MATHEMATIK I-II FÜR INF, SWT

WOLF–PATRICK DÜLL

UNIVERSITÄT STUTTGART

WINTERSEMESTER 2016/17, SOMMERSEMESTER 2017

Teil I

Grundlagen

1 Logik

Aussagen und die formale mathematische Sprache

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.

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 w(A), 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 $\rightarrow$ B: "aus A folgt B", "A impliziert B", "wenn A, dann B"

A	B	A$\rightarrow$ B
1	1	1
1	0	0
0	1	1
0	0	1

Die Zuweisung w(A $\rightarrow$ B)=1 in der dritten und vierten Zeile nennt man auch "ex falso quodlibet".
Ist A $\rightarrow$ B wahr, also w(A $\rightarrow$ 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.

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

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

A	B	$\neg A$	$\neg$ B	$A \rightarrow B$	$\neg B \rightarrow \neg A$	$A \rightarrow B \Leftrightarrow \neg B \rightarrow \neg A$
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$\rightarrow B \Leftrightarrow \neg(A \wedge \neg$ B)

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

7) (A$\rightarrow B) \wedge (\neg A \rightarrow B) \Leftrightarrow B$ (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) $(A \rightarrow B) \wedge (B \rightarrow C) \Rightarrow A \rightarrow C$ (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.

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 NAND B) NAND (A NAND B),

$\neg(A \wedge B) \Leftrightarrow$ ((A NAND B) NAND (A NAND B)) NAND ((A NAND B) NAND (A 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.

Verneinung von Quantoren:

Es gilt:

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

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

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

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.

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.

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:

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

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.

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

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.

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).

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.

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. $(m_1, m_2) \neq (m_2, m_1)$, falls $m_1 \neq m_2$ ist. Insbesondere gilt $(m_1, m_2) = (m_1', m_2') \Leftrightarrow m_1 = m_1' \wedge m_2 = m_2'$.

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

a) $\bigcup\limits_{k=1}^{n} M_k = M_1 \cup M_2 \cup \cdots \cup M_n = \{x : \exists j \in \{1, \dots, n\} : x \in M_j\}$
die Vereinigung der Mengen $M_1, \dots, M_n$.

b) $\bigcap\limits_{k=1}^{n} M_k = M_1 \cap M_2 \cap \cdots \cap M_n = \{x : \forall j \in \{1, \dots, n\} : x \in M_j\}$
die Schnittmenge von $M_1, \dots, M_n$.

c) $\prod\limits_{k=1}^{n} M_k = M_1 \times M_2 \times \cdots \times M_n = \{ \underbrace{(x_1, \dots, x_n)}_{\text{geordnetes n-Tupel}} : \forall j \in \{1, \dots, n\} : x_j \in M_j\}$
das kartesische Produkt von $M_1, \dots, M_n$. Ist $M_1 = M_2 = \cdots = M_n = M$, dann schreibt man auch M^n statt $\underbrace{M \times \cdots \times M}_{n-fach}$.

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) $L \setminus (M \cap N) = (L \setminus M) \cup (L \setminus N)$

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.

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 $f : M_1 \to M_2, x \mapsto y$.

Eine Relation $R \subseteq M \times M$ heißt Äquivalenzrelation (auf M), falls
(i) $\forall a \in M : aRa$ (R ist reflexiv)
(ii) $\forall a, b \in M : aRb \Leftrightarrow bRa$ (R ist symmetrisch)
(iii) $\forall a, b, c \in M : aRb \wedge bRc \Rightarrow aRc$ (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. $\sim$).

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 $\cdot$)*

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

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

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

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

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

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

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

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

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

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

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).*

Definition *(Summenzeichen, rekursive Definition)*

$$\sum_{k=1}^{1} a_k := a_1$$

$$\sum_{k=1}^{n+1} a_k := \left(\sum_{k=1}^{n} a_k\right) + a_{n+1}$$

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 $\mathbb{R}$)*

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

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*

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

Ist $x \neq 0$, dann sei außerdem $x^{-n} := \dfrac{1}{x^n}$.

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

Satz 3.10

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

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

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)*

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) $a > 0 \Rightarrow \frac{1}{a} > 0$

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

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

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

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

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

Definition *(Intervalle)*

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

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

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

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.

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.*

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

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

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

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.* $\sqrt{2} \in \mathbb{R} \backslash \mathbb{Q}$.

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:*

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

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

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

4 Zahlentheorie

4.1 Kombinatorik

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

a) *Jedes r-Tupel $(a_1, \ldots, a_r) \in M^r$ 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 $(a_1, \ldots, a_r) \in M^r$ 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 $\{a_1, \ldots, a_r\} \in \mathscr{P}(M)$ 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 $(a_1, \ldots, a_r) \in M^r$ mit $a_1 \leq a_2 \leq \cdots \leq a_r$ 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)

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

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

b) (Fakultät)

$$0! := 1, \quad n! := \prod_{k=1}^{n} k, \quad n \in \mathbb{N}, \qquad also\ n! = 1 \cdot 2 \cdot \ldots \cdot n.$$

c) (Binomialkoeffizient)

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

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

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

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.

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. $(a+b)^4 = a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + b^4$.

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.

Definition

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

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

der größte gemeinsame Teiler von a *und* b*.*

b) *Für* $m,n \in \mathbb{N}$ *heißt*

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

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

Hilfssatz 4.5

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

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

$$
\begin{aligned}
& n = q_1 m + r_1 \wedge r_1 \leq m - 1, \\
& m = q_2 r_1 + r_2 \wedge r_2 \leq r_1 - 1 \leq m - 2, \\
& r_1 = q_3 r_2 + r_3 \wedge r_3 \leq r_2 - 1, \\
& \vdots \\
& r_{N-2} = q_N r_{N-1} + r_N \wedge r_N \leq r_{N-1} - 1, \\
& r_{N-1} = q_{N+1} r_N + 0,
\end{aligned}
$$

und es gilt $r_N =$*ggT*(m, n)*.*

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:* $p | mn \Rightarrow p | m \vee p | n$*.*

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.*

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.*

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.

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

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

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.

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.

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* $\sum_{k=0}^{n} a_k x^k$ *mit* $a_k \in R$ *für* $k = 0,\ldots,n$ *heißt Polynom und eine Funktion* $p : R \to R,\ x \mapsto \sum_{k=0}^{n} a_k x^k$ *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 $a_0, \dots a_n, b_0, \dots, b_m \in \mathbb{R}$ 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 $k = 0, \dots, n$.

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 $a_0, \dots, a_n, b_0, \dots, b_m$ 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 $p(x) = \sum_{k=0}^{n} a_k x^k,\ q(x) = \sum_{k=0}^{m} b_k x^k,\ n > m$ und seien $\alpha, \beta \in \mathbb{R}$. Dann gilt*

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

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

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

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.

Satz 5.6 *Jedes reelle Polynom vom* $Grad \geq 1$ *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 :*

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

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

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

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

Satz 5.9 *Sei $p(x) = \sum_{k=0}^{n} a_k x^k$ 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 $b \mid a_n$.*

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

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

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 $\frac{1}{2}n(n+1)$.

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

$$p(x) = \sum_{k=0}^{n} y_k L_k(x) \quad \textit{mit}$$

$$L_k(x) = \frac{(x-x_0)\cdot \ldots \cdot (x-x_{k-1})\cdot (x-x_{k+1})\cdot \ldots \cdot (x-x_n)}{(x_k-x_0)\cdot \ldots \cdot (x_k-x_{k-1})\cdot (x_k-x_{k+1})\cdot \ldots \cdot (x_k-x_n)}$$

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.

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 $^\circ$ (Grad) ist dann definiert durch $(\text{Bogenmaß} : \pi) \cdot 180^\circ$.

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

$$\sin\varphi := y \quad (\text{Sinus von } \varphi),$$

$$\cos\varphi := x \quad (\text{Kosinus von } \varphi)$$

sowie

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

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

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

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

$$\sin(-\varphi) = -\sin\varphi, \qquad \cos(-\varphi) = \cos\varphi$$
$$\sin(\tfrac{\pi}{2}-\varphi) = \sin(\tfrac{\pi}{2}+\varphi), \qquad \cos(\tfrac{\pi}{2}-\varphi) = -\cos(\tfrac{\pi}{2}+\varphi).$$

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

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

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

$$\sin(\alpha+\beta) = \sin\alpha\cdot\cos\beta + \cos\alpha\cdot\sin\beta,$$
$$\cos(\alpha+\beta) = \cos\alpha\cdot\cos\beta - \sin\alpha\cdot\sin\beta.$$

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

$$1 - \frac{1}{2}\varphi^2 \leq \cos\varphi \leq 1,$$
$$\cos\varphi \cdot |\varphi| \leq |\sin\varphi| \leq |\varphi|.$$

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

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

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

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

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

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

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.

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) $z \cdot \overline{z} = |z|^2,\; \dfrac{1}{z} = \dfrac{\overline{z}}{|z|^2}$ *(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:

$$||z| - |w|| \leq |z - w|$$

Bemerkung

Die Abbildung $z \mapsto \bar{z}$ 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: $\arg(\mathrm{z}) := \varphi$ (Argument von z).

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*

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

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

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

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 $k = 0, 1, \ldots, n-1$.

Definition *Ein Polynom*

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

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* $w^2 = b^2 - 4ac$.

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

Korollar 6.10 *Jedes komplexe Polynom*

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

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.*

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

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

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.*

Teil II
Lineare Algebra

7 Lineare Gleichungssysteme

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

wobei

1. $l \leq \min\{m,n\}$,
2. $j_1 < j_2 < ... < j_l$,
3. $\forall i \in \{1,...,l\} \forall j \in \{j_i,...,n\} : c_{ij} \in K$,
4. $\forall i \in \{1,...,n\} : d_i \in K$,
5. $\forall i \in \{1,...,l\} : c_{ij_i} \neq 0$.

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.

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 $\cdot : K \times V \to V$, $(\lambda, x) \mapsto \lambda x = \lambda \cdot x$, so dass gilt*

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

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.*

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

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

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

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

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*

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

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

1. *U ist ein Untervektorraum.*

2. *$U \neq \emptyset \wedge \forall x,y \in U \, \forall \lambda, \mu \in K : \lambda x + \mu y \in U$.*

Definition

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

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

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*

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

wobei

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

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

$(N1)$	$\|x\| \geq 0 \wedge (\|x\| = 0 \Leftrightarrow x = 0)$	*(Positive Definitheit)*
$(N2)$	$\|\lambda x\| = \|\lambda\| \, \|x\|$	*(Homogenität)*
$(N3)$	$\|x + y\| \leq \|x\| + \|y\|$	*(Dreiecksungleichung)*

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

$(M1)$	$d(x,y) \geq 0 \wedge (d(x,y) = 0 \Leftrightarrow x = y)$	*(Positive Definitheit)*
$(M2)$	$d(x,y) = d(y,x)$	*(Symmetrie)*
$(M3)$	$d(x,y) \leq d(x,z) + d(z,y)$	*(Dreiecksungleichung)*

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

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

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*

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

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*

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

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

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

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*.*

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*

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

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

$$a \times b = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \times \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} a_2 b_3 - b_2 a_3 \\ a_3 b_1 - b_3 a_1 \\ a_1 b_2 - b_1 a_2 \end{pmatrix}.$$

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*

$$E = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3 : n_1x_1 + n_2x_2 + n_3x_3 - d = 0 \right\} = \{ x \in \mathbb{R}^3 : \langle n, x \rangle = d \} \tag{1}$$

gilt. Außerdem gilt:

$$\forall a, b \in E : \langle n, b - a \rangle = 0 \tag{2}$$

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 $n_1x_1 + n_2x_2 + n_3x_3 - d = 0$ 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 $n_1x_1 + n_2x_2 + n_3x_3 = 0$ als auch $-n_1x_1 - n_2x_2 - n_3x_3 = 0$ 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 $\left\{\pm \frac{q \times r}{\|q \times r\|}\right\}$ 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

1) $\mathbb{C}^n = \left\{ \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} : z_1, \ldots, z_n \in \mathbb{C} \right\}$ 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 $v_1, \dots, v_n \in M$ gibt, so dass $x = \sum_{i=1}^{n} \lambda_i v_i$ mit $\lambda_i \in K$ für $i = 1, \dots, n$ gilt. x heißt dann auch eine Linearkombination von $v_1, \dots, v_n$.*

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 $\{v_1, \dots, v_n\}$ von M gilt:*

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

In diesem Fall heißen auch die Vektoren $v_1, \dots, v_n$ 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.*

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* $v_i \exists! (\lambda_1, \ldots, \lambda_n) \in K^n : x = \sum_{i=1}^{n} \lambda_i v_i \wedge \forall i \in \{1, \ldots, n\} : \lambda_i \neq 0.$

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.* $\wedge \forall B' \in \mathscr{P}(V) : B \subseteq B' \wedge B'$ *l.u.* $\Rightarrow B' = B.$

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

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* $b_i \wedge$
$\exists! (\lambda_1, \ldots, \lambda_n) \in K^n : x = \sum_{i=1}^{n} \lambda_i b_i \wedge \forall i \in \{1, \ldots, n\} : \lambda_i \neq 0.$
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 $B = \{b_1, \dots, b_n\}$ und $M = \{v_1, \dots, v_m\}$ 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.*

Korollar 8.16 *Sei V ein K-Vektorraum mit einer endlichen Basis $\{b_1, \dots, b_n\}$.*
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: $dim V = \infty$.

Satz 8.17 *Jeder Vektorraum besitzt eine Basis.*

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:

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

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 $\Leftrightarrow$ $Kern(L) = \{0\}$.

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:*

$$dim V = dim Kern(L) + dim Bild(L)$$

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* $y \in B$.

9.2 Lineare Abbildungen und Matrizen

Defintion *Sei K ein Körper und* $m, n \in \mathbb{N}$. *Eine Abbildung* $A : \{1, \dots, m\} \times \{1, \dots, n\} \to K, (i, j) \mapsto a_{ij} \in K$ *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*

$$Ax = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \sum\limits_{j=1}^{n} a_{1j} x_j \\ \vdots \\ \sum\limits_{j=1}^{n} a_{mj} x_j \end{pmatrix}.$$

Satz 9.5 *Sei V ein K-Vektorraum mit endlicher Basis* $B = \{b_1, \ldots, b_n\}$, *W ein K-Vektorraum mit endlicher Basis* $C = \{c_1, \ldots, c_m\}$.

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 $B = \{b_1, \ldots, b_n\}$ zur Bezeichnung einer geordneten Basis, d.h. eines n-Tupels $(b_1, \ldots, b_n) \in V^n$ mit der Eigenschaft, dass $\{b_1, \ldots, b_n\}$ eine Basis von V ist. Die Indizes legen also die Reihenfolge der Basisvektoren fest. Entsprechendes gilt für $C = \{c_1, \ldots, c_m\}$.

Bezeichnung

Sei $M \in K^{m\times n}$. Dann sei $L_M : K^n \to K^m$ die eindeutig bestimmte lineare Abbildung, für die $M = M_{L_M}^{\mathscr{E}_m \mathscr{E}_n}$, wobei $\mathscr{E}_j = \{e_1, \ldots, e_j\}$ die geordnete kanonische Basis von K^j mit $j \in \{m, n\}$ sei.

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.*

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:

$$M_{T^{-1}}^{B_V,B_V} = (M_T^{B_V,B_V})^{-1}.$$

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* $GL(n,K)$*.*

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.

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* $(AB)^{-1} = B^{-1}A^{-1}$*.*

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 $M_L^{B',B'}=S^{-1}M_L^{B,B}S$.

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. $Ax = E_n$,*

$$\begin{array}{rcccccc}
a_{11}x_1 + \cdots + a_{1n}x_n & = & 1 & \textit{bzw.} & 0 & \cdots & \textit{bzw.} \quad 0 \\
a_{21}x_1 + \cdots + a_{2n}x_n & = & 0 & \text{"} & 1 & \cdots & \text{"} \quad 0 \\
 & \vdots & & & & & \\
a_{n1}x_1 + \cdots + a_{nn}x_n & = & 0 & \text{"} & 0 & \cdots & \text{"} \quad 1
\end{array}$$

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 $a \in K \setminus \{0\}$ 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 $a \in K \setminus \{0\}$ oder Addition eines Vielfachen einer Spalte zu einer anderen Spalte.*

Satz 9.13 *Sei $A \in K^{m \times n}$. Dann gilt:*

$$Zeilenrang\ von\ A \ = \ Spaltenrang\ von\ A$$

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:*

$$\begin{array}{lrcl}
(SP1) & \langle x,x\rangle & \geq & 0 \quad \wedge \quad (\langle x,x\rangle = 0 \Leftrightarrow x = 0), \\
(SP2) & \langle x,y\rangle & = & \langle y,x\rangle, \\
(SP3) & \langle \lambda x + y, z\rangle & = & \lambda\langle x,z\rangle + \langle y,z\rangle.
\end{array}$$

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:*

$$\begin{array}{lrcl}
(SP1) & \langle x,x\rangle & \geq & 0 \quad \wedge \quad (\langle x,x\rangle = 0 \Leftrightarrow x = 0), \\
(SP2) & \langle x,y\rangle & = & \overline{\langle y,x\rangle}, \\
(SP3) & \langle \lambda x + y, z\rangle & = & \lambda\langle x,z\rangle + \langle y,z\rangle.
\end{array}$$

V zusammen mit einem komplexen Skalarprodukt heißt (komplexer) Skalarproduktraum oder unitärer Vektorraum.

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 $x, y \in V \setminus \{0\}$ *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.*

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:*

$$\begin{array}{ll} (N1) & \|x\| \geq 0 \quad \wedge \quad (\|x\| = 0 \Leftrightarrow x = 0), \\ (N2) & \|\lambda x\| = |\lambda|\,\|x\|, \\ (N3) & \|x+y\| \leq \|x\| + \|y\| \quad (Dreiecksungleichung). \end{array}$$

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.

Satz 10.4 (Parallelogrammgleichung) *Sei $(V, \langle\cdot,\cdot\rangle)$ ein Skalarproduktraum und $\|\cdot\|$ die von $\langle\cdot,\cdot\rangle$ induzierte Norm. Dann gilt:*

$$\forall x, y \in V : \quad \|x+y\|^2 + \|x-y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right).$$

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.

Definition *Sei V ein reeller Skalarproduktraum und seien $x,y \in V\setminus\{0\}$. 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*

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

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*

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

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.*

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.* $\Leftrightarrow M_\sigma^B$ *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.*

Satz 10.16 *Jedes Orthogonalsystem $(v_i)_{i\in I}$ ist linear unabhängig.*

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*

$$\forall n \in I : \ Span\{v_1, \ldots, v_n\} = Span\{w_1, \ldots, w_n\}.$$

$(v_i)_{i\in I}$ kann durch das Gram-Schmidtsche Orthogonalisierungsverfahren konstruiert werden:

1. Schritt: Sei $v_1 := \dfrac{1}{\|w_1\|} w_1$.

n-ter Schritt: Sei $\{v_1, \ldots, v_{n-1}\}$ 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* $\{v_1, \ldots, v_n\}$. *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)*

Satz 10.19 *Es seien* $(V, \langle\cdot,\cdot\rangle)$ *ein Skalarproduktraum mit induzierter Norm* $\|\cdot\|$, U *ein endlich-dimensionaler Unterraum von* V, $\{u_1, \ldots, u_n\}$ *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.

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 $\{v_1, \dots, v_n\} \subseteq V$ ist $\{Lv_1, \dots, Lv_n\}$ 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.* $\mathbb{C}^n$.

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* $A^{-1} = \overline{A}^{\mathsf{T}}$.

Satz 10.23

1. $O(n) := \{A \in GL(n,\mathbb{R}) : A^{-1} = A^{\mathsf{T}}\}$ *ist bzgl. der Matrizenmultiplikation eine Gruppe, d. h.* $O(n)$ *ist eine Untergruppe von* $GL(n,\mathbb{R})$ *– die sogenannte orthogonale Gruppe.*

2. $U(n) := \{A \in GL(n,\mathbb{C}) : A^{-1} = \overline{A}^{\mathsf{T}}\}$ *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 $(a_1,\ldots,a_n) \in V^n$. Dann heißt $\det(a_1,\ldots,a_n)$ das orientierte (n-dimensionale) Volumen des von den Vektoren $a_1,\ldots,a_n$ aufgespannten Spats und $|\det(a_1,\ldots,a_n)|$ das (n-dimensionale) Volumen des von $a_1,\ldots,a_n$ 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 $a_1,\ldots,a_n$. Dann ist $\det(A) := \det(a_1,\ldots,a_n)$.

Kurzschreibweise: Statt $\det\left(\begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix}\right)$ *schreibt man auch* $\begin{vmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nn} \end{vmatrix}$.

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.

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 $\det(A) = \prod_{i=1}^{n} a_{ii}$.

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* $\quad \det(A) = a_{11}a_{22} - a_{12}a_{21}$.

Satz 11.4 *Sei* $A \in K^{n\times n}$ *Dann ist äquivalent:*

(i) Die Zeilenvektoren von A *sind linear unabhängig.*

(ii) $\det(A) \neq 0$.

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:*

$$\det(AB) = \det(A)\det(B).$$

Bemerkung

Im Allgemeinen gilt aber nicht $\det(A+B) = \det(A) + \det(B)$.

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:

$$\det(A) = \det(B).$$

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.

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*

$$\det(A^\mathsf{T}) = \det(A).$$

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 $i \in \{1,\ldots,n\}$:*

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij}) \quad \textit{(Entwicklung nach der } i\textit{-ten Zeile).}$$

Korollar 11.12 *Seien A, A_{ij} wie in Satz 11.11. Dann gilt für jedes $j \in \{1,\ldots,n\}$:*

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij}) \quad \textit{(Entwicklung nach } j\textit{-ter Spalte).}$$

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.*

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 $det(M_{Id}^{B',B}) > 0$ 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.

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 λ
$\Leftrightarrow v \in Kern(L - \lambda Id) \setminus \{0\}$

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 λ
$\Leftrightarrow x \in Kern(A - \lambda E_n) \setminus \{0\}$

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.

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.*

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*.*

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* $\overline{\lambda}$*.*

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* $Sp(A) = \sum_{i=1}^{n} a_{ii}$ *(die sogenannte Spur von* A*).*

Korollar 12.7 *Seien $\lambda_1, \ldots, \lambda_k$ die (paarweise verschiedenen) Eigenwerte von $A \in \mathbb{K}^{n \times n}$. Dann gilt:*

(i) $\prod_{i=1}^{k} \lambda_i^{n_a(\lambda_i)} = det(A)$

(ii) $\sum_{i=1}^{k} n_a(\lambda_i)\, \lambda_i = Sp(A)$

Satz 12.8 *Seien $A, B \in \mathbb{K}^{n \times n}$ ähnlich. Dann gilt: $\chi_A = \chi_B$.*

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).

Korollar 12.10 *Sei λ_i Eigenwert von $A \in \mathbb{K}^{n \times n}$. Dann gilt $n_g(\lambda_i) \leq n_a(\lambda_i)$.*

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* $\{d_{jj} : j = 1, \ldots, n\} = \sigma(A)$.

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.*

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.

Satz 12.16 *Sei $A \in \mathbb{K}^{n\times n}$ selbstadjungiert. Dann existiert eine Orthonormalbasis von $\mathbb{K}^n$ aus Eigenvektoren 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 :*

$$A \text{ ist} \begin{cases} \text{positiv definit} \\ \text{positiv semidefinit} \\ \text{negativ definit} \\ \text{negativ semidefinit} \end{cases} \Leftrightarrow \ \sigma(A) \subseteq \begin{cases} \mathbb{R}^+ \\ \mathbb{R}_0^+ \\ \mathbb{R}^- \\ \mathbb{R}_0^- \end{cases}$$

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 $q_A : \mathbb{R}^n \to \mathbb{R}\,,\, x \longmapsto x^T A x$ 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.

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*

$$M_{L_A}^{B,B} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & -\sin\varphi \\ 0 & \sin\varphi & \cos\varphi \end{pmatrix}$$ *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*

$$M_{L_A}^{B,B} = \begin{pmatrix} J(k_{11},\lambda_1) & & & \dots & & & 0 \\ & \ddots & & & & & \\ & & J(k_{1l_1},\lambda_1) & & & & \\ & & & \ddots & & & \\ & & & & J(k_{m1},\lambda_m) & & \\ & & & & & \ddots & \\ 0 & \dots & & & & & J(k_{ml_m},\lambda_m) \end{pmatrix}$$

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.*

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*

$$\forall \varepsilon > 0 \; \exists n_\varepsilon \in \mathbb{N} \; \forall n > n_\varepsilon : | a_n - a | < \varepsilon \qquad (K)$$

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.*

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.

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) $(a \geq 0\ \wedge\ \forall n \in \mathbb{N}:\ a_n \geq 0)\ \Rightarrow\ \forall m \in \mathbb{N}:\ \lim\limits_{n\to\infty}\sqrt[m]{a_n} = \sqrt[m]{a}$

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

$$\lim_{n\to\infty} a_n \leq \lim_{n\to\infty} b_n \,.$$

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*

$$\lim_{n\to\infty} b_n = a.$$

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*

$$\lim_{n\to\infty} a_n = \sup\{a_n : n \in \mathbb{N}\}\,.$$

Korollar 13.6 *Jede monoton fallende, nach unten beschränkte reelle Folge* $(a_n)_{n\in\mathbb{N}}$ *ist konvergent und es gilt*

$$\lim_{n\to\infty} a_n = \inf\{a_n : n \in \mathbb{N}\}\,.$$

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: $x = \liminf\limits_{n\to\infty} x_n = \underline{\lim\limits_{n\to\infty}} x_n$.

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.*

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* $\lim\limits_{n\to\infty} x_{n_k}=x$*, dann gilt auch* $\lim\limits_{n\to\infty} x_n = x$*.*

Satz 13.11 $\mathbb{R}$ *ist vollständig, d.h. jede Cauchy-Folge in* $\mathbb{R}$ *hat eine Grenzwert in* $\mathbb{R}$*.*

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 $m = \min\{n \in \mathbb{N} : n > -x\}$.*

1. *Die Folge $\left(\left(1+\frac{x}{n}\right)^n\right)_{n\in\mathbb{N}}$ ist ab dem m-ten Folgenglied monoton wachsend.*
2. *Die Folge $\left(\left(1-\frac{x}{n}\right)^{-n}\right)_{n\in\mathbb{N}}$, definiert für $n \geq m$, ist monoton fallend.*
3. $\left\{\left[\left(1+\frac{x}{n}\right)^n, \left(1-\frac{x}{n}\right)^{-n}\right] : n \in \mathbb{N} \wedge n > |x|\right\}$ *ist eine Intervallschachtelung und somit existiert*
$$e^x := \lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n = \lim_{n\to\infty}\left(1-\frac{x}{n}\right)^{-n}.$$

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) $\forall x \in \mathbb{R} : e^{-x} = \frac{1}{e^x}$

(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) $\forall x \in \mathbb{R} : e^x \geq 1 + x \wedge (e^x = 1 + x \Leftrightarrow x = 0)$

(viii) $\forall x, y \in \mathbb{R} : x < y \implies e^x < e^y$

(ix) $\forall x \in \mathbb{R} : |x| < 1 \implies |e^x - 1| \leq \frac{|x|}{1 - |x|}$

(x) $\lim_{n \to \infty} e^{\frac{1}{n}} = 1 = \lim_{n \to \infty} e^{-\frac{1}{n}}$.

(xi) $\forall m \in \mathbb{N}_0 : \lim_{n \to \infty} n^m e^{-n} = 0$.

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) $\forall a, b \in \mathbb{R}^+ : \ln \frac{a}{b} = \ln a - \ln b$

(vi) $\forall a \in \mathbb{R}^+ \forall n \in \mathbb{N} : \ln a^n = n \cdot \ln a$

(vii) $\forall x \in \mathbb{R}^+ : \ln x \leq x - 1 \wedge (\ln x = x - 1 \Leftrightarrow x = 1)$

(viii) $\forall m \in \mathbb{N} : \lim_{n \to \infty} n^{-\frac{1}{m}} \ln n = 0$

(ix) $\forall a \in \mathbb{R}^+ \forall x \in \mathbb{Q} : a^x = e^{x \cdot \ln a}$

Definition (Reelle Potenzen) *Für $a \in \mathbb{R}^+$ und $x \in \mathbb{R}$ sei $a^x := e^{x \cdot \ln a}$.*

Satz 13.15 *1.* $\forall a \in \mathbb{R}^+ \forall x, y \in \mathbb{R} : a^{x+y} = a^x a^y$

2. $\forall a \in \mathbb{R}^+ \forall x, y \in \mathbb{R} : (a^x)^y = a^{xy}$

3. $\forall a, b \in \mathbb{R}^+ \forall x \in \mathbb{R} : (ab)^x = a^x b^x$

4. 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: $\log_a : \mathbb{R}^+ \to \mathbb{R}$, $x \mapsto \log_a x$.

5. $\forall a \in \mathbb{R}^+ \setminus \{1\} \forall x \in \mathbb{R}^+ : \log_a x = \dfrac{\ln x}{\ln a}$.

Definition

1. *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.*
2. *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:*

1. $\lim\limits_{n\to\infty} z_n = z \iff \lim\limits_{n\to\infty} x_n = x \wedge \lim\limits_{n\to\infty} y_n = y$
2. $(z_n)_{n\in\mathbb{N}}$ *ist Cauchy-Folge* $\iff$ $(x_n)_{n\in\mathbb{N}}$ *und* $(y_n)_{n\in\mathbb{N}}$ *sind Cauchy-Folgen.*

Satz 13.17 *Für komplexwertige Folgen gelten:*

1. *die Aussagen aus Satz 13.1.*
2. *die Aussagen aus Satz 13.2 (i)-(v), wobei $\alpha, \beta \in \mathbb{C}$ sein darf.*
3. $\lim\limits_{n\to\infty} z_n = z \implies \lim\limits_{n\to\infty} \overline{z_n} = \overline{z}$.
4. $\left((\exists C > 0 \forall n \in \mathbb{N} : |z_n| \leq C) \wedge \lim\limits_{n\to\infty} z_n = z\right) \implies |z| \leq C$.
5. *die Aussage aus 13.8.*
6. *die Aussage aus 13.9 für $\mathbb{C}$ statt $\mathbb{R}$.*
7. *die Aussagen aus 13.10.*
8. *die Aussage aus 13.11 für $\mathbb{C}$ statt $\mathbb{R}$.*

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* $s_n = \sum_{k=0}^{n} a_k$ *eine reelle (unendliche) Reihe.*

Kurzschreibweise: $\sum_{k=0}^{\infty} a_k$ *statt* $\left(\sum_{k=0}^{n} a_k\right)_{n\in\mathbb{N}_0}$

Die Folgenglieder s_n *heißen Partialsummen der Reihe.*

Konvergiert die Reihe $\sum_{k=0}^{\infty} a_k$ *gegen* $a \in \mathbb{R}$, *dann schreibt man* $\sum_{k=0}^{\infty} a_k = a$ *als Abkürzung für* $\lim_{n\to\infty} \sum_{k=0}^{n} a_k = a$.

Allgemeiner definiert man für beliebige $k_0 \in \mathbb{Z}$ *auch Reihen der Form*

$\sum_{k=k_0}^{\infty} a_k := \sum_{m=0}^{\infty} a_{m+k_0}$.

Satz 14.1 *Seien* $\sum_{k=0}^{\infty} a_k$, $\sum_{k=0}^{\infty} b_k$ *konvergente Reihen mit* $\sum_{k=0}^{\infty} a_k = a$ *und* $\sum_{k=0}^{\infty} b_k = b$*. Dann gilt:* $\forall \alpha, \beta \in \mathbb{R}$*:* $\sum_{k=0}^{\infty} \alpha a_k + \beta b_k = \alpha a + \beta b$

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* $\sum_{k=0}^{\infty} a_k$ *genau dann konvergent wenn* $\sum_{k=0}^{\infty} b_k$ *konvergiert. Die Grenzwerte der beiden Reihen können allerdings verschieden sein.*

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* $\sum_{k=0}^{\infty} (-1)^k a_k$ *und für deren Grenzwert* a *gilt:*

$$\forall n \in \mathbb{N}_0 : \sum_{k=0}^{2n+1} (-1)^k a_k \leq a \leq \sum_{k=0}^{2n} (-1)^k a_k \tag{3}$$

und

$$\forall n \in \mathbb{N}_0 : |a - \sum_{k=0}^{n} (-1)^k a_k| \leq a_{n+1}. \tag{4}$$

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) $\sum\limits_{k=0}^{\infty} a_k$ *ist konvergent* $\Leftrightarrow \forall \varepsilon > 0\ \exists n_\varepsilon\ \forall m > n > n_\varepsilon :$ $\left|\sum\limits_{k=n+1}^{m} a_k\right| < \varepsilon$

Korollar 14.5 $\sum\limits_{k=0}^{\infty} a_k$ *ist konvergent* $\Rightarrow (a_k)_{k\in\mathbb{N}_0}$ *ist eine Nullfolge*

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.*

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 *$\sum\limits_{k=0}^{\infty} a_k$ ist absolut konvergent. $\Leftrightarrow$ $\left(\sum\limits_{k=0}^{n} |a_k|\right)_{n\in\mathbb{N}_0}$ ist eine beschränkte Folge.*

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 $\sum\limits_{k=0}^{\infty} a_k$.*

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 $\sum\limits_{k=0}^{\infty} a_k$.*

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:*

$$|\sum_{k=0}^{\infty} a_k| \leq \sum_{k=0}^{\infty} |a_k| \leq \sum_{k=0}^{\infty} b_k$$

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.*

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 $\forall k \geq N : |\frac{a_{k+1}}{a_k}| \leq q$ gilt, dann ist $\sum\limits_{k=0}^{\infty} a_k$ absolut konvergent.

b) Falls ein $N > n_0$ existiert, so dass $\forall k \geq N : |\frac{a_{k+1}}{a_k}| \geq 1$ gilt, dann ist $\sum\limits_{k=0}^{\infty} a_k$ divergent.

Korollar 14.13 *a)* $\limsup_{k\to\infty} |\frac{a_{k+1}}{a_k}| \leq q < 1 \Rightarrow \sum\limits_{k=0}^{\infty} a_k$ *konvergiert absolut*

b) $\liminf_{k\to\infty} |\frac{a_{k+1}}{a_k}| > 1 \Rightarrow \sum\limits_{k=0}^{\infty} a_k$ *divergiert*

Bemerkung

Gilt weder die Voraussetzung von $14.12a)$ noch die Voraussetzung von $14.12b)$, z.B. im Falle $\lim\limits_{k\to\infty} |\frac{a_{k+1}}{a_k}| = 1$, 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 $a_k = \frac{1}{k+1}$.

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* $\forall k \geq N : \sqrt[k]{|a_k|} \geq 1$, *dann ist* $\sum\limits_{k=0}^{\infty} a_k$ *divergent.*

Korollar 14.15 *a)* $\limsup_{k\to\infty} \sqrt[k]{|a_k|} < 1 \Rightarrow \sum\limits_{k=0}^{\infty} a_k$ *konvergiert absolut.*

b) $\limsup_{k\to\infty} \sqrt[k]{|a_k|} > 1 \Rightarrow \sum\limits_{k=0}^{\infty} a_k$ *divergiert.*

Bemerkung

Gilt weder die Voraussetzung von $14.14a)$ noch die Voraussetzung von $14.14b)$, z.B. im Falle $\limsup_{k\to\infty} \sqrt[k]{|a_k|} = 1$, 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 $a_k = \frac{1}{k+1}$.

Satz 14.16 (Umordnungssatz) *Sei* $\sum_{k=0}^{\infty} a_k$ *absolut konvergent und* $\sigma : \mathbb{N}_0 \to \mathbb{N}_0$ *eine bijektive Abbildung.*
Dann ist die umgeordnete Reihe $\sum_{k=0}^{\infty} a_{\sigma(k)}$ *ebenfalls absolut konvergent und es gilt*

$$\sum_{k=0}^{\infty} a_{\sigma(k)} = \sum_{k=0}^{\infty} a_k$$

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.

Bemerkung
Im Allgemeinen gilt auch keine Verallgemeinerung des Assoziativgesetzes von endlich vielen auf unendlich viele Summanden. Denn es gilt

$\sum\limits_{n=0}^{\infty}(-1)^k = 1-1+1-1+\ldots$ ist divergent,

$\sum\limits_{n=0}^{\infty}((-1)^{2n}+(-1)^{2n+1}) = (1-1)+(1-1)+\ldots = 0,$

$1+\sum\limits_{n=1}^{\infty}((-1)^{2n-1}+(-1)^{2n}) = 1+(-1+1)+(-1+1)+\ldots = 1$

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*

$(\sum\limits_{k=0}^{\infty} a_k)(\sum\limits_{k=0}^{\infty} b_k) = \sum\limits_{n=0}^{\infty} c_n = a_0b_0 + (a_0b_1+a_1b_0) + (a_0b_2+a_1b_1+a_2b_0) + \ldots\,.$

Die Reihe $\sum\limits_{n=0}^{\infty} c_n$ *heißt das Cauchy-Produkt von* $\sum\limits_{k=0}^{\infty} a_k$ *und* $\sum\limits_{k=0}^{\infty} b_k$.

Definition

a) *Reihen der Form* $\sum\limits_{n=0}^{\infty} \frac{d_n}{10^n}$ *oder* $\sum\limits_{n=0}^{\infty} \frac{-d_n}{10^n}$ *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* $\pm \sum\limits_{n=0}^{\infty} \frac{d_n}{10^n}$.

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.

Bemerkung

Entsprechend kann man für $x \in \mathbb{R}$ auch eine g-adische Entwicklung

$$x = (x_N x_{N-1} \ldots x_0, x_{-1} x_{-2} \ldots)_g := \sum_{n=0}^{\infty} a_{N-n} g^{N-n}$$

mit $N \in \mathbb{N}$ und $x_j \in \{0, 1, \ldots, g-1\}$ 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* $\{n \in \mathbb{N} : n < m\}$ *gleichmächtig 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.*

Satz 14.19 $\mathbb{Q}$ *ist abzählbar unendlich.*

Satz 14.20 $\mathbb{R}$ *ist überabzählbar.*

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.

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 $lim_{n\to\infty} x_n = x$.*

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: $lim_{x\to x_0} f(x) = a$, falls $lim_{n\to\infty} f(x_n) = a$ für jede Folge $(x_n)_{n\in\mathbb{N}}$ mit $x_n \in D\backslash\{x_0\}$ für alle $n \in \mathbb{N}$ und $lim_{n\to\infty} x_n = x_0$.*

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: $lim_{x\to x_0^-} f(x) = a$, falls $lim_{n\to\infty} f(x_n) = a$ für jede Folge $(x_n)_{n\in\mathbb{N}}$ in $D\cap]-\infty, x_0[$ mit $lim_{n\to\infty} x_n = x_0$.*

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: $lim_{x\to x_0^+} f(x) = a$, falls $lim_{n\to\infty} f(x_n) = a$ für jede Folge $(x_n)_{n\in\mathbb{N}}$ in $D\cap]x_0, \infty[$ mit $lim_{n\to\infty} x_n = x_0$.*

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: $lim_{x\to\infty} f(x) = a$, falls $lim_{n\to\infty} f(x_n) = a$ für jede Folge $(x_n)_{n\in\mathbb{N}}$ in D mit $x_n \to \infty$ für $n \to \infty$.*

e) *Analog definiert man $lim_{x\to-\infty} f(x) = a$.*

Bemerkung
In manchen Büchern findet man eine leicht modifizierte Definition von $lim_{x\to x_0}f(x)=a$, bei der $lim_{n\to\infty}f(x_n)=a$ 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 $lim_{x\to 0}q(x)$ nicht existieren.

Bemerkung
Die Gesetze für konvergente Folgen aus Kapitel 13 übertragen sich sinngemäß auf Grenzwerte von Funktionen. Beispielsweise gilt:
Sei $f:D\to\mathbb{R}$, $g:D\to\mathbb{R}$, $lim_{x\to x_0}f(x)=a$ und $lim_{x\to x_0}g(x)=b$. 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 ($\varepsilon-\delta-$ 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) $lim_{x\to x_0}f(x)=a$

(ii) $\forall_{\varepsilon>0}\exists_{\delta>0}\forall_{x\in D\backslash\{x_0\}}:|x-x_0|<\delta\Rightarrow|f(x)-a|<\varepsilon$

Definition *Sei $D \subseteq \mathbb{R}$ und $f : D \to \mathbb{R}$ eine Funktion.*

a) *f heißt stetig in $x_0 \in D$, falls $lim_{n\to\infty} f(x_n) = f(x_0)$ für jede Folge $(x_n)_{n\in\mathbb{N}}$ in D mit $lim_{n\to\infty} x_n = x_0$.*

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 $lim_{x\to x_0} f(x) = f(x_0)$.*

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: $\exists_{\delta>0} \forall_{x\in D} : |x - x_0| < \delta \Rightarrow f(x) \neq 0$.*

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.*

Definition *Eine Funktion $f : D \to \mathbb{R}$ heißt stetig fortsetzbar in $x_0 \notin D$, falls $lim_{x \to x_0} f(x)$ existiert. In diesem Fall heißt*

$$\tilde{f} : D \cup \{x_0\} \to \mathbb{R}, \tilde{f}(x) = \begin{cases} f(x), & x \in D \\ lim_{x \to x_0} f(x), & x = x_0 \end{cases}$$

die stetige Fortsetzung von f auf $D \cup \{x_0\}$.

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 $[a,b]$.

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. $\exists_{x_1,x_2 \in I} \forall_{x \in I} : f(x_1) \leq f(x) \leq f(x_2)$.*

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.*

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.

Definition

a) *$z \in \mathbb{C}$ heißt Häufungspunkt von $M \subseteq \mathbb{C}$, falls eine Folge $(z_n)_{n\in\mathbb{N}}$ in $M \backslash \{z\}$ existiert mit $lim_{n\to\infty} z_n = z$.*

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: $lim_{z\to z_0} f(z) = a$, falls $lim_{n\to\infty} f(z_n) = a$ für jede Folge $(z_n)_{n\in\mathbb{N}}$ in $D \backslash \{z_0\}$ mit $lim_{n\to\infty} z_n = z_0$.*

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 $lim_{n\to\infty} f(z_n) = f(z_0)$ für jede Folge $(z_n)_{n\in\mathbb{N}}$ in D mit $lim_{n\to\infty} z_n = z_0$. 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.

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 $\frac{f(x)-f(x_0)}{x-x_0}$ 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 $\frac{y(t)-y(t_0)}{t-t_0}$ unabhängig von $t \neq t_0$.
   Beim freien Fall dagegen gilt $y(t) = y(t_0) - \frac{g}{2}(t-t_0)^2$ 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 $\frac{y(t)-y(t_0)}{t-t_0}$ beträgt, für $t \to t_0$ in die Momentangeschwidigkeit 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* $\frac{df}{dx}$ *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

$\frac{f(x)-f(x_0)}{x-x_0}$ heisst auch ein Differenzenquotient von f in x_0. Er beschreibt die Steigung der Sekante durch $(x_0, f(x_0))$ und $(x, f(x))$.

$\lim\limits_{x \to x_0} \frac{f(x)-f(x_0)}{x-x_0}$ heisst auch Differentialquotient von f in x. Wenn er existiert, dann beschreibt er die Steigung der Tangente an den Graphen von f in $(x_0, f(x_0))$. Die Gleichung der Tangente ist dann $y = f(x_0) + f'(x_0)(x - x_0)$.

Satz 16.1 *Ist f differenzierbar in x_0, dann ist f auch stetig in x_0.*

Bemerkung

Aus der Stetigkeit einer Funktion f in x_0 folgt im Allgemeinen aber 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* $\lim\limits_{x \to x_0} \frac{R(x,x_0)}{x-x_0} = 0$

Sind (i) bzw. (ii) erfüllt, dann ist $c = f'(x_0)$.

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 gilt*

$$(\alpha f + \beta g)'(x_0) = \alpha f'(x_0) + \beta g'(x_0)$$

Satz 16.4 (Produktregel) *Sind f und* g *in* x_0 *differenzierbar, dann auch* fg *mit*

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

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

$$\left(\frac{1}{g}\right)'(x_0) = -\frac{g'(x_0)}{(g(x_0))^2}$$

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*

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

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 $y \in f(I)$.

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 aufwändig durchzuführen.

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 $f^{(3)}$,
$\frac{d^n}{dx^n} f$ oder $\frac{d^n f}{dx^n}$ statt $f^{(n)}$.

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 $\frac{d}{dx} : f \to \frac{d}{dx}f$ 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.

Satz 16.8 (Leibniz-Regel) *Sind f, g n-mal differenzierbar in x_0, dann auch fg mit*

$$(f \cdot g)^{(n)}(x_0) = \sum_{k=0}^{n} \binom{n}{k} f^{(k)} g^{(n-k)}(x_0)$$

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 hat in x_0 ein lokales Extremum $\Rightarrow f'(x_0) = 0$.

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:*

$$\exists \xi \in]a,b[: f'(\xi) = 0$$

Satz 16.11 (Mittelwertsatz der Differentialrechnung) *Sei $f : [a,b] \to \mathbb{R}$ stetig auf $[a,b]$ und differenzierbar in $]a,b[$. Dann gilt:*

$$\exists \xi \in]a,b[: f'(\xi) = \frac{f(b) - f(a)}{b - a}$$

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 $x_0 \in \,]a,b[$.*

a) *Falls $\exists \varepsilon > 0\ \forall x \in \,]x_0 - \varepsilon, x_0[\, : f'(x) \geq 0 \ \wedge\ \forall x \in \,]x_0, x_0 + \varepsilon[\, : f'(x) \leq 0$, so hat f 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.

Falls $\exists \varepsilon > 0\ \forall x \in \,]x_0 - \varepsilon, x_0[\, : f'(x) \leq 0 \ \wedge\ \forall x \in \,]x_0, x_0 + \varepsilon[\, : f'(x) \geq 0$, so hat 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.*

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 x_0 und hat f' ein striktes lokales Extremum in x_0, dann heißt x_0 Wendepunkt von f.

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:*

$$\exists \xi \in\,]a,b[\,: \quad \frac{f(b)-f(a)}{g(b)-g(a)} = \frac{f'(\xi)}{g'(\xi)}.$$

Satz 16.16 (Regel von de l'Hospital) *Seien* $f, g :]a, b[\to \mathbb{R}$*, wobei* $-\infty \leq a < b \leq \infty$*, differenzierbar, es gelte* $\lim_{x \to a^+} f(x) = \lim_{x \to a^+} g(x) = 0$*,* $g(x) \neq 0 \wedge g'(x) \neq 0$ *für alle* $x \in]a, b[$ *und es existiere* $\lim_{x \to a^+} \frac{f'(x)}{g'(x)}$*. 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 $\lim_{x \to a^+}$ *durch* $\lim_{x \to b^-}$ *ersetzt.*

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.

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.

Definition
Sei $D \subseteq \mathbb{R}$, $f, g : D \to \mathbb{R}$ zwei Funktionen und x_0 ein Häufungspunkt von D. 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

$$\exists C, \delta > 0\ \forall x \in D \cap]x_0 - \delta, x_0 + \delta[: \quad |f(x)| \leq C|g(x)|.$$

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:*

$$f(x) - T_n(f, x, x_0) = o((x - x_0)^n) \qquad \text{für } x \to x_0,$$
$$f(x) - T_n(f, x, x_0) = O((x - x_0)^{n+1}) \quad \text{für } x \to x_0.$$

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.*

$$\lim_{n\to\infty} x_n = x_* \ \wedge\ \exists C > 0\ \forall n \in \mathbb{N} : |x_{n+1} - x_*| \leq C|x_n - x_*|^2.$$

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.*

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* $\lim_{n\to\infty} R_n(f,x,x_0) = 0$ *gilt.*

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

$$\lim_{n\to\infty} \sup_{x\in D} |f_n(x) - f(x)| = 0.$$

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)

$$\left[\left(\forall n \in \mathbb{N}_0\ \exists c_n \in \mathbb{R} : \sup_{x\in D} |f_n(x)| \le c_n\right) \ \wedge\ \sum_{n=0}^{\infty} c_n \text{ ist konvergent}\right]$$
$$\Rightarrow \exists f,g : D \to \mathbb{R} : \sum_{n=0}^{\infty} f_n \text{ konvergiert gleichmäßig auf } D \text{ gegen } f \ \wedge$$
$$\sum_{n=0}^{\infty} |f_n| \text{ konvergiert gleichmäßig auf } D \text{ gegen } g.$$

Satz 17.7 (Konvergenzverhalten von Potenzreihen) *Sei $\sum_{n=0}^{\infty} a_n(x-x_0)^n$ 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 $\sum_{n=0}^{\infty} a_n(x-x_0)^n$.

b) *Für den Konvergenzradius R gilt die Formel von Cauchy-Hadamard, nämlich*

$$R = \begin{cases} \dfrac{1}{\limsup_{n\to\infty} \sqrt[n]{|a_n|}}, & \text{falls } \limsup_{n\to\infty} \sqrt[n]{|a_n|} \text{ existiert und } \neq 0 \text{ ist,} \\ \infty, & \text{falls } \limsup_{n\to\infty} \sqrt[n]{|a_n|} = 0 \text{ ist,} \\ 0, & \text{falls } \limsup_{n\to\infty} \sqrt[n]{|a_n|} \text{ nicht existiert.} \end{cases}$$

c) *Falls $\lim_{n\to\infty} \frac{1}{\sqrt[n]{|a_n|}}$, $\lim_{n\to\infty} \frac{|a_n|}{|a_{n+1}|}$ oder $\limsup_{n\to\infty} \frac{|a_n|}{|a_{n+1}|}$ existieren, dann gilt $R = \lim_{n\to\infty} \frac{1}{\sqrt[n]{|a_n|}}$, bzw. $R = \lim_{n\to\infty} \frac{|a_n|}{|a_{n+1}|}$, bzw. $R = \limsup_{n\to\infty} \frac{|a_n|}{|a_{n+1}|}$. Falls $\limsup_{n\to\infty} \frac{|a_n|}{|a_{n+1}|}$ nicht existiert, dann gilt $R = \infty$.*

d) *Falls $R > 0$ und $0 < \rho < R$ ist, dann konvergiert $\sum_{n=0}^{\infty} a_n(x-x_0)^n$ gleichmäßig auf $\{x \in \mathbb{R} : |x-x_0| \leq \rho\}$.*

Bemerkungen

1) Ist der Konvergenzradius R einer Potenzreihe $\sum_{n=0}^{\infty} a_n(x-x_0)^n$ 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.

Satz 17.8 (Identitätssatz) *Seien* $f(x) = \sum_{n=0}^{\infty} a_n(x-x_0)^n$ *und* $g(x) = \sum_{n=0}^{\infty} b_n(x-x_0)^n$ *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* $\lim_{m\to\infty} x_m = x_0$ *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* $f = g$*.*

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.

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* $|x-x_0| < \min\{R_f, R_g\}$*:*

$$f(x)g(x) = \sum_{n=0}^{\infty} c_n(x-x_0)^n \quad \textit{mit } c_n = \sum_{k=0}^{n} a_k b_{n-k}.$$

Bemerkung

Die Sätze 17.6 – 17.8, Korollar 17.9 und Satz 17.10 übertragen sich sinngemäß auf komplexe Funktionenfolgen bzw. komplexe Potenzreihen.

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* $e^z := \sum_{n=0}^{\infty} \frac{1}{n!} z^n$ *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) $\forall y \in \mathbb{R} : \cos y = \frac{e^{iy}+e^{-iy}}{2} \ \wedge \ \sin y = \frac{e^{iy}-e^{-iy}}{2i}$.

(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* $\frac{d}{dz}e^z = e^z$ *für alle* $z \in \mathbb{C}$ *ist.*

(viii) $\forall z \in \mathbb{C} : e^z = \lim_{n\to\infty} \left(1 + \frac{z}{n}\right)^n = \lim_{n\to\infty} \left(1 - \frac{z}{n}\right)^{-n}$.

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) $\sin z = 0 \iff z \in \{k\pi : k \in \mathbb{Z}\}$,

$\cos z = 0 \iff z \in \left\{\left(\frac{1}{2} + k\right)\pi : k \in \mathbb{Z}\right\}$.

(iv) $\forall z \in \mathbb{C} : \sin(z + 2\pi) = \sin z$,

$\forall z \in \mathbb{C} : \cos(z + 2\pi) = \cos z$,

$\forall z \in \mathbb{C} : \sin(z + \pi) = -\sin z$,

$\forall z \in \mathbb{C} : \cos(z + \pi) = -\cos z$,

$\forall z \in \mathbb{C} : \sin(z + \frac{\pi}{2}) = \cos z$,

$\forall z \in \mathbb{C} : \cos(z + \frac{\pi}{2}) = -\sin z$,

$\forall z \in \mathbb{C} \setminus \left\{\left(\frac{1}{2} + k\right)\pi : k \in \mathbb{Z}\right\} : \tan(z + \pi) = \tan z$.

(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) $\forall z \in \mathbb{C} : \frac{d}{dz} \sin z = \cos z,$

$$\forall z \in \mathbb{C} : \tfrac{d}{dz} \cos z = -\sin z,$$

$$\forall z \in \mathbb{C} \setminus \left\{ \left(\tfrac{1}{2} + k\right)\pi : k \in \mathbb{Z} \right\} : \tfrac{d}{dz} \tan z = \frac{1}{\cos^2 z},$$

$$\forall z \in \mathbb{C} \setminus \{ k\pi : k \in \mathbb{Z} \} : \tfrac{d}{dz} \cot z = -\frac{1}{\sin^2 z}.$$

(xi) $\forall z \in \mathbb{C} : \sin z = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} z^{2n+1},$

$$\forall z \in \mathbb{C} : \cos z = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} z^{2n}.$$

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) $\forall z \in \mathbb{C} : \frac{d}{dz} \sinh z = \cosh z$,
      $\forall z \in \mathbb{C} : \frac{d}{dz} \cosh z = \sinh z$.

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.*

Satz 18.1 *a)* $f : [a,b] \to \mathbb{R}$ *ist eine Treppenfunktion*

$\Leftrightarrow$ *Es existieren endlich viele Intervalle* $I_k \subseteq [a,b],\ k = 1,\ldots,n$ *und Konstanten* $c_k \in \mathbb{R}$, *so dass* $f = \sum_{k=1}^{n} c_k \chi_{I_k}$ *fast überall gilt.*

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* $\tau([a,b])$ *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

$$\int_b^a f(x)dx \; := \; -\int_a^b f(x)dx,$$

$$\forall a \leq c \leq d \leq b : \qquad \int_c^d f(x)dx \; := \; \int_a^b (f \cdot \chi_{[c,d]})(x)dx.$$

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:

$$f = g \quad \text{f. ü.} \;\Rightarrow\; \int_a^b f(x)dx = \int_a^b g(x)dx.$$

Satz 18.2 *Seien f,g Treppenfunktionen auf $[a,b]$. Dann gilt*

$$a)\ \forall \alpha,\beta \in \mathbb{R}: \qquad \int_a^b (\alpha f + \beta g)(x)dx = \alpha \int_a^b f(x)dx + \beta \int_a^b g(x)dx$$

$$b)\ \forall x \in [a,b]: \qquad f(x) \leq g(x) \ \Rightarrow \int_a^b f(x)dx \leq \int_a^b g(x)dx$$

Definition *Sei $f : [a,b] \to \mathbb{R}$ eine beschränkte Funktion.*

a)

$$\underline{\int_a^b} f(x)dx := sup\{ \int_a^b \varphi(x)dx : \ \varphi \in \tau([a,b]) \wedge \forall x \in [a,b] : \varphi(x) \leq f(x)\}$$

heißt Unterintegral von f über $[a,b]$ und

$$\overline{\int_a^b} f(x)dx := inf\{ \int_a^b \varphi(x)dx : \ \varphi \in \tau([a,b]) \wedge \forall x \in [a,b] : \varphi(x) \geq f(x)\}$$

heißt Oberintegral von f über $[a,b]$.

b) f heißt Riemann-integrierbar über $[a,b]$, falls

$$\int_{\underline{a}}^{b} f(x)dx = \int_{a}^{\overline{b}} f(x)dx.$$

In diesem Fall heißt

$$\int_{a}^{b} f(x)dx = \int_{\underline{a}}^{b} f(x)dx = \int_{a}^{\overline{b}} f(x)dx$$

das (bestimmte) Riemann-Integral von f über $[a,b]$. Außerdem sei

$$\int_{b}^{a} f(x)dx \;:=\; -\int_{a}^{b} f(x)dx$$
$$\int_{a}^{a} f(x)dx \;:=\; 0.$$

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 $k = 1, \ldots, n$, $\mathscr{Z} := ((x_0, x_1, \ldots, x_n), (\xi_1, \ldots, \xi_n))$, $\mu(\mathscr{Z}) := \max_{1 \le k \le n}(x_k - x_{k-1})$ und $f : [a,b] \to \mathbb{R}$ eine beschränkte Funktion. Dann heißt*

$$R(f, \mathscr{Z}) := \sum_{k=1}^{n} f(\xi_k)(x_k - x_{k-1})$$

Riemannsche Summe von f über $[a,b]$ bezüglich $\mathscr{Z}$ mit Feinheit $\mu(\mathscr{Z})$.

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 $\mathscr{Z}$ mit $\mu(\mathscr{Z}) < \delta$ gilt*

$$\left| \int_a^b f(x)dx - R(f, \mathscr{Z}) \right| < \varepsilon.$$

Kurzschreibweise:

$$\lim_{\mu(\mathscr{Z}) \to 0} R(f, \mathscr{Z}) = \int_a^b f(x)dx.$$

Satz 18.4 *Eine Funktion $f : [a,b] \to \mathbb{R}$ ist genau dann Riemann-integrierbar, wenn zu jedem $\varepsilon > 0$ Treppenfunktionen $\varphi, \psi \in \tau([a,b])$ 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.*

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.*

$$\forall\varepsilon>0\ \exists\delta>0\ \forall x_0,x\in[a,b]:\qquad |\,x-x_0\,|<\delta\ \Rightarrow |\,f(x)-f(x_0)\,|<\varepsilon.$$

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*

$$(i)\quad \forall x\in[a,b]:\qquad \varphi(x)\le f(x)\le\psi(x)$$
$$(ii)\quad \forall x\in[a,b]:\qquad |\,\psi(x)-\varphi(x)\,|<\varepsilon$$

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

$$\int_a^b (\alpha f + \beta g)(x)dx = \alpha \int_a^b f(x)dx + \beta \int_a^b g(x)dx \quad (Linearit\ddot{a}t\ des\ Integrals)$$

(ii) $f \leq g \;\Rightarrow\; \int_a^b f(x)dx \leq \int_a^b g(x)dx \qquad (Monotonie\ des\ Integrals)$

(iii) Die Funktionen $f_+ : [a,b] \to \mathbb{R}$ mit $f_+(x) = \begin{cases} f(x), & x > 0 \\ 0, & sonst \end{cases}$,

$f_- : [a,b] \to \mathbb{R}$ mit $f_-(x) = \begin{cases} -f(x), & x < 0 \\ 0, & sonst \end{cases}$

und $|f| : [a,b] \to \mathbb{R}$ sind ebenfalls integrierbar und es gilt

$$\left| \int_a^b f(x)\, dx \right| \leq \int_a^b |f(x)|\, dx \leq (b-a) \sup_{x \in [a,b]} |f(x)|$$

(iv) Die Funktion fg ist ebenfalls integrierbar.

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 $g \equiv 1$

$$\exists \xi \in [a,b] : \qquad \int_a^b f(x)dx = f(\xi)(b-a).$$

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 $\lim_{n\to\infty} \int_a^b t_n(x)dx$ 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

$$f : [0,1] \to \mathbb{R}, \qquad x \to \begin{cases} 1, & \text{für } x = \frac{1}{n},\ n \in \mathbb{N} \\ 0, & sonst \end{cases}$$

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- u. Integralrechnung) *Sei $f \in C^0([a,b])$.*

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

$$\int_a^b f(x)dx \;=\; F(b) - F(a)$$

Kurzschreibweise: $$\int_a^b f(x)dx \;=\; F(x)\Big|_a^b$$

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.

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) $\int x dx = \frac{1}{2}x^2 + c, \quad c \in \mathbb{R}$,
$\int_1^2 x dx = \frac{1}{2}x^2\Big|_1^2 = 2 - \frac{1}{2} = \frac{3}{2}$

2) $\forall r \in \mathbb{R} \setminus \{-1\} : \ \int x^r dx = \frac{1}{r+1}x^{r+1} + c, \quad c \in \mathbb{R}$

3) $\forall a, b > 0 : \ \int_a^b \frac{1}{x} dx = ln\ x\Big|_a^b$
$\forall a, b < 0 : \ \int_a^b \frac{1}{x} dx = ln(-x)\Big|_a^b$
$\Rightarrow \ \int \frac{1}{x} dx = ln \mid x \mid + c, \quad c \in \mathbb{R}$
für $x \neq 0$, d.h. 0 liegt nicht im Integrationsintervall

4) $\int e^x dx = e^x + c, \quad c \in \mathbb{R}$

5) $\int sin\ x\ dx = -cos\ x + c, \quad c \in \mathbb{R}$

6) $\int cos\ x\ dx = sin\ x + c, \quad c \in \mathbb{R}$

7) $\int \frac{1}{x^2+1} dx = arctan\ x + c, \quad c \in \mathbb{R}$

8) $\int \frac{1}{\sqrt{1-x^2}} dx = arcsin\ x + c, \quad c \in \mathbb{R}$ für $\mid x \mid < 1$

9) $\int \frac{1}{cos^2 x} dx = tan\ x + c, \quad c \in \mathbb{R}$ für $cos\ x \neq 0$

10) $\int cosh\ x\ dx = sinh\ x + c, \quad c \in \mathbb{R}$

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

$$F:[0,2]\to\mathbb{R}\quad f(x)=\begin{cases}1, & x\in[0,1[\\ 2, & x\in[1,2]\end{cases}$$

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*

$$\int_a^b f(x)g'(x)dx = f(x)g(x)\mid_a^b-\int_a^b f'(x)g(x)dx$$

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*

$$\int_a^b f(y(x))\frac{dy}{dx}(x)\;dx = \int_{y(a)}^{y(b)} f(y)\;dy.$$

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*

$$\int_a^b f(x)\;dx = \int_{t(a)}^{t(b)} f(x(t))\frac{dx}{dt}(t)\;dt = \int_\alpha^\beta f(x(t))\,|\frac{dx}{dt}(t)|\;dt.$$

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*

$$\int_a^b f(x)\;dx = \int_{t(a)}^{t(b)} f(x(t))\frac{1}{\frac{dt}{dx}(x(t))}\;dt.$$

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

$$Q(z) = (z-\lambda)^n Q_1(z), \qquad \text{wobei}\, Q_1(\lambda) \neq 0.$$

Sei $a_1 = \dfrac{P(\lambda)}{Q_1(\lambda)}$.

Dann ist λ Nullstelle von $P - a_1Q_1$, also gibt es ein eindeutig bestimmtes Polynom P_1 mit

$$P(z) = a_1Q_1(z) + (z-\lambda)P_1(z).$$

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

$$Q(x) = a\prod_{j=1}^{m}(x-\lambda_j)^{n_j}\prod_{l=1}^{r}(x^2+b_lx+c_l)^{p_l}$$

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

$$\begin{aligned}\frac{P(x)}{Q(x)} \;=\; &\sum_{j=1}^{m}\Big(\frac{A_{1j}}{x-\lambda_j}+\frac{A_{2j}}{(x-\lambda_j)^2}+\ldots+\frac{A_{n_jj}}{(x-\lambda_j)^{n_j}}\Big)\\ &+\sum_{l=1}^{r}\Big(\frac{B_{1l}x+C_{1l}}{x^2+b_lx+c_l}+\frac{B_{2l}x+C_{2l}}{(x^2+b_lx+c_l)^2}+\ldots+\frac{B_{p_ll}x+C_{p_ll}}{(x^2+b_lx+c_l)^{p_l}}\Big) \qquad (*)\end{aligned}$$

gilt.

3. Schritt:

Bestimme die Zerlegung von $\tilde{q}$ in reelle Linearfaktoren und reellle 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

$$\frac{P(x)}{\frac{Q(x)}{(x-\lambda_s)^{n_s}}} = A_{n_s s} + (x-\lambda_s)(\ldots).$$

Setzt man in diese Gleichung λ_s für x ein, dann erhält man aus der dadurch entstehenden Gleichung den Wert von $A_{n_s s}$.
Nachdem man auf diese Weise die Koeffizienten $A_{n_j j}$ 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 h sowie die einzelnen Summanden der Partialbruchzerlegungen $\frac{\tilde{r}}{\tilde{q}}$. Dabei sind Integrale der folgenden Form zu berechnen:

$$\int \frac{A}{(x-\lambda)^n}\,dx, \qquad n \in \mathbb{N}$$

$$\int \frac{Bx+G}{(x^2+bx+c)^n}\,dx, \qquad n \in \mathbb{N}$$

Es gilt:

$$\int \frac{A}{(x-\lambda)^n}\,dx = \begin{cases} A\ln|x-\lambda|+C, & n=1 \\ \frac{A}{1-n}\frac{1}{(x-\lambda)^{n-1}}+C, & n>1 \end{cases}, \quad C \in \mathbb{R}$$

$$\int \frac{Bx+G}{(x^2+bx+c)^n}\,dx = \frac{B}{2}\int \frac{2x+b}{(x^2+bx+c)^n}\,dx + \left(G-\frac{Bb}{2}\right)\int \frac{1}{(x^2+bx+c)^n}\,dx$$

$$\int \frac{2x+b}{(x^2+bx+c)^n}\,dx = \begin{cases} \ln|x^2+bx+c|+C, & n=1 \\ \frac{1}{1-n}\frac{1}{(x^2+bx+c)^{n-1}}+C, & n>1 \end{cases}, \quad C \in \mathbb{R}$$

Es bleibt nun noch die Berechnung von

$$\int \frac{1}{(x^2+bx+c)^n}\,dx, \qquad n \in \mathbb{N},$$

wobei x^2+bx+c keine reellen Nullstellen hat und somit $D := b^2-4c < 0$ ist.
Durch quadratische Ergänzung erhält man

$$x^2+bx+c = \left(x+\frac{b}{2}\right)^2 + c - \frac{b^2}{4} = K^{-2}\left(1+\left(K\left(x+\frac{b}{2}\right)\right)^2\right)$$

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

$$I_n := \int \frac{1}{(1+t^2)^n}\,dt, \qquad n \in \mathbb{N}.$$

Dann gilt

$$I_1 = \arctan t + C, \quad C \in \mathbb{R}.$$

Und für alle $n > 1$:

$$\begin{aligned} I_n &= I_{n-1} - \int \frac{t^2}{(1+t^2)^n}\,dt = I_{n-1} - \int \frac{t}{2} \cdot \frac{2t}{(1+t^2)^n}\,dt \\ &= I_{n-1} + \frac{1}{2(n-1)} \cdot \frac{t}{(1+t^2)^{n-1}} + C - \frac{1}{2(n-1)} I_{n-1} \\ &= \frac{2n-3}{2n-2} I_{n-1} + \frac{1}{2n-2} \cdot \frac{t}{(1+t^2)^{n-1}} + C, \quad C \in \mathbb{R} \end{aligned}$$

Somit ergibt sich für alle $b, c \in \mathbb{R}$ mit $b^2 < 4c$:

$$\int \frac{1}{x^2+bx+c}\,dx = \frac{1}{\sqrt{c - \frac{b^2}{4}}} \arctan\left(\frac{x + \frac{b}{2}}{\sqrt{c - \frac{b^2}{4}}} \right) + C, \quad C \in \mathbb{R}$$

sowie

$$\begin{aligned} \int \frac{1}{(x^2+bx+c)^n}\,dx &= \frac{1}{(n-1)(4c-b^2)} \cdot \frac{2x+b}{(x^2+bx+c)^{n-1}} \\ &\quad + \frac{2(2n-3)}{(n-1)(4c-b^2)} \int \frac{1}{(x^2+bx+c)^{n-1}}\,dx, \end{aligned}$$

falls $n \in \mathbb{N} \setminus \{1\}$ ist.

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
$\int_a^\infty f(x)dx := \lim_{S\to\infty} \int_a^S f(x)dx$
bzw. $\int_{-\infty}^b f(x)dx := \lim_{S\to-\infty} \int_S^b f(x)dx$
bzw. $\int_{-\infty}^\infty f(x)dx := \int_{-\infty}^c f(x)dx + \int_c^\infty f(x)dx$
bzw. $\int_a^b f(x)dx := \lim_{S\to a^+} \int_S^b f(x)dx$
bzw. $\int_a^b f(x)dx := \lim_{S\to b^-} \int_a^S f(x)dx$
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. $\int_a^b f(x)dx$ ist konvergent (konvergiert). Anderenfalls sagt man, das jeweilige uneigentliche Integral ist divergent (divergiert).

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 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 $|\int_c^\infty f(x)dx| \leq \int_c^\infty |f(x)|dx \leq \int_c^\infty g(x)dx$.*

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.

Satz 18.18 (Integralkriterium für Reihen) *Sei* $f : [1, \infty[\to \mathbb{R}_0^+$ *monoton fallend. Dann gilt*

$$\sum_{n=1}^{\infty} f(n) \text{ konvergiert} \iff \int_1^{\infty} f(x)\,dx \text{ konvergiert}.$$

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 $\mathbb{R}$.

Vertauschbarkeit von Grenzübergängen, Differentiation 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*:*

$$\lim_{x\to x_0}\left(\lim_{n\to\infty} f_n(x)\right) = \lim_{n\to\infty}\left(\lim_{x\to x_0} f_n(x)\right) = f(x_0).$$

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* $f:[a,b]\to\mathbb{R}$. *Dann gilt*

$$\int_a^b f(x)\,dx = \lim_{n\to\infty}\int_a^b f_n(x)\,dx.$$

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

$$f_n:[0,1]\to\mathbb{R},\ x\mapsto \begin{cases} n^2x, & 0\le x\le \frac{1}{n}\\ 2n-n^2x, & \frac{1}{n}<x\le\frac{2}{n}\\ 0, & sonst\end{cases}$$

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*

$$f'(x) = g(x) = \lim_{n\to\infty} f_n'(x) \quad \textit{für alle } x \in \,]a,b[.$$

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

$$f_n : [-1,1] \to \mathbb{R}, x \mapsto \begin{cases} |x|, & x \in [-1, -\frac{1}{n+1}] \cup [\frac{1}{1+n}, 1] \\ \frac{n+1}{2}x^2 + \frac{1}{2(n+1)}, & x \in \,]-\frac{1}{n+1}, \frac{1}{n+1}[\end{cases}$$

   Dann gilt $f_n \in C^1[-1,1]$ und $(f_n)_{n\in\mathbb{N}_0}$ konvergiert gleimäßig 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 konvergiert gegen

$$g : [0,1] \to \mathbb{R}, x \mapsto \begin{cases} 0, & x \neq 1 \\ 1, & x = 1. \end{cases}$$

Satz 18.22 *Sei* $f(x) = \sum\limits_{n=0}^{\infty} a_n(x-x_0)^n$ *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 $\sum\limits_{n=0}^{\infty} \frac{a_n}{n+1}(x-x_0)^{n+1}$ *hat auch Konvergenzradius* R.

(iii) f *kann in* $B_R(x_0)$ *gliedweise abgeleitet werden, d.h.*

$$\forall x \in B_R(x_0) : \frac{d}{dx} \sum_{n=0}^{\infty} a_n(x-x_0)^n = \sum_{n=1}^{\infty} n a_n (x-x_0)^{n-1}.$$

Die gliedweise abgeleitete Potenzreihe $\sum\limits_{n=1}^{\infty} n a_n(x-x_0)^{n-1}$ *hat auch Konvergenzradius* R.

(iv) $f \in C^\infty(B_R(x_0))$, *wobei*

$\forall k \in \mathbb{N} \forall x \in B_R(x_0) : \frac{d^k}{dx^k} f(x) = \sum_{n=k}^{\infty} n \cdot (n-1) \cdot \ldots \cdot (n-k+1) \cdot a_n(x-x_0)^{n-k}$.

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.

Bemerkung

Für Funktionenreihen, die keine Potenzreihen sind, müssen die Aussagen von Satz 18.22 nicht unbedingt gelten. Denn beispielsweise hat die Funktionenreihe $\sum_{n=1}^{\infty} \frac{1}{n^2} sin(n^2 x)$ die konvergente Majorante $\sum_{n=1}^{\infty} \frac{1}{n^2}$ und konvergiert somit für alle $x \in \mathbb{R}$. Die gliedweise abgeleitete Reihe $\sum_{n=1}^{\infty} cos(n^2 x)$ 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: $\lim_{n\to\infty} cos(n^2 x) = 0$.

$$\Rightarrow \forall k \in \mathbb{N} : \lim_{n\to\infty} cos((kn)^2 x) = 0.$$

$$\begin{aligned}
\Rightarrow 0 &= \lim_{n\to\infty} cos((5n)^2 x) \\
&= \lim_{n\to\infty} cos((3n)^2 x + (4n)^2 x) \\
&= \lim_{n\to\infty} cos((3n)^2 x)\, cos((4n)^2 x) - sin((3n)^2 x)\, sin((4n)^2 x)) \\
&= \lim_{n\to\infty} cos((3n)^2 x)\, cos(4n)^2 x) - \sqrt{1-(cos((3n)^2 x))^2}\sqrt{1-(cos((4n)^2 x))^2} \\
&= -1 \qquad ↯
\end{aligned}$$

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) $\lim\limits_{k\to\infty} \|a_k - a\|_\infty = \lim\limits_{k\to\infty} \max\limits_{i=1,\dots,n} |\,(a_k - a)_i\,| = 0$

(iii) $\forall i \in \{1,\dots,n\} : \lim\limits_{k\to\infty} (a_k)_i = a_i$

Außerdem gilt

$$(a_k)_{k\in\mathbb{N}} \,\text{ist Cauchy} - \text{Folge in}\, \mathbb{R}^n \;\Leftrightarrow\; \begin{array}{l} ((a_k)_i)_{k\in\mathbb{N}} \,\text{sind Cauchy} - \text{Folgen in}\, \mathbb{R} \\ \text{für alle}\; i \in \{1,\dots,n\} \end{array}$$

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

$$\left((\exists C > 0\ \forall k \in \mathbb{N} : |a_k| \leq C) \wedge \lim_{k\to\infty} a_k = a \right) \Rightarrow |a| \leq C.$$

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 $\lim_{k\to\infty} f(a_k) = f(a_0)$ für jede Folge $(a_k)_{k\in\mathbb{N}}$ in D mit $\lim_{k\to\infty} a_k = a_0$.

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.

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

$$f : \mathbb{R}^2 \to \mathbb{R}\,,\ f(x,y) = \begin{cases} \frac{xy^2}{x^2+y^4} & , (x,y) \neq (0,0) \\ 0 & , (x,y) = (0,0) \end{cases}$$

Dann gilt $\forall c \in \mathbb{R} : \lim\limits_{x\to 0} f(x,cx) = \lim\limits_{x\to 0} \frac{c^2x^3}{x^2+c^4x^4} = 0 = f(0,0)$

sowie $\lim\limits_{y\to 0} f(0,y) = \lim\limits_{y\to 0} 0 = 0 = f(0,0)$, sodass f in $(0,0)$ längs jeder Geraden stetig ist.

Aber es gilt $\lim\limits_{x\to 0} f(x,\sqrt{x}) = \lim\limits_{x\to 0} \frac{x^2}{2x^2} = \frac{1}{2} \neq f(0,0)$, sodass $\lim\limits_{k\to\infty} \left(\frac{1}{k}, \frac{1}{\sqrt{k}}\right) = (0,0)$ und $\lim\limits_{k\to\infty} f\left(\frac{1}{k}, \frac{1}{\sqrt{k}}\right) = \frac{1}{2} \neq f(0,0)$ ist und daher f nicht stetig in $(0,0)$ ist.

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) $f'(x_0) := \lim\limits_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim\limits_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}$ *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 $\frac{df}{dx}$ 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.

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.

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}, $\frac{\partial f}{\partial x_i}$, $\frac{\partial}{\partial x_i} f$ 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.

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.

Definition *Sei $f : \mathbb{R}^n \supseteq D \to \mathbb{R}$ eine Funktion und x ein innerer Punkt von D. Dann heißen*

$$\partial_{x_i}^2 f(x) = f_{x_i x_i}(x) = \frac{\partial^2 f}{\partial x_i^2}(x) = \frac{\partial^2}{\partial x_i^2} f(x) = \partial_i^2 f(x) \;:=\; \partial_{x_i}\left(\partial_{x_i} f\right)(x)\,, \quad i \in \{1, \ldots, n\}\,,$$

$$\begin{aligned} \partial_{x_i}\partial_{x_j} f(x) = f_{x_i x_j}(x) = \frac{\partial^2 f}{\partial x_i \partial x_j}(x) = \frac{\partial^2}{\partial x_i \partial x_j} f(x) \;&=\; \partial_i \partial_j f(x) \\ &:=\; \partial_{x_i}\left(\partial_{x_j} f\right)(x)\,, i, j \in \{1, \ldots, n\} \end{aligned}$$

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

$$\frac{\partial^k}{\partial x_{i_k} \partial x_{i_{k-1}} \ldots \partial x_{i_1}} f = \frac{\partial}{\partial x_{i_k}} \left(\frac{\partial^{k-1} f}{\partial x_{i_{k-1}} \ldots \partial x_{i_1}} \right)$$

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 $\partial_v f$ sowie die partiellen Ableitungen k-ter Ordnung von f.

Definition *Sei* $f : \mathbb{R}^n \supseteq D \to \mathbb{R}^m$, $(x_1,\dots,x_n) \mapsto \begin{pmatrix} f_1(x_1,\dots,x_n) \\ \vdots \\ f_m(x_1,\dots,x_n) \end{pmatrix}$.

a) Sei x_0 *ein innerer Punkt von* D. f *heißt (total) differenzierbar in* x_0*, falls gilt:*
Es existiert eine Matrix $A \in \mathbb{R}^{m\times n}$*, sodass*

$$\lim_{x\to x_0}\Big(\frac{f(x)-f(x_0)}{|x-x_0|} - A\frac{x-x_0}{|x-x_0|}\Big) = 0$$

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.*

Satz 20.2 *Sei $f : \mathbb{R}^n \supseteq D \to \mathbb{R}^m$ und $x_0 \in D$. Dann ist äquivalent:*

1. *f ist (total) differenzierbar in x_0.*
2. *$\exists A \in \mathbb{R}^{m\times n} \forall x \in D : \ f(x) = f(x_0) + A(x - x_0) + R(x, x_0)$ mit $\lim\limits_{x\to x_0} \frac{R(x,x_0)}{|x-x_0|} = 0$.*

Sind (i) bzw. (ii) erfüllt, dann ist $A = f'(x_0)$.

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*

1. $f'(x_0) = \begin{pmatrix} \partial_{x_1} f_1(x_0) & \cdots & \partial_{x_n} f_1(x_0) \\ \vdots & & \vdots \\ \partial_{x_1} f_m(x_0) & \cdots & \partial_{x_n} f_m(x_0) \end{pmatrix}$
2. $\forall v \in \mathbb{R}^n : \partial_v f(x_0) = f'(x_0)v$

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

$$f:\mathbb{R}^2 \to \mathbb{R} \;,\; f(x,y) = \begin{cases} \frac{xy^2}{x^2+y^4} & ,\,(x,y) \neq (0,0) \\ 0 & ,\,(x,y) = (0,0) \end{cases} \qquad \text{gilt}$$

$$\forall v = (v_1,v_2) \in \mathbb{R}^2 : \partial_v f(0,0) = \begin{cases} \frac{v_2^2}{v_1} & ,\,v_1 \neq 0 \\ 0 & ,\,v_1 = 0 \end{cases},$$

aber f ist, wie wir in Abschnitt 19 gesehen haben, nicht stetig in $(0,0)$. Und für

$$g:\mathbb{R}^2 \to \mathbb{R} \;,\; g(x,y) = \begin{cases} \frac{y^3-x^2y}{x^2+y^2} & ,\,(x,y) \neq (0,0) \\ 0 & ,\,(x,y) = (0,0) \end{cases} \qquad \text{gilt}$$

$$\forall v \in \mathbb{R}^2 : \partial_v g(0,0) = \begin{cases} \frac{v_2^3-v_1^2v_2}{v_1^2+v_2^2} & ,\,v \neq 0 \\ 0 & ,\,v = 0 \end{cases}.$$

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 ist g nach 20.3 $b)$ nicht differenzierbar in $(0,0)$.

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.*

Satz 20.5 *Sei $f : \mathbb{R}^n \supseteq D \to \mathbb{R}$ (total) differenzierbar in D. Dann gilt*

1. *Existiert eine reguläre parametrisierte Kurve $\gamma \in C^1(]-\varepsilon,\varepsilon[\,,D)$ 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.

2. *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.

Satz 20.6 *1. 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* $\frac{f_i}{g_i}$ *in* x_0 *(total) differenzierbar mit*

$$\left(\frac{f_i}{g_i}\right)'(x_0) = \frac{1}{(g_i(x_0))^2} \left(g_i(x_0) f_i'(x_0) - f_i(x_0) g_i'(x_0)\right)$$

2. *(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*

$$(h \circ g)'(x_0) = h'(g(x_0))g'(x_0) \quad (im\ Sinne\ der\ Matrizenmultiplikation)$$

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*

$$\exists \vartheta \in]0,1[: f(b) - f(a) = f'(a + \vartheta(b-a))(b-a)$$

Bemerkung

Sei $f : \mathbb{R}^n \supseteq D \to \mathbb{R}^m$ *(total) differenzierbar in* D *und* $\{a + t(b-a) : t \in [0,1]\} \subseteq D$*. Dann gilt nach Satz 20.7:*

$$\forall i \in \{1, \ldots, m\}\ \exists \vartheta_i \in]0,1[: f_i(b) - f_i(a) = f_i'(a + \vartheta_i(b-a))(b-a)$$

Im Allgemeinen gilt aber $\vartheta_i \neq \vartheta_j$ *für* $i \neq j$ *und dann existiert kein* $\vartheta \in]0,1[$ *mit* $f(b) - f(a) = f'(a + \vartheta(b-a))(b-a)$.

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.

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

$$f(x,y) = \begin{cases} xy\frac{x^2-y^2}{x^2+y^2} & , (x,y) \neq (0,0) \\ 0 & , (x,y) = (0,0) \end{cases}$$

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,*

$$Hf(x_0) = \begin{pmatrix} \partial^2_{x_1} f(x_0) & \cdots & \partial_{x_1}\partial_{x_n} f(x_0) \\ \vdots & & \vdots \\ \partial_{x_n}\partial_{x_1} f(x_0) & \cdots & \partial^2_{x_n} f(x_0) \end{pmatrix}$$

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:*

$$C^k(M,\mathbb{R}^m) := \{f : M \to \mathbb{R}^m : \, f \text{ ist } k\text{−mal stetig differenzierbar in} \, M\}$$

$f : \mathbb{R}^n \supseteq D \to \mathbb{R}^m$ *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:*

$$C^\infty(M,\mathbb{R}^m) := \{f : M \to \mathbb{R}^m : \, f \text{ ist unendlich oft differenzierbar in} \, M\}$$

Bemerkung

1) 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 $k \geq 1$.
2) Durch wiederholte Anwendung von Satz 20.9 ergibt sich, dass für alle $f \in C^k(M,\mathbb{R}^m)$ die Reihenfolge der partiellen Differentiation in M bis zur k-ten Ordnung keine Rolle spielt.

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*

$$
\begin{aligned}
|\alpha| \; &:= \; \sum_{i=1}^{n} \alpha_i \,, \\
\alpha! \; &:= \; \prod_{i=1}^{n} \alpha_i! \,, \\
\beta = \alpha \; &:\Leftrightarrow \; \forall i \in \{1, \ldots, n\} : \beta_i = \alpha_i \,, \\
\beta \leq \alpha \; &:\Leftrightarrow \; \forall i \in \{1, \ldots, n\} : \beta_i \leq \alpha_i \,, \\
\alpha + \beta \; &:= \; (\alpha_1 + \beta_1, \ldots, \alpha_n + \beta_n) \,, \\
\alpha - \beta \; &:= \; (\alpha_1 - \beta_1, \ldots, \alpha_n - \beta_n \,, \quad \textit{falls } \beta \leq \alpha \,, \\
\binom{\alpha}{\beta} \; &:= \; \frac{\alpha!}{\beta!(\alpha - \beta)!} = \prod_{i=1}^{n} \binom{\alpha_i}{\beta_i} \,, \quad \textit{falls } \beta \leq \alpha \,, \\
\forall x \in \mathbb{R}^n : x^\alpha \; &:= \; \prod_{i=1}^{n} x_i^{\alpha_i} \,, \\
\forall D \subseteq \mathbb{R}^n \forall f \in C^{|\alpha|}(D, \mathbb{R}^m) : \partial^\alpha f \; &:= \; \partial_{x_1}^{\alpha_1} \partial_{x_2}^{\alpha_2} \ldots \partial_{x_n}^{\alpha_n} f \,,
\end{aligned}
$$

wobei $\partial_{x_i}^0 g := g$ ist.

Satz 20.10 *1. (Binomischer Satz)*

$$\forall a, b \in \mathbb{R}^n \, \forall \alpha \in \mathbb{N}_0^n : \, (a+b)^\alpha = \sum_{\beta \le \alpha} \binom{\alpha}{\beta} a^\beta b^{\alpha-\beta}$$

2. (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 $|\alpha| \le k$ gilt

$$\partial^\alpha (fg) = \sum_{\beta \le \alpha} \binom{\alpha}{\beta} \partial^\beta f \, \partial^{\alpha-\beta} g$$

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*

$$\exists \vartheta \in]0,1[: \; f(x) = \sum_{|\alpha| \le k} \frac{\partial^\alpha f(x_0)}{\alpha!}(x-x_0)^\alpha + \sum_{|\alpha|=k+1} \frac{\partial^\alpha f(x_0 + \vartheta(x-x_0))}{\alpha!}(x-x_0)^\alpha.$$

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*

$$f(x) = \sum_{|\alpha|\le k} \frac{\partial^\alpha f(x_0)}{\alpha!}(x-x_0)^\alpha + O(|x-x_0|^{k+1}) = \sum_{|\alpha|\le k} \frac{\partial^\alpha f(x_0)}{\alpha!}(x-x_0)^\alpha + o(|x-x_0|^k)$$

für $x \to x_0$.

Bemerkungen

1. Für $k = 0$ liefert der Satz von Taylor wieder den Mittelwertsatz, für $k = 1$ ergibt sich $f(x) = f(x_0) + \langle \nabla f(x_0), x-x_0\rangle_{\mathbb{R}^n} + \frac{1}{2}\langle x-x_0, Hf(x_0+\vartheta(x-x_0))(x-x_0)\rangle_{\mathbb{R}^n}$ mit $\vartheta \in]0,1[$ und für $k = 2$ erhalten wir $f(x) = f(x_0) + \langle \nabla f(x_0), x-x_0\rangle_{\mathbb{R}^n} + \frac{1}{2}\langle x-x_0, Hf(x_0)(x-x_0)\rangle_{\mathbb{R}^n} + o(|x-x_0|^2)$ 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.

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* $f : \mathbb{R}^n \supset D \to \mathbb{R}$ *eine Funktion,* $x_0 \in D$ *und* f *(total) differenzierbar in* x_0. *Dann gilt:* f *hat in* x_0 *ein lokales Extremum* $\Rightarrow f'(x_0) = 0$ *(also* $(\partial_{x_1} f(x_0), \ldots, \partial_{x_n} f(x_0)) = (0, \ldots, 0)$*).*

Definition
Ist $f : \mathbb{R}^n \supset D \to \mathbb{R}$ *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* $D \subset \mathbb{R}^n$, $f \in C^2(D,\mathbb{R})$ *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 $D \subset \mathbb{R}^n$, $f \in C^2(D,\mathbb{R})$ 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 $f_1(x,y) = x^2 + x^4$ (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 $f_3(x,y) = x^2 + x^3$ (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.

21 Mehrdimensionale Integralrechnung

Definition *Jede nichtleere Menge B der Form*

$$B = \{(x_1,\ldots,x_n) \in \mathbb{R}^n : a_1 \le x_1 \le b_1,\, a_2(x_1) \le x_2 \le b_2(x_1),\\ a_3(x_1,x_2) \le x_3 \le b_3(x_1,x_2),\ldots,a_n(x_1,\ldots,x_{n-1}) \le x_n \le b_n(x_1,\ldots,x_{n-1})\}$$

mit $a_1, b_1 \in \mathbb{R}$ und stetigen Funktionen a_i, b_i für $i = 2,\ldots,n$ 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 $B_{n+1} = \{(x_1,\ldots,x_n,y) \in \mathbb{R}^{n+1} : (x_1,\ldots,x_n) \in B_n \wedge a(x_1,\ldots,x_n) \le y \le b(x_1,\ldots,x_n)\}$, 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*

$$I(x_1,\ldots,x_n) = \int_{a(x_1,\ldots,x_n)}^{b(x_1,\ldots,x_n)} f(x_1,\ldots,x_n,y)dy$$

stetig.

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* $i \in \{1,\ldots,n\}$*, dann existiert auch* $\partial_{x_i}I \in C^0(B_n,\mathbb{R})$ *und es gilt*

$$\begin{aligned}
\partial_{x_i} I(x_1,\ldots,x_n) &= \partial_{x_i} \int_{a(x_1,\ldots,x_n)}^{b(x_1,\ldots,x_n)} f(x_1,\ldots,x_n,y)\,dy \\
&= \int_{a(x_1,\ldots,x_n)}^{b(x_1,\ldots,x_n)} \partial_{x_i} f(x_1,\ldots,x_n,y)\,dy \\
&\quad + f(x_1,\ldots,x_n,b(x_1,\ldots,x_n))\,\partial_{x_i} b(x_1,\ldots,x_n) \\
&\quad - f(x_1,\ldots,x_n,a(x_1,\ldots,x_n))\,\partial_{x_i} a(x_1,\ldots,x_n).
\end{aligned}$$

Definition

a) *Für $I \subseteq \mathbb{R}^n$ mit $I \neq \varnothing$ 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 $\prod_{i=1}^n]a, b_i[\subseteq \mathbb{R}^n$ 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 $Q \setminus \underset{j \in \prod_{i=1}^n \{1,\ldots,m_i\}}{U} Q_j$ können beliebig sein).

Die Menge aller Treppenfunktionen auf Q sei mit $\mathscr{T}(Q)$ 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 $Q = \prod_{i=1}^n [a_i, b_i]$, $f \in \mathscr{T}(Q)$ und $f = \sum\limits_{j \in \prod_{i=1}^n \{1,\ldots,m_i\}} c_j \chi_{Q_j}$ 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)

$$\underline{\int_Q} f(x_1,\dots,x_n)dx_1\dots dx_n := sup\{\int_Q \varphi(x_1,\dots,x_n)dx_1\dots dx_n : \varphi \in \mathscr{T}(Q) \wedge \varphi \le f\}$$

heißt Unterintegral von f über Q und

$$\overline{\int_Q} f(x_1,\dots,x_n)dx_1\dots dx_n := inf\{\int_Q \varphi(x_1,\dots,x_n)dx_1\dots dx_n : \varphi \in \mathscr{T}(Q) \wedge \varphi \ge f\}$$

heißt Oberintegral von f über Q.

b) *f heißt Riemann-integrierbar über Q, falls*

$$\underline{\int_Q} f(x_1,\dots,x_n)dx_1\dots dx_n = \overline{\int_Q} f(x_1,\dots,x_n)dx_1\dots dx_n.$$

In diesem Fall heißt

$$\int_Q f(x_1,\dots,x_n)dx_1\dots dx_n := \underline{\int_Q} f(x_1,\dots,x_n)dx_1\dots dx_n = \overline{\int_Q} f(x_1,\dots,x_n)dx_1\dots dx_n$$

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 stetige Funktionen auf beschr. abgeschl. Quadern)*
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 bzgl. der einzelnen Koordinaten berechnet werden, wobei die Reihenfolge der Integration egal ist.

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

$$\int_{a_{\sigma(1)}}^{b_{\sigma(1)}} \dots \left(\int_{a_{\sigma(n)}}^{b_{\sigma(n)}} f(x_1, \dots, x_n)\, dx_{\sigma(n)} \right) \dots dx_{\sigma(1)}$$

existiert, bei Änderung der Integrationsreihenfolge sich aber der Wert des iterierten Integrals ändert und das Integral $\int_Q f(x_1, \dots, x_n)\, dx_1 \dots dx_n$ nicht existiert.
Ein Beispiel hierfür ist die Funktion $f : [0,1] \times [0,1] \to \mathbb{R}$ mit

$$f(x,y) = \begin{cases} \frac{x^2-y^2}{(x^2+y^2)^2} & , (x,y) \neq (0,0) \\ 0 & , (x,y) = (0,0) \end{cases}.$$

Unter Ausnutzung von $\frac{x^2-y^2}{(x^2+y^2)^2} = \partial^2_{xy} arctan(\frac{x}{y})$ 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*

$$\tilde{f}(x_1,\ldots,x_n) = \begin{cases} f(x_1,\ldots,x_n) & ,\ (x_1,\ldots,x_n) \in M \\ 0 & ,\ (x_1,\ldots,x_n) \in Q \setminus M \end{cases}.$$

f *heißt integrierbar über* M*, falls* $\tilde{f}$ *integrierbar über* Q *ist. In diesem Fall heißt*

$$\int_M f(x_1,\ldots,x_n)\,dx_1 \ldots dx_n := \int_Q \tilde{f}(x_1,\ldots,x_n)\,dx_1 \ldots dx_n$$

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:*

$$\begin{aligned}
&\int_B f(x_1,\ldots,x_n)\,dx_1\ldots dx_n \\
=&\int_{a_1}^{b_1}\left(\int_{a_2(x_1)}^{b_2(x_1)}\left(\ldots\left(\int_{a_{n-1}(x_1,\ldots,x_{n-2})}^{b_{n-1}(x_1,\ldots,x_{n-2})}\left(\int_{a_n(x_1,\ldots,x_{n-1})}^{b_n(x_1,\ldots,x_{n-1})} f(x_1,\ldots,x_n)\,dx_n\right)dx_{n-1}\right)\ldots\right)dx_2\right)dx_1
\end{aligned}$$

wobei $a_1,\ldots,a_n$, $b_1,\ldots,b_n$ 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 $\int_B f(x_1,\ldots,x_n)\,dx_1\ldots dx_n$.
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 $f = g$ fast überall, dann gilt

$$\int_B g(x_1,\ldots,x_n)\,dx_1\ldots dx_n = \int_B f(x_1,\ldots,x_n)\,dx_1\ldots dx_n.$$

Definition *Sei $B \subseteq \mathbb{R}^n$ ein Normalbereich. Dann heißt*

$$V_n(B) := \int_B 1 \, dx_1 \dots dx_n$$

das n-dimensionale Volumen von B.

Satz 21.5 *Sei $r : [a,b] \to \mathbb{R}_0^+$ stetig und*

$$B = \left\{ (x,y,z) \in \mathbb{R}^3 \, : \, a \le x \le b \wedge y^2 + z^2 \le (r(x))^2 \right\}$$

Dann gilt

$$V_3(B) = \pi \int_a^b (r(x))^2 \, dx.$$

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*

$$\int_M f(x_1, \dots, x_n) \, dx_1 \dots dx_n = \int_{g^{-1}(M)} f\left(g(u_1, \dots, u_n)\right) \left| \det\left(g'(u_1, \dots, u_n)\right) \right| du_1 \dots du_n.$$

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

$$\int_{g^{-1}(M)} f\,(g(u_1,\ldots,u_n))\,|\det(g'(u_1,\ldots,u_n))|\,du_1\ldots du_n$$

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*

$$\begin{cases} y(x_0) = y_0 \\ y'(x_0) = y_1 \\ \vdots \\ y^{(n-1)}(x_0) = y_{n-1} \end{cases} \qquad (2)$$

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 einer Teilmenge von $\mathbb{R}$, 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

$$g(x, y(x), y'(x), \ldots, y^{(n)}(x)) = 0 , \tag{3}$$

wobei x ebenfalls aus einer Teilmenge von $\mathbb{R}$ 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 $\binom{1}{y'}$ vorgegeben. Gesucht ist eine differenzierbare Funktion y, deren Graph in jedem seiner Punkte (x, y) die Steigung y' und damit $\binom{1}{y'}$ als Tangentialvektor hat.

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.

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

$$\begin{cases} y' = g(x)\ h(y) \\ y(x_0) = y_0 \end{cases}$$

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.

Bemerkungen

1) Die sogenannte Ähnlichkeitsdifferentialgleichung

$$y' = f\left(\frac{y}{x}\right), \qquad x \neq 0$$

kann mit Hilfe der Substitution

$$u(x) := \frac{y(x)}{x}$$

umgeformt werden zu einer Differentialgleichung mit getrennten Variablen der Form

$$u' = \frac{1}{x}(f(u) - u).$$

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

$$y' = f(\frac{ax + by + c}{\alpha x + \beta x + \gamma}), \alpha x + \beta x + \gamma \neq 0$$

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.

23 Lineare gewöhnliche Differentialgleichungen

Definition *Sei $I \subseteq \mathbb{R}$ ein offenes Intervall und $a, b : I \to \mathbb{R}$ 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

$$y_{hom}(x) = ce^{\int_{x_0}^{x} a(t)dt}, c \in \mathbb{R}, x \in I.$$

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

$$y(x) = \left(y_0 + \int_{x_0}^{x} e^{-\int_{x_0}^{t} a(s)ds}\, b(t)\, dt \right) e^{\int_{x_0}^{x} a(t)dt} = y_0\, e^{\int_{x_0}^{x} a(t)dt} + \int_{x_0}^{x} e^{\int_{t}^{x} a(s)ds}\, b(t)\, dt.$$

Die allgemeine Lösung der inhomogenen linearen Differentialgleichung $y' = a(x)y + b(x)$ ist

$$y_{inh}(x) = y_{hom}(x) + y_p(x), \qquad x \in I,$$

wobei y_{hom} wie in Satz 23.1 ist und

$$y_p(x) = \int_{x_0}^{x} e^{\int_{t}^{x} a(s)ds}\, b(t)\, dt.$$

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.

Bemerkungen

1) Die sogenannte Bernoulli-Differentialgleichung

$$y' = a(x)y + b(x)y^{\alpha}, \quad \alpha \neq 1$$

kann mit Hilfe der Substitution

$$u(x) := (y(x))^{1-\alpha}$$

umgeformt werden zu einer inhomogenen linearen Differentialgleichung der Form

$$u' = (1-\alpha)a(x)y + (1-\alpha)b(x).$$

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

$$u(x) = \frac{1}{y(x) - y_p(x)}$$

um zu einer inhomogenen linearen Differentialgleichung der Form

$$u' = -(a(x) + 2b(x)y_p(x))u - b(x).$$

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 $a_0, \ldots, a_{n-1}, b : I \to \mathbb{K}$ 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 $a_0, \ldots, a_{n-1}$ 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 $P(\lambda) = \sum_{k=0}^{n} a_k \lambda^k$ vom Grad $n \geq 1$ und jede stetige Funktion $b : I \to \mathbb{K}$, wobei $I \subseteq \mathbb{R}$ ein offenes Intervall ist, ist $P(\frac{d}{dx})y = b(x)$ mit $P(\frac{d}{dx}) = \sum_{k=0}^{n} a_k \frac{d^k}{dx^k}$, 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*

$$\left(\frac{d}{dx} - \lambda\right)^k \left(f(x)e^{\lambda x}\right) = \frac{d^k f}{dx^k}(x)e^{\lambda x}.$$

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*

$$\forall x \in \mathbb{R}\ P\left(\frac{d}{dx}\right)\left(g(x)e^{\lambda x}\right) = h(x)e^{\lambda x},$$

wobei h ebenfalls ein reelles oder komplexes Polynom vom Grad k ist.

Satz 23.5 *Sei* $P(\lambda) = \sum_{k=0}^{n} a_k \lambda^k$ *ein reelles oder komplexes Polynom vom Grad* $n \geq 1$, $\{\lambda_j : 1 \leq j \leq r\}$ *die Menge der Nullstellen von* P *und* k_j *die Vielfachheit der Nullstellen* λ_j, *d.h., es sei* $P(\lambda) = a_n \prod_{j=1}^{r} (\lambda - \lambda_j)^{k_j}$. *Dann gilt:*

a) Die Funktionen

$$y_{jm}(x) := x^m e^{\lambda_j x}, \quad 1 \leq j \leq r,\ 0 \leq m \leq k_j - 1,$$

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* $x^m e^{\mu x}$, $x^m e^{\overline{\mu} x}$, $0 \leq m \leq k_\mu$, *aus dem Fundamentalsystem aus a) durch die Funktionen*

$$x^m e^{\alpha x} \cos(\beta x),\ x^m e^{\alpha x} \sin(\beta x), \quad 0 \leq m \leq k_\mu,$$

dann erhält man ein reelles Fundamentalsystem von $P\left(\frac{d}{dx}\right) y = 0$.

c) Sei $x_0 \in \mathbb{R}$ *und seien* $y_0, \ldots, y_{n-1} \in \mathbb{C}$. *Dann besitzt das Anfangswertproblem*

$$\begin{cases} P\left(\frac{d}{dx}\right) y = 0 \\ \frac{d^k}{dx^k} y(x_0) = y_k, \quad k = 0, \ldots, n-1 \end{cases}$$

genau eine Lösung $y : \mathbb{R} \to \mathbb{C}$. *Ist* P *reell und sind* $y_0, \ldots, y_{n-1} \in \mathbb{R}$, *dann ist* y *reellwertig.*

Satz 23.6 *Sei* $P(\lambda) = \sum_{k=0}^{n} a_k \lambda^k$ *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
$P\left(\frac{d}{dx}\right) y = e^{\mu x}$.

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*
$P\left(\frac{d}{dx}\right) y = f(x) e^{\mu x}$
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*
$P\left(\frac{d}{dx}\right) y = f(x) e^{\mu x}$
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.

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 $Ly = \alpha b_1 + \beta b_2$.

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

$$(u')' + \left(2\frac{y_1'(x)}{y_1(x)} + a(x)\right)u' = 0$$

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* $\lambda_1, \dots, \lambda_n \in \mathbb{K}$. *Dann bilden die vektorwertigen Funktionen*

$$\begin{pmatrix} e^{\lambda_1 x} \\ 0 \\ \vdots \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ e^{\lambda_2 x} \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ \vdots \\ \vdots \\ 0 \\ e^{\lambda_n x} \end{pmatrix}$$

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.

Korollar 23.9 *a) Existiert eine Basis $a_1,\dots,a_n$ von $\mathbb{K}^n$ aus Eigenvektoren der Matrix $A \in \mathbb{K}^{n\times n}$ zu den jeweiligen Eigenwerten $\lambda_i, i = 1,\dots,n$, dann bilden die (vektorwertigen) Funktionen*

$$y_i : \mathbb{R} \to \mathbb{K}^n, \quad y_i(x) = e^{\lambda_i x} a_i, \qquad\qquad i = 1,\dots,n,$$

ein Fundamentalsystem von

$$y' = Ay$$

und für alle $x_0 \in \mathbb{R}$ und $y_0 \in \mathbb{K}^n$ besitzt das Anfangswertproblem

$$\begin{cases} y' = Ay\,, \\ y(x_0) = y_0\,, \end{cases}$$

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

$$e^{\mu x} a_{\mu,i}, \quad e^{\overline{\mu} x} a_{\overline{\mu},i}, \qquad\qquad i = 1, \ldots, k_\mu,$$

aus dem Fundamentalsystem aus $a)$ durch die Funktionen

$$Re(e^{\mu x} a_{\mu,i}), \quad Im(e^{\mu x} a_{\mu,i}), \qquad\qquad i = 1, \ldots, k_\mu,$$

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.

Satz 23.10 *Sei E_n die Einheitsmatrix in $\mathbb{K}^{n\times n}$ und $N_n \in \mathbb{K}^{n\times n}$ mit*

$$N_n = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ & & & \ddots & 1 \\ 0 & \dots & & 0 & 0 \end{pmatrix}$$ *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 $y : \mathbb{R} \to \mathbb{K}^n$.

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.

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:

$$\underbrace{M_k}_{=(m_{kij})_{i,j=1,\dots,n}} \xrightarrow{k \to \infty} \underbrace{M}_{=(m_{ij})_{i,j=1,\dots,n}} \quad :\Leftrightarrow \quad \forall i, j = 1, \dots, n : m_{kij} \xrightarrow{k \to \infty} m_{ij}$$

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.

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 $\psi_0 \in L_I$:*

$$L_I = \psi_0 + L_H.$$

Satz 23.12 (Variation der Konstanten) *Mit den Bezeichungen von Satz 23.11 gilt: Sei $\{\varphi_1(x), \dots, \varphi_n(x)\}$ ein Fundamentalsystem des homogenen Differentialgleichungssystems $y' = A(x)y$ und*

$$\Phi(x) := (\varphi_1(x), \dots, \varphi_n(x))$$

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.

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*

$$y([x_0 - h, x_0 + h]) \subseteq \{y \in \mathbb{R}^m : |y - y_0| \leq b\}.$$

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 $L \leq \sup_{\substack{(x,y) \in B \\ i=1,\ldots,m}} |\partial_{y_i} f(x,y)|$, 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.

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*

$$\begin{cases} y' = A(x)y + b(x) \\ y(x_0) = y_0 \end{cases}$$

für alle $x_0 \in I$ und alle $y_0 \in \mathbb{K}^m$ genau eine Lösung $y \in C^1(I, \mathbb{K}^m)$.

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

$$\left\{ y_1 = \begin{pmatrix} y_{11} \\ \vdots \\ y_{m1} \end{pmatrix}, \dots, y_1 = \begin{pmatrix} y_{m1} \\ \vdots \\ y_{mm} \end{pmatrix} \right\} \subseteq L_H$$

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.

Satz 24.4 *y ist genau dann Lösung der gewöhnlichen Differentialgleichung n-ter Ordnung*

$$y^{(n)} = f(x, y, y', \dots, y^{(n-1)})$$

wenn $\begin{pmatrix} y \\ y' \\ \vdots \\ y^{(n-1)} \end{pmatrix}$ *Lösung des Systems von n gewöhnlichen Differentialgleichungen erster Ordnung*

$$\begin{aligned} y_0' &= y_1 \\ y_1' &= y_2 \\ &\vdots \\ y_{n-2}' &= y_{n-1} \\ y_{n-1}' &= f(x, y_0, y_1, \dots, y_{n-1}) \end{aligned}$$

ist.

Korollar 24.5 *Sei* $f : \mathbb{R} \times \mathbb{R}^m \supseteq D \to \mathbb{R}$, $(x, Y) = (x, y, y_1, \dots, y_{n-1}) \mapsto f(x, Y) = f(x, y, y_1, \dots, y_{n-1})$ *eine Funktion, die auf* $B := [x_0 - a, a_0 + a] \times \{Y \in \mathbb{R}^n : |Y - Y_0| \leq b\}$ *mit* $x_0 \in \mathbb{R}$, $a, b \in \mathbb{R}^+$ *und* $Y_0 \in R^n$ *stetig und Lipschitz-stetig bezüglich* Y *ist. Dann existiert ein* $h > 0$, *so dass das Anfangswertpoblem*

$$
\begin{aligned}
y^{(n)} &= f(x, y, y', \dots, y^{(n-1)}) \\
y(x_0) &= y_0 \\
y'(x_0) &= (y')_0 \\
&\vdots \\
y^{(n-1)}(x_0) &= (y^{(n-1)})_0
\end{aligned}
$$

genau eine Lösung $y \in C^n([x_0 - h, x_0 + h], \mathbb{R})$ *für alle* $(y_0, (y')_0, \dots, (y^{(n-1)})_0) \in \mathbb{R}^n$ *besitzt.*

Korollar 24.6 *Sei $I \subseteq \mathbb{R}$ ein Intervall, $n \in \mathbb{N}$, $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$, $a_n \in K \setminus \{0\}$ und seien $a_0, \ldots, a_{n-1}, b : I \to \mathbb{K}$ stetige Funktionen. Dann besitzt das lineare Anfangswertpoblem n-ter Ordnung*

$$\left(\sum_{k=0}^{n} a_k(x) \frac{d^k}{dx^k} \right) y = b(x), \qquad \frac{d^k}{dx^k} y(x_0) = y_k, \qquad k = 0, \ldots, n-1$$

für alle $x_0 \in I$ und alle $y_0, \ldots, y_{n-1} \in \mathbb{K}$ 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 $\{\varphi_1, \ldots, \varphi_n\} \subseteq L_H$ ist genau dann ein Fundamentalsystem der homogenen Gleichung wenn für ein und damit für alle $x \in I$ die Wronskideterminante

$$W(x) := \det \begin{pmatrix} \varphi_1(x) & \varphi_2(x) & \ldots & \varphi_n(x) \\ \varphi_1'(x) & \varphi_2'(x) & \ldots & \varphi_n'(x) \\ \vdots & \vdots & \ldots & \vdots \\ \varphi_1^{(n-1)}(x) & \varphi_2^{(n-1)}(x) & \ldots & \varphi_n^{(n-1)}(x) \end{pmatrix}$$

von Null verschieden ist.

Satz 24.7 *Sei f wie in 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

$$\forall x \in [x_0-h, x_0+h] : |y_1(x) - y_2(x)| \le e^{L|x-x_0|} |y_{10} - y_{20}| \ .$$

Bemerkung

Für $f(x,y) = Ly$ gilt "$=$" in der Abschätzung aus Satz 24.7. Für manche Funktionen dagegen, wie zum Beispiel $f(x,y) = (1+x^2)^{-1}y$, ist die Abschätzung eher grob und kann verfeinert werden.

Definition *Sei (X,d) ein metrischer Raum, $A, M, \mathscr{O} \subset X$, $x_0 \in X$ 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) $\mathscr{O}$ heißt offen, falls

$$\forall x \in \mathscr{O} \, : \, \exists \varepsilon > 0 \, : \, B_\varepsilon(x) \subset \mathscr{O}$$

c) A heißt abgeschlossen, falls $X \backslash A$ offen ist.

d) $\overset{\circ}{M} := \{x \in M \, : \, \exists \varepsilon > 0 \, : \, B_\varepsilon(x) \subset M\}$ heißt das Innere von M.

e) $\overline{M} := X \backslash \, (X \backslash \overset{\circ}{M})$ heißt Abschluss von M.

f) $\partial M := \overline{M} \backslash \, \overset{\circ}{M}$ heißt Rand von M.

Definition *Eine Folge $(x_n)_{n\in\mathbb{N}}$ in (X,d) heißt konvergent gegen $x \in X$ für $n \to \infty$ (Kurzschreibweise: $x_n \longrightarrow x$ für $n \longrightarrow \infty$ oder auch $\lim_{n\to\infty} x_n = x$), 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.

Bemerkungen

1) In $(\mathbb{R}^m, ||\cdot||_2)$ gilt

$$\begin{aligned} x_n \xrightarrow{n\to\infty} x \quad &\iff \quad \sqrt{\sum_{i=1}^{m}((x_n)_i - x_i)^2} && \xrightarrow{n\to\infty} && 0 \\ &\iff \quad \forall i \in \{1,\ldots,m\} \;:\; (x_n)_i && \xrightarrow{n\to\infty} && x_i. \end{aligned}$$

Dies ist der bereits bekannte Konvergenzbegriff in $\mathbb{R}^m$.

2) Sei $X = C^0([0,1])$, $d(x,y) = \max_{0\le t\le 1} |x(t) - y(t)|$. Dann gilt

$$\begin{aligned} x_n \xrightarrow{n\to\infty} x \quad &\iff \quad \max_{0\le t\le 1} |x_n(t) - x(t)| \xrightarrow{n\to\infty} 0 \\ &\iff \quad \forall \varepsilon > 0\, \exists n_\varepsilon \in \mathbb{N}\, \forall n \ge n_\varepsilon : \max_{0\le t\le 1} |x_n(t) - x(t)| < \varepsilon \\ &\iff \quad \forall \varepsilon > 0\, \exists n_\varepsilon \in \mathbb{N}\, \forall n \ge n_\varepsilon : \forall t \in [0,1] : |x_n(t) - x(t)| < \varepsilon. \end{aligned}$$

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 gilt

$$x_n \xrightarrow{n\to\infty} x \quad \Longleftrightarrow \left(\int_0^1 |x_n(t)-x(t)|^p \mathrm{d}t\right)^{\frac{1}{p}} \xrightarrow{n\to\infty} 0$$
$$\Longleftrightarrow \forall \varepsilon > 0 \, \exists n_\varepsilon \in \mathbb{N} \, \forall n \geq n_\varepsilon : \int_0^1 |x_n(t)-x(t)|^p \mathrm{d}t < \varepsilon.$$

Dies nennt man auch Konvergenz im p-ten Mittel.

Satz 24.8 *Sei $(X, ||\cdot||)$ ein normierter Raum, $(x_n)_{n\in\mathbb{N}}, (y_n)_{n\in\mathbb{N}} \subset X, (\alpha_n)_{n\in\mathbb{N}} \subset \mathbb{K}$ mit $x_n \to x, y_n \to y$ und $\alpha_n \to \alpha$ für $n \longrightarrow \infty$. Dann gilt*

$$\alpha_n x_n + y_n \xrightarrow{n\to\infty} \alpha x + y.$$

Satz 24.9 *Sei (X,d) ein metrischer Raum, $M \subset X$. Dann gilt*

$$\overline{M} = \{x \in X \, : \, \exists (x_n)_{n\in\mathbb{N}} \subset M : \, x_n \to x\}.$$

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 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: $(x_n)_{\mathbb{N}} \subset \mathbb{Q}$ mit $\lim_{n\to\infty} = \sqrt{2}$, z.B. $x_n =$ die Dezimaldarstellung von $\sqrt{2}$ bis zur n-ten Stelle.

3) $C^0([a,b])$ mit $||x|| := \sup_{a\leq t\leq b} |x(t)|$ 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 $||x|| := \left(\int_a^b |x(t)|^p \mathrm{d}t\right)^{\frac{1}{p}}$ ist nicht vollständig. Gegenbeispiel: $(x_n)_{n\in\mathbb{N}} \subset C^0([0,1])$ mit

$$x_n(t) := \begin{cases} n^\alpha, t \leq \frac{1}{n} \\ t^{-\alpha}, t > \frac{1}{n} \end{cases}, \quad p = 2,\ 0 < \alpha < \frac{1}{2}$$

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*.*

Definition *Seien* $(X,d_X),(Y,d_Y)$ *metrische Räume. Eine Abbildung* $T : X \longrightarrow Y$ *heißt stetig in* $x_0 \in X$*, wenn* $\forall \varepsilon > 0 \exists \delta = \delta(\varepsilon, x_0) > 0$*, sodass* $d_Y(T(x), T(x_0)) < \varepsilon$ *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* $T(x_n) \stackrel{n\to\infty}{\longrightarrow} T(x)$*.*

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 $F : X \longrightarrow X$ eine Kontraktion, d.h.*

$$\exists \kappa \in (0,1) \forall x,y \in X \ : \ d(F(x),F(y)) \leq \kappa d(x,y).$$

Dann besitzt F einen eindeutigen Fixpunkt $x^\star$, d.h. $x^\star = F(x^\star)$.

Mit Hilfe des Banachschen Fixpunktsatzes können wir den Satz von Picard-Lindelöf beweisen.

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 $x_0 < x_1 < \ldots < x_n = x_0 + h$ 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 $k = 0, \ldots, n-1$. 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.