der x_1-Achse und $f : B \to \mathbb{R}$ stetig. Dann gilt:*

$$\begin{aligned} &\int_B f(x_1, \dots, x_n)\ dx_1 \dots dx_n \\ = &\int_{a_1}^{b_1} \left(\int_{a_2(x_1)}^{b_2(x_1)} \left(\dots \left(\int_{a_{n-1}(x_1,\dots,x_{n-2})}^{b_{n-1}(x_1,\dots,x_{n-2})} \left(\int_{a_n(x_1,\dots,x_{n-1})}^{b_n(x_1,\dots,x_{n-1})} f(x_1, \dots, x_n)\, dx_n \right) dx_{n-1} \right) \dots \right) dx_2 \right) dx_1, \end{aligned}$$

wobei $a_1, \dots, a_n$, $b_1, \dots, b_n$ wie in der Definition von B sind.
Ist B ein Normalbereich in $\mathbb{R}^n$ bezüglich einer anderen Koordinatenachse, dann ergibt sich eine analoge Aussage für

$$\int_B f(x_1, \dots, x_n)\ dx_1 \dots dx_n.$$

Kann B als Normalbereich in $\mathbb{R}^n$ bezüglich verschiedener Koordinatenachsen aufgefasst werden, dann erhält man unabhängig von der konkreten Wahl einer dieser Koordinatenachsen denselben Wert für

$$\int_B f(x_1, \ldots, x_n)\ dx_1 \ldots\ dx_n.$$

Ist $g : B \to \mathbb{R}$ eine weitere Funktion mit $f = g$ f. ü., dann gilt

$$\int_B g(x_1, \ldots, x_n)\ dx_1 \ldots\ dx_n = \int_B f(x_1, \ldots, x_n)\ dx_1 \ldots\ dx_n.$$

Beweisstrategie
Durch ähnliche Argumentation wie im Beweis von Satz 21.3, wobei ebenfalls der Satz 21.1 (in diesem Fall mit nicht-konstanten a, b) ausgenutzt wird. □

Definition *Sei $B \subseteq \mathbb{R}^n$ ein Normalbereich. Dann heißt*

$$V_n(B) := \int_B 1\ dx_1 \ldots\ dx_n$$

das n-dimensionale Volumen von B.

Beispiel
Sei

$$\begin{aligned} B &= \left\{(x, y) \in \mathbb{R}^2 \,:\, 0 \le x \le 1\,,\, 0 \le y \le 2x\right\} \\ &= \left\{(x, y) \in \mathbb{R}^2 \,:\, 0 \le y \le 2\,,\, \frac{1}{2}y \le x \le 1\right\} \end{aligned}$$

B ist Normalbereich sowohl bezüglich der x-Achse als auch bezüglich der y-Achse. Daher gilt

$$\begin{aligned} \int_B 2xy\ dxdy &= \int_0^1 \left(\int_0^{2x} 2xy\ dy\right) dx \\ &= \int_0^1 \left[xy^2\right]_{y=0}^{y=2x}\ dx \\ &= \int_0^1 4x^3\ dx \;=\; \left[x^4\right]_{x=0}^{x=1} \;=\; 1 \end{aligned}$$

sowie

$$\begin{aligned} \int_B 2xy\ dxdy &= \int_0^2 \left(\int_{\frac{1}{2}y}^1 2xy\ dx\right) dy \\ &= \int_0^2 \left[yx^2\right]_{x=\frac{1}{2}y}^{x=1}\ dy \\ &= \int_0^2 y - \frac{1}{4}y^3\ dy \;=\; \left[\frac{1}{2}y^2 - \frac{1}{16}y^4\right]_{y=0}^{y=2} \;=\; 1 \end{aligned}$$

Beachte: Beim Ändern der Integrationsreihenfolge müssen die Integrationsgrenzen richtig abgeändert werden.

Satz 21.5 *Sei* $r : [a, b] \to \mathbb{R}_0^+$ *stetig und*

$$B = \left\{(x, y, z) \in \mathbb{R}^3 \, : \, a \leq x \leq b \, \wedge \, y^2 + z^2 \leq (r(x))^2\right\}.$$

Dann gilt

$$V_3(B) = \pi \int_a^b (r(x))^2 \; dx$$

Beweisstrategie
Stelle entweder B als Normalbereich bezüglich der x-Achse dar und wende Satz 21.4 an oder approximiere B mit immer größer werdender Genauigkeit durch Zylinder.

Beispiel
Sei B wie in Satz 21.5 mit $r : [0, h] \to \mathbb{R}_0^+$, $r(x) = \dfrac{R}{h}x$, wobei $h, R \in \mathbb{R}^+$ seien. Dann beschreibt B einen Kegel mit Radius R und Höhe h und es gilt

$$V_3(B) = \pi \int_0^h \frac{R^2}{h^2} x^2 \; dx = \pi \left[\frac{1}{3}\frac{R^2}{h^2}x^3\right]_0^h = \frac{1}{3}\pi R^2 h$$

Satz 21.6 (Transformationssatz) *Sei* $g : \mathbb{R}^n \supseteq \Omega \to g(\Omega) \subseteq \mathbb{R}^n$ *ein Diffeomorphismus, d.h.,* g *ist bijektiv und stetig differenzierbar in* Ω *und die Umkehrfunktion* g^{-1} *ist stetig differenzierbar in* $g(\Omega)$*. Sei* $M \subseteq g(\Omega)$ *beschränkt und* f *integrierbar über* M*. Dann gilt*

$$\int_M f(x_1, \ldots, x_n) \, dx_1 \ldots \, dx_n = \int_{g^{-1}(M)} f\left(g(u_1, \ldots, u_n)\right) |\det\left(g'(u_1, \ldots, u_n)\right)| \, du_1 \ldots \, du_n.$$

Beweisidee
Da g stetig differenzierbar ist, gilt

$$\forall u, v \in \Omega \, : \, g(v) = g(u) + g'(u)(v - u) + o(v - u)$$

für $v \to u$.
Die Abbildung $v \mapsto g(u) + g'(u)(v - u)$ bildet jeden n-dimensionalen Spat in Ω mit n-dimensionalem Volumen V_n auf einen n-dimensionalen Spat mit n-dimensionalem Volumen $|\det(g'(u))| V_n$ ab, siehe Kapitel 11. Folglich gilt für $(x_1, \ldots, x_n) \in M$ mit $(x_1, \ldots, x_n) = g(u_1, \ldots, u_n)$ und $Q = \prod\limits_{i=1}^{n} [x_i - \varepsilon, x_i + \varepsilon]$:

$$f(x_1, \ldots, x_n) V_n(Q) = f(g(u_1, \ldots, u_n)) |\det(g'(u_1, \ldots, u_n))| V_n(g^{-1}(Q)) + o(\varepsilon^n)$$

für $\varepsilon \to 0$.

Da sich jedes bestimmte Integral als Grenzwert von Riemannschen Summen, deren Feinheit gegen 0 geht, darstellen lässt, erhält man daraus (nach einer längeren und technischen Argumentation) die Behauptung.

Beispiel
Sei $M = \{(x,y) \in \mathbb{R}^2 \,:\, y \geq 0 \,\wedge\, 1 \leq x^2 + y^2 \leq 4\}$ und $f : \mathbb{R}^2 \to \mathbb{R}$ mit $f(x,y) = e^{-(x^2+y^2)}$. Dann ist f integrierbar über M.
Sei $\Omega =]1-\varepsilon, 2+\varepsilon[\times]-\varepsilon, \pi+\varepsilon[$ für ein festes $\varepsilon \in]0,1[$ und $g : \Omega \to g(\Omega)$ mit $(r,\varphi) \mapsto \begin{pmatrix} r\cos\varphi \\ r\sin\varphi \end{pmatrix}$.
Dann kann man direkt nachrechnen, dass g ein Diffeomorphismus ist mit $M \subseteq g(\Omega)$, $g^{-1}(M) = [1,2] \times [0,\pi]$ und
$g'(r,\varphi) = \begin{pmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{pmatrix} \implies |\det(g'(r,\varphi))| = r$.
Daher gilt nach dem Transformationssatz:

$$\begin{aligned}
\int\limits_M e^{-(x^2+y^2)}\ dxdy &= \int\limits_{[1,2]\times[0,\pi]} e^{-r^2} r\ drd\varphi \\
&= \int_1^2 \left(\int_0^\pi re^{-r^2}\ d\varphi \right) dr \\
&= \pi \int_1^2 re^{-r^2}\ dr = \frac{\pi}{2} \int_1^4 e^{-s}\ ds \\
&= -\frac{\pi}{2} \left[e^{-s}\right]_1^4 = \frac{\pi}{2}(e^{-1} - e^{-4})\,.
\end{aligned}$$

Die Abbildung g nennt man auch Parametrisierung von $g(\Omega)$ durch Polarkoordinaten.

Bemerkung
Zur Berechnung des Integrals einer Funktion f über krummlinig berandete Mengen M ist es häufig hilfreich, eine geeignete Parametrisierung g von M zu finden, so dass

$$\int_{g^{-1}(M)} f\left(g(u_1, \ldots, u_n)\right) |\det\left(g'(u_1, \ldots, u_n)\right)|\ du_1 \ldots\ du_n$$

einfacher zu berechnen ist als

$$\int_M f(x_1, \ldots, x_n)\ dx_1 \ldots\ dx_n.$$

Am häufigsten verwendet werden Parametrisierungen durch Polarkoordinaten, Zylinderkoordinaten oder Kugelkoordinaten und manchmal auch Parametrisierungen durch parabolische oder hyperbolische Koordinaten. Für nähere Details sei auf die Literatur verwiesen.

Teil V

Gewöhnliche Differentialgleichungen

22 Elementar lösbare gewöhnliche Differentialgleichungen erster Ordnung

Definition

a) Sei $f : \mathbb{R}^n \supseteq D \to \mathbb{R}$ eine Funktion. Die Gleichung

$$y^{(n)}(x) = f(x, y(x), y'(x), \ldots, y^{(n-1)}(x)) \qquad (1)$$

heißt (explizite) gewöhnliche Differentialgleichung n-ter Ordnung.
Kurzschreibweise: $y^{(n)} = f(x, y, y', \ldots, y^{(n-1)})$

b) Jede n-mal differenzierbare Funktion $y : I \to \mathbb{R}$, wobei $I \subseteq \mathbb{R}$ ein Intervall ist, für die $\{(x, y(x), y'(x), \ldots, y^{(n-1)}(x)) : x \in I\} \subseteq D$ ist und die für alle $x \in I$ die Gleichung (1) löst, heißt Lösung der gewöhnlichen Differentialgleichung (1).
Eine Lösung von (1) heißt global, falls $I = \mathbb{R}$ ist, und lokal, falls $I \neq \mathbb{R}$ ist.
Die Menge aller Lösungen von (1) heißt allgemeine Lösung von (1).

c) Sei $(x_0, y_0, y_1, \ldots, y_{n-1}) \in D$. Dann heißen die Gleichungen

$$\begin{cases} y(x_0) = y_0\,, \\ y'(x_0) = y_1\,, \\ \quad\vdots \\ y^{(n-1)}(x_0) = y_{n-1}\,, \end{cases} \qquad (2)$$

Anfangsbedingungen für Lösungen von (1) und das Gleichungssystem (1), (2) ein Anfangswertproblem für (1).

d) Eine Lösung von (1), die auch (2) löst, heißt Lösung des Anfangswertproblems (1), (2).

Bemerkungen

1) Sind in (1) die Wertebereiche von $f, y, y', \ldots, y^{(n)}$ nicht $\mathbb{R}$, sondern $\mathbb{R}^m$, und ist x weiterhin aus einem Intervall $I \subseteq \mathbb{R}$, dann erhält man ein System von m gewöhnlichen Differentialgleichungen n-ter Ordnung. Anfangsbedingungen und Lösungen sind für Systeme von gewöhnlichen Differentialgleichungen analog zum Fall einer gewöhnlichen Differentialgleichung definiert.

2) Eine Gleichung der Form

$$g(x, y(x), y'(x), \ldots, y^{(n)}(x)) = 0\,, \qquad (3)$$

wobei x ebenfalls aus einem Intervall $I \subseteq \mathbb{R}$ ist, nennt man implizite gewöhnliche Differentialgleichung n-ter Ordnung bzw. implizites System gewöhnlicher Differentialgleichungen n-ter Ordnung.

3) Ersetzt man in (1) oder (3) x durch $(x_1, \ldots, x_k) \in \mathbb{R}^k$ und die eindimensionalen Ableitungen von y durch partielle Ableitungen, dann erhält man eine partielle Differentialgleichung bzw. ein System partieller Differentialgleichungen.

Geometrische Interpretation
Sei $f : \mathbb{R}^2 \supseteq D \to \mathbb{R}$. Dann bestimmt die Differentialgleichung $y' = f(x,y)$ ein sogenanntes Richtungsfeld in D, d.h., in jedem Punkt $(x,y) \in D$ wird durch $y' = f(x,y)$ eine Steigung y' und damit ein Vektor $\begin{pmatrix} 1 \\ y' \end{pmatrix}$ vorgegeben. Gesucht ist eine differenzierbare Funktion y, deren Graph in jedem seiner Punkte (x,y) die Steigung y' und damit $\begin{pmatrix} 1 \\ y' \end{pmatrix}$ als Tangentialvektor hat.

Beispiel
$y' = -\dfrac{x}{y}$ für $(x,y) \in \mathbb{R} \times \mathbb{R}^+$

Es gilt $\begin{pmatrix} 1 \\ y' \end{pmatrix} = \begin{pmatrix} 1 \\ -\frac{x}{y} \end{pmatrix} \perp \begin{pmatrix} x \\ y \end{pmatrix}$ für alle $(x,y) \in \mathbb{R} \times \mathbb{R}^+$.

Die Lösungen sind somit alle Halbkreise der Form

$$y(x) = \sqrt{c - x^2}, \quad |x| < \sqrt{c},\ c \in \mathbb{R}^+.$$

Definition *Seien $I, J \subseteq \mathbb{R}$ offene Intervalle und $g : I \to \mathbb{R},\ h : J \to \mathbb{R}$ stetige Funktionen. Die Differentialgleichung*

$$y' \;=\; g(x)\, h(y)$$

heißt gewöhnliche Differentialgleichung mit getrennten Variablen.

Die Lösungen von gewöhnlichen Differentialgleichungen mit getrennten Variablen kann man folgendermaßen bestimmen.

1. Fall: Sei $y_0 \in J$ mit $h(y_0) = 0$.
Dann ist $y : I \to \mathbb{R}$ mit $y(x) = y_0$ Lösung von $y' \;=\; g(x)\, h(y)$.

2. Fall: Sei $y_0 \in J$ mit $h(y_0) \neq 0$.
Dann gilt: Für jede differenzierbare Funktion $y : I \supseteq I_\star \to \mathbb{R}$ mit $y(x_0) = y_0$ existiert aufgrund der Stetigkeit von y und von h ein offenes Intervall I_y mit $x_0 \in I_y$, so dass für alle $x \in I_y$ entweder $h(y(x)) > 0$ oder $h(y(x)) < 0$ gilt. Außerdem gilt für alle

$x, \widetilde{x} \in I_y$:

$$
\begin{aligned}
& \frac{dy}{dx}(x) = g(x)\ h(y(x)) \ \wedge\ y(x_0) = y_0 \\
\Leftrightarrow\ & \frac{\frac{dy}{dx}(x)}{h(y(x))} = g(x) \ \wedge\ y(x_0) = y_0 \\
\Leftrightarrow\ & \int_{x_0}^{\widetilde{x}} \frac{\frac{dy}{dx}(x)}{h(y(x))}\ dx = \int_{x_0}^{\widetilde{x}} g(x)\ dx \ \wedge\ y(x_0) = y_0 \\
\Leftrightarrow\ & \int_{y_0}^{y(\widetilde{x})} \frac{1}{h(y)}\ dy = \int_{x_0}^{\widetilde{x}} g(x)\ dx
\end{aligned}
$$

Sei $G : I_y \to G(I_y),\ x \mapsto \displaystyle\int_{x_0}^{x} g(t)\ dt$. Außerdem sei $J_y := y(I_y)$ und $Q : J_y \to Q(J_y)$, $y \mapsto \displaystyle\int_{y}^{y_0} \frac{1}{h(t)}\ dt$. Dann ist $Q'(y) = \dfrac{1}{h(y)}$ und somit für alle $y \in J_y$ entweder $Q'(y) > 0$ oder $Q'(y) < 0$. Folglich ist Q entweder streng monoton steigend oder streng monoton fallend auf ganz I_y, wobei $Q'(y) \neq 0$ ist.
Daher ist Q nach Satz 16.7 invertierbar mit stetig differenzierbarer Umkehrfunktion $Q^{-1} : Q(J_y) \to J_y$. Also gilt :

$$
\begin{aligned}
& y'(x) = g(x)\, h(y(x)) \ \wedge\ y(x_0) = y_0 \\
\Leftrightarrow\ & y(x) = Q^{-1}(G(x)) \quad \text{für alle } \ x \in (G^{-1} \circ Q \circ y)(I_y) \subseteq I_y
\end{aligned}
$$

Wir erhalten also:

Satz 22.1 *Seien $I, J \subseteq \mathbb{R}$ offene Intervalle und $g : I \to \mathbb{R},\ h : J \to \mathbb{R}$ stetige Funktionen. Außerdem sei $x_0 \in I$ und $y_0 \in J$ mit $h(y_0) \neq 0$.*
Dann existiert ein offenes Intervall $\widetilde{I} \subseteq I$ mit $x_0 \in \widetilde{I}$, so dass

$$
\begin{cases}
\quad y' = g(x)\, h(y)\,, \\
y(x_0) = y_0\,,
\end{cases}
$$

auf $\widetilde{I}$ eine eindeutige Lösung $y : \widetilde{I} \to \mathbb{R}$ besitzt.
Diese Lösung ergibt sich durch Auflösen der Gleichung

$$
\int_{y_0}^{y(x)} \frac{1}{h(t)}\ dt = \int_{x_0}^{x} g(t)\ dt
$$

nach $y(x)$.

Bemerkungen

1) Eine alternative Strategie zur Bestimmung der Lösung aus Satz 22.1 ist, die Gleichung
$$\int \frac{1}{h(y)}\, dy = \int g(x)\, dx + c$$
nach y aufzulösen und dann c so zu wählen, dass $y(x_0) = y_0$ gilt.

2) Die Lösung aus Satz 22.1 kann man allerdings nicht in jedem Fall in geschlossener Form angeben.

Beispiele

1) $\begin{cases} y' = -\dfrac{x}{y}, & (x,y) \in \mathbb{R} \times \mathbb{R}^+, \\ y(x_0) = y_0, & (x_0, y_0) \in \mathbb{R} \times \mathbb{R}^+. \end{cases}$

Es ist $\dfrac{1}{y_0} \neq 0$ für alle $y_0 \in \mathbb{R}^+$. Daher erhalten wir:
$$\begin{aligned} & \int_{y_0}^{y(x)} t\, dt = \int_{x_0}^{x} -t\, dt \\ \Leftrightarrow\quad & \frac{1}{2}(y(x))^2 - \frac{1}{2}y_0^2 = \frac{1}{2}x_0^2 - \frac{1}{2}x^2 \\ \Leftrightarrow\quad & (y(x))^2 = x_0^2 + y_0^2 - x^2 \\ \overset{y>0}{\Leftrightarrow}\quad & y(x) = \sqrt{x_0^2 + y_0^2 - x^2} \qquad \text{für} \quad |x| < \sqrt{x_0^2 + y_0^2} \end{aligned}$$

2) Bewegungsgleichung des freien Falls:
$$\begin{cases} h''(t) = -g, & h > 0, \\ h(t_0) = h_0, & h_0 > 0, \\ h'(t_0) = v_0, & v_0 \in \mathbb{R}, \end{cases}$$
wobei $h(t)$ die Höhe zur Zeit t, g die Erdbeschleunigung und v_0 die Geschwindigkeit zur Zeit t_0 ist. Es gilt
$$\begin{aligned} h'(t) &= v_0 - \int_{t_0}^{t} g\, ds = v_0 - g(t - t_0) \\ h(t) &= h_0 + \int_{t_0}^{t} v_0 - g(s - t_0)\, ds \\ &= h_0 + v_0(t - t_0) - \frac{1}{2}g(t - t_0)^2 \end{aligned}$$
für alle $t \in I_{max}$, wobei I_{max} das größtmögliche Intervall ist mit $t_0 \in I_{max}$ und $h > 0$ auf ganz I_{max}.

3) $\begin{cases} y'(t) = \alpha y(t), & \alpha \in \mathbb{R} \setminus \{0\},\ y \geq 0, \\ y(t_0) = y_0, & y_0 \geq 0. \end{cases}$

Dieses Anfangswertproblem modelliert zahlreiche Wachstums-, Zerfalls-, Diffusions- und Wärmeleitungsprozesse in der Natur.
Sei $y_0 = 0$. Dann gilt $y(t) = 0$ für alle $t \in \mathbb{R}$.
Sei $y_0 > 0$ und $y(t) > 0$. Dann erhalten wir:

$$\begin{aligned} & \int \frac{1}{y}\ dy = \int \alpha\ dt + C \\ \overset{y>0}{\Leftrightarrow}\quad & lny = \alpha t + C \\ \Leftrightarrow\quad & y = e^C e^{\alpha t} \end{aligned}$$

Die Anfangsbedingung ist genau dann erfüllt wenn $e^C = y_0 e^{-\alpha t_0}$ ist. Also folgt

$$y(t) = y_0 e^{\alpha(t-t_0)} \qquad \text{für alle}\ \ t \in \mathbb{R}.$$

4) $\begin{cases} y'(t) = -\frac{1}{y(t)}, & y > 0, \\ y(0) = y_0, & y_0 > 0. \end{cases}$

Wir erhalten

$$\begin{aligned} & \int y\ dy = -\int 1\ dx + C \\ \Leftrightarrow\quad & \frac{1}{2} y^2 = -x + C \\ \overset{y>0}{\Leftrightarrow}\quad & y = \sqrt{2(C-x)} \end{aligned}$$

Aufgrund der Anfangsbedingung folgt

$$y(x) = \sqrt{y_0^2 - 2x}\ \ \text{für alle}\ x < \frac{1}{2} y_0^2.$$

5) $\begin{cases} y'(x) = y^2(x), & y > 0, \\ y(0) = y_0. \end{cases}$

Wir erhalten

$$\begin{aligned} & \int y^{-2}\ dy = \int 1\ dx + C \\ \Leftrightarrow\quad & -\frac{1}{y} = x + C \\ \Leftrightarrow\quad & y = -\frac{1}{x+C} \end{aligned}$$

Aufgrund der Anfangsbedingung folgt

$$y(x) = \frac{1}{y_0^{-1} - x}\ \text{für alle}\ x < y_0^{-1}.$$

6) $h'(t) = -\sqrt{h(t)}, \quad h(t) \geq 0. \qquad\qquad (*)$
Diese Gleichung modelliert die Veränderung der Höhe h des Wasserspiegels in einem auslaufenden Wassertank.
Für $h(t) > 0$ erhalten wir

$$\begin{aligned} & \int h^{-\frac{1}{2}} \, dh = -\int 1 \, dt + C \\ \Leftrightarrow \quad & 2\sqrt{h} = -t + C \\ \overset{h \geq 0}{\Leftrightarrow} \quad & h = \frac{1}{4}(C-t)^2 \end{aligned}$$

Also ist für jedes $C \in \mathbb{R}$ die Funktion

$$h(t) = \frac{1}{4}(C-t)^2$$

für alle $t < C$ eine Lösung von $(*)$. Setzt man h für $t \geq C$ durch 0 fort, dann erhält man eine Lösung von $(*)$, die für alle $t \in \mathbb{R}$ existiert.
Folglich bildet

$$h(t) = \begin{cases} \frac{1}{4}(C-t)^2, & t < C, \\ 0, & t \geq C, \end{cases}$$

zusammen mit

$$h(t) = 0 \quad \text{für alle } t \in \mathbb{R}$$

die allgemeine Lösung von (*). Infolgedessen hat das Anfangswertproblem

$$\begin{cases} h'(t) = -\sqrt{h(t)}, & h(t) \geq 0, \\ h(t_0) = h_0, \end{cases}$$

für $h_0 > 0$ eine eindeutige Lösung, für $h_0 = 0$ unendlich viele Lösungen.

Bemerkungen

1) Die sogenannte Ähnlichkeitsdifferentialgleichung

$$y' = f\left(\frac{y}{x}\right), \quad x \neq 0,$$

kann mit Hilfe der Substitution

$$u(x) := \frac{y(x)}{x}$$

zu einer Differentialgleichung mit getrennten Variablen der Form

$$u' = \frac{1}{x}(f(u) - u)$$

umgeformt werden. Hat man diese Gleichung gelöst, dann erhält man durch Rücksubstitution

$$y(x) := xu(x)$$

die Lösungen der Ähnlichkeitsdifferentialgleichung.

2) Ebenso kann eine Differentialgleichung der Form

$$y' = f(ax + by + c)$$

mit Hilfe der Substitution

$$u(x) := ax + by(x) + c$$

zur Gleichung

$$u' = a + bf(u)$$

umgeformt werden.

3) Allgemeiner können Gleichungen der Form

$$y' = f\Big(\frac{ax + by + c}{\alpha x + \beta x + \gamma}\Big), \quad \alpha x + \beta x + \gamma \neq 0,$$

mit Hilfe geeigneter Substitutionen zu Gleichungen mit getrennten Variablen umgeformt werden. Für Details sei auf die Literatur verwiesen.

Bemerkung
Manchmal ist es sinnvoll, gewöhnliche Differentialgleichungen lokal durch Einsetzen eines Potenzreihenansatzes in die Differentialgleichung, gliedweisem Differenzieren und anschließendem Koeffizientenvergleich zu lösen.

Beispiel
$(x - 2x^2)y' = 2xy$
Einsetzen von $y(x) = \sum\limits_{n=0}^{\infty} a_n x^n$ in die Gleichung und gliedweises Differenzieren liefert:
$0 = (x - 2x^2)\sum\limits_{n=1}^{\infty} n a_n x^{n-1} - 2x \sum\limits_{n=0}^{\infty} a_n x^n = \sum\limits_{n=1}^{\infty} n(a_n - 2a_{n-1})x^n$
Koeffizientenvergleich liefert:
$a_0 \in \mathbb{R}$ ist beliebig wählbar,
$\forall n \in \mathbb{N} : a_n = 2a_{n-1}$,
und somit $\forall n \in \mathbb{N}_0 : a_n = 2^n a_0$.
Daher folgt
$y(x) = \sum\limits_{n=0}^{\infty} 2^n a_0 x^n = \dfrac{a_0}{1 - 2x}$
mit Konvergenzradius $R = \frac{1}{2}$ für $a_0 \neq 0$ und $R = \infty$ für $a_0 = 0$.
Setzt man $y(x) = \dfrac{a_0}{1 - 2x}$ mit $a_0 \neq 0$ in die Differentialgleichung ein, dann erhält man, dass $y(x) = \dfrac{a_0}{1 - 2x}$ die Gleichung $(x - 2x^2)y' = 2xy$ sogar auf jedem Intervall $I \subseteq \mathbb{R}$ mit $\frac{1}{2} \notin I$ löst.
(Dasselbe Ergebnis würde man bei diesem speziellen Beispiel auch erhalten, wenn man die Differentialgleichung durch Trennung der Variablen, d.h., mit Hilfe von Satz 22.1, lösen würde.)

23 Lineare gewöhnliche Differentialgleichungen

Definition *Sei $I \subseteq \mathbb{R}$ ein offenes Intervall und seien $a, b : I \to \mathbb{R}$ stetige Funktionen. Die Differentialgleichung*

$$y' = a(x)y + b(x)$$

heißt explizite reelle lineare gewöhnliche Differentialgleichung erster Ordnung.
Ist $b(x) = 0$ für alle $x \in I$, dann nennt man die lineare Differentialgleichung homogen, anderenfalls inhomogen.

Satz 23.1 *Mit den Bezeichnungen der obigen Definition gilt: Sei $x_0 \in I$ und $y_0 \in \mathbb{R}$. Dann existiert genau eine Lösung $y : I \to \mathbb{R}$ des Anfangswertproblems*

$$\begin{cases} y' = a(x)y\,, \\ y(x_0) = y_0\,, \end{cases}$$

nämlich

$$y(x) = y_0\, e^{\int_{x_0}^{x} a(t)\,dt}\,.$$

Die allgemeine Lösung der homogenen linearen Differentialgleichung $y' = a(x)y$ ist

$$y_{hom}(x) = c\, e^{\int_{x_0}^{x} a(t)\,dt}\,, \quad c \in \mathbb{R},\ x \in I.$$

Beweis
1. Beweismöglichkeit:
Man fasst die Gleichung $y' = a(x)y$ als Gleichung mit getrennten Variablen auf und löst sie wie in Kapitel 22. Dies liefert die Behauptung.

2. Beweismöglichkeit:
Existenz der Lösung des Anfangswertproblems:
Man setzt $y(x) = y_0\, e^{\int_{x_0}^{x} a(t)\,dt}$ in das Anfangswertproblem ein und rechnet direkt nach, dass es sich um eine Lösung handelt.
Eindeutigkeit der Lösung des Anfangswertproblems:
Angenommen, $\varphi : I \to \mathbb{R}$ sei eine weitere Lösung des Anfangswertproblems.
Sei $w : I \to \mathbb{R}$, $w(x) = e^{-\int_{x_0}^{x} a(t)\,dt}$ und $\psi := \varphi w$. Dann gilt für alle $x \in I$:
$\psi'(x) = \varphi'(x)w(x) + \varphi(x)w'(x) = a(x)\varphi(x)w(x) - \varphi(x)a(x)w(x) = 0.$
Somit ist ψ auf I konstant, so dass
$\psi(x) = \psi(x_0) = \varphi(x_0)w(x_0) = y_0$
für alle $x \in I$ gilt. Dies impliziert aber
$\forall x \in I : \varphi(x) = y_0\, e^{\int_{x_0}^{x} a(t)\,dt} = y(x).$
Da $y_0 \in \mathbb{R}$ beliebig gewählt werden kann, folgt aus der eben bewiesenen Lösungsformel für das Anfangswertproblem die behauptete Lösungsformel für die allgemeine Lösung. $\square$

Beispiel
$y' = 2xy, \quad x \in \mathbb{R}$
Es gilt:
$y_{hom}(x) = c\, e^{\int_0^x 2t\, dt} = ce^{x^2}, \quad c \in \mathbb{R},$
für alle $x \in \mathbb{R}$ und
$y(x) = y_0 e^{-x_0^2} e^{x^2}$
ist die eindeutige Lösung von
$\begin{cases} y' = 2xy, & x \in \mathbb{R}, \\ y(x_0) = y_0. \end{cases}$

Satz 23.2 *Mit den Bezeichnungen der obigen Definition gilt: Sei $x_0 \in I$ und $y_0 \in \mathbb{R}$. Dann existiert genau eine Lösung $y : I \to \mathbb{R}$ des Anfangswertproblems*

$$\begin{cases} y' = a(x)y + b(x), \\ y(x_0) = y_0, \end{cases}$$

nämlich

$$y(x) = \Big(y_0 + \int_{x_0}^{x} e^{-\int_{x_0}^{t} a(s)\, ds} b(t)\; dt\Big) e^{\int_{x_0}^{x} a(t)\, dt} = y_0 e^{\int_{x_0}^{x} a(t)\, dt} + \int_{x_0}^{x} e^{\int_t^x a(s)\, ds} b(t)\; dt.$$

Die allgemeine Lösung der inhomogenen linearen Differentialgleichung $y' = a(x)y + b(x)$ ist

$$y_{inh}(x) = y_{hom}(x) + y_p(x), \quad x \in I,$$

wobei y_{hom} wie in Satz 23.1 ist und

$$y_p(x) = \int_{x_0}^{x} e^{\int_t^x a(s)\, ds} b(t)\; dt.$$

Beweis
Wir machen den Ansatz
$y(x) = c(x)\, e^{\int_{x_0}^{x} a(t)\, dt}.$
(Da $e^{\int_{x_0}^{x} a(t)\, dt} \neq 0$ für alle $x \in I$ ist, lässt sich jede Funktion $y : I \to \mathbb{R}$ in dieser Form schreiben.)
Einsetzen des Ansatzes in die Differentialgleichung liefert:
$(c'(x) + a(x)c(x))\, e^{\int_{x_0}^{x} a(t)\, dt} = a(x)c(x)e^{\int_{x_0}^{x} a(t)\, dt} + b(x)$
$\iff c'(x) = e^{-\int_{x_0}^{x} a(t)\, dt} b(x)$
$\iff c(x) = c(x_0) + \int_{x_0}^{x} e^{-\int_{x_0}^{t} a(s)\, ds} b(t)\, dt$
Unter Berücksichtigung der Anfangsbedingung $y(x_0) = y_0$ erhalten wir daraus die behauptete Lösungsformel für das Anfangswertproblem und aus dieser folgt die behauptete Lösungsformel für die allgemeine Lösung. $\square$

Bemerkungen

1) Die Lösungsformel für das Anfangswertproblem der inhomogenen linearen gewöhnlichen Differentialgleichung nennt man auch Variation der Konstanten, weil der konstante Faktor y_0 vor $e^{\int_{x_0}^{x} a(t)\,dt}$ in der Lösungsformel für das zugehörige homogene Anfangswertproblem durch einen von x abhängigen Faktor ersetzt wird.

2) Die Funktion y_p aus Satz 23.2 ist eine spezielle (partikuläre) Lösung der inhomogenen linearen Differentialgleichung $y' = a(x)y + b(x)$, nämlich diejenige Lösung, die $y(x_0) = 0$ erfüllt. Aus Satz 23.2 folgt, dasss man y_p in der Lösungsformel für y_{inh} durch jede andere spezielle Lösung von $y' = a(x)y + b(x)$ ersetzen kann und die Lösungsformel für y_{inh} trotzdem gültig bleibt.

Beispiel

$y' = 2xy + x, \quad x \in \mathbb{R}$

Es gilt:

$$y_p(x) = \int_0^x e^{\int_t^x 2s\,ds}\, t\,dt = \int_0^x e^{x^2-t^2}\, t\,dt = \frac{1}{2}e^{x^2}\int_0^{x^2} e^{-\tau}d\tau = \frac{1}{2}(e^{x^2}-1)$$

ist eine spezielle Lösung von $y' = 2xy + x$. Somit erhalten wir:

$$\begin{aligned} y_{inh}(x) &= y_{hom}(x) + y_p(x) \\ &= ce^{x^2} + \frac{1}{2}(e^{x^2}-1), \quad c \in \mathbb{R} \\ &= \tilde{c}e^{x^2} - \frac{1}{2}, \quad \tilde{c} \in \mathbb{R} \end{aligned}$$

für alle $x \in \mathbb{R}$. Folglich ist

$$y(x) = \Big(y_0 + \frac{1}{2}\Big)e^{-x_0^2}e^{x^2} - \frac{1}{2}$$

die eindeutige Lösung von

$$\begin{cases} y' = 2xy + x, \quad x \in \mathbb{R}, \\ y(x_0) = y_0. \end{cases}$$

Bemerkungen

1) Die sogenannte Bernoulli-Differentialgleichung

$$y' = a(x)y + b(x)y^{\alpha}, \quad \alpha \neq 1$$

kann mit Hilfe der Substitution

$$u(x) := (y(x))^{1-\alpha}$$

zu einer inhomogenen linearen Differentialgleichung der Form

$$u' = (1-\alpha)a(x)u + (1-\alpha)b(x)$$

umgeformt werden.

2) Die sogenannte Riccati-Differentialgleichung

$$y' = a(x)y + b(x)y^2 + c(x)$$

ist nicht in jedem Fall explizit lösbar. Kennt man aber eine spezielle Lösung y_p, dann kann man die allgemeine Lösung berechnen. Dazu formt man die Riccati-Differentialgleichung mit Hilfe der Substitution

$$u(x) = \frac{1}{y(x) - y_p(x)}$$

zu einer inhomogenen linearen Differentialgleichung der Form

$$u' = -(a(x) + 2b(x)y_p(x))u - b(x)$$

um.

Definition *Sei $n \in \mathbb{N}$, $I \subseteq \mathbb{R}$ ein offenes Intervall, $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$, $a_n \in \mathbb{K}\backslash\{0\}$ und $a_0, ..., a_{n-1}, b : I \to \mathbb{K}$ stetige Funktionen. Die Differentialgleichung*

$$Ly := a_n \frac{d^n}{dx^n} y + a_{n-1}(x) \frac{d^{n-1}}{dx^{n-1}} y + \cdots + a_1(x) \frac{d}{dx} y + a_0(x) y = b(x)$$

heißt explizite lineare gewöhnliche Differentialgleichung n-ter Ordnung.
Sind $a_0, ..., a_{n-1}$ konstant, dann nennt man die Gleichung eine lineare gewöhnliche Differentialgleichung n-ter Ordnung mit konstanten Koeffizienten.
Ist $b(x) = 0$ für alle $x \in I$, dann nennt man die lineare Differentialgleichung homogen, anderenfalls inhomogen.

Bemerkung
Für jedes reelle oder komplexe Polynom $P(\lambda) = \sum\limits_{k=0}^{n} a_k \lambda^k$ vom Grad $n \geq 1$ und jede stetige Funktion $b : I \to \mathbb{K}$, wobei $I \subseteq \mathbb{R}$ ein offenes Intervall ist, ist

$$P\Big(\frac{d}{dx}\Big) y = b(x)$$

mit

$$P\Big(\frac{d}{dx}\Big) = \sum_{k=0}^{n} a_k \frac{d^k}{dx^k},$$

wobei $\frac{d^0}{dx^0} = 1$ sei, eine lineare gewöhnliche Differentialgleichung n-ter Ordnung mit konstanten Koeffizienten.

Satz 23.3 *Mit den Bezeichnungen der obigen Definition ist $L : C^n(I) \to C^0(I)$, $y \mapsto Ly$ eine $\mathbb{K}$-lineare Abbildung und die Menge aller Lösungen der homogenen linearen Differentialgleichung $Ly = 0$ ein $\mathbb{K}$-Vektorraum.*

Definition *Jede Basis des $\mathbb{K}$-Vektorraums aller Lösungen der homogenen linearen Differentialgleichung $Ly = 0$ heißt ein Fundamentalsystem von $Ly = 0$.*
Ist $\mathbb{K} = \mathbb{R}$, dann spricht man von reellen Fundamentalsystemen, ist $\mathbb{K} = \mathbb{C}$, dann spricht man von komplexen Fundamentalsystemen.

Bemerkung
Jede Lösung von $Ly = 0$ lässt sich auf genau eine Weise als Linearkombination von Elementen eines Fundamentalsystems von $Ly = 0$ darstellen.

Hilfssatz 23.4 *a) Sei $k \in \mathbb{N}$, $\lambda \in \mathbb{C}$ und $I \subseteq \mathbb{R}$ ein offenes Intervall. Dann gilt für jede Funktion $f : I \to \mathbb{C}$, die k-mal differenzierbar auf I ist:*

$$\left(\frac{d}{dx} - \lambda\right)^k \big(f(x)e^{\lambda x}\big) = \frac{d^k f}{dx^k}(x)\, e^{\lambda x}.$$

b) Für jedes reelle oder komplexe Polynom P und jedes $\lambda \in \mathbb{C}$ gilt

$$P\left(\frac{d}{dx}\right) e^{\lambda x} = P(\lambda)e^{\lambda x}.$$

Ist insbesondere λ eine Nullstelle von P, dann ist $y(x) = e^{\lambda x}$ eine Lösung von $P\left(\frac{d}{dx}\right) y = 0$.

c) Sei P ein reelles oder komplexes Polynom und $\lambda \in \mathbb{C}$ mit $P(\lambda) \neq 0$. Für jedes reelle oder komplexe Polynom g vom Grad k gilt

$$\forall x \in \mathbb{R}: \; P\Big(\frac{d}{dx}\Big)\big(g(x)e^{\lambda x}\big) = h(x)e^{\lambda x},$$

wobei h ebenfalls ein reelles oder komplexes Polynom vom Grad k ist.

Beweisskizze
Man zeigt die Behauptungen durch direktes Nachrechnen und Induktion nach k.

Satz 23.5 *Sei $P(\lambda) = \sum\limits_{k=0}^{n} a_k \lambda^k$ ein reelles oder komplexes Polynom vom Grad $n \geq 1$ und sei $\{\lambda_j : 1 \leq j \leq r\}$ die Menge der Nullstellen von P und k_j die Vielfachheit der Nullstellen λ_j, d.h., es sei $P(\lambda) = a_n \prod\limits_{j=1}^{r} (\lambda - \lambda_j)^{k_j}$. Dann gilt:*

a) Die Funktionen

$$y_{jm}(x) := x^m e^{\lambda_j x}, \qquad 1 \leq j \leq r,\; 0 \leq m \leq k_j - 1,$$

bilden ein Fundamentalsystem von der Differentialgleichung $P\left(\frac{d}{dx}\right) y = 0$.

b) Ist P ein reelles Polynom. Ersetzt man für jedes Paar zueinander konjugierter nicht reeller Nullstellen ($\mu = \alpha + i\beta$, $\overline{\mu} = \alpha - i\beta$) von P mit der zugehörigen Vielfachheit k_μ von μ (und von $\overline{\mu}$) die Funktionen

$$x^m e^{\mu x}, \quad x^m e^{\overline{\mu} x}, \qquad\qquad 0 \leq m \leq k_\mu,$$

aus dem Fundamentalsystem aus a) durch die Funktionen

$$x^m e^{\alpha x} \cos(\beta x), \quad x^m e^{\alpha x} \sin(\beta x), \qquad\qquad 0 \leq m \leq k_\mu,$$

dann erhält man ein reelles Fundamentalsystem von $P\left(\frac{d}{dx}\right) y = 0$.

c) Sei $x_0 \in \mathbb{R}$ und seien $y_0, \cdots, y_{n-1} \in \mathbb{C}$. Dann besitzt das Anfangswertproblem

$$\begin{cases} P\Big(\dfrac{d}{dx}\Big)y = 0, \\ \dfrac{d^k}{dx^k}\, y(x_0) = y_k, \qquad k = 0, \dots n-1, \end{cases}$$

genau eine Lösung $y : \mathbb{R} \to \mathbb{C}$. Ist P ein reelles Polynom und sind $y_0, \cdots y_{n-1} \in \mathbb{R}$, dann ist y reellwertig.

Beweisskizze
Man zeigt die Behauptungen durch Nachrechnen unter Ausnutzung von Hilfssatz 23.4 und Induktion nach r.

Satz 23.6 *Sei $P(\lambda) = \sum\limits_{k=0}^{n} a_k \lambda^k$ ein reelles oder komplexes Polynom vom Grad $n \geq 1$. Dann gilt*

a) Ist $\mu \in \mathbb{C}$ und $P(\mu) \neq 0$, dann ist

$$y_p(x) := \frac{1}{P(\mu)}\, e^{\mu x}$$

eine spezielle Lösung der inhomogenen Differentialgleichung

$$P\Big(\frac{d}{dx}\Big)y = e^{\mu x}.$$

b) Ist $\mu \in \mathbb{C}$ mit $P(\mu) \neq 0$ und f ein reelles oder komplexes Polynom vom Grad m, dann besitzt die Differentialgleichung

$$P\Big(\frac{d}{dx}\Big)y = f(x)e^{\mu x}$$

eine spezielle Lösung der Form

$$y_p(x) = g(x)e^{\mu x},$$

wobei g ebenfalls ein reelles oder komplexes Polynom vom Grad m ist.

c) Ist $\mu \in \mathbb{C}$ eine k-fache Nullstelle von P und f ein reelles oder komplexes Polynom vom Grad $m \geq 0$, dann besitzt die Differentialgleichung

$$P\Big(\frac{d}{dx}\Big)y = f(x)e^{\mu x}$$

eine spezielle Lösung der Form

$$y_p(x) = x^k h(x)e^{\mu x},$$

wobei h ein reelles oder komplexes Polynom vom Grad m ist.

Beweisskizze

a) Beweis durch Nachrechnen unter Ausnutzung von 23.4 *b)*.

b) Beweis durch Induktion nach m unter Ausnutzung von 23.4 *c)*.

c) Beweis durch Nachrechnen unter Ausnutzung von 23.6 *b)* und 23.4 *a)*.

Wir führen als Beispiel die Details des Beweises von *c)* vor:
Ist $\mu \in \mathbb{C}$ eine k-fache Nullstelle von P, dann gilt

$$P(\lambda) = Q(\lambda)(\lambda - \mu)^k,$$

wobei $Q(\mu) \neq 0$ ist. Nach 23.6 *b)* existiert ein reelles oder komplexes Polynom g vom Grad m, so dass

$$Q\Big(\frac{d}{dx}\Big)(g(x)e^{\mu x}) = f(x)e^{\mu x}$$

gilt. Es gibt ein Polynom H mit $H(x) = x^k h(x)$, wobei h ein reelles oder komplexes Polynom vom Grad m ist und $\frac{d^k}{dx^k}H = g$. Nach 23.4 *a)* gilt dann

$$\Big(\frac{d}{dx} - \mu\Big)^k (H(x)e^{\mu x}) = g(x)e^{\mu x}.$$

Daher erhalten wir

$$\begin{aligned} P\Big(\frac{d}{dx}\Big)(x^k h(x)e^{\mu x}) &= Q\Big(\frac{d}{dx}\Big)\Big(\Big(\frac{d}{dx} - \mu\Big)^k (H(x)e^{\mu x})\Big) \\ &= Q\Big(\frac{d}{dx}\Big)(g(x)e^{\mu x}) \\ &= f(x)e^{\mu x}. \end{aligned}$$

Satz 23.7 *a) Ist y_{hom} die allgemeine Lösung der homogenen linearen gewöhnlichen Differentialgleichung $Ly = 0$ und y_p eine spezielle Lösung der inhomogenen Gleichung $Ly = b$, dann ist $y_{inh} = y_{hom} + y_p$ die allgemeine Lösung von $Ly = b$.*

b) Sind y_{p_i}, $i = 1, 2$, jeweils spezielle Lösungen von $Ly = b_i$ und $\alpha, \beta \in \mathbb{C}$, dann ist $\alpha y_{p_1} + \beta y_{p_2}$ eine spezielle Lösung von $Ly = \alpha b_1 + \beta b_2$.

Beispiele

1) $\begin{cases} y'' - y = x^2, \quad x \in \mathbb{R} \\ y(0) = 1, \\ y'(0) = 0. \end{cases}$

Einsetzen des Ansatzes $y(x) = e^{\lambda x}$ in die zugehörige homogene Gleichung $y'' - y = 0$ liefert das sogenannte charakteristische Polynom

$$\begin{aligned} & \lambda^2 - 1 = 0 \\ \Leftrightarrow \quad & \lambda_{1/2} = \pm 1 \ \wedge \ \lambda_{1/2} \text{ sind einfache Nullstellen} \end{aligned}$$

Aufgrund von Satz 23.5 ist daher

$$y_{hom}(x) = \alpha e^x + \beta e^{-x}, \qquad \alpha, \beta \in \mathbb{R} \text{ oder } \mathbb{C}$$

die allgemeine Lösung der homogenen Gleichung.
Es ist $x^2 = x^2 e^{0x}$ und 0 ist keine Nullstelle von $P(\lambda) = \lambda^2 - 1$.
Daher besitzt nach Satz 23.6 *b*) die inhomogene Gleichung $y'' - y = x^2$ eine spezielle Lösung der Form

$$y_p(x) = a_2 x^2 + a_1 x + a_0.$$

Einsetzen von y_p in die inhomogene Gleichung liefert

$$2a_2 - a_2 x^2 - a_1 x - a_0 = x^2$$

und durch Koeffizientenvergleich folgt

$$a_2 = -1, \ a_1 = 0, \ a_0 = 2a_2 = -2.$$

Also ist

$$y_p(x) = -x^2 - 2$$

eine spezielle Lösung der inhomogenen Gleichung und

$$\begin{aligned} y_{inh}(x) &= y_{hom}(x) + y_p(x) \\ &= \alpha e^x + \beta e^{-x} - x^2 - 2, \qquad \alpha, \beta \in \mathbb{R} \text{ oder } \mathbb{C} \end{aligned}$$

die allgemeine Lösung der inhomogenen Gleichung. Für die Lösung y des obigen Anfangswertproblems folgt:

$$\begin{aligned} y(x) &= \alpha_0 e^x + \beta_0 e^{-x} - x^2 - 2 \\ y'(x) &= \alpha_0 e^x - \beta_0 e^{-x} - 2x \end{aligned}$$

mit noch zu bestimmenden α_0, β_0. Aufgrund der Anfangsbedingungen folgt:

$$\begin{cases} y(0) = \alpha_0 + \beta_0 - 2 = 1 \\ y'(0) = \alpha_0 - \beta_0 = 0 \end{cases} \quad \Leftrightarrow \quad \alpha_0 = \beta_0 = \frac{3}{2}$$

Also ist

$$y(x) = \frac{3}{2}e^x + \frac{3}{2}e^{-x} - x^2 - 2$$

die Lösung des obigen Anfangswertproblems.

2) Der harmonische Oszillator erfüllt die gewöhnliche Differentialgleichung

$$my'' + ry' + ky = 0,$$

wobei $k, m > 0$ und $r \geq 0$ ist.

Der Ansatz: $y(t) = e^{\lambda t}$ liefert

$$m\lambda^2 + r\lambda + k = 0 \;\Leftrightarrow\; \lambda_{1/2} = \frac{1}{2m}(-r \pm \sqrt{r^2 - 4mk})$$

1. Fall: $D := r^2 - 4mk < 0$

$$\lambda_{1/2} = -\frac{r}{2m} \pm i\frac{\sqrt{-D}}{2m}$$

$$\begin{aligned} y(t) &= c_1 e^{-\frac{r}{2m}} \cos\left(\frac{\sqrt{-D}}{2m}\,t\right) + c_2 e^{-\frac{r}{2m}} \sin\left(\frac{\sqrt{-D}}{2m}\,t\right), \qquad c_1, c_2 \in \mathbb{R} \\ &= c e^{-\frac{r}{2m}} \cos(\omega t + \varphi), \qquad c \in \mathbb{R},\ \varphi \in [0, 2\pi[\,,\ \omega = \sqrt{\frac{k}{m} - \frac{r^2}{4m^2}} \end{aligned}$$

2. Fall: $D > 0$

$$\lambda_{1/2} = -\frac{r}{2m} \pm \frac{\sqrt{D}}{2m} < 0 \quad \text{(wegen } k, m > 0\text{)}$$

$$y(t) = c_1 e^{\lambda_1 t} + c_2 e^{\lambda_2 t}, \qquad c_1, c_2 \in \mathbb{R}$$

Beachte: Für $r = 0$ ist dieser Fall nicht möglich.

3. Fall: $D = 0$

$$\lambda_1 = \lambda_2 = -\frac{r}{2m} < 0,$$

$$y(t) = (c_1 + c_2 t)e^{-\frac{r}{2m}t}, \qquad c_1, c_2 \in \mathbb{R}$$

Beachte: Für $r = 0$ ist dieser Fall nicht möglich.

3) Die Stromstärke I in einem Wechselstromkreis mit einer Spule der Induktivität $L > 0$, einem Kondensator der Kapazität $C > 0$ und einem Ohmschen Widerstand $R \geq 0$ in Reihenschaltung und mit einer angelegten Wechselspannung $U(t) = U_0 \cos(\omega t)$ erfüllt die inhomogene Differentialgleichung

$$LI'' + RI' + C^{-1}I = U'.$$

Allgemeine Lösung $I_{hom}(t)$ der zugehörigen homogenen Gleichung: siehe 2), wobei $I_{hom}(t) = y(t)$, $m = L$, $r = R$ und $k = C^{-1}$ ist.

Komplexifizierung: $U^*(t) := U_0 e^{i\omega t} = U_0(\cos(\omega t) + i\sin(\omega t))$
$\Rightarrow (U^*)'(t) = i\omega U_0 e^{i\omega t}$

Ansatz für eine partikuläre Lösung der inhomogenen Gleichung:
Betrachte das charakteristische Polynom $L\lambda^2 + R\lambda + C^{-1}$.

1. Fall: $i\omega$ ist keine Nullstelle des charakteristischen Polynoms
Ansatz: $I_p^*(t) = I_0^* e^{i\omega t}, \quad I_0^* \in \mathbb{C}$
Einsetzen in die Differentialgleichung liefert
$(-\omega^2 L + i\omega R + C^{-1})I_0^* e^{i\omega t} = i\omega U_0 e^{i\omega t}$

$$\begin{aligned}
\Leftrightarrow \quad I_0^* &= \frac{i\omega U_0}{-\omega^2 L + i\omega R + C^{-1}} = \frac{i\omega U_0}{i\omega(R - \frac{1}{i}\omega L + \frac{1}{i}\frac{1}{\omega C})} = \frac{U_0}{R + i(\omega L - \frac{1}{\omega C})} \\
&= \frac{U_0(R - i(\omega L - \frac{1}{\omega C}))}{R^2 + (\omega L - \frac{1}{\omega C})^2} \\
&= \frac{U_0}{\sqrt{R^2 + (\omega L - \frac{1}{\omega C})^2}}\, e^{-i\arctan\left(\frac{\omega L - (\omega C)^{-1}}{R}\right)}
\end{aligned}$$

Eine reelle partikuläre Lösung der inhomogenen Gleichung ist damit

$$I_p(t) = Re(I_0^* e^{i\omega t}) = \frac{U_0}{\sqrt{R^2 + (\omega L - \frac{1}{\omega C})^2}} \cos(\omega t - \alpha)$$

mit $\alpha = \arctan\left(\frac{\omega L - \frac{1}{\omega C}}{R}\right)$.

Die allgemeine reelle Lösung der inhomogenen Gleichung ist dann

$$I_{inh}(t) = I_{hom}(t) + I_p(t).$$

Unterfall: $R > 0$
In diesem Fall haben die Nullstellen des charakteristischen Polynoms negativen Realteil. Daher sind bei jeder beliebigen speziellen Lösung $I(t)$ der inhomogenen Gleichung für hinreichend große t eventuelle Anteile der allgemeinen homogenen Lösung $I_{hom}(t)$ so weit abgeklungen, dass sie praktisch nicht mehr zu sehen sind. Folglich ist $I(t) \approx I_p(t)$ für hinreichend große t, wobei $I_p(t)$ wie oben ist.

Unterfall: $R = 0$ (kein Widerstand)
Sei $\omega_0 := \frac{1}{\sqrt{LC}}$. Dann gilt:

$$I_{inh}(t) = c_1 \cos(\omega_0 t) + c_2 \sin(\omega_0 t) + \frac{U_0}{\omega L - \frac{1}{\omega C}} \sin(\omega t), \qquad c_1, c_2 \in \mathbb{R}.$$

Eine in physikalischer Hinsicht wichtige spezielle Lösung der inhomogenen Gleichung ist in diesem Fall

$$I(t) = \frac{U_0}{\omega L - \frac{1}{\omega C}} (\sin(\omega t) - \sin(\omega_0 t)) = \left(\frac{2U_0}{\omega L - \frac{1}{\omega C}} \sin\left(\frac{\omega - \omega_0}{2} t\right)\right) \cos\left(\frac{\omega + \omega_0}{2} t\right).$$

Diese Lösung beschreibt eine amplitudenmodulierte Schwingung. Ist dabei $\frac{\omega}{\omega_0} \in \mathbb{Q}$, dann handelt es sich um eine Schwebung.
Eine physikalische Charakterisierung aller spezieller Lösungen der inhomogenen Gleichung im Unterfall $R = 0$ ist ziemlich aufwendig.

2. Fall: $i\omega$ ist Nullstelle des charakteristischen Polynoms
$\overset{2)}{\Rightarrow} R = 0 \overset{2)}{\Rightarrow} \omega = \omega_0 = \frac{1}{\sqrt{LC}}$ und $i\omega$ ist einfache Nullstelle
Ansatz: $I_p^*(t) = I_1^* \, t \, e^{i\omega_0 t}$
$\Rightarrow \ (I_p^*)'(t) = (I_1^* + i\omega_0 I_1^* t) e^{i\omega_0 t}, \ (I_p^*)''(t) = (2i\omega_0 I_1^* - \omega_0^2 I_1^* t) e^{i\omega_0 t}$
Einsetzen in die Differentialgleichung liefert
$(2i\omega_0 L - \underbrace{\omega_0^2 L}_{=C^{-1}} t + C^{-1} t) I_1^* e^{i\omega_0 t} = i\omega_0 U_0 e^{i\omega_0 t}$
$\Leftrightarrow \ I_1^* = \frac{U_0}{2L}$
Die allgemeine reelle Lösung der inhomogenen Gleichung ist dann

$$I_{inh}(t) = c_1 \cos(\omega_0 t) + c_2 \sin(\omega_0 t) + \frac{U_0}{2L} t \cos(\omega_0 t), \quad c_1, c_2 \in \mathbb{R}.$$

In diesem Fall beschreibt die allgemeine reelle Lösung das Auftreten von Resonanz.

Bemerkung
Für eine lineare gewöhnliche Differentialgleichung n-ter Ordnung, bei der nicht alle Koeffizienten konstant sind, gibt es, wenn $n \geq 2$ ist, im Allgemeinen keine explizite Lösungsformel.
Ist allerdings bei einer expliziten homogenen linearen gewöhnlichen Differentialgleichung zweiter Ordnung

$$y'' + a(x)y' + b(x)y = 0$$

eine Lösung y_1 bekannt (zum Beispiel durch Raten oder mit Hilfe eines Potenzreihenansatzes gefunden), dann erhält man auf jedem offenen Intervall, auf dem $y_1(x) \neq 0$ gilt, eine zweite, von y_1 linear unabhängige Lösung y_2 und somit ein Fundamentalsystem. y_2 ist von der Form

$$y_2(x) = u(x)y_1(x),$$

wobei u eine nicht-konstante Lösung der Gleichung

$$(u')' + \Big(2\frac{y_1'(x)}{y_1(x)} + a(x)\Big)u' = 0$$

ist, welche mit Hilfe von Satz 23.1 bestimmt werden kann.
Entsprechend erhält man bei Kenntnis einer Lösung y_1 einer expliziten homogenen linearen gewöhnlichen Differentialgleichung n-ter Ordnung mit $n > 2$ auf jedem offenen Intervall, auf dem $y_1(x) \neq 0$ ist, mittels des Ansatzes $y_2(x) = u(x)y_1(x)$ eine explizite homogene lineare gewöhnliche Differentialgleichung $(n-1)$-ter Ordnung für u'. Für jede Lösung $u' \neq 0$ ist dann $y_2 = uy_1$ eine von y_1 linear unabhängige Lösung der Gleichung n-ter Ordnung.

Definition *Sei $I \subseteq \mathbb{R}$ ein offenes Intervall, $n \in \mathbb{N}$, $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$ und* $A : I \to \mathbb{K}^{n\times n}$, $x \mapsto A(x) = (a_{ij}(x))_{i,j=1,\dots,n}$ *sowie* $b : I \to \mathbb{K}^n$, $x \mapsto b(x) = \begin{pmatrix} b_1(x) \\ \cdot \\ \cdot \\ b_n(x) \end{pmatrix}$ *stetig (d.h., alle Funktionen a_{ij}, b_j seien stetig). Dann heißt*

$$y' = A(x)y + b(x)$$

ein System expliziter linearer gewöhnlicher Differentialgleichungen erster Ordnung.
Sind alle a_{ij} konstant, dann nennt man das System ein System mit konstanten Koeffizienten.
Sind alle $b_j = 0$ auf I, dann nennt man das System homogen, anderenfalls inhomogen.

Bemerkung
Satz 23.3 und die Definition eines Fundamentalsystems übertragen sich sinngemäß auf den Fall von linearen Systemen.

Satz 23.8 *a) Sei* $A = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{pmatrix}$ *mit $\lambda_1, ..., \lambda_n \in \mathbb{K}$. Dann bilden die vektorwertigen Funktionen*

$$\begin{pmatrix} e^{\lambda_1 x} \\ 0 \\ \vdots \\ \vdots \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ e^{\lambda_2 x} \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ \vdots \\ \vdots \\ 0 \\ e^{\lambda_n x} \end{pmatrix}$$

ein Fundamentalsystem von

$$y' = Ay.$$

b) *Sei $A \in \mathbb{K}^{n\times n}$ und $S \in GL(n, \mathbb{K})$. Dann ist $y : \mathbb{R} \to \mathbb{K}^n$ genau dann eine Lösung von*

$$y' = Ay$$

wenn die Funktion $S^{-1}y : \mathbb{R} \to \mathbb{K}^n$ eine Lösung von

$$\tilde{y}' = (S^{-1}AS)\tilde{y}$$

ist.

Beweis

$a)$ durch direktes Nachrechnen.

$b)$ Es gilt

$$y' = Ay \;\Leftrightarrow\; (S^{-1}y)' = S^{-1}y' = S^{-1}Ay = (S^{-1}AS)S^{-1}y\,. \qquad \square$$

Korollar 23.9 *a) Existiert eine Basis $a_1, ..., a_n$ von $\mathbb{K}^n$ aus Eigenvektoren der Matrix $A \in \mathbb{K}^{n\times n}$ zu den jeweiligen Eigenwerten $\lambda_i, i = 1, ..., n$, dann bilden die (vektorwertigen) Funktionen*

$$y_i : \mathbb{R} \to \mathbb{K}^n, \quad y_i(x) = e^{\lambda_i x} a_i, \qquad i = 1, ..., n,$$

ein Fundamentalsystem von

$$y' = Ay$$

und für alle $x_0 \in \mathbb{R}$ und $y_0 \in \mathbb{K}^n$ besitzt das Anfangswertproblem

$$\begin{cases} \quad y' = Ay\,, \\ y(x_0) = y_0\,, \end{cases}$$

genau eine Lösung $y : \mathbb{R} \to \mathbb{K}^n$.

b) *Ist die Matrix A aus a) reellwertig und ersetzt man für jedes Paar $\mu = \alpha + i\beta$, $\overline{\mu} = \alpha - i\beta$) nicht reeller Eigenwerte von A mit der zugehörigen Vielfachheit k_μ von μ (und von $\overline{\mu}$) die Funktionen*

$$e^{\mu x} a_{\mu,i}, \quad e^{\overline{\mu} x} a_{\overline{\mu},i}, \qquad i = 1, ..., k_\mu,$$

aus dem Fundamentalsystem aus a) durch die Funktionen

$$Re(e^{\mu x}a_{\mu,i}), \quad Im(e^{\mu x}a_{\mu,i}), \qquad\qquad i = 1, ..., k_\mu,$$

dann erhält man ein reelles Fundamentalsystem von

$$y' = Ay$$

und für alle $x_0 \in \mathbb{R}$ und $y_0 \in \mathbb{R}^n$ ist die Lösung des Anfangswertproblems

$$\begin{cases} \quad y' = Ay\,, \\ y(x_0) = y_0\,, \end{cases}$$

reellwertig.

Beweis
durch Nachrechnen unter Ausnutzung von 23.8. □

Beispiele

1) $$\begin{cases} y' = \begin{pmatrix} 8 & -2 \\ -2 & 5 \end{pmatrix} y\,, \\ \\ y(0) = \begin{pmatrix} 3 \\ 1 \end{pmatrix}. \end{cases}$$

Mit den bekannten Methoden aus der linearen Algebra erhält man, dass

$\begin{pmatrix} 2 \\ -1 \end{pmatrix}$ ein EV von $\begin{pmatrix} 8 & -2 \\ -2 & 5 \end{pmatrix}$ zum EW 9,

$\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ ein EV von $\begin{pmatrix} 8 & -2 \\ -2 & 5 \end{pmatrix}$ zum EW 4

und $\left\{ \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix} \right\}$ eine Basis von $\mathbb{R}^2$ ist.

Also bilden die Funktionen

$y_1(x) = e^{9x} \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \quad y_2(x) = e^{4x} \begin{pmatrix} 1 \\ 2 \end{pmatrix}$

ein Fundamentalsystem der obigen Differentialgleichung.
Sei y die Lösung des obigen Anfangswertproblems. Dann gilt:

$y(x) = \alpha y_1(x) + \beta y_2(x) \ \wedge \ y(0) = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$

$\Rightarrow \ \alpha \begin{pmatrix} 2 \\ -1 \end{pmatrix} + \beta \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \end{pmatrix} \ \Leftrightarrow \ \alpha = \beta = 1$

Also gilt: $y(x) = e^{9x} \begin{pmatrix} 2 \\ -1 \end{pmatrix} + e^{4x} \begin{pmatrix} 1 \\ 2 \end{pmatrix}$.

2) $y' = \begin{pmatrix} 0 & -\omega \\ \omega & 0 \end{pmatrix} y$ mit $\omega \in \mathbb{R} \setminus \{0\}$

Die lineare Algebra liefert

$\begin{pmatrix} 1 \\ -i \end{pmatrix}$ ist ein EV von $\begin{pmatrix} 0 & -\omega \\ \omega & 0 \end{pmatrix}$ zum EW $i\omega$,

$\begin{pmatrix} 1 \\ i \end{pmatrix}$ ist ein EV von $\begin{pmatrix} 0 & -\omega \\ \omega & 0 \end{pmatrix}$ zum EW $-i\omega$

und $\left\{ \begin{pmatrix} 1 \\ -i \end{pmatrix}, \begin{pmatrix} 1 \\ i \end{pmatrix} \right\}$ ist eine Basis von $\mathbb{C}^2$

Also ist

$\left\{ e^{i\omega x} \begin{pmatrix} 1 \\ -i \end{pmatrix}, e^{-i\omega x} \begin{pmatrix} 1 \\ i \end{pmatrix} \right\}$ ein komplexes Fundamentalsystem und

$\left\{ \begin{pmatrix} \cos(\omega x) \\ \sin(\omega x) \end{pmatrix}, \begin{pmatrix} -\sin(\omega x) \\ \cos(\omega x) \end{pmatrix} \right\}$ ein reelles Fundamentalsystem

der obigen Gleichung.

Satz 23.10 *Sei E_n die Einheitsmatrix in $\mathbb{K}^{n \times n}$ und $N_n \in \mathbb{K}^{n \times n}$ mit*

$$N_n = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ & & & \ddots & 1 \\ 0 & \dots & & 0 & 0 \end{pmatrix} \text{ und } \lambda \in \mathbb{K}.$$

Dann bilden

$$e^{\lambda x} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ \vdots \\ \vdots \\ 0 \end{pmatrix}, \; e^{\lambda x} \begin{pmatrix} x \\ 1 \\ 0 \\ \vdots \\ \vdots \\ 0 \end{pmatrix}, \; e^{\lambda x} \begin{pmatrix} \frac{1}{2}x^2 \\ x \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \; \dots, \; e^{\lambda x} \begin{pmatrix} \frac{1}{(n-1)!}\, x^{n-1} \\ \frac{1}{(n-2)!}\, x^{n-2} \\ \vdots \\ \vdots \\ x \\ 1 \end{pmatrix}$$

ein Fundamentalsystem von

$$y' = (\lambda E_n + N_n) y$$

und für alle $x_0 \in \mathbb{R}$ und $y_0 \in \mathbb{K}^n$ besitzt das Anfangswertproblem

$$\begin{cases} y' = (\lambda E_n + N_n) y\,, \\ y(x_0) = y_0\,, \end{cases}$$

genau eine Lösung $y : \mathbb{R} \to \mathbb{K}^n$.

Beweisskizze
$y(x) = e^{\lambda x} w(x)$ ist genau dann Lösung von $y' = (\lambda E_n + N_n)y$ wenn w Lösung von $w' = N_n w$ ist und für jede Lösung $w = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}$ gilt

$$\begin{aligned} w_n(x) &= c_n \\ w_{n-1}(x) &= c_{n-1} + c_n x \\ &\vdots \\ w_1(x) &= c_1 + c_2 x + \frac{1}{2} c_3 x^2 + \ldots + \frac{1}{(n-1)!} c_n x^{n-1} \end{aligned}$$

mit $c_1, ..., c_n \in \mathbb{K}$, was direkt nachgerechnet werden kann.

Bemerkung
Da $\lambda E_n + N_n$ ein Jordan-Block ist und nach Satz 12.22 jede Matrix $A \in \mathbb{K}^{m \times m}$ ähnlich zu einer Matrix ist, die aus Jordan-Blöcken aufgebaut ist, kann man mit Hilfe von Satz 23.10 und Satz 23.8 a), b) ein Fundamentalsystem von

$$y' = Ay$$

konstruieren und für das Anfangswertproblem

$$\begin{cases} y' = Ay\,, \\ y(x_0) = y_0\,, \end{cases}$$

mit $x_0 \in \mathbb{R}$, $y_0 \in \mathbb{K}^m$ die eindeutige Lösung bestimmen.
Ist A reell, dann kann man auf analoge Weise wie in Korollar 23.9 aus einem komplexen Fundamentalsystem ein reelles konstruieren.
Ist A reell und $y_0 \in \mathbb{R}^m$, dann ist die Lösung des obigen Anfangswertproblems ebenfalls reell.

Beispiel
$y' = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{pmatrix} y$

Dann ist nach Satz 23.10 und Satz 23.8 a)

$\left\{ e^x \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, e^{3x} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, e^{3x} \begin{pmatrix} 0 \\ x \\ 1 \end{pmatrix} \right\}$

ein Fundamentalsystem dieser Gleichung.

Am Ende von Kapitel 12 haben wir berechnet, dass

$$S^{-1}\begin{pmatrix}1&0&0\\0&4&1\\0&-1&2\end{pmatrix}S = \begin{pmatrix}1&0&0\\0&3&1\\0&0&3\end{pmatrix}$$

mit $S = \begin{pmatrix}1&0&0\\0&1&\frac{1}{2}\\0&-1&\frac{1}{2}\end{pmatrix}$ gilt.

Schreibt man das obige Fundamentalsystem als sogenannte Fundamentalmatrix $\Phi = \Phi(x)$, d.h.

$$\Phi(x) = \begin{pmatrix}e^x&0&0\\0&e^{3x}&xe^{3x}\\0&0&e^{3x}\end{pmatrix},$$

dann ist nach Satz 23.8 b) die Matrix

$$S\Phi(x)$$

eine Fundamentalmatrix von

$$y' = \begin{pmatrix}1&0&0\\0&4&1\\0&-1&2\end{pmatrix}y.$$

Bemerkung
Alternativ zu der oben vorgestellten Vorgehensweise zur Lösung von Differentialgleichungssystemen der Form $y' = Ay$ mit $A \in \mathbb{K}^{n\times n}$ kann man Folgendes machen. Man führt die sogenannte Matrixexponentialfunktion

$$e^A := \sum_{k=0}^{\infty} \frac{1}{k!}A^k$$

ein, wobei $A^0 = E_n$ ist und folgender Konvergenzbegriff für Matrizen zugrundegelegt wird:

$$\underbrace{M_k}_{=(m_{kij})_{i,j=1,\dots,n}} \xrightarrow{k\to\infty} \underbrace{M}_{=(m_{ij})_{i,j=1,\dots,n}} \quad :\Leftrightarrow \quad \forall i,j = 1,...,n:\ m_{kij} \xrightarrow{k\to\infty} m_{ij}$$

Dann kann man durch Verallgemeinerung der Argumentationen aus den entsprechenden Beweisen der reellen und komplexen Exponentialfunktion zeigen, dass e^A für alle Matrizen $A \in \mathbb{K}^{n\times n}$ konvergiert und

$$y(x) = e^{(x-x_0)A}y_0\,, \quad x \in \mathbb{R}$$

die eindeutige Lösung von

$$\begin{cases} y' = Ay\,, \\ y(x_0) = y_0\,, \end{cases}$$

ist. Aus diesem Resultat kann man außerdem sämtliche Aussagen aus 23.8 − 23.10 herleiten.

Der folgende Satz zeigt den Zusammenhang zwischen den Lösungen von beliebigen inhomogenen linearen gewöhnlichen Differentialgleichungen und den zugehörigen homogenen Gleichungen.

Satz 23.11 *Sei $I \subseteq \mathbb{R}$ ein offenes Intervall und seien $A : I \to \mathbb{K}^{n\times n}, b : I \to \mathbb{K}^n$ stetige Abbildungen. Wir bezeichnen mit L_H den Vektorraum aller Lösungen $\varphi : I \to \mathbb{K}$ des homogenen Differentialgleichungssystems $y' = A(x)y$ und mit L_I die Menge aller Lösungen $\psi : I \to \mathbb{K}^n$ des inhomogenen Differentialgleichungssystems $y' = A(x)y + b(x)$. Dann gilt für beliebiges $\psi_0 \in L_I$:*

$$L_I = \psi_0 + L_H.$$

Beweis

a) Wir zeigen zunächst $L_I \subseteq \psi_0 + L_H$. Sei $\psi \in L_I$ beliebig vorgegeben. Wir setzen $\varphi(x) := \psi(x) - \psi_0(x)$. Dann gilt

$$\begin{aligned} \varphi'(x) &= \psi'(x) - \psi_0'(x) \;=\; (A(x)\psi(x) + b(x)) - (A(x)\psi_0(x) + b(x)) \\ &= A(x)(\psi(x) - \psi_0(x)) \;=\; A(x)\varphi(x), \end{aligned}$$

d.h. $\varphi \in L_H$. Da $\psi(x) = \psi_0(x) + \varphi(x)$ ist, folgt $\psi \in \psi_0 + L_H$.

b) Wir zeigen jetzt $\psi_0 + L_H \subseteq L_I$. Sei $\psi \in \psi_0 + L_H$, d.h. $\psi(x) = \psi_0(x) + \varphi(x)$ mit $\varphi \in L_H$. Dann gilt

$$\begin{aligned} \psi'(x) &= \psi_0'(x) + \varphi'(x) \;=\; (A(x)\psi_0(x) + b(x)) + A(x)\varphi(x) \\ &= A(x)(\psi_0(x) + \varphi(x)) + b(x) \;=\; A(x)\psi(x) + b(x), \end{aligned}$$

d.h. $\psi \in L_I$. □

Um die allgemeine Lösung von inhomogenen linearen gewöhnlichen Differentialgleichungen zu erhalten ist also neben der allgemeinen Lösung der zugehörigen homogenen Gleichungen nur die Kenntnis einer einzigen speziellen Lösung der jeweiligen inhomogenen Gleichung notwendig. Eine solche kann man durch die sogenannte Methode der Variation der Konstanten bekommen.

Satz 23.12 (Variation der Konstanten) *Seien A und b wie in Satz 23.11. Sei $\{\varphi_1(x), \ldots, \varphi_n(x)\}$ ein Fundamentalsystem des homogenen Differentialgleichungssystems $y' = A(x)y$ und*

$$\Phi(x) := (\varphi_1(x), \ldots, \varphi_n(x))$$

eine Fundamentalmatrix. Dann erhält man eine Lösung $\psi : I \to \mathbb{K}^n$ des inhomogenen Differentialgleichungssystems $y' = A(x)y + b(x)$ durch den Ansatz

$$\psi(x) = \Phi(x)u(x),$$

wobei $u : I \to \mathbb{K}^n$ eine differenzierbare Funktion mit $\Phi(x)u'(x) = b(x)$ ist, d.h., es gilt

$$u(x) = \int_{x_0}^{x} (\Phi(t))^{-1} b(t)\, dt + C, \quad C \in \mathbb{K}^n.$$

Hierbei wird die vektorwertige Funktion $\Phi^{-1}b$ komponentenweise integriert.

Beweis
Zunächst gilt

$$\Psi'(x) = (\varphi_1'(x), \ldots, \varphi_n'(x)) = (A(x)\varphi_1(x), \ldots, A(x)\varphi_n(x)) = A(x)\Phi(x).$$

Aus $\psi(x) = \Phi(x)u(x)$ folgt

$$\begin{aligned} \psi'(x) &= \Phi'(x)u(x) + \Phi(x)u'(x) \quad = \quad A(x)\Phi(x)u(x) + \Phi(x)u'(x) \\ &= A(x)\psi(x) + \Phi(x)u'(x). \end{aligned}$$

Also ist ψ eine Lösung des inhomogenen Differentialgleichungssystems $y' = A(x)y + b(x)$ genau dann, wenn $\Phi(x)u'(x) = b(x)$ gilt. $\square$

Beispiel
Wir behandeln das Differentialgleichungssystem

$$\begin{cases} y_1' &= -y_2 \\ y_2 &= y_1 + x, \end{cases}$$

oder in Matrixschreibweise

$$y' = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} y + \begin{pmatrix} 0 \\ x \end{pmatrix}.$$

In einem früheren Beispiel haben wir bereits berechnet, dass

$$\left\{ \begin{pmatrix} \cos x \\ \sin x \end{pmatrix}, \begin{pmatrix} -\sin x \\ \cos x \end{pmatrix} \right\}$$

ein Fundamentalsystem des homogenen Gleichungssystems

$$y' = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} y$$

ist. Mit $\Phi(x) = \begin{pmatrix} \cos x & -\sin x \\ \sin x & \cos x \end{pmatrix}$ ist $(\Phi(x))^{-1} = \begin{pmatrix} \cos x & \sin x \\ -\sin x & \cos x \end{pmatrix}$. Also ist

$$(\Phi(x))^{-1} b(x) = \begin{pmatrix} \cos x & \sin x \\ -\sin x & \cos x \end{pmatrix} \begin{pmatrix} 0 \\ x \end{pmatrix} = \begin{pmatrix} x \sin x \\ x \cos x \end{pmatrix}.$$

Damit ergibt sich

$$u(x) = \int_{x_0}^{x} \begin{pmatrix} t \sin t \\ t \cos t \end{pmatrix} dt + C = \begin{pmatrix} -x \cos x + \sin x \\ x \sin x + \cos x \end{pmatrix},$$

wobei wir den konstanten Vektor C geeignet gewählt haben. Eine spezielle Lösung des inhomogenen Differentialgleichungssystems ist somit gegeben durch

$$\psi(x) = \Phi(x) u(x) = \begin{pmatrix} \cos x & -\sin x \\ \sin x & \cos x \end{pmatrix} \begin{pmatrix} -x \cos x + \sin x \\ x \sin x + \cos x \end{pmatrix} = \begin{pmatrix} -x \\ 1 \end{pmatrix}.$$

Dass ψ das inhomogene Differentialgleichungssystem löst, kann man auch sofort direkt verifizieren. Die allgemeine Lösung des inhomogenen Systems ist somit gegeben durch

$$\psi(x) = \begin{pmatrix} -x \\ 1 \end{pmatrix} + \alpha \begin{pmatrix} \cos x \\ \sin x \end{pmatrix} + \beta \begin{pmatrix} -\sin x \\ \cos x \end{pmatrix}, \quad \alpha, \beta \in \mathbb{R} \text{ bzw. } \mathbb{C}.$$

24 Existenz und Eindeutigkeit von Lösungen gewöhnlicher Differentialgleichungen

Satz 24.1 (Existenz- und Eindeutigkeitssatz von Picard-Lindelöf)
Sei $\mathbb{R} \times \mathbb{R}^m \supseteq D \to \mathbb{R},\ (x,y) \mapsto f(x,y)$ eine Funktion mit den folgenden Eigenschaften:

1. *f ist stetig auf $B := [x_0 - a, x_0 + a] \times \{y \in \mathbb{R}^m : |y - y_0| \leq b\}$, wobei $x_0 \in \mathbb{R}$, $a, b \in \mathbb{R}^+$ und $y_0 \in \mathbb{R}^m$ sind.*

2. *f ist auf B Lipschitz-stetig bezüglich y, d. h.*

$$\exists L \geq 0\ \forall (x,y),(x,\tilde{y}) \in B : |f(x,y) - f(x,\tilde{y})| \leq L|y - \tilde{y}|.$$

Dann besitzt das Anfangswertproblem

$$\begin{cases} y' = f(x,y) \\ y(x_0) = y_0 \end{cases}$$

genau eine Lösung $y \in C^1([x_0 - h, x_0 + h], \mathbb{R}^m)$, wobei $h = \min\{a, \frac{b}{M}, \frac{1}{2L}\}$ mit $M = \max\limits_{(x,y)\in B} |f(x,y)|$ ist (für $L = 0$ bzw. $M = 0$ sei $\frac{1}{2L} = \infty$ bzw. $\frac{b}{M} = \infty$), und es gilt

$$y([x_0 - h, x_0 + h]) \subseteq \{y \in \mathbb{R}^m : |y - y_0| \leq b\}.$$

Der Beweis von Satz 24.1 folgt am Ende dieses Kapitels.

Bemerkung
Ist $f : \mathbb{R} \times \mathbb{R}^m \supseteq D \to \mathbb{R},\ (x,y) \mapsto f(x,y)$ eine Funktion, für welche die partiellen Ableitungen $\partial_{y_1} f, \partial_{y_2} f, \ldots, \partial_{y_m} f$, wobei $y = (y_1, \ldots, y_m)$ sei, auf ganz D existieren und dort stetig sind, dann ist f auf jeder Menge $B \subseteq D$ mit $B = [x_0 - a, x_0 + a] \times \{y \in \mathbb{R}^m : |y - y_0| \leq b\}$, $x_0 \in \mathbb{R}$, $a, b \in \mathbb{R}^+$, $y_0 \in \mathbb{R}^m$ Lipschitz-stetig bezüglich y mit $L \leq \sup\limits_{\substack{(x,y)\in B \\ i=1,\ldots,m}} |\partial_{y_i} f(x,y)|$, was aus dem Mittelwertsatz der Differentialrechnung folgt.

Bemerkung
Hat eine Funktion f alle Eigenschaften aus Satz 24.1 außer der Lipschitz-Stetigkeit, dann besitzt das Anfangswertproblem aus Satz 24.1 trotzdem eine Lösung $y \in C^1([x_0 - h, x_0 + h], \mathbb{R}^m)$ mit $h = \min\{a, \frac{b}{M}\}$ und $y([x_0 - h, x_0 + h]) \subseteq \{y \in \mathbb{R}^m : |y - y_0| \leq b\}$. Dies ist der Existenzsatz von Peano. Es kann aber in diesem Fall mehrere Lösungen geben.

Beispiele

1) $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R},\ f(x,y) = |y|^n,\ n > 0$. f ist stetig auf jeder Menge B mit

$$B = [x_0 - a, x_0 + a] \times [y_0 - b, y_0 + b],\ x_0, y_0 \in \mathbb{R},\ a, b \in \mathbb{R}^+,$$

und es ist $M = \max\{|y_0 - b|^n, |y_0 + b|^n\}$. Ist $n \geq 1$, dann ist f auf B auch Lipschitz-stetig bezüglich y mit $L \leq n \max\{|y_0 - b|^{n-1}, |y_0 + b|^{n-1}\}$. Für $0 < n < 1$ gilt das nur, falls $0 \notin [y_0 - b, y_0 + b]$. Also besitzt das Anfangswertproblem

$$\begin{cases} \quad y' = |y|^n \\ y(x_0) = y_0 \end{cases}$$

für $n \geq 1$ genau eine Lösung $y \in C^1([x_0 - h, x_0 + h], \mathbb{R})$ mit $h = \min\{a, \frac{b}{M}, \frac{1}{2L}\}$. Ersetzt man in den Voraussetzungen von Satz 24.1 x_0 durch $x_0 \pm h$ und y_0 durch $y(x_0 \pm h)$ und wendet Satz 24.1 erneut an, dann kann man die Lösung y fortsetzen und erhält genau eine Lösung $y \in C^1([x_0 - h - \tilde{h}, x_0 + h + \tilde{\tilde{h}}], \mathbb{R})$. Durch Wiederholung dieser Vorgehensweise erhält man die eindeutige maximale Lösung $y \in C^1(]r_-, r_+[\,, \mathbb{R})$ des obigen Anfangswertproblems. Dabei gilt entweder $r_- = -\infty$ oder $|y(x)| \to \infty$ für $x \to (r_-)^+$ sowie entweder $r_+ = \infty$ oder $|y(x)| \to \infty$ für $x \to (r_+)^-$.

Für $n = 1$ ist beispielsweise $r_\pm = \pm\infty$, für $n = 2$ und $y_0 > 0$ ist $r_- = -\infty$ und $r_+ = y_0^{-1}$ (siehe Kapitel 22, Beispiele 3) und 5)).

Für $0 < n < 1$ dagegen ist das obige Anfangswertproblem, wie man direkt nachrechnen kann, lösbar, aber im Allgemeinen nicht eindeutig lösbar (vergleiche das ähnliche Beispiel 6) aus Kapitel 22).

2) Das Räuber-Beute-Modell von Lotka-Volterra

$$\begin{cases} \quad y_1' = ay_1 - by_1y_2 \\ \quad y_2' = -cy_2 + dy_1y_2 \\ y_1(t_0) = y_{10} \\ y_2(t_0) = y_{20} \end{cases}$$

mit $a, b, c, d \in \mathbb{R}^+$ und $y_{10}, y_{20} \in \mathbb{R}_0^+$, wobei y_1 die Beute-Population und y_2 die Räuber-Population beschreibt, erfüllt, wie man direkt nachrechnen kann, die Voraussetzungen des Satzes 24.1 und besitzt daher eine eindeutige Lösung für beliebig vorgegebene Anfangsbedingungen $y_{10}, y_{20} \in \mathbb{R}_0^+$.

Korollar 24.2 *Sei $I \subseteq \mathbb{R}$ ein Intervall, $m \in \mathbb{N}$, $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$ und $A : I \to \mathbb{K}^{m \times m}$, $b : I \to \mathbb{K}^m$ stetig. Dann besitzt das lineare Anfangswertproblem*

$$\begin{cases} \quad y' = A(x)y + b(x) \\ y(x_0) = y_0 \end{cases}$$

für alle $x_0 \in I$ und alle $y_0 \in \mathbb{K}^m$ genau eine Lösung $y \in C^1(I, \mathbb{K}^m)$.

Beweisskizze
Für jedes beschränkte und abgeschlossene Intervall $J \subseteq I$ ergibt sich durch Nachrechnen, dass $f(x,y) = A(x)y + b(x)$ unter den gegebenen Voraussetzungen auf $J \times \mathbb{K}^n$ nicht nur stetig, sondern auch Lipschitz-stetig bezüglich y ist. Durch wiederholte Anwendung von Satz 24.1 erhält man dann die Behauptung. .

Bemerkung
Man beachte, dass im Gegensatz zum allgemeinen Fall von Satz 24.1 im linearen Fall von Korollar 24.2 die Lösung y auf ganz I existiert.

Korollar 24.3 *Sei $A : I \to \mathbb{K}^{m\times m}$ wie in Korollar 24.2 und L_H der $\mathbb{K}$-Vektorraum aller Lösungen der homogenen linearen Differentialgleichung*

$$y' = A(x)y.$$

Dann ist $\dim L_H = m$ und für jede m-elementige Menge

$$\left\{ y_1 = \begin{pmatrix} y_{11} \\ \vdots \\ y_{m1} \end{pmatrix}, \dots, y_m = \begin{pmatrix} y_{1m} \\ \vdots \\ y_{mm} \end{pmatrix} \right\} \subseteq L_H$$

und die zugehörige $m \times m$-Matrix $\Phi = (y_{ij})_{i,j=1,\dots,m}$ ist äquivalent:

(i) $y_1, \dots, y_m$ *sind linear unabhängig in* L_H.

(ii) $\exists x_0 \in I : y_1(x_0), \dots, y_m(x_0)$ *sind linear unabhängig in* $\mathbb{K}^m$.

(iii) $\forall x_0 \in I : y_1(x_0), \dots, y_m(x_0)$ *sind linear unabhängig in* $\mathbb{K}^m$.

(iv) $\{y_1, \dots, y_m\}$ *ist ein Fundamentalsystem von* $y' = A(x)y$.

(v) $\exists x_0 \in I : W(x_0) := \det \Phi(x_0) \neq 0$.

(vi) $\forall x_0 \in I : W(x_0) := \det \Phi(x_0) \neq 0$.

(vii) Φ *ist eine Fundamentalmatrix von* $y' = A(x)y$.

Bezeichnung: $W(x_0)$ *heißt Wronski-Determinante.*

Beweisskizze
(i) $\Rightarrow$ *(iii)*: Seien $y_1, \dots, y_m$ linear unabhängig in L_H und $x_0 \in I$, $a_1, \dots, a_m \in \mathbb{K}$ mit $\sum_{i=1}^{m} a_i y_i(x_0) = 0$. $\Rightarrow$ $y = \sum_{i=1}^{m} a_i y_i$ löst

$$\begin{cases} y' = A(x)y \\ y(x_0) = 0 \end{cases}$$

$$\overset{24.2}{\Rightarrow} y = 0 \overset{(i)}{\Rightarrow} a_1 = \cdots = a_m = 0 \Rightarrow (iii).$$

Die restlichen Implikationen sind entweder offensichtlich (z. B. $(iii) \Rightarrow (ii) \Rightarrow (i)$) oder unmittelbare Folgerungen aus Standardresultaten der Linearen Algebra, siehe Teil II.

Bleibt zu zeigen, dass $\dim L_H = m$. Sei $\{e_i : i = 1, \ldots, m\}$ die kanonische Basis von $\mathbb{K}^m$. Aufgrund von Korollar 24.2 existiert genau eine Lösung $y = \varphi_i$ von

$$\begin{cases} y' = A(x)y \\ y(x_0) = e_i \end{cases}$$

für $x_0 \in I$ und $i = 1, \ldots, m$ und da $y' = A(x)y$ linear ist, ist $y = \sum\limits_{i=1}^{m} a_i\varphi_i$ eindeutige Lösung von

$$\begin{cases} y' = A(x)y \\ y(x_0) = \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} \in \mathbb{K}^m \end{cases}$$

und folglich $\{\varphi_i : i = 1, \ldots, m\}$ ein Erzeugendensystem von L_H, welches wegen der Äquivalenz $(i) \iff (iii)$ linear unabhängig ist. Also ist $\dim L_H = m$.

Satz 24.4 *y ist genau dann Lösung der gewöhnlichen Differentialgleichung n-ter Ordnung*

$$y^{(n)} = f(x, y, y', \ldots, y^{(n-1)})$$

wenn $\begin{pmatrix} y \\ y' \\ \vdots \\ y^{(n-1)} \end{pmatrix}$ *Lösung des Systems von n gewöhnlichen Differentialgleichungen erster Ordnung*

$$\begin{cases} y_0' = y_1 \\ y_1' = y_2 \\ \quad\vdots \\ y_{n-2}' = y_{n-1} \\ y_{n-1}' = f(x, y_0, y_1, \ldots, y_{n-1}) \end{cases}$$

ist.

Beweis

durch direktes Nachrechnen. □

Korollar 24.5 *Sei $f : \mathbb{R} \times \mathbb{R}^n \supseteq D \to \mathbb{R}$, $(x, Y) = (x, y, y_1, \dots, y_{n-1}) \mapsto f(x, Y) = f(x, y, y_1, \dots, y_{n-1})$ eine Funktion, die auf $B := [x_0 - a, x_0 + a] \times \{Y \in \mathbb{R}^n : |Y - Y_0| \leq b\}$ mit $x_0 \in \mathbb{R}$, $a, b \in \mathbb{R}^+$ und $Y_0 \in \mathbb{R}^n$ stetig und Lipschitz-stetig bezüglich Y ist. Dann existiert ein $h > 0$, so dass das Anfangswertpoblem*

$$\begin{cases} \qquad y^{(n)} = f(x, y, y', \dots, y^{(n-1)}) \\ \quad y(x_0) = y_0 \\ \quad y'(x_0) = (y')_0 \\ \qquad \vdots \\ y^{(n-1)}(x_0) = (y^{(n-1)})_0 \end{cases}$$

für alle $(y_0, (y')_0, \dots, (y^{(n-1)})_0) \in \mathbb{R}^n$ genau eine Lösung $y \in C^n([x_0 - h, x_0 + h], \mathbb{R})$ besitzt.

Korollar 24.6 *Sei $I \subseteq \mathbb{R}$ ein Intervall, $n \in \mathbb{N}$, $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$, $a_n \in \mathbb{K} \setminus \{0\}$ und seien $a_0, \dots, a_{n-1}, b : I \to \mathbb{K}$ stetige Funktionen. Dann besitzt das lineare Anfangswertpoblem n-ter Ordnung*

$$\begin{cases} \left(\sum_{k=0}^{n} a_k(x) \dfrac{d^k}{dx^k} \right) y = b(x) \\ \qquad \dfrac{d^k}{dx^k} y(x_0) = y_k, \qquad k = 0, \dots, n-1 \end{cases}$$

für alle $x_0 \in I$ und alle $y_0, \dots, y_{n-1} \in \mathbb{K}$ genau eine Lösung $y \in C^n(I, \mathbb{K})$.
Für den $\mathbb{K}$-Vektorraum L_H aller Lösungen der homogenen Differentialgleichung

$$\left(\sum_{k=0}^{n} a_k(x) \frac{d^k}{dx^k} \right) y = 0$$

gilt $\dim L_H = n$ *und jede n-elementige Menge $\{\varphi_1, \dots, \varphi_n\} \subseteq L_H$ ist genau dann ein Fundamentalsystem der homogenen Gleichung wenn für ein und damit für alle $x \in I$ die Wronski-Determinante*

$$W(x) := \det \begin{pmatrix} \varphi_1(x) & \varphi_2(x) & \dots & \varphi_n(x) \\ \varphi_1'(x) & \varphi_2'(x) & \dots & \varphi_n'(x) \\ \vdots & \vdots & \dots & \vdots \\ \varphi_1^{(n-1)}(x) & \varphi_2^{(n-1)}(x) & \dots & \varphi_n^{(n-1)}(x) \end{pmatrix}$$

von Null verschieden ist.

Beispiel

$$y'' - \frac{1}{2x} y' + \frac{1}{2x^2} y = 0, \quad x \in \mathbb{R}^+.$$

Durch Einsetzen des Ansatzes $y(x) = x^\alpha$, $\alpha \in \mathbb{R}$, kann man herausfinden, dass $\varphi_1(x) = x$ und $\varphi_2(x) = \sqrt{x}$ Lösungen der Differentialgleichung sind.
Es gilt

$$\begin{aligned} W(x) &= \det \begin{pmatrix} \varphi_1(x) & \varphi_2(x) \\ \varphi_1'(x) & \varphi_2'(x) \end{pmatrix} = \det \begin{pmatrix} x & \sqrt{x} \\ 1 & \frac{1}{2\sqrt{x}} \end{pmatrix} \\ &= -\frac{1}{2}\sqrt{x} \neq 0 \end{aligned}$$

für alle $x \in \mathbb{R}^+$. Somit ist $\{\varphi_1, \varphi_2\}$ ein Fundamentalsystem und

$$y_{hom}(x) = ax + b\sqrt{x}, \quad a, b \in \mathbb{K}$$

die allgemeine Lösung der obigen Gleichung.

Satz 24.7 *Sei f wie in Satz 24.1 und seien y_i, $i = 1, 2$, die eindeutigen Lösungen von*

$$\begin{cases} y_i' = f(x, y_i) \\ y_i(x_0) = y_{i0}. \end{cases}$$

Dann gilt:

$$\forall x \in [x_0 - h, x_0 + h] : \quad |y_1(x) - y_2(x)| \leq e^{L|x-x_0|} |y_{10} - y_{20}|.$$

Beweisstrategie
Die Behauptung ergibt sich aufgrund der Lösungsformel für lineare gewöhnliche Differentialgleichungen mit konstanten Koeffizienten und der Monotonie von Integralen.

Bemerkung
Für $f(x, y) = Ly$ gilt "=" in der Abschätzung aus Satz 24.7.
Für manche Funktionen, wie zum Beispiel $f(x, y) = \dfrac{y}{1 + x^2}$, ist die Abschätzung eher grob und kann verfeinert werden.

Wir stellen nun den Beweis des Satzes von Picard-Lindelöf vor. Als Erstes skizzieren wir die Beweisstrategie:

1. Schritt: Das Anfangswertproblem

$$\begin{cases} y' = f(x, y) \\ y(x_0) = y_0 \end{cases}$$

kann in Form der äquivalenten Integralgleichung

$$y(x) = y_0 + \int_{x_0}^{x} f(s, y(s))\, ds$$

geschrieben werden.

2. Schritt: Approximation der Lösung dieser Integralgleichung durch die sogenannte Picard-Iteration

$$\begin{aligned} y_1(x) &:= y_0 + \int_{x_0}^{x} f(s, y_0)\, ds \\ y_2(x) &:= y_0 + \int_{x_0}^{x} f(s, y_1(s))\, ds \\ &\vdots \\ y_{n+1}(x) &:= y_0 + \int_{x_0}^{x} f(s, y_n(s))\, ds \\ &\vdots \end{aligned}$$

3. Schritt: Nachweis, dass die Funktionenfolge $(y_n)_{n\in\mathbb{N}}$ gleichmäßig auf $[x_0-h, x_0+h]$ gegen eine Funktion y konvergiert.

4. Schritt: Nachweis, dass diese Funktion y die eindeutige Lösung des Anfangswertproblems auf $[x_0 - h, x_0 + h]$ ist.

Der 3. Schritt lässt sich wesentlich eleganter durchführen, wenn man einen verallgemeinerten Konvergenzbegriff verwendet, der für beliebige metrische Räume gilt. Wir führen im Folgenden diesen Konvergenzbegriff ein.

Definition *Sei (X, d) ein metrischer Raum, $A, M, \mathcal{O} \subseteq X$, $x_0 \in X$ und $\varepsilon > 0$.*

a) $B_\varepsilon(x_0) := \{x \in X : \ d(x, x_0) < \varepsilon\}$ *heißt offene ε-Kugel um x_0. Ist X ein normierter Raum, dann kann in dieser Definition $d(x, x_0)$ durch $\|x - x_0\|$ ersetzt werden.*

b) $\mathcal{O}$ *heißt offen, falls*

$$\forall x \in \mathcal{O} \ : \ \exists \varepsilon > 0 \ : \ B_\varepsilon(x) \subseteq \mathcal{O}$$

c) A *heißt abgeschlossen, falls* $X \backslash A$ *offen ist.*

d) $\overset{\circ}{M} := \{x \in M \ : \ \exists \varepsilon > 0 \ : \ B_\varepsilon(x) \subseteq M\}$ *heißt das Innere von M.*

e) $\overline{M} := X \backslash \, (X \overset{\circ}{\backslash M})$ *heißt Abschluss von M.*

f) $\partial M := \overline{M} \backslash \overset{\circ}{M}$ *heißt Rand von M.*

Beispiele

1) Sei $(X, d) = (\mathbb{R}^2, \| \cdot \|_2)$. Dann ist $B_1(0) = \overset{\circ}{B_1}(0)$ offen.

   $\overline{B_1(0)} = \{x \in \mathbb{R}^2 \ : \ \|x\|_2 \leq 1\}$ ist abgeschlossen.

   Außerdem gilt, dass $\partial B_1(0) = \{x \in \mathbb{R}^2 \ : \ \|x\|_2 = 1\}$ ist.

2) $(X,d) = (\mathbb{R}, |\cdot|)$. Dann ist $M = \bigcup_{n\in\mathbb{N}} [\frac{1}{2n}, \frac{1}{2n-1}]$ ist nicht offen und nicht abgeschlossen. Es ist $\overline{M} = M \cup \{0\}$ und $\partial M = \{\frac{1}{m} : m \in \mathbb{N}\} \cup \{0\}$.

Definition *Eine Folge $(x_n)_{n\in\mathbb{N}}$ in (X,d) heißt konvergent gegen $x \in X$ für $n \to \infty$ (Kurzschreibweise: $x_n \to x$ für $n \to \infty$ oder auch $\lim\limits_{n\to\infty} x_n = x$), falls*

$$\lim_{n\longrightarrow\infty} d(x_n, x) = 0,$$

also falls

$$\forall \varepsilon > 0\, \exists n_\varepsilon \in \mathbb{N}\, \forall n \geq n_\varepsilon \ : \ d(x_n, x) < \varepsilon.$$

Bemerkung
Der Grenzwert einer Folge $(x_n)_{n\in\mathbb{N}}$ ist eindeutig bestimmt. Nehmen wir an, neben x sei y noch ein weiterer Grenzwert. Dann gilt

$$\begin{array}{ccccccc} 0 & \leq & d(x,y) & \leq & d(x,x_n) & + & d(x_n,y) \\ & & & = & \underbrace{d(x_n,x)}_{\overset{n\to\infty}{\longrightarrow} 0} & + & \underbrace{d(x_n,y)}_{\overset{n\to\infty}{\longrightarrow} 0} \end{array}$$

und somit $d(x,y) = 0$, also $x = y$.

Bemerkungen

1) In $(\mathbb{R}^m, \|\cdot\|_2)$ gilt

$$\begin{array}{ccccc} x_n \xrightarrow{n\to\infty} x & \Longleftrightarrow & \sqrt{\displaystyle\sum_{i=1}^{m}((x_n)_i - x_i)^2} & \xrightarrow{n\to\infty} & 0 \\ & \Longleftrightarrow & \forall i \in \{1,\ldots,m\} \ : \ (x_n)_i & \xrightarrow{n\to\infty} & x_i. \end{array}$$

Dies ist der bereits bekannte Konvergenzbegriff in $\mathbb{R}^m$.

2) Sei $X = C^0([0,1])$, $d(x,y) = \sup\limits_{0\leq t\leq 1} |x(t) - y(t)|$. Dann ist (X,d) ein metrischer Raum und es gilt

$$\begin{array}{cl} x_n \xrightarrow{n\to\infty} x \Longleftrightarrow & \sup\limits_{0\leq t\leq 1} |x_n(t) - x(t)| \xrightarrow{n\to\infty} 0 \\ \Longleftrightarrow & \forall \varepsilon > 0\, \exists n_\varepsilon \in \mathbb{N}\, \forall n \geq n_\varepsilon : \ \sup\limits_{0\leq t\leq 1} |x_n(t) - x(t)| \ < \ \varepsilon \\ \Longleftrightarrow & \forall \varepsilon > 0\, \exists n_\varepsilon \in \mathbb{N}\, \forall n \geq n_\varepsilon : \ \forall t \in [0,1] : \ |x_n(t) - x(t)| \ < \ \varepsilon. \end{array}$$

Dies ist genau die gleichmäßige Konvergenz.

3) Sei $X = C^0([0,1])$, $d(x,y) = \left(\int_0^1 |x(t)-y(t)|^p \, dt\right)^{\frac{1}{p}}, 1 \leq p < \infty$. Dann ist (X,d) ein metrischer Raum und es gilt

$$\begin{aligned} x_n \xrightarrow{n\to\infty} x \quad &\Longleftrightarrow \left(\int_0^1 |x_n(t)-x(t)|^p \, dt\right)^{\frac{1}{p}} \xrightarrow{n\to\infty} 0 \\ &\Longleftrightarrow \forall \varepsilon > 0 \, \exists n_\varepsilon \in \mathbb{N} \, \forall n \geq n_\varepsilon : \int_0^1 |x_n(t)-x(t)|^p \, dt \; < \; \varepsilon. \end{aligned}$$

Dies nennt man auch Konvergenz im p-ten Mittel.

Satz 24.8 *Sei $(X, \|\cdot\|)$ ein normierter Raum, $(x_n)_{n\in\mathbb{N}}, (y_n)_{n\in\mathbb{N}} \subseteq X, (\alpha_n)_{n\in\mathbb{N}} \subseteq \mathbb{K}$ mit $x_n \to x, y_n \to y$ und $\alpha_n \to \alpha$ für $n \longrightarrow \infty$. Dann gilt*

$$\alpha_n x_n + y_n \xrightarrow{n\to\infty} \alpha x + y.$$

Beweis

$$\begin{aligned} \|\alpha x + y - (\alpha_n x_n + y_n)\| &= \|\alpha x - \alpha_n x + \alpha_n x - \alpha_n x_n + y - y_n\| \\ &\leq \|\alpha x - \alpha_n x\| + \|\alpha_n (x - x_n)\| + \|y - y_n\| \\ &\leq \underbrace{|\alpha - \alpha_n|}_{\to 0} \underbrace{\|x\|}_{<\infty} + \underbrace{|\alpha_n|}_{<\infty} \underbrace{\|x - x_n\|}_{\to 0} + \underbrace{\|y - y_n\|}_{\to 0} \quad \xrightarrow{n\to\infty} \quad 0 \end{aligned}$$

□

Satz 24.9 *Sei (X,d) ein metrischer Raum, $M \subseteq X$. Dann gilt*

$$\overline{M} \; = \; \{x \in X \; : \; \exists (x_n)_{n\in\mathbb{N}} \subseteq M : \; x_n \to x\}.$$

Beweis
Sei $\tilde{M} := \{x \in X \; : \; \exists (x_n)_{n\in\mathbb{N}} \subseteq M : \; x_n \to x\}$. Zu zeigen ist, dass $\tilde{M} = \overline{M}$.

$$\begin{aligned} x \in \tilde{M} \quad &\Longleftrightarrow \quad \forall \varepsilon > 0 \; \exists x_{n_\varepsilon} \in M : \; x_{n_\varepsilon} \in B_\varepsilon(x) \\ &\Longleftrightarrow \quad \forall \varepsilon > 0 : \; B_\varepsilon(x) \cap M \; \neq \emptyset \\ &\Longleftrightarrow \quad \forall \varepsilon > 0 : \; \neg \left(B_\varepsilon(x) \subseteq X \backslash M\right) \\ &\Longleftrightarrow \quad \neg \left(\exists \varepsilon > 0 : \; B_\varepsilon(x) \subseteq X \backslash M\right) \\ &\Longleftrightarrow \quad x \notin \overset{\circ}{(X \backslash M)} \\ &\Longleftrightarrow \quad x \in \overline{M} \end{aligned}$$

□

Definition *Eine Folge $(x_n)_{n\in\mathbb{N}}$ in einem metrischen Raum (X,d) heißt Cauchy-Folge, wenn*

$$\forall \varepsilon > 0 \; \exists n_\varepsilon \in \mathbb{N} \; \forall n, m \geq n_\varepsilon : \; d(x_n, x_m) < \varepsilon.$$

Satz 24.10 *Jede konvergente Folge in einem metrischen Raum (X, d) ist eine Cauchy-Folge.*

Die Umkehrung von Satz 24.10 gilt nicht immer.

Definition *Ein metrischer Raum heißt vollständig, wenn jede Cauchy-Folge in X gegen ein Element von X konvergiert. Ein vollständiger metrischer Raum heißt auch Fréchet-Raum. Ein vollständiger normierter Raum heißt auch Banachraum. Ein vollständiger Skalarproduktraum heißt auch Hilbertraum.*

Bemerkungen

1) $(\mathbb{R}^n, \|\cdot\|_p)$ ist ein Banachraum und für p=2 auch ein Hilbertraum.

2) $(\mathbb{Q}, |\cdot|)$ ist nicht vollständig. Gegenbeispiel: $(x_n)_{\mathbb{N}} \subseteq \mathbb{Q}$ mit $\lim\limits_{n\to\infty} = \sqrt{2}$, z.B. $x_n =$ die Dezimaldarstellung von $\sqrt{2}$ bis zur n-ten Stelle.

3) $C^0([a, b])$ mit $\|x\| := \sup\limits_{a\leq t\leq b} |x(t)|$ ist ein Banachraum. Das folgt aus der Vollständigkeit von $\mathbb{R}$ und aus der Tatsache, dass der gleichmäßige Limes von stetigen Funktionen wieder eine stetige Funktion ist.

4) $C^0([a, b])$ mit $\|x\| := \left(\int_a^b |x(t)|^p \, dt\right)^{\frac{1}{p}}$ ist nicht vollständig. Gegenbeispiel: $(x_n)_{n\in\mathbb{N}} \subseteq C^0([0, 1])$ mit

$$x_n(t) := \begin{cases} n^\alpha, \, t \leq \dfrac{1}{n} \\ t^{-\alpha}, \, t > \dfrac{1}{n} \end{cases}, \quad p = 2, \; 0 < \alpha < \frac{1}{2},$$

ist eine nicht konvergente Cauchy-Folge.

Satz 24.11 *Sei (X, d) ein vollständiger metrischer Raum und A eine nichtleere Teilmenge von X. Dann ist äquivalent:*

(i) (A, d) ist ein vollständiger metrischer Raum.

(ii) A ist eine abgeschlossene Teilmenge von X.

Beweisstrategie
Die Behauptung folgt aus den Definitionen von Vollständigkeit und Abgeschlossenheit unter Ausnutzung von Satz 24.9.

Definition *Seien $(X, d_X), (Y, d_Y)$ metrische Räume. Eine Abbildung $T : X \longrightarrow Y$ heißt stetig in $x_0 \in X$, wenn $\forall \varepsilon > 0 \, \exists \delta = \delta(\varepsilon, x_0) > 0$, so dass $d_Y(T(x), T(x_0)) < \varepsilon$ für alle $x \in X$ mit $d_X(x, x_0) < \delta$ gilt. T heißt stetig in $M \subseteq X$, falls T stetig in jedem $x \in M$ ist.*

Satz 24.12 *Für $T : (X, d_X) \longrightarrow (Y, d_Y)$ sind äquivalent:*

(i) T ist stetig für alle $x \in X$.

(ii) T ist folgenstetig für alle $x \in X$, d.h., $x_n \stackrel{n\to\infty}{\longrightarrow} x$ impliziert $T(x_n) \stackrel{n\to\infty}{\longrightarrow} T(x)$.

Beweisstrategie
Analog zum reellen Fall.

Bemerkung
Mit Hilfe des Konvergenzbegriffes in metrischen Räumen kann man auch die Begriffe der Differenzierbarkeit und der Integrierbarkeit für Banachraum-wertige Abbildungen verallgemeinern. Für Details sei auf die Literatur über Funktionalanalysis verwiesen.

Satz 24.13 (Banachscher Fixpunktsatz) *Sei (X, d) ein vollständiger metrischer Raum und $F : X \to X$ eine Kontraktion, d.h.*

$$\exists \kappa \in (0,1)\, \forall x, y \in X \;:\; d(F(x), F(y)) \le \kappa d(x, y).$$

Dann besitzt F einen eindeutigen Fixpunkt $x^\star$, d.h. $x^\star = F(x^\star)$.

Beweis
Wir zeigen zunächst die Eindeutigkeit: Angenommen, es existieren zwei Fixpunkte $x^\star \neq y^\star$. Dann gilt

$$d(x^\star, y^\star) \;=\; d(F(x^\star), F(y^\star)) \;\le\; \kappa d(x^\star, y^\star)$$

mit $\kappa \in (0,1)$. Daraus folgt $d(x^\star, y^\star) = 0$ und somit $x^\star = y^\star$.
Nun zur Existenz: Sei x_0 ein beliebig gewählter Startpunkt in X. Wir definieren die Folge

$$x_{n+1} \;:=\; F(x_n).$$

Dann gilt

$$\begin{aligned} \forall m \ge n : \qquad d(x_m, x_n) \qquad &\le \qquad \sum_{j=n}^{m-1} d(x_{j+1}, x_j) \quad \le \quad \sum_{j=n}^{m-1} \kappa^j d(x_1, x_0) \\ &\le \qquad d(x_1, x_0) \underbrace{\sum_{j=n}^{\infty} \kappa^j}_{= \frac{\kappa^n}{1-\kappa} \stackrel{n\to\infty}{\longrightarrow} 0} \end{aligned}$$

$$\Longrightarrow \qquad \forall \varepsilon > 0\, \exists n_\varepsilon \in \mathbb{N}\, \forall m, n > n_\varepsilon : d(x_m, x_n) \,<\, \varepsilon.$$

Also ist $(x_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge. Da X vollständig ist, existiert $x^\star \in X$, so dass $\lim\limits_{n\to\infty} = x^\star$. Da F eine Kontraktion ist, ist F stetig in X und daher

$$F(x^\star) \;=\; F(\lim_{n\to\infty} x_n) \;=\; \lim_{n\to\infty} \underbrace{F(x_n)}_{= x_{n+1}} \;=\; \lim_{n\to\infty} x_{n+1} \;=\; \lim_{n\to\infty} x_n \;=\; x^\star. \qquad \square$$

Mit Hilfe des Banachschen Fixpunktsatzes beweisen wir nun den Satz von Picard-Lindelöf:

Beweis von Satz 24.1
Wir wenden den Banachschen Fixpunktsatz auf die Abbildung $F : \mathcal{M} \longrightarrow \mathcal{M}$ mit

$$(F(y))(x) = y_0 + \int_{x_0}^{x} f(s, y(s))\ ds$$

im vollständigen metrischen Raum

$$\mathcal{M} := \left\{ y \in C^0([x_0 - h, x_0 + h], \mathbb{R}^m)\ :\ d(y, y_0) = \sup_{x \in [x_0-h, x_0+h]} |y(x) - y_0| \leq b \right\}$$

an. Wir zeigen nun:

i) $F(\mathcal{M}) \subseteq \mathcal{M}$, denn es ist für alle $y \in C^0([x_0 - h, x_0 + h], \mathbb{R}^m)$:

$$F(y) \in C^0([x_0 - h, x_0 + h], \mathbb{R}^m)$$

und außerdem gilt

$$d(F(y), y_0) = \sup_{x \in [x_0-h, x_0+h]} \left| \int_{x_0}^{x} f(s, y(s))\ ds \right| \leq hM \leq b,$$

also gilt $f(\mathcal{M}) \subseteq \mathcal{M}$.

ii) F ist eine Kontraktion, da

$$\begin{aligned} d(F(y), F(\tilde{y})) &= \sup_{x \in [x_0-h, x_0+h]} \left| \int_{x_0}^{x} f(s, y(s)) - f(s, \tilde{y}(s))\ ds \right| \\ &\leq \sup_{x \in [x_0-h, x_0+h]} \int_{x_0}^{x} |f(s, y(s)) - f(s, \tilde{y}(s))|\ ds \\ &\leq \sup_{x \in [x_0-h, x_0+h]} \int_{x_0}^{x} L|y(s) - \tilde{y}(s)|\ ds \\ &\leq Lh \sup_{x \in [x_0-h, x_0+h]} |y(x) - \tilde{y}(x)| \\ &= Lh\, d(y, \tilde{y})\ \leq\ \frac{1}{2} d(y, \tilde{y}) \end{aligned}$$

aufgrund der Definition von h gilt. F erfüllt also die Voraussetzungen des Banachschen Fixpunktsatzes. Somit besitzt F einen <u>eindeutigen</u> Fixpunkt $y^\star \in \mathcal{M}$ mit

$$y^\star(x) = y_0 + \int_{x_0}^{x} f(s, y^\star(s))\ ds. \tag{1}$$

Wegen $y^\star \in \mathcal{M}$ ist die rechte Seite dieser Gleichung und damit auch $y^\star \in C^1([x_0 - h, x_0 + h], \mathbb{R}^m)$. Damit können wir die Gleichung (1) einmal nach x ableiten und erhalten, dass $y^\star$ eindeutige Lösung des Anfangswertproblems ist. □

Bemerkungen

1) Die Picard-Iteration kann nicht nur abstrakt für Beweise genutzt werden, sondern auch konkret verwendet werden zur näherungsweisen Berechnung von Lösungen gewöhnlicher Differentialgleichungen, sowohl mit als auch ohne Einsatz von Computern. In der Praxis sind allerdings häufig andere Verfahren besser geeignet. Das einfachste Beispiel für ein solches Verfahren ist das sogenannte explizite Euler-Verfahren (Polygonzugverfahren). Bei diesem Verfahren wählt man zur näherungsweisen Berechnung von

$$y(x) = y_0 + \int_{x_0}^{x} f(s, y(s))\, ds$$

eine Zerlegung $x_0 < x_1 < ... < x_n = x_0 + h$ von $[x_0, x_0 + h]$ und berechnet

$$\begin{aligned} y(x_0) &:= y_0\,, \\ y(x_1) &:= y(x_0) + (x_1 - x_0) f(x_0, y(x_0))\,, \\ y(x_2) &:= y(x_1) + (x_2 - x_1) f(x_1, y(x_1))\,, \\ &\;\vdots \\ y(x_0 + h) &:= y(x_{n-1}) + (x_0 + h - x_{n-1}) f(x_{n-1}, y(x_{n-1})) \end{aligned}$$

und dann

$$y(x_k + \vartheta(x_{k+1} - x_k)) \;:=\; y(x_k) + \vartheta(y(x_{k+1}) - y(x_k))$$

für $\vartheta \in]0, 1[$ und $k = 0, ..., n - 1$. Entsprechendes macht man für das Intervall $[x_0 - h, x_0]$.
Das explizite Euler-Verfahren kann auch abstrakt für Beweise genutzt werden, beispielsweise für den Beweis des Existenzsatzes von Peano.
Für weitere Näherungsverfahren, zum Beispiel das implizite Eulerverfahren, das Runge-Kutta-Verfahren oder das Crank-Nicolson-Verfahren, deren Genauigkeit, deren Vor- und Nachteile sowie deren Verwendungsmöglichkeiten sei auf die Literatur über numerische Mathematik verwiesen.

2) In manchen Fällen ist es sinnvoller, anstelle einer expliziten oder einer näherungsweisen Berechnung von Lösungen gewöhnlicher Differentialgleichungen lieber Aussagen über qualitative Eigenschaften der Lösungen (zum Beispiel Konvergenzverhalten, Periodizität, ...) herzuleiten. Für nähere Details sei auf die Literatur über dynamische Systeme verwiesen.

3) Die in diesem Abschnitt diskutierten Resultate und Methoden lassen sich zum Teil auch auf partielle Differentialgleichungen übertragen. Für nähere Details sei auf die Literatur über partielle Differentialgleichungen verwiesen.