

Analysis I/II
HS 2020/FS2021

Martin U. Schmidt

Inhaltsverzeichnis

1	Mengen und Abbildungen	7
1.1	Mengen	7
1.2	Operationen von Mengen	8
1.3	Relationen	9
1.4	Abbildungen	9
1.5	Komposition von Abbildungen	10
2	Reelle Zahlen	11
2.1	Axiome von angeordneten Körpern	11
2.2	Die natürlichen Zahlen $\mathbb{N} \subset \mathbf{K}$	18
2.3	Mächtigkeit von Mengen	20
2.4	$\mathbb{R}$ und das Vollständigkeitsaxiom	22
2.5	Die erweiterte Zahlengerade $\bar{\mathbb{R}}$	26
2.6	Die Peano Axiome	27
2.7	Der Körper der komplexen Zahlen	28
3	Zahlenfolgen	31
3.1	Konvergenz	31
3.2	Konvergenzprinzipien	35
3.3	Häufungspunkte	38
3.4	Beispiele	40
4	Reihen	43
4.1	Konvergenzkriterien	43
4.2	Dezimalbruchdarstellung von reellen Zahlen	47
4.3	Addition, Multiplikation, Umordnung	48
4.4	Sinus und Cosinus	55
5	Stetigkeit	57
5.1	Teilmengen von $\mathbb{K}$	57
5.2	Vollständigkeit und Kompaktheit	58
5.3	Stetigkeit	59

6	Stetige Funktionen $f : \mathbb{R} \rightarrow \mathbb{R}$	63
6.1	Umkehrfunktionen	63
6.2	Exponentialfunktion und Logarithmus	65
6.3	Trigonometrische Funktionen	66
6.4	Fundamentalsatz der Algebra	69
7	Differenzierbare Funktionen $f : \mathbb{R} \rightarrow \mathbb{K}$	71
7.1	Definition der Ableitung	71
7.2	Rechenregeln der Ableitung	72
7.3	Mittelwertsatz und Monotonie	75
7.4	Regel von de L'Hopital	77
7.5	Konvexität und Ableitungen	79
7.6	Taylorreihen	82
8	Das Integral von Funktionen $f : \mathbb{R} \rightarrow \mathbb{K}$	85
8.1	Regelfunktionen	85
8.2	Technik des Integrierens	87
8.3	Riemannintegrale Funktionen	92
8.4	Kriterium von Riemann	94
8.5	Differentiation und Integration	97
8.6	Uneigentliches Integral	98
8.7	Die Konvergenz von Taylorreihen	101
9	Metrische Räume und Banachräume	107
9.1	Metrik und Norm	107
9.2	Vollständigkeit und Kompaktheit	112
9.3	Stetigkeit	114
9.4	Funktionenräume	119
9.5	Lineare Operatoren	125
10	Ableitungen in höheren Dimensionen	129
10.1	Ableitungen von $f : X \rightarrow Y$	129
10.2	Schranksatz	131
10.3	Partielle Ableitungen	133
10.4	Höhere Ableitungen	139
11	Nichtlineare Analysis	143
11.1	Der Banachsche Fixpunktsatz	143
11.2	Das Lösen von nichtlinearen Gleichungen	146
11.3	Lagrangemultiplikatoren	151

12 Das Lebesgueintegral auf dem $\mathbb{R}^d$	155
12.1 Treppenfunktionen	155
12.2 Lebesgueintegrale Funktionen auf dem $\mathbb{R}^d$	157
12.3 Das Riemann- und das Lebesgueintegral	162
12.4 Konvergenzsätze	165
12.5 Der Satz von Fubini	168
12.6 Jacobis Transformation von Maßen	170
12.7 Integration über Untermannigfaltigkeiten	173
12.8 Der Gaußsche Satz	176
12.9 Maßtheorie	178
12.10 Die Räume $L^p(\mathbb{R}^d)$	182

Kapitel 1

Mengen und Abbildungen

1.1 Mengen

Georg Cantor (1845-1918) hat den Begriff der Menge definiert als „eine Zusammenfassung von wohlbestimmten und wohlunterschiedenen Objekten unserer Anschauung oder unseres Denkens zu einem Ganzen“. Diese Objekte werden Elemente der Menge genannt. Um ein Objekt a als Element der Menge A zu kennzeichnen, schreiben wir $a \in A$. Ist a dagegen kein Element der Menge A , so schreiben wir $a \notin A$.

Wir können Mengen dadurch beschreiben, dass wir alle ihre Elemente angeben, also z.B. ist

$$A = \{a\}$$

die Menge, die nur das Element a enthält und B

$$B = \{a, b, c\}$$

die Menge, die drei Elemente a, b und c enthält. Dabei kann auch ein Element einer Menge wieder eine Menge sein:

$$C = \{a, \{a\}\}.$$

$$M = \{x \in X \mid x \text{ hat die Eigenschaft } p\} \text{ oder nur } M = \{x \mid x \text{ hat die Eigenschaft } p\}$$

bezeichnet die Menge aller Elemente x (von der Menge X), die die Eigenschaft p haben.¹

A ist eine Teilmenge von B , wenn alle Elemente von A auch Elemente von B sind. In Symbolen $A \subset B$.

¹Bei dieser Beschreibung muss man allerdings Vorsicht walten lassen, um die Russellsche Antinomie zu vermeiden. Lässt man nämlich die Menge aller Mengen zu, die sich nicht selbst als Element enthalten, so wird nicht entscheidbar, ob diese Menge sich selbst als Element enthält oder nicht. Als Ausweg wird in der axiomatischen Mengenlehre die Frage, ob eine Menge Element einer Menge ist, nicht in allen Fällen als sinnvoll zugelassen.

1.2 Operationen von Mengen

Die Vereinigung zweier Mengen A und B ist die Menge $A \cup B$ aller Elemente, die in mindestens einer der beiden Mengen A und B enthalten sind. Der Durchschnitt zweier Mengen A und B ist die Menge $A \cap B$ aller Elemente, die sowohl Element von A als auch Element von B sind. Die Differenzmenge $A \setminus B$ ist die Menge aller Elemente von A , die nicht Element von B sind. Das kartesische Produkt der Mengen A und B ist die Menge aller geordneten Paare (a, b) von Elementen von A und B .

$$A \times B = \{(a, b) \mid a \in A \text{ und } b \in B\}$$

Die Potenzmenge $\mathcal{P}(M)$ einer Menge M ist die Menge aller Teilmengen von M . Dabei ist die leere Menge $\emptyset$ stets ein Element der Potenzmenge.

z.B. $\mathcal{P}(\{a, \{a\}\}) = \{\emptyset, \{a\}, \{\{a\}\}, \{a, \{a\}\}\}$.

Wenn wir nur Teilmengen einer vorgegebenen Menge M betrachten, dann wird für eine solche Menge $A \in \mathcal{P}(M)$ die Menge $M \setminus A$ auch als das Komplement A^c bezeichnet. Diese Operationen erfüllen die folgenden Regeln:

(Idempotenz)	$A \cup A = A$	$A \cap A = A$
(Kommutativität)	$A \cup B = B \cup A$	$A \cap B = B \cap A$
(Assoziativität)	$A \cap (B \cap C) = (A \cap B) \cap C$	$A \cup (B \cup C) = (A \cup B) \cup C$
(Distributivität)	$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$	$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
(de Morgan)	$(A \cap B)^c = A^c \cup B^c$	$(A \cup B)^c = A^c \cap B^c$

$A \subset A \cup B$		$A \cap B \subset A$
$A \cup \emptyset = A$		$A \cap \emptyset = \emptyset$
$A \setminus A = \emptyset$		$A \setminus \emptyset = A$
$(A^c)^c = A$		$A \subset B \iff B^c \subset A^c$
$(A \subset B \text{ und } B \subset A)$	$\iff$	$A = B$
$A \cup B = B$	$\iff$	$A \subset B$
$A \cap B = A$	$\iff$	$A \subset B$
$(A \subset C \text{ und } B \subset C)$	$\iff$	$(A \cup B) \subset C$
$(C \subset A \text{ und } C \subset B)$	$\iff$	$C \subset (A \cap B)$

$$A \times \emptyset = \emptyset \times A = \emptyset$$

$$(A \times C) \cup (B \times C) = (A \cup B) \times C \quad (A \times C) \cap (B \times C) = (A \cap B) \times C.$$

1.3 Relationen

Als Relationen auf einer Menge bezeichnet man Aussagen über mehrere Elemente der Menge, die entweder wahr oder falsch sind. Wir wollen hier nur Aussagen über zwei Elemente einer Menge A betrachten. Jede solche Relation beschreiben wir durch eine Teilmenge R aller geordneten Paare in $A \times A$, in der wir alle die Paare zusammenfassen, für die die Aussage der Relation wahr ist. Wir sagen dann, dass $(a, b) \in A \times A$ diese Relation erfüllt, wenn (a, b) zu der Teilmenge R gehört. Andernfalls erfüllt (a, b) die Relation nicht. Wir führen jetzt folgende Eigenschaften einer Relation ein:

Reflexivität: für alle $a \in A$ erfüllt (a, a) die Relation.

Symmetrie: falls (a, b) die Relation erfüllt, dann auch (b, a) .

Transitivität: falls (a, b) und (b, c) die Relation erfüllen, dann auch (a, c) .

Definition 1.1. *Eine Äquivalenzrelation ist eine Relation, die reflexiv, symmetrisch und transitiv ist.*

Eine Äquivalenzrelation definiert dann sogenannte Äquivalenzklassen. Für jedes $a \in A$ ist die Äquivalenzklasse $[a]$ von a die Teilmenge aller Elemente $b \in A$, so dass (a, b) die Relation erfüllt. Aus der Transitivität folgt, dass für jedes $b \in [a]$ die entsprechende Äquivalenzklasse $[b]$ in $[a]$ enthalten ist. Aus der Symmetrie folgt dann $a \in [b]$ und wegen der Transitivität wieder $[a] \subset [b]$. Wegen der Symmetrie und der Transitivität folgt aus $b \in [a]$ also $[a] = [b]$. Wenn zwei Äquivalenzklassen $[a]$ und $[b]$ beide ein Element $c \in A$ enthalten, dann gilt also $[a] = [c] = [b]$. Also sind zwei Äquivalenzklassen entweder disjunkt oder gleich. Wegen der Reflexivität ist jedes Element in einer Äquivalenzklasse enthalten. Also zerfällt A in eine disjunkte Vereinigung von Äquivalenzklassen, d.h. jedes Element von A gehört zu genau einer Äquivalenzklasse.

1.4 Abbildungen

Eine Abbildung $f : X \rightarrow Y$ ist eine Zuordnungsvorschrift, die jedem Element x von X genau ein Element $f(x)$ aus Y zuordnet:

$$f : X \rightarrow Y, \quad x \mapsto f(x)$$

Die Menge X der Argumente wird Definitionsbereich genannt und die Menge Y , in die abgebildet wird, Wertebereich. Das Bild einer Teilmenge $A \subset X$ ist die Teilmenge aller Elemente y des Wertebereichs Y , auf die ein Argument in A abgebildet wird:

Bild $f[A] = \{y \in Y \mid \exists x \in A \text{ mit } f(x) = y\}$. $\exists$ steht für “es gibt (mindestens) ein”

Definition 1.2. Eine Abbildung $f : X \rightarrow Y$, $x \mapsto f(x)$ heißt

(i) *injektiv*, wenn je zwei verschiedene Elemente $x, x' \in X$ auch auf verschiedene Elemente von Y abgebildet werden:

$$\forall x, x' \in X \text{ folgt aus } x \neq x' \text{ auch } f(x) \neq f(x'). \quad \forall \text{ steht für "für alle"}$$

(ii) *surjektiv*, wenn das Bild $f[X]$ der ganze Wertebereich Y ist.

$$\forall y \in Y \quad \exists x \in X \text{ mit } f(x) = y.$$

(iii) *bijektiv*, wenn sie injektiv und surjektiv ist.

Für jede Teilmenge $B \subset Y$ ist das Urbild von B unter f die Menge aller Elemente von X , die nach B abgebildet werden:

$$f^{-1}[B] = \{x \in X \mid f(x) \in B\}.$$

Für eine bijektive Abbildung $f : X \rightarrow Y$, $x \mapsto f(x)$ besteht für jedes $y \in Y$ das Urbild $f^{-1}[\{y\}]$ nur aus genau einem Element. Also existiert auch die Umkehrabbildung

$$f^{-1} : Y \rightarrow X, \quad y \mapsto f^{-1}(y) \text{ mit } f^{-1}[\{y\}] = \{f^{-1}(y)\}.$$

Offenbar gilt dann $f^{-1}(f(x)) = x$ für alle $x \in X$ und $f(f^{-1}(y)) = y$ für alle $y \in Y$.

1.5 Komposition von Abbildungen

Seien $f : X \rightarrow Y$, $x \mapsto f(x)$ und $g : Y \rightarrow Z$, $y \mapsto g(y)$ Abbildungen, dann definiert $g \circ f : X \rightarrow Z$, $x \mapsto g(f(x))$ die sogenannte Komposition (Verkettung) von f und g .

Satz 1.3. Die Komposition von Abbildungen ist assoziativ, d.h. für Abbildungen $f : X \rightarrow Y$, $x \mapsto f(x)$, $g : Y \rightarrow Z$, $y \mapsto g(y)$ und $h : Z \rightarrow V$, $z \mapsto h(z)$ gilt

$$h \circ (g \circ f) = (h \circ g) \circ f.$$

Sei $f : X \rightarrow Y$, $x \mapsto f(x)$ eine Abbildung und seien $\mathbb{1}_X : X \rightarrow X$, $x \mapsto x$ und $\mathbb{1}_Y : Y \rightarrow Y$, $y \mapsto y$ die identischen Abbildungen von den Mengen X und Y . Dann gilt

$$f \circ \mathbb{1}_X = f = \mathbb{1}_Y \circ f = f$$

Ist die Abbildung f bijektiv, dann gilt auch $f \circ f^{-1} = \mathbb{1}_Y$ und $f^{-1} \circ f = \mathbb{1}_X$.

Beweis:

$(h \circ (g \circ f))(x) = h(g(f(x))) = (h \circ g) \circ f(x)$	$\forall x \in X$	
$f \circ \mathbb{1}_X(x) = f(x) = (\mathbb{1}_Y \circ f)(x)$	$\forall x \in X$	
$f \circ f^{-1}(y) = f(f^{-1}(y)) = y = \mathbb{1}_Y(y)$	$\forall y \in Y$	
$f^{-1} \circ f(x) = f^{-1}(f(x)) = x = \mathbb{1}_X(x)$	$\forall x \in X$	q.e.d.

Kapitel 2

Reelle Zahlen

2.1 Axiome von angeordneten Körpern

Zunächst charakterisieren wir angeordnete Körper $\mathbf{K}$ durch folgende Axiome:

A1. Axiome der Addition

A2. Axiome der Multiplikation

A3. Distributivgesetz

A4. Ordnungsaxiome

Es gibt mehrere angeordnete Körper. Der wichtigste ist der Körper der reellen Zahlen $\mathbb{R}$. Er wird später als der eindeutige angeordnete Körper charakterisiert, der zusätzlich das **A5.** Vollständigkeitsaxiom erfüllt.

A1. Axiome der Addition 2.1. *Es gibt eine Operation*

$+: \mathbf{K} \times \mathbf{K} \rightarrow \mathbf{K}, \quad (x, y) \mapsto x + y$ mit

- (i) *Kommutativgesetz:* $x + y = y + x$ für alle $x, y \in \mathbf{K}$.
- (ii) *Assoziativgesetz:* $x + (y + z) = (x + y) + z$ für alle $x, y, z \in \mathbf{K}$.
- (iii) *Existenz der Null:* es gibt eine Zahl $0 \in \mathbf{K}$ mit $x + 0 = x$ für alle $x \in \mathbf{K}$
- (iv) *Existenz des Negativen:* zu jedem $x \in \mathbf{K}$ gibt es ein $-x \in \mathbf{K}$ mit $x + (-x) = 0$.

A2. Axiome der Multiplikation 2.2. *Es gibt eine Operation*

$\cdot: \mathbf{K} \times \mathbf{K} \rightarrow \mathbf{K}, \quad (x, y) \mapsto x \cdot y$ mit

- (i) *Kommutativgesetz:* $x \cdot y = y \cdot x$ für alle $x, y \in \mathbf{K}$.
- (ii) *Assoziativgesetz:* $x \cdot (y \cdot z) = (x \cdot y) \cdot z$ für alle $x, y, z \in \mathbf{K}$.

- (iii) *Existenz der Eins:* es gibt eine Zahl $1 \in \mathbf{K}$, $1 \neq 0$ mit $x \cdot 1 = x$ für alle $x \in \mathbf{K}$
- (iv) *Existenz des Inversen:* zu jedem $x \in \mathbf{K} \setminus \{0\}$ gibt es ein $x^{-1} \in \mathbf{K}$ mit $x \cdot x^{-1} = 1$.

A3. Distributivgesetz 2.3.

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

Definition 2.4. Allgemein heißt eine Menge $\mathbf{K}$, die die Axiome **A1-A3** erfüllt Körper. Für Körper gelten daher auch alle Folgerungen aus **A1-A3**. Es gibt viele Körper.

Beispiel 2.5. Der kleinste Körper besteht aus zwei Elementen $\mathbb{Z}_2 = \{0, 1\}$. Die Operationen $+$ und $\cdot$ sind dann definiert durch:

$$0 + 0 = 0 \quad 0 + 1 = 1 \quad 0 \cdot 0 = 0 \quad 0 \cdot 1 = 0$$

$$1 + 0 = 1 \quad 1 + 1 = 0 \quad 1 \cdot 0 = 0 \quad 1 \cdot 1 = 1$$

Zeige dass diese Definitionen von $+$ und $\cdot$ auf $\mathbb{Z}_2$ die Axiome **A1-A3** erfüllen und umgekehrt durch **A1-A3** eindeutig bestimmt sind. In einem angeordneten Körper $\mathbf{K}$ soll aber $1 + 1 > 0$, also $1 + 1 \neq 0$ gelten, so dass wir noch weitere Axiome benötigen um angeordnete Körper zu charakterisieren.

Wir benützen folgende Abkürzungen:

$$\begin{array}{llll} x - y = x + (-y) & xy = x \cdot y & -xy = -(x \cdot y) & \forall x, y \in \mathbf{K} \\ \frac{1}{x} = x^{-1} & \frac{y}{x} = y \cdot x^{-1} & & \forall x \in \mathbf{K} \setminus \{0\}, y \in \mathbf{K} \\ x + y + z = x + (y + z) & xyz = x \cdot (y \cdot z) & xy + z = (x \cdot y) + z & \forall x, y, z \in \mathbf{K} \end{array}$$

Satz 2.6 (Folgerungen aus **A1**).

- (i) Falls $x + y = x + z$, dann $y = z$
- (ii) Falls $x + y = x$, dann $y = 0$
- (iii) Falls $x + y = 0$, dann $y = -x$
- (iv) $-(-x) = x$

Bemerkung 2.7. (i) heißt Kürzungsregel.

(ii) zeigt, dass die Null eindeutig durch die Eigenschaft $x + 0 = x$ bestimmt ist.

(iii) zeigt, dass das Negative $(-x)$ eindeutig durch $x + (-x) = 0$ bestimmt ist.

Beweis: (i) Sei $x + y = x + z$, dann folgern wir

$$\begin{aligned}
 y &= y + 0 &= 0 + y &= (x - x) + y \\
 &= (-x + x) + y &= -x + (x + y) &= -x + (x + z) &= (-x + x) + z \\
 &= (x - x) + z &= 0 + z &= z + 0 &= z.
 \end{aligned}$$

(ii) Sei $x + y = x$, dann gilt $x + y = x + 0$. Also folgt aus (i) $y = 0$.

(iii) Sei $x + y = 0$, dann gilt $x + y = x + (-x)$. Also folgt aus (i) $y = -x$.

(iv) $-x + x = x + (-x) = 0$. Also folgt aus (iii) $x = -(-x)$.

q.e.d.

Satz 2.8 (Folgerungen aus **A2**).

(i) Falls $x \neq 0$ und $xy = xz$, dann $y = z$

(ii) Falls $x \neq 0$ und $xy = x$, dann $y = 1$

(iii) Falls $x \neq 0$ und $x \cdot y = 1$, dann $y = x^{-1}$

(iv) Falls $x \neq 0$ und $x^{-1} \neq 0$, dann $(x^{-1})^{-1} = x$

Bemerkung 2.9. Die Folgerungen sind analog zu denen aus **A1**. Wieder heißt (i) Kürzungsregel, (ii) impliziert wieder die Eindeutigkeit der Eins und (iii) die Eindeutigkeit des Inversen.

Beweis: (i) Sei $x \neq 0$ und $xy = xz$, dann folgern wir

$$y = 1y = \left(x \frac{1}{x}\right) y = \left(\frac{1}{x} x\right) y = \frac{1}{x}(xy) = \frac{1}{x}(xz) = \left(\frac{1}{x} x\right) z = \left(x \frac{1}{x}\right) z = 1z = z.$$

(ii) Sei $x \neq 0$ und $xy = x$, dann gilt $xy = x \cdot 1$. Also folgt aus (i) $y = 1$.

(iii) Sei $x \neq 0$ und $xy = 1$, dann gilt $xy = x \cdot x^{-1}$. Also folgt aus (i) $y = x^{-1}$.

(iv) Sei $x \neq 0$. Dann ist $x^{-1}x = xx^{-1} = 1$. Also folgt (iv) aus (iii).

q.e.d.

Satz 2.10 (Folgerungen aus **A1-A3**).

(i) $x \cdot 0 = 0$ für alle x . Insbesondere gilt $x^{-1} \neq 0$ für alle $x \neq 0$.

(ii) $x \cdot y = 0 \iff x = 0$ oder $y = 0$

(iii) $(-x)y = -xy = x(-y)$.

(iv) $(-1) \cdot x = -x$

(v) $(-x)(-y) = xy$

(vi) $x \neq 0, y \neq 0$ dann $(xy)^{-1} = y^{-1}x^{-1}$.

Bemerkung 2.11. Die Null hat kein multiplikatives Inverses, sonst wäre wegen (i) $0 = 0 \cdot 0^{-1} = 1$. Daraus und aus (i) würde $x = x \cdot 1 = x \cdot 0 = 0$ für alle x folgen.

Beweis: (i) $x \cdot 0 + x \cdot 0 = x \cdot (0 + 0) = x \cdot 0$. Aus (ii) von Satz 2.6 folgt dann $x \cdot 0 = 0$
(ii) Sei $x \cdot y = 0$ und $x \neq 0$. Dann gilt wegen (i) $x \cdot y = 0 = x \cdot 0$. Also folgt aus Satz 2.8 (i) $y = 0$.

(iii) $0 = 0 \cdot y = (x + (-x))y = xy + (-x)y$. Aus (iii) im Satz 2.6 folgt dann

$$(-x)y = -xy = -yx = (-y)x = x(-y).$$

(iv) Setze in (iii) $x = 1$.

(v) Wegen (iii) und (iv) im Satz 2.6 gilt $(-x)(-y) = -(x \cdot (-y)) = -(-xy) = xy$.

(vi) $1 = xy(xy)^{-1} = (xy)^{-1}xy = ((xy)^{-1}x)y = y((xy)^{-1}x)$. Wegen Satz 2.8 (iii) folgt $y^{-1} = (xy)^{-1}x$, also $y^{-1}x^{-1} = ((xy)^{-1}x)x^{-1} = (xy)^{-1}(xx^{-1}) = (xy)^{-1}$. **q.e.d.**

A4. Ordnungsaxiome 2.12. Es gibt eine Relation $<$ in $\mathbf{K}$ mit drei Eigenschaften:

(i) *Totalität der Ordnung:* Für je zwei Zahlen $x, y \in \mathbf{K}$ gilt genau eine der drei folgenden Relationen $x < y$ oder $x = y$ oder $y < x$.

(ii) *Transitivität:* $x < y$ und $y < z \implies x < z$

(iii) *Monotonie:* $x < y \implies \begin{cases} x + c < y + c & \text{für alle } c \in \mathbf{K} \\ x \cdot c < y \cdot c & \text{für alle } 0 < c \in \mathbf{K} \end{cases}$

Definition 2.13. Ein Körper, der das Axiom A4 erfüllt, heißt angeordneter Körper.

Die rationalen Zahlen sind ein angeordneter Körper, und jeder angeordnete Körper enthält die rationalen Zahlen als Unterkörper. Wir benutzten folgende Abkürzungen:

$$x > y \iff y < x \qquad x \leq y \iff (x < y \text{ oder } x = y) \iff y \geq x.$$

Satz 2.14 (Folgerungen aus A1-A4).

(i) $0 < x \implies -x < 0$ und $x < 0 \implies -x > 0$

(ii) $x < y \iff 0 < y - x$

(iii) $x < y$ und $a < 0 \implies ya < xa$

(iv) $x \neq 0 \implies x \cdot x = x^2 > 0$

(v) $x > 0 \implies \frac{1}{x} > 0$

(vi) $0 < x < y \implies 0 < \frac{1}{y} < \frac{1}{x}$

Bemerkung 2.15. Da $1 \cdot 1 = 1$ folgt aus (iv) $1 > 0$ und $1 + 1 > 0 + 1 = 1 > 0$. Also ist Beispiel 2.5 kein angeordneter Körper. Außerdem folgt aus (i) $-1 < 0$. Also gilt $-1 < x^2$ für jedes $x \in \mathbf{K}$. Also gibt es keine Zahl $x \in \mathbf{K}$ mit $x^2 = -1$.

Beweis: (i) Sei $0 < x$. Dann folgt mit Monotonie $0 + (-x) < x + (-x)$, also $-x < 0$. Sei $x < 0$. Dann folgt mit Monotonie $x + (-x) < -x$ also auch $0 < -x$.

(ii) Sei $x < y$. Dann folgt mit Monotonie $x - x < y - x$, also auch $0 < y - x$.

Sei umgekehrt $0 < y - x$. Dann folgt mit Monotonie $x < y$.

(iii) Sei $x < y$ und $a < 0$. Dann folgt aus (i) $-a > 0$. Also gilt wegen Monotonie $-xa = x \cdot (-a) < y \cdot (-a) = -ya \iff 0 < xa - ya \iff ya < xa$.

(iv) Sei $x > 0$. Dann folgt wegen Monotonie $x^2 > 0 \cdot x = 0$.

Sei $x < 0$. Dann folgt aus (i) $-x > 0$ und mit Monotonie $x^2 = (-x) \cdot (-x) > 0 \cdot (-x) = 0$.

(v) Sei $x > 0$. Dann ist $x \neq 0$ und $\frac{1}{x} \neq 0$. Aus (iv) folgt $\frac{1}{x} \cdot \frac{1}{x} > 0$. Mit Monotonie folgt dann $\frac{1}{x} = x \cdot \frac{1}{x} \cdot \frac{1}{x} > 0 \cdot \frac{1}{x} \cdot \frac{1}{x} = 0$.

(vi) Sei $0 < x < y$. Dann ist $x > 0$ und $y > 0$ also wegen (v) auch $\frac{1}{x} > 0$ und $\frac{1}{y} > 0$. Dann folgt mit Monotonie $\frac{1}{y} = \frac{1}{y} \cdot \frac{1}{x} \cdot x < \frac{1}{y} \cdot \frac{1}{x} \cdot y = \frac{1}{x} \cdot \frac{1}{y} \cdot y = \frac{1}{x}$. **q.e.d.**

Satz 2.16 (Arithmetisches Mittel). Seien $x, y \in \mathbf{K}$, dann gilt

$$x < y \implies x < \frac{x+y}{2} < y$$

Also liegt zwischen zwei verschiedenen Zahlen immer eine weitere.

Beweis: Aus $x < y$ folgt mit Monotonie $x + x < x + y < y + y$. Weil aber $x + x = x(1+1) = 2x$ und $2 = 1+1 > 0$ folgt dann $2x < x + y < 2y$ und $x < \frac{x+y}{2} < y$ **q.e.d.**

Übungsaufgabe 2.17. Es gelten auch folgende Regeln:

(i) $a < b$ und $c < d \implies a + c < b + d$

(ii) $0 < a < b$ und $0 < c < d \implies ac < bd$

(iii) $ab > 0 \iff$ entweder $a > 0, b > 0$ oder $a < 0, b < 0$

(iv) $ab < 0 \iff$ entweder $a > 0, b < 0$ oder $a < 0, b > 0$

Definition 2.18 (Betrag). Der Betrag einer Zahl $x \in \mathbf{K}$ ist die Zahl

$$|x| = \begin{cases} x & \text{falls } x \geq 0 \\ -x & \text{falls } x < 0 \end{cases} \iff \begin{cases} 2x \geq 0 \\ 2x < 0 \end{cases} \iff \begin{cases} x \geq -x \\ x < -x \end{cases} = \max\{x, -x\}.$$

$$\max\{x, y\} = \begin{cases} x & \text{falls } y \leq x \\ y & \text{falls } y > x \end{cases} \quad \forall x, y \in \mathbf{K}.$$

Aus der Definition folgt

$$\begin{array}{lll}
 |x| \geq 0 & \text{und} & |x| = 0 \iff x = 0. \\
 -|x| \leq x \leq |x| & \iff & x \leq |x| \text{ und } -x \leq |x| \\
 |-x| = |x| & \text{denn} & |-x| = \max\{-x, x\}.
 \end{array}$$

Satz 2.19 (Eigenschaften des Betrags). *Für alle $x, y \in \mathbf{K}$ gilt:*

(i) $|x| \geq 0$ und $|x| = 0 \iff x = 0$

(ii) $|x \cdot y| = |x| \cdot |y|$

(iii) $|x + y| \leq |x| + |y|$

Beweis: (i) haben wir schon gesehen.

(ii) Wegen Satz 2.10 (iii) und (v) ändern sich beide Seiten nicht wenn wir x durch $-x$ bzw. y durch $-y$ ersetzen. Für $x \geq 0$ und $y \geq 0$ ist die Aussage klar.

(iii) Zwei Fälle: $x + y \geq 0$: $|x + y| = x + y \leq |x| + y \leq |x| + |y|$ wegen Monotonie.

$x + y < 0$: $|x + y| = -(x + y) = -x - y \leq |x| - y \leq |x| + |y|$ wegen Monotonie. **q.e.d.**

Korollar 2.20. $||x| - |y|| \leq |x - y|$

Beweis: $|x| = |x - y + y| \leq |x - y| + |y| \implies |x| - |y| \leq |x - y|$.

Vertausche x und $y \implies |y| - |x| \leq |y - x| = |x - y|$. Also gilt $||x| - |y|| \leq |x - y|$. **q.e.d.**

Definition 2.21 (Abstand).

Der Abstand $d(x, y)$ zweier Zahlen $x, y \in \mathbf{K}$ ist die nicht negative Zahl $d(x, y) = |x - y|$.

Satz 2.22 (Eigenschaften des Abstands).

Der Abstand $d : \mathbf{K} \times \mathbf{K} \rightarrow \mathbf{K}$, $(x, y) \mapsto d(x, y)$ hat folgende Eigenschaften:

(i) $d(x, y) \geq 0$ und $d(x, y) = 0 \iff x = y$

(ii) $d(x, y) = d(y, x)$

(iii) $d(x, y) \leq d(x, z) + d(z, y)$ für alle $x, y, z \in \mathbf{K}$.

Beweis: folgt aus den Folgerungen der Definition und Satz 2.19.

(iii) $|x - y| = |x - z + z - y| \leq |x - z| + |z - y|$. **q.e.d.**

Definition 2.23. Sei $M \subset \mathbf{K}$ eine nicht leere Teilmenge von Zahlen.

(i) M heißt nach oben beschränkt, falls es eine Zahl $\beta \in \mathbf{K}$ gibt, so dass $x \leq \beta$ für alle $x \in M$ gilt. Dann heißt β obere Schranke von M .

(ii) M heißt nach unten beschränkt, falls es eine Zahl $\alpha \in \mathbf{K}$ gibt, so dass $\alpha \leq x$ für alle $x \in M$ gilt. Dann heißt α untere Schranke von M .

(iii) M heißt beschränkt, wenn M nach oben und unten beschränkt ist.

Definition 2.24 (Maximum und Minimum). Für eine nicht leere nach oben beschränkte Menge $M \subset \mathbf{K}$ heißt eine obere Schranke m von M , die in M liegt, Maximum. Wir schreiben dann $m = \max M$. Analog heißt eine untere Schranke m einer nicht leeren nach unten beschränkten Menge M , die in M liegt, Minimum. Wir schreiben dann $m = \min M$.

Definition 2.25 (Supremum und Infimum einer Menge). Ein Minimum s der Menge aller oberen Schranken von einer nicht leeren nach oben beschränkte Menge $M \subset \mathbf{K}$ heißt Supremum. Wir schreiben dann $s = \sup M$. Analog heißt ein Maximum t der Menge aller unteren Schranken einer nach unten beschränkten Menge M Infimum. Wir schreiben dann $t = \inf M$.

Bemerkung 2.26. Wenn eine Menge ein Maximum bzw. Minimum besitzt, ist dieses eindeutig, weil es weder größer noch kleiner als ein anderes Maximum bzw. Minimum sein kann. Also sind auch Supremum und Infimum eindeutig, wenn sie existieren.

Folgende Charakterisierung erleichtert manche Beweise:

$$\begin{aligned} s = \sup M &\iff s \text{ ist obere Schranke von } M \text{ und } \forall \epsilon > 0 \quad \exists x \in M \text{ mit } s - \epsilon < x. \\ t = \inf M &\iff t \text{ ist untere Schranke von } M \text{ und } \forall \epsilon > 0 \quad \exists x \in M \text{ mit } x < t + \epsilon. \end{aligned}$$

Insbesondere ist jedes Maximum auch ein Supremum und jedes Minimum ein Infimum.

Warnung 2.27. Das Supremum $\sup M$ bzw. das Infimum $\inf M$ muss im Unterschied zum Maximum bzw. Minimum kein Element von M sein.

Die folgenden Teilmengen eines angeordneten Körpers $\mathbf{K}$ heißen Intervalle. Nachdem wir $\mathbb{R}$ eingeführt haben, bezeichnen diese Symbole nur diese Teilmengen von $\mathbb{R}$:

$$\begin{array}{ll} [a, b] = \{x \in \mathbf{K} \mid a \leq x \leq b\} & \text{für alle } a \leq b \in \mathbf{K} \\ [a, b) = \{x \in \mathbf{K} \mid a \leq x < b\} & \text{für alle } a < b \in \mathbf{K} \\ (a, b] = \{x \in \mathbf{K} \mid a < x \leq b\} & \text{für alle } a < b \in \mathbf{K} \\ (a, b) = \{x \in \mathbf{K} \mid a < x < b\} & \text{für alle } a < b \in \mathbf{K} \\ (-\infty, b] = \{x \in \mathbf{K} \mid x \leq b\} & \text{für alle } b \in \mathbf{K} \\ (-\infty, b) = \{x \in \mathbf{K} \mid x < b\} & \text{für alle } b \in \mathbf{K} \\ [a, \infty) = \{x \in \mathbf{K} \mid a \leq x\} & \text{für alle } a \in \mathbf{K} \\ (a, \infty) = \{x \in \mathbf{K} \mid a < x\} & \text{für alle } a \in \mathbf{K} \end{array}$$

$\mathbf{K}^+ = (0, \infty)$ positive Zahlen

$\mathbf{K}^- = (-\infty, 0)$ negative Zahlen

$\mathbf{K}_0^+ = [0, \infty)$ nichtnegative Zahlen

$\mathbf{K}_0^- = (-\infty, 0]$ nichtpositive Zahlen

Offenbar ist b eine obere Schranke von (a, b) . Andererseits gibt es wegen Satz 2.16 für jedes $x < b$ ein $y = \max\{\frac{x+b}{2}, \frac{a+b}{2}\}$ mit $x < y$ und $y \in (a, b)$. Also gibt es keine obere Schranke von (a, b) , die kleiner ist als b . Damit ist $b = \sup(a, b)$. Analog gilt:

$$\begin{aligned} a = \inf[a, b] &= \inf[a, b] &= \inf(a, b] &= \inf(a, b) &= \inf[a, \infty) &= \inf(a, \infty) \\ b = \sup[a, b] &= \sup[a, b) &= \sup(a, b) &= \sup(a, b) &= \sup(-\infty, b] &= \sup(-\infty, b) \end{aligned}$$

$(-\infty, b]$ und $(-\infty, b)$ sind nicht nach unten beschränkt
 $[a, \infty)$ und (a, ∞) sind nicht nach oben beschränkt

Also existiert $\max[a, b] = \max(a, b] = \max(-\infty, b] = b$
 $\min[a, b] = \min[a, b) = \min[a, \infty) = a$,
während $[a, b)$ und (a, b) und $(-\infty, b)$ kein Maximum besitzen
und $(a, b]$ und (a, b) und (a, ∞) kein Minimum.

2.2 Die natürlichen Zahlen $\mathbb{N} \subset \mathbf{K}$

Wir wollen die natürlichen Zahlen als Teilmenge eines angeordneten Körpers $\mathbf{K}$ charakterisieren. Aufgrund von **A2** gibt es das Element $1 \in \mathbb{N} \subset \mathbf{K} \setminus \{0\}$, das wegen (ii) in Satz 2.8 eindeutig ist. Aus (iv) im Satz 2.14 folgt $1 > 0$ aus $1 = 1 \cdot 1$. Wegen der Monotonie ist dann die Nachfolgerabbildung streng monoton:

$$\begin{aligned} \mathbf{K} &\rightarrow \mathbf{K}, & n &\mapsto n + 1 > n \\ 1 &\mapsto 1 + 1 = 2 > 1, & 2 &\mapsto 2 + 1 = 3 > 2, & \dots & n &\mapsto n + 1 > n, & \dots \end{aligned}$$

Definition 2.28. Eine Menge $M \subset \mathbf{K}$ heißt induktiv, falls

- (i) $1 \in M$ (Einselement).
- (ii) $a \in M \implies a + 1 \in M$ (invariant unter der Nachfolgerabbildung).

$\mathbf{K}$ selber ist offenbar induktiv oder auch $[1, \infty)$. Offenbar ist der Durchschnitt von induktiven Mengen wieder induktiv.

Definition 2.29 (Natürliche Zahlen). $\mathbb{N}$ ist die kleinste induktive Teilmenge von $\mathbf{K}$, also der Durchschnitt aller induktiven Teilmengen von $\mathbf{K}$.

Satz 2.30. Es gilt $\mathbb{N} = \{1\} \cup \{n \in \mathbf{K} \mid n - 1 \in \mathbb{N}\} = \{1\} \cup \text{Bild der Einschränkung der Nachfolgerabbildung auf } \mathbb{N}$.

Beweis: Wegen $(1+1) - 1 = 1$, $(n+1) - 1 = (n-1) + 1$ und weil $\mathbb{N}$ eine induktive Menge ist, ist auch $S = \{1\} \cup \{n \in \mathbf{K} \mid n-1 \in \mathbb{N}\}$ eine induktive Menge. Also folgt $S \supset \mathbb{N}$. Weil $\mathbb{N}$ eine induktive Menge ist, folgt $S \subset \mathbb{N}$ aus $n = (n-1) + 1$. **q.e.d.**

Satz 2.31 (Prinzip der vollständigen Induktion). *Eine Teilmenge $S \subset \mathbb{N}$ erfülle*
 (i) $1 \in S$ (ii) $a \in S \implies a+1 \in S$. *Dann ist $S = \mathbb{N}$.*

Beweis: S ist offenbar eine induktive Menge. Also gilt $\mathbb{N} \subset S$. Andererseits ist S eine Teilmenge von $\mathbb{N}$. Also folgt $\mathbb{N} = S$. **q.e.d.**

Beweis durch vollständige Induktion: Um eine Aussage $A(n)$ über alle natürlichen Zahlen $n \in \mathbb{N}$ zu beweisen genügt es also zu zeigen:

- (i) die Aussage $A(1)$ ist richtig.
- (ii) Falls die Aussage $A(n)$ richtig ist, dann auch $A(n+1)$.

Ein solcher Beweis heißt Beweis durch vollständige Induktion.

Satz 2.32 (Bernoulli Ungleichung). *Sei x^n für alle $x \in \mathbf{K}$ und $n \in \mathbb{N}$ das n -fache Produkt von x mit sich selber und $x^0 = 1$. Für $x \in \mathbf{K} \setminus \{0\}$ sei $x^{-n} = 1/(x^n)$. Dann gilt*

$$(1+x)^n \geq 1+nx \quad \text{für alle } x \in (-1, \infty) \text{ und alle } n \in \mathbb{N}.$$

Gleichheit gilt nur für $x = 0$ oder $n = 1$.

Beweis durch vollständige Induktion: Für $x = 0$ oder $n = 1$ gilt offenbar die Gleichheit. Für jede natürliche Zahl $n \in \mathbb{N}$ sei $A(n)$ die Aussage

$$\text{Für alle } x \in (-1, 0) \cup (0, \infty) \quad \text{gilt} \quad (1+x)^{n+1} > 1 + (n+1)x.$$

- (i) Wenn $x \neq 0$ dann folgt $(1+x)^2 = 1+2x+x^2 > 1+2x$. Also gilt $A(1)$.
- (ii) Es gelte $A(n)$. Dann folgern wir für $x \in (-1, 0) \cup (0, \infty)$ aufgrund der Monotonie:

$$\text{Wegen } x > -1 \text{ gilt } 1+x > 0, \quad \text{und wegen } x \neq 0 \text{ folgt daraus}$$

$$(1+x)(1+x)^{n+1} > (1+x)(1+(n+1)x) = 1+(n+2)x+(n+1)x^2 > 1+(n+2)x.$$

Also gilt $A(n+1)$. **q.e.d.**

Satz 2.33. *Für alle $n \in \mathbb{N}$ gibt es keine Zahl in $\mathbb{N}$ zwischen $n-1$ und n .*

Beweis durch vollständige Induktion:

- (i) $[1, \infty)$ ist eine induktive Menge. Also ist $\mathbb{N} \cap (0, 1) = \mathbb{N} \cap [1, \infty) \cap (0, 1) = \emptyset$.
- (ii) Wir nehmen an, dass es für ein $n \in \mathbb{N}$ keine natürliche Zahl in $(n-1, n)$ gibt. Wenn es eine Zahl $m \in \mathbb{N} \cap (n, n+1)$ gibt, dann liegt m wegen Satz 2.30 auch in $\{1\} \cup \{x \in \mathbf{K} \mid x-1 \in \mathbb{N}\}$. Wegen $1 \leq n < m$ ist m nicht gleich 1. Also liegt m dann in $\{n \in \mathbf{K} \mid n-1 \in \mathbb{N}\}$. Wegen der Monotonie liegt $m-1$ dann in $\mathbb{N} \cap (n-1, n)$, was der Annahme widerspricht. Also gibt es keine natürliche Zahl in $\mathbb{N} \cap (n, n+1)$. **q.e.d.**

Für alle $n \in \mathbb{N}$ ist also $\{m \in \mathbb{N} \mid m \leq n+1\} = \{m \in \mathbb{N} \mid m \leq n\} \cup \{n+1\}$. Für alle $n \in \mathbb{N}$ schreiben wir deshalb $\{1, \dots, n\} = \{m \in \mathbb{N} \mid m \leq n\}$.

Satz 2.34 (Wohlordnungsprinzip). *Für jede nichtleere Menge $M \subset \mathbb{N}$ existiert $\min M$.*

Beweis: Wenn $n \in M \subset \mathbb{N}$, dann ist offenbar jedes Minimum von $M \cap \{1, \dots, n\}$ auch ein Minimum von M und umgekehrt. Deshalb zeigen wir mit vollständiger Induktion, dass für alle $n \in \mathbb{N}$ jede nichtleere Teilmenge $M \subset \{1, \dots, n\}$ ein Minimum besitzt:

- (i) Jede nichtleere Teilmenge von $\{1\}$ ist gleich $\{1\}$ mit Minimum $\min\{1\} = 1$.
- (ii) Für jede nichtleere Teilmenge $M \subset \{1, \dots, n+1\}$ ist entweder $M \cap \{1, \dots, n\}$ nichtleer oder $M = \{n+1\}$. Wenn also jede nichtleere Teilmenge von $\{1, \dots, n\}$ ein Minimum hat, dann auch jede nichtleere Teilmenge von $\{1, \dots, n+1\}$. **q.e.d.**

Für zwei verschiedene angeordnete Körper $\mathbf{K}$ und $\mathbf{K}'$ gibt es eine eindeutige Abbildung, die $1_{\mathbf{K}}$ auf $1_{\mathbf{K}'}$ abbildet, $1_{\mathbf{K}} + 1_{\mathbf{K}}$ auf $1_{\mathbf{K}'} + 1_{\mathbf{K}'}$, $1_{\mathbf{K}} + 1_{\mathbf{K}} + 1_{\mathbf{K}}$ auf $1_{\mathbf{K}'} + 1_{\mathbf{K}'} + 1_{\mathbf{K}'}$ usw. Wegen dem Prinzip der vollständigen Induktion ist der Definitionsbereich $\mathbb{N}_{\mathbf{K}}$, und das Bild ist $\mathbb{N}_{\mathbf{K}'}$. Also sind die natürlichen Zahlen in dem Sinne eindeutig, dass $\mathbb{N}_{\mathbf{K}}$ auf eindeutige Weise mit $\mathbb{N}_{\mathbf{K}'}$ identifiziert wird, wenn $1_{\mathbf{K}}$ mit $1_{\mathbf{K}'}$ und Nachfolger mit Nachfolgern identifiziert werden. Analoge Aussagen gelten für folgende Mengen:

Definition 2.35 (ganze Zahlen $\mathbb{Z}$, rationale Zahlen $\mathbb{Q}$).

$$\mathbb{N}_0 = \mathbb{N} \cup \{0\}, \quad \mathbb{Z} = \mathbb{N}_0 \cup \{x \in K \mid -x \in \mathbb{N}\}, \quad \mathbb{Q} = \left\{ \frac{m}{n} \in \mathbf{K} \mid m \in \mathbb{Z}, n \in \mathbb{N} \right\}.$$

2.3 Mächtigkeit von Mengen

Definition 2.36. *Zwei Mengen heißen gleichmächtig, wenn es eine bijektive Abbildung zwischen ihnen gibt.*

Diese Relation von Mengen ist eine Äquivalenzrelation. Wir zeigen mit vollständiger Induktion, dass $\{1, \dots, n\}$ für alle $n \in \mathbb{N}$ nicht gleichmächtig zu einer echten Teilmenge von sich ist: (i) Die einzige echte Teilmenge von $\{1\}$ ist leer. (ii) Eine bijektive Abbildung f von $\{1, \dots, n+1\}$ auf eine echte Teilmenge von $\{1, \dots, n+1\}$ bildet $\{1, \dots, n\}$ bijektiv auf eine echte Teilmenge von $\{1, \dots, n+1\} \setminus \{f(n+1)\}$ ab, also auf eine zu einer echten Teilmenge von $\{1, \dots, n\}$ gleichmächtige Menge. Das zeigt die Aussage, also insbesondere, dass $\{1, \dots, n\}$ und $\{1, \dots, m\}$ nur für $n = m$ gleichmächtig sind.

Wenn wir $1 \in \mathbb{N}$ mit der Äquivalenzklasse von $\{\emptyset\}$ indentifizieren und für Mengen die Nachfolgerabbildung $A \mapsto A \cup \{A\}$ definieren, werden die natürlichen Zahlen mit den Äquivalenzklassen folgender Mengen identifiziert:

$$\begin{array}{lll} 1 & \leftrightarrow & \{\emptyset\} \\ 2 & \leftrightarrow & \{\emptyset, \{\emptyset\}\} \\ 3 & \leftrightarrow & \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\} \\ 4 & \leftrightarrow & \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}, \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}\} \\ \vdots & & \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}, \dots, \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}, \dots, \}\} \end{array}$$

Die Menge aller Äquivalenzklassen kann man als eine Erweiterung von $\mathbb{N}$ betrachten.

Definition 2.37. Eine nicht leere Menge A heißt

endlich, falls es ein $n \in \mathbb{N}$ gibt, so dass A gleichmächtig ist zu $\{1, 2, \dots, n\}$,

unendlich, falls sie nicht endlich ist,

abzählbar, falls sie gleichmächtig ist zu $\mathbb{N}$,

höchstens abzählbar, wenn sie endlich oder abzählbar ist,

überabzählbar, wenn sie weder endlich noch abzählbar ist.

Satz 2.38. (i) Jede nichtleere Teilmenge von $\mathbb{N}$ ist höchstens abzählbar.

(ii) Eine Menge A ist genau dann höchstens abzählbar, wenn es eine surjektive Abbildung von $\mathbb{N}$ auf A gibt.

Beweis: (i) Mit dem Wohlordnungsprinzip nummerieren wir der Größe nach mit dem kleinsten Element anfangend jede Teilmenge $M \subset \mathbb{N}$ durch. Für alle $m \in M$ ist $\{1, \dots, m\} \cap M$ gleichmächtig zu $\{1, \dots, n\}$ für ein $n \in \{1, \dots, m\}$. Wenn M nach oben unbeschränkt ist, erhalten wir eine bijektive Abbildung von $\mathbb{N}$ nach M und M ist abzählbar. Andernfalls ist $M \subset \{1, \dots, m\}$ für ein $m \in M$ und damit endlich.

(ii) Sei A eine höchstens abzählbare Menge. Wenn A endlich ist, gibt es ein $n \in \mathbb{N}$, so dass A gleichmächtig ist zu $\{1, 2, \dots, n\}$. Die Abbildung $\mathbb{N} \rightarrow \{1, 2, \dots, n\}$, $m \mapsto \min\{m, n\}$ ist eine surjektive Abbildung, so dass dann auch eine surjektive Abbildung auf A existiert. Wenn A abzählbar ist, gibt es eine bijektive Abbildung von $\mathbb{N}$ auf A .

Sei umgekehrt A eine Menge und f eine surjektive Abbildung von $\mathbb{N}$ auf A . Dann existiert eine Abbildung $g : A \rightarrow \mathbb{N}$, $a \mapsto \min f^{-1}[\{a\}]$, die offenbar eine bijektive Abbildung von A auf eine Teilmenge von $\mathbb{N}$ definiert. Also ist A gleichmächtig zu einer Teilmenge von $\mathbb{N}$ und damit wegen (i) höchstens abzählbar. **q.e.d.**

Übungsaufgabe 2.39. Zeige, dass jede endliche Teilmenge der reellen Zahlen ein Maximum und ein Minimum besitzt.

Satz 2.40. (i) Die Menge $\mathbb{N} \times \mathbb{N}$ ist abzählbar.

(ii) Eine höchstens abzählbare Vereinigung von höchstens abzählbaren Mengen ist wieder höchstens abzählbar.

Beweis: (i) Wir definieren die sogenannte Diagonalnummerierung als die Abbildung

$$f : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}, \quad (x, y) \mapsto f(x, y) = y + \frac{(x + y - 2)(x + y - 1)}{2} = y + \sum_{n=0}^{x+y-2} n.$$

Diese Abbildung ist injektiv. Gilt nämlich $x + y < x' + y'$ so folgt aus

$$f(x, y) - y \leq f(x', y') - y' - (x' + y' - 2) \quad \text{und} \quad y - y' < y \leq x + y - 1 \leq x' + y' - 2$$

$$f(x, y) \leq f(x', y') + y - y' - (x' + y' - 2) < f(x', y').$$

Gilt aber $x + y = x' + y'$ so gilt auch $f(x, y) - f(x', y') = y - y'$.

Also ist $\mathbb{N} \times \mathbb{N}$ gleichmächtig zum Bild von f . Wegen $f(x, y) \geq y$ ist das Bild nicht beschränkt und nicht endlich. Wegen (i) vom vorangehenden Satz ist $\mathbb{N} \times \mathbb{N}$ abzählbar.

(ii) Wegen (ii) im vorangehenden Satz gibt es eine surjektive Abbildung $n \mapsto A_n$ von $\mathbb{N}$ auf die höchstens abzählbar vielen Mengen. Wegen (ii) im vorangehenden Satz gibt es dann für alle $n \in \mathbb{N}$ eine surjektive Abbildung $f_n : \mathbb{N} \rightarrow A_n$. Dann ist

$$F : \mathbb{N} \times \mathbb{N} \rightarrow \bigcup_{n \in \mathbb{N}} A_n, \quad (n, m) \mapsto f_n(m)$$

eine surjektive Abbildung. Also folgt (ii) aus (i) und dem vorangehenden Satz. **q.e.d.**

Korollar 2.41. $\mathbb{Q}$ ist abzählbar.

Beweis: $\mathbb{Z} \times \mathbb{N} \rightarrow \mathbb{Q}, \quad (m, n) \mapsto \frac{m}{n}$ ist eine surjektive Abbildung. $\mathbb{Z}$ ist abzählbar, also ist $\mathbb{Q}$ höchstens abzählbar. $\mathbb{Q}$ ist aber nicht endlich und damit abzählbar. **q.e.d.**

2.4 $\mathbb{R}$ und das Vollständigkeitsaxiom

A5. Vollständigkeitsaxiom 2.42. Für jede nicht leere nach oben beschränkte Teilmenge M eines angeordneten Körpers $\mathbf{K}$ existiert das Supremum $s = \sup M \in \mathbf{K}$.

Wir werden sehen, dass die rationalen Zahlen $\mathbb{Q}$ **A1-A4** erfüllen aber nicht **A5**. Man kann zeigen, dass zwei angeordnete Körper, für die **A5** gilt, isomorph sind.

Definition 2.43. Sei $\mathbb{R}$ der angeordnete Körper, in dem **A5** gilt.

Satz 2.44 (Folgerungen aus **A5**).

- (i) Jede nicht leere nach unten beschränkte Menge M besitzt ein Infimum $\inf M \in \mathbb{R}$.
- (ii) Eine nicht leere nach oben beschränkte Menge M besitzt genau dann ein Maximum, wenn $\sup M \in M$. In diesem Fall ist $\max M = \sup M$.
- (iii) Eine nicht leere nach unten beschränkte Menge M besitzt genau dann ein Minimum, wenn $\inf M \in M$. In diesem Fall ist $\min M = \inf M$.

Beweis: (i) Weil die Ordnungsrelation nicht symmetrisch ist, erhalten wir dadurch, dass wir in einer Aussage alle auftauchenden Ordnungsrelationen umkehren, eine andere Aussage. Zum Beispiel erhalten wir die Definition der unteren Schranke, indem wir in der Definition der oberen Schranke alle Ordnungsrelationen umdrehen. Dann erhalten wir aber auch die Definition des Infimums, indem wir in der Definition des Supremums alle Ordnungsrelationen umkehren. Weil $x < y \iff -y < -x$, sind also Aussagen

über Ordnungsrelationen zwischen reellen Zahlen äquivalent zu den analogen Aussagen, in denen wir alle Ordnungsrelationen umdrehen und alle reellen Zahlen durch ihre Negativen ersetzen¹. Insbesondere besitzt die Menge M genau dann ein Infimum, wenn die Menge $-M = \{x \in \mathbb{R} \mid -x \in M\}$ ein Supremum besitzt und es gilt $\inf M = -\sup -M$. Also folgt (i) aus dem Vollständigkeitsaxiom **A5**.

(ii) Wenn $\sup M \in M$, dann ist $\sup M$ eine obere Schranke von M , die Element von M ist. Wenn $\sup M \notin M$, dann gilt sogar $x < \sup M \leq s$ für alle $x \in M$ und alle oberen Schranken s von M . Also gibt es in M keine obere Schranke von M .

(iii) analog zu (ii).

q.e.d.

Satz 2.45 (Archimedes-Eudoxos).

(i) Für jede reelle Zahl $x \in \mathbb{R}$ gibt es eine natürliche Zahl $n > x$ (Archimedes).

(ii) Für jedes $\epsilon > 0$ gibt es ein $n \in \mathbb{N}$ mit $\frac{1}{n} < \epsilon$ (Eudoxos).

Beweis: (i) Es reicht zu zeigen, dass $\mathbb{N}$ keine obere Schranke hat. Wenn $\mathbb{N}$ eine obere Schranke hat, dann existiert $\sup \mathbb{N}$, und $\exists n \in \mathbb{N}$ mit $n > \sup \mathbb{N} - 1$, weil $\sup \mathbb{N} - 1$ keine obere Schranke von $\mathbb{N}$ ist. Dann gilt $\mathbb{N} \ni n + 1 > \sup \mathbb{N}$, was ein Widerspruch ist.

(ii) Nach (i) gibt es ein $n > \frac{1}{\epsilon}$. Aus Satz 2.14 (v)-(vi) folgt $\frac{1}{n} < \epsilon$.

q.e.d.

Bemerkung 2.46. In einem angeordneten Körper sind die beiden Eigenschaften (i) und (ii) aus Satz 2.45 wegen Satz 2.14 (v)-(vi) äquivalent. Ein angeordneter Körper heißt archimedisch, wenn (i) oder (ii) gilt. In jedem angeordneten Körper sind die rationalen Zahlen archimedisch. Es gibt auch nicht archimedische angeordnete Körper.

Wegen dem Teil (i) von Archimedes enthalten die Intervalle (a, b) rationale Zahlen:

Satz 2.47. Sei $a < b$. Dann existiert $r \in \mathbb{Q}$ mit $a < r < b$.

Beweis: Für alle $n \in \mathbb{N}$ ist $a < \frac{m}{n} < b$ äquivalent zu $na < m < nb$. Wegen $b - a > 0$ gibt es nach Satz 2.45 (i) ein $n \in \mathbb{N}$ mit $n > \frac{1}{b-a}$. Dann gilt $na < nb$ und $nb - na > 1$.

Wir nehmen zunächst $a \geq 0$ an. Sei m die kleinste natürliche Zahl, die größer ist als na . Wegen Satz 2.30 sind die natürlichen Zahlen enthalten in $\{1\} \cup \{x \in \mathbb{R} \mid x - 1 \in \mathbb{N}\}$. Also ist $m - 1$ entweder eine natürliche Zahl oder gleich Null. Dann folgt

$$m - 1 \leq na < m \iff na < m \leq na + 1 < nb.$$

Als nächstes nehmen wir $b \leq 0$ an. Sei $-m$ die kleinste natürliche Zahl, die größer ist als $-nb$. Wieder ist $-m - 1$ entweder eine natürliche Zahl oder Null. Also folgt

$$-m - 1 \leq -nb < -m \iff na < nb - 1 \leq m < nb.$$

¹Wenn diese Aussagen aber algebraische Operationen benutzen, die nicht verträglich sind mit der Abbildung jeder reellen Zahl auf ihre Negative, wie z. B. das Produkt zweier reeller Zahlen, dann müssen bei dieser Ersetzung auch die algebraischen Operationen entsprechend geändert werden, also z. B. das Produkt in das negative des Produktes.

Wenn $a < 0$ und $b > 0$ ist, wählen wir $m = 0$.

q.e.d.

Wir haben sogar gezeigt, dass $\forall a < b$ mit $b - a > 1$ das Intervall (a, b) eine ganze Zahl enthält. Also enthält jedes Intervall mindestens eine und damit sogar unendlich viele rationale Zahlen. Insbesondere gibt es für jede reelle Zahl x und jedes $\epsilon > 0$ eine rationale Zahl $r \in (x - \epsilon, x + \epsilon)$, die dann $d(x, r) < \epsilon$ erfüllt. Wir sagen, dass $\mathbb{Q}$ dicht in $\mathbb{R}$ liegt. Die rationalen Zahlen erfüllen als Unterkörper der reellen Zahlen die Axiome **A1-A4**. Wir werden gleich sehen, dass sie nicht das Vollständigkeitsaxiom **A5** erfüllen.

Satz 2.48 (Quadratwurzeln). *Für alle $a > 0$ gibt es genau ein $b > 0$ mit $b^2 = a$.*

Wir schreiben $b = \sqrt{a} = a^{\frac{1}{2}}$.

Beweis der Eindeutigkeit: Aus $0 < x < y$ folgt aufgrund der Monotonie $x^2 < xy < y^2$. Also gibt es höchstens ein $b > 0$ mit $b^2 = a$.

Beweis der Existenz: Die Menge $M = \{x \in [0, \infty) \mid x^2 < a\}$ enthält 0. Dann gilt

$$x^2 > x(a+1) > (a+1)^2 = a^2 + 2a + 1 > 2a > a \quad \text{für } x > a+1.$$

Also ist $a+1$ eine obere Schranke von M . Sei $b = \sup M$. Aus $0 \in M$ folgt $0 \leq b$.

Wenn $b^2 < a$, folgt $0 < a - b^2$. Für $h = \min\{1, \frac{a-b^2}{2b+2}\}$ gilt also $0 < h \leq 1$ und $h(2b+2) \leq a - b^2$ und damit auch

$$(b+h)^2 = b^2 + 2bh + h^2 \leq b^2 + 2bh + h = b^2 + h(2b+1) < b^2 + h(2b+2) \leq b^2 + a - b^2 = a.$$

Also ist $b < b+h \in M$ im Widerspruch zu $b = \sup M$.

Wenn $a < b^2$, folgt $0 < b$, $-b^2 < a - b^2 < 0$ und $-\frac{1}{2} < \frac{a-b^2}{2b^2} < 0$. Also folgt aus der Bernoulli Ungleichung

$$\left(b \left(1 + \frac{a-b^2}{2b^2}\right)\right)^2 = b^2 \left(1 + \frac{a-b^2}{2b^2}\right)^2 \geq b^2 \left(1 + 2\frac{a-b^2}{2b^2}\right) = a.$$

Dann folgt aus $x > b(1 + \frac{a-b^2}{2b^2})$ wieder $x^2 > b^2(1 + \frac{a-b^2}{2b^2})^2 \geq a$ und damit auch $x \notin M$.

Also ist $b(1 + \frac{a-b^2}{2b^2}) < b$ eine obere Schranke von M im Widerspruch zu $b = \sup M$.

Weil also weder $b^2 < a$ noch $a < b^2$ gilt, folgt $b^2 = a$ und $b \neq 0$.

q.e.d.

Insbesondere gibt es also genau ein $\sqrt{2} \in \mathbb{R}$. Wir werden in Beispiel 3.9 für jedes $n \in \mathbb{N}$ zeigen, dass es für jedes $a > 0$ genau ein $b > 0$ gibt mit $b^n = a$.

Lemma 2.49. $\sqrt{2} \notin \mathbb{Q}$. *Insbesondere erfüllt $\mathbb{Q}$ nicht das Vollständigkeitsaxiom **A5**.*

Beweis: Wir zeigen, dass es nicht zwei natürliche Zahlen $n, m \in \mathbb{N}$ geben kann mit $(\frac{m}{n})^2 = 2$. Wenn es zwei solche Zahlen gibt, dann sei n_0 das Minimum der Teilmenge

$$M = \{n \in \mathbb{N} \mid \exists m \in \mathbb{N} \text{ mit } m^2 = 2n^2\}.$$

Also gibt es ein $m_0 \in \mathbb{N}$ mit $m_0^2 = 2n_0^2$. Dann enthält $(\frac{m_0-1}{2}, \frac{m_0+2}{2})$ eine positive ganze Zahl m_1 . Aus $m_0 \in (2m_1 - 2, 2m_1 + 1)$ folgt entweder $m_0 = 2m_1 - 1$ oder $m_0 = 2m_1$.

Im ersten Fall widerspricht $n_0^2 = 2m_1^2 - m_1 + \frac{1}{2} \in (2m_1^2 - m_1, 2m_1^2 - m_1 + 1)$ dem Satz 2.33. Im zweiten Fall ist $n_0^2 = 2m_1^2$, also $m_1 \in M$. Dann folgt $n_0 = \inf M \leq m_1$ und $n_0^2 \leq m_1^2$ im Widerspruch zu $m_1^2 < 2m_1^2 = n_0^2$. **q.e.d.**

Satz 2.50 (Intervallschachtelungsprinzip). *Seien $I_n = [a_n, b_n] \subset \mathbb{R}$ für $n = 1, 2, \dots$ abgeschlossene Intervalle mit $a_n < b_n$ und den beiden folgenden Eigenschaften:*

- (i) $I_{n+1} \subset I_n$ für alle $n \in \mathbb{N}$.
- (ii) Für jedes $\epsilon > 0$ gibt es ein $n \in \mathbb{N}$, so dass $b_n - a_n < \epsilon$.

Dann enthält der Durchschnitt $\bigcap_{n \geq 1} I_n = \{x\}$ genau ein $x \in \mathbb{R}$.

Dieser Satz ist für offene Intervalle falsch. Z.B. $\bigcap_{n \geq 1} (0, \frac{1}{n}) = \emptyset$ nach Satz 2.45.

Beweis: Wegen (i) gilt

$$a_1 \leq a_2 \leq \dots \leq a_n \leq a_{n+1} \leq \dots \leq b_{n+1} \leq b_n \leq \dots \leq b_2 \leq b_1$$

Also besteht die Menge $A = \{a_1, a_2, \dots\}$ aus unteren Schranken der Menge $B = \{b_1, b_2, \dots\}$ und umgekehrt die Menge B aus oberen Schranken der Menge A . Dann liegen $x = \sup A$ und $y = \inf B$ in der Schnittmenge der oberen Schranken von A mit den unteren Schranken von B , was gleich $\bigcap_{n \geq 1} I_n$ ist. Dann folgt $x \leq y$ und $\bigcap_{n \geq 1} I_n = [x, y]$. Andererseits gibt es wegen (ii) für alle $\epsilon > 0$ ein $n \in \mathbb{N}$ mit $y - x \leq b_n - a_n < \epsilon$. Dann muss aber $0 \leq y - x \leq \inf \{\epsilon \mid \epsilon > 0\} = 0$ und deshalb $y = x$ gelten. **q.e.d.**

Satz 2.51. *In jedem angeordneten archimedischen Körper folgt aus dem Intervallschachtelungsprinzip im Satz 2.50 das Vollständigkeitsaxiom A5.*

Beweis*: Sei M eine nicht leere nach oben beschränkte Menge, a_1 ein Element von M und $b_1 > a_1$ eine obere Schranke von M . Wir definieren induktiv eine Folgen $a_1 \leq a_2 \leq \dots$ von Elementen von M und eine Folge $b_1 \geq b_2 \geq \dots$ von oberen Schranken von M . Wenn $\frac{a_n + b_n}{2}$ keine obere Schranke von M ist, dann gibt es ein Element $a_{n+1} \in [\frac{a_n + b_n}{2}, b_n] \cap M$ und wir setzen $b_{n+1} = b_n$. Wenn $\frac{a_n + b_n}{2}$ eine obere Schranke von M ist, setzen wir $a_{n+1} = a_n$ und $b_{n+1} = \frac{a_n + b_n}{2}$. Offenbar gilt

$$0 \leq b_{n+1} - a_{n+1} \leq 2^{-1}(b_n - a_n) \leq 2^{-2}(b_{n-1} - a_{n-1}) \leq \dots \leq 2^{-n}(b_1 - a_1).$$

Aus der Bernoulli Ungleichung folgt $2^n > n$ und wegen Satz 2.45 (ii) gibt es für alle $\epsilon > 0$ ein $n \in \mathbb{N}$ mit $2^{-n}(b_1 - a_1) < \frac{b_1 - a_1}{n} < \epsilon$. Wegen dem Intervallschachtelungsprinzip besteht die Schnittmenge der Intervalle $\bigcap_{n \in \mathbb{N}} [a_n, b_n]$ aus einer Zahl $s \in \mathbb{R}$.

Für jede Zahl $x > s$ gibt es ein $n \in \mathbb{N}$ mit $\frac{x-s}{b_1 - a_1} > \frac{1}{n} > 2^{-n}$. Daraus folgt $b_{n+1} \leq a_{n+1} + 2^{-n}(b_1 - a_1) < s + (x - s) = x$. Weil b_{n+1} eine obere Schranke von M ist, folgt $x \notin M$. Also ist s eine obere Schranke von M . Für jede Zahl $x < s$ gibt es ein $n \in \mathbb{N}$ mit $\frac{s-x}{b_1 - a_1} > \frac{1}{n} > 2^{-n}$. Daraus folgt $a_{n+1} \geq b_{n+1} - 2^{-n}(b_1 - a_1) > s - (s - x) = x$. Wegen $a_{n+1} \in M$ ist x keine obere Schranke von M . Also ist $s = \sup M$. **q.e.d.**

Deshalb können die reellen Zahlen auch als angeordneter archimedischer Körper charakterisiert werden, in dem das Intervallschachtelungsprinzip gilt.

Satz 2.52. *Die Menge der reellen Zahlen ist überabzählbar.*

Beweis: Wir nehmen an, dass $\mathbb{R}$ abzählbar ist. Sei $(x_n)_{n \in \mathbb{N}}$ eine Durchnumerierung von $\mathbb{R}$. Wähle $a, b \in \mathbb{R}$ mit $a < b$. Sei $I_1 = [a, b]$. Wir definieren induktiv eine Folge $(I_n)_{n \in \mathbb{N}}$ von Intervallen. Für alle $n \in \mathbb{N}$ wählen wir in $I_n = [a_n, b_n]$ als I_{n+1} eins der beiden schnittfremden Teilintervalle $[a_n, a_n + \frac{b_n - a_n}{3}]$ und $[b_n - \frac{b_n - a_n}{3}, b_n]$, das x_n nicht enthält. Für alle $n \in \mathbb{N}$ liegt dann x_n nicht in I_{n+1} . Wegen der Bernoulli Ungleichung 2.32 gilt

$$b_{n+1} - a_{n+1} = \frac{b_n - a_n}{3} = \dots = \frac{b_1 - a_1}{3^n} \leq \frac{b - a}{1 + 2n} < \frac{b - a}{2n} < \epsilon \text{ für } \epsilon > 0 \text{ und } n > \frac{b - a}{2\epsilon}.$$

Wegen dem Intervallschachtelungsprinzip enthält $\bigcap_{n \in \mathbb{N}} I_n$ ein Element, das nicht zu $\{x_n \mid n \in \mathbb{N}\}$ gehört. Das widerspricht der Annahme, dass $\mathbb{R}$ abzählbar ist. **q.e.d.**

Auch die Potenzmenge der natürlichen Zahlen ist nicht abzählbar. Wir können sie mit den Folgen identifizieren, die Werte in $\mathbb{Z}_2 = \{0, 1\}$ annehmen, also der Menge aller Abbildungen von $\mathbb{N}$ nach $\mathbb{Z}_2$. Diese Folgen werden wir mithilfe der dyadischen Entwicklung mit der Teilmenge $[0, 1) \subset \mathbb{R}$ identifizieren. Diese ist gleichmächtig zu $\mathbb{R}$.

2.5 Die erweiterte Zahlengerade $\bar{\mathbb{R}}$

Definition 2.53. *Die erweiterte Zahlengerade besteht aus $\bar{\mathbb{R}} = \mathbb{R} \cup \{+\infty\} \cup \{-\infty\}$.*

Mit ∞ bezeichnen wir auch $+\infty$. Die Ordnungsrelation läßt sich auf $\bar{\mathbb{R}}$ durch $-\infty < x < \infty$ für alle $x \in \mathbb{R}$ fortsetzen und erfüllt offensichtlich auch (i) und (ii) des Ordnungsaxioms. Die Operationen $+$ und $\cdot$ lassen sich nicht auf $\bar{\mathbb{R}}$ fortsetzen, so dass **A1-A4** gilt: Für alle $x, y \in \mathbb{R}$ würde aus $y - x < \infty$ wegen der Monotonie $y < \infty + x$ folgen. Deshalb müsste $\infty + x = \infty$ gelten. Aus Satz 2.6 (ii) würde $x = 0$ folgen. $\bar{\mathbb{R}}$ ist also kein angeordneter Körper.

Die Definitionen von oberen und unteren Schranken, von Supremum und Infimum, und von Maximum und Infimum übertragen sich auf die erweiterte Zahlengerade. Offenbar ist ∞ das Maximum und $-\infty$ das Minimum von $\bar{\mathbb{R}}$. Insbesondere ist ∞ eine obere Schranke und $-\infty$ eine untere Schranke von jeder Teilmenge von $\bar{\mathbb{R}}$. Weil ∞ die einzige obere Schranke einer nicht leeren Teilmenge M von $\mathbb{R} \subset \bar{\mathbb{R}}$ ist, die in $\mathbb{R}$ keine obere Schranke hat, ist ∞ dann das Supremum von M als Teilmenge von $\bar{\mathbb{R}}$. Analog ist $-\infty$ das Infimum einer nicht leeren Teilmenge von $\mathbb{R} \subset \bar{\mathbb{R}}$, die keine untere Schranke in $\mathbb{R}$ hat. In $\bar{\mathbb{R}}$ gilt $\sup \emptyset = -\infty$ und $\inf \emptyset = \infty$. Daraus folgt, dass jede Teilmenge von $\bar{\mathbb{R}}$ ein Supremum und ein Infimum hat. Außerdem gilt wieder, dass eine Teilmenge $M \subset \bar{\mathbb{R}}$ genau dann ein Maximum bzw. Infimum hat, wenn $\sup M \in M$ bzw. $\inf M \in M$ gilt. In diesen Fällen ist wieder $\max M = \sup M$ bzw. $\min M = \inf M$. Wir benutzen die Symbole $\sup$, $\inf$, $\max$ und $\min$ sowohl für Teilmengen von $\mathbb{R}$ als auch für Teilmengen von $\bar{\mathbb{R}}$, so dass aus dem Zusammenhang klar werden muss, was genau gemeint ist.

2.6 Die Peano Axiome

Es gibt andere Möglichkeiten die reellen Zahlen einzuführen. Man kann zuerst die natürlichen Zahlen einführen und danach der Reihe nach die ganzen Zahlen, die rationalen Zahlen und die reellen Zahlen. Dabei kann man die natürlichen Zahlen im Rahmen der Mengenlehre einführen oder sie durch die Peano Axiome charakterisieren:

1. **Peano Axiom:** $1 \in \mathbb{N}$ (Einselement).
2. **Peano Axiom:** $\forall n \in \mathbb{N} \quad \exists$ genau einen Nachfolger $n' \in \mathbb{N}$ (Nachfolgerabbildung).
3. **Peano Axiom:** Für alle $n \in \mathbb{N}$ gilt $n' \neq 1$ (Eins $\notin$ Bild der Nachfolgerabbildung).
4. **Peano Axiom:** Für alle $n, m \in \mathbb{N}$ folgt $n = m$ aus $n' = m'$ (Injektivität der Nachfolgerabbildung).
5. **Peano Axiom:** Jede induktive Teilmenge $S \subset \mathbb{N}$, d.h. jede Teilmenge mit folgenden Eigenschaften umfaßt die natürlichen Zahlen $\mathbb{N} \subset S$:
 - (i) $1 \in S$.
 - (ii) Falls $n \in S$, dann ist auch $n' \in S$.

Man kann die natürlichen Zahlen durch diese Axiome charakterisieren und damit die reellen Zahlen einführen. Dafür muss viel Schlussfolgerungsarbeit geleistet werden, bevor die reellen Zahlen eingeführt sind (vgl. E. Landau: Grundlagen der Analysis). Wir skizzieren in diesem Abschnitt, wodurch sich diese zweite Einführung der reellen Zahlen, von der von uns gewählten unterscheidet, und stellen sie in ihren Grundzügen vor.

Die Addition von natürlichen Zahlen $n + m$ wird $\forall n, m \in \mathbb{N}$ so eingeführt, dass (i) $n + 1 = n'$ für alle $n \in \mathbb{N}$ und (ii) $n + m' = (n + m)'$ für alle $n, m \in \mathbb{N}$ gilt. Für $n = 1$ gilt $1 + 1 = 1'$ und für alle $m \in \mathbb{N}$ folgt aus $1 + m = m'$ auch $1 + m' = (m')' = (1 + m)'$. Wegen dem 5. Peano Axiom ist also $1 + m$ für alle $m \in \mathbb{N}$ durch m' definiert. Wenn für ein $n \in \mathbb{N}$ durch diese Bedingungen $n + m$ für alle $m \in \mathbb{N}$ definiert ist, dann folgt aus (i) $n' + 1 = (n')' = (n + 1)'$. Wenn $n' + m$ für ein $m \in \mathbb{N}$ definiert ist, dann folgt aus (ii) $n' + m' = (n' + m)'$ also ist dann wegen dem 5. Peano Axiom $n' + m$ für alle $m \in \mathbb{N}$ definiert. Also ist wegen dem 5. Peano Axiom die Menge aller $n \in \mathbb{N}$, für die $n + m$ durch diese Bedingungen für alle $m \in \mathbb{N}$ definiert ist gleich $\mathbb{N}$. Deshalb ist dadurch die Addition zwischen allen natürlichen Zahlen definiert. Mit Hilfe der Peano Axiome kann man dann folgern, dass die Addition kommutativ und assoziativ ist.

Die Multiplikation von natürlichen Zahlen nm wird $\forall n, m \in \mathbb{N}$ so eingeführt, dass (i) $n1 = n$ für alle $n \in \mathbb{N}$ und (ii) $nm' = n \cdot m + n$ für alle $n, m \in \mathbb{N}$ gilt. Wieder kann man aus den Peano Axiomen folgern, dass diese Bedingungen eine Multiplikation $n \cdot m$ für alle $n, m \in \mathbb{N}$ definiert. Danach zeigt man, dass auch die Multiplikation kommutativ, assoziativ und zusammen mit der Addition distributiv ist.

Die Ordnungsrelation wird auf den natürlichen Zahlen durch $n < n + m$ und $m < n + m \quad \forall n, m \in \mathbb{N}$ eingeführt. Sie erfüllt **A4** wobei alle $n \in \mathbb{N}$ als positiv gelten.

Ganze Zahlen: Die ganzen Zahlen werden als formale Differenzen $n - m$ eingeführt. Wegen $n - m = \tilde{n} - \tilde{m} \Leftrightarrow n + \tilde{m} = \tilde{n} + m$ sind das Äquivalenzklassen in $\mathbb{N} \times \mathbb{N}$

$$(n, m) \sim (\tilde{n}, \tilde{m}) \quad \Longleftrightarrow \quad n + \tilde{m} = \tilde{n} + m \quad \forall (n, m), (\tilde{n}, \tilde{m}) \in \mathbb{N} \times \mathbb{N}$$

Die Addition, Multiplikation und Ordnungsrelation von $\mathbb{N}$ wird auf $\mathbb{Z}$ fortgesetzt.

Rationale Zahlen: Die rationalen Zahlen werden als formale Brüche $\frac{m}{n}$ mit $(m, n) \in \mathbb{Z} \times \mathbb{N}$ eingeführt. Wegen $\frac{m}{n} = \frac{\tilde{m}}{\tilde{n}} \Leftrightarrow m\tilde{n} = \tilde{m}n$ sind das Äquivalenzklassen in $\mathbb{Z} \times \mathbb{N}$

$$(m, n) \sim (\tilde{m}, \tilde{n}) \quad \Longleftrightarrow \quad m\tilde{n} = \tilde{m}n \quad \forall (m, n), (\tilde{m}, \tilde{n}) \in \mathbb{Z} \times \mathbb{N}$$

Addition, Multiplikation, Ordnungsrelation wird auf $\mathbb{Q}$ definiert, so dass **A1-A4** gilt.

Reelle Zahlen werden als Dedekindsche Schnitte eingeführt, also Mengen $A \subset \mathbb{Q}$ mit

(i) $A \neq \emptyset, \mathbb{Q}$ (ii) $p \in A$ und $q < p \Rightarrow q \in A$. (iii) A hat kein Maximum.

Dabei entspricht $a \in \mathbb{R}$ dem Dedekindschen Schnitt $A = (-\infty, a) \cap \mathbb{Q}$ mit $a = \sup A$.

Addition, Multiplikation, Ordnungsrelation wird auf $\mathbb{R}$ definiert, so dass **A1-A5** gilt.

2.7 Der Körper der komplexen Zahlen

Motivation: Wir hatten aus der Ordnungsrelation gefolgert, dass $x^2 > 0$ gilt, falls $x \neq 0$. Deshalb existieren in den reellen Zahlen keine Quadratwurzeln von negativen Zahlen. Erweitert man die reellen Zahlen durch eine Quadratwurzel i von -1 , so erhält man die komplexen Zahlen, in denen sich dann alle algebraischen Gleichungen lösen lassen.

Definition 2.54 (Komplexe Zahlen). Die Komplexen Zahlen $\mathbb{C}$ ist die Menge $\mathbb{R} \times \mathbb{R}$ aller geordneten Paare (x, y) von reellen Zahlen zusammen mit den Operationen

$$\begin{aligned} + : \quad & \mathbb{C} \times \mathbb{C} \rightarrow \mathbb{C}, & ((x, y), (u, v)) & \mapsto (x, y) + (u, v) = (x + u, y + v) \\ \cdot : \quad & \mathbb{C} \times \mathbb{C} \rightarrow \mathbb{C}, & ((x, y), (u, v)) & \mapsto (xu - yv, xv + yu) \end{aligned}$$

Mit dieser Definition erfüllt $\mathbb{C}$ die Körperaxiome **A1-A3**. Dabei ist

$$\text{Null :} \quad 0_{\mathbb{C}} = (0, 0) \quad \text{Eins :} \quad 1_{\mathbb{C}} = (1, 0)$$

$$\text{negatives Element :} \quad -(x, y) = (-x, -y)$$

$$\text{inverses Element :} \quad (x, y)^{-1} = \left(\frac{x}{x^2 + y^2}, \frac{-y}{x^2 + y^2} \right) \quad \text{für } (x, y) \neq (0, 0).$$

Wir bezeichnen komplexe Zahlen meistens durch einen Buchstaben, üblicherweise z . Die Null und die Eins bezeichnen wir auch durch 0 und 1, so dass aus dem Zusammenhang

klar werden muss, ob es sich um die Null bzw. Eins der reellen oder komplexen Zahlen handelt. Außerdem benutzen wir dieselben Abkürzungen wie bei den reellen Zahlen:

$$z + (-z) = z - z = 0 \qquad z^{-1} = \frac{1}{z} \qquad z \cdot \frac{1}{z} = 1 \qquad \text{usw.}$$

Da die komplexen Zahlen eine Erweiterung der reellen Zahlen sind, wollen wir $\mathbb{R}$ als Teilmenge von $\mathbb{C}$ auffassen. Dafür definieren wir eine injektive Abbildung

$$\mathbb{R} \hookrightarrow \mathbb{C}, \quad x \mapsto (x, 0)$$

Offenbar ist diese Abbildung verträglich mit den Operationen $+$ und $\cdot$ von $\mathbb{R}$ und $\mathbb{C}$:

$$\begin{aligned} (x, 0) + (y, 0) &= (x + y, 0) & (x, 0) \cdot (y, 0) &= (xy, 0) \\ (0, 0) &= 0 & (1, 0) &= 1 \\ -(x, 0) &= (-x, 0) & (x, 0)^{-1} &= (x^{-1}, 0) \end{aligned}$$

Definition 2.55 (Imaginäre Einheit). $\iota = (0, 1) \in \mathbb{C}$ mit $\iota^2 = (-1, 0) = -1$.

Wir schreiben also $(x, y) = x + \iota y$. Dann ergeben sich die Operationen $+$ und $\cdot$

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

Für die komplexe Zahl $z = x + \iota y$ heißt

$$\begin{aligned} x &= \Re(z) \text{ Realteil von } z & y &= \Im(z) \text{ Imaginärteil von } z \\ z &\text{ heißt reell, falls } \Im(z) = 0 & z &\text{ heißt imaginär, falls } \Re(z) = 0 \end{aligned}$$

Definition 2.56 (komplexe Konjugation). Die Abbildung $\mathbb{C} \rightarrow \mathbb{C}$, $z = x + \iota y \mapsto \bar{z} = x - \iota y$ heißt *komplexe Konjugation* oder einfach nur *Konjugation*

Die komplexen Zahlen erhalten wir aus den reellen Zahlen, indem wir reelle Vielfache einer Wurzel aus -1 hinzufügen. Weil aber $(-1) \cdot (-1) = 1$ ist das Negative einer Wurzel aus -1 wieder eine Wurzel aus -1 . Welche dieser beiden Wurzeln aus -1 wir zur imaginären Einheit machen ist aber eine Konvention. Das legt nahe, dass die Konjugation ein Endomorphismus der komplexen Zahlen, d.h. eine bijektive Abbildung, die mit den Operationen $+$ und $\cdot$ verträglich ist, die die reellen Zahlen invariant läßt:

Satz 2.57. (i) $(\bar{\bar{z}}) = z$

(ii) $\overline{z + w} = \bar{z} + \bar{w}$

(iii) $\overline{z \cdot w} = \bar{z} \cdot \bar{w}$

(iv) $z + \bar{z} = 2\Re z$ und $z - \bar{z} = 2i\Im(z)$

(v) $z \cdot \bar{z}$ ist reell und nicht negativ.

Beweis: (i)-(iv) nachrechnen. (v) $(x + iy)(x - iy) = x^2 + y^2 \geq 0$. **q.e.d.**

Definition 2.58 (Betrag). *Der Betrag einer komplexen Zahl $z = x + iy$ ist definiert als die reelle nicht negative Wurzel aus $z \cdot \bar{z}$, d.h. $|\cdot| : \mathbb{C} \rightarrow \mathbb{R}$, $z \mapsto |z| = \sqrt{z \cdot \bar{z}}$.*

Satz 2.59 (Eigenschaften des Betrags).

(i) $|z| \geq 0$ und $|z| = 0 \iff z = 0$

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

(iii) $|z + w| \leq |z| + |w|$ (Dreiecksungleichung)

(iv) $||z| - |w|| \leq |z - w|$

(v) $|\bar{z}| = |z|$

(vi) $|\Re(z)| \leq |z|$ und $|\Im(z)| \leq |z|$

Beweis: (i) Für $x \neq 0$ oder $y \neq 0$ gilt $|x + iy| = \sqrt{x^2 + y^2} > 0$ und andernfalls $|0| = 0$.

(ii) folgt wegen der Eindeutigkeit der Wurzel aus $|z \cdot w|^2 = z w \overline{z w} = z \bar{z} w \bar{w} = (|z| |w|)^2$.

(v) folgt wegen der Eindeutigkeit der Wurzel aus $|\bar{z}|^2 = \bar{z} z = |z|^2$.

(vi) Für $z = 0$ gilt die Aussage. Sei also $z = x + iy \neq 0$. Dann gilt $x^2 \leq x^2 + y^2$. Aus $0 < \sqrt{x^2 + y^2} < x$ folgt aber mit Monotonie $x^2 + y^2 < x \sqrt{x^2 + y^2} < x^2$. Also gilt $x \leq \sqrt{x^2 + y^2}$ und damit $|\Re(z)| \leq |z|$. Vertauschen von x und y ergibt $|\Im(z)| \leq |z|$.

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

Aus $0 < |z| + |w| < |z + w|$ folgt mit Monotonie $(|z| + |w|)^2 < (|z| + |w|)|z + w| < |z + w|^2$. Also gilt $|z + w| \leq |z| + |w|$.

(iv) folgt aus (i)-(iii) genau wie im reellen Fall. **q.e.d.**

Definition 2.60 (Abstand). *Der Abstand zweier komplexer Zahlen ist die nicht negative Zahl $d : \mathbb{C} \times \mathbb{C} \rightarrow \mathbb{R}$, $(z, w) \mapsto d(z, w) = |z - w|$*

Aus den Eigenschaften des Betrages folgt genau wie für die reellen Zahlen

Satz 2.61 (Eigenschaften des Abstandes).

(i) $d(z, w) \geq 0$ und $d(z, w) = 0 \iff z = w$

(ii) $d(z, w) = d(w, z)$

(iii) $d(z, w) \leq d(z, u) + d(u, w)$. (Dreiecksungleichung).

Wir veranschaulichen die komplexen Zahlen in der zweidimensionalen Ebene. Der Abstand ist der euklidische Abstand zwischen den entsprechenden Punkten der Ebene.

Kapitel 3

Zahlenfolgen

3.1 Konvergenz

Im folgenden werden wir des öfteren Aussagen vorstellen, die sowohl für die reellen Zahlen als auch für die komplexen Zahlen gelten. Wir benutzen dann das Symbol $\mathbb{K}$ um entweder die reellen oder die komplexen Zahlen zusammen mit den entsprechenden Abbildungen und Operationen zu bezeichnen. Die Elemente von $\mathbb{K}$ wollen wir dann einfach Zahlen nennen. Eine Folge ist eine Abbildung $\mathbb{N} \rightarrow \mathbb{K}$, $n \mapsto a_n$. Wir bezeichnen sie mit $(a_n)_{n \in \mathbb{N}}$. Wir interessieren uns vorwiegend für die Grenzwerte solcher Zahlenfolgen.

Definition 3.1. Die Folge $(a_n)_{n \in \mathbb{N}}$ heißt konvergent, wenn es eine Zahl $a \in \mathbb{K}$ gibt und für jedes $\epsilon > 0$ ein $N \in \mathbb{N}$, so dass $|a_n - a| < \epsilon$ für alle natürlichen Zahlen $n \geq N$ gilt. Die Zahl a heißt dann Grenzwert der Folge $(a_n)_{n \in \mathbb{N}}$.

Wir schreiben dann $\lim_{n \rightarrow \infty} a_n = a$ oder $\lim a_n = a$ oder auch $a_n \rightarrow a$ für $n \rightarrow \infty$.

Beispiel 3.2. $\lim_{n \rightarrow \infty} \frac{1}{n} = 0$ (Folgen, die gegen 0 konvergieren heißen Nullfolgen).

Beweis: Nach dem Satz von Archimedes-Eudoxos gibt es für jedes $\epsilon > 0$ ein $N \in \mathbb{N}$ mit $\frac{1}{N} < \epsilon$. Dann gilt wegen Satz 2.14 (vi) für alle $n \geq N$ auch $|\frac{1}{n} - 0| < \epsilon$. **q.e.d.**

Satz 3.3. (i) Der Grenzwert einer konvergenten Folge ist eindeutig.

(ii) Eine komplexe Folge konvergiert genau dann, wenn die entsprechenden Folgen der Realteile und Imaginärteile konvergieren.

(iii) Für jede konvergente Zahlenfolge $(a_n)_{n \in \mathbb{N}}$ ist $\{|a_n| \mid n \in \mathbb{N}\} \subset \mathbb{R}$ beschränkt.

Beweis: (i) Seien a und b zwei Grenzwerte einer konvergenten Folge $(a_n)_{n \in \mathbb{N}}$. Für ein $\epsilon > 0$ gilt dann $|a_n - a| < \frac{\epsilon}{2}$ für alle $n \geq N$ und $|a_n - b| < \frac{\epsilon}{2}$ für alle $n \geq M$. Dann folgt

$$0 \leq |a - b| \leq |a - a_n| + |a_n - b| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \text{ für alle } n \geq \max\{N, M\}.$$

Dann gilt auch $0 \leq |a - b| \leq \inf(0, \infty) = 0$. Also ist $|a - b| = 0$ und damit auch $a = b$.

(ii) Die Realteile und Imaginärteile einer komplexen Folge $(z_n)_{n \in \mathbb{N}}$ bilden zwei reelle Folgen $(x_n)_{n \in \mathbb{N}}$ und $(y_n)_{n \in \mathbb{N}}$ mit $z_n = x_n + iy_n$ für alle $n \in \mathbb{N}$. Sei $z = x + iy$ die Zerlegung einer komplexen Zahl in Realteil und Imaginärteil. Wegen Satz 2.59 (vi)

$$\max\{|x_n - x|, |y_n - y|\} \leq |z_n - z|$$

konvergieren die Real- bzw. Imaginärteile einer konvergenten komplexen Folge gegen den Realteil bzw. Imaginärteil des Grenzwertes. Konvergieren umgekehrt die Folgen $(x_n)_{n \in \mathbb{N}}$ bzw. $(y_n)_{n \in \mathbb{N}}$ gegen x bzw. y , dann gibt es für jedes $\epsilon > 0$ zwei natürliche Zahlen $N, M \in \mathbb{N}$ so dass $|x_n - x| < \frac{\epsilon}{2}$ für alle $n \geq N$ gilt, und $|y_n - y| < \frac{\epsilon}{2}$ für alle $n \geq M$. Dann gilt für alle $n \geq \max\{N, M\}$

$$|x_n + iy_n - (x + iy)| \leq |x_n - x| + |i(y_n - y)| = |x_n - x| + |i| \cdot |y_n - y| < \epsilon.$$

Also konvergiert eine komplexe Folge mit ihren Realteilen und Imaginärteilen.

(iii) Sei $(a_n)_{n \in \mathbb{N}}$ eine konvergente Folge mit Grenzwert a . Dann gibt es ein $N \in \mathbb{N}$, so dass $|a_n - a| < 1$ für alle $n \geq N$ gilt und damit auch $|a_n| \leq |a_n - a + a| < 1 + |a|$. Daraus folgt $|a_n| \leq \max\{|a| + 1, |a_1|, |a_2|, \dots, |a_{N-1}|\}$ für alle $n \in \mathbb{N}$. **q.e.d.**

Eine Folge $(a_n)_{n \in \mathbb{N}}$ heißt divergent, wenn sie nicht konvergiert, wenn es also kein $a \in \mathbb{K}$ gibt, so dass $\lim_{n \rightarrow \infty} a_n = a$. Wenn es für eine reelle Folge für jedes $b \in \mathbb{R}$ ein $N \in \mathbb{N}$ gibt, so dass für alle natürlichen Zahlen $n \geq N$ gilt $a_n > b$ bzw. $a_n < b$ dann schreiben wir $\lim_{n \rightarrow \infty} a_n = \infty$ bzw. $\lim_{n \rightarrow \infty} a_n = -\infty$. Weil in beiden Fällen die Folgen nicht beschränkt sind, können sie wegen dem vorangehenden Satz nicht konvergieren. Es gilt offenbar $\lim_{n \rightarrow \infty} n = \infty$ und $\lim_{n \rightarrow \infty} -n = -\infty$.

Satz 3.4. (i) Sei $|x| < 1$ dann gilt $\lim_{n \rightarrow \infty} x^n = 0$

(ii) Sei $|x| > 1$ dann ist $(x^n)_{n \in \mathbb{N}}$ divergent.

(iii) Sei $x = 1$ dann gilt $\lim_{n \rightarrow \infty} x^n = 1$

(iv) Sei $x \in (1, \infty)$ dann gilt $\lim_{n \rightarrow \infty} x^n = \infty$.

(v) Sei $|x| = 1$ und $x \neq 1$ dann ist $(x^n)_{n \in \mathbb{N}}$ divergent.

Beweis: (i) Für $x = 0$ gilt $\lim_{n \rightarrow \infty} x^n = 0$. Sei also $0 < |x| < 1$. Dann ist $1 < \frac{1}{|x|} = 1 + y$ mit $y = \frac{1 - |x|}{|x|}$. Aus der Bernoulli Ungleichung folgt $\frac{1}{|x|^n} = (1 + y)^n \geq 1 + ny > ny = n \frac{1 - |x|}{|x|}$ und $|x|^n < \frac{1}{n} \frac{|x|}{1 - |x|}$. Aus dem Satz von Archimedes-Eudoxos folgt dann, dass es für jedes $\epsilon > 0$ ein $N \in \mathbb{N}$ gibt, so dass $\frac{1}{N} < \epsilon \cdot \frac{1 - |x|}{|x|}$. Daraus folgt für alle $n \geq N$

$$|x^n - 0| = |x|^n < \frac{1}{n} \frac{|x|}{1 - |x|} \leq \frac{1}{N} \frac{|x|}{1 - |x|} < \epsilon.$$

(iii) Wegen $1^n = 1$ ist $\lim_{n \rightarrow \infty} 1^n = 1$.

(iv) Sei $x > 1$. Dann ist $y = x - 1 > 0$. Für jedes $b \in \mathbb{R}$ gibt es ein $N \in \mathbb{N}$ mit $N > \frac{b}{x-1}$. Dann folgt für alle $n \geq N$ aufgrund der Bernoulli Ungleichung

$$x^n = (1 + y)^n \geq 1 + ny > ny = n(x - 1) \geq N(x - 1) > b.$$

Also gilt $\lim_{n \rightarrow \infty} x^n = \infty$.

(ii) Für $|x| > 1$ gilt wegen (iv) $\lim_{n \rightarrow \infty} |x^n| = \infty$. Also ist $(x^n)_{n \in \mathbb{N}}$ nicht beschränkt.

(v) Für alle $y \in \mathbb{K}$ und alle natürlichen Zahlen n gilt

$$|x^n - y| + |y - x^{n+1}| \geq |x^n - y + y - x^{n+1}| = |(1 - x)x^n| = |1 - x| \cdot |x|^n.$$

Für $|x| = 1$ gilt also $\max\{|x^n - y|, |x^{n+1} - y|\} \geq \frac{|1 - x|}{2}$.

Also kann es für $x \neq 1$ keinen Grenzwert geben.

q.e.d.

Satz 3.5 (Rechenregeln). *Für konvergente Zahlenfolgen $(x_n)_{n \in \mathbb{N}}$ und $(y_n)_{n \in \mathbb{N}}$ gilt*

(i) $\lim_{n \rightarrow \infty} (x_n + y_n) = \lim_{n \rightarrow \infty} x_n + \lim_{n \rightarrow \infty} y_n$

(ii) $\lim_{n \rightarrow \infty} \lambda x_n = \lambda \lim_{n \rightarrow \infty} x_n$ für alle $\lambda \in \mathbb{K}$.

(iii) $\lim_{n \rightarrow \infty} x_n y_n = (\lim_{n \rightarrow \infty} x_n)(\lim_{n \rightarrow \infty} y_n)$

(iv) Wenn $x_n \neq 0$ für alle $n \in \mathbb{N}$ und $\lim_{n \rightarrow \infty} x_n \neq 0$, dann gilt auch $\lim_{n \rightarrow \infty} \frac{1}{x_n} = \frac{1}{\lim_{n \rightarrow \infty} x_n}$.

(v) $\lim_{n \rightarrow \infty} |x_n| = |\lim_{n \rightarrow \infty} x_n|$.

(vi) Wenn zwei reelle konvergente Folgen $(x_n)_{n \in \mathbb{N}}$ und $(y_n)_{n \in \mathbb{N}}$ für alle $n \in \mathbb{N}$ $x_n \leq y_n$ erfüllen, dann gilt auch $\lim_{n \rightarrow \infty} x_n \leq \lim_{n \rightarrow \infty} y_n$.

(vii) Wenn zwei reelle konvergente Folgen $(x_n)_{n \in \mathbb{N}}$ und $(y_n)_{n \in \mathbb{N}}$ für alle $n \in \mathbb{N}$ $x_n \leq y_n$ erfüllen und den gleichen Grenzwert haben, dann gilt für jede reelle Folge $(z_n)_{n \in \mathbb{N}}$, die $x_n \leq z_n \leq y_n$ für alle $n \in \mathbb{N}$ erfüllt, $\lim_{n \rightarrow \infty} z_n = \lim_{n \rightarrow \infty} x_n = \lim_{n \rightarrow \infty} y_n$.

Beweis: Seien $x = \lim_{n \rightarrow \infty} x_n$ und $y = \lim_{n \rightarrow \infty} y_n$. (i) Für jedes $\epsilon > 0$ gibt es natürliche Zahlen $N, M \in \mathbb{N}$, so dass $|x_n - x| < \frac{\epsilon}{2}$ für alle $n \geq N$ und $|y_n - y| < \frac{\epsilon}{2}$ für alle $n \geq M$ gilt. Dann gilt $|x_n + y_n - (x + y)| \leq |x_n - x| + |y_n - y| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$ für alle $n \geq \max\{N, M\}$. Also konvergiert $(x_n + y_n)_{n \in \mathbb{N}}$ gegen $x + y$.

(ii) Für $\lambda = 0$ gilt $\lim_{n \rightarrow \infty} \lambda x_n = 0 = \lambda \cdot \lim_{n \rightarrow \infty} x_n$. Sei also $\lambda \neq 0$ und damit $0 < |\lambda|$. Dann gibt es für jedes $\epsilon > 0$ ein $N \in \mathbb{N}$, so dass $|x_n - x| < \frac{\epsilon}{|\lambda|}$ für alle $n \geq N$ gilt. Daraus folgt:

$$|\lambda x_n - \lambda x| = |\lambda| \cdot |x_n - x| < \epsilon \quad \text{für alle } n \geq N.$$

(iii) Weil $(x_n)_{n \in \mathbb{N}}$ konvergiert, gibt es ein $\lambda > 0$ mit $|x_n| \leq \lambda$ für alle $n \in \mathbb{N}$. Also gibt es für alle $\epsilon > 0$ zwei natürliche Zahlen $N, M \in \mathbb{N}$, so dass $|x_n - x| < \frac{\epsilon}{2|y|+1}$ für $n \geq N$ gilt und $|y_n - y| < \frac{\epsilon}{2\lambda}$ für alle $n \geq M$. Dann gilt für alle $n \geq \max\{N, M\}$

$$|x_n y_n - xy| = |x_n y_n - x_n y + x_n y - xy| \leq |x_n| \cdot |y_n - y| + |x_n - x| \cdot |y| < \frac{\epsilon}{2} + \frac{\epsilon|y|}{2|y|+1} < \epsilon.$$

(iv) Aufgrund der Voraussetzungen gibt es ein $N \in \mathbb{N}$, so dass $|x_n - x| < \frac{|x|}{2}$ für alle $n \geq N$ gilt. Aus $|x| = |x - x_n + x_n| \leq |x - x_n| + |x_n|$ folgt dann

$$|x_n| \geq |x| - |x - x_n| > \frac{|x|}{2} \quad \text{und} \quad \frac{1}{|x_n|} < \frac{2}{|x|} \quad \text{für alle } n \geq N.$$

Dann gibt es für jedes $\epsilon > 0$ ein $M \in \mathbb{N}$, so dass $|x_n - x| < \frac{\epsilon|x|^2}{2}$ für alle $n \geq M$ gilt. Dann gilt für alle $n \geq \max\{N, M\}$ auch

$$\left| \frac{1}{x_n} - \frac{1}{x} \right| = \left| \frac{x - x_n}{x_n x} \right| = \frac{|x - x_n|}{|x_n| \cdot |x|} < \frac{\epsilon \cdot |x|^2}{2} \frac{2}{|x|} \frac{1}{|x|} = \epsilon$$

(v) folgt aus der Ungleichung $||x_n| - |x|| \leq |x_n - x|$.

(vi) Für alle $\epsilon > 0$ gibt es zwei natürliche Zahlen N und M , so dass $|x_n - x| < \frac{\epsilon}{2}$ für alle $n \geq N$ gilt und $|y_n - y| < \frac{\epsilon}{2}$ für alle $n \geq M$. Dann gilt für alle $n \geq \max\{N, M\}$

$$x - y \leq (x - x_n) + (y_n - y) + (x_n - y_n) < \epsilon$$

Also ist $x - y \leq \inf(0, \infty) = 0$. Daraus folgt $x \leq y$.

(vii) Für alle $\epsilon > 0$ gibt es zwei natürliche Zahlen $N, M \in \mathbb{N}$, so dass $|x_n - x| < \epsilon$ für alle $n \geq N$ gilt und $|y_n - x| < \epsilon$ für alle $n \geq M$. Für alle $n \geq \max\{N, M\}$ gilt dann

$$-\epsilon < x_n - x \leq z_n - x \leq y_n - x < \epsilon$$

Daraus folgt $|z_n - x| < \epsilon$.

q.e.d.

Offenbar gilt für alle $n \in \mathbb{N}_0$

$$(1 + x + x^2 + \dots + x^n)(1 - x) = (1 + x + \dots + x^n) - (x + x^2 + \dots + x^{n+1}) = 1 - x^{n+1}$$

$$\text{Also gilt} \quad 1 + x + \dots + x^n = \frac{1 - x^{n+1}}{1 - x} \quad \text{für } x \neq 1 \text{ und } n \in \mathbb{N}_0.$$

Aus Satz 3.4 folgt, dass die Folge $y_n = 1 + x + \dots + x^n$ für $|x| < 1$ gegen $\frac{1}{1-x}$ konvergiert.

Satz 3.6.
$$\lim_{n \rightarrow \infty} \sum_{m=0}^n x^m = \frac{1}{1-x} \quad \text{für } |x| < 1. \quad \text{q.e.d.}$$

3.2 Konvergenzprinzipien

Wir stellen 3 Methoden vor, um zu entscheiden, ob eine Folge konvergiert oder nicht.

Definition 3.7 (Monotonie). *Eine reelle Folge $(a_n)_{n \in \mathbb{N}}$ heißt:*

monoton wachsend, wenn $a_{n+1} \geq a_n$ für alle $n \in \mathbb{N}$.

streng monoton wachsend, wenn $a_{n+1} > a_n$ für alle $n \in \mathbb{N}$.

monoton fallend, wenn $a_{n+1} \leq a_n$ für alle $n \in \mathbb{N}$.

streng monoton fallend, wenn $a_{n+1} < a_n$ für alle $n \in \mathbb{N}$.

Satz 3.8 (Monotonieprinzip). **(i)** *Eine monoton wachsende (fallende) beschränkte reelle Folge $(a_n)_{n \in \mathbb{N}}$ ist konvergent, und es gilt*

$$\lim_{n \rightarrow \infty} a_n = \sup\{a_n \mid n \in \mathbb{N}\} \quad \text{bzw.} \quad \lim_{n \rightarrow \infty} a_n = \inf\{a_n \mid n \in \mathbb{N}\}.$$

(ii) *Für eine monoton wachsende (fallende) unbeschränkte reelle Folge gilt*

$$\lim_{n \rightarrow \infty} a_n = \infty \quad \text{bzw.} \quad \lim_{n \rightarrow \infty} a_n = -\infty.$$

Beweis: **(i)** Sei $(a_n)_{n \in \mathbb{N}}$ eine monoton wachsende (fallende) beschränkte reelle Folge. Dann existiert $\sup\{a_n \mid n \in \mathbb{N}\}$ bzw. $\inf\{a_n \mid n \in \mathbb{N}\}$. Für alle $\epsilon > 0$ gibt es $N \in \mathbb{N}$ mit

$$\begin{aligned} \sup\{a_n \mid n \in \mathbb{N}\} - \epsilon &< a_N \leq a_m \leq \sup\{a_n \mid n \in \mathbb{N}\} && \text{für alle } m \geq N && \text{bzw.} \\ \inf\{a_n \mid n \in \mathbb{N}\} &\leq a_m \leq a_N < \inf\{a_n \mid n \in \mathbb{N}\} + \epsilon && \text{für alle } m \geq N. \end{aligned}$$

Daraus folgt $|a_m - \sup\{a_n \mid n \in \mathbb{N}\}| < \epsilon$ bzw. $|a_m - \inf\{a_n \mid n \in \mathbb{N}\}| < \epsilon$.

(ii) Sei $(a_n)_{n \in \mathbb{N}}$ eine monoton wachsende (fallende) unbeschränkte reelle Folge. Für jedes $b \in \mathbb{R}$ gibt es $N \in \mathbb{N}$ mit $a_m \geq a_N > b$ bzw. $a_m \leq a_N < b$ für alle $m \geq N$. **q.e.d.**

Beispiel 3.9 (Existenz und Konstruktion der k -ten Wurzel). *Für alle $a > 0$ und alle $k \in \mathbb{N} \setminus \{1\}$ definieren wir die Folge $(a_n)_{n \in \mathbb{N}_0}$ rekursiv durch*

$$a_0 = 1 + \frac{a-1}{k} \quad a_{n+1} = a_n \left(1 + \frac{a - a_n^k}{k \cdot a_n^k} \right) \quad \text{für alle } n \in \mathbb{N}_0.$$

Für $a = 1$ ist dann $a_n = 1$ für alle $n \in \mathbb{N}$. Für $a \neq 1$ gilt für alle $n \in \mathbb{N}$

$$0 < a_n < a_0, \quad a_n < a_{n-1} \quad \text{und} \quad a < a_n^k.$$

Beweis durch vollständige Induktion in n :

(i) Für $a > 0$ und $a \neq 1$ ist $-1 < -\frac{1}{k} < \frac{a-1}{k} \neq 0$. Wegen der Bernoulli Ungleichung gilt $a_0^k = \left(1 + \frac{a-1}{k}\right)^k > 1 + a - 1 = a$. Es folgt $-1 < -\frac{1}{k} < \frac{a-a_0^k}{ka_0^k} < 0$. Also gilt $0 < a_1 < a_0$

und mit der Bernoulli Ungleichung $a_1^k = a_0^k \left(1 + \frac{a-a_0^k}{ka_0^k}\right)^k > a_0^k \left(1 + k \frac{a-a_0^k}{ka_0^k}\right) = a$.

(ii) Wir nehmen an es gilt $0 < a_n < a_0$, $a_n < a_{n-1}$ und $a < a_n^k$ für ein $n \in \mathbb{N}$. Dann folgt $-1 < -\frac{1}{k} < \frac{a-a_n^k}{k \cdot a_n^k} < 0$. Daraus folgt $0 < a_{n+1} < a_n < a_0$ und wegen der Bernoulli

Ungleichung: $a_{n+1}^k = a_n^k \left(1 + \frac{a-a_n^k}{ka_n^k}\right)^k > a_n^k \left(1 + k \frac{a-a_n^k}{ka_n^k}\right) = a$. **q.e.d.**

Die Folge $(a_n)_{n \in \mathbb{N}}$ ist monoton fallend und beschränkt. Wegen dem Monotonieprinzip konvergiert sie gegen $b \geq 0$. Wir multiplizieren die Rekursionsgleichung mit ka_n^{k-1} :

$$a_{n+1} \cdot ka_n^{k-1} = ka_n^k + a - a_n^k.$$

Bilden wir links und rechts den Grenzwert $n \rightarrow \infty$ so erhalten wir mit Satz 3.5

$$kb^k = (k-1)b^k + a \text{ oder auch } b^k = a.$$

Also existiert eine positive Zahl b mit $b^k = a$. Diese Folge konvergiert sehr schnell. Außerdem sind für rationale a alle Folgenglieder rational.

Satz 3.10. (i) Für jede positive Zahl $a > 0$ und jede rationale Zahl r gibt es genau eine positive Zahl a^r . Für $r=0$ setzen wir $a^0 = 1$.

(ii) Für jede positive rationale Zahl $r > 0$ und $0 < a < b$ gilt auch $0 < a^r < b^r$.

(iii) Für jede negative rationale Zahl $r < 0$ und $0 < a < b$ gilt auch $0 < b^r < a^r$.

Beweis: (i) Für $r = \frac{p}{q}$ definieren wir $a^r = \sqrt[q]{a^p}$. Offenbar ist a^r eine positive Lösung der Gleichungen $(a^r)^{qn} = (a^p)^n = a^{pn}$ für alle $n \in \mathbb{N}$. Wir zeigen, dass für $a, b \in (0, \infty)$ und für $n \in \mathbb{N}$ die Ungleichung $a < b$ äquivalent ist zu $a^n < b^n$. Für $n \in \mathbb{N} \setminus \{1\}$ gilt

$$b^n - a^n = (b-a)(b^{n-1} + b^{n-2}a + \dots + ba^{n-2} + a^{n-1}).$$

Also folgt aus $0 < a < b$ auch $a^n < b^n$ und aus $0 < b \leq a$ auch $a^n \leq b^n$. Also ist für $a > 0$ und $b > 0$ die Ungleichung $a < b$ äquivalent zu der Ungleichung $a^n < b^n$. Also gibt es für alle $n \in \mathbb{N}$ und alle $r = \frac{p}{q} > 0$ genau eine positive Lösung a^r der Gleichung $(a^r)^{qn} = a^{pn}$. Für $r < 0$ ist $a^r = \frac{1}{a^{-r}}$ die entsprechende positive Zahl.

(ii) Sei $r = \frac{p}{q} > 0$. Dann ist für $a, b \in (0, \infty)$ die Ungleichung $a < b$ äquivalent zu $a^p < b^p$ und das wiederum äquivalent zu $a^{\frac{p}{q}} < b^{\frac{p}{q}}$.

(iii) Sei $r = -\frac{p}{q} < 0$. Dann folgt (iii) aus (ii) wegen Satz 2.14 (vi). **q.e.d.**

Definition 3.11 (Teilfolge). Eine Teilfolge einer Folge $(a_n)_{n \in \mathbb{N}}$ ist eine Folge $(b_m)_{m \in \mathbb{N}}$ von der Form $b_m = a_{n_m}$, wobei $0 < n_1 < n_2 < \dots$ eine streng monoton wachsende Folge von natürlichen Zahlen ist.

Z.B. hat die divergente Folge $a_n = (-1)^n$ zwei konvergente Teilfolgen $b_m = a_{2m} = 1$ und $c_m = a_{2m+1} = -1$.

Satz 3.12 (monotone Teilfolgen). *Jede reelle Folge enthält eine monotone Teilfolge.*

Beweis: Sei $(a_n)_{n \in \mathbb{N}}$ eine reelle Folge und A die Menge $A = \{n \in \mathbb{N} \mid a_n \geq a_m \forall m > n\}$. Wenn A eine unendliche Menge ist, dann sei $n_1 < n_2 < \dots$ eine Abzählung der Elemente von A . Die Teilfolge $(b_m)_{m \in \mathbb{N}} = (a_{n_m})_{m \in \mathbb{N}}$ ist dann monoton fallend.

Wenn A eine endliche Menge ist besitzt sie ein Maximum N . Dann gibt es also zu jedem $n > N$ ein $m > n$, so dass $a_m > a_n$ ist. Also definieren wir induktiv eine Teilfolge $(b_m)_{m \in \mathbb{N}}$, so dass $b_1 = a_{N+1}$ und für alle $m \in \mathbb{N}$ $b_{m+1} > b_m$ gilt. Diese Folge ist streng monoton steigend. Also besitzt $(a_n)_{n \in \mathbb{N}}$ entweder eine monoton fallende oder eine streng monoton steigende Teilfolge. Umgekehrt gilt auch, dass $(a_n)_{n \in \mathbb{N}}$ entweder eine monoton steigende oder eine streng monoton fallende Teilfolge besitzt. **q.e.d.**

Satz 3.13 (Auswahlprinzip von Bolzano–Weierstraß). *Jede beschränkte reelle Folge besitzt eine konvergente Teilfolge.*

Beweis: Wegen dem vorangehenden Satz besitzt jede Folge $(a_n)_{n \in \mathbb{N}}$ eine monotone Teilfolge $(b_n)_{n \in \mathbb{N}}$. Wenn die ursprüngliche Folge beschränkt ist, ist auch die Teilfolge beschränkt. Diese konvergiert dann wegen dem Monotonieprinzip. **q.e.d.**

Korollar 3.14. *Jede beschränkte komplexe Folge besitzt eine konvergente Teilfolge.*

Beweis: Wegen $\max\{|x|, |y|\} \leq |x + iy|$ sind die reellen Folgen der Realteile und Imaginärteile einer komplexen beschränkten Folge beschränkt. Wegen dem Auswahlprinzip von Bolzano–Weierstraß besitzt die Folge der Realteile eine konvergente Teilfolge, und dann auch die entsprechende Teilfolge der Imaginärteile. Wegen Satz 3.3 (ii) konvergiert die der zweiten Teilfolge entsprechende komplexe Teilfolge. **q.e.d.**

Definition 3.15 (Cauchyfolge). *Eine Folge $(a_n)_{n \in \mathbb{N}}$ heißt Cauchyfolge, wenn es für jedes $\epsilon > 0$ ein $N \in \mathbb{N}$ gibt, so dass $|a_n - a_m| < \epsilon$ für alle $n, m \geq N$ gilt.*

Satz 3.16 (Kriterium von Cauchy). *Eine Zahlenfolge konvergiert genau dann, wenn sie eine Cauchyfolge ist.*

Beweis: Sei $(a_n)_{n \in \mathbb{N}}$ eine konvergente Folge. Dann gibt es für jedes $\epsilon > 0$ ein $N \in \mathbb{N}$, so dass alle $m \geq N$ die Ungleichung $|a_m - \lim_{n \rightarrow \infty} a_n| < \frac{\epsilon}{2}$ erfüllen. Also gilt auch

$$|a_m - a_l| \leq |a_m - \lim_{n \rightarrow \infty} a_n| + |\lim_{n \rightarrow \infty} a_n - a_l| < \epsilon \quad \text{für alle } m, l \geq N.$$

Sei umgekehrt $(a_n)_{n \in \mathbb{N}}$ eine Cauchyfolge. Dann gibt es ein $N \in \mathbb{N}$, so dass für alle $m \geq N$ gilt $|a_m - a_N| < 1$, und damit auch $|a_m| \leq |a_N| + |a_m - a_N| < |a_N| + 1$. Für alle $n \in \mathbb{N}$ gilt dann aber $|a_n| \leq \max\{|a_1|, \dots, |a_{N-1}|, |a_N| + 1\}$. Deshalb ist die Folge a_n beschränkt und besitzt wegen dem Auswahlprinzip von Bolzano–Weierstraß eine

konvergente Teilfolge $(b_n)_{n \in \mathbb{N}}$. Sei $a = \lim_{n \rightarrow \infty} b_n$. Dann gibt es für jedes $\epsilon > 0$ zwei natürliche Zahlen $N, M \in \mathbb{N}$, so dass alle $n, m \geq N$ die Ungleichung $|a_n - a_m| < \frac{\epsilon}{2}$ erfüllen und alle $n \geq M$ die Ungleichung $|b_n - a| < \frac{\epsilon}{2}$. Weil $(b_n)_{n \in \mathbb{N}}$ eine Teilfolge von $(a_n)_{n \in \mathbb{N}}$ ist, folgt $|a_n - a| \leq |a_n - b_n| + |b_n - a| < \epsilon$ für alle $n \geq \max\{N, M\}$. **q.e.d.**

Satz 3.17. *Für einen angeordneten archimedischen Körper ist folgendes äquivalent:*

Vollständigkeitsaxiom A5.

(i) **aus dem Monotonieprinzip.**

Auswahlprinzip von Bolzano–Weierstraß.

Jede Cauchyfolge konvergiert.

Intervallschachtelungsprinzip.

Beweis*: Wegen der Beweise des Monotonieprinzips, des Auswahlprinzips von Bolzano–Weierstraß und des Kriteriums von Cauchy und wegen Satz 2.51 genügt es zu zeigen, dass aus dem Kriterium von Cauchy das Vollständigkeitsaxiom **A5** folgt.

Sei M eine nicht leere nach oben beschränkte Menge. Sei $a_1 \in M$ und $b_1 > a_1$ eine obere Schranke von M . Wir konstruieren wie im Beweis von Satz 2.51 eine Folge $a_1 \leq a_2 \leq \dots$ in M und eine Folge $b_1 \geq b_2 \geq \dots$ von oberen Schranken von M mit

$$0 \leq \max\{b_{n+1} - b_{n+m}, a_{n+m} - a_{n+1}\} \leq b_{n+1} - a_{n+1} \leq 2^{-n}(b_1 - a_1) \text{ für alle } m \in \mathbb{N}.$$

Dann sind $(a_n)_{n \in \mathbb{N}}$ und $(b_n)_{n \in \mathbb{N}}$ Cauchyfolgen und konvergieren gegen den gleichen Grenzwert s . Für alle $x \in M$ gilt $x \leq b_n$ für alle $n \in \mathbb{N}$. Aus Satz 3.5 (vi) folgt $x \leq \lim_{n \rightarrow \infty} b_n = s$. Also ist s eine obere Schranke von M . Für alle $x < s = \lim_{n \rightarrow \infty} a_n$ gibt es ein $a_n > x$. Deshalb ist s das Minimum der oberen Schranken von M . **q.e.d.**

Insbesondere sind die reellen Zahlen auch als der angeordnete archimedische Körper charakterisiert, in dem jede Cauchyfolge konvergiert. Wir werden die reellen Zahlen später auch als Äquivalenzklasse von Cauchyfolgen von rationalen Zahlen auffassen, wobei zwei Cauchyfolgen äquivalent heißen, wenn ihre Differenz eine Nullfolge ist.

3.3 Häufungspunkte

Definition 3.18 (Häufungspunkt). *Die Grenzwerte von konvergenten Teilfolgen heißen Häufungspunkte. Bei reellen Folgen sind in $\mathbb{R}$ zusätzlich $+\infty$ bzw. $-\infty$ Häufungspunkte, wenn es Teilfolgen mit diesen Grenzwerten in $\mathbb{R}$ gibt.*

Satz 3.19 (Limes superior und Limes inferior). *Ist die Menge der reellen Häufungspunkte einer reellen Folge nicht leer und nach oben (unten) in $\mathbb{R}$ beschränkt, so besitzt sie ein Maximum (Minimum) in $\mathbb{R}$.*

Beweis: Sei die Menge der Häufungspunkte der reellen Folge $(a_n)_{n \in \mathbb{N}}$ in $\mathbb{R}$ nach oben beschränkt. Sei $a \in \mathbb{R}$ das Supremum der Häufungspunkte, und b_m für jedes $m \in \mathbb{N}$ ein Häufungspunkt $b_m \in (a - \frac{1}{2m}, a]$. Wir definieren induktiv eine Teilfolge $(c_m)_{m \in \mathbb{N}}$ von $(a_n)_{n \in \mathbb{N}}$ mit $|c_m - b_m| < \frac{1}{2m}$ für alle $m \in \mathbb{N}$. Für alle $m \in \mathbb{N}$ gilt $|c_m - a| \leq |c_m - b_m| + |b_m - a| < \frac{1}{m}$. Also ist a ein Häufungspunkt von $(a_n)_{n \in \mathbb{N}}$ und damit das Maximum der Häufungspunkte. Der Beweis für das Minimum ist analog. **q.e.d.**

Definition 3.20. Für eine nach oben (unten) beschränkte reelle Folge $(a_n)_{n \in \mathbb{N}}$ heißt das Supremum (bzw. Infimum) der Häufungspunkte Limes superior bzw. inferior. Wir bezeichnen es mit $\overline{\lim} a_n = \limsup_{n \rightarrow \infty} a_n$ bzw. $\underline{\lim} a_n = \liminf_{n \rightarrow \infty} a_n$. Wenn $(a_n)_{n \in \mathbb{N}}$ nicht nach oben (unten) beschränkt ist, dann sei $\overline{\lim} a_n = \infty$ bzw. $\underline{\lim} a_n = -\infty$.

Satz 3.21. (i) $\bar{a}$ ist genau dann der Limes superior einer nach oben beschränkten reellen Folge $(a_n)_{n \in \mathbb{N}}$, wenn für alle $\epsilon > 0$ unendlich viele Elemente der Folge $a_n > \bar{a} - \epsilon$ erfüllen, aber höchstens endlich viele $a_n > \bar{a} + \epsilon$.

(ii) $\underline{a}$ ist genau dann der Limes inferior einer nach unten beschränkten reellen Folge $(a_n)_{n \in \mathbb{N}}$, wenn für alle $\epsilon > 0$ unendlich viele Elemente $a_n < \underline{a} + \epsilon$ erfüllen, aber höchstens endlich viele $a_n < \underline{a} - \epsilon$.

Beweis: Wir beweisen wieder nur (i), weil (ii) analog zu beweisen ist. Sei $(a_n)_{n \in \mathbb{N}}$ eine nach oben beschränkte reelle Folge. Wegen dem Auswahlprinzip von Bolzano–Weierstraß ist jede untere Schranke einer Teilfolge von $(a_n)_{n \in \mathbb{N}}$ auch eine untere Schranke von mindestens einem Häufungspunkt. Aus der Charakterisierung der Zahl $\bar{a}$ in (i) folgt also, dass $\bar{a} - 2\epsilon$ keine obere Schranke, aber $\bar{a} + \epsilon$ eine obere Schranke der Häufungspunkte von $(a_n)_{n \in \mathbb{N}}$ ist. Umgekehrt folgt sie daraus, dass $\bar{a} - \epsilon$ keine obere Schranke, aber $\bar{a}$ eine obere Schranke der Häufungspunkte ist. Damit ist ihre Gültigkeit für alle $\epsilon > 0$ äquivalent dazu, dass $\bar{a}$ das Supremum der Häufungspunkte ist. **q.e.d.**

Korollar 3.22. Eine reelle Folge konvergiert genau dann, wenn sie beschränkt ist und wenn $\overline{\lim} a_n = \underline{\lim} a_n$ gilt, sie also nur einen Häufungspunkt hat.

Weil $\overline{\lim} a_n$ und $\underline{\lim} a_n$ das Maximum und das Minimum der Häufungspunkte sind, ist $\overline{\lim} a_n = \underline{\lim} a_n$ äquivalent zu der Bedingung, dass es nur einen Häufungspunkt gibt.

Beweis: Wenn $(a_n)_{n \in \mathbb{N}}$ beschränkt ist und $\overline{\lim} a_n = \underline{\lim} a_n = a$ gilt, dann folgt aus dem vorangehenden Satz, dass es für jedes $\epsilon > 0$ ein N gibt, so dass $\underline{\lim} a_n - \epsilon \leq a_n \leq \overline{\lim} a_n + \epsilon$ für alle $n \geq N$ gilt. Also gilt $|a_n - a| \leq \epsilon$ und $(a_n)_{n \in \mathbb{N}}$ konvergiert gegen a .

Wenn $(a_n)_{n \in \mathbb{N}}$ konvergiert, dann konvergiert jede Teilfolge gegen $a = \lim_{n \rightarrow \infty} a_n$. Also besitzt $(a_n)_{n \in \mathbb{N}}$ nur einen Häufungspunkt und es gilt $\overline{\lim} a_n = \underline{\lim} a_n = a$. **q.e.d.**

Korollar 3.23. Seien $(a_n)_{n \in \mathbb{N}}$ und $(b_n)_{n \in \mathbb{N}}$ nach unten (bzw. oben) beschränkte reelle Folgen, die für alle $n \in \mathbb{N}$ $a_n \leq b_n$ erfüllen. Dann gilt $\underline{\lim} a_n \leq \underline{\lim} b_n$ bzw. $\overline{\lim} a_n \leq \overline{\lim} b_n$.

Beweis: Sei $(d_n)_{n \in \mathbb{N}}$ eine Teilfolge von $(b_n)_{n \in \mathbb{N}}$, die gegen $\lim b_n$ (bzw. $(c_n)_{n \in \mathbb{N}}$ eine Teilfolge von $(a_n)_{n \in \mathbb{N}}$ die gegen $\lim a_n$) konvergiert. Dann gibt es eine Teilfolge $(c_n)_{n \in \mathbb{N}}$ von $(a_n)_{n \in \mathbb{N}}$ (bzw. $(d_n)_{n \in \mathbb{N}}$ von $(b_n)_{n \in \mathbb{N}}$), die für alle $n \in \mathbb{N}$ auch $c_n \leq d_n$ erfüllt. Diese Teilfolge ist beschränkt und besitzt eine konvergente Teilfolge. Wir können also ohne Beschränkung der Allgemeinheit annehmen, dass $(c_n)_{n \in \mathbb{N}}$ (bzw. $(d_n)_{n \in \mathbb{N}}$) konvergiert. Dann ist aber $\lim_{n \rightarrow \infty} c_n$ ein Häufungspunkt von $(a_n)_{n \in \mathbb{N}}$ (bzw. $\lim_{n \rightarrow \infty} d_n$ ein Häufungspunkt von $(b_n)_{n \in \mathbb{N}}$). Also gilt

$$\liminf a_n \leq \lim_{n \rightarrow \infty} c_n \leq \lim_{n \rightarrow \infty} d_n = \liminf b_n$$

(bzw. $\liminf a_n = \lim_{n \rightarrow \infty} c_n \leq \lim_{n \rightarrow \infty} d_n \leq \limsup b_n$). **q.e.d.**

3.4 Beispiele

(i) Für alle $x \in \mathbb{C}$ ist $\lim_{n \rightarrow \infty} \frac{x^n}{n!} = 0$.

Beweis: Wegen dem Satz von Archimedes–Endoxos gibt es ein $N \in \mathbb{N}$ mit $|x| < N$. Dann gilt für alle $n \geq N$

$$\left| \frac{x^n}{n!} \right| = \frac{|x|^{N-1}}{(N-1)!} \cdot \frac{|x|^{n+1-N}}{N \cdots n} \leq \frac{|x|^{N-1}}{(N-1)!} \left(\frac{|x|}{N} \right)^{n-N} \cdot \frac{|x|}{n} < \frac{|x|^N}{(N-1)!} \cdot \frac{1}{n}$$

Weil aber $\frac{1}{n}$ eine Nullfolge ist, konvergiert dann $\left| \frac{x^n}{n!} \right|$ auch gegen Null. **q.e.d.**

(ii) Für alle positiven rationalen Zahlen $r > 0$ gilt $\lim_{n \rightarrow \infty} \frac{1}{n^r} = 0$.

Beweis: Nach dem Satz von Archimedes–Eudoxos, gibt es für alle $\epsilon > 0$ ein $N \in \mathbb{N}$ mit $N > \frac{1}{\epsilon^{\frac{1}{r}}}$. Für alle $n \geq N$ folgt $\frac{1}{n^r} \leq \frac{1}{N^r} < \epsilon$. Also konvergiert $\frac{1}{n^r}$ nach Null. **q.e.d.**

(iii) Für alle $x \in (0, \infty)$ gilt $\lim_{n \rightarrow \infty} \sqrt[n]{x} = 1$.

Beweis: Sei zunächst $x \geq 1$. Sei also $y_n = \sqrt[n]{x} - 1 \geq 0$. Wegen der Bernoulli Ungleichung gilt dann $x = (1 + y_n)^n \geq 1 + ny_n$. Daraus folgt $0 \leq y_n \leq \frac{x-1}{n}$. Dann konvergiert y_n aber gegen Null. $\lim_{n \rightarrow \infty} \sqrt[n]{x} = 1 + \lim_{n \rightarrow \infty} y_n = 1$.

Sei jetzt $0 < x < 1$. Dann ist $\frac{1}{x} > 1$. Also gilt $\lim_{n \rightarrow \infty} \sqrt[n]{x} = \frac{1}{\lim_{n \rightarrow \infty} \sqrt[n]{\frac{1}{x}}} = 1$. **q.e.d.**

Binomische Formel 3.24. Für alle $n \in \mathbb{N}$ und alle Zahlen $x, y \in \mathbb{K}$ gilt

$$(x + y)^n = \sum_{k=0}^n \binom{n}{k} x^k y^{n-k}, \text{ wobei } \binom{n}{k} = \frac{n(n-1) \cdots (n-k+1)}{k!} \text{ und } 0! = 1 = \binom{n}{0}.$$

Beweis: durch vollständige Induktion:

(i) Offenbar ist $\binom{1}{0} = \binom{1}{1} = 1$, und die binomische Formel ist für $n = 1$ richtig.

(ii) Wenn die binomische Formel für $n \in \mathbb{N}$ gilt, dann folgt

$$\begin{aligned}
 (x+y)^{n+1} &= (x+y)(x+y)^n = (x+y) \sum_{k=0}^n \binom{n}{k} x^k y^{n-k} \\
 &= \sum_{k=1}^{n+1} \binom{n}{k-1} x^k y^{n+1-k} + \sum_{k=0}^n \binom{n}{k} x^k y^{n+1-k} \\
 &= \sum_{k=0}^{n+1} \left(k \frac{n(n-1) \cdots (n-k+2)}{k!} + \frac{n(n-1) \cdots (n-k+1)}{k!} \right) x^k y^{n+1-k} \\
 &= \sum_{k=0}^{n+1} \left(\frac{(n+1)n \cdots (n-k+2)}{k!} \right) x^k y^{n+1-k}.
 \end{aligned}$$

Also gilt sie auch für $n+1$.

q.e.d.

(iv) $\lim_{n \rightarrow \infty} \sqrt[n]{n} = 1$.

Beweis: Sei $y_n = \sqrt[n]{n} - 1 \geq 0$. Wegen der binomischen Formel gilt für alle $n \in \mathbb{N}$

$$n = (\sqrt[n]{n})^n = (1 + y_n)^n = \sum_{k=0}^n \binom{n}{k} y_n^k$$

Für alle $n \geq 2$ folgt $n \geq 1 + \frac{n(n-1)}{2} y_n^2 \iff y_n^2 \leq \frac{2}{n} \iff y_n \leq \frac{\sqrt{2}}{\sqrt{n}}$.

Wegen (ii) ist dann y_n eine Nullfolge.

q.e.d.

(v) $\lim_{n \rightarrow \infty} \sqrt[n]{n!} = \infty$.

Beweis: Wegen (i) gibt es für alle $x \in (0, \infty)$ ein N , so dass für alle $n \geq N$ gilt $\frac{x^n}{n!} < 1$. Dann gilt auch $\frac{x}{\sqrt[n]{n!}} < 1 \iff x < \sqrt[n]{n!}$. Also folgt $\lim_{n \rightarrow \infty} \sqrt[n]{n!} = \infty$.

q.e.d.

Satz 3.25. Sei $(a_n)_{n \in \mathbb{N}}$ eine reelle positive Folge, dann gilt

$$\underline{\lim} \left(\frac{a_{n+1}}{a_n} \right) \leq \underline{\lim} \sqrt[n]{a_n} \qquad \overline{\lim} \sqrt[n]{a_n} \leq \overline{\lim} \left(\frac{a_{n+1}}{a_n} \right).$$

Wenn die Folge $\sqrt[n]{a_n}$ also einen reellen Häufungspunkt hat, dann hat auch die Folge $\frac{a_{n+1}}{a_n}$ einen reellen Häufungspunkt. Und wenn gilt $\overline{\lim} \left(\frac{a_{n+1}}{a_n} \right) < \infty$ dann auch $\overline{\lim} \sqrt[n]{a_n} < \infty$.

Beweis: Wenn $\underline{\lim} \left(\frac{a_{n+1}}{a_n} \right) = 0$ ist, ist die erste Aussage richtig. Für $\underline{\lim} \left(\frac{a_{n+1}}{a_n} \right) > 0$ gibt es wegen Satz 3.21 (ii) ein $a > 0$, und für alle $0 < \epsilon < a$ ein $N \in \mathbb{N}$ gibt, so dass alle $n \geq N$ auch $\frac{a_{n+1}}{a_n} \geq a - \epsilon$ erfüllen. Dann gilt für alle $n > N$

$$\frac{a_n}{a_N} = \frac{a_{N+1}}{a_N} \cdots \frac{a_n}{a_{n-1}} \geq (a - \epsilon)^{n-N} \implies a_n \geq a_N \frac{(a - \epsilon)^n}{(a - \epsilon)^N} \implies \sqrt[n]{a_n} \geq \sqrt[n]{\frac{a_N}{(a - \epsilon)^N}} \cdot (a - \epsilon).$$

Wegen (iii) und Korollar 3.23 gilt dann $\underline{\lim} \sqrt[n]{a_n} \geq (a - \epsilon)$. Weil dies für alle $\epsilon > 0$ gilt, folgt auch $\underline{\lim} \sqrt[n]{a_n} \geq a$. Also ist $\underline{\lim} \sqrt[n]{a_n} \geq a$ nicht kleiner als der kleinste Häufungspunkt von $(\frac{a_{n+1}}{a_n})_{n \in \mathbb{N}}$. Und für $\lim \frac{a_{n+1}}{a_n} = \infty$ gilt auch $\lim \sqrt[n]{a_n} = \infty$.

Für die Folge $(a_n^{-1})_{n \in \mathbb{N}}$ gilt entsprechend $\underline{\lim} \frac{a_n}{a_{n+1}} \leq \underline{\lim} \sqrt[n]{a_n^{-1}} = \underline{\lim} (\sqrt[n]{a_n})^{-1}$. Wegen Satz 2.14 (vi) und Satz 3.5 (iv) ist für eine positive Folge $(b_n)_{n \in \mathbb{N}}$ $\underline{\lim} b_n^{-1} = (\overline{\lim} b_n)^{-1}$ (mit $\frac{1}{\infty} = 0$ und $\frac{1}{0} = \infty$). Also folgt die zweite Ungleichung aus Satz 2.14 (vi). **q.e.d.**

(vi) $\left(\sum_{k=0}^n \frac{1}{k!} \right)_{n \in \mathbb{N}}$ ist streng monoton wachsend und durch 3 beschränkt: Für $n > 2$ gilt

$$\frac{5}{2} < \sum_{k=0}^n \frac{1}{k!} \leq 2 + \sum_{k=2}^n \frac{1}{k(k-1)} = 2 + \sum_{k=2}^n \left(\frac{1}{k-1} - \frac{1}{k} \right) = 2 + \sum_{k=1}^{n-1} \frac{1}{k} - \sum_{k=2}^n \frac{1}{k} = 2 + 1 - \frac{1}{n}.$$

Also konvergiert diese Folge. Der Grenzwert heißt Eulersche Zahl $e \in (\frac{5}{2}, 3]$.

(vii) $\lim_{n \rightarrow \infty} \left(1 + \frac{1}{n} \right)^n = e$.

Beweis: Wegen der binomischen Formel gilt für alle $n \in \mathbb{N}$

$$\left(1 + \frac{1}{n} \right)^n = \sum_{k=0}^n \binom{n}{k} \frac{1}{n^k} = \sum_{k=0}^n \frac{n(n-1) \cdots (n-k+1)}{n^k} \cdot \frac{1}{k!} \leq \sum_{k=0}^n \frac{1}{k!} \leq e.$$

Für alle $k \in \mathbb{N}$ gilt $\lim_{n \rightarrow \infty} \binom{n}{k} \frac{1}{n^k} = \lim_{n \rightarrow \infty} \frac{n(n-1) \cdots (n-k+1)}{n^k} \cdot \frac{1}{k!} = \frac{1}{k!}$.

Also gilt für alle $m \in \mathbb{N}$ $\underline{\lim} \left(1 + \frac{1}{n} \right)^n \geq \sum_{k=0}^m \lim_{n \rightarrow \infty} \binom{n}{k} \frac{1}{n^k} = \sum_{k=0}^m \frac{1}{k!}$.

Wegen $\sup \{ \sum_{k=0}^m \frac{1}{k!} \mid m \in \mathbb{N} \} = e$ folgt $e \leq \underline{\lim} \left(1 + \frac{1}{n} \right)^n \leq \overline{\lim} \left(1 + \frac{1}{n} \right)^n \leq e$. **q.e.d.**

(viii) $\lim_{n \rightarrow \infty} \frac{n}{\sqrt[n]{n!}} = e$.

Beweis: Sei $(a_n)_{n \in \mathbb{N}}$ die positive Folge $\frac{n}{n!}$. Dann gilt $\frac{a_{n+1}}{a_n} = \frac{(n+1)^{n+1}}{(n+1)n^n} = \left(1 + \frac{1}{n} \right)^n$. Wegen (vii) folgt $\lim_{n \rightarrow \infty} \frac{a_{n+1}}{a_n} = e$. Dann folgt aus Satz 3.25

$$e = \underline{\lim} \frac{a_{n+1}}{a_n} \leq \underline{\lim} \frac{n}{\sqrt[n]{n!}} \leq \overline{\lim} \frac{n}{\sqrt[n]{n!}} \leq \overline{\lim} \frac{a_{n+1}}{a_n} = e.$$

Daraus folgt $\lim_{n \rightarrow \infty} \frac{n}{\sqrt[n]{n!}} = e$. **q.e.d.**

Kapitel 4

Reihen

4.1 Konvergenzkriterien

Definition 4.1. Sei $(a_n)_{n \in \mathbb{N}}$ eine Folge. Dann heißt die Folge $(s_n)_{n \in \mathbb{N}}$ mit $s_n = \sum_{j=1}^n a_j$ die zu $(a_n)_{n \in \mathbb{N}}$ gehörende Reihe. Diese Folge bezeichnen wir mit $(\sum a_n)_{n \in \mathbb{N}}$.

Wenn die Reihe $(\sum a_n)_{n \in \mathbb{N}}$ konvergiert, dann bezeichnen wir den Grenzwert mit $\lim_{n \rightarrow \infty} \sum_{j=1}^n a_j = \sum_{n=1}^{\infty} a_n = \sum_{n \in \mathbb{N}} a_n$. Analog definieren wir $\sum_{n=m}^{\infty} a_n = \lim_{n \rightarrow \infty} \sum_{j=m}^n a_j$.

Beispiel 4.2. (i) Geometrische Reihe $(\sum q^n)_{n \in \mathbb{N}_0}$. Für $q \neq 1$ hatten wir berechnet:

$\sum_{j=0}^n q^j = \frac{1-q^{n+1}}{1-q}$. Dann folgt aus dem letzten Abschnitt

Für $|q| < 1$ ist $(\sum q^n)_{n \in \mathbb{N}_0}$ konvergent: $\sum_{n=0}^{\infty} q^n = \frac{1}{1-q}$.

Für $|q| \geq 1$ ist $(\sum q^n)_{n \in \mathbb{N}_0}$ divergent.

Für reelles $q \geq 1$ ist $(\sum q^n)_{n \in \mathbb{N}_0}$ divergent: $\sum_{n=0}^{\infty} q^n = \infty$.

(ii) Die ζ -Funktion ist definiert als $\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$ der Grenzwert (wenn er existiert) der Reihe $(\sum \frac{1}{n^s})_{n \in \mathbb{N}}$. Zunächst ist diese Reihe nur für alle rationalen Zahlen $s \in \mathbb{Q}$ definiert. Für $s = 1$ ist $(\sum \frac{1}{n^s})_{n \in \mathbb{N}}$ divergent.

Beweis: Diese Reihe ist monoton wachsend. Also ist nur die Frage, ob sie beschränkt ist oder nicht. Für alle $n \in \mathbb{N}$ gilt aber $\frac{1}{n+1} + \frac{1}{n+2} + \dots + \frac{1}{n+n} \geq \frac{n}{2n} = \frac{1}{2}$. Also

sind für alle $n \in \mathbb{N}_0$ jeweils die Summen $\sum_{j=1}^{2^n} \frac{1}{j} = 1 + \sum_{m=0}^{n-1} \sum_{j=2^m+1}^{2^{m+1}} \frac{1}{j} \geq 1 + \frac{n}{2}$. **q.e.d.**

(iii) Für $k \in \mathbb{N}$ konvergiert die Reihe $\left(\sum \frac{1}{n(n+1)\cdots(n+k)}\right)_{n \in \mathbb{N}}$ gegen $\sum_{n=1}^{\infty} \frac{1}{n(n+1)\cdots(n+k)} = \frac{1}{k \cdot k!}$.

Beweis:

$$\begin{aligned} \sum_{n=1}^m \frac{1}{n(n+1)\cdots(n+k)} &= \sum_{n=1}^m \frac{1}{k} \frac{n+k-n}{n(n+1)\cdots(n+k)} \\ &= \frac{1}{k} \left(\sum_{n=1}^m \frac{1}{n \cdots (n+k-1)} - \sum_{n=2}^{m+1} \frac{1}{n \cdots (n+k-1)} \right) = \frac{1}{k} \left(\frac{1}{k!} - \frac{1}{(m+1) \cdots (m+k)} \right) \end{aligned}$$

Die Folge $\frac{1}{(m+1)\cdots(m+k)} < \frac{1}{m^k}$ konvergiert aber gegen Null. **q.e.d.**

Wir wenden das Cauchy Kriterium und das Monotonieprinzip auf Reihen an.

Satz 4.3 (Cauchy Kriterium für Reihen). *Die Reihe $(\sum a_n)_{n \in \mathbb{N}}$ konvergiert genau dann, wenn es für jedes $\epsilon > 0$ ein N gibt, so dass $\left|\sum_{j=n}^m a_j\right| < \epsilon$ für alle $m \geq n > N$ gilt. Insbesondere ist $(a_n)_{n \in \mathbb{N}}$ eine Nullfolge, wenn die Reihe $(\sum a_n)_{n \in \mathbb{N}}$ konvergiert.* **q.e.d.**

Satz 4.4 (Monotonieprinzip für Reihen). *Sei $(a_n)_{n \in \mathbb{N}}$ eine Folge von nicht negativen reellen Zahlen. Dann konvergiert die Reihe $(\sum a_n)_{n \in \mathbb{N}}$ genau dann, wenn sie beschränkt ist. Für den Grenzwert gilt dann $\sum_{n=1}^{\infty} a_n = \sup\{\sum_{n=1}^m a_n \mid m \in \mathbb{N}\}$.* **q.e.d.**

Definition 4.5 (absolut konvergent). *Die Reihe $(\sum a_n)_{n \in \mathbb{N}}$ heißt absolut konvergent, wenn die Reihe $(\sum |a_n|)_{n \in \mathbb{N}}$ konvergiert.*

Satz 4.6. *Jede absolut konvergente Reihe konvergiert. Und es gilt $\left|\sum_{n=1}^{\infty} a_n\right| \leq \sum_{n=1}^{\infty} |a_n|$.*

Beweis: Aus der Dreiecksungleichung folgt $\left|\sum_{j=n}^m a_j\right| \leq \sum_{j=n}^m |a_j|$ für alle $m \geq n$. Also ist die Reihe einer absolut konvergenten Reihe eine Cauchyfolge und konvergiert. Insbesondere gilt für alle $m \in \mathbb{N}$ auch $|\sum_{n=1}^m a_n| \leq \sum_{n=1}^m |a_n|$. Dann erfüllen auch die Grenzwerte $|\sum_{n=1}^{\infty} a_n| \leq \sum_{n=1}^{\infty} |a_n|$. **q.e.d.**

Aus dem Monotonieprinzip und Satz 3.5 folgt das

Satz 4.7 (Majoranten Kriterium). *Die Folgen von nicht negativen reellen Zahlen $(a_n)_{n \in \mathbb{N}}$ und $(b_n)_{n \in \mathbb{N}}$ erfüllen $b_n \leq a_n$ für alle $n \in \mathbb{N}$.*

- (i) *Wenn außerdem $(\sum a_n)_{n \in \mathbb{N}}$ konvergiert, dann konvergiert auch $(\sum b_n)_{n \in \mathbb{N}}$*
- (ii) *Wenn außerdem $(\sum b_n)_{n \in \mathbb{N}}$ divergiert, dann gilt $\sum_{n=1}^{\infty} b_n = \infty = \sum_{n=1}^{\infty} a_n$.* **q.e.d.**

Beispiel 4.8. Für alle $n, k \in \mathbb{N}$ ist $\frac{1}{(n+k)^{k+1}} \leq \frac{1}{n \cdots (n+k)}$. Also folgt aus der Konvergenz von $(\sum \frac{1}{n \cdots (n+k)})$ auch die Konvergenz von $(\sum \frac{1}{n^{k+1}})_{n \in \mathbb{N}}$.

Satz 4.9 (Wurzeltest). Sei $(a_n)_{n \in \mathbb{N}}$ eine Folge und sei $\alpha = \overline{\lim} \sqrt[n]{|a_n|}$, wenn $(|a_n|)_{n \in \mathbb{N}}$ nach oben beschränkt ist und ansonsten $\alpha = \infty$.

- (i) Falls $\alpha < 1$, dann konvergiert $(\sum a_n)_{n \in \mathbb{N}}$ absolut.
- (ii) Falls $\alpha > 1$, dann divergieren $(\sum a_n)_{n \in \mathbb{N}}$ und $(\sum |a_n|)_{n \in \mathbb{N}}$.

Im Fall $\overline{\lim} \sqrt[n]{|a_n|} = 1$ kann die Reihe $(\sum a_n)_{n \in \mathbb{N}}$ sowohl konvergent als auch divergent sein. So ist z.B. $\lim_{n \rightarrow \infty} \sqrt[n]{\frac{1}{n}} = 1 = \lim_{n \rightarrow \infty} \sqrt[n]{\frac{1}{n^2}}$. Aber die Reihe $(\sum \frac{1}{n})_{n \in \mathbb{N}}$ ist divergent, während die Reihe $(\sum \frac{1}{n^2})_{n \in \mathbb{N}}$ konvergent ist.

Beweis: (i) Sei $\alpha < 1$. Dann gibt es für jedes $\alpha < \beta < 1$ aufgrund von Satz 3.21 (i) ein $N \in \mathbb{N}$, so dass für alle $n \geq N$ gilt $\sqrt[n]{|a_n|} \leq \beta \iff |a_n| \leq \beta^n$. Weil aber $(\sum \beta^n)_{n \in \mathbb{N}}$ konvergiert, ist auch $(\sum a_n)_{n \in \mathbb{N}}$ absolut konvergent.

(ii) Sei $\alpha > 1$. Dann gibt es wieder aufgrund von Satz 3.21 (i) unendlich viele $\sqrt[n]{|a_n|} > 1$. Also kann die Folge $(|a_n|)_{n \in \mathbb{N}}$ nicht gegen Null konvergieren. Dann sind Reihen $(\sum a_n)_{n \in \mathbb{N}}$ und $(\sum |a_n|)_{n \in \mathbb{N}}$ keine Cauchyfolgen, also divergent. **q.e.d.**

Satz 4.10 (Exponentialfunktion). Für alle $x \in \mathbb{K}$ definieren wir

$$\exp(x) : \mathbb{K} \rightarrow \mathbb{K}, \quad x \mapsto \exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + \sum_{n=1}^{\infty} \frac{x^n}{n!}$$

Aufgrund des Beispiels (v) im letzten Kapitel ist $\lim_{n \rightarrow \infty} \sqrt[n]{n!} = \infty$. Für $\epsilon > 0$ gibt es also

$$N \in \mathbb{N} \quad \text{mit} \quad \sqrt[n]{n!} > \frac{|x| + 1}{\epsilon} \quad \text{für} \quad n \geq N \quad \implies \quad \sqrt[n]{\frac{|x^n|}{n!}} = \frac{|x|}{\sqrt[n]{n!}} < \frac{|x|\epsilon}{|x| + 1} < \epsilon.$$

Deshalb ist $\lim_{n \rightarrow \infty} \sqrt[n]{\frac{|x|^n}{n!}} = 0$ und die Reihe $(\sum \frac{x^n}{n!})_{n \in \mathbb{N}_0}$ konvergiert absolut.

Satz 4.11 (Quotiententest). Sei $(a_n)_{n \in \mathbb{N}}$ eine Folge in $\mathbb{K} \setminus \{0\}$.

- (i) Falls $\overline{\lim} |\frac{a_{n+1}}{a_n}| < 1$, dann konvergiert $(\sum a_n)_{n \in \mathbb{N}}$ absolut.
- (ii) Falls es ein $N \in \mathbb{N}$ gibt, so dass $|\frac{a_{n+1}}{a_n}| \geq 1$ alle $n \geq N$ erfüllen (das folgt wegen Satz 3.21 (ii) aus $\underline{\lim} |\frac{a_{n+1}}{a_n}| > 1$), dann divergieren $(\sum a_n)_{n \in \mathbb{N}}$ und $(\sum |a_n|)_{n \in \mathbb{N}}$.

Beweis: (i) Wegen $\overline{\lim} \sqrt[n]{|a_n|} \leq \overline{\lim} |\frac{a_{n+1}}{a_n}|$ (Satz 3.25) folgt (i) aus dem Wurzeltest.

(ii) Aus $|\frac{a_{n+1}}{a_n}| \geq 1$ folgt $|a_{n+1}| = |\frac{a_{n+1}}{a_n}| \cdot |a_n| \geq |a_n|$. Also ist $(|a_n|)_{n \in \mathbb{N}}$ keine Nullfolge und, die Reihen $(\sum a_n)_{n \in \mathbb{N}}$ und $(\sum |a_n|)_{n \in \mathbb{N}}$ sind divergent. **q.e.d.**

Beispiel 4.12. Für $\overline{\lim} \left| \frac{a_{n+1}}{a_n} \right| \geq 1$ und $\underline{\lim} \left| \frac{a_{n+1}}{a_n} \right| \leq 1$ kann die Reihe konvergieren und divergieren: Die Reihe zu $a_n = n^{-(l+(-1)^n)}$ für $n \in \mathbb{N}$ divergiert für $l = 2$ und konvergiert für $l = 3$ und hat in beiden Fällen $\overline{\lim} \left| \frac{a_{n+1}}{a_n} \right| = \infty$ und $\underline{\lim} \left| \frac{a_{n+1}}{a_n} \right| = 0$.

Satz 4.13 (Cauchy's Verdichtungssatz). Sei $(a_n)_{n \in \mathbb{N}}$ eine nicht negative monoton fallende Folge. Dann konvergiert $(\sum a_n)_{n \in \mathbb{N}}$ genau dann, wenn $(\sum 2^n a_{2^n})_{n \in \mathbb{N}}$ konvergiert.

Beweis: Für alle $n \in \mathbb{N}$ sei $s_n = \sum_{j=1}^n a_j$ und $t_n = \sum_{j=0}^n 2^j a_{2^j}$. Wegen der Monotonie von $(a_n)_{n \in \mathbb{N}}$ gilt für alle $j \in \mathbb{N}$:

$$a_{2^j} + a_{2^j+1} + \dots + a_{2^{j+1}-1} \leq 2^j a_{2^j} \leq 2(a_{2^{j-1}+1} + a_{2^{j-1}+2} + \dots + a_{2^j})$$

und für $j = 0$ gilt: $a_1 \leq a_1 \leq 2a_1$. Deshalb gilt für alle $n \in \mathbb{N}$

$$0 \leq s_{2^{n+1}-1} \leq t_n \leq 2s_{2^n}.$$

Also ist die Beschränktheit von der Folge $(t_n)_{n \in \mathbb{N}}$ äquivalent zu der von $(s_n)_{n \in \mathbb{N}}$. **q.e.d.**

Die Reihe $(\sum \frac{1}{n^s})_{n \in \mathbb{N}}$ ist für $s \in \mathbb{Q} \setminus \{0\}$ genau dann konvergent, wenn die Reihe

$$\left(\sum \frac{2^n}{(2^n)^s} \right)_{n \in \mathbb{N}_0} = \left(\sum 2^{(1-s) \cdot n} \right)_{n \in \mathbb{N}_0}$$

konvergent ist, also für $s > 1$. $\implies$ Für alle $s \in \mathbb{Q}$ mit $s > 1$ ist $\zeta(s)$ wohl definiert.

Satz 4.14 (Alternierende Reihe von Leibniz). Sei $(a_n)_{n \in \mathbb{N}_0}$ eine monoton fallende reelle Nullfolge. Dann konvergiert $(\sum (-1)^n a_n)_{n \in \mathbb{N}_0}$

Beweis: Wegen dem Monotonieprinzip sind die Glieder einer monoton fallenden Nullfolge nicht negativ. Sei $s_n = \sum_{m=0}^n (-1)^m a_m$ für alle $n \in \mathbb{N}_0$. Aus der Monotonie folgt:

$$s_1 \leq s_3 \leq \dots \leq s_{2n+1} \leq \dots \leq s_{2n} \leq \dots \leq s_2 \leq s_0.$$

Also ist die Folge $(s_{2n+1})_{n \in \mathbb{N}_0}$ monoton wachsend und beschränkt und $(s_{2n})_{n \in \mathbb{N}_0}$ monoton fallend und beschränkt. Dann konvergieren aber beide Folgen und es gilt

$$\lim_{n \rightarrow \infty} s_{2n} - \lim_{n \rightarrow \infty} s_{2n+1} = - \lim_{n \rightarrow \infty} (-1)^{2n+1} a_{2n+1} = \lim_{n \rightarrow \infty} a_{2n+1} = 0$$

Also konvergiert die Reihe $(\sum (-1)^n a_n)_{n \in \mathbb{N}}$.

q.e.d.

Also ist $\sum_{n=1}^{\infty} \frac{(-1)^n}{n}$ konvergent, aber nicht absolut konvergent, weil $\sum_{n=1}^{\infty} \frac{1}{n} = \infty$.

4.2 Dezimalbruchdarstellung von reellen Zahlen

Als Ziffern wählen wir $Z = \{0, 1, \dots, 9\}$ (bzw. $\{0, 1, \dots, p-1\}$ mit $p \in \mathbb{N} \setminus \{1\}$). Für jede Folge $(z_n)_{n \in \mathbb{N}}$ mit Werten in Z definieren wir die Zahlenfolge $(\sum x_n)_{n \in \mathbb{N}}$ mit $x_n = \frac{z_n}{p^n}$ für alle $n \in \mathbb{N}$. Dann ist $(\sum x_n)_{n \in \mathbb{N}}$ wegen dem Majorantenkriterium und wegen

$$\sum_{n=1}^{\infty} \frac{p-1}{p^n} = \frac{p-1}{p} \sum_{n=0}^{\infty} \frac{1}{p^n} = \frac{p-1}{p} \frac{1}{1-\frac{1}{p}} = 1$$

absolut konvergent. Also definiert $x = \sum_{n=1}^{\infty} x_n$ eine reelle Zahl in $[0, 1]$. Sei jetzt M die Menge aller Folgen $M = \{(z_n)_{n \in \mathbb{N}} \mid (z_n)_{n \in \mathbb{N}} \text{ konvergiert nicht gegen } p-1\}$. Wir zeigen jetzt, dass die Abbildung $M \rightarrow [0, 1), (z_n)_{n \in \mathbb{N}} \mapsto \sum_{n=0}^{\infty} \frac{z_n}{p^n}$ surjektiv und injektiv ist

Bemerkung 4.15. Wir können auch fordern, dass $(z_n)_{n \in \mathbb{N}}$ nicht gegen Null konvergiert. Reelle Zahlen, deren Ziffernfolge gegen 0 konvergiert, haben auch eine Dezimalbruchdarstellung, deren Ziffernfolge gegen 9 konvergiert, z.B. $\frac{1}{2} = 0,500\dots = 0,499\dots$

Surjektiv: Sei $x \in [0, 1)$. Wir definieren $(z_n)_{n \in \mathbb{N}}$ induktiv, so dass für jedes $n \in \mathbb{N}$

$$0 \leq \frac{z_n}{p^n} \leq x - \sum_{m=1}^{n-1} \frac{z_m}{p^m} < \frac{z_n + 1}{p^n} \leq \frac{1}{p^{n-1}} \Leftrightarrow x \in \left[\frac{z_n}{p^n} + \sum_{m=1}^{n-1} \frac{z_m}{p^m}, \frac{z_n + 1}{p^n} + \sum_{m=1}^{n-1} \frac{z_m}{p^m} \right)$$

gilt. Es folgt $0 \leq x - \sum_{m=1}^n \frac{z_m}{p^m} < \frac{1}{p^n}$ und $\sum_{n=1}^{\infty} \frac{z_n}{p^n} = x$. Weil für alle $n \in \mathbb{N}$

$$\sum_{m=n+1}^{\infty} \frac{p-1}{p^m} = \frac{p-1}{p^{n+1} \left(1 - \frac{1}{p}\right)} = \frac{1}{p^n} \text{ gilt, konvergiert } (z_n)_{n \in \mathbb{N}} \text{ nicht gegen } p-1.$$

Injektiv: Seien $(z_n)_{n \in \mathbb{N}}$ und $(w_n)_{n \in \mathbb{N}}$ Ziffernfolgen, mit $\sum_{n=1}^{\infty} \frac{z_n}{p^n} = \sum_{n=1}^{\infty} \frac{w_n}{p^n}$. Sei $n \in \mathbb{N}$ der kleinste Index mit $z_n \neq w_n$, also $\sum_{m=n}^{\infty} \frac{z_m - w_m}{p^m} = 0$ und o.B.d.A. $z_n > w_n$.

$$\Rightarrow \frac{z_n - w_n}{p^n} = \sum_{m=n+1}^{\infty} \frac{w_m - z_m}{p^m} \leq \sum_{m=n+1}^{\infty} \frac{p-1}{p^m} = \frac{p-1}{p^{n+1}} \sum_{m=0}^{\infty} \frac{1}{p^m} = \frac{p-1}{p^{n+1}} \frac{1}{1-\frac{1}{p}} = \frac{1}{p^n}.$$

Also ist $z_n - w_n \leq 1$ und damit $z_n = w_n + 1$. Für alle $m > n$ gilt dann $w_m - z_m = p-1 \Rightarrow z_m = 0$ und $w_m = p-1 \Rightarrow \lim_{m \rightarrow \infty} w_m = p-1$ und $(w_n)_{n \in \mathbb{N}} \notin M$. **q.e.d.**

Übungsaufgabe 4.16. Wir ordnen den Teilmengen $A \subset \mathbb{N}$ die Folgen mit Werten in $\{0, 1\}$ zu, die nur auf den Elementen von A 1 sind. Zeige mithilfe der bijektiven Abbildung $M \simeq [0, 1)$ für $p = 2$, dass die Potenzmenge von $\mathbb{N}$ gleichmächtig zu $(0, 1)$ und zu $\mathbb{R}$ ist (Tipp: Zeige zuerst mit Satz 2.40, dass $(0, 1)$ gleichmächtig ist zu $(0, 1) \cup \mathbb{N}_0$).

4.3 Addition, Multiplikation, Umordnung

Aus dem Satz 3.5 folgt

Satz 4.17 (Rechenregeln für Reihen). *Die Reihen $(\sum a_n)_{n \in \mathbb{N}}$ und $(\sum b_n)_{n \in \mathbb{N}}$ seien konvergent, dann konvergieren auch die Reihen*

$$\left(\sum (a_n + b_n)\right)_{n \in \mathbb{N}} \quad \text{und} \quad \left(\sum \lambda a_n\right)_{n \in \mathbb{N}} \quad \text{für alle } \lambda \in \mathbb{K}. \quad \text{q.e.d.}$$

Definition 4.18. *Für gegebene Reihen $(\sum a_n)_{n \in \mathbb{N}_0}$ und $(\sum b_n)_{n \in \mathbb{N}_0}$ heißt die Reihe $(\sum c_n)_{n \in \mathbb{N}_0}$ mit $c_n = \sum_{k=0}^n a_k b_{n-k}$ für $n \in \mathbb{N}_0$ das (Cauchy-)Produkt der beiden Reihen.*

Diese Definition stammt aus der Theorie der Potenzreihen. Das Produkt der beiden Potenzreihen $(\sum a_n x^n)_{n \in \mathbb{N}_0}$ und $(\sum b_n x^n)_{n \in \mathbb{N}_0}$ ist die Potenzreihe $(\sum c_n x^n)_{n \in \mathbb{N}_0}$, d.h. wir haben alle Summanden des Produktes mit gleichen Potenzen zusammengefasst.

Satz 4.19 (Konvergenz des Produktes). *Wenn die Reihe $(\sum a_n)_{n \in \mathbb{N}_0}$ absolut konvergiert und die Reihe $(\sum b_n)_{n \in \mathbb{N}_0}$ konvergiert, dann konvergiert das Produkt $(\sum c_n)_{n \in \mathbb{N}_0}$ der beiden Reihen. Wenn auch $(\sum b_n)_{n \in \mathbb{N}_0}$ absolut konvergiert, dann auch $(\sum c_n)_{n \in \mathbb{N}_0}$.*

Beweis: Mit $c_n = \sum_{k=0}^n a_k b_{n-k}$, $A_n = \sum_{k=0}^n a_k$, $B_n = \sum_{k=0}^n b_k$ und $C_n = \sum_{k=0}^n c_k$ für $n \in \mathbb{N}_0$ gilt

$$\begin{aligned} C_n &= a_0 b_0 + (a_0 b_1 + a_1 b_0) + \dots + (a_0 b_n + \dots + a_n b_0) \\ &= a_0 B_n + a_1 B_{n-1} + \dots + a_n B_0 \\ &= a_0 (B - \beta_n) + a_1 (B - \beta_{n-1}) + \dots + a_n (B - \beta_0) \end{aligned}$$

Hierbei ist $B = \lim_{n \rightarrow \infty} B_n = \sum_{n=0}^{\infty} b_n$ und $\beta_n = B - B_n = \sum_{k=n+1}^{\infty} b_k$. Daraus ergibt sich

$$C_n = A_n \cdot B - (a_0 \beta_n + a_1 \beta_{n-1} + \dots + a_n \beta_0).$$

Also gilt $\lim_{n \rightarrow \infty} C_n = \lim_{n \rightarrow \infty} A_n B$, wenn $a_0 \beta_n + \dots + a_n \beta_0$ im Grenzwert $n \rightarrow \infty$ gegen Null konvergiert. Es gibt tatsächlich $\alpha, \beta > 0$ und für jedes $\epsilon > 0$ $N, M \in \mathbb{N}$ mit

$$|\beta_n| \leq \beta \quad \text{und} \quad \sum_{k=0}^n |a_k| \leq \alpha \quad \text{für } n \in \mathbb{N}_0, \quad |\beta_n| < \frac{\epsilon}{2\alpha} \quad \text{für } n \geq N, \quad |a_n| < \frac{\epsilon}{2N\beta} \quad \text{für } n \geq M.$$

Dann gilt für alle $n \geq N + M - 1$, also $n - N + 1 \geq M$:

$$\begin{aligned} |a_0 \beta_n + \dots + a_n \beta_0| &\leq |a_0 \beta_n + \dots + a_{n-N} \beta_N| + |a_{n-N+1} \beta_{N-1} + \dots + a_n \beta_0| \\ &< \frac{\epsilon}{2\alpha} \sum_{k=0}^{n-N} |a_k| + N\beta \cdot \frac{\epsilon}{2N\beta} \leq \epsilon. \end{aligned}$$

Also konvergiert $(\sum c_n)_{n \in \mathbb{N}_0}$ gegen das Produkt der Grenzwerte von $(\sum a_n)_{n \in \mathbb{N}_0}$ und $(\sum b_n)_{n \in \mathbb{N}_0}$ wenn beide konvergieren und eine absolut konvergiert. Wenn beide Reihen absolut konvergieren, dann gilt für alle $N \in \mathbb{N}_0$

$$\sum_{n=0}^N |c_n| \leq \left(\sum_{n=0}^N |a_n| \right) \left(\sum_{n=0}^N |b_n| \right) \leq \left(\sum_{n=0}^{\infty} |a_n| \right) \left(\sum_{n=0}^{\infty} |b_n| \right).$$

Also ist dann das Produkt absolut konvergent.

q.e.d.

Beispiel 4.20. Das Quadrat der Reihe $\left(\sum \frac{(-1)^n}{\sqrt{n+1}} \right)_{n \in \mathbb{N}_0}$ ist nicht konvergent, obwohl die Reihe als Beispiel einer alternierenden Reihe nach Leibniz konvergiert, aber nicht absolut konvergiert. Die Koeffizienten des Quadrates sind gegeben durch:

$$\sum_{k=0}^n \frac{(-1)^k (-1)^{n-k}}{\sqrt{k+1} \sqrt{n-k+1}} = (-1)^n \sum_{k=0}^n \frac{1}{\sqrt{k+1} \sqrt{n-k+1}}$$

Es gilt aber $\sum_{k=0}^n \frac{1}{\sqrt{k+1} \sqrt{n-k+1}} \geq \sum_{k=0}^n \frac{1}{\sqrt{n+1} \sqrt{n+1}} = 1$. Also ist das Quadrat der Reihe keine Cauchyfolge.

q.e.d.

Satz 4.21 (Eigenschaften der Exponentialfunktion).

- (i) für alle $x, y \in \mathbb{K}$ gilt $\exp(x+y) = \exp(x) \exp(y)$.
- (ii) Für alle $x \in \mathbb{K}$ ist $\exp(x) \neq 0$ und für alle $x \in \mathbb{R}$ sogar $\exp(x) > 0$.
- (iii) Für alle $x \in \mathbb{R}$ und alle $r \in \mathbb{Q}$ ist $\exp(rx) = (\exp(x))^r$.
- (iv) Für alle $x \in \mathbb{C}$ ist $\exp(\bar{x}) = \overline{\exp(x)}$.
- (v) Für alle $x \in \mathbb{R}$ ist $|\exp(ix)| = 1$.

Die Zahl $\exp(1)$ wird Eulersche Zahl genannt und mit e bezeichnet. Wegen (iii) gilt dann für alle $r \in \mathbb{Q}$: $\exp(r) = \exp(r \cdot 1) = e^r$. Deshalb schreiben wir auch e^x für $\exp(x)$.

Beweis: (i) Wir hatten schon gesehen, dass die Reihe $\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}$ für alle $x \in \mathbb{K}$ absolut konvergiert. Dann ist das Produkt von $\exp(x) \cdot \exp(y) = \sum_{n=0}^{\infty} \sum_{k=0}^n \frac{x^k y^{n-k}}{k!(n-k)!}$. Wegen der binomischen Formel gilt

$$\sum_{k=0}^n \frac{x^k y^{n-k}}{k!(n-k)!} = \frac{1}{n!} \sum_{k=0}^n \binom{n}{k} x^k y^{n-k} = \frac{(x+y)^n}{n!}.$$

Dann folgt $\exp(x)\exp(y) = \sum_{n=0}^{\infty} \frac{(x+y)^n}{n!} = \exp(x+y)$.

(ii) Wegen (i) gilt $\exp(x)\exp(-x) = 1$. Also besitzt $\exp(x)$ für alle $x \in \mathbb{K}$ ein Inverses und ist ungleich Null. Wegen (i) gilt $\exp(x) = \left(\exp\left(\frac{x}{2}\right)\right)^2 > 0$ für alle $x \in \mathbb{R}$.

(iii) Offenbar ist für alle $n \in \mathbb{N}, m \in \mathbb{Z}$ $(\exp(\frac{x \cdot m}{n}))^n = \exp(x \cdot m) = (\exp(x))^m$. Also gilt auch wegen (ii) $\exp(\frac{xm}{n}) = (\exp(x))^{\frac{m}{n}}$.

(iv) $\exp(\bar{x}) = \sum_{n=0}^{\infty} \frac{\bar{x}^n}{n!} = \overline{\sum_{n=0}^{\infty} \frac{x^n}{n!}} = \overline{\exp(x)}$ wegen Satz 2.57, Satz 3.3 (ii) und Satz 4.17.

(v) Für $x \in \mathbb{R}$ gilt $\exp(ix)\overline{\exp(ix)} = \exp(ix)\exp(-ix) = 1$. **q.e.d.**

Wir können jetzt für jede Zahl $y > 0$, für die es ein $x \in \mathbb{R}$ gibt, so dass $y = \exp(x)$ gilt, und für jedes $z \in \mathbb{R}$ die Zahl $y^z = \exp(zx)$ definieren. Wir werden später sehen, dass wir so für alle $y \in (0, \infty)$ und alle $z \in \mathbb{R}$ y^z definieren können.

Definition 4.22 (Umordnen von Reihen). Sei $\tau : \mathbb{N} \rightarrow \mathbb{N}$ eine bijektive Abbildung von den natürlichen Zahlen auf sich selber. Dann heißt die Reihe $(\sum a_{\tau(n)})_{n \in \mathbb{N}}$ eine Umordnung der Reihe $(\sum a_n)_{n \in \mathbb{N}}$.

Analog könnten wir auch die Umordnung von Folgen bilden. Letztere sind aber weniger interessant, weil jede Umordnung einer konvergenten Folge wieder gegen den gleichen Grenzwert konvergiert (Übungsaufgabe). Dagegen gilt dies bei Reihen nicht.

Satz 4.23. Konvergiert eine Reihe $(\sum a_n)_{n \in \mathbb{N}}$ absolut, so konvergiert auch jede Umordnung $(\sum a_{\tau(n)})_{n \in \mathbb{N}}$ absolut. In diesem Fall gilt
$$\sum_{n=1}^{\infty} a_n = \sum_{n=1}^{\infty} a_{\tau(n)}.$$

Beweis: Sei also τ eine gegebene bijektive Abbildung $\tau : \mathbb{N} \rightarrow \mathbb{N}$. Wenn $(\sum a_n)_{n \in \mathbb{N}}$ absolut konvergent, dann gibt es für jedes $\epsilon > 0$ ein $N \in \mathbb{N}$, so dass $\sum_{k=n}^m |a_k| < \epsilon$ für alle $m \geq n > N$ gilt. Dann gibt es auch ein $M = \max\{\tau^{-1}(1), \dots, \tau^{-1}(N)\} \in \mathbb{N}$, so dass $\tau(n) > N$ für alle $n > M$ gilt. Dann folgt für alle $m \geq n > M$

$$\sum_{k=n}^m |a_{\tau(k)}| \leq \sum_{k=N+1}^{\max\{\tau(n), \tau(n+1), \dots, \tau(m)\}} |a_k| < \epsilon.$$

Also ist $(\sum |a_{\tau(n)}|)_{n \in \mathbb{N}}$ eine Cauchyfolge und die Umordnung konvergiert sogar absolut.

Mit denselben N und M in Abhängigkeit von $\epsilon > 0$ gilt für alle $m > M \geq N$

$$\left| \sum_{k=1}^m a_{\tau(k)} - \sum_{k=1}^m a_k \right| \leq \sum_{k=N+1}^{\max\{\tau(1), \tau(2), \dots, \tau(m), m\}} |a_k| < \epsilon.$$

Also konvergiert $(\sum a_{\tau(n)})_{n \in \mathbb{N}}$ gegen den gleichen Grenzwert wie $(\sum a_n)_{n \in \mathbb{N}}$. **q.e.d.**

Satz 4.24 (Riemann). *Sei $(\sum a_n)_{n \in \mathbb{N}}$ eine konvergente reelle Reihe, die nicht absolut konvergiert, und $\alpha \leq \beta$ zwei reelle Zahlen. Dann gibt es eine Umordnung $(\sum a_{\tau(n)})_{n \in \mathbb{N}}$, die als Reihe beschränkt ist und für die α der Limes inferior der Reihe ist und β der Limes superior. Wenn $\alpha \neq \beta$ konvergiert die Reihe also nicht.*

Beweis*: Weil $(\sum a_n)_{n \in \mathbb{N}}$ konvergiert, ist die Folge $(a_n)_{n \in \mathbb{N}}$ eine Nullfolge. Wir betrachten im folgenden die beiden Teilfolgen aller nichtnegativen Elemente $(b_n)_{n \in \mathbb{N}}$ und aller negativen Elemente $(c_n)_{n \in \mathbb{N}}$. Weil $(\sum a_n)_{n \in \mathbb{N}}$ nicht absolut konvergiert, und $(\sum 2b_n)_{n \in \mathbb{N}}$ genau dann konvergiert, wenn $(\sum a_n + |a_n|)_{n \in \mathbb{N}}$ konvergiert, bzw. $(\sum 2c_n)_{n \in \mathbb{N}}$ genau dann konvergiert, wenn $(\sum a_n - |a_n|)_{n \in \mathbb{N}}$ konvergiert, konvergieren die beiden monotonen Reihen $(\sum b_n)_{n \in \mathbb{N}}$ und $(\sum c_n)_{n \in \mathbb{N}}$ nicht und sind auch nicht beschränkt.

Wir setzen die umgeordnete Folge $(a_{\tau(n)})_{n \in \mathbb{N}}$ abwechselnd jeweils der Reihe nach aus den Folgen $(b_n)_{n \in \mathbb{N}}$ und $(c_n)_{n \in \mathbb{N}}$ zusammen. Wenn die Summe aller bisherigen Folgenglieder größer ist als β , dann fahren wir solange mit Folgengliedern aus c_n fort, bis die Summe kleiner als α ist. Wenn die Summe aller bisherigen Folgenglieder kleiner als α ist, dann fahren wir solange mit Folgengliedern aus b_n fort, bis die Summe größer als β ist. Wenn $0 \in [\alpha, \beta]$ starten wir mit Folgengliedern aus b_n , bis die Summe größer als β ist. Für jedes $\epsilon > 0$ gibt es nur endlich viele Glieder von $(a_n)_{n \in \mathbb{N}}$, deren Absolutbetrag nicht kleiner als ϵ ist. Deshalb gibt es ein $N \in \mathbb{N}$, so dass die Summen $\sum_{k=1}^n a_{\tau(k)}$ für alle $n \geq N$ in $(\alpha - \epsilon, \beta + \epsilon)$ liegen. Es gibt unendlich viele Glieder von $(\sum a_{\tau(n)})_{n \in \mathbb{N}}$ die kleiner als α sind und unendlich viele, die größer als β sind. Die umgeordnete Reihe $(\sum a_{\tau(n)})_{n \in \mathbb{N}}$ hat also als Limes inferior α und als Limes superior β . **q.e.d.**

Weil für jedes $\epsilon > 0$ und $x \in [\alpha, \beta]$ unendlich viele Glieder von $(\sum a_{\tau(n)})_{n \in \mathbb{N}}$ in $(x - \epsilon, x + \epsilon)$ liegen, ist die Menge der Häufungspunkte dieser Reihe gleich $[\alpha, \beta]$.

Definition 4.25. *Sei $(a_n)_{n \in \mathbb{N}_0}$ eine Folge, dann heißt $(\sum a_n x^n)_{n \in \mathbb{N}_0}$ die entsprechende Potenzreihe mit Koeffizienten $(a_n)_{n \in \mathbb{N}_0}$.*

Aus dem Wurzeltest folgt sofort

Satz 4.26 (Konvergenzradius von Potenzreihen). *Seien $(a_n)_{n \in \mathbb{N}_0}$ die Koeffizienten der Potenzreihe $(\sum a_n x^n)_{n \in \mathbb{N}_0}$ und $\alpha = \lim_{n \rightarrow \infty} \sqrt[n]{|a_n|}$ wenn $(\sqrt[n]{|a_n|})_{n \in \mathbb{N}_0}$ nach oben beschränkt ist und sonst $\alpha = \infty$, und sei $R = \frac{1}{\alpha}$ ($R = 0$ für $\alpha = \infty$ und $R = \infty$ für $\alpha = 0$).*

(i) *Für $|x| < R$ konvergiert $(\sum a_n x^n)_{n \in \mathbb{N}_0}$ absolut.*

(ii) *Für $|x| > R$ divergiert $(\sum a_n x^n)_{n \in \mathbb{N}_0}$.*

Wenn $\alpha < \infty$ definiert die Potenzreihe also folgende Funktion:

$$f : \{x \in \mathbb{K} \mid |x| < R\} \rightarrow \mathbb{K}, \quad x \mapsto \sum_{n=0}^{\infty} a_n x^n. \quad \mathbf{q.e.d.}$$

Beispiel 4.27. (i) $\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}$, $\alpha = \overline{\lim} \sqrt[n]{\frac{1}{n!}} = \frac{1}{\lim_{n \rightarrow \infty} \sqrt[n]{n!}} = 0 \Rightarrow R = \infty$.

(ii) $\left(\sum_{n \in \mathbb{N}_0} x^n\right)$ also $\alpha = \overline{\lim} \sqrt[n]{1} = 1 \Rightarrow R = 1$. Für $|x| < 1$ gilt $\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}$.

(iii) $\left(\sum_{n \in \mathbb{N}} \frac{x^n}{n}\right)$ also $\alpha = \overline{\lim} \sqrt[n]{\frac{1}{n}} = \frac{1}{\lim_{n \rightarrow \infty} \sqrt[n]{n}} = 1 \Rightarrow R = 1$.

Für $|x| < 1$ ist die Potenzreihe also konvergent, aber für $x = 1$ divergent und für $x = -1$ konvergent (alternierende Reihe von Leibniz).

(iv) $\left(\sum_{n \in \mathbb{N}} \frac{x^n}{n^2}\right)$ also $\alpha = \overline{\lim} \sqrt[n]{\frac{1}{n^2}} = \left(\frac{1}{\lim_{n \rightarrow \infty} \sqrt[n]{n}}\right)^2 = 1 \Rightarrow R = 1$.

Für $|x| \leq 1$ also konvergent und für $|x| > 1$ divergent.

Satz 4.28 (Eigenschaften von Potenzreihenfunktionen).

(i) Seien $(\sum a_n x^n)_{n \in \mathbb{N}_0}$ und $(\sum b_n x^n)_{n \in \mathbb{N}_0}$ Potenzreihen mit Konvergenzradius R_1 bzw. R_2 . Dann konvergieren für $|x| < \min\{R_1, R_2\}$ die Summe $(\sum (a_n + b_n) x^n)_{n \in \mathbb{N}_0}$ und das Cauchyprodukt $(\sum c_n x^n)_{n \in \mathbb{N}}$ der beiden Potenzreihen und es gilt

$$\sum_{n=0}^{\infty} (a_n + b_n) x^n = \sum_{n=0}^{\infty} a_n x^n + \sum_{n=0}^{\infty} b_n x^n \text{ und } \sum_{n=0}^{\infty} c_n x^n = \left(\sum_{n=0}^{\infty} a_n x^n\right) \left(\sum_{n=0}^{\infty} b_n x^n\right).$$

(ii) Für $r < R_1$ gibt es ein $M(r) \in (0, \infty)$, so dass $\left|\sum_{n=0}^{\infty} a_n x^n\right| \leq M(r)$ für $|x| \leq r$ gilt.

(iii) Für $r < R_1$ und für $\epsilon > 0$ gibt es ein $N \in \mathbb{N}$, so dass für $|x| \leq r$ folgendes gilt:

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

(iv) Für $r < R_1$ gibt es ein $L > 0$, so dass für x, y mit $|x|, |y| \leq r$ folgendes gilt:

$$\left|\sum_{n=0}^{\infty} a_n x^n - \sum_{n=0}^{\infty} a_n y^n\right| \leq L|x - y|.$$

Beweis: (i) Folgt aus den Rechenregeln für Reihen und der Konvergenz des Produktes.

(ii) Für $|x| \leq r$ gilt $\left|\sum_{n=0}^{\infty} a_n x^n\right| \leq \sum_{n=0}^{\infty} |a_n| r^n = M(r) < \infty$.

(iii) Weil die Reihe $\sum_{n=0}^{\infty} |a_n| r^n$ absolut konvergiert gibt es für jedes $\epsilon > 0$ ein $N \in \mathbb{N}$, so dass $\left| \sum_{n=0}^{\infty} |a_n| - \sum_{n=0}^N |a_n| \right| = \sum_{n=N+1}^{\infty} |a_n| r^n < \epsilon$ gilt. Dann folgt für $|x| \leq r$

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

(iv) $(x^n - y^n) = (x - y)(x^{n-1} + x^{n-2}y + \dots + xy^{n-2} + y^{n-1})$. Für $|x| \leq r$ und $|y| \leq r$ folgt also $|x^n - y^n| \leq |x - y| n r^{n-1}$. Weil aber gilt

$$\overline{\lim} \sqrt[n]{n|a_n|} = \overline{\lim} \left(\sqrt[n]{n} \cdot \sqrt[n]{|a_n|} \right) = \left(\lim_{n \rightarrow \infty} \sqrt[n]{n} \right) \overline{\lim} \sqrt[n]{|a_n|} = \overline{\lim} \sqrt[n]{|a_n|},$$

haben die Reihen $(\sum a_n x^n)_{n \in \mathbb{N}_0}$ und $(\sum n \cdot a_n x^{n-1})_{n \in \mathbb{N}_0}$ den gleichen Konvergenzradius. Also gilt für $|x| \leq r$ und $|y| \leq r$

$$\left| \sum_{n=0}^{\infty} a_n x^n - \sum_{n=0}^{\infty} a_n y^n \right| \leq \sum_{n=0}^{\infty} |a_n| \cdot |x^n - y^n| \leq |x - y| L \text{ mit } L = \sum_{n=1}^{\infty} |a_n| \cdot n r^{n-1}. \quad \mathbf{q.e.d.}$$

Satz 4.29 (Identitätssatz für Potenzreihenfunktionen).

- (i) Sei $\sum_{n=0}^{\infty} a_n x^n$ eine Potenzreihe mit Konvergenzradius $R > 0$, die nicht identisch verschwindet, dann hat die Potenzreihenfunktion für ein $0 < r < R$ in $\{x \in \mathbb{K} \mid 0 < |x| < r\}$ keine Nullstelle.
- (ii) Sei $\sum_{n=0}^{\infty} a_n x^n$ eine Potenzreihenfunktion mit Konvergenzradius $R > 0$. Für alle x_0 mit $|x_0| < R$ und alle $n \in \mathbb{N}_0$ konvergiert $b_n = \sum_{k=0}^{\infty} \binom{n+k}{k} a_{n+k} x_0^k$. Außerdem ist der Konvergenzradius der Potenzreihenfunktion $\sum_{n=0}^{\infty} b_n x^n$ nicht kleiner als $R - |x_0|$ und für alle $|x| < R - |x_0|$ gilt auch $\sum_{n=0}^{\infty} b_n x^n = \sum_{n=0}^{\infty} a_n (x + x_0)^n$.
- (iii) Seien $\sum_{n=0}^{\infty} a_n x^n$ und $\sum_{n=0}^{\infty} b_n x^n$ zwei Potenzreihenfunktionen, deren Konvergenzradien größer sind als $r > 0$. Falls $\{x \in \mathbb{K} \mid |x| \leq r\}$ unendliche viele verschiedene Elemente enthält, an denen die beiden Potenzreihenfunktionen übereinstimmen, dann gilt $a_n = b_n$ für alle $n \in \mathbb{N}_0$, d.h. sie stimmen als Potenzreihen überein.

Beweis: (i) Sei $N \in \mathbb{N}_0$ der kleinste Index, so dass $a_N \neq 0$. Wenn alle anderen Koeffizienten $(a_n)_{n > N}$ verschwinden, hat die Potenzreihe höchstens Nullstellen bei $x = 0$. Andernfalls gilt für ein $0 < r < R$ und alle $|x| \leq r$

$$\left| \sum_{n=0}^{\infty} a_n x^n - a_N x^N \right| \leq \sum_{n=N+1}^{\infty} |a_n| \cdot |x|^n \leq |x|^{N+1} \left(\sum_{n=0}^{\infty} |a_{n+N+1}| r^n \right).$$

Also gilt für alle Nullstellen x_m mit $|x_m| \leq r$ der Potenzreihenfunktionen

$$|a_N| \cdot |x_m|^N \leq |x_m|^{N+1} \left(\sum_{n=0}^{\infty} |a_{n+N+1}| r^n \right).$$

Wenn $|x_m| \neq 0$ ist folgt daraus $0 < |a_N| \leq |x_m|(\sum_{n=0}^{\infty} |a_{n+N+1}|r^n)$. Also hat die Potenzreihenfunktion keine Nullstelle in

$$\left\{ x \in \mathbb{K} \mid 0 < |x| < \min \left\{ r, |a_N| \left(\sum_{n=0}^{\infty} |a_{n+N+1}|r^n \right)^{-1} \right\} \right\}.$$

(ii) Für $x_0 = 0$ trivial. Sei $0 < |x_0| = r < R$. Dann ist für alle $0 < s < R - r$ die Reihe

$$\sum_{n=0}^{\infty} a_n(r+s)^n = \sum_{n=0}^{\infty} \sum_{k=0}^n a_n \binom{n}{k} r^k s^{n-k} < \infty$$

absolut summierbar. Dann gilt für alle $m = n - k \in \mathbb{N}_0$ und $K \in \mathbb{N}_0$:

$$s^m \sum_{k=0}^K |a_{m+k}| \binom{m+k}{k} r^k \leq \sum_{n=0}^{\infty} \sum_{k=0}^n |a_n| \binom{n}{k} r^k s^{n-k} < \infty.$$

Für alle $m = n - k \in \mathbb{N}_0$ konvergiert also $b_m = \sum_{k=0}^{\infty} a_{m+k} \binom{m+k}{k} x_0^k$ absolut. Und es gilt

$$\sum_{m=0}^M s^m \sum_{k=0}^{\infty} |a_{m+k}| \binom{m+k}{k} r^k \leq \sum_{n=0}^{\infty} \sum_{k=0}^n |a_n| \binom{n}{k} r^{n-k} s^k < \infty$$

für alle $M \in \mathbb{N}_0$. Also ist auch $\sum_{m=0}^{\infty} |b_m| s^m \leq \sum_{n=0}^{\infty} |a_n| (r+s)^n < \infty$.

Also ist $\sum_{m=0}^{\infty} b_m x^m$ für alle $|x| < R - r$ konvergent, und der Konvergenzradius nicht kleiner als $R - r = R - |x_0|$. Für $|x| < R - |x_0|$ und alle $M, K \in \mathbb{N}_0$ gilt dann

$$\left| \sum_{m=0}^M x^m \sum_{k=0}^K a_{m+k} \binom{m+k}{k} x_0^k - \sum_{n=0}^{M+K} a_n (x+x_0)^n \right| \leq \sum_{n=\min\{M+1, K+1\}}^{M+K} |a_n| (|x| + |x_0|)^n.$$

Also folgt im Grenzwert $K \rightarrow \infty$

$$\left| \sum_{m=0}^M b_m x^m - \sum_{n=0}^{\infty} a_n (x+x_0)^n \right| \leq \sum_{n=M+1}^{\infty} |a_n| (|x| + |x_0|)^n < \infty.$$

und im Grenzwert $M \rightarrow \infty$ auch $\sum_{m=0}^{\infty} b_m x^m - \sum_{n=0}^{\infty} a_n (x+x_0)^n = 0$.

(iii) Sei $(x_m)_{m \in \mathbb{N}}$ eine Folge von paarweise verschiedenen Nullstellen der Potenzreihenfunktion $x \mapsto \sum_{n=0}^{\infty} (a_n - b_n)x^n$ in $\{x \in \mathbb{K} \mid |x| \leq r\}$. Dann hat die Folge $(x_m)_{m \in \mathbb{N}}$ einen Häufungspunkt x_0 mit $|x_0| \leq r$ und für alle $\epsilon > 0$ unendlich viele Glieder in $\{x \in \mathbb{K} \mid |x - x_0| < \epsilon\}$. Sei R das Minimum der Konvergenzradien von $(\sum a_n)_{n \in \mathbb{N}_0}$

und $(\sum b_n)_{n \in \mathbb{N}_0}$. Dann hat aufgrund der Voraussetzung und wegen (ii) die Potenzreihe $\sum_{n=0}^{\infty} (a_n - b_n)(y+x_0)^n$, als Potenzreihe in y mindestens den Konvergenzradius $R-r > 0$, und für alle $0 < \epsilon \leq R-r$ unendlich viele Nullstellen auf $\{y \in \mathbb{C} \mid |y| < \epsilon\}$. Also verschwindet wegen (i) diese Potenzreihe in y identisch. Dann stimmen wegen (ii) die beiden Potenzreihen $\sum_{n=0}^{\infty} a_n x^n$ und $\sum_{n=0}^{\infty} b_n x^n$ auf dem Ball $\{x \in \mathbb{K} \mid |x-x_0| < R-r\}$ überein. Also gibt es eine Folge $(\tilde{x}_m)_{m \in \mathbb{N}}$ von paarweise verschiedenen Nullstellen von $x \mapsto \sum_{n=0}^{\infty} (a_n - b_n)x^n$, die gegen ein $\tilde{x}_0$ mit $|\tilde{x}_0| = \max\{r - (R-r), 0\} < r$ konvergiert. Dabei ist $R - |\tilde{x}_0| > R-r$. Sei $\frac{R}{R-r} < N \in \mathbb{N}$. Indem wir die beiden Potenzreihe immer wieder an einer Stelle mit minimalem Radius in dem Bereich entwickeln, in dem wir schon die Gleichheit beider Potenzreihen gezeigt haben, erhalten wir nach höchstens N Schritten, dass beide Potenzreihen auf einer Nullfolge übereinstimmen. Wegen (i) sind dann die beiden Potenzreihen $(\sum a_n x^n)_{n \in \mathbb{N}_0}$ und $(\sum b_n x^n)_{n \in \mathbb{N}_0}$ identisch. **q.e.d.**

4.4 Sinus und Cosinus

Definition 4.30. Für alle $x \in \mathbb{K}$ sei

$$\cos(x) = \frac{\exp(ix) + \exp(-ix)}{2} \quad \sin(x) = \frac{\exp(ix) - \exp(-ix)}{2i}$$

Also gilt für reelle x $\cos(x) = \Re(\exp(ix))$ $\sin(x) = \Im(\exp(ix))$
und für alle $x \in \mathbb{K}$ die **Eulersche Formel**: $\exp(ix) = \cos(x) + i \sin(x)$.
Außerdem gilt für alle $x \in \mathbb{K}$

$$\cos^2(x) + \sin^2(x) = \frac{\exp(2ix) + 2 + \exp(-2ix)}{4} - \frac{\exp(2ix) - 2 + \exp(-2ix)}{4} = 1,$$

$$\cos(-x) = \cos(x) \quad \text{und} \quad \sin(-x) = -\sin(x).$$

Satz 4.31 (Additionstheorem). Für alle $x, y \in \mathbb{K}$ gilt:

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

Beweis: $\exp(i(x+y)) = \exp(ix)\exp(iy)$.

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

Ersetzen wir x und y durch $-x$ und $-y$ und benutzen $\cos(-x) = \cos(x)$ und $\sin(-x) = -\sin(x)$, dann erhalten wir

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

Die Summe und die Differenz dieser beiden Gleichungen ergibt

$$\begin{aligned}\cos(x+y) &= \cos(x)\cos(y) - \sin(x)\sin(y) \\ \sin(x+y) &= \sin(x)\cos(y) + \cos(x)\sin(y).\end{aligned}\quad \text{q.e.d.}$$

Durch Einsetzen von (x, y) und $(x, -y)$ erhalten wir für alle $x, y \in \mathbb{K}$

$$\begin{aligned}\cos(x+y) + \cos(x-y) &= 2\cos(x)\cos(y) \\ \cos(x-y) - \cos(x+y) &= 2\sin(x)\sin(y) \\ \sin(x+y) + \sin(x-y) &= 2\sin(x)\cos(y)\end{aligned}$$

Satz 4.32 (Potenzreihen von Sinus und Cosinus).

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

Beweis: Weil $i^2 = -1$ gilt für alle $k \in \mathbb{N}_0$ $i^{2k} = (-1)^k$ und $i^{2k+1} = i(-1)^k$. Also gilt

$$\begin{aligned}(ix)^{2k} + (-ix)^{2k} &= 2(-1)^k x^{2k}, & (ix)^{2k+1} + (-ix)^{2k+1} &= 0, \\ (ix)^{2k} - (-ix)^{2k} &= 0, & (ix)^{2k+1} - (-ix)^{2k+1} &= 2i(-1)^k x^{2k+1}\end{aligned}$$

für alle $x \in \mathbb{K}$ und damit auch

$$\begin{aligned}\cos(x) &= \frac{\exp(ix) + \exp(-ix)}{2} = \sum_{n=0}^{\infty} \frac{(ix)^n + (-ix)^n}{2 \cdot n!} = \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k}}{(2k)!} \\ \sin(x) &= \frac{\exp(ix) - \exp(-ix)}{2i} = \sum_{n=0}^{\infty} \frac{(ix)^n - (-ix)^n}{2i \cdot n!} = \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k+1}}{(2k+1)!}\end{aligned}\quad \text{q.e.d.}$$

Kapitel 5

Stetigkeit

5.1 Teilmengen von $\mathbb{K}$

In diesem Abschnitt betrachten wir Teilmengen X von $\mathbb{K}$, also von $\mathbb{R}$ oder $\mathbb{C}$. Der Absolutbetrag $|x - y|$ der Differenz zweier Elemente $x, y \in X$ definiert einen Abstand. Wir hatten in den Sätzen 2.22 und 2.61 folgende Eigenschaften hergeleitet:

- (i) Für alle $x, y \in X$ ist $|x - y| \geq 0$ und $|x - y| = 0 \iff x = y$ (Positivität).
- (ii) Für alle $x, y \in X$ ist $|x - y| = |y - x|$ (Symmetrie).
- (iii) Für alle $x, y, z \in X$ ist $|x - y| \leq |x - z| + |z - y|$ (Dreiecksungleichung).

Definition 5.1 (offener Ball, Umgebung, offene Menge). Sei $X \subset \mathbb{K}$. Ein offener Ball in X mit Zentrum $x \in X$ und Radius $r > 0$ ist die Menge $B(x, r) = \{y \in X \mid |x - y| < r\}$. Eine Umgebung eines Punktes $x \in X$ ist eine Menge $U \subset X$, die für ein $\epsilon > 0$ den Ball $B(x, \epsilon)$ enthält. Eine offene Menge $O \subset X$ ist eine Teilmenge, die eine Umgebung aller ihrer Punkte ist, d.h. für alle $x \in O$ gibt es ein $\epsilon > 0$, so dass $B(x, \epsilon) \subset O$.

Beispiel 5.2. In $\mathbb{R}$ bestehen die Bälle $B(x, r)$ aus den Intervallen $(x - r, x + r)$. In $\mathbb{C}$ bestehen die Bälle $B(x, r)$ aus Kreisscheiben um x mit Radius r ohne den Rand.

In $[0, \infty)$ ist $[0, 1)$ eine Umgebung von 0, weil sie der offene Ball $B(0, 1)$ ist. In $\mathbb{R}$ ist $[0, 1)$ keine Umgebung von 0, weil sie keinen Ball $B(0, r)$ mit $r > 0$ von $\mathbb{R}$ enthält.

Alle offene Bälle $B(x, r)$ sind offenbar Umgebungen von x . Sei $y \in B(x, r)$. Dann ist $|x - y| < r$. Sei $z \in B(y, r - |x - y|)$. Dann gilt $|x - z| \leq |x - y| + |y - z| < r$, also auch $B(y, r - |x - y|) \subset B(x, r)$. Deshalb sind die offenen Bälle offene Mengen.

Insbesondere ist $[0, 1)$ in $[0, \infty)$ offen aber in $\mathbb{R}$ nicht.

Offenbar ist eine beliebige Vereinigung von offenen Mengen offen. Seien O und O' zwei offene Mengen und $x \in O \cap O'$. Dann gibt es $r > 0$ und $r' > 0$ so dass $B(x, r) \subset O$ und $B(x, r') \subset O'$. Also ist $B(x, \min\{r, r'\}) = B(x, r) \cap B(x, r') \subset O \cap O'$. Also ist $O \cap O'$ offen. Damit ist auch die Schnittmenge von endlich vielen offenen Mengen offen.

Definition 5.3 (abgeschlossene Mengen, Abschluss). *Sei $X \subset \mathbb{K}$. Die abgeschlossenen Teilmengen von X sind die Komplemente der offenen. Der Abschluss $\bar{A}$ einer Teilmenge $A \subset X$ ist die Schnittmenge aller abgeschlossenen Teilmengen, die A enthalten.*

Wegen der Regel von de Morgan, sind beliebige Schnitte und endliche Vereinigungen von abgeschlossenen Teilmengen von X abgeschlossen. Deshalb ist der Abschluss einer beliebigen Teilmenge von X abgeschlossen und der Abschluss einer abgeschlossenen Teilmenge gleich der Menge.

Ein Punkt $x \in X$ gehört genau dann zu dem Abschluss $\bar{A}$ in X , wenn es keine offene Teilmenge von X gibt, die x enthält aber mit A schnittfremd ist. Dies ist äquivalent dazu, dass für alle $n \in \mathbb{N}$ die offenen Bälle $B(x, \frac{1}{n})$ in X ein Element a_n aus A enthalten, oder auch dazu, dass es eine Folge $(a_n)_{n \in \mathbb{N}}$ in A gibt, die gegen x konvergiert. Dabei konvergiert eine Folge $(x_n)_{n \in \mathbb{N}}$ in X , wenn es einen Grenzwert in X gibt, gegen den sie konvergiert. Damit haben wir gezeigt:

Lemma 5.4. *Der Abschluss $\bar{A}$ einer Teilmenge $A \subset X$ besteht aus den Grenzwerten von allen Folgen in A , die in X konvergieren. $A \subset X$ ist genau dann abgeschlossen, wenn die Grenzwerte von allen Folgen in A , die in X konvergieren, in A liegen. **q.e.d.***

Wegen Satz 3.5 ist in $\mathbb{K}$ der Abschluss der offenen Bälle $\overline{B(x, r)} = \{y \mid |y - x| \leq r\}$.

5.2 Vollständigkeit und Kompaktheit

Wenn eine Folge $(x_n)_{n \in \mathbb{N}}$ in X konvergiert, dann auch in $\mathbb{K}$. Deshalb gilt

Satz 5.5. *In $X \subset \mathbb{K}$ ist jede konvergente Folge eine Cauchyfolge. **q.e.d.***

Definition 5.6. *$X \subset \mathbb{K}$ heißt vollständig, wenn jede Cauchyfolge in X konvergiert.*

Wegen dem Vollständigkeitsaxiom sind $\mathbb{R}$ und $\mathbb{C}$ vollständig. Wegen Lemma 5.4 ist $A \subset \mathbb{K}$ genau dann vollständige, wenn A in $\mathbb{K}$ abgeschlossen ist.

Weil jede reelle Zahl der Grenzwert einer Folge von rationalen Zahlen ist, sind die reellen Zahlen der Abschluss der rationalen Zahlen $\mathbb{Q} \subset \mathbb{R}$. Anstelle unserer axiomatischen Charakterisierung der reellen Zahlen können wir also die reellen Zahlen auch aus den rationalen Zahlen konstruieren als Äquivalenzklassen von rationalen Cauchyfolgen, wobei zwei Cauchyfolgen als äquivalent gelten, wenn ihre Differenz eine Nullfolge ist.

Definition 5.7 (Kompakte Mengen). *Eine Teilmenge $X \subset \mathbb{K}$ heißt kompakt, wenn jede Folge in X eine in X konvergente Teilfolge besitzt.*

Definition 5.8. *Eine Teilmenge $X \subset \mathbb{K}$ heißt beschränkt, wenn für ein $x \in X$, die Menge der Abstände $\{|x - y| \mid y \in X\}$ beschränkt ist.*

Wegen der Dreiecksungleichung ist diese Bedingung äquivalent dazu, dass für jedes $x \in \mathbb{K}$ die Menge der Abstände $\{|x - y| \mid y \in X\}$ beschränkt ist.

Satz 5.9 (Heine-Borel). *Eine Teilmenge $X \subset \mathbb{K}$ ist genau dann kompakt, wenn sie in $\mathbb{K}$ abgeschlossen und beschränkt ist.*

Beweis: Wegen dem Auswahlprinzip von Bolzano Weierstraß besitzt jede Folge $(x_n)_{n \in \mathbb{N}}$ in einer beschränkten Menge $X \subset \mathbb{K}$ eine in $\mathbb{K}$ konvergente Teilfolge. Der Grenzwert einer Folge in einer kompakten Teilmenge $X \subset \mathbb{K}$, die in $\mathbb{K}$ konvergiert, muss in X liegen. Wegen Lemma 5.4 ist dann eine beschränkte Teilmenge $X \subset \mathbb{K}$ genau dann kompakt, wenn sie abgeschlossen ist. Wenn es für endlich viele paarweise disjunkte Bälle $B(x_1, 1), \dots, B(x_n, 1)$ einer Teilmenge $X \subset \mathbb{K}$ kein $x_{n+1} \in X$ gibt, so dass auch $B(x_1, 1), \dots, B(x_{n+1}, 1)$ paarweise disjunkt sind, dann ist X in $B(x_1, 2) \cup \dots \cup B(x_n, 2)$ enthalten und beschränkt. Also gibt es in unbeschränkten X Folgen $(x_n)_{n \in \mathbb{N}}$ mit paarweise disjunkten Bällen $B(x_n, 1)$. Für $x \in X$ folgt $x \in B(x_n, 1)$ aus $x_n \in B(x, 1)$. Also hat eine solche Folge $(x_n)_{n \in \mathbb{N}}$ in X keinen Häufungspunkt x . **q.e.d.**

Korollar 5.10. *Teilmengen $A \subset X$ einer kompakten Teilmenge $X \subset \mathbb{K}$ sind genau dann kompakt, wenn sie in X abgeschlossen sind.*

Beweis: Teilmengen A einer beschränkten Menge $X \subset \mathbb{K}$ sind beschränkt. Wegen Lemma 5.4 ist eine Teilmenge A einer in $\mathbb{K}$ abgeschlossenen Menge $X \subset \mathbb{K}$ genau dann in X abgeschlossen, wenn sie es in $\mathbb{K}$ ist. Also folgt die Aussage aus Satz 5.9. **q.e.d.**

Korollar 5.11. *Die kompakten Teilmengen von $\mathbb{R}$ besitzen Minimum und Maximum.*

Beweis: Für jede beschränkte nicht leere Teilmenge $A \subset \mathbb{R}$ und für jedes $n \in \mathbb{N}$ ist $\sup A - \frac{1}{n}$ keine obere Schranke von A und $\inf A + \frac{1}{n}$ keine untere Schranke. Deshalb gibt es ein $a_n \in (\sup A - \frac{1}{n}, \sup A] \cap A$ und ein $b_n \in [\inf A, \inf A + \frac{1}{n}) \cap A$. Die Folgen $(a_n)_{n \in \mathbb{N}}$ und $(b_n)_{n \in \mathbb{N}}$ konvergieren gegen $\sup A$ bzw. $\inf A$. Also liegen $\sup A$ und $\inf A$ im Abschluss von A . Kompakte Teilmengen besitzen Minimum und Maximum. **q.e.d.**

Beispiel 5.12. *In $\mathbb{K}$ sind alle abgeschlossenen Bälle $\overline{B(x, r)}$ für $r \geq 0$ kompakt. Dasselbe gilt in $X \subset \mathbb{K}$ genau dann, wenn X in $\mathbb{K}$ abgeschlossen ist.*

5.3 Stetigkeit

Definition 5.13. *Seien X und Y jeweils eine Teilmenge entweder von $\mathbb{R}$ oder von $\mathbb{C}$. Eine Abbildung $f : X \rightarrow Y$, $x \mapsto f(x)$ heißt stetig in $x \in X$, wenn es für jedes $\epsilon > 0$ ein $\delta > 0$ gibt, so dass $|f(x) - f(y)| < \epsilon$ für alle $y \in X$ gilt, die $|x - y| < \delta$ erfüllen. Die Abbildung f heißt stetig, wenn sie in allen Punkten von X stetig ist.*

Stetig im Punkt x heißt also, dass alle Punkte, die hinreichend nahe bei x liegen, auf Werte abgebildet werden, die beliebig nahe bei $f(x)$ liegen.

Satz 5.14. *Für eine Abbildung $f : X \rightarrow Y$, $x \mapsto f(x)$ zwischen zwei Teilmengen X und Y jeweils von $\mathbb{R}$ oder $\mathbb{C}$ ist folgendes äquivalent: (i) f ist stetig in x .*

(ii) Das Urbild jeder Umgebung von $f(x)$ ist eine Umgebung von x .

(iii) Für gegen x konvergente Folgen $(x_n)_{n \in \mathbb{N}}$ in X konvergiert $(f(x_n))_{n \in \mathbb{N}}$ gegen $f(x)$.

Beweis: (i) $\Leftrightarrow$ (ii): Umgebungen von x bzw. $f(x)$ sind Mengen, die $B(x, \delta)$ mit $\delta > 0$ bzw. $B(f(x), \epsilon)$ mit $\epsilon > 0$ enthalten. Also ist (ii) äquivalent dazu, dass das Urbild jeder Menge, die $B(f(x), \epsilon)$ für ein $\epsilon > 0$ enthält, $B(x, \delta)$ mit $\delta > 0$ enthält und dazu, dass $f^{-1}[B(f(x), \epsilon)]$ für alle $\epsilon > 0$ ein $B(x, \delta)$ mit $\delta > 0$ enthält. Das ist äquivalent zu (i).

(ii) $\Leftrightarrow$ (iii): Folgen $(x_n)_{n \in \mathbb{N}}$ und $(f(x_n))_{n \in \mathbb{N}}$ konvergieren genau dann gegen x bzw. $f(x)$, wenn jede Umgebung von x bzw. $f(x)$ alle bis auf endlich viele Folgenglieder enthält. Wenn also (ii) gilt und $(x_n)_{n \in \mathbb{N}}$ gegen x konvergiert, dann konvergiert $(f(x_n))_{n \in \mathbb{N}}$ gegen $f(x)$ und (iii) folgt aus (ii). Wenn umgekehrt das Urbild einer Umgebung von $f(x)$ keine Umgebung von x enthält, dann gibt es für alle $n \in \mathbb{N}$ ein $x_n \in B(x, \frac{1}{n})$, so dass $(f(x_n))_{n \in \mathbb{N}}$ im Komplement der Umgebung von $f(x)$, also für ein $\epsilon > 0$ nicht in $B(f(x), \epsilon)$ liegt. Dann konvergiert $(x_n)_{n \in \mathbb{N}}$ gegen x , aber $f(x_n)$ nicht gegen $f(x)$. **q.e.d.**

Korollar 5.15. Für eine Funktion $f : X \rightarrow Y$ zwischen zwei Teilmengen X und Y jeweils von $\mathbb{R}$ oder $\mathbb{C}$ ist folgendes äquivalent: (i) f ist stetig.

(ii) Für in X konvergente Folge $(x_n)_{n \in \mathbb{N}}$ konvergiert $(f(x_n))_{n \in \mathbb{N}}$ gegen $f(\lim_{n \rightarrow \infty} x_n)$.

(iii) Das Urbild jeder offenen Teilmenge von Y ist offen in X .

(iv) Das Urbild jeder abgeschlossenen Teilmenge von Y ist abgeschlossen in X .

Beweis: Wegen dem vorangehenden Satz sind (i) und (ii) äquivalent. Weil eine Menge genau dann offen ist, wenn sie eine Umgebung von allen ihren Punkten ist, zeigt der vorangehende Satz, dass aus (i) bzw. (ii) auch (iii) folgt. Weil jede Umgebung eines Punktes auch eine offene Umgebung des Punktes enthält, folgt wieder wegen dem vorangehenden Satz aus (iii) auch (i) bzw. (ii). Weil nun die abgeschlossenen Mengen gerade die Komplemente der offenen Mengen sind und das Urbild eines Komplementes gerade gleich dem Komplement des Urbildes ist, ist (iii) zu (iv) äquivalent. **q.e.d.**

Korollar 5.16. Die Komposition zweier stetiger Abbildungen $f : X \rightarrow Y$ und $g : Y \rightarrow Z$ ist stetig. Die analoge punktweise Aussage gilt auch.

Beweis: $g(f(\lim_{n \rightarrow \infty} x_n)) = g(\lim_{n \rightarrow \infty} f(x_n)) = \lim_{n \rightarrow \infty} g(f(x_n))$ für $(x_n)_{n \in \mathbb{N}}$ konvergent. **q.e.d.**

Korollar 5.17. Das Bild einer kompakten Menge unter einer stetigen Abbildung $f : X \rightarrow Y$ zwischen zwei Teilmengen X und Y jeweils von $\mathbb{R}$ oder $\mathbb{C}$ ist kompakt.

Beweis: Sei $f : X \rightarrow Y$, $x \mapsto f(x)$ eine stetige Abbildung zwischen zwei Teilmengen X und Y jeweils von $\mathbb{R}$ oder $\mathbb{C}$ und $A \subset X$ eine kompakte Menge. Für jede Folge $(y_n)_{n \in \mathbb{N}}$ in $f[A]$ gibt es eine Folge $(x_n)_{n \in \mathbb{N}}$ in A mit $f(x_n) = y_n$ für alle $n \in \mathbb{N}$. Weil A kompakt ist, besitzt $(x_n)_{n \in \mathbb{N}}$ eine konvergente Teilfolge. Die entsprechende Teilfolge von $(y_n)_{n \in \mathbb{N}} = (f(x_n))_{n \in \mathbb{N}}$ konvergiert wegen der Stetigkeit von f . Also besitzt jede Folge in $f[A]$ eine konvergente Teilfolge. **q.e.d.**

Korollar 5.18. *Sei f eine bijektive stetige Abbildung von einer kompakten Teilmenge X auf eine Teilmenge Y jeweils von $\mathbb{R}$ oder $\mathbb{C}$. Dann ist die Umkehrabbildung stetig.*

Beweis: Wegen dem vorangehenden Korollar ist das Bild $f[X] = Y$ kompakt und wegen Korollar 5.10 das Bild $f[A]$ jeder abgeschlossenen Teilmenge $A \subset X$ abgeschlossen. Wegen $(f^{-1})^{-1}[A] = f[A]$ folgt die Aussage aus Korollar 5.15 (iv). **q.e.d.**

Beispiel 5.19. (i) *Auf jeder Teilmenge $X \subset \mathbb{K}$ ist die identische Abbildung $\mathbb{1}_X$ stetig.*
(ii) *Die konstante Abbildung, die alle $x \in X$ auf einen Punkt y abbildet ist stetig.*
(iii) *Aus den Rechenregeln für Folgen Satz 3.5 folgt, dass für jede Teilmenge X von $\mathbb{R}$ oder $\mathbb{C}$ und alle (bei $x \in X$) stetigen Abbildungen $f, g : X \rightarrow \mathbb{K}$ auch die Abbildung*

$$f + g : X \rightarrow \mathbb{K}, \quad x \mapsto f(x) + g(x) \quad f \cdot g : X \rightarrow \mathbb{K}, \quad x \mapsto f(x) \cdot g(x)$$

(bei x) stetig ist. Das gilt auch (bei $x \in X$ bzw. $x \in X \setminus f^{-1}[\{0\}]$) für die Abbildungen

$$-f : X \rightarrow \mathbb{K}, \quad x \mapsto -f(x) \quad \text{und} \quad \frac{1}{f} : X \setminus f^{-1}[\{0\}] \rightarrow \mathbb{K} \setminus \{0\}, \quad x \mapsto \frac{1}{f(x)}.$$

Definition 5.20 (Gleichmäßige Stetigkeit, Lipschitzstetigkeit). *Eine Abbildung $f : X \rightarrow Y$, $x \mapsto f(x)$ zwischen den beiden Teilmengen X und Y jeweils von $\mathbb{R}$ oder $\mathbb{C}$ heißt gleichmäßig stetig auf einer Teilmenge $A \subset X$, wenn es für alle $\epsilon > 0$ ein $\delta > 0$ gibt, so dass $|f(x) - f(y)| < \epsilon$ für alle $x, y \in A$ mit $|x - y| < \delta$ gilt.*

Die Abbildung heißt lipschitzstetig auf A , wenn es eine Konstante $L > 0$ (Lipschitzkonstante) gibt, so dass $|f(x) - f(y)| \leq L|x - y|$ für alle $x, y \in A$ gilt.

Beispiel 5.21. (i) *Eine Potenzreihenfunktion $f : x \mapsto \sum_{n \in \mathbb{N}_0} a_n x^n$ mit Konvergenzradius R ist nach Satz 4.28 (iv) für alle $0 < r < R$ auf $\overline{B(0, r)}$ lipschitzstetig. Auf der Vereinigung $B(0, R) = \bigcup_{0 < r < R} B(0, r)$ der offenen Bälle $B(0, r)$ ist f dann stetig.*
(ii) *Wegen Korollar 2.20 und Satz 2.59 (iv) ist $x \mapsto |x|$ lipschitzstetig mit $L = 1$.*

Offenbar ist jede lipschitzstetige Abbildung auch gleichmäßig stetig und jede gleichmäßig stetige Abbildung auch stetig. Es gilt auch folgende Umkehrung:

Satz 5.22. *Eine Abbildung $f : X \rightarrow Y$ von einer kompakten Teilmenge X von $\mathbb{R}$ oder $\mathbb{C}$ auf eine Teilmenge $Y \subset \mathbb{K}$ ist genau dann stetig, wenn sie gleichmäßig stetig ist.*

Beweis: Sei X kompakt und $f : X \rightarrow Y$, $x \mapsto f(x)$ eine Abbildung. Wenn f nicht gleichmäßig stetig ist, dann gibt es ein $\epsilon > 0$ und für alle $n \in \mathbb{N}$ zwei Punkte x_n und y_n in X mit $|x_n - y_n| < \frac{1}{n}$ und $|f(x_n) - f(y_n)| \geq \epsilon$. Die Folge $(x_n - y_n)_{n \in \mathbb{N}}$ ist also eine Nullfolge. Weil X kompakt ist, besitzt $(x_n)_{n \in \mathbb{N}}$ eine in X konvergente Teilfolge, also mindestens einen Häufungspunkt in X . Dann konvergiert auch die entsprechende Teilfolge von $(y_n)_{n \in \mathbb{N}}$ gegen den gleichen Grenzwert. Wegen $|f(x_n) - f(y_n)| \geq \epsilon$ konvergieren die entsprechenden Teilfolgen von $(f(x_n))_{n \in \mathbb{N}}$ und $(f(y_n))_{n \in \mathbb{N}}$ nicht gegen den gleichen Grenzwert, und f ist an allen Häufungspunkten von $(x_n)_{n \in \mathbb{N}}$ nicht stetig. Wenn f umgekehrt gleichmäßig stetig ist, dann ist f auch stetig. **q.e.d.**

Definition 5.23. Eine Folge von Funktionen $(f_n)_{n \in \mathbb{N}}$ von X nach $\mathbb{K}$ heißt

punktweise konvergent, wenn die Folgen $(f_n(x))_{n \in \mathbb{N}}$ für jedes $x \in X$ konvergieren. Die Grenzwerte definieren eine Funktion $f : X \rightarrow \mathbb{K}$, $x \mapsto \lim_{n \rightarrow \infty} f_n(x)$.

gleichmäßig konvergent, wenn es eine Funktion $f : X \rightarrow \mathbb{K}$, $x \mapsto f(x)$ gibt, und für alle $\epsilon > 0$ ein $N \in \mathbb{N}$, so dass $|f_n(x) - f(x)| < \epsilon$ für $n \geq N$ und $x \in X$ gilt.

Gleichmäßig konvergente Folgen sind punktweise konvergent, aber nicht umgekehrt.

Beispiel 5.24. Die Folge von Funktionen $(f_n)_{n \in \mathbb{N}}$ mit $f_n : [0, 1] \rightarrow \mathbb{R}$, $x \mapsto x^n$ konvergiert wegen Satz 3.4 punktweise gegen die Funktion

$$f : [0, 1] \rightarrow \mathbb{R}, \quad x \mapsto \begin{cases} 0 & \text{für } x \in [0, 1) \\ 1 & \text{für } x = 1 \end{cases}$$

Für alle $n \in \mathbb{N}$ gilt $\lim_{m \rightarrow \infty} (1 - 1/m)^n = 1$ und damit $\sup\{|f_n(x) - f(x)| \mid x \in [0, 1]\} = \sup\{x^n \mid x \in [0, 1)\} = 1$. Also konvergiert die Folge $(f_n)_{n \in \mathbb{N}}$ nicht gleichmäßig gegen f .

Definition 5.25. Sei X eine Teilmenge von $\mathbb{R}$ oder $\mathbb{C}$. Die Menge aller stetigen Funktionen von X nach $\mathbb{K}$ bezeichnen wir mit $C(X, \mathbb{K})$. Eine Funktion $f : X \rightarrow \mathbb{K}$, $x \mapsto f(x)$ heißt beschränkt, wenn das Bild $f[X]$ eine beschränkte Menge ist. $B(X, \mathbb{K})$, bezeichne die Menge aller beschränkten Funktionen auf X . $C(X, \mathbb{K})$ und $B(X, \mathbb{K})$ sind offenbar $\mathbb{K}$ -Algebren. Auf $B(X, \mathbb{K})$ bezeichne $\|\cdot\|_\infty$ folgende Abbildung:

$$\|\cdot\|_\infty : B(X, \mathbb{K}) \rightarrow \mathbb{R}, \quad f \mapsto \|f\|_\infty = \sup\{|f(x)| \mid x \in X\}.$$

Satz 5.26. Auf einer kompakten Teilmenge $X \subset \mathbb{K}$ sind alle stetigen reellen Funktionen beschränkt. Das Bild einer solchen Funktion besitzt Minimum und Maximum.

Beweis: Wegen Satz 5.17 ist das Bild jeder kompakten Teilmenge $X \subset \mathbb{K}$ unter einer stetigen Abbildung kompakt und wegen Heine-Borel beschränkt. Für reelle Funktionen besitzt es wegen Korollar 5.11 ein Maximum und ein Minimum. **q.e.d.**

Dieser Satz hat viele Anwendungen, z.B. den Fundamentalsatz der Algebra.

Satz 5.27. Sei X eine Teilmenge von $\mathbb{R}$ oder $\mathbb{C}$. Dann ist der Grenzwert f einer gleichmäßig konvergenten Folge $(f_n)_{n \in \mathbb{N}}$ in $C(X, \mathbb{K})$ auch stetig.

Beweis: Für alle $\epsilon > 0$ gibt es ein N , so dass $|f(y) - f_n(y)| \leq \frac{\epsilon}{3}$ für alle $n \geq N$ und alle $y \in X$ gilt. Dann gibt es für alle $x \in X$ ein $\delta > 0$, so dass $|f_N(x) - f_N(y)| < \frac{\epsilon}{3}$ für alle $y \in X$ mit $|x - y| < \delta$ gilt. Dann folgt für diese x und y

$$|f(x) - f(y)| \leq |f(x) - f_N(x)| + |f_N(x) - f_N(y)| + |f_N(y) - f(y)| < \epsilon. \quad \mathbf{q.e.d.}$$

Kapitel 6

Stetige Funktionen $f : \mathbb{R} \rightarrow \mathbb{R}$

6.1 Umkehrfunktionen

Satz 6.1 (Zwischenwertsatz). *Sei $f : [a, b] \rightarrow \mathbb{R}$, $x \mapsto f(x)$ stetig. Dann enthält das Bild $f[[a, b]]$ das abgeschlossene Intervall $[\min\{f(a), f(b)\}, \max\{f(a), f(b)\}]$.*

Beweis: Für $f(a) = f(b)$ ist die Aussage trivial. Sei y eine reelle Zahl im offenen Intervall zwischen $f(a)$ und $f(b)$ und $A = f^{-1}[(-\infty, y]]$ für $f(a) < f(b)$ bzw. $A = f^{-1}[[y, \infty)]$ für $f(a) > f(b)$. Diese Menge enthält a , ist beschränkt und wegen der Stetigkeit von f abgeschlossen. Also ist A kompakt und besitzt ein Maximum $x = \max A$ mit $f(x) \leq y$ für $f(a) < f(b)$ bzw. $f(x) \geq y$ für $f(a) > f(b)$. Weil y zwischen $f(a)$ und $f(b)$ liegt, ist $x < b$ und für alle $z \in (x, b]$ liegt $f(z)$ auf der selben Seite von y wie $f(b)$. Sei $(x_n)_n$ eine Folge in $(x, b]$, die gegen x konvergiert. Dann gilt $\lim_{n \rightarrow \infty} f(x_n) = f(x)$ und $f(x) \geq y$ für $f(a) < f(b)$ bzw. $f(x) \leq y$ für $f(a) > f(b)$. Also ist $f(x) = y$ und das Bild von f enthält neben $f(a)$ und $f(b)$ alle reellen Zahlen zwischen $f(a)$ und $f(b)$. **q.e.d.**

Korollar 6.2. *Sei $I \subset \mathbb{R}$ ein Intervall und $f : I \rightarrow \mathbb{R}$ stetig. Dann ist $f[I]$ ein Intervall.*

Beweis: Sei $J = f[I]$. Für $a < b$ und $a, b \in J$ folgt aus dem Zwischenwertsatz $[a, b] \subset J$. Dann enthält J alle reelle Zahlen, die weder eine untere noch eine obere Schranke von J sind, also $(\inf J, \sup J)$. Jenachdem, ob $\min J$ bzw. $\max J$ existiert, ist dann J eines der Intervalle $(\inf J, \sup J)$, $[\min J, \sup J)$, $(\inf J, \max J]$ oder $[\min J, \max J]$. Dabei ist $\inf J = -\infty$ bzw. $\sup J = \infty$ wenn J keine untere bzw. obere Schranke hat. **q.e.d.**

Die Injektivität von solchen Funktionen ist äquivalent zur strengen Monotonie:

Definition 6.3 (Monotonie). *Eine Funktion $f : X \rightarrow \mathbb{R}$, $x \mapsto f(x)$ auf einer Teilmenge X von $\mathbb{R}$ heißt*

monoton wachsend, wenn $f(x) \leq f(x')$ für alle $x, x' \in X$ mit $x \leq x'$ gilt

streng monoton wachsend, wenn $f(x) < f(x')$ für alle $x, x' \in X$ mit $x < x'$ gilt.

monoton fallend, wenn $f(x) \geq f(x')$ für alle $x, x' \in X$ mit $x \leq x'$ gilt.

streng monoton fallend, wenn $f(x) > f(x')$ für alle $x, x' \in X$ mit $x < x'$ gilt.

Satz 6.4. Eine stetige reelle Funktion f auf einem Intervall ist genau dann injektiv, wenn f entweder streng monoton wachsend oder streng monoton fallend ist.

Beweis: Eine auf einem Intervall stetige Funktion $f : I \rightarrow \mathbb{R}$ ist genau dann streng monoton, wenn $f(b) - f(a)$ für alle $a < b \in I$ nicht verschwindet und das Vorzeichen unabhängig von der Wahl von $a < b \in I$ ist, also auch wenn für alle $a < b < c \in I$ die beiden Differenzen $f(b) - f(a)$ und $f(c) - f(b)$ nicht verschwinden und untereinander das gleiche Vorzeichen haben, also $f(b)$ zwischen $f(a)$ und $f(c)$ liegt. Gilt das für $a < b < c$ nicht, dann nimmt f wegen dem Zwischenwertsatz die Werte nahe bei $f(b)$, die auf der gleichen Seite wie $f(a)$ und $f(c)$ liegen, die also sowohl zwischen $f(a)$ und $f(b)$ als auch zwischen $f(b)$ und $f(c)$ liegen, sowohl in (a, b) als auch in (b, c) an. Dann ist f nicht injektiv. Umgekehrt folgt aus der strengen Monotonie die Injektivität. **q.e.d.**

Korollar 6.5. Sei f eine reelle, stetige und injektive Funktion auf einem Intervall I . Dann ist die Umkehrfunktion $f^{-1} : f[I] \rightarrow I$ stetig.

Beweis: Jedes $x \in I$ besitzt eine Umgebung in I , die ein abgeschlossenes beschränktes Intervall ist. Das Bild solcher kompakter Intervalle ist wegen dem vorangehenden Satz und Korollar 6.2 eine Umgebung von $f(x)$ in dem Intervall $f[I]$. Dann ist f^{-1} wegen Korollar 5.18 auf dieser Umgebung von $f(x)$ stetig. Also ist f^{-1} auf $f[I]$ stetig. **q.e.d.**

Beispiel 6.6. Für $k \in \mathbb{N}$ ist $[0, \infty) \rightarrow [0, \infty), x \mapsto x^k$ streng monoton wachsend. Dann ist die Umkehrabbildung $[0, \infty) \rightarrow [0, \infty), x \mapsto x^{\frac{1}{k}}$ stetig und streng monoton wachsend. Dasselbe gilt für die Abbildung $[0, \infty) \rightarrow [0, \infty), x \mapsto x^{\frac{p}{q}}$ mit Umkehrabbildung $[0, \infty) \rightarrow [0, \infty), x \mapsto x^{\frac{q}{p}}, p, q \in \mathbb{N}$. Für negative $\frac{p}{q}$ sind diese Abbildungen von $(0, \infty)$ nach $(0, \infty)$ streng monoton fallend.

Satz 6.7. Sei f eine monoton wachsende (fallende) Funktion von einem Intervall I nach $\mathbb{R}$. Dann ist die Menge aller Unstetigkeitsstellen von f höchstens abzählbar.

Beweis*: Wir betrachten monoton wachsende Funktionen. Für monoton fallende Funktionen verläuft der Beweis analog. Für jeden inneren Punkt ξ von I sei $f(\xi_-) = \sup\{f(x) \mid x \in I \text{ und } x < \xi\}$ und $f(\xi_+) = \inf\{f(x) \mid x \in I \text{ und } \xi < x\}$. Wenn $f(\xi_-) = f(\xi_+)$, dann gibt es für jedes $\epsilon > 0$ Punkte $x_-, x_+ \in I$ mit $x_- < \xi < x_+$ und

$$f(\xi_-) - \epsilon < f(x_-) \leq f(\xi_-) = f(\xi_+) \leq f(x_+) < f(\xi_+) + \epsilon.$$

Wegen der Monotonie gilt $f(\xi_-) \leq f(\xi) \leq f(\xi_+)$. Dann gilt für alle $x \in [x_-, x_+]$ auch

$$-\epsilon < f(x_-) - f(\xi_-) \leq f(x) - f(\xi) \leq f(x_+) - f(\xi_+) < \epsilon.$$

Also ist f bei solchen ξ stetig. Wenn f bei ξ stetig ist gilt $f(\xi_-) = f(\xi) = f(\xi_+)$. Die Unstetigkeitsstellen im Inneren von I bestehen also aus den ξ mit $f(\xi_-) < f(\xi_+)$. Wegen

der Monotonie sind die offenen Intervall $(f(\xi_-), f(\xi_+))$ paarweise disjunkt. Wir wählen in jedem eine rationale Zahl und erhalten eine injektive Abbildung von den Unstetigkeitsstellen im Inneren von I nach $\mathbb{Q}$. Damit sind die Unstetigkeitsstellen gleichmächtig zu einer Teilmenge der rationalen Zahlen also höchstens abzählbar. **q.e.d.**

6.2 Exponentialfunktion und Logarithmus

Satz 6.8 (Eigenschaften von $\exp$). **(i)** $e^0 = \exp(0) = 1$ und $e^1 = \exp(1) = e$.

(ii) $e^x > \sum_{k=0}^n \frac{x^k}{k!}$ für jedes $n \in \mathbb{N}_0$ und $x > 0$.

(iii) Für $x, y \in \mathbb{R}$ folgt $e^x < e^y$ aus $x < y$.

(iv) $\exp : \mathbb{R} \rightarrow (0, \infty)$, $x \mapsto e^x$ ist bijektiv.

Beweis: **(i)** und **(ii)** folgen aus der Definition.

(iii) Aus $x < y$ folgt $y - x > 0$. Dann gilt wegen **(ii)** $e^{y-x} > 1$. Wegen Satz 4.21 **(i)** und **(ii)** gilt dann $e^y - e^x = (e^{y-x} - 1)e^x > 0$. Also folgt $e^x < e^y$.

(iv) Offenbar ist die Funktion wegen **(iii)** injektiv. Wegen $e > 1$ und Satz 3.4 gibt es für jedes $y \in (0, \infty)$ zwei $n, m \in \mathbb{N}$ mit $e^{-n} < y < e^m$. Wegen dem Zwischenwertsatz gehört dann y zum Bild von $[-n, m] \rightarrow (0, \infty)$, $x \mapsto e^x$. **q.e.d.**

Definition 6.9 (natürlicher Logarithmus). Die Umkehrfunktion von $\exp : \mathbb{R} \rightarrow (0, \infty)$, $x \mapsto e^x$ heißt natürlicher Logarithmus: $\ln : (0, \infty) \rightarrow \mathbb{R}$, $x \mapsto \ln x$.

Wegen Korollar 6.5 ist der Logarithmus stetig.

Satz 6.10 (Eigenschaften von $\ln$). **(i)** $\ln(1) = 0$.

(ii) $\ln(e) = 1$.

(iii) $\ln(xy) = \ln(x) + \ln(y)$ für alle $x, y \in (0, \infty)$.

(iv) $a^r = e^{\ln(a) \cdot r}$ für alle $r \in \mathbb{Q}$ und $a \in (0, \infty)$.

(v) $\ln(e^{\ln(a)x}) = x \ln(a)$ für alle $a \in (0, \infty)$, $x \in \mathbb{R}$.

(vi) Für $x, y \in (0, \infty)$ folgt $\ln(x) < \ln(y)$ aus $x < y$.

(vii) $\ln : (0, \infty) \rightarrow \mathbb{R}$, $x \mapsto \ln(x)$ ist bijektiv.

Beweis: **(i)** $\Leftrightarrow e^0 = 1$ und **(ii)** $\Leftrightarrow e^1 = e$ und **(iii)** $\Leftrightarrow e^{\ln(x) + \ln(y)} = (e^{\ln(x)})(e^{\ln(y)})$.

(iv) Für $r = \frac{p}{q}$ mit $p \in \mathbb{Z}$ und $q \in \mathbb{N}$ gilt $(e^{\ln(a)\frac{p}{q}})^q = e^{\ln(a)p} = a^p$ und $e^{\ln(a)\frac{p}{q}} > 0$. Wegen der Eindeutigkeit der q -ten Wurzel folgt $e^{\ln(a)\frac{p}{q}} = a^{\frac{p}{q}}$.

(v) ist offensichtlich.

(vi) folgt aus **(iii)** des vorhergehenden Satzes.

(vii) folgt aus **(iv)** des vorhergehenden Satzes. **q.e.d.**

Definition 6.11. Für alle $a > 0$ und alle $x \in \mathbb{R}$ sei $a^x = e^{x \ln(a)}$

Satz 6.12 (Eigenschaften von a^x). **(i)** $a^{x+y} = a^x a^y$ für alle $a \in (0, \infty), x, y \in \mathbb{R}$.

(ii) $(a^x)^y = a^{xy}$ für alle $a \in (0, \infty), x, y \in \mathbb{R}$.

(iii) Für $a > 1$ und $x, y \in \mathbb{R}$ folgt $a^x < a^y$ aus $x < y$.

(iv) Für $0 < a < 1$ und $x, y \in \mathbb{R}$ folgt $a^x > a^y$ aus $x < y$.

(v) Für $a \in (0, \infty) \setminus \{1\}$ ist $a^x : \mathbb{R} \rightarrow (0, \infty)$, $x \mapsto a^x$ bijektiv und stetig.

Beweis: **(i)** $a^{x+y} = e^{x \ln(a) + y \ln(a)} = e^{x \ln(a)} e^{y \ln(a)} = a^x a^y$.

(ii) $(a^x)^y = e^{y \ln(a^x)} = e^{y \cdot x \ln(a)} = a^{xy}$.

(iii) Für $a > 1$ ist $\ln(a) > 0$. Also folgt $x \ln(a) < y \ln(a)$ und $a^x < a^y$ aus $x < y$.

(iv) Für $a < 1$ ist $\ln(a) < 0$. Also folgt $x \ln(a) > y \ln(a)$ und $a^x > a^y$ aus $x < y$.

(v) Für $a \in (0, \infty) \setminus \{1\}$ ist $\ln(a) \neq 0$. Also ist $\mathbb{R} \rightarrow \mathbb{R}$, $x \mapsto \ln(a)x$ bijektiv und stetig, und damit auch $\mathbb{R} \rightarrow (0, \infty)$, $x \mapsto \exp(\ln(a)x)$. **q.e.d.**

Definition 6.13 (des Logarithmus zur Basis a). Für alle $a \in (0, \infty) \setminus \{1\}$ sei $\log_a : (0, \infty) \rightarrow \mathbb{R}$, $x \mapsto \log_a(x) = \frac{\ln(x)}{\ln(a)}$ die Umkehrfunktion von a^x .

Satz 6.14 (Eigenschaften des Logarithmus zur Basis a). **(i)** $\log_a(1) = 0$

(ii) $\log_a(a) = 1$

(iii) $\log_a(xy) = \log_a(x) + \log_a(y)$

(iv) Für $a > 1$ und $x, y \in (0, \infty)$ folgt $\log_a(x) < \log_a(y)$ aus $x < y$.

(v) Für $0 < a < 1$ und $x, y \in (0, \infty)$ folgt $\log_a(x) > \log_a(y)$ aus $x < y$.

(vi) $\log_a : (0, \infty) \rightarrow \mathbb{R}$, $x \mapsto \log_a(x)$ ist bijektiv und stetig.

Beweis analog zum Beweis der Eigenschaften von $\ln$.

q.e.d.

6.3 Trigonometrische Funktionen

Satz 6.15. **(i)** Für alle $x \in [-5, 5] \setminus \{0\}$ gilt $1 - \frac{x^2}{2} < \cos x < 1 - \frac{x^2}{2} + \frac{x^4}{24}$.

(ii) Für alle $x \in [-6, 6] \setminus \{0\}$ gilt $1 - \frac{x^2}{6} < \frac{\sin x}{x} < 1 - \frac{x^2}{6} + \frac{x^4}{120}$.

(iii) $\cos : [0, \sqrt{6}] \rightarrow \mathbb{R}$, $x \mapsto \cos(x)$ ist streng monoton fallend.

(iv) $\cos$ hat auf $[0, 2]$ genau eine Nullstelle, die wir mit $\frac{\pi}{2}$ bezeichnen.

(v) $\sin\left(\frac{\pi}{2}\right) = 1, \quad \exp\left(i\frac{\pi}{2}\right) = i.$

(vi) $\cos(n\pi) = (-1)^n \quad \sin(n\pi) = 0$ für alle $n \in \mathbb{N}$.

(vii) $\cos\left((n + \frac{1}{2})\pi\right) = 0 \quad \sin\left((n + \frac{1}{2})\pi\right) = (-1)^n$ für alle $n \in \mathbb{N}$.

(viii) $\cos(x + n\pi) = (-1)^n \cos(x)$ und $\sin(x + n\pi) = (-1)^n \sin(x)$

(ix) $\cos\left(x + \left(n + \frac{1}{2}\right)\pi\right) = -(-1)^n \sin(x)$ und $\sin\left(x + \left(n + \frac{1}{2}\right)\pi\right) = (-1)^n \cos(x).$

(x) $\sin: \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \rightarrow [-1, 1], \quad x \mapsto \sin(x)$ ist streng monoton steigend und bijektiv.

(xi) $\cos: [0, \pi] \rightarrow [-1, 1], \quad x \mapsto \cos(x)$ ist streng monoton fallend und bijektiv.

Beweis:(i) Für alle $x \in [-5, 5]$ und $k \in \mathbb{N} \setminus \{1\}$ gilt $0 \leq \frac{x^2}{(2k+1)(2k+2)} < 1$. Also ist für $x \in [-5, 5] \setminus \{0\}$ die Folge $\left(\frac{x^{2k}}{(2k)!}\right)_{k \in \mathbb{N}_0 \setminus \{0,1\}}$ streng monoton fallend und positiv. Für diese $x \in [-5, 5] \setminus \{0\}$ folgt dann $1 - \frac{x^2}{2} < \cos(x) < 1 - \frac{x^2}{2} + \frac{x^4}{24}$ wie im Beweis zu Satz 4.14.

(ii) Für $x \in [-6, 6]$ und $k \in \mathbb{N} \setminus \{1\}$ gilt $0 \leq \frac{x^2}{(2k+2)(2k+3)} < 1$. Also ist für $x \in [-6, 6] \setminus \{0\}$ die Folge $\left(\frac{x^{2k}}{(2k+1)!}\right)_{k \in \mathbb{N}_0 \setminus \{0,1\}}$ streng monoton fallend und positiv. Für diese $x \in [-6, 6] \setminus \{0\}$ folgt wieder $1 - \frac{x^2}{6} < \frac{\sin x}{x} < 1 - \frac{x^2}{6} + \frac{x^4}{120}$ wie im Beweis von Satz 4.14.

(iii) Für $x, y \in [0, \sqrt{6}]$ mit $x < y$ gilt $0 < \frac{y \pm x}{2} \leq \sqrt{6}$. Aus (ii) und dem Additionstheorem folgt $\cos(x) - \cos(y) = \cos\left(\frac{y+x}{2} - \frac{y-x}{2}\right) - \cos\left(\frac{y+x}{2} + \frac{y-x}{2}\right) = 2 \sin\left(\frac{y+x}{2}\right) \sin\left(\frac{y-x}{2}\right) > 0$.

(iv) $\sin$ und $\cos$ sind wegen Beispiel 5.21 stetig auf ganz $\mathbb{R}$. Wegen (i) ist $\cos(2) < 1 - 2 + \frac{2}{3} = -\frac{1}{3}$. Dann folgt aus dem Zwischenwertsatz, dass es eine Nullstelle in $[0, 2]$ gibt. Wegen (iii) kann es höchstens eine Nullstelle geben.

(v) Wegen $\sin^2(x) + \cos^2(x) = 1$ folgt aus (iv) $\sin^2\left(\frac{\pi}{2}\right) = 1$ und aus (ii) $\sin\left(\frac{\pi}{2}\right) > 0$. Also gilt $\sin\left(\frac{\pi}{2}\right) = 1$. Dann folgt aus der Eulerschen Formel $\exp(i\frac{\pi}{2}) = i$.

(vi) Wegen (v) folgt aus der Eulerschen Formel $\exp(ni\pi) = (i)^{2n} = (-1)^n$, also $\cos(n\pi) = (-1)^n$ und $\sin(n\pi) = 0$.

(vii) Wegen (v) folgt aus der Eulerschen Formel: $\exp\left((n + \frac{1}{2})i\pi\right) = (-1)^n i$ also $\cos\left((n + \frac{1}{2})\pi\right) = 0$ und $\sin\left((n + \frac{1}{2})\pi\right) = (-1)^n$.

(viii) Wegen (vi) folgt aus der Eulerschen Formel $\cos(x + n\pi) + i \sin(x + n\pi) = \exp(ix + ni\pi) = \exp(ix)(-1)^n = (-1)^n(\cos(x) + i \sin(x))$, also (viii).

(ix) Wegen (vii) folgt aus der Eulerschen Formel $\cos\left(x + \left(n + \frac{1}{2}\right)\pi\right) + i \sin\left(x + \left(n + \frac{1}{2}\right)\pi\right) = \exp\left(ix + \left(n + \frac{1}{2}\right)i\pi\right) = \exp(ix)(-1)^n i = (-1)^n(i \cos(x) - \sin(x))$, also (ix).

(x) Aus (ix) folgt $\sin(x) = \begin{cases} -\cos\left(x + \frac{\pi}{2}\right) & \text{für } x \in \left[-\frac{\pi}{2}, 0\right] \\ -\sin(-x) = \cos\left(\frac{\pi}{2} - x\right) & \text{für } x \in \left[0, \frac{\pi}{2}\right]. \end{cases}$

Dann folgt (x) aus (iii).

(xi) Aus (viii) folgt $\cos(x) = \begin{cases} \cos(x) & \text{für } x \in \left[0, \frac{\pi}{2}\right] \\ \cos(-x) = -\cos(\pi - x) & \text{für } x \in \left[\frac{\pi}{2}, \pi\right]. \end{cases}$

Dann folgt (xi) aus (iii).

q.e.d.

Die Umkehrfunktionen von (xi) und (x) heißen

$$\begin{array}{lll} \text{Arcuscosinus} & \arccos : [-1, 1] \rightarrow [0, \pi], & x \mapsto \arccos(x), \\ \text{Arcussinus} & \arcsin : [-1, 1] \rightarrow [-\frac{\pi}{2}, \frac{\pi}{2}], & x \mapsto \arcsin(x). \end{array}$$

Arcuscosinus ist wegen (xi) streng monoton fallend und Arcussinus wegen (x) streng monoton wachsend. Beide sind wegen Korollar 6.5 stetig. Wegen (ix) gilt $\sin(-x) = -\sin(x) = \cos(\frac{\pi}{2} + x)$ und $\arccos(x) = \frac{\pi}{2} + \arcsin(-x) = \frac{\pi}{2} - \arcsin(x)$.

Definition 6.16 (Tangens und Cotangens).

$$\tan : \mathbb{R} \setminus \left(\frac{\pi}{2} + \mathbb{Z}\pi\right) \rightarrow \mathbb{R}, \quad x \mapsto \frac{\sin(x)}{\cos(x)} \quad \cot : \mathbb{R} \setminus (\mathbb{Z}\pi) \rightarrow \mathbb{R}, \quad x \mapsto \frac{\cos(x)}{\sin(x)}.$$

$$\text{Beachte } \tan(x + \pi) = \frac{\sin(x+\pi)}{\cos(x+\pi)} = \frac{-\sin(x)}{-\cos(x)} = \tan(x) \text{ und } \cot(x + \pi) = \cot(x).$$

Satz 6.17. (i) $\tan : (-\frac{\pi}{2}, \frac{\pi}{2}) \rightarrow \mathbb{R}$ ist streng monoton steigend, stetig und bijektiv.

(ii) $\cot : (0, \pi) \rightarrow \mathbb{R}, \quad x \mapsto \cot(x)$ ist streng monoton fallend, stetig und bijektiv.

Beweis: Auf $x \in (0, \frac{\pi}{2})$ ist $\sin$ streng monoton steigend und $\cos$ streng monoton fallend und beide positiv. Also ist $\tan$ streng monoton steigend und positiv. Aus $\tan(-x) = -\tan(x)$ folgt, dass $\tan$ auf $(-\frac{\pi}{2}, \frac{\pi}{2})$ streng monoton steigend ist.

Für $x \notin \frac{\pi}{2}\mathbb{Z}$ gilt $\cot(x) = \frac{1}{\tan(x)}$ und für $x \notin \pi\mathbb{Z}$ wegen Satz 6.15 (ix) $\tan(x - \frac{\pi}{2}) = -\cot(x)$. Also ist $\cot$ auf $(0, \pi)$ streng monoton fallend.

Beide Funktionen $\tan$ und $\cot$ sind wegen Beispiel 5.19 (iii) stetig. Dann sind die Folgen $(\tan(\frac{1}{n}))_{n \in \mathbb{N}}, (-\tan(\pi - \frac{1}{n}))_{n \in \mathbb{N}}, (\cot(\frac{\pi}{2} - \frac{1}{n}))_{n \in \mathbb{N}}$ und $(-\cot(-\frac{\pi}{2} + \frac{1}{n}))_{n \in \mathbb{N}}$ identische streng monoton fallende positive Nullfolgen. Dann gilt auch

$$\lim_{n \rightarrow \infty} \tan(-\frac{\pi}{2} + \frac{1}{n}) = \lim_{n \rightarrow \infty} \cot(\pi - \frac{1}{n}) = -\infty \quad \lim_{n \rightarrow \infty} \cot(\frac{1}{n}) = \lim_{n \rightarrow \infty} \tan(\frac{\pi}{2} - \frac{1}{n}) = \infty.$$

Also sind $\tan : (-\frac{\pi}{2}, \frac{\pi}{2}) \rightarrow \mathbb{R}$ und $\cot : (0, \pi) \rightarrow \mathbb{R}$ wegen Satz 6.1 surjektiv. **q.e.d.**

Die Umkehrfunktionen von $\tan : (-\frac{\pi}{2}, \frac{\pi}{2}) \rightarrow \mathbb{R}$ und $\cot : (0, \pi) \rightarrow \mathbb{R}$ heißen

$$\begin{array}{ll} \text{Arcustangens} : \mathbb{R} \rightarrow (-\frac{\pi}{2}, \frac{\pi}{2}), & x \mapsto \arctan(x), \\ \text{Arcuscotangens} : \mathbb{R} \rightarrow (0, \pi), & x \mapsto \text{arccot}(x). \end{array}$$

Beide Funktionen sind wegen Satz 6.17 streng monoton und wegen Korollar 6.5 stetig. Wegen Satz 6.15 (ix) gilt $\cot(x + \frac{\pi}{2}) = \tan(-x)$ und $\text{arccot}(x) = \frac{\pi}{2} - \arctan(x)$.

6.4 Fundamentalsatz der Algebra

Wir übertragen unsere Erkenntnisse auf die komplexe Exponentialfunktion.

Satz 6.18 (Polardarstellung von $z \in \mathbb{C}$). *Jede komplexe Zahl hat die Darstellung:*

$$z = re^{i\varphi} \text{ mit } r = |z| \text{ und } \varphi \in \mathbb{R}.$$

Für $z \neq 0$ ist φ bis auf Addition von $2\pi n$ eindeutig und heißt Argument von z .

Beweis: Der Fall $z = 0$ ist trivial. Sei $z = x + iy \in \mathbb{C} \setminus \{0\}$. Für $y \geq 0$ sei $r = |z| = \sqrt{x^2 + y^2}$ und $\varphi = \arccos(\frac{x}{r}) \in [0, \pi]$. Dann gilt $x = r \cos(\varphi)$ und $r \sin(\varphi) \geq 0$. Aus $\frac{y^2}{x^2 + y^2} + \frac{x^2}{x^2 + y^2} = 1$ folgt $y = r \sin(\varphi)$. Wegen der Eulerschen Formel gilt dann

$$z = re^{i\varphi} = r \cos(\varphi) + ir \sin(\varphi) = x + iy.$$

Für $y < 0$ sei $r = |z| = \sqrt{x^2 + y^2}$ und $\varphi = -\arccos(\frac{x}{r}) \in [-\pi, 0]$. Wegen $x = r \cos(-\varphi) = r \cos(\varphi)$ und $r \sin(\varphi) \leq 0$ folgt wieder

$$z = re^{i\varphi} = r \cos(\varphi) + ir \sin(\varphi) = x + iy.$$

Für $(r, \varphi), (s, \vartheta) \in [0, \infty) \times \mathbb{R}$ folgt $r = |re^{i\varphi}| = |se^{i\vartheta}| = s$ aus $re^{i\varphi} = se^{i\vartheta}$ und für $r = s \neq 0$ auch $e^{i\varphi}e^{-i\vartheta} = e^{i(\varphi-\vartheta)} = 1$, was äquivalent ist zu $\varphi - \vartheta = 2\pi n$. **q.e.d.**

Die Multiplikation mit $e^{i\varphi}$ wirkt auf $z = x + iy$ wie $\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$. Diese lineare Abbildung der komplexen Zahlenebene interpretieren wir als Drehung, weil sie die Länge $\sqrt{x^2 + y^2} = |z| = |e^{i\varphi}z|$ und das Skalarprodukt erhält und Determinante 1 hat. Weil sie 1 auf $e^{i\varphi}$ abbildet, entspricht sie dem Winkel zwischen 1 und $e^{i\varphi}$. Im rechtwinkligen Dreieck ist mit diesem Winkel und auf 1 normierter Hypothense $\cos(\varphi)$ die gerichtete Länge der Ankathete und $\sin(\varphi)$ die gerichtete Länge der Gegenkathete.

Für kleine $\varphi > 0$ gilt $\varphi - \frac{\varphi^3}{6} < \sin(\varphi) < L(\varphi) < \sin(\varphi) + 1 - \cos(\varphi) < \varphi + \frac{\varphi^2}{2}$ für die Länge $L(\varphi)$ des Kreissegmentes zwischen 1 und $e^{i\varphi}$. Wir unterteilen dieses Kreissegment in n gleichgroße Kreissegmente: $e^{i\varphi} = (e^{i\frac{\varphi}{n}})^n$. Das ergibt $L(\varphi) = \lim_{n \rightarrow \infty} nL(\frac{\varphi}{n}) = \varphi$. Also entspricht φ einem Winkel, dessen Kreissegment im Einheitskreis die Länge φ hat.

Korollar 6.19. $\exp : \mathbb{C} \rightarrow \mathbb{C} \setminus \{0\}$ ist surjektiv und $\exp(z) = \exp(z') \Leftrightarrow z - z' \in 2\pi i\mathbb{Z}$.

Beweis: Sei $z = x + iy \in \mathbb{C}$. Dann gilt $e^z = e^x e^{iy}$. Also folgt das Korollar aus dem Satz 6.18 und der Bijektivität von $\exp : \mathbb{R} \rightarrow (0, \infty)$, Satz 6.8 (iv). **q.e.d.**

Korollar 6.20. Für $z \in \mathbb{C} \setminus \{0\}$ und $n \in \mathbb{N}$ gibt es genau n $w_1, \dots, w_n \in \mathbb{C}$ mit $w^n = z$.

Beweis: Seien (r, φ) die Polarkoordinaten von z . Dann müssen die Polarkoordinaten $(s, \vartheta) \in (0, \infty) \times \mathbb{R}$ der Lösungen von $w^n = z$ die Gleichungen $n\vartheta = \varphi + 2\pi m$ mit $m \in \mathbb{Z}$ und $s^n = r$ erfüllen. Also sind die Lösungen gegeben durch $s = \sqrt[n]{r}$ und $\vartheta_m = \frac{\varphi}{n} + \frac{2\pi m}{n}$, wobei zwei Lösungen (s, ϑ_m) und $(s, \vartheta_{m'})$ genau dann übereinstimmen, wenn $\frac{m-m'}{n} \in \mathbb{Z}$. Also ergeben $m = 0, \dots, n-1$ alle Lösungen. **q.e.d.**

Satz 6.21 (Fundamentalsatz der Algebra). *Jedes komplexe Polynom $p(z) = a_n z^n + \dots + a_0$ mit $a_n \neq 0$ und $n \in \mathbb{N}$ hat mindestens eine Nullstelle auf $z \in \mathbb{C}$.*

Beweis: Wir zeigen zuerst, dass $z \mapsto |p(z)|$ ein Minimum z_0 hat. Danach zeigen wir mit dem Spezialfall in dem vorangehenden Korollar, dass beim Minimum $p(z_0) = 0$ gilt.

Für $|z| \geq R = 1 + 2 \left| \frac{a_{n-1}}{a_n} \right| + \dots + 2 \left| \frac{a_0}{a_n} \right| \geq 1$ gilt

$$\begin{aligned} \left| \frac{p(z)}{z^n} \right| &= \left| a_n + \frac{a_{n-1}}{z} + \dots + \frac{a_0}{z^n} \right| = \left| a_n + \frac{a_{n-1}}{z} - \dots - \frac{a_0}{z^n} \right| = \left| a_n \left(\frac{a_{n-1}}{a_n z} + \dots + \frac{a_0}{a_n z^n} \right) \right| \\ &\geq |a_n| - |a_n| \left| \frac{a_{n-1}}{a_n z} + \dots + \frac{a_0}{a_n z^n} \right| \geq |a_n| \left(1 - \frac{\left| \frac{a_{n-1}}{a_n} \right| + \dots + \left| \frac{a_0}{a_n} \right|}{|z|} \right) > |a_n| \frac{1}{2}. \end{aligned}$$

Also ist $|p(z)| > \frac{|a_n|}{2} |z|^n \geq \frac{|a_n|}{2} |z| > \frac{|a_n|}{2} 2 \frac{|a_0|}{|a_n|} = |a_0| = |p(0)|$ für alle $z \notin B(0, R)$. Auf der kompakten Menge $\overline{B(0, R)}$ nimmt $z \mapsto |p(z)|$ wegen Satz 5.26 das Minimum bei einem z_0 an. Dieses liegt in $B(0, R)$ und ist dann das Minimum auf ganz $z \in \mathbb{C}$.

Wir zeigen jetzt $p(z_0) = 0$. Andernfalls sei $p(y + z_0) = b_n y^n + \dots + b_0$ das entsprechende Polynom in $y = z - z_0$ mit $b_n = a_n \neq 0$ und $b_0 = p(z_0) \neq 0$. Sei m das kleinste $m > 0$ mit $b_m \neq 0$. Für $|y| \leq r = \frac{1}{1 + 2 \left| \frac{b_{m+1}}{b_m} \right| + \dots + 2 \left| \frac{b_n}{b_m} \right|} \leq 1$ gilt dann

$$|b_{m+1} y^{m+1} + \dots + b_n y^n| \leq |b_m| |y|^m \left(\left| \frac{b_{m+1}}{b_m} \right| |y| + \dots + \left| \frac{b_n}{b_m} \right| |y| \right) \leq \frac{|b_m| |y|^m}{2}.$$

Also gilt $|p(z_0 + y)| \leq |b_0 + b_m y^m| + \frac{|b_m| |y|^m}{2}.$

Sei w eine Lösung der Gleichung $w^m b_m = -b_0$. Dann gilt für alle $t \in \mathbb{C}$ mit $|tw| \leq r$

$$|p(z_0 + tw)| \leq |b_0| |1 - t^m| + \frac{|b_0|}{2} |t|^m.$$

Insbesondere gilt $|p(z_0 + tw)| \leq |b_0| \left(1 - \frac{t^m}{2} \right) < |b_0|$ für alle $0 < t \leq \min \left\{ 1, \frac{r}{|w|} \right\}$.

Also sind alle $z_0 \in \mathbb{C}$ mit $p(z_0) \neq 0$ keine Minima von $z \mapsto |p(z)|$. **q.e.d.**

Korollar 6.22. *Jedes komplexe Polynom vom Grad $n \in \mathbb{N}$ zerfällt in ein Produkt von Polynomen ersten Grades.*

Beweis durch vollständige Induktion: (i) für $n = 1$ ist die Aussage trivial.

(ii) Die Aussage gelte für $n \in \mathbb{N}$. Sei p ein beliebiges Polynom $(n+1)$ -ten Grades. Wegen dem Fundamentalsatz der Algebra hat p eine Nullstelle bei $z_0 \in \mathbb{C}$. Wir schreiben p als $a_{n+1}(u + z_0)^{n+1} + \dots + a_1(u + z_0) + a_0 = b_{n+1}u^{n+1} + \dots + b_1u + p(z_0)$, und erhalten p als Produkt von $u = z - z_0$ mit dem Polynom $b_{n+1}(z - z_0)^n + \dots + b_1$, das wegen der Induktionsvoraussetzung in ein Produkt von Polynomen ersten Grades zerfällt. **q.e.d.**

Kapitel 7

Differenzierbare Funktionen

$$f : \mathbb{R} \rightarrow \mathbb{K}$$

7.1 Definition der Ableitung

Definition 7.1. Sei $f : X \rightarrow \mathbb{K}$ eine Funktion auf einer Teilmenge X von $\mathbb{R}$ (bzw. $\mathbb{C}$), die eine Umgebung von $x_0 \in X$ ist. Dann heißt f im Punkt x_0 (komplex) differenzierbar, wenn es ein $f'(x_0) \in \mathbb{K}$ gibt, so dass die folgende Funktion stetig in x_0 ist:

$$X \rightarrow \mathbb{K}, \quad x \mapsto \begin{cases} \frac{f(x)-f(x_0)}{x-x_0} & \text{für } x \neq x_0 \\ f'(x_0) & \text{für } x = x_0 \end{cases}.$$

Wenn X offen und f in allen $x \in X$ differenzierbar ist, heißt f differenzierbar und die Funktion $f' : X \rightarrow \mathbb{K}, \quad x \mapsto f'(x) = \frac{df}{dx}(x)$ heißt Ableitung von f .

Satz 7.2. Sei f im Punkt x_0 differenzierbar, dann ist f im Punkt x_0 auch stetig.

Beweis:

$$f(x) = f(x_0) + (x - x_0) \cdot \begin{cases} \frac{f(x)-f(x_0)}{x-x_0} & \text{für } x \neq x_0 \\ f'(x_0) & \text{für } x = x_0 \end{cases}.$$

Wir benutzen das Kriterium in Satz 5.14 (iii). Dann folgt die Aussage aus der Stetigkeit von $x \mapsto x - x_0$ und den Rechenregeln für Folgen in Satz 3.5 (i) und (iii). **q.e.d.**

Definition 7.3. Das Differential von f im Punkt x_0 ist die lineare Abbildung $df(x_0) : \mathbb{R} \rightarrow \mathbb{K}, \quad h \mapsto f'(x_0)h$. Die Gerade $\{(x, y) \in \mathbb{R}^2 \mid y = f(x_0) + (x - x_0)f'(x_0)\}$ heißt für $\mathbb{K} = \mathbb{R}$ Tangente an den Graphen von f im Punkt $(x_0, f(x_0))$. Dabei ist

$$\text{Graph}(f) = \{(x, y) \in \mathbb{R}^2 \mid y = f(x)\}.$$

Die Sekante durch zwei Punkte $(x_0, f(x_0)) \neq (x_1, f(x_1))$ des Graphen ist gegeben durch

$$\{(x, y) \in \mathbb{R}^2 \mid y = f(x_0) + \frac{x-x_0}{x_1-x_0}(f(x_1) - f(x_0))\}.$$

Im Grenzwert $x_1 \rightarrow x_0$ nähert sich die Sekante durch die Punkte $(x_0, f(x_0))$ und $(x_1, f(x_1))$ an die Tangente an den Graphen von f im Punkt $(x_0, f(x_0))$ an.

Beispiel 7.4. (i) $f(x) = |x|$. Für $x_0 = 0$ ist $\frac{f(x) - f(0)}{x - 0} = \begin{cases} 1 & \text{für } x > 0 \\ -1 & \text{für } x < 0. \end{cases}$ Also ist f im Punkt $x_0 = 0$ stetig aber nicht differenzierbar.

(ii) $f(x) = c \Rightarrow \frac{f(x) - f(x_0)}{x - x_0} = 0$ für alle $x \neq x_0$. Also ist f differenzierbar mit $f'(x) = 0$.

(iii) $f(x) = x \Rightarrow \frac{f(x) - f(x_0)}{x - x_0} = 1$ für alle $x \neq x_0$. Also ist f differenzierbar mit $f'(x) = 1$.

(iv) $f(x) = x^n$ für $n \in \mathbb{N}$. $\Rightarrow \frac{f(x) - f(x_0)}{x - x_0} = x^{n-1} + x_0 x^{n-2} + \dots + x_0^{n-1}$ für alle $x \neq x_0$. Also ist f differenzierbar mit $f'(x) = nx^{n-1}$.

(v) Sei $f(x) = \sum_{n=0}^{\infty} a_n x^n$ eine Potenzreihenfunktion mit Konvergenzradius $R > 0$. Dann ist $\frac{f(x) - f(0)}{x - 0} = \sum_{n=1}^{\infty} a_n x^{n-1} = \sum_{n=0}^{\infty} a_{n+1} x^n$ auch ein Potenzreihe mit Konvergenzradius R und wegen Satz 4.28 (iv) stetig. Deshalb ist f in $x = 0$ differenzierbar mit $f'(0) = a_1$. Aus Satz 4.29 (ii) folgt, dass f für alle $x \in B(0, R)$ in x differenzierbar und die Ableitung $f'(x)$ gegeben ist durch die Potenzreihe $f'(x) = \sum_{n=0}^{\infty} \binom{n+1}{1} a_{n+1} x^n = \sum_{n=1}^{\infty} n a_n x^{n-1}$ mit Konvergenzradius R .

(vi)

$$\exp'(x) = \sum_{n=1}^{\infty} \frac{n x^{n-1}}{n!} = \sum_{k=0}^{\infty} \frac{x^k}{k!} = \exp(x).$$

(vii)

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

(viii)

$$\cos'(x) = \sum_{k=1}^{\infty} \frac{(-1)^k 2k x^{2k-1}}{(2k)!} = - \sum_{l=0}^{\infty} \frac{(-1)^l x^{2l+1}}{(2l+1)!} = -\sin(x).$$

7.2 Rechenregeln der Ableitung

Satz 7.5. Seien $f, g : X \rightarrow \mathbb{K}$ in x_0 differenzierbar und $\lambda \in \mathbb{K}$. Dann sind auch die Funktionen λf , $f + g$ und $f \cdot g$ in x_0 differenzierbar und es gilt

$$\begin{aligned} (\lambda f)'(x_0) &= \lambda f'(x_0) & (f + g)'(x_0) &= f'(x_0) + g'(x_0) \\ (f \cdot g)'(x_0) &= f'(x_0)g(x_0) + f(x_0) \cdot g'(x_0) & & \text{(Leibnizregel).} \end{aligned}$$

Wenn $f(x_0) \neq 0$ dann ist $X' = f^{-1}[\mathbb{K} \setminus \{0\}]$ wegen Satz 7.2 eine Umgebung von x_0 und $\frac{1}{f} : X' \rightarrow \mathbb{K}$, $x \mapsto \frac{1}{f(x)}$ in x_0 differenzierbar mit $\left(\frac{1}{f}\right)'(x_0) = -\frac{f'(x_0)}{f^2(x_0)}$.

Beweis: Für $x \neq x_0$ gilt

$$\begin{aligned} \frac{\lambda f(x) - \lambda f(x_0)}{x - x_0} &= \lambda \frac{f(x) - f(x_0)}{x - x_0} && \text{und} \\ \frac{f(x) + g(x) - (f(x_0) + g(x_0))}{x - x_0} &= \frac{f(x) - f(x_0)}{x - x_0} + \frac{g(x) - g(x_0)}{x_0 - x_0} && \text{und} \\ \frac{f(x)g(x) - f(x_0)g(x_0)}{x - x_0} &= \frac{f(x) - f(x_0)}{x - x_0}g(x) + \frac{g(x) - g(x_0)}{x - x_0}f(x_0) && \text{und} \\ \left(\frac{1}{f(x)} - \frac{1}{f(x_0)}\right) \frac{1}{x - x_0} &= -\frac{f(x) - f(x_0)}{f(x)f(x_0)(x - x_0)}. \end{aligned}$$

Also folgt die Aussage aus Beispiel 5.19 und Satz 7.2.

q.e.d.

Satz 7.6 (Kettenregel). Seien $f : X \rightarrow Y$ und $g : Y \rightarrow \mathbb{K}$ Funktionen und $X \subset \mathbb{R}$ eine Umgebung von x_0 und $Y \subset \mathbb{R}$ eine Umgebung von $f(x_0)$. Wenn f in x_0 differenzierbar ist und g in $f(x_0)$, dann ist $g \circ f$ in x_0 differenzierbar mit

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

Beweis: $\frac{g(f(x)) - g(f(x_0))}{x - x_0} = \frac{g(f(x)) - g(f(x_0))}{f(x) - f(x_0)} \cdot \frac{f(x) - f(x_0)}{x - x_0}$, wobei wir den linken Faktor für $f(x) = f(x_0) = y_0$ durch $g'(y_0)$ ersetzen. Dieser linke Faktor ist die Komposition von $x \mapsto f(x)$ mit $y \mapsto \frac{g(y) - g(y_0)}{y - y_0}$ für $y \neq y_0$ und $y_0 \rightarrow g'(y_0)$, also wegen Satz 5.16 und Satz 7.2 in x_0 stetig. Also folgt die Behauptung aus Beispiel 5.19. **q.e.d.**

Satz 7.7 (Ableitung der Umkehrfunktion). Sei $f : X \rightarrow Y$ eine bijektive Funktion von einer Umgebung $X \subset \mathbb{R}$ von x_0 auf eine Umgebung $Y \subset \mathbb{R}$ von $y_0 = f(x_0)$. Wenn f in x_0 differenzierbar ist mit $f'(x_0) \neq 0$ und entweder f auf einer Umgebung von x_0 oder f^{-1} in y_0 stetig ist, dann ist auch f^{-1} in y_0 differenzierbar mit $(f^{-1})'(y_0) = \frac{1}{f'(x_0)}$.

Beweis: $\frac{f^{-1}(y) - f^{-1}(y_0)}{y - y_0} = \frac{x - x_0}{f(x) - f(x_0)}$ für $y = f(x)$ und $y_0 = f(x_0)$. Die erste der beiden folgenden Funktionen ist die Komposition von $y \rightarrow f^{-1}(y)$ mit der zweiten

$$y \mapsto \begin{cases} \frac{f^{-1}(y) - f^{-1}(y_0)}{y - y_0} & \text{für } y \neq y_0 \\ \frac{1}{f'(x_0)} & \text{für } y = y_0 \end{cases} \quad x \mapsto \begin{cases} \frac{x - x_0}{f(x) - f(x_0)} & \text{für } x \neq x_0 \Leftrightarrow f(x) \neq f(x_0) \\ \frac{1}{f'(x_0)} & \text{für } x = x_0 \end{cases}.$$

Der Satz folgt aus Korollar 5.16, Beispiel 5.19 (iii) und Korollar 6.5.

q.e.d.

Beispiel 7.8. (i) $\ln(0, \infty) \rightarrow \mathbb{R}, \quad x \mapsto \ln(x)$

$$\ln'(x) = \frac{1}{\exp'(y)} = \frac{1}{\exp(y)} = \frac{1}{x} \text{ mit } \exp(y) = x.$$

(ii) $\arcsin : [-1, 1] \rightarrow [-\frac{\pi}{2}, \frac{\pi}{2}], \quad x \mapsto \arcsin(x)$

$$\arcsin'(x) = \frac{1}{\cos(y)} = \frac{1}{\sqrt{1-x^2}} \text{ mit } \sin(y) = x \text{ und } x^2 \neq 1.$$

(iii) $\arccos : [-1, 1] \rightarrow [0, \pi], \quad x \mapsto \arccos(x)$

$$\arccos'(x) = \frac{-1}{\sin(y)} = \frac{-1}{\sqrt{1-x^2}} \text{ mit } \cos(y) = x \text{ und } x^2 \neq 1.$$

(iv) $\cdot^\alpha : (0, \infty) \rightarrow \mathbb{R}, \quad x \mapsto x^\alpha, \alpha \in \mathbb{R}.$

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

(v) $a : \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto a^x, a \in (0, \infty).$

$$(a)'(x) = \exp(x \cdot \ln(a))' = \exp(x \ln(a)) \cdot \ln(a) = \ln(a) \cdot a^x.$$

(vi) Quotientenregel. Seien f und g in x_0 differenzierbar und $g(x_0) \neq 0$. Dann ist $\frac{f}{g}$ in x_0 differenzierbar und es gilt

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

(vii) $x \mapsto \tan(x) = \frac{\sin(x)}{\cos(x)}$

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

(viii) $x \mapsto \cot(x) = \frac{\cos(x)}{\sin(x)}$

$$\cot'(x) = \frac{-\sin(x)\sin(x) - \cos(x)\cos(x)}{\sin^2(x)} = -1 - \cot^2(x) = \frac{-1}{\sin^2(x)}.$$

(ix) $\arctan : \mathbb{R} \rightarrow \left(\frac{\pi}{2}, \frac{\pi}{2}\right), \quad x \mapsto \arctan(x)$

$$\arctan'(x) = \frac{1}{1 + \tan^2(y)} = \frac{1}{1 + x^2} \text{ mit } \tan(y) = x.$$

(x) $\operatorname{arccot} : \mathbb{R} \rightarrow (0, \pi), \quad x \mapsto \operatorname{arccot}(x)$

$$\operatorname{arccot}'(x) = \frac{-1}{1 + \cot^2(y)} = \frac{-1}{1 + x^2} \text{ mit } \cot(y) = x.$$

(xi) $\log_a(0, \infty) \rightarrow \mathbb{R}, \quad x \mapsto \log_a(x) \quad \log'_a(x) = \left(\frac{\ln(x)}{\ln(a)}\right)' = \frac{1}{x \ln(a)}.$

(xii) $x^x : (0, \infty) \rightarrow (0, \infty), \quad x \mapsto x^x$

$$(x^x)' = \exp(x \cdot \ln(x))' = \exp(x \cdot \ln(x)) \left(\ln(x) + x \frac{1}{x} \right) = (\ln(x) + 1) \cdot x^x.$$

(xiii) $f(x) = \begin{cases} x^2 \sin\left(\frac{1}{x}\right) & \text{für } x \neq 0 \\ 0 & \text{für } x = 0. \end{cases}$

$$f'(x) = \begin{cases} \lim_{x \rightarrow 0 \pm} \frac{x^2 \sin\left(\frac{1}{x}\right)}{x} = \lim_{x \rightarrow 0 \pm} x \sin\left(\frac{1}{x}\right) = 0 & \text{für } x = 0 \\ 2x \sin\left(\frac{1}{x}\right) - \cos\left(\frac{1}{x}\right) & \text{für } x \neq 0. \end{cases}$$

Diese Funktion ist zwar differenzierbar, aber f' ist im Punkt $x = 0$ nicht stetig.

7.3 Mittelwertsatz und Monotonie

Definition 7.9 (lokale Maxima und Minima). Eine Funktion $f : (a, b) \rightarrow \mathbb{R}, \quad x \mapsto f(x)$ hat bei $x_0 \in (a, b)$ ein lokales Maximum bzw. Minimum, falls es ein $\delta > 0$ gibt, so dass $f(x) \leq f(x_0)$ bzw. $f(x) \geq f(x_0)$ für alle $x \in (x_0 - \delta, x_0 + \delta) \setminus \{x_0\}$ gilt. Bei strikten Ungleichungen sprechen wir von strikten lokalen Maxima und Minima.

Wenn für eine bei x_0 differenzierbaren Funktion $f'(x_0)$ nicht verschwindet, dann gibt es ein $\delta > 0$, so dass $\left| \frac{f(x) - f(x_0)}{x - x_0} - f'(x_0) \right| < |f'(x_0)|$ für alle $x \in (x_0 - \delta, x_0 + \delta)$ gilt. Dort hat $\frac{f(x) - f(x_0)}{x - x_0}$ das gleiche Vorzeichen wie $f'(x_0)$. Dann gilt entweder $f(x) < f(x_0)$ für $x \in (x_0 - \delta, x_0)$ und $f(x) > f(x_0)$ für $x \in (x_0, x_0 + \delta)$ oder $f(x) > f(x_0)$ für $x \in (x_0 - \delta, x_0)$ und $f(x) < f(x_0)$ für $x \in (x_0, x_0 + \delta)$. Das zeigt

Lemma 7.10. Ein differenzierbarer lokaler Extremwert ist eine Nullstelle von f' . **q.e.d.**

Definition 7.11 (kritischer Punkt und kritischer Wert). *Eine Nullstelle der Ableitung einer differenzierbaren Funktion heißt kritischer Punkt. Der entsprechende Funktionswert heißt kritischer Wert.*

Kandidaten für die lokalen Extremwerte einer stetigen Funktion $f : [a, b] \rightarrow \mathbb{R}$ sind

- (i) Kritische Punkte in (a, b) ,
- (ii) Randpunkte, also entweder a oder b ,
- (iii) und Punkte in (a, b) an denen f nicht differenzierbar ist.

Satz 7.12 (Satz von Rolle). *Sei $f : [a, b] \rightarrow \mathbb{R}$ stetig und auf (a, b) differenzierbar. Falls $f(a) = f(b)$, dann existiert ein $x_0 \in (a, b)$ mit $f'(x_0) = 0$.*

Beweis: Wegen Satz 5.26 gibt es $x_1, x_2 \in [a, b]$ mit $f(x_1) \leq f(x) \leq f(x_2)$ für alle $x \in [a, b]$. Wenn x_1 und x_2 beide am Rand liegen, $x_1, x_2 \in \{a, b\}$, dann muss f konstant gleich $f(a) = f(b)$ sein. Also gilt dann $f'(x) = 0$ für alle $x \in (a, b)$. Andernfalls muss es einen lokalen Extremwert in (a, b) geben, an dem wegen Lemma 7.10 $f' = 0$ ist. **q.e.d.**

Satz 7.13 (Verallgemeinerter Mittelwertsatz). *Seien $f, g : [a, b] \rightarrow \mathbb{R}$ stetig und auf (a, b) differenzierbar. Dann existiert ein $x_0 \in (a, b)$ mit*

$$(f(b) - f(a))g'(x_0) = (g(b) - g(a))f'(x_0), \quad \frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(x_0)}{g'(x_0)} \text{ für } g(b) \neq g(a), g'(x_0) \neq 0.$$

Die Tangente $(f(x_0), g(x_0)) + \mathbb{R}(f'(x_0), g'(x_0))$ an $\{(f(x), g(x)) \in \mathbb{R}^2 \mid x \in [a, b]\}$ in $(f(x_0), g(x_0))$ verläuft also parallel zu der Geraden durch $(f(a), g(a))$ und $(f(b), g(b))$.

Beweis: $x \mapsto (f(b) - f(a))g(x) - (g(b) - g(a))f(x)$ erfüllt die Voraussetzungen von Satz 7.12. Die Ableitung ist Null für $f(b) - f(a))g'(x_0) = (g(b) - g(a))f'(x_0)$. **q.e.d.**

Satz 7.14 (Mittelwertsatz). *Sei $f : [a, b] \rightarrow \mathbb{R}$, $x \mapsto f(x)$ stetig und auf (a, b) differenzierbar. Dann existiert ein $x_0 \in (a, b)$ mit $f'(x_0) = \frac{f(b) - f(a)}{b - a}$.*

Beweis: Wende den verallgemeinerten Mittelwertsatz auf f und $\mathbb{1}_{[a, b]}$ an. **q.e.d.**

Satz 7.15 (Schränkensatz). *Eine auf (a, b) differenzierbare stetige Funktion $f : [a, b] \rightarrow \mathbb{R}$ ist genau dann Lipschitzstetig, wenn für alle $x \in (a, b)$ die Lipschitzkonst. $L \geq |f'(x)|$.*

Beweis: Für $x < y \in [a, b]$ gibt es wegen dem Mittelwertsatz ein $x_0 \in (x, y)$ mit $f(y) - f(x) = f'(x_0)(y - x)$. Dann folgt $|f(y) - f(x)| \leq L|y - x|$ aus $|f'(x_0)| \leq L$. Umgekehrt folgt $|f'(x)| = \lim_{n \rightarrow \infty} \left| \frac{f(x_n) - f(x)}{x_n - x} \right| \leq L$ für $x_n \rightarrow x$ aus $|f(y) - f(x)| \leq L|y - x|$. **q.e.d.**

Satz 7.16 (Ableitung und Monotonie). *Sei $f : [a, b] \rightarrow \mathbb{R}$ stetig und auf (a, b) differenzierbar. Dann gilt*

- (i) $f'(x) = 0$ für alle $x \in (a, b) \iff f$ ist konstant.

- (ii) $f'(x) \geq 0$ für alle $x \in (a, b) \iff f$ ist monoton steigend.
- (iii) $f'(x) \leq 0$ für alle $x \in (a, b) \iff f$ ist monoton fallend.
- (iv) $f'(x) \geq 0$ für alle $x \in (a, b)$ und der Abschluss der Menge $\{x \in (a, b) \mid f'(x) > 0\}$ ist $[a, b] \iff f$ ist streng monoton steigend.
- (v) $f'(x) \leq 0$ für alle $x \in (a, b)$ und der Abschluss der Menge $\{x \in (a, b) \mid f'(x) < 0\}$ ist $[a, b] \iff f$ ist streng monoton fallend.

Beweis: Weil eine Funktion genau dann konstant ist, wenn sie monoton steigend und monoton fallend ist, folgt (i) aus (ii) und (iii). Wir beweisen nur (ii) und (iv). Für $a \leq x < y \leq b$ erfüllt $f : [x, y] \rightarrow \mathbb{R}$ die Voraussetzungen des Mittelwertsatzes. Aus $f'(x_0) \geq 0$ für alle $x_0 \in (a, b)$ folgt $f(y) - f(x) \geq 0$, und f ist monoton wachsend. Umgekehrt folgt $f(x_0) > f(x)$ für ein $x > x_0$ aus $f'(x_0) < 0$ und f ist nicht monoton steigend. Es folgt (ii). Monoton wachsende f sind genau dann streng monoton, wenn es kein $a \leq x < y \leq b$ gibt mit $f(x) = f(y)$. Auf $[x, y]$ ist dann f konstant und $f'(z) = 0$ für $z \in (x, y)$. Weil jede offene Menge ein solches Intervall enthält folgt (iv). **q.e.d.**

Korollar 7.17 (Isolierte kritische Punkte). Sei $f : (a, b) \rightarrow \mathbb{R}$ differenzierbar und $x_0 \in (a, b)$ ein kritischer Punkt.

- (i) f' bei x_0 differenzierbar mit $f''(x_0) > 0 \implies x_0$ ist ein striktes lokales Minimum.
- (ii) f' bei x_0 differenzierbar mit $f''(x_0) < 0 \implies x_0$ ist ein striktes lokales Maximum.
- (iii) Wenn es ein $\delta > 0$ gibt, so dass $f'(x) \leq 0$ für $x \in (x_0 - \delta, x_0)$ und $f'(x) \geq 0$ für $x \in (x_0, x_0 + \delta)$ gilt, dann ist x_0 ein lokales Minimum.
- (iv) Wenn es ein $\delta > 0$ gibt, so dass $f'(x) \geq 0$ für $x \in (x_0 - \delta, x_0)$ und $f'(x) \leq 0$ für $x \in (x_0, x_0 + \delta)$ gilt, dann ist x_0 ein lokales Maximum. **q.e.d.**

Beispiel 7.18. Die Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$, $x \mapsto (x+1)e^{-x}$ hat die Ableitung $f'(x) = (1 - (x+1))e^{-x} = -xe^{-x}$. Also ist sie auf $(-\infty, 0]$ streng monoton wachsend und auf $[0, \infty)$ streng monoton fallend. Insbesondere ist $f(0) = 1$ ein globales Maximum. Also gilt $x+1 \leq e^x$ für alle $x \in \mathbb{R}$ und $y \leq e^{y-1}$ für alle $y = x+1 \in \mathbb{R}$.

7.4 Regel von de L'Hopital

Definition 7.19 (Grenzwert von Funktionswerten). Für eine Funktion $f : (a, b) \rightarrow \mathbb{K}$ existiert der Grenzwert $\lim_{x \rightarrow a+} f(x)$ genau dann, wenn für ein $f(a) \in \mathbb{K}$ die Funktion

$$x \mapsto \begin{cases} f(x) & \text{für } x \in (a, b) \\ f(a) & \text{für } x = a \end{cases} \quad \text{stetig bei } x = a \text{ ist. Wir schreiben dann } \lim_{x \rightarrow a+} f(x) = f(a).$$

Der analoge Grenzwert $x \rightarrow b$ wird mit $\lim_{x \rightarrow b-} f(x)$ bezeichnet. Aufgrund der Definition der Stetigkeit existiert der Grenzwert $\lim_{x \rightarrow a+} f(x)$ also genau dann, wenn es ein $f(a) \in \mathbb{K}$ und für jedes $\epsilon > 0$ ein $\delta > 0$ gibt, mit dem $f(x) \in B(f(a), \epsilon)$ aus $x \in B(a, \delta) \cap (a, b)$ folgt. Wegen Satz 5.14 ist das äquivalent dazu, dass für jede Folge $(x_n)_{n \in \mathbb{N}}$ in (a, b) , die gegen a konvergiert, die Folge $(f(x_n))_{n \in \mathbb{N}}$ gegen $f(a)$ konvergiert.

Satz 7.20 (1. Regel von de L'Hopital). *Seien $\infty < a < b < \infty$ und f und g auf (a, b) differenzierbare Funktionen mit $\lim_{x \rightarrow a+} f(x) = 0 = \lim_{x \rightarrow a+} g(x)$. Wenn der Grenzwert $\lim_{x \rightarrow a+} \frac{f'(x)}{g'(x)}$ existiert, dann existiert auch $\lim_{x \rightarrow a+} \frac{f(x)}{g(x)}$ und es gilt $\lim_{x \rightarrow a+} \frac{f(x)}{g(x)} = \lim_{x \rightarrow a+} \frac{f'(x)}{g'(x)}$.*

Bemerkung 7.21. *Wenn die Grenzwerte $\lim_{x \rightarrow a+} f'(x)$ und $\lim_{x \rightarrow a+} g'(x)$ existieren und der zweite nicht verschwindet, dann existiert auch $\lim_{x \rightarrow a+} \frac{f'(x)}{g'(x)}$ mit $\lim_{x \rightarrow a+} \frac{f(x)}{g(x)} = \frac{\lim_{x \rightarrow a+} f'(x)}{\lim_{x \rightarrow a+} g'(x)}$.*

Beweis: Wenn der Grenzwert $\lim_{x \rightarrow a+} \frac{f'(x)}{g'(x)}$ existiert, ist $g'(x) \neq 0$ für $x \in (a, b')$ mit $a < b' \leq b$. Aus dem Mittelwertsatz folgt $g(x) = g(x) - g(a) \neq 0$ für $x \in (a, b')$. Die auf $[a, b')$ stetig fortgesetzten Funktionen f und g erfüllen die Voraussetzungen des Verallgemeinerten Mittelwertsatzes. Deshalb gibt es für jedes $x \in (a, b')$ ein $x_0 \in (a, x)$ so dass $\frac{f(x)}{g(x)} = \frac{f(x) - f(a)}{g(x) - g(a)} = \frac{f'(x_0)}{g'(x_0)}$ gilt. Wenn also der Grenzwert $\lim_{x \rightarrow a+} \frac{f'(x)}{g'(x)}$ existiert, dann existiert auch der Grenzwert $\lim_{x \rightarrow a+} \frac{f(x)}{g(x)} = \lim_{x \rightarrow a+} \frac{f'(x)}{g'(x)}$. **q.e.d.**

Satz 7.22 (2. Regel von de L'Hopital). *Unter derselben Voraussetzung wie bei der 1. Regel von de L'Hopital, nur gelte $\lim_{x \rightarrow a+} g(x) = \infty$ statt $\lim_{x \rightarrow a+} f(x) = \lim_{x \rightarrow a+} g(x) = 0$, dann gilt dieselbe Schlussfolgerung.*

Beweis: Wenn $\lim_{x \rightarrow a+} \frac{f'(x)}{g'(x)}$ existiert, gibt es $b' \in (a, b)$ mit $g'(x) \neq 0$ für $x \in (a, b')$. Für $a < x < y < b'$ folgt $g(y) \neq g(x)$ aus dem Mittelwertsatz, und wegen dem verallgemeinerten Mittelwertsatz gibt es $x_0 \in (x, y)$ mit $\frac{f(x) - f(y)}{g(x) - g(y)} = \frac{f'(x_0)}{g'(x_0)}$. Wenn f und g die Voraussetzungen der 2. Regel von de L'Hopital erfüllen, dann gibt es für jedes $\epsilon > 0$ ein $y \in (a, b')$, so dass es für alle $x_0 \in (a, y)$ gilt $|\frac{f'(x_0)}{g'(x_0)} - \alpha| < \frac{\epsilon}{2}$ mit $\alpha = \lim_{x \rightarrow a+} \frac{f'(x)}{g'(x)}$. Dann folgt $\alpha - \frac{\epsilon}{2} < \frac{f(x) - f(y)}{g(x) - g(y)} < \alpha + \frac{\epsilon}{2}$ für alle $x \in (a, y)$. Wegen $\lim_{x \rightarrow a+} g(x) = \infty$ gibt es ein $y_0 \in (a, y)$ so dass für alle $x \in (a, y_0)$ gilt $g(x) > \max\{g(y), 0\}$. Daraus folgt

$$\begin{aligned} \left(\alpha - \frac{\epsilon}{2}\right)(g(x) - g(y)) + f(y) &< f(x) < \left(\alpha + \frac{\epsilon}{2}\right)(g(x) - g(y)) + f(y) \quad \text{oder} \\ \left(\alpha - \frac{\epsilon}{2}\right) + \frac{f(y) - g(y)(\alpha - \frac{\epsilon}{2})}{g(x)} &< \frac{f(x)}{g(x)} < \left(\alpha + \frac{\epsilon}{2}\right) + \frac{f(y) - g(y)(\alpha + \frac{\epsilon}{2})}{g(x)}. \end{aligned}$$

Wegen $\lim_{x \rightarrow a+} g(x) = \infty$ gibt es dann auch ein $y_1 \in (a, y_0)$, so dass für alle $x \in (a, y_1)$ gilt $\alpha - \epsilon < \frac{f(x)}{g(x)} < \alpha + \epsilon$. Also gilt $\lim_{x \rightarrow a+} \frac{f(x)}{g(x)} = \lim_{x \rightarrow a+} \frac{f'(x)}{g'(x)}$. **q.e.d.**

Die analogen Aussagen für die Grenzwerte $\lim_{x \rightarrow b-}$ gelten natürlich auch. Grenzwerte der Form $\lim_{x \rightarrow -\infty+} f(x)$ bzw. $\lim_{x \rightarrow \infty-} f(x)$ definieren wir als die Grenzwerte $\lim_{x \rightarrow 0-} f(1/x)$ bzw. $\lim_{x \rightarrow 0+} f(1/x)$. Wegen der Kettenregel gilt dann

$$\frac{df(\frac{1}{x})}{dx} \left(\frac{dg(\frac{1}{x})}{dx} \right)^{-1} = \frac{-1}{x^2} f'(\frac{1}{x}) \left(\frac{-1}{x^2} g'(\frac{1}{x}) \right)^{-1} = f'(\frac{1}{x}) / g'(\frac{1}{x}).$$

Deshalb gelten die analogen Aussagen auch für diese Grenzwerte.

7.5 Konvexität und Ableitungen

Definition 7.23. Eine reelle Funktion auf einem Intervall heißt konvex bzw. streng konvex, wenn für alle $a \neq b$ im Definitionsbereich und alle $t \in (0, 1)$ folgendes gilt

$$f((1-t)a + tb) \leq (1-t)f(a) + tf(b) \quad \text{bzw.} \quad f((1-t)a + tb) < (1-t)f(a) + tf(b).$$

Satz 7.24. Für eine reelle Funktion f auf einem Intervall I ist folgendes äquivalent:

(i) f ist konvex

(ii) Für $[a, b] \subset I$ und $x \in (a, b)$ gilt $\frac{f(x) - f(a)}{x - a} \leq \frac{f(b) - f(a)}{b - a}$.

(iii) Für $[a, b] \subset I$ und $x \in (a, b)$ gilt $\frac{f(b) - f(a)}{b - a} \leq \frac{f(b) - f(x)}{b - x}$.

(iv) Für $[a, b] \subset I$ und $x \in (a, b)$ gilt $\frac{f(x) - f(a)}{x - a} \leq \frac{f(b) - f(x)}{b - x}$.

Die analogen Äquivalenzen zu streng konvex gelten, wenn $\leq$ durch $<$ ersetzt wird.

Beweis: Wir können wegen der Symmetrie $(a, b, t) \leftrightarrow (b, a, 1-t)$ in (i) $a < b$ annehmen. Dann sei $x = (1-t)a + tb \in (a, b) \Leftrightarrow t = \frac{x-a}{b-a} \in (0, 1)$. Also ist (i) äquivalent zu

$$f(x) \leq \frac{b-x}{b-a} f(a) + \frac{x-a}{b-a} f(b) \quad \Leftrightarrow \quad (b-a)f(x) \leq (b-x)f(a) + (x-a)f(b).$$

Ersetzen wir entweder $(b-x) = (b-a) - (x-a)$, oder $(x-a) = (b-a) - (b-x)$ oder $(b-a) = (b-x) + (x-a)$, dann ist diese Ungleichung äquivalent zu

$$\begin{aligned} (b-a)(f(x) - f(a)) &\leq (x-a)(f(b) - f(a)) && \Leftrightarrow && \text{(ii)} \\ (b-a)(f(x) - f(b)) &\leq (b-x)(f(a) - f(b)) && \Leftrightarrow && \text{(iii)} \\ (b-x)(f(x) - f(a)) &\leq (x-a)(f(b) - f(x)) && \Leftrightarrow && \text{(iv)}. \end{aligned}$$

Die analogen Aussagen für streng konvex lassen sich genauso beweisen, wenn wir alle Ungleichungen $\leq$ durch $<$ ersetzen. **q.e.d.**

Korollar 7.25. Für eine stetige reelle Funktion auf einem Intervall I , die im Inneren von I differenzierbar ist, ist folgendes äquivalent:

- (i) f ist (streng) konvex
- (ii) f' ist (streng) monoton wachsend

Beweis: (i) $\Rightarrow$ (ii): Seien $a < x < b$ Punkte im Inneren von I . Die Grenzwerte $x \rightarrow a+$ in (ii) und $x \rightarrow b-$ in (iii) aus Satz 7.24 zeigen $f'(a) \leq \frac{f(b)-f(a)}{b-a} \leq f'(b)$ und damit (ii). (ii) $\Rightarrow$ (i): Für $[a, b] \subset I$ und $x \in (a, b)$ gibt es wegen dem Mittelwertsatz $y \in (a, x)$ und $z \in (x, b)$ mit $\frac{f(x)-f(a)}{x-a} = f'(y) \leq f'(z) = \frac{f(b)-f(x)}{b-x}$. Aus Satz 7.24 (iv) folgt (i). **q.e.d.**

Korollar 7.26. Für eine stetige reelle Funktion auf einem Intervall, die im Inneren zweimal differenzierbar ist, ist folgendes äquivalent

- (i) f ist (streng) konvex.
- (ii) $f''(x) \geq 0$ im Inneren des Intervalls (und der Abschluss der Menge $\{x \mid f''(x) > 0\}$ ist das ganze Intervall).

Dieses Korollar folgt sofort aus Korollar 7.25 und Satz 7.16.

q.e.d.

Wenn wir die Ungleichungen zwischen den Funktionswerten alle umdrehen, so erhalten wir die analogen Aussagen für konkave Funktionen. Also ist eine Funktion f genau dann (streng) konkav, wenn die negative Funktion $-f$ (streng) konvex ist.

Übungsaufgabe 7.27. Zeige, dass die Umkehrfunktion einer (streng) konvexen bijektiven (streng) monoton wachsenden Funktion (streng) konkav ist.

Beispiel 7.28. (i) $f(x) = x^\alpha$, $\alpha > 1 \Rightarrow f'' = \alpha(\alpha-1)x^{\alpha-2}$. Also ist f streng konvex.

(ii) $f(x) = x^\alpha$, $0 < \alpha < 1 \Rightarrow f'' = \alpha(1-\alpha)x^{\alpha-2}$. Also ist f streng konkav.

(iii) $\exp : \mathbb{R} \rightarrow (0, \infty)$, $x \mapsto \exp(x) \Rightarrow \exp'' = \exp$. Also ist $\exp$ streng konvex.

(iv) $\ln : (0, \infty) \rightarrow \mathbb{R}$, $x \mapsto \ln(x) \Rightarrow \ln''(x) = -\frac{1}{x^2}$. Also ist $\ln$ streng konkav.

Satz 7.29 (Ungleichung von Jensen)*. Sei f eine reelle konvexe Funktion auf einem Intervall. Seien $x_1, \dots, x_n$ im Definitionsbereich und $\lambda_1, \dots, \lambda_n$ positive Zahlen, die $\lambda_1 + \dots + \lambda_n = 1$ erfüllen. Dann gilt

$$f(\lambda_1 x_1 + \dots + \lambda_n x_n) \leq \lambda_1 f(x_1) + \dots + \lambda_n f(x_n).$$

Wenn f streng konvex ist, dann gilt Gleichheit nur für $x_1 = x_2 = \dots = x_n$.

Beweis*: durch vollständige Induktion:

- (i) Für $n = 1$ muss $\lambda_1 = 1$ sein, so dass die Aussage klar ist.

- (ii) Die Aussage gelte für $n \in \mathbb{N}$. Seien $x_1, \dots, x_{n+1}$ und $\lambda_1, \dots, \lambda_{n+1}$ wie gefordert. Dann definieren wir $\lambda = \lambda_1 + \dots + \lambda_n$ und $x = \frac{\lambda_1}{\lambda}x_1 + \dots + \frac{\lambda_n}{\lambda}x_n$. Also gilt $\lambda_{n+1} = 1 - \lambda$ und $\frac{\lambda_1}{\lambda} + \dots + \frac{\lambda_n}{\lambda} = 1$. Dann folgt aus der Induktionsvoraussetzung $f(x) \leq \frac{\lambda_1}{\lambda}f(x_1) + \dots + \frac{\lambda_n}{\lambda}f(x_n)$. Wenn f streng konvex ist, dann gilt Gleichheit nur für $x_1 = \dots = x_n$. Weil f konvex ist folgt aber $f(\lambda x + (1 - \lambda)x_{n+1}) \leq \lambda f(x) + (1 - \lambda)f(x_{n+1}) \leq \lambda_1 f(x_1) + \dots + \lambda_{n+1} f(x_{n+1})$. Wenn f streng konvex ist, dann gilt Gleichheit wieder nur für $x_{n+1} = x = x_1 = \dots = x_n$. **q.e.d.**

Korollar 7.30 (Ungleichung arithmetisches-geometrisches Mittel)*. Seien $\lambda_1, \dots, \lambda_n$ positive Zahlen mit $\lambda_1 + \dots + \lambda_n = 1$. Dann gilt für positive Zahlen $x_1, \dots, x_n$

$$x_1^{\lambda_1} x_2^{\lambda_2} \cdots x_n^{\lambda_n} \leq \lambda_1 x_1 + \dots + \lambda_n x_n.$$

Insbesondere gilt $\sqrt[n]{x_1 \cdots x_n} \leq \frac{x_1 + \dots + x_n}{n}$. Gleichheit gilt nun für $x_1 = x_2 = \dots = x_n$.

Beweis*: $-\ln$ ist streng konvex. Also folgt aus Jensen's Ungleichung

$$\begin{aligned} -\ln(\lambda_1 x_1 + \dots + \lambda_n x_n) &\leq -\lambda_1 \ln x_1 - \dots - \lambda_n \ln x_n \\ \iff \lambda_1 \ln x_1 + \dots + \lambda_n \ln x_n &\leq \ln(\lambda_1 x_1 + \dots + \lambda_n x_n) \end{aligned}$$

Wegen der Monotonie von $\exp$ folgt:

$$x_1^{\lambda_1} \cdots x_n^{\lambda_n} = \exp(\lambda_1 \ln x_1 + \dots + \lambda_n \ln x_n) \leq \lambda_1 x_1 + \dots + \lambda_n x_n. \quad \mathbf{q.e.d.}$$

Ersetzen wir $x_1, \dots, x_n$ durch $y_1^{1/\lambda_1}, \dots, y_n^{1/\lambda_n}$ so erhalten wir

Korollar 7.31* Seien $\lambda_1, \dots, \lambda_n$ positive Zahlen mit $\lambda_1 + \dots + \lambda_n = 1$. Dann gilt

$$y_1 \cdots y_n \leq \lambda_1 y_1^{1/\lambda_1} + \dots + \lambda_n y_n^{1/\lambda_n}$$

für alle $y_1, \dots, y_n \in (0, \infty)$. Gleichheit gilt nur für $y_1^{1/\lambda_1} = y_2^{1/\lambda_2} = \dots = y_n^{1/\lambda_n}$. **q.e.d.**

Korollar 7.32 (Youngsche Ungleichung). Seien $p, q \in (0, \infty)$ mit $\frac{1}{p} + \frac{1}{q} = 1$. Dann gilt

$$xy \leq \frac{1}{p}x^p + \frac{1}{q}y^q \quad \text{für alle } x, y \in (0, \infty) \text{ und Gleichheit nur } x^p = y^q.$$

Beweis: Wegen der strengen Monotonie von $\ln$ ist diese Ungleichung äquivalent zu

$$\ln(xy) = \frac{1}{p} \ln(a) + \frac{1}{q} \ln(b) \leq \ln\left(\frac{1}{p}a + \frac{1}{q}b\right) = \ln\left(\frac{1}{p}x^p + \frac{1}{q}y^q\right) \quad \text{mit } a = x^p \text{ und } b = y^q.$$

Weil $\ln$ streng konkav ist, gilt diese Ungleichung und Gleichheit nur für $a = b$. **q.e.d.**

7.6 Taylorreihen

Auf offenen Intervallen I (Teilmenge von $\mathbb{R}$) ist die Ableitung f' einer differenzierbaren Funktion f wieder eine Funktion auf I . Die Bildung der Ableitung ist also eine lineare Abbildung $\frac{d}{dx}$, die differenzierbaren Funktionen auf I , Funktionen auf I zuordnet. Wenn die Ableitung wieder differenzierbar ist, können wir diese Abbildung nochmal anwenden und erhalten $(\frac{d}{dx})^2 f = f''$ die zweite Ableitung von f . Durch n -faches Anwenden erhalten wir gegebenenfalls dann die n -te Ableitung $(\frac{d}{dx})^n f = f^{(n)}$.

Definition 7.33. Für alle $n \in \mathbb{N}_0$ und alle offenen Teilmengen $I \subset \mathbb{R}$ sei $C^n(I)$ die Menge aller n -mal stetig differenzierbaren reellen Funktionen auf I , und $C^\infty(I)$ die Menge aller beliebig oft differenzierbaren reellen Funktionen auf I .

$$C(I) = C^0(I) \supset C^1(I) \supset \dots \supset C^n(I) \supset \dots \supset C^\infty(I)$$

Beispiel 7.34. (i) $\exp \in C^\infty(\mathbb{R})$, weil $\exp^{(n)} = \exp$.

(ii) für alle $n \in \mathbb{N}_0$ ist $x \mapsto x^n \in C^\infty(\mathbb{R})$, weil $(x^n)^{(n)} = n!$ und $(x^n)^{(m)} = 0$ für $m > n$.

(iii) für alle $n \in \mathbb{N}$ ist $x \mapsto x^{-n} \in C^\infty(\mathbb{R} \setminus \{0\})$, weil $(x \mapsto x^{-n})^{(m)} =$

$$x \mapsto \frac{(-n)(-n-1)\dots(-n-m+1)}{x^{n+m}} = (-1)^m \frac{(n+m-1)(n+m-2)\dots n}{x^{n+m}}.$$

(iv) $\ln \in C^\infty((0, \infty))$ weil $\ln^{(m)}(x) = \frac{(-1)^{m-1}(m-1)!}{x^m}$ für $m \geq 1$ und mit $0! = 1$.

Übungsaufgabe 7.35. Zeige mit vollständiger Induktion für alle $n \in \mathbb{N}$:

(i) $\frac{d^n}{dx^n} f \cdot g = \sum_{k=0}^n \binom{n}{k} f^{(k)} g^{(n-k)}$ für alle $f, g \in C^n(I)$ (Verallgemeinerte Leibnizregel).

(ii) $C^n(I)$ ist eine Unteralgebra von $C(I)$.

Aus der Rechenregel und der Kettenregel folgt auch

Korollar 7.36. (i) Die Komposition von n -mal (stetig) differenzierbaren Funktionen ist wieder n -mal (stetig) differenzierbar.

(ii) Die Umkehrfunktion einer n -mal (stetig) differenzierbaren bijektiven Funktion ist n -mal (stetig) differenzierbar, wenn die Ableitung keine Nullstellen hat. **q.e.d.**

Definition 7.37. (Taylor-Polynom) Sei $f : I \rightarrow \mathbb{R}$ in x_0 n -mal differenzierbar, d.h. es gibt eine offene Umgebung $O \subset I$ von x_0 , so dass die Einschränkung von f auf O in $C^{n-1}(O)$ liegt, und $f^{(n-1)}$ in x_0 differenzierbar ist. Dann heißt

$$T_{n,x_0}(x) = f(x_0) + (x - x_0)f'(x_0) + \frac{(x - x_0)^2}{2}f^{(2)}(x_0) + \dots + \frac{(x - x_0)^n}{n!}f^{(n)}(x_0)$$

das Taylorpolynom von f der Ordnung n in x_0 .

Offenbar hat das Taylorpolynom der Ordnung n in x_0 die gleichen Ableitungen bis zur Ordnung n wie f an dem Punkt x_0 . Es ist das eindeutig bestimmte Polynom vom Grad n , das an der Stelle x_0 die Ableitungen $f(x_0), f^{(1)}(x_0), \dots, f^{(n)}(x_0)$ besitzt:

$$p(x) = c_0 + c_1(x - x_0) + \dots + c_n(x - x_0)^n \implies p(x_0) = c_0, p^{(1)}(x_0) = c_1, \dots, p^{(n)}(x_0) = n!c_n.$$

Satz 7.38 (Taylorformel). *Sei $f \in C^n((a, b))$ auf (a, b) $n+1$ -mal differenzierbar. Dann gibt es für jedes $x_0 \neq x \in (a, b)$ ein $\xi \in (x_0, x)$ bzw. $\xi \in (x, x_0)$, so dass*

$$f(x) = T_{n,x_0}(x) + \frac{f^{(n+1)}(\xi)}{(n+1)!}(x - x_0)^{n+1} \quad \text{gilt.}$$

Beweis: Sei $x \in (a, b)$ und $g(t) = \sum_{k=0}^n \frac{f^{(k)}(t)}{k!}(x - t)^k$ für $t \in (a, b)$. Dann gilt

$$g'(t) = \sum_{k=0}^n \frac{f^{(k+1)}(t)}{k!}(x - t)^k - \sum_{k=1}^n \frac{f^{(k)}(t)}{(k-1)!}(x - t)^{k-1} = \frac{f^{(n+1)}(t)}{n!}(x - t)^n.$$

Außerdem sei $h(t) = (x - t)^{n+1}$ und $h'(t) = -(n+1)(x - t)^n$. Dann folgt aus dem verallgemeinerten Mittelwertsatz, dass es ein $\xi \in (x_0, x)$ bzw. $\xi \in (x, x_0)$ gibt mit

$$(g(x) - g(x_0))h'(\xi) = (h(x) - h(x_0))g'(\xi).$$

Es gilt aber $g(x) - g(x_0) = f(x) - T_{n,x_0}(x)$ und $h(x) - h(x_0) = -(x - x_0)^{n+1}$. Also folgt $f(x) - T_{n,x_0}(x) = \frac{f^{(n+1)}(\xi)(x - \xi)^n(x - x_0)^{n+1}}{n!(n+1)(x - \xi)^n} = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x - x_0)^{n+1}$. **q.e.d.**

Definition 7.39 (Taylorreihe). *Für $f \in C^\infty((a, b))$ und $x_0 \in (a, b)$ heißt die Potenzreihe $\sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n$ Taylorreihe von f in x_0 .*

Für $f \in C^\infty((a, b))$ und $x_0, x \in (a, b)$ gibt es drei Möglichkeiten:

- (i) Die Taylorreihe von f in x_0 konvergiert an dem Punkt x gegen $f(x)$.
- (ii) Die Taylorreihe von f in x_0 konvergiert an dem Punkt x , aber nicht gegen $f(x)$.
- (iii) Die Taylorreihe von f in x_0 konvergiert an dem Punkt x nicht.

Korollar 7.40. *Sei $f \in C^\infty((a, b))$ und $x \neq x_0 \in (a, b)$. Dann konvergiert die Taylorreihe von f in x_0 an dem Punkt x gegen $f(x)$, wenn auf $\xi \in (x_0, x)$ bzw. (x, x_0) die Folge $(\frac{|f^{(n)}(\xi)|}{n!}|x - x_0|^n)_{n \in \mathbb{N}_0}$ gleichmäßig gegen Null konvergiert.* **q.e.d.**

Satz 7.41. *Für jede Potenzreihenfunktion $f(x) = \sum_{n=0}^{\infty} a_n x^n$ mit Konvergenzradius $R > 0$ und jedes $|x_0| < R$ hat die Taylorreihe von $f(x)$ in x_0 einen Konvergenzradius nicht kleiner als $R - |x_0|$, und konvergiert auf $B(x_0, R - |x_0|)$ gegen f .*

Beweis: Wegen Beispiel 7.4 (v) stimmen bei $x_0 = 0$ die Ableitungen von f bis zur Ordnung N mit den entsprechenden Ableitungen von $\sum_{n=0}^N a_n x^n$ überein. Deshalb sind $T_{n,0} = \sum_{n=0}^N a_n x^n$ die Taylorpolynome von f bei $x_0 = 0$ und f ist dort die Taylorreihe. Dann folgt die Aussage aus dem Identitätssatz für Potenzreihenfunktionen (ii). **q.e.d.**

Definition 7.42. Eine Funktion $f \in C^\infty(I)$ heißt reellanalytisch bei x_0 , falls die Taylorreihe bei x_0 einen Konvergenzradius größer als Null hat und auf einer Umgebung von x_0 gegen $f(x)$ konvergiert. Sie heißt reellanalytisch, wenn das für alle $x_0 \in I$ gilt.

Also sind alle Potenzreihenfunktionen im Inneren ihres Konvergenzbereiches reellanalytisch. Umgekehrt sind alle reellanalytischen Funktionen Potenzreihenfunktionen.

Korollar 7.43.* Zwei reellanalytische Funktionen $f, g \in C^\infty((a, b))$ stimmen genau dann auf (a, b) überein, wenn ihre Taylorreihen für ein $x_0 \in (a, b)$ übereinstimmen.

Beweis*: Die folgenden Mengen sind abgeschlossen und offen, also $(a, x_0]$ bzw. $[x_0, b)$:

$$\{x \in (a, x_0] \mid f(t) = g(t) \forall t \in [x, x_0]\}, \quad \{y \in [x_0, b) \mid f(t) = g(t) \forall t \in [x_0, y]\}. \quad \mathbf{q.e.d.}$$

Beispiel 7.44. (i) Alle Polynome und die Funktionen $\exp, \sin, \cos, x \mapsto a^x$ sind reellanalytische Funktionen auf ganz $\mathbb{R}$.

(ii) Wegen Satz 4.28 (iv) gibt es für jede Potenzreihenfunktion f mit $f(0) \neq 0$ und Konvergenzradius $R > 0$ ein $r > 0$, so dass $|\frac{f(0)-f(x)}{f(0)}| < \frac{Lr}{|f(0)|} \leq 1$ für alle $x \in B(0, r)$ gilt. Offenbar ist $(\frac{f(0)-f(x)}{f(0)})^n$ für alle $n \in \mathbb{N}$ das Produkt von x^n mit einer Potenzreihe in x . Dann ist auch $\frac{1}{f(0)} \sum_{n=0}^{\infty} (\frac{f(0)-f(x)}{f(0)})^n$ eine Potenzreihe, die für $x \in B(0, r)$ absolut gegen $\frac{1}{f(0)} \frac{1}{1-(f(0)-f(x))/f(0)} = \frac{1}{f(x)}$ konvergiert, weil $x \mapsto \sum_{n=0}^{\infty} x^n = \frac{1}{1-x}$ auf $x \in B(0, 1)$ stetig ist. Aus dem Identitätssatz für Potenzreihenfunktionen folgt, dass der Quotient zweier Potenzreihenfunktionen reellanalytisch ist, solange beide Potenzreihenfunktionen absolut konvergieren und der Nenner nicht verschwindet. Also sind alle rationalen Funktionen und $\tan$ und $\cot$ auf dem Definitionsbereich reellanalytisch.

$$\text{(iii) Für } f(x) = \begin{cases} e^{-\frac{1}{x}} & \text{für } x > 0, \\ 0 & \text{für } x \leq 0 \end{cases} \text{ gilt } f^{(n)}(x) = \begin{cases} p_n(\frac{1}{x})e^{-\frac{1}{x}} & \text{für } x > 0 \\ 0 & \text{für } x < 0 \end{cases}$$

mit einem Polynom p_n vom Grad $2n$, das induktiv definiert ist durch

$$p_{n+1}(\frac{1}{x}) = \frac{1}{x^2}(p_n(\frac{1}{x}) - p'_n(\frac{1}{x})), \quad p_0 = 1.$$

Wegen der 2. Regel von L'Hopital und der Monotonie von e^x gilt für $\alpha, \beta > 0$:

$$\lim_{x \rightarrow 0+} \frac{\ln x}{x^{-\alpha}} = \lim_{x \rightarrow 0+} -\frac{x^\alpha}{\alpha} = 0, \quad \lim_{x \rightarrow 0+} \frac{\exp(-x^{-\alpha})}{x^\beta} = \lim_{x \rightarrow 0+} \exp(-x^{-\alpha} (1 + \beta \frac{\ln x}{x^{-\alpha}})) = 0.$$

Für alle $n \in \mathbb{N}$ ist dann $x \mapsto \begin{cases} \frac{1}{x^n} e^{-\frac{1}{x}} & \text{für } x > 0 \\ 0 & \text{für } x \leq 0 \end{cases}$ stetig, und wegen dem Mittelwertsatz

$f \in C^\infty(\mathbb{R})$. Die Taylorreihe von f verschwindet bei $x_0 = 0$ mit allen Ableitungen. Also ist f bei $x_0 = 0$ nicht reellanalytisch aber auf $\mathbb{R} \setminus \{0\}$.

Kapitel 8

Das Integral von Funktionen

$$f : \mathbb{R} \rightarrow \mathbb{K}$$

8.1 Regelfunktionen

Zunächst definieren wir das Integral für sogenannte Treppenfunktionen.

Definition 8.1 (Treppenfunktion). Sei $[a, b] \subset \mathbb{R}$ ein kompaktes Intervall. Für jedes Intervall $I \subset [a, b]$ einschließlich der Intervalle, die nur aus einer Zahl bestehen, heißt

$$\chi_I : [a, b] \rightarrow \mathbb{K}, \quad x \mapsto \begin{cases} 1 & \text{falls } x \in I \\ 0 & \text{sonst} \end{cases}$$

charakteristische Funktion von I . Ihre endlichen Linearkombinationen in $B([a, b], \mathbb{K})$ heißen Treppenfunktionen und bilden eine Unteralgebra. Die Grenzwerte in $B([a, b], \mathbb{K})$ von gleichmäßig konvergenten Folgen von Treppenfunktionen heißen Regelfunktionen.

Für jedes Teilintervall $I \subset [a, b]$ ist $[a, b] \setminus I$ entweder ein Intervall oder eine disjunkte Vereinigung von zwei Intervallen, von denen jedes einen der beiden Randpunkte a oder b enthält. Dann ist $[a, b]$ eine disjunkte Vereinigung von zwei oder drei Intervallen, von denen eines I ist. Weil die Schnittmenge von zwei Intervallen entweder leer oder ein Intervall ist, ist $[a, b]$ auch für endlich viele Teilintervalle $I_i \subset [a, b]$ eine disjunkte Vereinigung von endlich vielen disjunkten Intervalle $J_j \subset [a, b]$, so dass jedes I_i die disjunkte Vereinigung von endlich vielen J_j ist.

Satz 8.2 (Definition des Integrals). Für jede Treppenfunktion f auf $[a, b]$ und jede Zerlegung von $[a, b]$ in eine endliche disjunkte Vereinigung von Intervallen J_j , auf denen f jeweils konstant ist, hängt folgendes Integral und nicht von der Wahl der J_j ab:

$$\int_a^b f(x) dx = \sum_{y \in f[[a, b]]} y \sum_{J_j \subset f^{-1}[\{y\}]} |J_j|, \text{ wobei } |J_j| \text{ die Intervalllänge von } J_j \text{ ist.}$$

Beweis: Für jedes $y \in f[[a, b]]$ ist die Summe der Intervalllängen von $f^{-1}[\{y\}]$ unabhängig von der Zerlegung in eine disjunkte Vereinigung von Intervallen. **q.e.d.**

Satz 8.3 (Integral und Hauptsatz für Regelfunktionen).

- (i) Eine Funktion $f : [a, b] \rightarrow \mathbb{K}$ ist genau dann eine Regelfunktion, wenn es Funktionen $v : (a, b] \rightarrow \mathbb{K}$ and $w : [a, b) \rightarrow \mathbb{K}$ gibt, und für jedes $x \in [a, b]$ und jedes $\epsilon > 0$ ein $\delta > 0$, so dass $|f(y) - v(x)| < \epsilon$ für alle $y \in (x - \delta, x) \cap [a, b]$ und $|f(y) - w(x)| < \epsilon$ für alle $y \in (x, x + \delta) \cap [a, b]$ gilt, d.h. $\lim_{y \rightarrow x-} f(y) = v(x)$ and $\lim_{y \rightarrow x+} f(y) = w(x)$. Die Regelfunktionen bilden eine Unter algebra von $B([a, b], \mathbb{K})$, die $C([a, b], \mathbb{K})$ enthält und unter $f \mapsto |f|$ invariant ist.
- (ii) Jede Regelfunktion f hat höchstens abzählbar viele Unstetigkeitsstellen.
- (iii) Das Integral setzt sich zu einer eindeutigen linearen Abbildung von der Unter algebra aller Regelfunktionen in $B([a, b], \mathbb{K})$ nach $\mathbb{K}$ fort, die folgendes erfüllt:

$$\left| \int_a^b f(x) dx \right| \leq \int_a^b |f(x)| dx \leq (b - a) \|f\|_\infty.$$

- (iv) Für eine Regelfunktion f ist $F : [a, b] \rightarrow \mathbb{K}$ mit $F(y) = \int_a^y f(x) dx$ lipschitzstetig mit Lipschitzkonstante $\|f\|_\infty$ und genau bei den $x \in (a, b)$ differenzierbar, bei denen $v(x) = w(x)$ gilt. Dort gilt $F'(x) = v(x)$. Also ist F genau dann auf (a, b) differenzierbar, wenn $F' \in C([a, b], \mathbb{K})$ und wir $f = F'$ wählen können.

Beweis: (i): Sei $f : [a, b] \rightarrow \mathbb{K}$ keine Regelfunktion. Dann gibt es für ein $\epsilon > 0$ keine Treppenfunktion g auf $[a, b]$, die folgendes erfüllt:

$$|f(x) - g(x)| < \epsilon \quad \text{für alle } x \in [a, b].$$

Aus Treppenfunktionen g auf $[a, \frac{a+b}{2}]$ und auf $[\frac{a+b}{2}, b]$ die diese Ungleichung auf ihren Definitionsbereichen erfüllen, können wir eine solche Treppenfunktion g auf $[a, b]$ zusammensetzen. Also gibt es eine Folge von Intervallen $([a_n, b_n])_{n \in \mathbb{N}}$ auf denen keine solche Treppenfunktion existiert, so dass $[a_{n+1}, b_{n+1}]$ jeweils entweder $[a_n, \frac{a_n+b_n}{2}]$ oder $[\frac{a_n+b_n}{2}, b_n]$ und $[a_1, b_1] = [a, b]$ ist. Wegen dem Intervallschachtelungsprinzip enthält $\bigcap_{n \in \mathbb{N}} [a_n, b_n]$ genau ein $x \in [a, b]$. Dann ist die Bedingung in (i) bei x nicht erfüllt, weil für hinreichend großes n sonst die Funktion g , die auf $(-\infty, x) \cap [a_n, b_n]$ gleich $v(x)$, auf $(x, \infty) \cap [a_n, b_n]$ gleich $w(x)$, und bei x gleich $f(x)$ ist, obige Ungleichung auf $x \in [a_n, b_n]$ erfüllt. Jede Funktion f , die die Bedingung in (i) erfüllt, ist also eine Regelfunktion.

Für jede Regelfunktion $f \in B([a, b], \mathbb{K})$ gibt es für jedes $\epsilon > 0$ eine Treppenfunktion $g \in B([a, b], \mathbb{K})$ mit $\|f - g\|_\infty < \frac{\epsilon}{2}$. Für jedes $x \in [a, b]$ wählen wir $\delta > 0$ so, dass g auf $(x - \delta, x) \cap [a, b]$ und $(x, x + \delta) \cap [a, b]$ konstant ist. Dann ist der Abstand zwischen zwei Funktionswerten von f auf $(x - \delta, x) \cap [a, b]$ oder $(x, x + \delta) \cap [a, b]$ jeweils kleiner als ϵ . Das

gilt für jedes $\epsilon > 0$ mit einem geeignet gewählten $\delta > 0$. Für gegen x konvergierende Folgen $(x_n)_{n \in \mathbb{N}}$ in $(-\infty, x) \cap [a, b]$ bzw. in $(x, \infty) \cap [a, b]$ sind $(f(x_n))_{n \in \mathbb{N}}$ Cauchyfolgen mit Grenzwerten $v(x)$ bzw. $w(x)$. Dann gilt (i) für f . Die letzte Aussage von (i) folgt. **(ii):** Wenn x nicht zu den abzählbar vielen Unstetigkeitsstellen einer gleichmäßig gegen f konvergierenden Folge von Treppenfunktionen gehört, dann gibt es für jedes $\epsilon > 0$ ein $\delta > 0$ so dass die Abstände zwischen zwei Funktionswerten von f auf $(x - \delta, x + \delta) \cap [a, b]$ kleiner als ϵ sind. Deshalb ist f an höchstens abzählbar vielen Punkten unstetig.

(iii): Für Treppenfunktionen folgt die Ungleichung in (iii) aus der Dreiecksungleichung. Für eine Regelfunktion f konvergieren dann die Integrale von allen Folgen von Treppenfunktionen, die gleichmäßig gegen f konvergieren, gegen die gleiche Zahl. Deshalb setzt sich das Integral eindeutig zu einer Abbildung von den Regelfunktionen nach $\mathbb{K}$ fort, und diese Fortsetzung ist auch linear und erfüllt die Ungleichung in (iii).

(iv): Für $x < y \in [a, b]$ gilt $|F(y) - F(x)| \leq \int_x^y |f(t)| dt \leq |y - x| \cdot \|f\|_\infty$. Also ist F Lipschitzstetig. Aus $|f(t) - z| < \epsilon$ für $t \in (x, y) \subset [a, b]$ folgt

$$\left| \frac{F(y) - F(x)}{y - x} - z \right| \leq \frac{1}{|y - x|} \int_x^y |f(t) - z| dt < \epsilon.$$

Wegen (i) sind mit $z = v(y)$ bzw. $z = w(x)$ die Grenzwerte $\lim_{x \uparrow y} -$ bzw. $\lim_{y \downarrow x} +$ Null. Also ist F genau bei den $x \in (a, b)$ differenzierbar, bei denen $v(x) = w(x) = F'(x)$ gilt. Gilt das für alle $x \in (a, b)$, dann definieren wir $\tilde{f}(x) = v(x) = w(x)$ für $x \in (a, b)$ und $\tilde{f}(a) = w(a)$ und $\tilde{f}(b) = v(b)$. Für jedes $x \in [a, b]$ und jede Folge $(\tilde{x}_n)_{n \in \mathbb{N}}$ in $[a, x) \cup (x, b]$ mit $\lim \tilde{x}_n = x$ konvergiert dann $(f(\tilde{x}_n))_{n \in \mathbb{N}}$ gegen $\tilde{f}(x)$. Für jedes $n \in \mathbb{N}$ gibt es ein $x_n \in [a, x) \cup (x, b]$ mit $\max\{|x_n - x|, |f(x_n) - \tilde{f}(\tilde{x}_n)|\} \leq |\tilde{x}_n - x|$. Dann konvergiert $(x_n)_{n \in \mathbb{N}}$ gegen x , und $(\tilde{f}(\tilde{x}_n))_{n \in \mathbb{N}}$ mit $(f(x_n))_{n \in \mathbb{N}}$ gegen $\tilde{f}(x)$. Also ist $\tilde{f}$ stetig. Das entsprechende $\tilde{F}$ in (iv) erfüllt $\tilde{F}' = \tilde{f} = F'$ und $\tilde{F}(a) = F(a)$, also $\tilde{F} = F$. **q.e.d.**

Übungsaufgabe 8.4. Zeige, dass reelle monotone Funktionen Regelfunktionen sind.

8.2 Technik des Integrierens

In Definition (8.16) wird das Integral auf eine Unteralgebra $\mathcal{R}[a, b]$ von $B([a, b], \mathbb{R})$ erweitert, die die Regelfunktionen enthält. Wenn für ein $f \in \mathcal{R}[a, b]$ ein $F \in C([a, b], \mathbb{R})$ existiert, das auf (a, b) differenzierbar ist mit $F' = f$, dann gilt für alle $[x, y] \subset [a, b]$

$$\int_x^y f(t) dt = F(y) - F(x), \quad \int_y^x f(t) dt = - \int_x^y f(t) dt = F(x) - F(y).$$

Definition 8.5 (Stammfunktion). Auf einer offenen Teilmenge von $\mathbb{R}$ heißt eine differenzierbare Funktion F mit $F' = f$ Stammfunktion von f und wird $\int f(x) dx$ bezeichnet. Die Differenz zweier Stammfunktionen von f ist auf einem Intervall konstant.

Beispiel 8.6. (i) $\int x^\alpha dx = \frac{x^{\alpha+1}}{\alpha+1} + C$ für $\alpha \neq -1$ und entweder $\alpha \in \mathbb{N}$ oder $x \in (0, \infty)$.

(ii) $\int \frac{1}{x} dx = \ln |x| + C$ für $x \neq 0$.

(iii) $\int e^x dx = e^x + C$.

(iv) $\int a^x dx = \frac{a^x}{\ln(a)} + C$ für $a \in (0, \infty) \setminus \{1\}$.

(v) $\int \cos(x) dx = \sin(x) + C$.

(vi) $\int \sin(x) dx = -\cos(x) + C$.

(vii) $\int \tan(x) dx = -\ln |\cos(x)| + C$ für $x \notin \{(n + \frac{1}{2})\pi \mid n \in \mathbb{Z}\}$.

(viii) $\int \cot(x) dx = \ln |\sin(x)| + C$ für $x \notin \{n\pi \mid n \in \mathbb{Z}\}$.

(ix) $\int \frac{1}{1+x^2} dx = \arctan(x) + C$.

(x) $\int \frac{1}{\sqrt{1-x^2}} dx = \arcsin(x) + C$ für $x \in [-1, 1]$.

Substitutionsregel 8.7. Sei $f \in C([a, b])$ und $\phi : [\alpha, \beta] \rightarrow [a, b]$ eine stetige auf (α, β) differenzierbare Funktion, so dass $\phi' \in \mathcal{R}[\alpha, \beta]$. Dann gilt

$$\int f(\phi(t))\phi'(t)dt = \left(\int f(x)dx\right) \circ \phi + C, \quad \int_{\alpha}^{\beta} f(\phi(t))\phi'(t)dt = \int_{\phi(\alpha)}^{\phi(\beta)} f(x)dx.$$

Beweis: Wegen dem Hauptsatz 8.3 hat f eine stetig differenzierbare Stammfunktion F . Also ist $(F \circ \phi)' = (F' \circ \phi) \cdot \phi'$. Weil $\mathcal{R}[a, b]$ eine Unter algebra von $B([a, b], \mathbb{R})$ ist, die $C([a, b])$ enthält, folgt $(F' \circ \phi) \cdot \phi' = (f \circ \phi) \cdot \phi' \in \mathcal{R}[\alpha, \beta]$ und die Aussage. **q.e.d.**

Wenn ϕ bijektiv ist, können wir auch umgekehrt schließen:

Korollar 8.8 (Transformation der Variablen). Sei $\phi : [\alpha, \beta] \rightarrow [a, b]$ bijektiv, stetig und auf (α, β) differenzierbar mit $\phi' \in \mathcal{R}[\alpha, \beta]$ und $f \in C([a, b])$. Dann gilt

$$\int f(x)dx = \left(\int f(\phi(t))\phi'(t)dt\right) \circ \phi^{-1} + C, \quad \int_a^b f(x)dx = \int_{\phi^{-1}(a)}^{\phi^{-1}(b)} f(\phi(t))\phi'(t)dt. \text{q.e.d.}$$

Beispiel 8.9. (i)

$$\int_a^b f(\alpha t + \beta)dt = \frac{1}{\alpha} \int_{\alpha a + \beta}^{\alpha b + \beta} f(x)dx$$

Insbesondere ist das Restglied der Taylorformel im Beweis von Satz 7.38

$$\begin{aligned} f(x) - T_{n,x_0}(x) &= g(x) - g(x_0) = \int_{x_0}^x g'(t)dt = \int_{x_0}^x \frac{f^{(n+1)}(t)}{n!} (x-t)^n dt \\ &= \frac{(x-x_0)^{n+1}}{n!} \int_0^1 f^{(n+1)}(x_0 + s(x-x_0))(1-s)^n ds \quad \text{mit } t = x_0 + s(x-x_0). \end{aligned}$$

(ii)

$$\int R\left(x, \sqrt[n]{ax+b}\right) dx = \int R\left(\frac{t^n-b}{a}, t\right) \frac{n}{a} t^{n-1} dt + C$$

für $n \in \mathbb{N}, a, b \in \mathbb{R}, a \neq 0$ und einer Funktion $R(\cdot, \cdot)$ in zwei Variablen. Wir substituieren $t = \sqrt[n]{ax+b} \Rightarrow ax+b = t^n \Rightarrow x = \frac{t^n-b}{a}$ und $dx = \frac{n}{a} t^{n-1} dt$.

(iii) mit den Funktionen $\sinh x = \frac{e^x - e^{-x}}{2}$ und $\cosh x = \frac{e^x + e^{-x}}{2}$ gilt

$$\int R\left(x, \sqrt{x^2+1}\right) dx = \int R(\sinh t, \cosh t) \cosh t dt + C$$

mit der Substitution $x = \sinh t, \sqrt{x^2+1} = \cosh t$ und $dx = \cosh t dt$.

(iv)

$$\int R\left(x, \sqrt{x^2-1}\right) dx = \pm \int R(\pm \cosh t, \sinh t) \sinh t dt + C$$

mit der Substitution $x = \pm \cosh t$, je nachdem ob x in $[1, \infty)$ oder $(-\infty, -1]$ liegt. Dann gilt $\sqrt{x^2-1} = \sinh t$ und $dx = \pm \sinh t dt$.

(v)

$$\int R\left(x, \sqrt{1-x^2}\right) dx = \mp \int R(\pm \cos t, \sin t) \sin t dt + C.$$

mit der Substitution $x = \pm \cos t, \sqrt{1-x^2} = \sin t$ und $dx = \mp \sin t dt$.

(vi)

$$\int R(\cos x, \sin x) dx = \int R\left(\frac{1-t^2}{1+t^2}, \frac{2t}{1+t^2}\right) \frac{2dt}{1+t^2} + C$$

mit der Substitution $t = \tan\left(\frac{x}{2}\right), x = 2 \arctan(t)$ und $dx = \frac{2dt}{1+t^2}$, so dass gilt

$$\frac{1-t^2}{1+t^2} = \frac{\cos^2\left(\frac{x}{2}\right) - \sin^2\left(\frac{x}{2}\right)}{\cos^2\left(\frac{x}{2}\right) + \sin^2\left(\frac{x}{2}\right)} = \cos(x) \text{ und } \frac{2t}{1+t^2} = \frac{2 \cos\left(\frac{x}{2}\right) \sin\left(\frac{x}{2}\right)}{\cos^2\left(\frac{x}{2}\right) + \sin^2\left(\frac{x}{2}\right)} = \sin(x).$$

(vii)

$$\int R(\cosh x, \sinh x) dx = \int R\left(\frac{t^2+1}{2t}, \frac{t^2-1}{2t}\right) \frac{dt}{t} + C$$

mit der Substitution $t = e^x, x = \ln(t)$ und $dx = \frac{dt}{t}$.

Partielle Integration 8.10. Seien $f, g \in C([a, b])$ auf (a, b) differenzierbar mit $f', g' \in \mathcal{R}[a, b]$. Dann gilt

$$\int f g' dx = f g - \int f' g dx + C, \quad \int_a^b f g' dx = f(b)g(b) - f(a)g(a) - \int_a^b f' g dx.$$

Beweis folgt aus dem Hauptsatz der Differentialrechnung und der Leibnizregel. **q.e.d.**

Beispiel 8.11. (i) $\int \ln(x) dx = x \ln(x) - \int x \frac{1}{x} dx = x(\ln(x) - 1) + C.$

(ii) $\int \sqrt{1-x^2} dx = x\sqrt{1-x^2} + \int \frac{x^2}{\sqrt{1-x^2}} dx + C = x\sqrt{1-x^2} - \int \frac{1-x^2}{\sqrt{1-x^2}} dx + \int \frac{1}{\sqrt{1-x^2}} dx + C$
 $\implies \int \sqrt{1-x^2} dx = \frac{x\sqrt{1-x^2}}{2} + \frac{\arcsin(x)}{2} + C \quad \text{also} \quad \int_{-1}^1 \sqrt{1-x^2} dx = \frac{\pi}{2}.$

(iii) $\int x^n e^x dx = x^n e^x - n \int x^{n-1} e^x dx$

(iv) $\int x^n \cos x dx = x^n \sin x - n \int x^{n-1} \sin x dx.$

(v) $\int x^n \sin x dx = -x^n \cos x + n \int x^{n-1} \cos x dx.$

Partialbruchzerlegung 8.12 (Integration von rationalen Funktionen $f(x) = \frac{P(x)}{Q(x)}$).

1. Faktorisierung des Nenners. In der Algebra $\mathbb{K}[x]$ der reellen bzw. komplexen Polynome heißt $q(x) \in \mathbb{K}[x]$ Teiler von $p(x) \in \mathbb{K}[x]$, wenn es ein $r(x) \in \mathbb{K}[x]$ mit $p(x) = q(x)r(x)$ gibt. Wegen dem Fundamentalsatz der Algebra zerfällt $Q(x) \in \mathbb{C}[x]$ in ein Produkt von Polynomen ersten Grades. $\mathbb{R}[x]$ ist in $\mathbb{C}[x]$ enthalten, und $Q(x) \in \mathbb{C}[x]$ ist genau dann reell, wenn $\overline{Q(x)} = Q(\bar{x})$ für alle $x \in \mathbb{C}$ gilt. Deshalb sind die komplexen Nullstellen von $Q(x) \in \mathbb{R}[x]$ entweder reell oder komplex konjugierte Paare. Dabei ist $(x - x_0)(x - \bar{x}_0) = x^2 - 2\Re(x_0)x + |x_0|^2 \in \mathbb{R}[x]$. Also zerfällt $Q(x)$ in $\mathbb{R}[x]$ in

$$Q(x) = C \prod_i (x - x_i)^{k_i} \prod_j (x^2 + p_j x + q_j)^{l_j} \quad \text{mit } p_j^2 - 4q_j < 0.$$

Wenn wir $P(x)$ und $Q(x)$ durch den Koeffizienten C von $Q(x)$ teilen, wird $C = 1$.

2. Polynomdivision. Wir beschreiben zuerst die Zerlegung von $\frac{P(x)}{Q(x)}$ und die Berechnung der Summanden. Lemma 8.13 zeigt dann, dass die Zerlegung immer möglich ist.

Wenn der Grad des Zählers $P(x)$ nicht kleiner ist als der Grad des Nenners, dann subtrahieren wir von $P(x)$ der Reihe nach das Produkt von solchen Monomen $S_l x^l$ mit $Q(x)$, so dass sich jeweils der Grad der Differenz um Eins erniedrigt. Damit können wir solange fortfahren, bis der Grad der Differenz niedriger ist als der von $Q(x)$. Dann haben wir das Polynom $S(x)$ und ein Polynom $R(x)$ bestimmt, dessen Grad kleiner ist als der von $Q(x)$, so dass $P(x) - S(x)Q(x) = R(x)$ bzw. $\frac{P(x)}{Q(x)} = S(x) + \frac{R(x)}{Q(x)}$ gilt. Weil im Grenzwert $|x| \rightarrow \infty$ alle anderen Summanden in Lemma 8.13 und $\frac{R(x)}{Q(x)}$ verschwinden, stimmt dieses Polynom $S(x)$ tatsächlich mit dem aus Lemma 8.13 überein.

Danach bestimmen wir die Koeffizienten a_{jl}, b_{jl} und c_{ik} mit dem Ansatz

$$\frac{R(x)}{Q(x)} = \sum_i \sum_{k=1}^{k_i} \frac{c_{ik}}{(x - x_i)^k} + \sum_j \sum_{l=1}^{l_j} \frac{a_{jl}x + b_{jl}}{(x^2 + p_j x + q_j)^l}$$

Wenn wir beide Seiten mit $Q(x)$ multiplizieren, erhalten wir eine Gleichung zwischen 2 reellen Polynomen. Durch Einsetzen von geeigneten Werten von x (Nullstellen von $Q(x)$) und durch Koeffizientenvergleich erhalten wir die Zahlen a_{jl}, b_{jl} und c_{ik} .

Lemma 8.13. (i) Eine rationale Funktion $\frac{P(x)}{Q(x)}$ mit komplexen Koeffizienten lässt sich schreiben als eine Summe eines komplexen Polynoms $S(x)$ und Summanden von der Form $\frac{c_{ik}}{(x-x_i)^k}$, wobei $(x-x_i)^k$ Teiler von $Q(x)$ sind und $c_{ik} \in \mathbb{C}$.

(ii) Eine reelle rationale Funktion $\frac{P(x)}{Q(x)}$ lässt sich schreiben als eine Summe eines reellen Polynoms $S(x)$ und Summanden von der Form $\frac{c_{ik}}{(x-x_i)^k}$ und $\frac{a_{jl}x+b_{jl}}{(x^2+p_jx+q_j)^l}$, wobei $(x-x_i)^k$ und $(x^2+p_jx+q_j)^l$ reelle Teiler von $Q(x)$ sind mit $a_{jl}, b_{jl}, c_{ik} \in \mathbb{R}$.

Beweis: (i) Sei x_i eine k -fache Nullstelle vom Nennerpolynom $Q(x)$, d.h. $Q(x) = (x-x_i)^k q(x)$ mit $q(x_i) \neq 0$. Dann hat $P(x) - \frac{P(x_i)}{q(x_i)} q(x)$ bei $x = x_i$ eine Nullstelle. Deshalb gibt es ein $p(x) \in \mathbb{C}[x]$ mit $P(x) = \frac{P(x_i)}{q(x_i)} q(x) + (x-x_i)p(x)$. Es folgt

$$\frac{P(x)}{Q(x)} = \frac{\frac{P(x_i)}{q(x_i)} q(x) + (x-x_i)p(x)}{(x-x_i)^l q(x)} = \frac{\frac{P(x_i)}{q(x_i)}}{(x-x_i)^l} + \frac{p(x)}{(x-x_i)^{l-1} q(x)}.$$

Der Grad des Nenners von dem zweiten Summanden ist dabei um Eins kleiner als der Grad von $Q(x)$. Indem wir diese Formel mehrfach bei allen Nullstellen von $Q(x)$ auf diesen Rest anwenden erhalten wir als letzten Summanden ein Polynom $S(x)$.

(ii) Für reelle Nullstellen x_i von $Q(x)$ sind die Koeffizienten in (i) reell. Deshalb genügt es ein analoges Vorgehen für Teiler $Q(x) = (x^2+p_jx+q_j)^l q(x)$ von $Q(x)$ anzugeben, wobei $q(x)$ an den komplexen Nullstellen $x_{1,2} = \frac{-p_j \pm \sqrt{p_j^2 - 4q_j}}{2}$ von $x^2+p_jx+q_j$ nicht verschwindet. Dort verschwindet $P(x) - (ax+b)q(x)$ genau dann, wenn

$$\begin{aligned} \frac{P(x_1)}{q(x_1)} &= ax_1 + b & \frac{P(x_2)}{q(x_2)} &= ax_2 + b \\ a &= \frac{1}{x_1-x_2} \left(\frac{P(x_1)}{q(x_1)} - \frac{P(x_2)}{q(x_2)} \right) & b &= \frac{1}{x_1-x_2} \left(x_1 \frac{P(x_2)}{q(x_2)} - x_2 \frac{P(x_1)}{q(x_1)} \right) \end{aligned}$$

gilt. Weil $P(x)$ und $q(x)$ reelle Koeffizienten haben, gilt $\overline{P(x)} = P(\bar{x})$ und $\overline{q(x)} = q(\bar{x})$. Dann folgt $\overline{\left(\frac{P(x_1)}{q(x_1)}\right)} = \frac{P(x_2)}{q(x_2)}$ aus $\bar{x}_1 = x_2$. Deshalb sind a und b reell mit

$$a = \frac{2}{\sqrt{4q_j-p_j^2}} \Im\left(\frac{P(x_1)}{q(x_1)}\right) \quad b = \Re\left(\frac{P(x_1)}{q(x_1)}\right) + \frac{p_j}{\sqrt{4q_j-p_j^2}} \Im\left(\frac{P(x_1)}{q(x_1)}\right).$$

Wieder gibt es ein $p(x) \in \mathbb{R}[x]$ mit $P(x) = (ax+b)q(x) + (x^2+p_jx+q_j)p(x)$ und

$$\frac{P(x)}{Q(x)} = \frac{ax+b}{(x^2+p_jx+q_j)^l} + \frac{p(x)}{(x^2+p_jx+q_j)^{l-1} q(x)}.$$

Nach mehrmaligem Anwenden erhalten wir als Rest ein reelles Polynom $S(x)$. **q.e.d.**

3. Termweise Integration. $\int \frac{dx}{(x-x_i)^k} = \begin{cases} \ln|x-x_i| + C & \text{für } k=1 \\ \frac{-1}{(k-1)(x-x_i)^{k-1}} + C & \text{sonst.} \end{cases}$

$$\int \frac{(ax+b)dx}{(x^2+px+q)^l} = \begin{cases} \frac{a}{2} \ln(x^2+px+q) + \left(b - \frac{ap}{2}\right) \int \frac{dx}{x^2+px+q} + C & \text{für } l = 1 \\ \frac{-a}{2(l-1)(x^2+px+q)^{l-1}} + \left(b - \frac{ap}{2}\right) \int \frac{dx}{(x^2+px+q)^l} + C & \text{sonst.} \end{cases}$$

$$\int \frac{dx}{(x^2+px+q)^l} = \begin{cases} \frac{2}{\sqrt{4q-p^2}} \arctan\left(\frac{2x+p}{\sqrt{4q-p^2}}\right) + C & l = 1 \\ \frac{2x+p}{(l-1)(4q-p^2)(x^2+px+q)^{l-1}} + \frac{2(2l-3)}{(l-1)(4q-p^2)} \int \frac{dx}{(x^2+px+q)^{l-1}} + C & l \neq 1. \end{cases}$$

Beispiel 8.14. $\int f(x)dx$ mit $f(x) = \frac{2x^5+x^4+x^2+2x-2}{x^4-1}$.

1. Faktorisierung des Nenners. $x^4 - 1 = (x-1)(x+1)(x^2+1)$.

2. Polynomdivision.

$$\begin{aligned} 2x^5 + x^4 + x^2 + 2x - 2 &= (2x+1)(x^4-1) + x^2 + 4x - 1 \\ &\quad - 2x(x^4-1) \\ &= x^4 + x^2 + 4x - 2 \\ &\quad - (x^4-1) \\ &= x^2 + 4x - 1 \end{aligned}$$

$$x^2 + 4x - 1 = c_1(x+1)(x^2+1) + c_2(x-1)(x^2+1) + (ax+b)(x-1)(x+1)$$

Einsetzen von $x = 1$ und $x = -1$ ergibt $4 = 4c_1$ und $-4 = -4c_2$ also $c_1 = 1$ und $c_2 = 1$. Koeffizientenvergleich von x^3 und x^0 ergibt $0 = c_1 + c_2 + a$ und $-1 = c_1 - c_2 - b$. Dann folgt $a = -2$, $b = 1$ und $f(x) = 2x + 1 + \frac{1}{x-1} + \frac{1}{x+1} + \frac{-2x+1}{x^2+1}$.

3. Termweise Integration.

$$\int f(x)dx = x^2 + x + \ln|x-1| + \ln|x+1| - \ln(x^2+1) + \arctan(x) + C.$$

8.3 Riemannintegrale Funktionen

In diesem Abschnitt erweitern wir das Integral auf eine größere Klasse von Funktionen. Dabei machen wir wesentlichen Gebrauch von der Monotonie des Integrals von reellen Treppenfunktionen: für Treppenfunktionen $f \leq g$ in $B([a, b], \mathbb{R})$ gilt $\int_a^b f dx \leq \int_a^b g dx$.

Definition 8.15 (Unterintegral und Oberintegral). Für $f \in B([a, b], \mathbb{R})$ heißt

$$\underline{\int} f = \sup \left\{ \int_a^b g(x)dx \mid g \text{ Treppenfunktion mit } g \leq f \right\} \quad \text{Unterintegral von } f \text{ und}$$

$$\overline{\int} f = \inf \left\{ \int_a^b g(x)dx \mid g \text{ Treppenfunktion mit } g \geq f \right\} \quad \text{Oberintegral von } f.$$

Offenbar gilt $\underline{\int} f \leq \overline{\int} f$. Unserer Zugang zum Riemannintegral über Treppenfunktionen weicht von dem Standardzugang etwas ab, beschreibt aber die gleichen Funktionen.

Definition 8.16 (Riemannintegral). Eine Funktion $f \in B([a, b], \mathbb{R})$ heißt riemannintegrabel, wenn $\int f = \overline{\int} f$ gilt. Diese Zahl heißt Riemannintegral $\int_a^b f dx$ von f über $[a, b]$. Die Menge aller riemannintegrablen Funktionen auf $[a, b]$ bezeichnen wir mit $\mathcal{R}[a, b]$. Für $f \in \mathcal{R}[a, b]$ definieren wir $\int_b^a f(x) dx = -\int_a^b f(x) dx$.

Der Flächeninhalt zwischen dem Graphen von f und der x -Achse liegt offenbar in dem Intervall $[\int f, \overline{\int} f]$ und ist für $f \in \mathcal{R}[a, b]$ das Riemannintegral $\int_a^b f(x) dx$.

Definition 8.17 (Partition). Eine Partition p von $[a, b]$ ist eine Zerlegung von $[a, b]$ in eine endliche Vereinigung von paarweise disjunkten Teilintervallen $[a, b] = I_1 \cup \dots \cup I_n$. Die Feinheit $\|p\|$ der Partition p ist das Maximum der Intervalllängen. Die Menge aller Partitionen von $[a, b]$ bezeichnen wir mit $\mathcal{P}[a, b]$.

Für jede Partition $p \in \mathcal{P}[a, b]$ bilden die Linearkombinationen von $\chi_{I_1}, \dots, \chi_{I_n}$ einen endlichdimensionalen Teilraum von $B([a, b], \mathbb{R})$. Für $f \in B([a, b], \mathbb{R})$ enthält dieser Teilraum ein kleinstes Element g mit $g \geq f$ und ein größtes Element h mit $h \leq f$:

$$g = \sum_i \sup_{x \in I_i} f(x) \chi_{I_i} \qquad h = \sum_i \inf_{x \in I_i} f(x) \chi_{I_i}$$

Ihre Integrale heißen Obersummen $S(f, p) = \int_a^b g dx$ und Untersummen $s(f, p) = \int_a^b h dx$.

Definition 8.18 (Verfeinerung). $p' \in \mathcal{P}[a, b]$ heißt Verfeinerung von $p \in \mathcal{P}[a, b]$, wenn alle Intervalle von p' in einem Intervall von p enthalten sind. Wir schreiben dann $p' \subset p$. Für $p_1, \dots, p_n \in \mathcal{P}[a, b]$ besteht die gemeinsame Verfeinerung $p_1 \cap \dots \cap p_n \in \mathcal{P}[a, b]$ aus allen nichtleeren Schnittmengen $I_1 \cap \dots \cap I_n$ von Intervallen $I_1 \in p_1, \dots, I_n \in p_n$.

Lemma 8.19. (i) Wenn $p' \subset p$ gilt $s(p, f) \leq s(p', f)$ und $S(p', f) \leq S(p, f)$.

(ii) Für $p, p' \in \mathcal{P}[a, b]$ gilt $s(p, f) \leq S(p', f)$.

Beweis:(i) Die den Partitionen p und p' entsprechenden Treppenfunktionen erfüllen mit $h \leq f \leq g$ und $h' \leq f \leq g'$ auch $h \leq g'$ und für $p' \subset p$ sogar $h \leq h' \leq f \leq g' \leq g$. Dann folgen (i) und (ii) aus der Monotonie des Integrals für Treppenfunktionen. **q.e.d.**

Satz 8.20 (Kriterium von Darboux). Eine Funktion $f \in B([a, b], \mathbb{R})$ ist genau dann riemannintegrabel, wenn es für jedes $\epsilon > 0$ ein $p \in \mathcal{P}[a, b]$ gibt mit $S(p, f) - s(p, f) < \epsilon$.

Beweis: Für $f \in \mathcal{R}[a, b]$ und $\epsilon > 0$ gibt es $p', p'' \in \mathcal{P}[a, b]$ mit $S(p'', f) - s(p', f) < \epsilon$. Für $p = p' \cap p''$ folgt $S(p, f) - s(p, f) \leq S(p'', f) - s(p', f) < \epsilon$ aus Lemma 8.19 (i).

Wenn es für jedes $\epsilon > 0$ ein $p_\epsilon \in \mathcal{P}[a, b]$ gibt mit $S(p_\epsilon, f) - s(p_\epsilon, f) < \epsilon$ dann ist $0 \leq \overline{\int} f - \int f \leq S(p_\epsilon, f) - s(p_\epsilon, f) \leq \inf_{\epsilon > 0} \epsilon = 0$. Also gilt dann auch $\int f = \overline{\int} f$. **q.e.d.**

Beispiel 8.21. (i) Für eine Regelfunktion f und eine Treppenfunktion g folgt $g - \epsilon \leq f \leq g + \epsilon$ aus $\|f - g\|_\infty \leq \epsilon$. Die entsprechende Partition $p \in \mathcal{P}[a, b]$ erfüllt $S(p, f) - s(p, f) \leq \int_a^b (g + \epsilon) dx - \int_a^b (g - \epsilon) dx \leq 2\epsilon(b - a)$. Also sind alle Regelfunktionen riemannintegabel.

(ii) Die Summe $\sum_{n \in \mathbb{N}} \chi_{I_n}$ mit $I_n = (\frac{1}{2n+1}, \frac{1}{2n})$ ist auf $[0, 1]$ riemannintegabel, aber keine Regelfunktion, weil sie bei $x = 0$ nicht die Bedingung (i) in Satz 8.3 erfüllt.

(iii) Für $f(x) = \begin{cases} 1 & \text{für } x \in [a, b] \cap \mathbb{Q} \\ 0 & \text{für } x \in [a, b] \text{ und } x \notin \mathbb{Q} \end{cases}$ ist auf allen Teilintervalle $I \subset [a, b]$ mit positiver Intervalllänge $\sup_{x \in I} f(x) = 1$ und $\inf_{x \in I} f(x) = 0$. Also ist $\underline{\int} f = 0$ und $\overline{\int} f = b - a$ und $f \notin \mathcal{R}[a, b]$.

8.4 Kriterium von Riemann

Definition 8.22 (Riemannsummen). Für $p \in \mathcal{P}[a, b]$ und Zwischenpunkte $\xi \in I_1 \times \dots \times I_n$ heißt $R(p, f, \xi) = \int_a^b \sum_{i=1}^n f(\xi_i) \chi_{I_i} dx$ Riemannsumme von f bezüglich p und ξ .

Aus der Definition von $s(p, f)$ und $S(p, f)$ folgt für $f \in B([a, b], \mathbb{R})$ und $p \in \mathcal{P}[a, b]$

$$\inf\{R(p, f, \xi) \mid \xi \in I_1 \times \dots \times I_n\} = s(p, f) \quad \sup\{R(p, f, \xi) \mid \xi \in I_1 \times \dots \times I_n\} = S(p, f).$$

Satz 8.23 (Kriterium von Riemann). $f \in B([a, b], \mathbb{R})$ ist genau dann riemannintegabel, wenn es ein $A \in \mathbb{R}$ und für jedes $\epsilon > 0$ ein $\delta > 0$ gibt, so dass $|R(p, f, \xi) - A| < \epsilon$ für alle $p \in \mathcal{P}[a, b]$ mit $\|p\| < \delta$ und Zwischenpunkte ξ gilt. Dann gilt $A = \int_a^b f(x) dx$.

Beweis: Für jede Funktion $f \in B([a, b], \mathbb{R})$, die das Kriterium von Riemann erfüllt, gibt es für alle $\epsilon > 0$ eine Partition $p \in \mathcal{P}[a, b]$ mit $S(p, f) - s(p, f) =$

$$= \sup\{R(p, f, \xi) \mid \xi \in I_1 \times \dots \times I_n\} - \inf\{R(p, f, \xi) \mid \xi \in I_1 \times \dots \times I_n\} \leq 2\epsilon.$$

Dann ist das Kriterium von Darboux erfüllt. Erfülle umgekehrt $f \in B([a, b], \mathbb{R})$ das Kriterium von Darboux. Dann gibt es für jedes $\epsilon > 0$ ein $p \in \mathcal{P}[a, b]$, so dass $S(p, f) - s(p, f) < \frac{\epsilon}{2}$ gilt. Seien $I_1, \dots, I_n$ die Teilintervalle von p mit positiven Intervalllängen $|I_1|, \dots, |I_n|$, und $\delta = \min\{\frac{\epsilon}{8n\|f\|_\infty + 1}, |I_1|, \dots, |I_n|\}$. Jedes Teilintervall einer Partition $p' \in \mathcal{P}[a, b]$ mit $\|p'\| < \delta$ ist entweder in einem Teilintervall von p enthalten, oder enthält Punkte am linken oder rechten Rand eines der Intervalle $I_1, \dots, I_n$. Also sind höchstens $2n$ Teilintervalle von p' nicht in einem Teilintervall von p enthalten. Es folgt

$$S(p', f) - s(p', f) \leq S(p, f) - s(p, f) + 2\|f\|_\infty \cdot 2n \cdot \delta < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Für alle Zwischenpunkte ξ von p' gilt $s(p', f) \leq R(p', f, \xi)$, $\int_a^b f(x)dx \leq S(p', f)$. Es folgt

$$\left| R(p', f, \xi) - \int_a^b f(x)dx \right| \leq S(p', f) - s(p', f) < \epsilon. \quad \mathbf{q.e.d.}$$

Stimmen von zwei Partitionen $p, p' \in \mathcal{P}[a, b]$ die Inneren der Intervalle überein, so gilt $R(p, f, \xi) = R(p', f, \xi)$ für gleiche Zwischenpunkte ξ . Deshalb ist unser Zugang zum Riemannintegral äquivalent zum Standardzugang.

Korollar 8.24. Sei $f \in \mathcal{R}[a, b]$. Dann gilt für alle $t \in [0, 1]$

$$\int_a^b f(x)dx = \lim_{n \rightarrow \infty} \frac{b-a}{n} \sum_{i=1}^n f\left(a + \frac{i-t}{n}(b-a)\right).$$

Beweis: Für die Partitionen $p_n = \{[a, a + \frac{1}{n}(b-a)), [a + \frac{1}{n}(b-a), a + \frac{2}{n}(b-a)), \dots, [a + \frac{n-1}{n}(b-a), b]\}$ bzw. $p_n = \{[a, a + \frac{1}{n}(b-a)], (a + \frac{1}{n}(b-a), a + \frac{2}{n}(b-a)], \dots, (a + \frac{n-1}{n}(b-a), b]\}$ in $\mathcal{P}[a, b]$ mit den Zwischenpunkten $\xi_i = a + \frac{i-t}{n}(b-a)$ für $i = 1, \dots, n$ ist $\|p_n\| = \frac{b-a}{n}$. Die Aussage folgt aus dem Kriterium von Riemann. $\mathbf{q.e.d.}$

Korollar 8.25*: Seien $f, g \in \mathcal{R}[a, b]$. Wenn f und g auf einer dichten Teilmenge von $[a, b]$ (z.B. $\mathbb{Q} \cap [a, b]$) übereinstimmen, dann gilt $\int_a^b f(x)dx = \int_a^b g(x)dx$.

Beweis*: Weil jedes Teilintervall von positiver Intervalllänge einer beliebigen Partition $p \in \mathcal{P}[a, b]$ immer Elemente einer dichten Teilmenge von $[a, b]$ enthält, können die Zwischenpunkte immer aus einer dichten Teilmenge gewählt werden. $\mathbf{q.e.d.}$

Satz 8.26 (Eigenschaften des Riemannintegrals).

- (i) $\mathcal{R}[a, b]$ ist eine Unteralgebra von $B([a, b], \mathbb{R})$ die die Regelfunktionen und $C([a, b])$ enthält. Die Abbildung $\mathcal{R}[a, b] \rightarrow \mathbb{R}$, $f \mapsto \int_a^b f dx$ ist $\mathbb{R}$ -linear.
- (ii) $\mathcal{R}[a, b]$ enthält die monotonen Funktionen, und mit $f \in \mathcal{R}[a, b]$ auch $|f| \in \mathcal{R}[a, b]$.
- (iii) Monotonie: Für $f, g \in \mathcal{R}[a, b]$ folgt aus $f \leq g$ (d.h. $f(x) \leq g(x)$ für alle $x \in [a, b]$) $\int_a^b f dx \leq \int_a^b g dx$. Insbesondere gilt $|\int_a^b f dx| \leq \int_a^b |f| dx \leq (b-a)\|f\|_\infty$.
- (iv) Normierung: $\int_a^b 1 dx = b - a$.
- (v) Stetigkeit: $f \in \mathcal{R}[a, b]$ und $g \in C(\mathbb{R})$, dann ist $g \circ f \in \mathcal{R}[a, b]$.
- (vi) Intervall Additivität: Für jedes $c \in (a, b)$ gilt:

$$f \in \mathcal{R}[a, b] \iff f \in \mathcal{R}[a, c] \cap \mathcal{R}[c, b] \text{ und } \int_a^b f dx = \int_a^c f dx + \int_c^b f dx.$$

- (vii) Sei $f \in \mathcal{R}[a, b]$ und $(a_n)_{n \in \mathbb{N}}$ und $(b_n)_{n \in \mathbb{N}}$ Folgen in $[a, b]$ mit $\lim_{n \rightarrow \infty} a_n = a$ und $\lim_{n \rightarrow \infty} b_n = b$. Dann konvergiert $(\int_{a_n}^{b_n} f dx)_{n \in \mathbb{N}}$ gegen $\int_a^b f dx$.
- (viii) Gleichmäßige Konvergenz und Stetigkeit des Riemannintegrals: Der Grenzwert f einer gleichmäßig konvergenten Folge $(f_n)_{n \in \mathbb{N}}$ in $\mathcal{R}[a, b]$ liegt auch in $\mathcal{R}[a, b]$ und die Folge $(\int_a^b f_n dx)_{n \in \mathbb{N}}$ konvergiert dann gegen $\int_a^b f dx$.

Beweis:(i) Für $f, g \in B([a, b], \mathbb{R})$ und $p \in [a, b]$ gilt

$$S(p, f + g) \leq S(p, f) + S(p, g) \quad \text{und} \quad -s(p, f + g) \leq -s(p, f) - s(p, g)$$

Daraus und aus $f(x)g(x) - f(y)g(y) = g(x)(f(x) - f(y)) + f(y)(g(x) - g(y))$ folgt

$$\begin{aligned} S(p, f + g) - s(p, f + g) &\leq S(p, f) - s(p, f) + S(p, g) - s(p, g) \\ S(p, fg) - s(p, fg) &\leq \|g\|_\infty (S(p, f) - s(p, f)) + \|f\|_\infty (S(p, g) - s(p, g)) \end{aligned}$$

Aus $f, g \in \mathcal{R}[a, b]$ folgt also wegen dem Darbouxkriterium $f + g, fg \in \mathcal{R}[a, b]$. Wegen Beispiel 8.21 (i) ist jede Regelfunktion riemannintegabel und damit wegen Satz 8.3 auch jede stetige Funktion. Die Linearität des Riemannintegrals folgt aus der Linearität des Integrals von Treppenfunktionen.

(ii) Für $p \in \mathcal{P}[a, b]$ seien $x_0 = a < x_1 < \dots < x_{n-1} < x_n = b$ die entsprechenden Teilintervallendpunkte. Für monoton steigende f folgt aus $\|p\| < \frac{\epsilon}{f(b) - f(a)}$

$$S(p, f) - s(p, f) \leq \sum_{i=1}^n (f(x_i) - f(x_{i-1}))(x_i - x_{i-1}) \leq \|p\| \sum_{i=1}^n f(x_i) - f(x_{i-1}) < \epsilon.$$

Das Kriterium von Darboux zeigt $f \in \mathcal{R}[a, b]$. Analoges gilt für monoton fallende f .

Für $f \in \mathcal{R}[a, b]$ seien $f^\pm(x) = \frac{1}{2}(f(x) \pm |f(x)|)$. Dann folgt

$$0 \leq \sup\{f^\pm(x) \mid x \in I_i\} - \inf\{f^\pm(x) \mid x \in I_i\} \leq \sup\{f(x) \mid x \in I_i\} - \inf\{f(x) \mid x \in I_i\}.$$

Also gilt $S(p, f^\pm) - s(p, f^\pm) \leq S(p, f) - s(p, f)$ für alle $p \in \mathcal{P}[a, b]$. Dann folgt $f^\pm \in \mathcal{R}[a, b]$ und damit auch $|f| = f^+ - f^- \in \mathcal{R}[a, b]$ aus dem Kriterium von Darboux.

(iii) Aus $f, g \in \mathcal{R}[a, b]$ mit $f \leq g$ folgt $\int_a^b f(x) dx = \underline{\int} f \leq \underline{\int} g = \int_a^b g(x) dx$.

(iv) Für $f = 1$ (konstant) gilt $S(p, 1) = s(p, 1) = b - a$ für alle $p \in \mathcal{P}[a, b]$.

(v) Sei $f \in \mathcal{R}[a, b]$ und $g \in C(\mathbb{R})$. Dann ist g auf $[-\|f\|_\infty, \|f\|_\infty]$ gleichmäßig stetig. Also gibt es für jedes $\epsilon > 0$ ein $\delta > 0$, so dass $|g(x) - g(x')| < \frac{\epsilon}{2(b-a)}$ aus $|x - x'| < \delta$ mit $x, x' \in [-\|f\|_\infty, \|f\|_\infty]$ folgt. Sei $\|g\|_\infty = \max\{|g(x)| \mid x \in [-\|f\|_\infty, \|f\|_\infty]\}$. Wegen dem Darbouxkriterium gibt es für $f \in \mathcal{R}[a, b]$ ein $p \in \mathcal{P}[a, b]$ mit $S(p, f) - s(p, f) < \frac{\epsilon \cdot \delta}{4\|g\|_\infty}$. Wir zerlegen die Summe $S(p, g \circ f) - s(p, g \circ f)$ in die Summe über Teilintervalle I_i , auf denen $\sup_{x \in I_i} f(x) - \inf_{x \in I_i} f(x) < \delta$ gilt, und die Summe über Teilintervalle I_j , auf denen $\sup_{x \in I_j} f(x) - \inf_{x \in I_j} f(x) \geq \delta$ gilt. Aus der Wahl von δ folgt, dass die erste

Summe nicht größer ist als $\frac{\epsilon}{2(b-a)} \cdot (b-a) = \frac{\epsilon}{2}$. Weil die Summe der Teilintervalllängen in der zweiten Summe nicht größer ist als $\frac{S(p,f)-s(p,f)}{\delta}$, ist die zweite Summe nicht größer als $(S(p,f) - s(p,f)) \frac{2\|g\|_\infty}{\delta} < \frac{\epsilon}{2}$. Also gilt $S(p, g \circ f) - s(p, g \circ f) < \epsilon$ und $g \circ f$ erfüllt das Darbouxkriterium.

(vi) Für jedes $p \in \mathcal{P}[a, b]$ entspricht die Verfeinerung $p \cap \{[a, c], \{c\}, (c, b]\} \in \mathcal{P}[a, b]$ einem Paar von Partitionen in $\mathcal{P}[a, c] \times \mathcal{P}[c, b]$. Dann folgt (v) aus dem Darbouxkriterium.

(vii) Wegen (vi) und (iii) gilt $|\int_{a_n}^{b_n} f(x)dx - \int_a^b f(x)dx| \leq (a_n - a + b - b_n)\|f\|_\infty$.

(viii) Für $p \in \mathcal{P}[a, b]$ folgt mit $f = (f - f_n) + f_n$ aus dem Beweis von (i)

$$\begin{aligned} S(p, f) - s(p, f) &\leq S(p, f - f_n) - s(p, f - f_n) + S(p, f_n) - s(p, f_n) \\ &\leq 2(b-a)\|f - f_n\|_\infty + S(p, f_n) - s(p, f_n). \end{aligned}$$

Für $\epsilon > 0$ wählen wir zuerst n so groß, dass $\|f - f_n\|_\infty < \frac{\epsilon}{4(b-a)}$ gilt, und dann $p \in \mathcal{R}[a, b]$ so dass $S(p, f_n) - s(p, f_n) < \frac{\epsilon}{2}$ gilt. Dann erfüllt f das Kriterium von Darboux.

Andererseits folgt für $f, f_n \in \mathcal{R}[a, b]$ aus der Monotonie $|\int_a^b f(x)dx - \int_a^b f_n(x)dx| \leq (b-a)\|f - f_n\|_\infty$. Also konvergiert $(\int_a^b f_n(x)dx)_{n \in \mathbb{N}}$ gegen $\int_a^b f(x)dx$. **q.e.d.**

8.5 Differentiation und Integration

Hauptsatz der Differential- und Integralrechnung 8.27. Sei $f \in \mathcal{R}[a, b]$ und F eine stetige Funktion auf $[a, b]$, die auf (a, b) differenzierbar ist mit $F' = f$. Dann gilt

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

Umgekehrt ist $F(y) = \int_a^y f(x)dx$ für $f \in \mathcal{R}[a, b]$ lipschitzstetig mit Lipschitzkonstante $\|f\|_\infty$ und bei den $x \in (a, b)$ differenzierbar mit $F'(x) = f(x)$, bei denen f stetig ist.

Beweis: Für alle $p \in \mathcal{P}[a, b]$ enthält wegen dem Mittelwertsatz jedes Teilintervall mit Abschluss $[x_{i-1}, x_i]$ einen Zwischenpunkt ξ_i mit $f(\xi_i)(x_i - x_{i-1}) = F(x_i) - F(x_{i-1})$ und $R(p, f, \xi) = F(b) - F(a)$. Aus dem Kriterium von Riemann folgt $\int_a^b f(x)dx = F(b) - F(a)$. Die zweite Aussage folgt wie im Satz 8.3 (iv) aus Satz 8.26 (iii). **q.e.d.**

Beispiel 8.28. (i) Sei $1 < \alpha < 2$ und $F(x) = \begin{cases} |x|^\alpha \sin \frac{1}{x} & \text{für } x \neq 0 \\ 0 & \text{für } x = 0. \end{cases}$

Dann ist F für $x \in \mathbb{R} \setminus \{0\}$ differenzierbar mit

$$F'(x) = \alpha \frac{|x|^\alpha}{x} \sin\left(\frac{1}{x}\right) - \frac{|x|^\alpha}{x^2} \cos\left(\frac{1}{x}\right).$$

Wegen $\left| \frac{F(x)-F(0)}{x} \right| = |x|^{\alpha-1} \left| \sin\left(\frac{1}{x}\right) \right| \leq |x|^{\alpha-1}$ ist F mit $F'(0) = 0$ auch bei $x = 0$ differenzierbar. Also gibt es differenzierbare Funktionen, deren Ableitungen auf allen Umgebungen eines Punktes nicht beschränkt sind, und die nicht riemannintegabel sind.

(ii) Sei $f(x) = \begin{cases} 1 & \text{für } x \geq 0 \\ -1 & \text{für } x < 0. \end{cases}$ Dann ist f auf allen kompakten Intervallen riemannintegabel. Offenbar gilt $F(x) = \int_0^x f(t)dt = |x|$. Also sind nicht alle Integrale von riemannintegablen Funktionen differenzierbar.

Satz 8.29 (Restglied der Taylorformel in Integralform). Sei $f \in C([a, b])$ auf (a, b) $(n+1)$ -mal differenzierbar mit auf $[a, b]$ stetig fortsetzbaren Ableitungen $f', \dots, f^{(n)}$ und $f^{(n+1)} \in \mathcal{R}[a, b]$. Dann gilt für alle $x_0, x \in [a, b]$

$$f(x) - T_{n,x_0}(x) = f(x) - \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k = \int_{x_0}^x \frac{f^{(n+1)}(t)}{n!} (x - t)^n dt.$$

Beweis: Wir definieren $g(t) = \sum_{k=0}^n \frac{f^{(k)}(t)}{k!} (x - t)^k$ für $x, t \in [a, b]$. Dann ist g auf (a, b) differenzierbar mit $g' \in \mathcal{R}[a, b]$ und es gilt wie im Beweis von Satz 7.38

$$f(x) - T_{n,x_0}(x) = g(x) - g(x_0) = \int_{x_0}^x g'(t)dt = \int_{x_0}^x \frac{f^{(n+1)}(t)}{n!} (x - t)^n dt. \quad \text{q.e.d.}$$

Satz 8.30 (Mittelwertsatz der Integralrechnung)*. $\inf_{x \in [a, b]} f(x) \leq \frac{1}{b-a} \int_a^b f dx \leq \sup_{x \in [a, b]} f(x)$ gilt für $f \in \mathcal{R}[a, b]$ und $\frac{1}{b-a} \int_a^b f dx = f(x_0)$ für $f \in C([a, b])$ mit einem $x_0 \in (a, b)$.

Beweis:* Wegen $\inf_{x \in [a, b]} f(x) \leq f \leq \sup_{x \in [a, b]} f(x)$ folgt die erste Aussage aus der Monotonie. Wenn f stetig ist folgt für $F(x) = \int_a^x f(t)dt$ aus dem Mittelwertsatz $f(x_0) = F'(x_0) = \frac{F(b)-F(a)}{b-a} = \frac{1}{b-a} \int_a^b f(x)dx$ mit $x_0 \in (a, b)$. q.e.d.

8.6 Uneigentliches Integral

Wir erweitern das Riemannintegral auf offene und unbeschränkte Intervalle.

Definition 8.31. Eine Funktion f heißt uneigentlich riemannintegabel auf dem offenen (nicht notwendigerweise beschränkten) Intervall $(a, b) \subset \mathbb{R}$, wenn f auf allen kompakten Teilintervallen riemannintegabel ist, und wenn für ein $c \in (a, b)$ beide Grenzwerte $\lim_{x \rightarrow a+} \int_x^c f(t)dt$ und $\lim_{x \rightarrow b-} \int_c^x f(t)dt$ existieren.

Beispiel 8.32. (i) $\int_1^\infty \frac{1}{x^\alpha} dx. \quad \int \frac{1}{x^\alpha} dx = \begin{cases} \frac{-1}{(\alpha-1)x^{\alpha-1}} + C & \text{für } \alpha \neq 1 \\ \ln|x| + C & \text{für } \alpha = 1 \end{cases}$.

Also existiert der Grenzwert $\lim_{x \rightarrow \infty} \int_1^x \frac{1}{t^\alpha} dt$ nur für $\alpha > 1$ mit $\int_1^\infty \frac{1}{x^\alpha} dx = \frac{1}{\alpha-1}$.

$$(ii) \int_0^1 \frac{1}{x^\alpha} dx. \quad \int \frac{1}{x^\alpha} dx = \begin{cases} \frac{-1}{(\alpha-1)x^{\alpha-1}} + C & \text{für } \alpha \neq 1 \\ \ln|x| + C & \text{für } \alpha = 1 \end{cases}.$$

Dann existiert der Grenzwert $\lim_{x \rightarrow 0+} \int_x^1 \frac{1}{x^\alpha} dx$ genau dann, wenn $\alpha < 1$ mit $\int_0^1 \frac{1}{x^\alpha} dx = \frac{1}{1-\alpha}$. Wegen (i) folgt dann, dass $\int_0^\infty \frac{1}{x^\alpha} dx$ für kein α existiert.

$$(iii) \int_{-\infty}^\infty \frac{1}{1+x^2} dx. \quad \int \frac{1}{1+x^2} dx = \arctan(x) + C.$$

Also folgt $\int_{-\infty}^\infty \frac{1}{1+x^2} dx = \pi$ aus $\lim_{x \rightarrow -\infty+} \int_x^0 \frac{1}{1+t^2} dt = \frac{\pi}{2} = \lim_{x \rightarrow \infty-} \int_0^x \frac{1}{1+t^2} dt$.

Verschiedene Kriterien helfen zu entscheiden, ob diese Grenzwerte existieren. Hier einige Kriterien für uneigentliche Integrale $\lim_{x \rightarrow b-} \int_a^x f(t) dt$.

Cauchy Kriterium: $\lim_{x \rightarrow b-} \int_a^x f(t) dt$ existiert genau dann, wenn es zu jedem $\epsilon > 0$ ein $c \in (a, b)$ gibt, so dass für alle $a < c < d < e < b$ gilt $|\int_d^e f(x) dx| < \epsilon$.

Monotoniekriterium: Wenn $f \geq 0$, dann existiert $\lim_{x \rightarrow b-} \int_a^x f(t) dt$ genau dann, wenn $F(x) = \int_a^x f(t) dt$ auf $x \in (a, b)$ beschränkt ist.

Majorantenkriterium: Wenn $f \geq 0$ und $f \leq g$, dann existiert $\lim_{x \rightarrow b-} \int_a^x f(t) dt$, wenn $\lim_{x \rightarrow b-} \int_a^x g(t) dt$ existiert.

Definition 8.33. Eine Funktion f auf einem offenen (unbeschränkten) Intervall heißt absolut riemannintegabel, wenn $|f|$ riemannintegabel ist.

Wegen der Dreiecksungleichung gilt dann $\left| \int_a^b f(x) dx \right| \leq \int_a^b |f(x)| dx$.

Alle absolut riemannintegablen Funktionen sind also auch riemannintegabel.

Satz 8.34 (Integralkriterium für Reihen). Für monoton fallendes $f : [1, \infty) \rightarrow (0, \infty)$ ist $(\sum_{k=1}^n f(k) - \int_1^n f(x) dx)_{n \in \mathbb{N}}$ eine konvergente monoton fallende Folge positiver Zahlen. Für alle $m < n \in \mathbb{N}$ gilt

$$f(n) + \int_m^n f(x) dx \leq \sum_{k=m}^n f(k) \leq f(m) + \int_m^n f(x) dx.$$

Die Reihe $(\sum f(n))_{n \in \mathbb{N}}$ konvergiert genau dann, wenn $\int_1^\infty f(x) dx < \infty$. Dann gilt:

$$\int_1^\infty f(x) dx \leq \sum_{n=1}^\infty f(n) \leq f(1) + \int_1^\infty f(x) dx.$$

Beweis: Für $m < n \in \mathbb{N}$ sei $p_{m,n} \in \mathcal{P}[m, n]$ die Partition $[m, m+1) \cup \dots \cup [n-1, n]$. Dann ist offenbar $\sum_{k=m+1}^n f(k) \leq s(p_{m,n}, f)$ und $S(p_{m,n}, f) = \sum_{k=m}^{n-1} f(k)$. Also gilt

$$\sum_{k=m+1}^n f(k) \leq \int_m^n f(x) dx \leq \sum_{k=m}^{n-1} f(k).$$

Für $n = m+1$ zeigt die linke Ungleichung $f(m+1) - \int_m^{m+1} f(x)dx \leq 0$ dass $(\sum_{k=1}^n f(k) - \int_1^n f(x)dx)_{n \in \mathbb{N}}$ monoton fällt. Durch Addieren von $f(m)$ bzw. $f(n)$ folgt die zweite Aussage. Mit $m = 1$ ist die vorangehende Folge nach unten durch $(f(n))_{n \in \mathbb{N}}$ beschränkt, also positiv, und wegen dem Monotonieprinzip konvergent. Im Grenzwert $n \rightarrow \infty$ folgt die dritte Aussage aus dem Monotonieprinzip und dem Majorantenkriterium. **q.e.d.**

Beispiel 8.35. (i) Der Grenzwert $(\sum_{n=1}^{\infty} \frac{1}{n^s})_{n \in \mathbb{N}}$ heißt Riemannsche ζ -Funktion und ist genau dann konvergent, wenn $\int_1^{\infty} \frac{1}{x^s} dx$ existiert. Also für $s > 1$. Dann gilt

$$\frac{1}{s-1} < \zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} < 1 + \frac{1}{s-1}.$$

(ii) Die Folge $\sum_{k=1}^n \frac{1}{k} - \ln(n)$ ist eine monoton fallende konvergente Folge positiver Zahlen. Der Grenzwert $\gamma = \lim_{n \rightarrow \infty} (\sum_{k=1}^n \frac{1}{k}) - \ln(n)$ wird Eulersche Konstante genannt. Bis heute ist nicht bekannt, ob er rational oder irrational ist.

(iii) Wegen (i) ist die Funktion $\zeta(s) = \sum_{n=1}^{\infty} n^{-s}$ für $s \in (1, \infty)$ konvergent. Die Folge

$$f_n(s) = \sum_{k=1}^n \frac{1}{k^s} - \int_1^n \frac{dx}{x^s}$$

ist wegen dem vorangehenden Satz auf $s \in (0, \infty)$ eine monoton fallende Folge von Funktionen. Weil die folgende Formel auch für $s = 1$ gilt

$$\int_1^n \frac{dx}{x^s} = \frac{n^{1-s} - 1}{1-s} = \frac{\exp((1-s) \ln n) - 1}{1-s} = \sum_{k=1}^{\infty} \frac{(\ln(n))^k (1-s)^{k-1}}{k!},$$

ist das eine Folge von stetigen Funktionen auf $s \in \mathbb{R}$. Wegen der rechten Ungleichung der zweiten Aussage von Satz 8.34, Satz 5.27 und weil $s \mapsto m^{-s}$ für alle $m \in \mathbb{N}$ auf $s \in (0, \infty)$ monoton fallend ist, konvergiert sie für alle $\epsilon > 0$ auf $[\epsilon, \infty)$ gleichmäßig gegen eine stetige Funktion. Auf $s \in (1, \infty)$ ist wegen (i) der Grenzwert gleich

$$\zeta(s) - \int_1^{\infty} \frac{dx}{x^s} = \zeta(s) - \frac{1}{s-1}.$$

Also ist $\lim_{s \rightarrow 1+} (\zeta(s) - \frac{1}{s-1}) = \gamma$, weil für $s = 1$ gilt

$$\lim_{n \rightarrow \infty} \left(\sum_{k=1}^n \frac{1}{k} - \int_1^n \frac{dx}{x} \right) = \lim_{n \rightarrow \infty} \left(\sum_{k=1}^n \frac{1}{k} - \ln(n) \right) = \gamma.$$

(iv) $\left(\sum_{n \in \mathbb{N}} \frac{1}{(n+1) \ln^s(n+1)} \right) < \infty \Leftrightarrow \int_2^{\infty} \frac{1}{x \ln^s(x)} dx = \int_{\ln(2)}^{\infty} \frac{1}{t^s} dt < \infty \Leftrightarrow s > 1.$

(v) Nach Euler ist die Γ -Funktion definiert durch
$$\Gamma(x) = \int_0^\infty e^{-t} t^{x-1} dt.$$

Dieses Integral ist an beiden Grenzen uneigentlich. Wir zerlegen es in $\int_0^\infty = \int_0^1 + \int_1^\infty$. Auf $t \in (0, 1]$ ist der Integrand beschränkt durch $e^{-1} t^{x-1} \leq e^{-t} t^{x-1} \leq t^{x-1}$. Deshalb konvergiert das erste Integral für $x - 1 > -1 \iff x > 0$. Wegen $e^{-t} t^{x-1} = \exp(-t + (x-1) \ln(t))$ und weil für alle $\epsilon > 0$ im Grenzwert $\lim_{t \rightarrow \infty}$ der Ausdruck $-\epsilon t + (x-1) \ln(t)$ negativ ist, konvergiert das zweite Integral für alle $x \in \mathbb{R}$. Also ist $\Gamma(x)$ für alle $x \in (0, \infty)$ definiert. Durch eine partielle Integration erhalten wir

$$\int_\epsilon^R e^{-t} t^x dt = -e^{-t} t^x \Big|_{t=\epsilon}^{t=R} + x \int_\epsilon^R e^{-t} t^{x-1} dt.$$

Im Grenzwert $\epsilon \rightarrow 0$ und $R \rightarrow \infty$ erhalten wir für alle $x \in (0, \infty)$ die Funktionalgleichung $\Gamma(x+1) = x\Gamma(x)$. Mit $\Gamma(1) = \int_0^\infty e^{-t} dt = 1$ folgt induktiv $\Gamma(n) = (n-1)!$.

8.7 Die Konvergenz von Taylorreihen

Satz 8.36 (Abelscher Grenzwertsatz). Wenn die Potenzreihe $(\sum a_n x^n)_{n \in \mathbb{N}_0}$ für ein $x \in \mathbb{K}$ konvergiert, dann konvergiert die Potenzreihenfunktion $t \mapsto \sum_{n \in \mathbb{N}_0} a_n (tx)^n$ auf $t \in [0, 1]$ gleichmäßig gegen eine stetige Funktion von $[0, 1]$ nach $\mathbb{K}$.

Beweis: Indem wir a_n durch $a_n x^n$ ersetzen können wir x weglassen. Zur Abkürzung setzen wir $S_{m,k} = \sum_{n=m+1}^{m+k} a_n$. Wegen dem Cauchy Kriterium für Reihen gibt es für jedes $\epsilon > 0$ ein $N \in \mathbb{N}$ so dass $|S_{m,k}| < \epsilon$ für alle $m \geq N$ und alle $k \in \mathbb{N}$ gilt. Dann folgt

$$\begin{aligned} \sum_{n=m+1}^{m+k} a_n t^n &= S_{m,1} t^{m+1} + (S_{m,2} - S_{m,1}) t^{m+2} + \dots + (S_{m,k} - S_{m,k-1}) t^{m+k} \\ &= S_{m,1} (t^{m+1} - t^{m+2}) + S_{m,2} (t^{m+2} - t^{m+3}) + \dots + S_{m,k-1} (t^{m+k-1} - t^{m+k}) + S_{m,k} t^{m+k}. \end{aligned}$$

Für $t \in [0, 1]$ sind die hinteren Faktoren $t^{m+1} - t^{m+2}, t^{m+2} - t^{m+3}, \dots, t^{m+k-1} - t^{m+k}, t^{m+k}$ alle nicht negativ und ihre Summe gleich $t^{m+1} \leq 1$. Aus $|S_{m,k}| < \epsilon$ folgt also

$$\left| \sum_{n=m+1}^{m+k} a_n t^n \right| < \epsilon (t^{m+1} - t^{m+2} + t^{m+2} - \dots - t^{m+k} + t^{m+k}) = \epsilon t^{m+1} \leq \epsilon \quad \text{für alle } t \in [0, 1].$$

Also konvergiert die Potenzreihe auf $t \in [0, 1]$ gleichmäßig, und damit wegen Satz 5.27 gegen eine stetige Funktion. **q.e.d.**

Beispiel 8.37. (i) Für alle $x_0 \in (0, \infty)$ hat $x \mapsto f(x) = -\sum_{n=1}^{\infty} \frac{(-x)^n}{nx_0^n}$ im Konvergenzbereich $|x| < x_0$ die Ableitung $f'(x) = \sum_{n=0}^{\infty} \frac{(-x)^n}{x_0^{n+1}} = \frac{1}{x+x_0}$. Also stimmt sie mit

$f(x) = \ln(x+x_0) - \ln(x_0)$ überein. Deshalb sind sowohl $\ln$ also auch $\log_a$ auf $(0, \infty)$ reellanalytisch. Insbesondere folgt aus dem Abelschen Grenzwertsatz $\sum_{n=1}^{\infty} \frac{(-1)^n}{n} = -\ln(2)$.

(ii) Für $\alpha \in \mathbb{R}$ und $x_0 \in (0, \infty)$ ($x_0 \in \mathbb{R} \setminus \{0\}$ für $\alpha \in \mathbb{Z} \cap (-\infty, 0)$) hat

$$x \mapsto f(x) = \sum_{n=0}^{\infty} \binom{\alpha}{n} x_0^{\alpha-n} x^n \text{ mit } \binom{\alpha}{n} = \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n(n-1)\cdots 1} \text{ und } \binom{\alpha}{0} = 1$$

wegen dem Quotiententest den Konvergenzradius $R = \frac{1}{\lim_{n \rightarrow \infty} \frac{|\alpha-n|}{|x_0|(n+1)}} = |x_0|$. Die

Ableitung ist $x \mapsto \sum_{n=1}^{\infty} \binom{\alpha}{n} nx_0^{\alpha-n} x^{n-1} = \alpha \sum_{n=0}^{\infty} \binom{\alpha-1}{n} x_0^{\alpha-1-n} x^n$. Wegen

$$\binom{\alpha-1}{n} + \binom{\alpha-1}{n-1} = \frac{(\alpha-1)(\alpha-2)\cdots(\alpha-n+1)}{n!} (\alpha-n+n) = \binom{\alpha}{n}, \quad \binom{\alpha-1}{0} = \binom{\alpha}{0}$$

ist $(x_0+x) \sum_{n=0}^{\infty} \binom{\alpha-1}{n} x_0^{\alpha-1-n} x^n = \sum_{n=0}^{\infty} \binom{\alpha}{n} x_0^{\alpha-n} x^n$. Dann erfüllt f die Differentialgleichung $(x_0+x)f' = \alpha f$ mit $f(0) = x_0^\alpha$. Also verschwindet die Ableitung von

$$g(x) = \frac{f(x)}{(x_0+x)^\alpha} \quad g'(x) = \frac{(x+x_0)f'(x) - \alpha f(x)}{(x+x_0)^{\alpha+1}} = 0 \quad \text{mit } g(0) = 1.$$

Dann folgt aus Satz 7.16 (i), dass $f(x) = (x+x_0)^\alpha$ für alle $|x| < x_0$ gilt. Also sind für $\alpha \in \mathbb{R}$ die Funktionen $x \mapsto x^\alpha$ auf $(0, \infty)$ und für $\alpha \in \mathbb{Z}$ auf $\mathbb{R} \setminus \{0\}$ reellanalytisch.

(iii) Die Ableitungen der Umkehrfunktionen der trigonometrischen Funktionen sind wegen (ii) im Inneren ihrer Definitionsbereiche reellanalytisch. Für alle $|x| < 1$ gilt

$$\begin{aligned} \arcsin'(x) &= \sum_{n=0}^{\infty} \binom{-\frac{1}{2}}{n} (-1)^n x^{2n} & \arccos'(x) &= -\sum_{n=0}^{\infty} \binom{-\frac{1}{2}}{n} (-1)^n x^{2n} \\ \arctan'(x) &= \sum_{n=0}^{\infty} (-1)^n x^{2n} & \operatorname{arccot}'(x) &= -\sum_{n=0}^{\infty} (-1)^n x^{2n}. \end{aligned}$$

Wegen Beispiel 7.4 (v) sind sie selber dann auch reellanalytisch. Für alle $|x| < 1$ gilt

$$\begin{aligned} \arcsin(x) &= \sum_{n=0}^{\infty} \binom{-\frac{1}{2}}{n} \frac{(-1)^n x^{2n+1}}{2n+1} & \arccos(x) &= \frac{\pi}{2} - \sum_{n=0}^{\infty} \binom{-\frac{1}{2}}{n} \frac{(-1)^n x^{2n+1}}{2n+1} \\ \arctan(x) &= \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{2n+1} & \operatorname{arccot}(x) &= \frac{\pi}{2} - \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{2n+1}. \end{aligned}$$

Wegen Beispiel 7.18 gilt $x \geq \ln(1+x)$ für $x > -1$. Dann folgt

$$\begin{aligned} \ln \left((-1)^n \binom{-\frac{1}{2}}{n} \right) &= \ln \left(\frac{1}{2} \frac{3}{4} \cdots \frac{2n-1}{2n} \right) = \ln \left(\left(1 - \frac{1}{2}\right) \left(1 - \frac{1}{4}\right) \cdots \left(1 - \frac{1}{2n}\right) \right) \\ &\leq -\frac{1}{2} \left(1 + \frac{1}{2} + \cdots + \frac{1}{n} \right) \leq -\frac{1}{2} \ln \left(\frac{2}{1} \frac{3}{2} \cdots \frac{n+1}{n} \right) = \ln \left(\frac{1}{\sqrt{n+1}} \right). \end{aligned}$$

Dann konvergieren aber die ersten beiden Potenzreihen auch für $x = \pm 1$ und die letzten beiden wegen der alternierenden Reihe von Leibniz. Wegen dem Abelschen Grenzwertsatz gelten die obigen Gleichungen dann auch für $x = \pm 1$. Insbesondere ist

$$\pi = 4 \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} = 2 \sum_{n=0}^{\infty} \binom{-\frac{1}{2}}{n} \frac{(-1)^n}{2n+1}.$$

Satz 8.38. Sei $(f_n)_{n \in \mathbb{N}}$ eine Folge von Funktionen in $C^1((a, b))$ eines beschränkten Intervalles (a, b) , die für ein $x_0 \in (a, b)$ punktweise konvergiert. Wenn die Folge $(f'_n)_{n \in \mathbb{N}}$ außerdem gleichmäßig gegen g konvergiert, dann konvergiert $(f_n)_{n \in \mathbb{N}}$ gleichmäßig gegen eine Funktion $f \in C^1((a, b))$ und es gilt $f' = g$.

Beweis: Wegen dem Mittelwertsatz gilt für alle $x, x_0 \in (a, b)$ und alle $n, m \in \mathbb{N}$

$$|f_n(x) - f_m(x)| \leq |f_n(x_0) - f_m(x_0)| + |x - x_0| \sup\{|f'_n(y) - f'_m(y)| \mid y \in (a, b)\}.$$

Wegen Satz 5.27 konvergiert $(f_n)_{n \in \mathbb{N}}$ gleichmäßig gegen eine Funktion $f \in C((a, b))$. Wegen dem Mittelwertsatz gibt es für alle $x \neq x_1 \in (a, b)$ eine Folge $(\xi_n)_{n \in \mathbb{N}}$ in (x, x_1) bzw. (x_1, x) , so dass $f_n(x) - f_n(x_1) = (x - x_1)f'_n(\xi_n)$ für alle $n \in \mathbb{N}$ gilt. Wegen dem Auswahlprinzipien von Bolzano-Weierstraß konvergiert eine Teilfolge von $(\xi_n)_{n \in \mathbb{N}}$ gegen $\xi \in [x, x_1]$ bzw. $\xi \in [x_1, x]$. Weil g wegen Satz 5.27 stetig ist und wegen

$$|f'_n(\xi_n) - g(\xi)| \leq |f'_n(\xi_n) - g(\xi_n)| + |g(\xi_n) - g(\xi)|,$$

konvergiert die entsprechende Teilfolge von $(f'_n(\xi_n))_{n \in \mathbb{N}}$ gegen $g(\xi)$. Also konvergiert $(f_n(x))_{n \in \mathbb{N}}$ gegen $f(x) = f(x_1) + (x - x_1)g(\xi)$. Aus der Stetigkeit von g folgt, dass f bei x_1 differenzierbar ist und $g(x_1)$ die Ableitung $f'(x_1)$ ist. **q.e.d.**

Korollar 8.39*. Sei $(\sum f'_n)_{n \in \mathbb{N}}$ eine gleichmäßig konvergente Reihe in $C((a, b))$ auf dem beschränkten Intervall (a, b) und $(\sum f_n(x_0))_{n \in \mathbb{N}}$ konvergent mit $x_0 \in (a, b)$. Dann konvergiert $(\sum f_n)_{n \in \mathbb{N}}$ gleichmäßig gegen $f \in C^1((a, b))$ und $(\sum f'_n)_{n \in \mathbb{N}}$ gegen f' . **q.e.d.**

Satz 8.40 (Satz von Borel). Sei $(a_n)_{n \in \mathbb{N}_0}$ eine beliebige Folge in $\mathbb{R}$. Dann gibt es eine Funktion, deren Taylorreihe bei $x_0 = 0$ gleich $\sum_{n=0}^{\infty} \frac{a_n}{n!} x^n$ ist, und die außerhalb von $(-2, 2)$ verschwindet. D.h. alle Potenzreihen sind Taylorreihen einer solchen Funktion.

Beweis*:

$$\text{Sei } h(x) = \begin{cases} 1 & \text{für } |x| \leq 1 \\ \exp\left(\exp\left(\frac{-1}{|x|-1}\right) \cdot \frac{-1}{2-|x|}\right) & \text{für } 1 < |x| < 2 \\ 0 & \text{für } 2 \leq |x| \end{cases}$$

Dann ist $h \in C^\infty(\mathbb{R})$ eine 'Hutfunktion', die außerhalb von $(-2, 2)$ verschwindet, und die auf $[-1, 1]$ gleich 1 ist. Für alle $n \in \mathbb{N}$ existiert dann eine Konstante $M_n > 0$

$$M_n = \max\{\|h_n\|_\infty, \|h'_n\|_\infty, \dots, \|h_n^{(n-1)}\|_\infty\}$$

mit $h_n : \mathbb{R} \rightarrow \mathbb{R}$, $x \mapsto x^n \cdot h(x)$. Dann sei für eine beliebige reelle Folge $(a_n)_{n \in \mathbb{N}_0}$

$$C_n = |a_n| M_n + 1 \quad \text{und} \quad f_n(x) = \frac{a_n}{n! C_n^n} h_n(C_n \cdot x) \quad \text{für alle } n \in \mathbb{N}_0.$$

Für alle $n, m \in \mathbb{N}_0$ gilt dann $f_n^{(m)}(0) = \begin{cases} 0 & \text{für } m \neq n \\ a_n & \text{für } m = n. \end{cases}$ Außerdem gilt

$$\|f_n^{(m)}\|_\infty = \frac{|a_n| C_n^m}{n! C_n^n} \|h_n^{(m)}\|_\infty \leq \frac{|a_n| M_n C_n^m}{n! C_n^n} < \frac{C_n^{m+1}}{n! C_n^n} \leq \frac{1}{n!} \quad \text{für alle } n > m \in \mathbb{N}_0.$$

Also konvergiert für alle $m \in \mathbb{N}_0$ $(\sum f_n^{(m)})_{n \in \mathbb{N}_0}$ gleichmäßig. Wegen Korollar 8.39 konvergieren also für alle $m \in \mathbb{N}_0$ die Reihen $(\sum f_n)_{n \in \mathbb{N}_0}$, $(\sum f'_n)_{n \in \mathbb{N}_0}$, $\dots$, $(\sum f_n^{(m)})_{n \in \mathbb{N}_0}$ gleichmäßig gegen $f, f', \dots, f^{(m)}$. Also ist der Grenzwert $f = \sum_{n=0}^\infty f_n$ eine Funktion in $C^\infty(\mathbb{R})$ und es gilt $f^{(m)}(0) = a_m$ für alle $m \in \mathbb{N}_0$. **q.e.d.**

Beispiel 8.41. (i)* Die Funktion $x \mapsto \sum_{n=0}^\infty \frac{x^n}{(n+1)!} = \frac{e^x - 1}{x}$ ist auf ganz $\mathbb{R}$ reellanalytisch und hat dort keine Nullstellen. Deshalb definiert $f : \mathbb{R} \rightarrow \mathbb{R}$, $x \mapsto \frac{x}{e^x - 1}$ eine Funktion in $C^\infty(\mathbb{R})$. Die Ableitungen bei $x = 0$ heißen Bernoulli Zahlen $B_0 = f(0) = 1, B_1 = f'(0) = -\frac{1}{2}, \dots$. Dann hat f die Taylorreihe $f(x) = \sum_{n=0}^\infty \frac{B_n x^n}{n!}$. Aus $(e^x - 1)f(x) = x$ folgt

$$\left(\sum_{n=1}^\infty \frac{x^n}{n!}\right) \left(\sum_{n=0}^\infty \frac{B_n x^n}{n!}\right) = \sum_{n=1}^\infty \sum_{k=0}^{n-1} \frac{B_k}{k!(n-k)!} x^n = x.$$

Wegen Beispiel 7.44 (ii) ist f reellanalytisch, und es gilt die folgende Rekursionsformel

$$\frac{1}{(n+1)!} \sum_{k=0}^n \binom{n+1}{k} B_k = 0 \quad \text{also} \quad B_n = -\frac{1}{(n+1)} \sum_{k=0}^{n-1} \binom{n+1}{k} B_k \quad \text{und} \quad B_0 = 1.$$

Aus $f(x) - f(-x) = \frac{x}{e^x - 1} + \frac{x}{e^{-x} - 1} = \frac{x}{e^x - 1} + \frac{x e^x}{1 - e^x} = -x$ folgt $B_{2n+1} = 0$ für alle $n \in \mathbb{N}$.

(ii)* Wegen Beispiel 7.44 (ii) ist $\tan$ reellanalytisch. Er hat bei $x_0 = 0$ die Taylorreihe

$$\begin{aligned}\tan(x) &= \frac{e^{ix} - e^{-ix}}{i(e^{ix} + e^{-ix})} = -i \frac{e^{2ix} + 1 - 2}{e^{2ix} + 1} = \frac{2i(e^{2ix} - 1)}{e^{4ix} - 1} - i = \frac{2i(e^{2ix} + 1) - 4i}{e^{4ix} - 1} - i = \\ &= \frac{1}{x} \left(\frac{2ix}{e^{2ix} - 1} + ix - \frac{4ix}{e^{4ix} - 1} - 2ix \right) = \sum_{n=1}^{\infty} \frac{2^{2n} - 4^{2n}}{(2n)!} (-1)^n B_{2n} x^{2n-1}.\end{aligned}$$

(iii)* Wegen Beispiel 7.44 (ii) ist $\cot$ reellanalytisch. Er hat bei $x_0 = 0$ die Taylorreihe

$$\begin{aligned}x \cot(x) &= ix \frac{e^{ix} + e^{-ix}}{e^{ix} - e^{-ix}} = ix \frac{e^{2ix} + 1}{e^{2ix} - 1} = ix \frac{2 + e^{2ix} - 1}{e^{2ix} - 1} = \frac{2ix}{e^{2ix} - 1} + ix = \\ &= 2ixB_1 + ix + \sum_{n=0}^{\infty} \frac{2^{2n}}{(2n)!} (-1)^n B_{2n} x^{2n} = \sum_{n=0}^{\infty} \frac{2^{2n}}{(2n)!} (-1)^n B_{2n} x^{2n}.\end{aligned}$$

Also folgt $\cot(x) = \frac{1}{x} + \sum_{n=1}^{\infty} \frac{2^{2n}}{(2n)!} (-1)^n B_{2n} x^{2n-1}$

(iv)* Wir betrachten die Reihe $f(x) = \frac{1}{x} + \sum_{n \in \mathbb{N}} \left(\frac{1}{x-n} + \frac{1}{x+n} \right) = \frac{1}{x} + \sum_{n \in \mathbb{N}} \frac{2x}{x^2 - n^2}$. Sie konvergiert für $\epsilon > 0$ auf $x \in B(0, \frac{1}{\epsilon}) \setminus \bigcup_{n \in \mathbb{N}} (B(n, n\epsilon) \cup B(-n, n\epsilon))$ gleichmäßig mit

$$\begin{aligned}f^2(x) &= \frac{1}{x^2} + \sum_{n \in \mathbb{N}} \frac{4}{x^2 - n^2} + \sum_{m, n \in \mathbb{N}} \frac{4x^2}{(x^2 - m^2)(x^2 - n^2)} \\ &= \frac{1}{x^2} + \sum_{n \in \mathbb{N}} \left(\frac{4}{x^2 - n^2} + \frac{4x^2}{(x^2 - n^2)^2} \right) + \sum_{m, n \in \mathbb{N}, n \neq m} \left(\frac{4x^2}{x^2 - n^2} - \frac{4x^2}{x^2 - m^2} \right) \frac{1}{n^2 - m^2} \\ &= \frac{1}{x^2} + \sum_{n \in \mathbb{N}} \left(\frac{4}{x^2 - n^2} + \frac{4x^2}{(x^2 - n^2)^2} \right) + \sum_{n \in \mathbb{N}} \left(\frac{8x^2}{x^2 - n^2} \sum_{m \in \mathbb{N} \setminus \{n\}} \frac{1}{n^2 - m^2} \right) \\ &= \frac{1}{x^2} + \sum_{n \in \mathbb{N}} \left(\frac{4}{x^2 - n^2} + \frac{4x^2}{(x^2 - n^2)^2} - \frac{6x^2 - 6n^2 + 6n^2}{(x^2 - n^2)n^2} \right) = -f'(x) - \alpha^2 \text{ mit } \alpha^2 = \sum_{n \in \mathbb{N}} \frac{6}{n^2}.\end{aligned}$$

Dabei haben wir folgendes benutzt: $2n \sum_{m \in \mathbb{N} \setminus \{n\}} \frac{1}{m^2 - n^2} = \sum_{m=1}^{n-1} \left(\frac{1}{m-n} - \frac{1}{m+n} \right) +$
 $+ \sum_{l=1-n}^n \sum_{k=1}^{\infty} \left(\frac{1}{l+2nk-n} - \frac{1}{l+2nk+n} \right) = - \sum_{m=1}^{n-1} \left(\frac{1}{m} + \frac{1}{m+n} \right) + \sum_{l=1}^{2n} \frac{1}{l} = \frac{1}{n} + \frac{1}{2n},$

und zuletzt Satz 8.38. Durch Integration über dx folgt

$$\operatorname{arccot}\left(\frac{f}{\alpha}\right) = - \int \frac{\frac{f'}{\alpha} dx}{\left(\frac{f}{\alpha}\right)^2 + 1} = -\alpha \int \frac{f' dx}{f^2 + \alpha^2} = \alpha \int dx = \alpha x + C.$$

Weil f Polstellen bei $x \in \mathbb{Z}$ und $\cot$ bei $\pi\mathbb{Z}$ hat folgt $\alpha = \pi$ und $C \in \pi\mathbb{Z}$ und damit

$$f(x) = \pi \cot(\pi x) \quad \cot(x) = \frac{1}{\pi} f\left(\frac{x}{\pi}\right) = \frac{1}{x} + \sum_{n \in \mathbb{N}} \frac{2x}{x^2 - n^2 \pi^2} \quad (\text{Partialbruchzerlegung von } \cot).$$

Durch mehrmaliges Anwenden von Satz 8.38 erhalten wir dann für alle $x \in \mathbb{R} \setminus \mathbb{Z}$

$$\begin{aligned} \cot(x) &= \frac{1}{\pi \frac{x}{\pi}} + \sum_{n \in \mathbb{N}} \left(\frac{1}{x - \pi n} + \frac{1}{x + \pi n} \right). \\ \cot'(x) &= -\frac{1}{x^2} - \sum_{n \in \mathbb{N}} \left(\frac{1}{(x - \pi n)^2} + \frac{1}{(x + \pi n)^2} \right). \\ &\vdots = \vdots \\ \cot^{(k)}(x) &= \frac{(-1)^k k!}{x^{k+1}} + \sum_{n \in \mathbb{N}} \left(\frac{(-1)^k k!}{(x - \pi n)^{k+1}} + \frac{(-1)^k k!}{(x + \pi n)^{k+1}} \right). \end{aligned}$$

Der Vergleich mit der in (iii) berechneten Taylorreihe bei $x_0 = 0$ ergibt

$$\begin{aligned} \zeta(2) &= \sum_{n \in \mathbb{N}} \frac{1}{n^2} = -\frac{\pi^2}{2} \frac{2^2 (-1) B_2}{2!} = \frac{\pi^2}{6}, \quad \zeta(4) = \sum_{n \in \mathbb{N}} \frac{1}{n^4} = -\frac{\pi^4}{2} \frac{2^4 B_4}{4!} = \frac{\pi^4}{90}, \\ \zeta(2k) &= \sum_{n \in \mathbb{N}} \frac{1}{n^{2k}} = -\frac{\pi^{2k}}{2} \frac{2^{2k} (-1)^k B_{2k}}{(2k)!} = \frac{(-1)^{k-1} 2^{2k-1} B_{2k} \pi^{2k}}{(2k)!} \quad \text{für } k \in \mathbb{N}. \end{aligned}$$

(v)* Für eine Nullfolge $(a_n)_{n \in \mathbb{N}}$ konvergiert das Produkt $\prod_{n \in \mathbb{N}} (1 + a_n)$ wegen den Eigenschaften von $\ln$ genau dann, wenn die Reihe $(\sum \ln(1 + a_n))_{n \in \mathbb{N}}$ konvergiert. Wegen Beispiel 8.37 (i) und Satz 7.38 gilt für alle $k \in \mathbb{N}_0$ und für alle x mit $|x| \leq \frac{1}{2}$

$$\left| \ln(1+x) - \sum_{n=1}^k \frac{(-1)^{n-1}}{n} x^n \right| \leq \frac{C_k}{(k+1)!} |x|^{k+1} \quad \text{mit} \quad C_k = \sup_{|x| \leq \frac{1}{2}} \frac{k!}{|1+x|^{k+1}} = 2^{k+1} k!.$$

Also konvergiert das Produkt $\prod_{n \in \mathbb{N}} (1 + a_n)$, wenn für ein $k \in \mathbb{N}_0$ die Reihen $(\sum a_n)_{n \in \mathbb{N}}, \dots, (\sum a_n^k)_{n \in \mathbb{N}}$ und $(\sum |a_n|^{k+1})_{n \in \mathbb{N}}$ konvergieren.

Insbesondere konvergiert das Produkt $f(x) = x \prod_{n \in \mathbb{N}} \left(1 - \frac{x^2}{n^2} \right)$ für alle $x \in \mathbb{K}$ mit

$$\frac{f'(x)}{f(x)} = \frac{1}{x} - \sum_{n \in \mathbb{N}} \frac{\frac{2x}{n^2}}{1 - \frac{x^2}{n^2}} = \frac{1}{x} + \sum_{n \in \mathbb{N}} \frac{2x}{x^2 - n^2} = \frac{1}{x} + \sum_{n \in \mathbb{N}} \left(\frac{1}{x - n} + \frac{1}{x + n} \right) = \pi \cot(\pi x).$$

Also verschwindet die Ableitung von $\frac{\pi f(x)}{\sin(\pi x)}$ und wegen $f'(0) = 1 = \sin'(0)$ folgt

$$f(x) = \frac{\sin(\pi x)}{\pi}, \quad \sin(x) = \pi f\left(\frac{x}{\pi}\right) = x \prod_{n \in \mathbb{N}} \left(1 - \frac{x^2}{n^2 \pi^2} \right) \quad (\text{Produktzerlegung von } \sin).$$

Für $x = \frac{1}{2}$ erhalten wir die sogenannte Formel von Wallis:

$$\frac{2}{1} \cdot \frac{2}{3} \cdot \frac{4}{3} \cdot \frac{4}{5} \cdot \frac{6}{5} \cdot \frac{6}{7} \cdots = \prod_{n=1}^{\infty} \frac{2n \cdot 2n}{(2n-1)(2n+1)} = \prod_{n=1}^{\infty} \left(1 - \frac{1}{4n^2} \right)^{-1} = \frac{\frac{\pi}{2}}{\sin(\frac{\pi}{2})} = \frac{\pi}{2}.$$

Kapitel 9

Metrische Räume und Banachräume

9.1 Metrik und Norm

Definition 9.1. (*Metrik auf einer Menge X*) Eine Metrik (oder Abstandsfunktion) ist eine Abbildung $d : X \times X \rightarrow \mathbb{R}$, $(x, y) \mapsto d(x, y)$ mit drei Eigenschaften

- (i) $d(x, y) \geq 0$ für alle $x, y \in X$ und $d(x, y) = 0 \Leftrightarrow x = y$ (Positivität).
- (ii) $d(x, y) = d(y, x)$ (Symmetrie).
- (iii) $d(x, y) \leq d(x, z) + d(z, y)$ für alle $x, y, z \in X$ (Dreiecksungleichung).

Beispiel 9.2. (i) auf jeder Menge X definiert $d(x, y) = \begin{cases} 0 & \text{für } x = y \\ 1 & \text{für } x \neq y \end{cases}$

die sogenannte diskrete Metrik.

- (ii) Auf $\mathbb{K}$ definiert $d(x, y) = |x - y|$ eine Metrik.
- (iii) Auf jeder nicht leeren Teilmenge $A \subset X$ eines metrischen Raumes (X, d) definiert die Einschränkung von d auf $A \times A \rightarrow \mathbb{R}$ eine Metrik.
- (iv) Auf dem kartesischen Produkt zweier metrischer Räume definiert die Summe beider Metriken eine Metrik. Sie heißt Metrik des kartesischen Produktes.
- (v) Die Einschränkung der Metrik (ii) auf die Vereinigung der inversen der natürlichen Zahlen mit $\{0\}$ definiert eine Metrik auf $\bar{\mathbb{N}} = \mathbb{N} \cup \{\infty\} \simeq \{\frac{1}{n} \mid n \in \mathbb{N}\} \cup \{0\}$:

$$d(n, m) = \frac{|n - m|}{nm} \quad d(\infty, n) = d(n, \infty) = \frac{1}{n} \quad d(\infty, \infty) = 0 \quad \text{für alle } n, m \in \mathbb{N}.$$

Die Menge $V = \mathbb{K}^n$ erfüllt mit $(x_1, \dots, x_n) + (y_1, \dots, y_n) = (x_1 + y_1, \dots, x_n + y_n)$ und $0 = (0, \dots, 0)$ die Axiome A1. Außerdem besitzt sie eine Skalarmultiplikation

$$\mathbb{K} \times V \rightarrow V, \quad (\lambda, (x_1, \dots, x_n)) \mapsto \lambda \cdot (x_1, \dots, x_n) = (\lambda x_1, \dots, \lambda x_n),$$

so dass sie einen Vektorraum über dem Körper $\mathbb{K}$ bildet:

Definition 9.3. Ein $\mathbb{K}$ -Vektorraum V ist eine Menge V zusammen mit Abbildungen:

$$+ : V \times V \rightarrow V, \quad (v, w) \mapsto v + w \quad \cdot : \mathbb{K} \times V \rightarrow V, \quad (\lambda, v) \mapsto \lambda \cdot v,$$

die die Axiome A1 der Addition und für alle $\lambda, \mu \in \mathbb{K}$ und $v, w \in V$ folgendes erfüllt:

$$\begin{aligned} \lambda \cdot (v + w) &= \lambda \cdot v + \lambda \cdot w, & (\lambda + \mu) \cdot v &= \lambda \cdot v + \mu \cdot v, \\ (\lambda \cdot \mu) \cdot v &= \lambda \cdot (\mu \cdot v), & 1 \cdot v &= v. \end{aligned}$$

Definition 9.4. Eine Norm auf einem reellen bzw. komplexen Vektorraum V ist eine Abbildung $\|\cdot\| : V \rightarrow \mathbb{R}, \quad x \mapsto \|x\|$ mit folgenden drei Eigenschaften:

(i) $\|x\| \geq 0$ für alle $x \in V$ und $\|x\| = 0 \iff x = 0$

(ii) $\|\lambda \cdot x\| = |\lambda| \cdot \|x\|$ für $\lambda \in \mathbb{K}$ und $x \in V$

(iii) $\|x + y\| \leq \|x\| + \|y\|$ für alle $x, y \in V$

Satz 9.5. Für jede Norm ist $d : V \times V \rightarrow \mathbb{R}, \quad (x, y) \mapsto d(x, y) = \|x - y\|$ eine Metrik.

Beweis:(i) folgt aus (i) der Definition einer Norm.

(ii) $d(y, x) = \|y - x\| = \|(-1)(x - y)\| = |-1| \|x - y\| = \|x - y\| = d(x, y)$

(iii) $d(x, y) = \|x - y\| = \|x - z + z - y\| \leq \|x - z\| + \|z - y\| = d(x, z) + d(z, y)$. **q.e.d.**

Im folgenden werden wir mit der Vorgabe einer Norm auf einem Vektorraum immer auch die entsprechende induzierte Metrik auf diesem Vektorraum betrachten. Deshalb fassen wir jeden normierten Vektorraum auch als metrischen Raum auf.

Definition 9.6. (Euklidische Norm und Metrik) Auf $\mathbb{R}^n$ definiert

$$\|(x_1, \dots, x_n)\| = \sqrt{|x_1|^2 + \dots + |x_n|^2} = \sqrt{x_1^2 + \dots + x_n^2}$$

die euklidische Norm. Die entsprechende Metrik heißt dann euklidische Metrik.

Die Eigenschaft (iii) heißt dabei Minkowski-Ungleichung:

$$\sqrt{(x_1 + y_1)^2 + \dots + (x_n + y_n)^2} \leq \sqrt{x_1^2 + \dots + x_n^2} + \sqrt{y_1^2 + \dots + y_n^2}.$$

Um diese zu beweisen zeigen wir zuerst die

Cauchy–Schwarzsche Ungleichung 9.7.

$$|x_1 y_1| + \dots + |x_n y_n| \leq \sqrt{x_1^2 + \dots + x_n^2} \sqrt{y_1^2 + \dots + y_n^2}$$

Beweis*: Wegen $|x_i y_i| = |x_i| \cdot |y_i|$, $x_i^2 = |x_i|^2$ und $y_i^2 = |y_i|^2$ genügt es offenbar

$$x_1 y_1 + \dots + x_n y_n \leq \sqrt{x_1^2 + \dots + x_n^2} \sqrt{y_1^2 + \dots + y_n^2}$$

zu zeigen. Das ist äquivalent zu

$$(x_1 y_1 + \dots + x_n y_n)^2 \leq (x_1^2 + \dots + x_n^2)(y_1^2 + \dots + y_n^2).$$

Für $y = (y_1, \dots, y_n) = 0$ ist die Aussage trivial. Für $y \neq 0$ gilt

$$\begin{aligned} \left\| \|y\|x - \frac{x_1 y_1 + \dots + x_n y_n}{\|y\|} y \right\|^2 &= \sum_{i=1}^n \left(\|y\|x_i - \frac{x_1 y_1 + \dots + x_n y_n}{\|y\|} y_i \right)^2 \\ &= \sum_{i=1}^n \left(\|y\|^2 x_i^2 + \frac{(x_1 y_1 + \dots + x_n y_n)^2}{\|y\|^2} y_i^2 - 2(x_1 y_1 + \dots + x_n y_n) x_i y_i \right) \\ &= \|y\|^2 \|x\|^2 + \frac{(x_1 y_1 + \dots + x_n y_n)^2}{\|y\|^2} \|y\|^2 - 2(x_1 y_1 + \dots + x_n y_n)^2 \\ &= \|y\|^2 \|x\|^2 - (x_1 y_1 + \dots + x_n y_n)^2. \end{aligned}$$

Weil diese Ausdrücke nicht negativ sind folgt $(x_1 y_1 + \dots + x_n y_n)^2 \leq \|x\|^2 \|y\|^2$. **q.e.d.**

Beweis der Minkowski Ungleichung:

$$\begin{aligned} (x_1 + y_1)^2 + \dots + (x_n + y_n)^2 &= \|x\|^2 + \|y\|^2 + 2(x_1 y_1 + \dots + x_n y_n) \\ &\leq \|x\|^2 + \|y\|^2 + 2\|x\| \cdot \|y\| \leq (\|x\| + \|y\|)^2. \quad \mathbf{q.e.d.} \end{aligned}$$

Analog definiert $\|(x_1, \dots, x_n)\| = \sqrt{x_1 \bar{x}_1 + \dots + x_n \bar{x}_n}$ eine Norm auf $\mathbb{C}^n$. Identifizieren wir $\mathbb{C}$ mit $\mathbb{R}^2$ (Realteil und Imaginärteil), dann ist $\mathbb{C}^n \simeq (\mathbb{R}^2)^n \simeq \mathbb{R}^{2n}$.

Übungsaufgabe 9.8. Die euklidische Norm auf $\mathbb{R}^{2n}$ induziert durch diese Identifikation auf $\mathbb{C}^n$ die Norm $\|(x_1, \dots, x_n)\| = \sqrt{x_1 \bar{x}_1 + \dots + x_n \bar{x}_n}$.

Definition 9.9. Für $x = (x_1, \dots, x_n) \in \mathbb{K}^n$ und $1 \leq p < \infty$ sei

$$\|x\|_p = (|x_1|^p + \dots + |x_n|^p)^{1/p}$$

Wir hatten in einer Übungsaufgabe gesehen, dass $\|x\|_p$ im Grenzwert $p \rightarrow \infty$ gegen

$$\|x\|_\infty = \max\{|x_1|, \dots, |x_n|\}$$

konvergiert. Das sogenannte hermitesche Skalarprodukt wird definiert als

$$\langle x, y \rangle = x_1 \bar{y}_1 + \dots + x_n \bar{y}_n \in \mathbb{K}, \quad \text{für } x, y \in \mathbb{K}^n.$$

Für $y \in \mathbb{R}^n$ setzen wir hierbei $(\bar{y}_1, \dots, \bar{y}_n) = \bar{y} = y$.

Satz 9.10. (Höldersche Ungleichung) Seien $p \geq 1$ und $q \geq 1$ mit $\frac{1}{p} + \frac{1}{q} = 1$.

Dann gilt $|\langle x, y \rangle| \leq |x_1 y_1| + \dots + |x_n y_n| \leq \|x\|_p \cdot \|y\|_q$ für alle $x, y \in \mathbb{K}^n$.

Für $p = q = 2$ erhalten wir wieder die Cauchy-Schwarzsche Ungleichung

$$|x_1 y_1| + \dots + |x_n y_n| \leq \|x\| \cdot \|y\|.$$

Beweis: Wenn $p = 1$ und $q = \infty$ gilt

$$|x_1 y_1| + \dots + |x_n y_n| \leq (|x_1| + \dots + |x_n|) \max\{|y_1|, \dots, |y_n|\}.$$

Den Fall $p = \infty$ und $q = 1$ erhalten wir durch vertauschen von x und y . Wir können also $1 < p, q < \infty$ annehmen, und dass $\|x\|_p \neq 0 \neq \|y\|_q$ gilt, weil die Ungleichung sonst offensichtlich ist. Dann folgt aus der Youngschen Ungleichung für alle $k = 1, \dots, n$

$$\frac{|x_k y_k|}{\|x\|_p \|y\|_q} = \frac{|x_k|}{\|x\|_p} \frac{|y_k|}{\|y\|_q} \leq \frac{1}{p} \frac{|x_k|^p}{\|x\|_p^p} + \frac{1}{q} \frac{|y_k|^q}{\|y\|_q^q}$$

Nach Summation über $k = 1, \dots, n$ erhalten wir

$$\frac{|x_1 y_1| + \dots + |x_n y_n|}{\|x\|_p \|y\|_q} \leq \frac{1}{p} + \frac{1}{q} = 1. \quad \mathbf{q.e.d.}$$

Satz 9.11. (Minkowski Ungleichung) Sei $p \geq 1$ und $x, y \in \mathbb{K}^n$, dann gilt

$$\|x + y\|_p \leq \|x\|_p + \|y\|_p$$

Korollar 9.12. Für alle $1 \leq p \leq \infty$ und $n \in \mathbb{N}$ ist $\|\cdot\|_p : \mathbb{K}^n \rightarrow \mathbb{R}$ eine Norm. **q.e.d.**

Beweis der Minkowski Ungleichung: Für $p = 1$ oder $p = \infty$ folgt sie aus der Dreiecksungleichung. Sei also $1 < p, q < \infty$ mit $\frac{1}{p} + \frac{1}{q} = 1 \Leftrightarrow p + q = pq \Leftrightarrow p = (p-1)q$.

$$\begin{aligned} |x_1 + y_1|^p + \dots + |x_n + y_n|^p &= |x_1 + y_1| \cdot |x_1 + y_1|^{p-1} + \dots + |x_n + y_n| \cdot |x_n + y_n|^{p-1} \\ &\leq (|x_1| + |y_1|)|x_1 + y_1|^{p-1} + \dots + (|x_n| + |y_n|)|x_n + y_n|^{p-1} \\ &\leq (\|x\|_p + \|y\|_p)(|x_1 + y_1|^{(p-1)q} + \dots + |x_n + y_n|^{(p-1)q})^{1/q} \\ &= (\|x\|_p + \|y\|_p)\|x + y\|_p^{p/q} \end{aligned}$$

Die zweite Zeile wurde ausmultipliziert und jeweils die Höldersche Ungleichung benutzt. Es folgt $\|x + y\|_p^p \leq (\|x\|_p + \|y\|_p)\|x + y\|_p^{p/q}$. Für $\|x + y\|_p = 0$ ist die Aussage trivial.

Aus $\|x + y\|_p \neq 0$ folgt $\|x + y\|_p = \|x + y\|_p^{p(1-\frac{1}{q})} = \|x + y\|_p^{(p-\frac{p}{q})} \leq \|x\|_p + \|y\|_p$. **q.e.d.**

Definition 9.13. (offener Ball, Umgebung, offene Menge) Ein offener Ball in (X, d) mit Zentrum $x \in X$ und Radius $r > 0$ ist die Menge $B(x, r) = \{y \in X \mid d(x, y) < r\}$. Eine Umgebung eines Punktes $x \in X$ ist eine Menge $O \subset X$, die für ein $r > 0$ einen Ball $B(x, r)$ enthält. Eine offene Menge $O \subset X$ ist eine Teilmenge, die eine Umgebung aller ihrer Punkte ist, d.h. für alle $x \in O$ gibt es ein $\epsilon > 0$ mit $B(x, \epsilon) \subset O$.

Beispiel 9.14. In $\mathbb{R}$ besteht der Ball $B(x, r)$ aus $(x - r, x + r)$. Im $\mathbb{R}^n$ besteht der Ball $B(x, r)$ aus allen Punkten, deren euklidischer Abstand zu x kleiner ist als r .

Alle offenen Bälle $B(x, r)$ sind tatsächlich Umgebungen von x : Für $y \in B(x, r)$ ist $d(x, y) < r$. Sei $z \in B(y, r - d(x, y))$. Dann gilt $d(x, z) \leq d(x, y) + d(y, z) < r$, also auch $B(y, r - d(x, y)) \subset B(x, r)$. Deshalb sind die offenen Bälle tatsächlich offen.

Offenbar ist die beliebige Vereinigung von offenen Mengen wieder offen. Für zwei offene Mengen O und O' und $x \in O \cap O'$ gibt es $r > 0$ und $r' > 0$ mit $B(x, r) \subset O$ und $B(x, r') \subset O'$. Also ist $B(x, \min\{r, r'\}) \subset B(x, r) \cap B(x, r') \subset O \cap O'$, und $O \cap O'$ offen. Damit ist auch die Schnittmenge von endlich vielen offenen Mengen wieder offen.

Definition 9.15. (abgeschlossene Mengen, Abschluss) Die Komplemente von offenen Mengen heißen abgeschlossen. Der Abschluss $\bar{A}$ einer Menge A ist die Schnittmenge aller abgeschlossenen Mengen, die A enthalten.

Wegen der Regel von de Morgan, sind beliebige Schnitte und endliche Vereinigungen von abgeschlossenen Mengen wieder abgeschlossen. Deshalb ist eine Menge genau dann abgeschlossen, wenn sie mit ihrem Abschluss übereinstimmt.

Definition 9.16. Zwei Normen $\|\cdot\|_1$ und $\|\cdot\|_2$ heißen äquivalent, wenn es Konstanten $C_1, C_2 > 0$ gibt, so dass für alle $v \in V$ folgendes gilt

$$\|v\|_1 \leq C_1 \|v\|_2 \quad \text{und} \quad \|v\|_2 \leq C_2 \|v\|_1.$$

Diese Relation zwischen Normen ist eine Äquivalenzrelation, denn aus

$$\begin{array}{llll} \|v\|_1 \leq C_1 \|v\|_2 & \|v\|_2 \leq C_2 \|v\|_1 & \|v\|_2 \leq C_3 \|v\|_3 & \|v\|_3 \leq C_4 \|v\|_2 \\ \text{folgt} & \|v\|_1 \leq C_1 C_3 \|v\|_3 & \text{und} & \|v\|_3 \leq C_4 C_2 \|v\|_1. \end{array}$$

Wenn $\|v\|_1 \leq C_1 \|v\|_2$ für alle $v \in V$ gilt, dann ist C_1 wegen der Positivität von $\|\cdot\|_1$ positiv und die Bälle bezüglich der Normen mit den gleichen Indizes erfüllen $B_2(w, r) \subset B_1(w, C_1 r)$ für alle $w \in V$ und $r > 0$. Also sind alle Umgebungen bezüglich $\|\cdot\|_1$ auch Umgebungen bezüglich $\|\cdot\|_2$. Wenn umgekehrt $B_1(0, 1)$ auch eine Umgebung von 0 bezüglich $\|\cdot\|_2$ ist, folgt $\frac{\|v\|_1}{C_1 \|v\|_2 + \epsilon} = \|\frac{v}{C_1 \|v\|_2 + \epsilon}\|_1 < 1$ wegen $\frac{v}{C_1 \|v\|_2 + \epsilon} \in B_2(0, \frac{1}{C_1})$ für $\epsilon > 0$ und $v \in V$ aus $B_2(0, \frac{1}{C_1}) \subset B_1(0, 1)$, und im Grenzwert $\epsilon \downarrow 0$ wieder $\|v\|_1 \leq C_1 \|v\|_2$. Also stimmen die Umgebungen und damit die offenen und abgeschlossenen Mengen zweier Normen auf V genau dann überein, wenn die Normen äquivalent sind.

Beispiel 9.17. Auf den Vektorräumen $\mathbb{K}^n$ haben wir für $1 \leq p \leq \infty$ die Normen

$$\|\cdot\|_p : \mathbb{K}^n \rightarrow \mathbb{R}, \quad v \mapsto \|v\|_p = \begin{cases} \left(\sum_{i=1}^n |v_i|^p \right)^{1/p} & \text{für } p < \infty \\ \max\{|v_1|, \dots, |v_n|\} & \text{für } p = \infty. \end{cases}$$

eingeführt. Sie sind alle äquivalent, weil für $1 \leq p < \infty$ und $v \in \mathbb{K}^n$ folgendes gilt:

$$\|v\|_\infty \leq \|v\|_p \leq (n \|v\|_\infty^p)^{1/p} = n^{1/p} \|v\|_\infty.$$

9.2 Vollständigkeit und Kompaktheit

Zunächst verallgemeinern wir einige Aussagen über Zahlenfolgen auf allgemeine Folgen in metrischen Räumen.

Definition 9.18. *(Folgen und Cauchyfolgen)* Eine Folge $(x_n)_{n \in \mathbb{N}}$ in einem metrischen Raum (X, d) ist eine Abbildung von $\mathbb{N}$ nach X , mit $n \mapsto x_n$.

Eine Folge $(x_n)_{n \in \mathbb{N}}$ in einem metrischen Raum (X, d) konvergiert gegen $x \in X$, wenn es für jedes $\epsilon > 0$ ein $N \in \mathbb{N}$ gibt, so dass $d(x_n, x) < \epsilon$ für alle $n \geq N$ gilt.

Eine Folge $(x_n)_{n \in \mathbb{N}}$ heißt Cauchyfolge, wenn es für jedes $\epsilon > 0$ ein $N \in \mathbb{N}$ gibt, so dass $d(x_n, x_m) < \epsilon$ für alle $n, m \geq N$ gilt.

Offenbar konvergiert eine Folge $(x_n)_{n \in \mathbb{N}}$ genau dann gegen einen Punkt x , wenn jede Umgebung von x alle bis auf endlich viele Folgenglieder enthält. Deshalb hängt der Begriff der Konvergenz nur von der Wahl der offenen Mengen ab.

Ein Punkt x gehört genau dann zu dem Abschluss $\bar{A}$, wenn es keine offene Menge gibt, die x enthält und einen leeren Schnitt mit A hat. Dies ist wiederum äquivalent dazu, dass es für jedes $n \in \mathbb{N}$ ein Element a_n in dem Ball $B(x, \frac{1}{n}) \cap A$ gibt, oder auch dazu, dass es eine Folge in A gibt, die gegen x konvergiert. Damit haben wir gezeigt:

Lemma 9.19. *Der Abschluss einer Teilmenge eines metrischen Raumes besteht aus allen Grenzwerten von Folgen innerhalb der Teilmenge, die in dem metrischen Raum konvergieren. Und eine Teilmenge ist genau dann abgeschlossen, wenn die Grenzwerte von allen konvergenten Folgen in der Teilmenge auch zu der Menge gehören.* **q.e.d.**

Wenn die Folge $(x_n)_{n \in \mathbb{N}}$ gegen x konvergiert, dann gibt es für jedes $\epsilon > 0$ ein $N \in \mathbb{N}$, so dass $d(x_n, x) < \frac{\epsilon}{2}$ und $d(x_m, x) < \frac{\epsilon}{2}$ für alle $n, m \geq N$ gilt. Dann gilt auch $d(x_n, x_m) \leq d(x_n, x) + d(x, x_m) < \epsilon$. Damit haben wir gezeigt:

Satz 9.20. *In einem metrischen Raum sind konvergente Folgen Cauchyfolgen.* **q.e.d.**

Definition 9.21. *Ein metrischer Raum (X, d) heißt vollständig, wenn auch jede Cauchyfolge konvergiert.*

Definition 9.22. *Ein vollständiger normierter Vektorraum heißt Banachraum.*

Wegen dem Vollständigkeitsaxiom sind $\mathbb{R}^n$ und $\mathbb{C}^n$ für alle $n \in \mathbb{N}$ Banachräume. Wegen Lemma 9.19 ist eine Teilmenge eines vollständigen metrischen Raumes genau dann ein vollständiger metrischer Raum, wenn sie abgeschlossen ist.

Definition 9.23. *(kompakt)* Eine Teilmenge der Menge der offenen Teilmengen von (X, d) heißt offene Überdeckung von X , wenn jedes Element von X in mindestens einer der offenen Mengen enthalten ist. Der metrische Raum heißt kompakt, wenn jede offene Überdeckung eine endliche Teilüberdeckung besitzt, d.h. jede Menge von offenen Teilmengen von X , die X überdeckt, enthält eine endliche Teilmenge, die X überdeckt.

Satz 9.24. *Für einen metrischen Raum (X, d) sind folgende Aussagen äquivalent:*

- (i) (X, d) ist kompakt.
- (ii) Jede Folge in (X, d) besitzt eine konvergente Teilfolge.
- (iii) (X, d) ist vollständig und für jedes $\epsilon > 0$ besitzt (X, d) eine endliche Überdeckung mit offenen Bällen vom Radius ϵ .

Beweis: (i) $\Rightarrow$ (ii): Sei $(x_n)_{n \in \mathbb{N}}$ eine Folge ohne Häufungspunkt. Dann sind für alle $n \in \mathbb{N}$ die Mengen $F_n = \{x_m \mid m \geq n\}$ abgeschlossen, weil ihr Abschluss neben den Elementen von F_n nur aus Häufungspunkten von $(x_n)_{n \in \mathbb{N}}$ besteht. Weil $\bigcap_{n \in \mathbb{N}} F_n$ nur Häufungspunkte von $(x_n)_{n \in \mathbb{N}}$ enthält, bilden die Mengen $(X \setminus F_n)_{n \in \mathbb{N}}$ eine offene Überdeckung von X . Offenbar ist der Schnitt von endlich vielen Mengen der Mengen $(F_n)_{n \in \mathbb{N}}$ nicht leer. Also besitzt die Überdeckung $(X \setminus F_n)_{n \in \mathbb{N}}$ keine endliche Teilüberdeckung.

(ii) $\Rightarrow$ (iii): Sei also (X, d) ein metrischer Raum, der (ii) erfüllt. Dann besitzt jede Cauchyfolge $(x_n)_{n \in \mathbb{N}}$ einen Häufungspunkt x . Dann gibt es für jedes $\epsilon > 0$ eine natürliche Zahl $N \in \mathbb{N}$, so dass alle $m, n \geq N$ auch $d(x_n, x_m) < \frac{\epsilon}{2}$ erfüllen. Weil x ein Häufungspunkt ist, gibt es ein $m \geq N$, so dass $d(x_m, x) < \frac{\epsilon}{2}$ gilt. Also gilt

$$d(x_n, x) \leq d(x_n, x_m) + d(x_m, x) < \epsilon \quad \text{für alle } n \geq N.$$

Also konvergiert $(x_n)_{n \in \mathbb{N}}$ gegen x und damit ist (X, d) vollständig. Sei $\epsilon > 0$ so gewählt, dass es keine endliche Überdeckung von X mit Bällen vom Radius ϵ gibt. Wir definieren induktiv eine Folge $(x_m)_{m \in \mathbb{N}}$, so dass $x_{n+1} \in X \setminus \bigcup_{m=1}^n B(x_m, \epsilon)$ für jedes $n \in \mathbb{N}$ gilt. Daraus folgt $d(x_n, x_m) \geq \epsilon$ für alle $n \neq m \in \mathbb{N}$. Also besitzt sie keine Teilfolge, die eine Cauchyfolge ist, und damit auch keinen Häufungspunkt im Widerspruch zu (ii).

(iii) $\Rightarrow$ (i): Wir nehmen an (X, d) erfüllt Bedingung (iii) und $(U_\lambda)_{\lambda \in L}$ sei eine Überdeckung von (X, d) , die keine endliche Teilüberdeckung besitzt. Dann definieren wir induktiv eine Folge $(x_n)_{n \in \mathbb{N}}$, so dass die Bälle $B(x_n, 2^{-n})$ nicht durch endlich viele $(U_\lambda)_{\lambda \in L}$ überdeckt werden und für alle $n \in \mathbb{N}$ die Bälle $B(x_{n+1}, 2^{-(n+1)})$ und $B(x_n, 2^{-n})$ nicht disjunkt sind. Weil nämlich (X, d) durch endlich viele Bälle vom Radius $2^{-(n+1)}$ überdeckt wird, aber $B(x_n, 2^{-n})$ nicht durch endlich viele $(U_\lambda)_{\lambda \in L}$ überdeckt wird, gibt es mindestens einen dieser endlich vielen Bälle vom Radius $2^{-(n+1)}$, der nichtleeren Schnitt mit $B(x_n, 2^{-n})$ hat und nicht durch endlich viele $(U_\lambda)_{\lambda \in L}$ überdeckt wird. Wegen

$$d(x_n, x_{n+1}) < 2^{-n} + 2^{-(n+1)} = \frac{3}{2 \cdot 2^n} \quad \text{und} \quad \sum_{n=m}^{\infty} \frac{3}{2 \cdot 2^n} = \frac{3}{2 \cdot 2^m} \cdot \frac{1}{1 - \frac{1}{2}} = \frac{3}{2^m}$$

ist die Folge $(x_n)_{n \in \mathbb{N}}$ eine Cauchyfolge. Der Grenzwert x erfüllt $d(x_n, x) \leq \frac{3}{2^n}$ für alle $n \in \mathbb{N}$ und gehört für ein $\lambda \in L$ zu der offenen Menge U_λ , die $B(x, \epsilon)$ für ein $\epsilon > 0$ enthält. Für genügend großes m ist dann $\frac{1}{2^{m-2}} \leq \epsilon$, und wegen $d(x_m, x) \leq \frac{3}{2^m}$ auch

$$B(x_m, 2^{-m}) \subset B(x, 2^{-m} + 3 \cdot 2^{-m}) = B(x, 2^{2-m}) \subset B(x, \epsilon).$$

Also wird der Ball $B(x_m, 2^{-m})$ durch U_λ für ein $\lambda \in L$ überdeckt im Widerspruch zu der Annahme, dass $(U_\lambda)_{\lambda \in L}$ keine endliche Teilüberdeckung besitzt. **q.e.d.**

Dieser Satz hat einige wichtige Folgerungen:

Korollar 9.25. (i) *Kompakte Mengen eines metrischen Raums sind abgeschlossen.*

(ii) *Abgeschlossene Teilmengen einer kompakten Menge sind wieder kompakt.*

Beweis:(i) Kompakte Mengen sind wegen (iii) des vorangehenden Satzes vollständig, und stimmen wegen Lemma 9.19 mit ihrem Abschluss überein.

(ii) Abgeschlossene Teilmengen einer kompakten Menge erfüllen wegen Lemma 9.19 wieder die Bedingung (ii) des vorangehenden Satzes. **q.e.d.**

Definition 9.26. *Eine Teilmenge A eines metrischen Raumes (X, d) heißt beschränkt, wenn für ein $x \in X$, die Menge der Abstände $\{d(x, y) | y \in A\}$ beschränkt ist.*

Wegen der Dreiecksungleichung ist diese Bedingung äquivalent dazu, dass für alle $x \in X$ die Mengen $\{d(x, y) | y \in A\}$ beschränkt sind, aber nicht uniform in $x \in X$.

Der zweite Teil des Beweises vom Satz 5.9 zeigt dass alle kompakten Teilmengen eines metrischen Raumes beschränkt sind. Im $\mathbb{K}^n$ konvergiert eine Folge genau dann, wenn für alle $i = 1, \dots, n$ die jeweilige Folge der i -ten Komponenten konvergieren. Deshalb überträgt sich auch der erste Teil des Beweises des Satzes 5.9 und zeigt

Satz 9.27. *In jedem metrischen Raum ist eine kompakte Teilmenge beschränkt.*

In $\mathbb{K}^n$ ist eine Teilmenge bezüglich einer der äquivalenten Normen $\|\cdot\|_p$ mit $p \in [1, \infty]$ genau dann kompakt, wenn sie abgeschlossen und beschränkt ist.

Beispiel 9.28. (i) *Die Intervalle $[a, b]$ sind kompakt.*

(ii) $\mathbb{N}$ aus Beispiel (vi) ist kompakt.

(iii) Sei $(a_n)_{n \in \mathbb{N}}$ konvergent mit Grenzwert a . Dann ist $\{a\} \cup \{a_n | n \in \mathbb{N}\}$ kompakt.

9.3 Stetigkeit

Definition 9.29. *Eine Abbildung $f : X \rightarrow Y$, $x \mapsto f(x)$ von einem metrischen Raum (X, d) in den metrischen Raum (Y, d) heißt stetig in $x \in X$, wenn es für jedes $\epsilon > 0$ ein $\delta > 0$ gibt, so dass alle $y \in B(x, \delta) \subset X$ auch $f(y) \in B(f(x), \epsilon) \subset Y$ erfüllen. Die Abbildung f heißt stetig, wenn sie in allen Punkten von X stetig ist.*

Stetig im Punkt x heißt also, dass alle Punkte, die hinreichend nahe bei x liegen, auf Werte abgebildet werden, die beliebig nahe bei $f(x)$ liegen.

Satz 9.30. *Für eine Abbildung $f : X \rightarrow Y$, $x \mapsto f(x)$ zwischen den metrischen Räumen (X, d) und (Y, d) ist folgendes äquivalent:*

- (i) f ist stetig in x .
- (ii) Das Urbild jeder Umgebung von $f(x)$ ist eine Umgebung von x .
- (iii) Für jede Folge $(x_n)_{n \in \mathbb{N}}$ in (X, d) , die gegen x konvergiert, konvergiert auch die Folge $(f(x_n))_{n \in \mathbb{N}}$ gegen $f(x)$.

Beweis: (i) $\Leftrightarrow$ (ii) Die Umgebungen von x sind gerade die Mengen, die einen δ -Ball um x enthalten. Also ist (ii) äquivalent zu der Aussage, dass das Urbild jedes ϵ -Balles um $f(x)$ einen δ -Ball um x enthält. Diese Aussage ist nur eine Umformulierung von (i).

(ii) $\Leftrightarrow$ (iii) Die Folgen $(x_n)_{n \in \mathbb{N}}$ und $(f(x_n))_{n \in \mathbb{N}}$ konvergieren genau dann gegen x bzw. $f(x)$, wenn jede Umgebung von x bzw. $f(x)$ alle bis auf endlich viele Folgenglieder enthält. Wenn also $(x_n)_{n \in \mathbb{N}}$ gegen x konvergiert und f (ii) erfüllt, dann konvergiert auch $(f(x_n))_{n \in \mathbb{N}}$ gegen $f(x)$. Also folgt aus (ii) auch (iii). Wenn es umgekehrt einen ϵ -Ball von $f(x)$ gibt, dessen Urbild keinen δ -Ball von x enthält, dann gibt es eine Folge $(x_n)_{n \in \mathbb{N}}$ von Punkten $x_n \in B(x, \frac{1}{n})$, so dass die Folge der entsprechenden Werte $f(x_n)$ im Komplement dieses ϵ -Balles von $f(x)$ liegt: $f(x_n) \notin B(f(x), \epsilon)$. Also konvergiert $(x_n)_{n \in \mathbb{N}}$ gegen x aber $f(x_n)$ nicht gegen $f(x)$. **q.e.d.**

Korollar 9.31. Für eine Abbildung $f : X \rightarrow Y$ zwischen den metrischen Räumen (X, d) und (Y, d) ist folgendes äquivalent:

- (i) f ist stetig.
- (ii) Das Bild $(f(x_n))_{n \in \mathbb{N}}$ jeder konvergenten Folge $(x_n)_{n \in \mathbb{N}}$ ist konvergent und es gilt

$$\lim_{n \rightarrow \infty} f(x_n) = f\left(\lim_{n \rightarrow \infty} x_n\right).$$

- (iii) Das Urbild jeder offenen Menge ist offen.
- (iv) Das Urbild jeder abgeschlossenen Menge ist abgeschlossen.

Beweis: Wegen dem vorangehenden Satz sind (i) und (ii) äquivalent. Weil eine Menge genau dann offen ist, wenn sie eine Umgebung von allen ihren Punkten ist, zeigt der vorangehende Satz, dass aus (i) bzw. (ii) auch (iii) folgt. Weil jede Umgebung eines Punktes auch eine offene Umgebung des Punktes enthält, folgt wieder wegen dem vorangehenden Satz aus (iii) auch (i) bzw. (ii). Weil nun die abgeschlossenen Mengen gerade die Komplemente der offenen Mengen sind und das Urbild eines Komplementes gerade gleich dem Komplement des Urbildes ist, ist (iii) zu (iv) äquivalent. **q.e.d.**

Korollar 9.32. Die Komposition zweier stetiger Abbildungen ist stetig. Die analoge punktweise Aussage gilt auch.

Beweis: Für die Verkettung von $g \circ f$ von stetigen Funktionen gilt wegen (ii) im Korollar 9.31 bzw. (iii) in Satz 9.30 für eine konvergente Folge $(x_n)_{n \in \mathbb{N}}$ die Gleichung

$$\lim_{n \rightarrow \infty} g(f(x_n)) = g\left(\lim_{n \rightarrow \infty} f(x_n)\right) = g\left(f\left(\lim_{n \rightarrow \infty} x_n\right)\right).$$

Beispiel 9.33. (i) Auf jedem metrischen Raum ist die identische Abbildung 1_X stetig.

(ii) Die konstante Abbildung, die alle $x \in X$ auf einen Punkt y abbildet ist stetig.

(iii) Wegen der Dreiecksungleichung gilt

$$d(x, y) \leq d(x, u) + d(u, y) \leq d(x, u) + d(u, v) + d(v, y).$$

Also gilt auch $d(x, y) - d(u, v) \leq d(x, u) + d(v, y)$. Durch vertauschen $(x, y) \leftrightarrow (u, v)$ und unter Benutzung der Symmetrie erhalten wir

$$d(u, v) - d(x, y) \leq d(x, u) + d(v, y) \Rightarrow |d(x, y) - d(u, v)| \leq d(x, u) + d(v, y).$$

Mit der Metrik aus dem Beispiel 9.2 (v) auf $X \times X$ ist $d : X \times X \rightarrow \mathbb{R}$ also stetig.

(iv) Auf einem normierten Vektorraum V folgt $\|v\| - \|w\| \leq \|v - w\|$ für $v, w \in V$ aus $\|v\| \leq \|v - w\| + \|w\|$. Das Vertauschen von v und w ergibt $|\|v\| - \|w\|| \leq \|v - w\|$. Also ist $\|\cdot\|$ auf jedem normierten Vektorraum eine stetige reelle Funktion.

(v) Wegen der Dreiecksungleichung sind für jeden normierten Vektorraum V

$$+ : V \times V \rightarrow V, \quad (v, w) \mapsto v + w \quad \cdot : \mathbb{K} \times V \rightarrow V, \quad (\lambda, v) \mapsto \lambda \cdot v$$

stetige Abbildungen. Das gilt auch für die Abbildungen

$$- : \mathbb{K} \rightarrow \mathbb{K}, \quad x \mapsto -x \quad \text{und} \quad {}^{-1} : \mathbb{K} \setminus \{0\} \rightarrow \mathbb{K} \setminus \{0\}, \quad x \mapsto x^{-1}.$$

(vi) Eine Folge $(x_n)_{n \in \mathbb{N}}$ in einem metrischen Raum (X, d) läßt sich genau dann zu einer stetigen Abbildung von $\bar{\mathbb{N}}$ nach X fortsetzen, wenn sie konvergiert. Dann wird ∞ auf $\lim_{n \rightarrow \infty} x_n$ abgebildet.

Korollar 9.34. Das Bild einer kompakten Menge unter einer stetigen Abbildung ist kompakt.

Beweis: Sei $f : X \rightarrow Y$, $x \mapsto f(x)$ eine stetige Abbildung und $A \subset X$ eine kompakte Menge. Dann ist das Urbild einer beliebig offenen Überdeckung von dem Bild

$$f[A] = \{y \in Y \mid \exists x \in A \text{ mit } f(x) = y\}$$

eine offene Überdeckung von A . Diese besitzt, wenn A kompakt ist, eine endliche Teilüberdeckung. Die entsprechende endliche Teilmenge der ursprünglichen offenen Überdeckung von $f[A]$ überdeckt auch $f[A]$. Also ist $f[A]$ kompakt. **q.e.d.**

Korollar 9.35. *Das Bild einer kompakten Menge unter einer stetigen Abbildung ist beschränkt. Das Bild einer solchen reellen Funktion besitzt Minimum und Maximum.*

Beweis: Sei (X, d) kompakt und $f : X \rightarrow Y$, $x \mapsto f(x)$ stetig. Wegen Korollar 9.34 ist $f[X]$ kompakt und wegen Satz 9.27 beschränkt. Wegen Korollar 5.11 besitzen die kompakten Teilmengen von $\mathbb{R}$ sowohl ein Minimum als auch ein Maximum. **q.e.d.**

Eine bijektive stetige Abbildung $f : X \rightarrow Y$ zwischen metrischen Räumen (X, d) und (Y, d) heißt homöomorph oder Homöomorphismus, wenn f^{-1} stetig ist.

Korollar 9.36. *Auf kompaktem X ist jedes stetige bijektive $f : X \rightarrow Y$ homöomorph.*

Beweis: Wegen Korollar 9.34 ist das Bild $f[X] = Y$ kompakt. Wegen Korollar 9.25 ist eine Teilmenge eines kompakten metrischen Raumes genau dann abgeschlossen, wenn sie kompakt ist. Weil das Urbild unter der Umkehrabbildung gleich dem Bild unter f ist, folgt die Aussage aus Korollar 9.34 und Korollar 9.31 (iv). **q.e.d.**

Satz 9.37. *Auf $\mathbb{K}^n$ sind alle Normen paarweise äquivalent.*

Beweis: Es genügt zu zeigen, dass alle Normen äquivalent sind zu $\|\cdot\|_1$. Sei also $\|\cdot\|$ eine beliebige Norm. Sei $e_1, \dots, e_n$ die Basis von $\mathbb{K}^n$, deren i -tes Element nur an der i -ten Stelle eine nicht verschwindende Komponente hat, die dann jeweils gleich Eins ist. Wegen der Dreiecksungleichung gilt dann für jedes $v = (v_1, \dots, v_n) \in \mathbb{K}^n$

$$\|v\| \leq |v_1| \cdot \|e_1\| + \dots + |v_n| \cdot \|e_n\| \leq \max\{\|e_1\|, \dots, \|e_n\|\} \|v\|_1.$$

Aus der Dreiecksungleichung folgt dann für alle $v, w \in \mathbb{K}^n$

$$||v\| - \|w\|| \leq \|v - w\| \leq \max\{\|e_1\|, \dots, \|e_n\|\} \|v - w\|_1.$$

Also ist die reelle Funktion $v \mapsto \|v\|$ stetig bezüglich der Norm $\|\cdot\|_1$, und wegen dem Satz 9.27 von Heine-Borel ist die Teilmenge $\{v \in \mathbb{K}^n \mid \|v\|_1 = 1\}$ mit der von $\|\cdot\|_1$ induzierten Metrik kompakt. Wegen Korollar 9.35 nimmt diese Funktion auf dieser Menge das Minimum C an. Wegen der Positivität von $\|\cdot\|$ gilt $C > 0$. Daraus folgt

$$C\|v\|_1 \leq \left\| \frac{v}{\|v\|_1} \right\| \|v\|_1 = \|v\| \leq \max\{\|e_1\|, \dots, \|e_n\|\} \|v\|_1 \text{ für alle } v \in \mathbb{K}^n \setminus \{0\}. \text{q.e.d.}$$

Definition 9.38. *(Gleichmäßige Stetigkeit, Lipschitzstetigkeit) Eine Abbildung $f : X \rightarrow Y$ zwischen metrischen Räumen heißt gleichmäßig stetig, wenn es für alle $\epsilon > 0$ ein $\delta > 0$ gibt, so dass $d(f(x), f(y)) < \epsilon$ für alle $x, y \in X$ mit $d(x, y) < \delta$ gilt.*

Die Abbildung heißt lipschitzstetig auf A , wenn es eine Konstante $L > 0$ (Lipschitzkonstante) gibt, so dass für alle $x, y \in A$ gilt $d(f(x), f(y)) \leq Ld(x, y)$.

Offenbar ist jede lipschitzstetige Abbildung auch gleichmäßig stetig und jede gleichmäßig stetige Abbildung auch stetig. Es gilt auch folgende Umkehrung:

Satz 9.39. Sei $f : X \rightarrow Y$, $x \mapsto f(x)$ eine stetige Abbildung zwischen metrischen Räumen. Dann ist f auf jeder kompakten Menge A auch gleichmäßig stetig.

Auf kompakten Mengen sind also gleichmäßige und einfache Stetigkeit äquivalent.

Beweis: Sei also $f : X \rightarrow Y$, $x \mapsto f(x)$ stetig und $A \subset X$ kompakt. Dann gibt es für jedes $\epsilon > 0$ und jedes $x \in A$ ein $\delta(x)$, so dass $f(y) \in B(f(x), \frac{\epsilon}{2})$ aus $y \in B(x, 2\delta(x))$ folgt. Wir wählen eine endliche Teilüberdeckung von der offenen Überdeckung $\{B(x, \delta(x)) \mid x \in A\}$ von A . Sei δ das Minimum der Radien dieser endlichen Teilüberdeckung. Dann gibt es für alle $y, z \in A$ mit $d(y, z) < \delta$ einen Ball $B(x, \delta(x))$ der endlichen Teilüberdeckung mit $y \in B(x, \delta(x))$. Dann folgt

$$d(x, z) \leq d(x, y) + d(y, z) < \delta(x) + \delta \leq 2\delta(x) \quad \text{also } z \in B(x, 2\delta(x)).$$

Daraus folgt $d(f(y), f(z)) \leq d(f(x), f(y)) + d(f(x), f(z)) < \epsilon$. **q.e.d.**

Übungsaufgabe 9.40. In dieser Aufgabe konstruieren wir $\mathbb{R}$ als die Vervollständigung von dem angeordneten Körper $\mathbb{Q}$. Sei dazu $\mathfrak{C}$ die Menge der Cauchyfolgen in $\mathbb{Q}$.

Für zwei Cauchyfolgen $(x_n)_{n \in \mathbb{N}}, (\tilde{x}_n)_{n \in \mathbb{N}} \in \mathfrak{C}$ definieren wir die Relation

$$(x_n) \sim (\tilde{x}_n)_{n \in \mathbb{N}} \iff (x_n - \tilde{x}_n)_{n \in \mathbb{N}} \text{ ist Nullfolge in } \mathbb{Q}.$$

- (i) Zeige, dass $\sim$ eine Äquivalenzrelation auf $\mathfrak{C}$ ist.
- (ii) Wir bezeichnen die Menge der Äquivalenzklassen $\mathfrak{C}/\sim$ suggestiv mit $\mathbb{R}$. Zeige, dass die folgendermaßen definierte Addition und Multiplikation

$$[(x_n)_{n \in \mathbb{N}}] + [(y_n)_{n \in \mathbb{N}}] := [(x_n + y_n)_{n \in \mathbb{N}}], \quad [(x_n)_{n \in \mathbb{N}}] \cdot [(y_n)_{n \in \mathbb{N}}] := [(x_n \cdot y_n)_{n \in \mathbb{N}}]$$
 jeweils wohldefiniert sind, d.h. dass erstens die Addition und die Multiplikation von zwei Cauchyfolgen selbst eine Cauchyfolge ist und zweitens, dass diese Verknüpfungen nicht von den jeweiligen Repräsentanten (x_n) und (y_n) der Äquivalenzklassen abhängen. (Tipp: Benutze: Cauchyfolgen sind beschränkt)
- (iii) Zeige, dass $\mathbb{R} = \mathfrak{C}/\sim$ mit diesen Verknüpfungen die Körperaxiome A1–A3 erfüllt. (Tipp: Benutze, dass $\mathbb{Q}$ die Axiome A1–A3 erfüllt.)
- (iv) Zeige, dass folgende Relation auf $\mathbb{R}$ eine wohldefinierte Ordnungsrelation ist, die das Axiom A4 erfüllt:

$$[(x_n)_{n \in \mathbb{N}}] > [(y_n)_{n \in \mathbb{N}}] \iff \exists N \in \mathbb{N} : x_n - y_n \geq \frac{1}{N} \quad \forall n \geq N.$$

- (v) Definiere die Einbettungsabbildung,

$$\Phi : \mathbb{Q} \hookrightarrow \mathbb{R}, \quad q \mapsto [q],$$

wobei für jedes $q \in \mathbb{Q}$, $[q]$ die konstante Folge $q_n = q$ bezeichnet. Diese Abbildung ist offenbar injektiv. Zeige, dass $\mathbb{R}$ archimedisch ist. Hierbei kann ohne Beweis benutzt werden, dass die natürlichen Zahlen in $\mathbb{R}$ das Bild von $\mathbb{N} \subset \mathbb{Q}$ unter Φ sind. Zeige in einem zweiten Schritt, dass das Bild $\Phi[\mathbb{Q}]$ dicht in $\mathbb{R}$ liegt.

- (vi) Zeige, dass $\mathbb{R}$ vollständig ist. (Tipp: Nimm eine Cauchyfolge $(\xi_n)_{n \in \mathbb{N}}$ in $\mathbb{R}$ als gegeben an. Da das Bild $\Phi[\mathbb{Q}]$ dicht in $\mathbb{R}$ liegt, existiert für jedes $n \in \mathbb{N}$ jeweils ein $x_n \in \mathbb{Q}$ mit $\Phi(-\frac{1}{n}) < \Phi(x_n) - \xi_n < \Phi(\frac{1}{n})$. Zeige, dass $(x_n)_{n \in \mathbb{N}}$ eine Cauchyfolge in $\mathbb{Q}$ ist, und $(\xi_n)_{n \in \mathbb{N}}$ in $\mathbb{R}$ gegen das entsprechende $\xi := [(x_n)_{n \in \mathbb{N}}]$ konvergiert.)

9.4 Funktionenräume

In diesem Abschnitt sei (Y, d) ein metrischer Raum, ein normierter Vektorraum oder eine normierte Algebra und später auch (X, d) ein metrischer Raum:

Definition 9.41. Eine normierte Algebra ist ein normierter Vektorraum V mit einer assoziativen und distributiven Multiplikation $\cdot : V \times V \rightarrow V$, die folgendes erfüllt:

$$\begin{aligned} (v + v'') \cdot v' &= v \cdot v' + v'' \cdot v & (\lambda v) \cdot v' &= \lambda(v \cdot v') & \|v \cdot v'\| &\leq \|v\| \cdot \|v'\| \\ v \cdot (v' + v'') &= v \cdot v' + v \cdot v'' & v \cdot (\lambda v') &= \lambda(v \cdot v') & \text{für alle } v, v', v'' \in V, \lambda \in \mathbb{K}. \end{aligned}$$

Wenn V vollständig ist, heißt V Banachalgebra.

Wir betrachten in diesem Abschnitt Mengen von Abbildungen von X nach Y . Wenn Y ein normierter Vektorraum ist, können wir solche Abbildungen punktweise miteinander addieren und mit Elementen von $\mathbb{K}$ multiplizieren, und wenn Y eine Algebra ist, auch punktweise miteinander multiplizieren:

$$\begin{aligned} f + g : X &\rightarrow Y, & x &\mapsto f(x) + g(x), & \lambda f : X &\rightarrow Y, & x &\mapsto \lambda f(x) \\ f \cdot g : X &\rightarrow Y, & x &\mapsto f(x) \cdot g(x). \end{aligned}$$

Die Addition erfüllt die Axiome A1 und mit der Skalarmultiplikation das Distributivgesetz. Dadurch wird die Menge aller Abbildungen in einen Vektorraum Y zu einem Vektorraum, und zu einer Algebra, wenn Y eine Algebra ist. Das Inverse einer Funktion f in einer Algebra mit Eins $\mathbb{1} \in Y$ existiert nur, wenn $f(x)$ für alle $x \in X$ invertierbar ist. Indem wir die Elemente von $\mathbb{K}$ mit den entsprechenden Vielfachen der Eins identifizieren, wird die Skalarmultiplikation zu einem Spezialfall der Multiplikation.

Definition 9.42. Eine Folge von Funktionen $(f_n)_{n \in \mathbb{N}}$ von X nach Y heißt

punktweise konvergent, wenn die Folgen $(f_n(x))_{n \in \mathbb{N}}$ für jedes $x \in X$ konvergieren. Die Grenzwerte definieren wieder eine Funktion $f : X \rightarrow Y$, $x \mapsto \lim_{n \rightarrow \infty} f_n(x)$.

gleichmäßig konvergent, wenn es eine Funktion $f : X \rightarrow Y$, $x \mapsto f(x)$ gibt, und für alle $\epsilon > 0$ ein $N \in \mathbb{N}$, so dass $d(f_n(x), f(x)) < \epsilon$ für $n \geq N$ und $x \in X$ gilt.

Offenbar ist jede gleichmäßig konvergente Folge (f_n) auch punktweise konvergent, aber nicht umgekehrt (siehe Beispiel 5.24).

Definition 9.43. Eine Abbildung $f : X \rightarrow Y$ in einen metrischen Raum Y heißt beschränkt, wenn das Bild $f[X]$ in Y beschränkt ist. $B(X, Y)$, bezeichne die Menge aller beschränkten Abbildungen von X nach Y . Auf $B(X, Y)$ bezeichne d folgende Abbildung:

$$d : B(X, Y) \times B(X, Y) \rightarrow \mathbb{R}, \quad (f, g) \mapsto d(f, g) = \sup\{d(f(x), g(x)) \mid x \in X\}.$$

Wenn Y ein normierter Vektorraum ist, dann bezeichne $\|\cdot\|_\infty$ folgende Abbildung:

$$\|\cdot\|_\infty : B(X, Y) \rightarrow \mathbb{R}, \quad f \mapsto \|f\|_\infty = \sup\{\|f(x)\| \mid x \in X\}.$$

Satz 9.44. (i) Für einen metrischen Raum Y ist d eine Metrik auf $B(X, Y)$.

(ii) Wenn Y ein normierter Vektorraum (Algebra) ist, ist $B(X, Y)$ ein normierter Vektorraum (Algebra) mit Norm $\|\cdot\|_\infty$, die die Metrik aus (i) induziert.

(iii) Wenn Y ein vollständiger metrischer Raum ist, dann auch $(B(X, Y), d)$.

Beweis: (i) und (ii) folgen aus den Eigenschaften der Metrik bzw. Norm $\|\cdot\|$ von Y , und weil wegen der Dreiecksungleichung und wegen $\|v \cdot w\| \leq \|v\| \cdot \|w\|$ die Summe und das Produkt zweier beschränkter Abbildungen wieder beschränkt ist.

(iii) Sei $(f_n)_{n \in \mathbb{N}}$ eine Cauchyfolge in $B(X, Y)$. Für alle $\epsilon > 0$ gibt es ein $N \in \mathbb{N}$ mit

$$d(f_n(x), f_m(x)) \leq d(f_n, f_m) < \frac{\epsilon}{2} \quad \text{für alle } n, m \geq N \text{ und alle } x \in X.$$

Dann sind für alle $x \in X$ die Folgen $(f_n(x))_{n \in \mathbb{N}}$ Cauchyfolgen. Also konvergieren sie punktweise gegen eine Funktion $f : X \rightarrow Y$, $x \mapsto f(x)$. Für obiges ϵ und jedes $x \in X$ gibt es ein $M(x) \in \mathbb{N}$, so dass $d(f_m(x), f(x)) < \frac{\epsilon}{2}$ für alle $m \geq M(x)$ gilt. Damit folgt

$$d(f_n(x), f(x)) \leq d(f_n(x), f_{\max\{N, M(x)\}}(x)) + d(f_{\max\{N, M(x)\}}(x), f(x)) < \epsilon$$

für alle $x \in X$ und $n \geq N$. Also konvergiert $(f_n)_{n \in \mathbb{N}}$ gleichmäßig gegen f . Aus $\sup\{d(f_N(x), f(x)) \mid x \in X\} \leq \epsilon$ und $f_N \in B(X, Y)$ folgt $f \in B(X, Y)$. **q.e.d.**

Definition 9.45. Für metrische Räume X und Y sei $C_b(X, Y)$ sei der Unterraum von $B(X, Y)$ aller stetigen und beschränkten Funktionen von X nach Y .

Satz 9.46. (i) Für metrische Räume X und Y ist $C_b(X, Y)$ abgeschlossen in $B(X, Y)$.

(ii) Wenn Y ein normierter Vektorraum (Algebra) ist, dann auch $C_b(X, Y)$.

(iii) Wenn Y vollständig ist, dann auch $C_b(X, Y)$.

Beweis: Wegen dem vorangehenden Satz und Lemma 9.19 genügt es zu zeigen, dass für jede Folge in $C_b(X, Y)$, die als Folge in $B(X, Y)$ konvergiert, der Grenzwert in $C_b(X, Y)$ liegt. Sei $(f_n)_{n \in \mathbb{N}}$ eine Folge in $C_b(X, Y)$, die in $B(X, Y)$ gegen f konvergiert. Dann

gibt es für jedes $\epsilon > 0$ ein $n \in \mathbb{N}$ mit $d(f_n, f) < \frac{\epsilon}{3}$. Weil f_n stetig bei $x \in X$ ist gibt es ein $\delta > 0$, so dass $d(f_n(x), f_n(y)) < \frac{\epsilon}{3}$ für alle $y \in B(x, \delta)$ gilt. Dann folgt

$$d(f(x), f(y)) \leq d(f(x), f_n(x)) + d(f_n(x), f_n(y)) + d(f_n(y), f(y)) < \epsilon.$$

Also ist f bei $x \in X$ stetig.

q.e.d.

Die gleichmäßige Konvergenz der Folge $(f_n)_{n \in \mathbb{N}}$ ist notwendig (siehe Beispiel 5.24). Wenn Y ein Banachraum ist, dann sind sowohl $B(X, Y)$ als auch $C_b(X, Y)$ Banachräume. Wenn Y eine Banachalgebra ist wie z.B. $\mathbb{K}$, dann sind auch $B(X, Y)$ und $C_b(X, Y)$ Banachalgebren. Der Fall $Y = \mathbb{K}$ wird im folgenden noch öfter vorkommen. Jetzt können wir die Vervollständigungen aller metrischen Räume leicht konstruieren:

Satz 9.47. *Sei X ein metrischer Raum und $x_0 \in X$. Für alle $x \in X$ gehört dann*

$$\begin{aligned} I(x) : X &\rightarrow \mathbb{R} & y &\mapsto d(x, y) - d(x_0, y) \quad \text{zu } C_b(X, \mathbb{R}). \\ \text{Die Abbildung } I : X &\rightarrow C_b(X, \mathbb{R}) & x &\mapsto I(x) \end{aligned}$$

ist eine isometrische Abbildung von X nach $C_b(X, \mathbb{R})$, d.h. es gilt $d(I(x), I(y)) = d(x, y)$ für alle $x, y \in X$. Für jede gleichmäßig stetige Abbildung f von X in einen vollständigen metrischen Raum Y , gibt es eine stetige Abbildung $g : \overline{I[X]} \rightarrow Y$ auf dem Abschluss des Bildes von I in $C_b(X, \mathbb{R})$, so dass f gleich $g \circ I$ ist (vergleiche Übungsaufgabe 9.40).

Beweis: Wegen Beispiel 9.33 (iii) sind die reellen Funktionen $I(x)$ für alle $x \in X$ stetig. Wegen der Dreiecksungleichung gilt für $x, y, z \in X$

$$\begin{aligned} |d(x, z) - d(y, z)| &= \max\{d(x, z) - d(y, z), d(y, z) - d(x, z)\} \leq \\ &\leq \max\{d(x, y) + d(y, z) - d(y, z), d(y, x) + d(x, z) - d(x, z)\} = d(x, y) \quad \text{und} \\ d(x, y) &= d(x, y) - d(y, y) \leq d(I(x), I(y)) = \sup\{|d(x, z) - d(y, z)| \mid z \in X\} \leq d(x, y). \end{aligned}$$

Mit $I(x_0) = 0$ und $y = x_0$ folgt $\|I(x)\|_\infty = d(x, x_0)$ und $I(x) \in C_b(X, \mathbb{R})$. Also ist I eine isometrische Abbildung. Sei $h \in \overline{I[X]} \subset C_b(X, \mathbb{R})$ und $f : X \rightarrow Y$ eine gleichmäßig stetige Abbildung in einen vollständigen metrischen Raum Y . Für jedes $\epsilon > 0$ gibt es ein $\delta > 0$, so dass $d(f(x), f(y)) < \epsilon$ aus $d(x, y) < \delta$ folgt. Also haben für $u \in \overline{I[X]}$ alle Elemente von $\{f(x) \in Y \mid x \in X \text{ mit } d(I(x), u) < \frac{\delta}{2}\}$ paarweise einen Abstand kleiner als ϵ . Deshalb sind $(f(x_n))_{n \in \mathbb{N}}$ für alle Folgen $(x_n)_{n \in \mathbb{N}}$ in X , für die $(I(x_n))_{n \in \mathbb{N}}$ gegen ein u konvergiert, Cauchyfolgen in Y , die alle gegen das gleiche Element von Y konvergieren. Wir definieren $g(u)$ als $\lim_{n \rightarrow \infty} f(x_n)$. Wenn $(I(x_n))_{n \in \mathbb{N}}$ und $(I(y_n))_{n \in \mathbb{N}}$ in $\overline{I[X]}$ gegen u bzw. v konvergieren mit $d(u, v) < \delta$, dann ist $\lim_{n \rightarrow \infty} d(x_n, y_n) = \lim_{n \rightarrow \infty} d(I(x_n), I(y_n)) < \delta$, also $d(g(u), g(v)) = \lim_{n \rightarrow \infty} d(f(x_n), f(y_n)) \leq \epsilon$, und g stetig. **q.e.d.**

Satz 9.48. (Satz von Stone–Weierstraß) *Sei (X, d) ein kompakter metrischer Raum und $A \subset C_b(X, \mathbb{R})$ eine Unteralgebra, die $1 \in C_b(X, \mathbb{R})$ enthält und die Punkte trennt, d.h. für alle $x \neq y \in X$ gibt es $f \in A$ mit $f(x) \neq f(y)$. Dann ist $\bar{A} = C_b(X, \mathbb{R})$.*

Lemma 9.49. Auf $[0, 1] \subset \mathbb{R}$ konvergiert die induktiv definierte Folge von Polynomen $p_{n+1}(x) = p_n(x) + \frac{1}{2}(x - p_n^2(x))$ mit $p_0 = 0$, gleichmäßig gegen die Funktion $x \mapsto \sqrt{x}$.

Beweis : Wir zeigen zunächst mit vollständiger Induktion, dass $0 \leq p_n(x) \leq p_{n+1}(x)$ und $0 \leq p_n^2(x) \leq x$ für $x \in [0, 1]$ und alle $n \in \mathbb{N}_0$ gilt. Beides ist für $n = 0$ offensichtlich.

$$\begin{aligned} x - p_{n+1}^2(x) &= x - p_n^2(x) - p_n(x)(x - p_n^2(x)) - \frac{1}{4}(x - p_n^2(x))^2 \\ &= (x - p_n^2(x)) \left(1 - p_n(x) - \frac{1}{4}(x - p_n^2(x)) \right) \\ &= (x - p_n^2(x)) \left(\left(1 - \frac{p_n(x)}{2} \right)^2 - \frac{x}{4} \right) \end{aligned}$$

Aus der Induktionsvoraussetzung folgt $0 \leq p_n(x) \leq 1$, also $\frac{1}{2} \leq 1 - \frac{p_n(x)}{2} \leq 1$ und $(1 - \frac{p_n(x)}{2})^2 \geq \frac{1}{4} \geq \frac{x}{4}$. Das zeigt die Induktion. Also ist $(p_n(x))_{n \in \mathbb{N}_0}$ für $x \in [0, 1]$ monoton wachsend mit $0 \leq p_n^2(x) \leq x$. Dann gilt $(1 - \frac{p_n(x)}{2})^2 - \frac{x}{4} \leq 1 - \frac{x}{4}$ und deshalb $0 \leq x - p_n^2(x) \leq x \cdot (1 - \frac{x}{4})^n$. Wegen $\frac{1}{1 - \frac{x}{4}} = \sum_{n=0}^{\infty} (\frac{x}{4})^n \geq 1 + \frac{x}{4}$ folgt dann aus der Bernoulli Ungleichung $\frac{1}{(1 - \frac{x}{4})^n} \geq 1 + \frac{nx}{4}$ und $0 \leq x - p_n^2(x) \leq \frac{x}{1 + \frac{nx}{4}} < \frac{4}{n}$. Also konvergiert $(p_n^2(x))_{n \in \mathbb{N}}$ auf $x \in [0, 1]$ gleichmäßig gegen x . Die Funktion $[0, 1] \rightarrow [0, 1]$, $x \mapsto \sqrt{x}$ ist die Umkehrfunktion von $[0, 1] \rightarrow [0, 1]$, $x \mapsto x^2$. Weil die zweite Funktion stetig ist, ist wegen Korollar 5.18 die erste stetig und wegen Satz 5.22 sogar gleichmäßig stetig. Dann konvergiert die Folge $(p_n(x) = \sqrt{p_n^2(x)})_{n \in \mathbb{N}}$ gleichmäßig gegen $\sqrt{x}$. **q.e.d.**

Beweis des Satzes von Stone–Weierstraß: Wegen Lemma 9.49 konvergiert eine Folge $(p_n)_{n \in \mathbb{N}}$ von Polynomen auf $[0, 1]$ gleichmäßig gegen $x \mapsto \sqrt{x}$. Für jedes $f \in A$ konvergiert dann $(p_n(\frac{f^2}{\|f\|_{\infty}^2}))_{n \in \mathbb{N}}$ in $C_b(X, \mathbb{R})$ gegen $\sqrt{(\frac{f}{\|f\|_{\infty}})^2}$. Also gehört $|f| = \|f\|_{\infty} \sqrt{(\frac{f}{\|f\|_{\infty}})^2}$ zum Abschluss $\bar{A}$ von A . Aus Korollar 2.20 folgt mit $\||f| - |g|\|_{\infty} \leq \|f - g\|_{\infty}$ für alle $f, g \in C_b(X, \mathbb{R})$ die Stetigkeit von $f \mapsto |f|$. Wegen der Stetigkeit von $+$ und $\cdot$ auf $C_b(X, \mathbb{R})$ ist $\bar{A}$ eine Algebra mit $|f| \in \bar{A}$ für $f \in \bar{A}$. Für $f, g \in \bar{A}$ gehören

$$\max(f, g) = \frac{1}{2}(f + g + |f - g|) \quad \text{und} \quad \min(f, g) = \frac{1}{2}(f + g - |f - g|)$$

zu $\bar{A}$. Weil A die Punkte von X trennt, gibt es für alle $x \neq y \in X$ und alle $\alpha, \beta \in \mathbb{R}$ ein Element $f \in A$ mit $f(x) = \alpha$ und $f(y) = \beta$. Sei nämlich g eine Funktion mit $g(x) \neq g(y)$. Dann ist $f = \alpha + \frac{\beta - \alpha}{g(y) - g(x)}(g - g(x))$ eine solche Funktion.

Sei jetzt $f \in C_b(X, \mathbb{R})$ eine fest vorgegebene Funktion und $\epsilon > 0$. Dann gibt es für alle $x, y \in X$ eine Funktion $g_{x,y} \in \bar{A}$ die bei x und y mit f übereinstimmt. Dann gibt es ein $\delta_{x,y} > 0$, so dass $g_{x,y}(z) < f(z) + \epsilon$ für alle $z \in B(y, \delta_{x,y})$ gilt. Durch Übergang zu einer endlichen Teilüberdeckung von $\{B(y, \delta_{x,y}) \mid y \in X\}$ und dem Minimum der entsprechenden Funktionen $g_{x,y} \in \bar{A}$ gibt es eine Funktion $g_x \in \bar{A}$, die $g_x(x) = f(x)$

und $g_x < f + \epsilon$ erfüllt. Wegen der Stetigkeit von f und g_x gibt es für alle $x \in X$ ein $\delta_x > 0$, so dass $f(y) - \epsilon < g_x(y)$ für alle $y \in B(x, \delta_x)$ gilt. Durch Übergang zu einer endlichen Teilüberdeckung von $\{B(x, \delta_x) \mid x \in X\}$ und dem Maximum der entsprechenden Funktionen g_x finden wir schließlich eine Funktion g in $\bar{A}$, die $f - \epsilon < g < f + \epsilon$ auf X erfüllt. Weil ϵ beliebig ist folgt $f \in \bar{A}$. **q.e.d.**

Für jeden metrischen Raum X besitzt die Algebra $C_b(X, \mathbb{C})$ folgende komplexe Konjugation, die jedes $f \in C_b(X, \mathbb{C})$ auf $\bar{f} \in C_b(X, \mathbb{C})$ mit $\bar{f}(x) = \overline{f(x)}$ abbildet. Diese Abbildung ist ein Algebromorphismus, d.h. sie ist linear und erhält das Produkt.

Korollar 9.50*: Sei (X, d) ein kompakter metrischer Raum und $A \subset C_b(X, \mathbb{C})$ eine Unteralgebra, die $1 \in C_b(X, \mathbb{C})$ und für jedes $f \in A$ auch die komplex konjugierte Funktion $\bar{f} \in A$ enthält und die Punkte trennt, d.h. für alle $x \neq y \in X$ gibt es $f \in A$ mit $f(x) \neq f(y)$. Dann ist $\bar{A} = C_b(X, \mathbb{C})$.

Beweis*: Jedes $f \in A$ ist die Summe einer reellen Funktion $\frac{1}{2}(f + \bar{f}) \in A$ und des Produktes von i mit einer reellen Funktion $\frac{i}{2}(\bar{f} - f) \in A$. Also folgt die Aussage aus dem Satz von Stone-Weierstraß. **q.e.d.**

Satz 9.51* (Satz von Dini): Auf einem kompakten metrischen Raum (X, d) konvergiert eine monotone Funktionenfolge $(f_n)_{n \in \mathbb{N}}$ von stetigen reellen Funktionen gleichmäßig, wenn sie punktweise gegen eine stetige Funktion f konvergiert.

Beweis*: Sei $(f_n)_{n \in \mathbb{N}}$ eine monoton wachsende Folge in $C_b(X, \mathbb{R})$, die punktweise gegen $f \in C_b(X, \mathbb{R})$ konvergiert. Dann gibt es zu jedem $\epsilon > 0$ und $x \in X$ ein $n(x) \in \mathbb{N}$, so dass $f(x) - f_{n(x)}(x) < \frac{\epsilon}{3}$ gilt. Da $f_{n(x)}$ und f stetig sind gibt es ein $\delta(x)$, so dass

$$|f_{n(x)}(x) - f_{n(x)}(y)| < \frac{\epsilon}{3} \quad \text{und} \quad |f(x) - f(y)| < \frac{\epsilon}{3} \quad \text{für alle } y \in B(x, \delta(x)) \text{ gilt.}$$

Dann gilt dort auch $f(y) - f_{n(x)}(y) < \epsilon$. Wähle eine endliche Teilüberdeckung von $\{B(x, \delta(x)) \mid x \in X\}$. Dann gilt für $m \geq \text{Maximum der entsprechenden } n(x)$

$$f(y) - f_m(y) \leq f(y) - f_{n(x)}(y) < \epsilon$$

auf den Mengen der Teilüberdeckung. Das zeigt die gleichmäßige Konvergenz. **q.e.d.**

Definition 9.52* (relativkompakt): Eine Teilmenge eines metrischen Raumes heißt relativkompakt, wenn der Abschluss kompakt ist.

Lemma 9.53*: Eine Teilmenge A eines metrischen Raumes (X, d) ist genau dann relativkompakt, wenn jede Folge in A eine in X konvergente Teilfolge besitzt.

Beweis*: Wenn A relativkompakt ist, dann besitzt wegen Satz 9.24 jede Folge in A eine konvergente Teilfolge, deren Grenzwert im Abschluss $\bar{A}$ liegt. Hat umgekehrt jede Folge in A eine konvergente Teilfolge, dann gibt es wegen Lemma 9.19 für jede Folge $(x_n)_{n \in \mathbb{N}}$

im Abschluss von A auch eine Folge $(a_n)_{n \in \mathbb{N}}$ in A mit $d(x_n, a_n) < \frac{1}{n}$. Dann konvergiert die jeder konvergenten Teilfolge von $(a_n)_{n \in \mathbb{N}}$ entsprechende Teilfolge von $(x_n)_{n \in \mathbb{N}}$ gegen den gleichen Grenzwert wie die entsprechende Teilfolge von $(a_n)_{n \in \mathbb{N}}$. Wegen Satz 9.24 ist dann der Abschluss von A kompakt. **q.e.d.**

Satz 9.54* (Arzela–Ascoli) Sei X ein kompakter und Y ein vollständiger metrischer Raum. Eine Teilmenge $\mathcal{F} \subset C_b(X, Y)$ ist genau dann relativkompakt, wenn

- (i) für jedes $x \in X$ die Menge $\{f(x) \mid f \in \mathcal{F}\}$ relativkompakt ist, und
- (ii) für jedes $x \in X$ die Menge $\mathcal{F}$ gleichgradig stetig ist in x , d.h. für jedes $x \in X$ und jedes $\epsilon > 0$ gibt es ein $\delta > 0$ mit $f(x') \in B(f(x), \epsilon)$ für alle $x' \in B(x, \delta)$, $f \in \mathcal{F}$.

Beweis*: Zunächst zeigen wir, dass wenn die Menge $\mathcal{F}$ die Bedingungen (i)–(ii) erfüllt, jede Folge $(f_n)_{n \in \mathbb{N}}$ in $\mathcal{F}$ eine in $C_b(X, Y)$ konvergente Teilfolge $(g_n)_{n \in \mathbb{N}}$ besitzt. Dafür zeigen wir zuerst, dass $\mathcal{F}$ auf X sogar gleichmäßig gleichgradig stetig ist. Für jedes $\epsilon > 0$ und jedes $y \in X$ gibt es wegen (ii) ein $\delta_y > 0$, so dass $d(f(x), f(y)) < \frac{\epsilon}{2}$ für alle $f \in \mathcal{F}$ aus $d(x, y) < 2\delta_y$ folgt. Wegen der Kompaktheit von X hat die Überdeckung $\{B(y, \delta_y) \mid y \in X\}$ eine endliche Teilüberdeckung $X = B(y_1, \delta_1) \cup \dots \cup B(y_N, \delta_N)$. Sei δ das Minimum von $\delta_1, \dots, \delta_N$. Dann enthält für alle Paare $x, x' \in X$ mit $d(x, x') < \delta$ einer der Bälle $B(y_1, \delta_1), \dots, B(y_N, \delta_N)$ den einen Punkt x . Damit sind beide in einem der Bälle $B(y_1, 2\delta_1), \dots, B(y_N, 2\delta_N)$ enthalten. Daraus folgt $d(f(x), f(x')) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$ für alle $f \in \mathcal{F}$. Also ist $\mathcal{F}$ auf ganz X gleichmäßig gleichgradig stetig.

Für jedes $n \in \mathbb{N}$ wird X durch endlich viele Bälle mit Radius $\frac{1}{n}$ überdeckt. Sei $(x_l)_{l \in \mathbb{N}}$ eine Abzählung der Zentren aller dieser Bälle, die als Folge in X dicht liegt. Wegen (i) ist $A_l = \overline{\{f_n(x_l) \mid n \in \mathbb{N}\}}$ für alle $l \in \mathbb{N}$ eine kompakte Teilmenge von Y . Wir definieren induktiv eine Teilfolge von $(g_n)_{n \in \mathbb{N}}$ von $(f_n)_{n \in \mathbb{N}}$ und eine Folge $(a_l)_{l \in \mathbb{N}}$ in Y , so dass $d(g_n(x_l), a_l) < \frac{1}{n}$ für alle $l \in \mathbb{N}$ und alle $n \geq l$ gilt. Dafür wählen wir zunächst einen Häufungspunkt a_1 von $(f_n(x_1))_{n \in \mathbb{N}}$ und eine Teilfolge $(g_n)_{n \in \mathbb{N}}$ von $(f_n)_{n \in \mathbb{N}}$, so dass $d(g_n(x_1), a_1) \leq \frac{1}{n}$ für alle $n \in \mathbb{N}$ gilt. Induktiv wählen wir für jedes $L \in \mathbb{N} \setminus \{1\}$ einen Häufungspunkt a_L von $(g_n(x_L))_{n \in \mathbb{N}}$ und ersetzen alle Folgenglieder von $(g_n)_{n \in \mathbb{N}}$ mit Indizes $\geq L$ durch eine Teilfolge von $(g_n)_{n \geq L}$, so dass $d(g_n(x_L), a_L) < \frac{1}{n}$ für alle $n \geq L$ gilt. Dann gilt $d(g_n(x_l), a_l) < \frac{1}{n}$ für alle $l = 1, \dots, L$ und $n \geq l$.

Für ein $\epsilon > 0$ wählen wir ein $\delta > 0$, so dass $d(g_n(x), g_n(x')) < \frac{\epsilon}{3}$ für alle $n \in \mathbb{N}$ und alle $x, x' \in X$ mit $d(x, x') < \delta$ gilt. Die Überdeckung $(B(x_l, \delta))_{l \in \mathbb{N}}$ von X besitzt eine endliche Teilüberdeckung. Also gibt es ein $M \in \mathbb{N}$, so dass an den Zentren der Bälle der Teilüberdeckung $d(g_m(x_l), g_n(x_l)) < \frac{\epsilon}{3}$ für alle $m, n \geq M$ gilt. Dann folgt

$$d(g_m(x), g_n(x)) \leq d(g_m(x), g_m(x_l)) + d(g_m(x_l), g_n(x_l)) + d(g_n(x_l), g_n(x)) < \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon$$

für alle $x \in X$ und alle $m, n \geq M$. Also ist $(g_n)_{n \in \mathbb{N}}$ in $C_b(X, Y)$ eine Cauchyfolge und konvergiert wegen Satz 9.46 (iii) in $C_b(X, Y)$. Wegen Lemma 9.53 ist $\mathcal{F}$ relativkompakt.

Wenn umgekehrt $\mathcal{F}$ relativkompakt ist, dann besitzt wegen Lemma 9.53 mit jeder Folge in $\mathcal{F}$ für jedes $x \in X$ auch die Folge der entsprechenden Funktionswerte eine konvergente Teilfolge. Also erfüllt $\mathcal{F}$ die Bedingung (i).

Außerdem gibt es für jedes $x \in X$ und $\epsilon > 0$ endlich viele $f_1, \dots, f_k$ im Abschluss von $\mathcal{F}$, so dass $B(f_1, \frac{\epsilon}{3}) \cup \dots \cup B(f_k, \frac{\epsilon}{3})$ den Abschluss von $\mathcal{F}$ überdeckt. Weil $f_1, \dots, f_k$ stetig sind, gibt es $\delta_1, \dots, \delta_k > 0$, so dass $f_i(x') \in B(f_i(x), \frac{\epsilon}{3})$ aus $x' \in B(x, \delta_i)$ für $i = 1, \dots, k$ folgt. Für alle $x' \in B(x, \min\{\delta_1, \dots, \delta_k\})$ und $f \in \mathcal{F}$ gilt für ein f_i

$$d(f(x'), f(x)) \leq d(f(x'), f_i(x')) + d(f_i(x'), f_i(x)) + d(f_i(x), f(x)) < \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3}. \quad \mathbf{q.e.d.}$$

9.5 Lineare Operatoren

Die Ableitung einer Funktion von mehreren Veränderlichen ist eine lineare Abbildung. Zur Vorbereitung der Differentialrechnung von Funktionen mehrerer Veränderlicher behandeln wir in diesem Abschnitt solche linearen Abbildungen zwischen normierten Vektorräumen. Dabei betrachten wir wieder Vektorräume über dem Körper $\mathbb{K}$.

Definition 9.55. Eine Abbildung $A : V \rightarrow W$ von einem Vektorraum V in einen Vektorraum W heißt *linear*, wenn für alle $v, w \in V$ und $\lambda \in \mathbb{K}$ gilt

$$A(v + w) = Av + Aw \quad \text{und} \quad A(\lambda v) = \lambda Av.$$

Satz 9.56. Seien V und W normierte Vektorräume und $A : V \rightarrow W$ eine lineare Abbildung. Dann ist folgendes äquivalent:

- (i) A ist stetig in 0.
- (ii) A ist stetig.
- (iii) A ist gleichmäßig stetig.
- (iv) Es gibt ein $C > 0$, so dass für alle $v \in V$ gilt $\|Av\| \leq C\|v\|$.
- (v) A ist auf $B(0, 1)$ beschränkt, d.h. $\|Av\| \leq C$ für alle $\|v\| < 1$ mit $0 < C < \infty$.

Beweis: (i) $\Rightarrow$ (v): Wenn A in 0 stetig ist, dann enthält das Urbild jedes Balles $B(0, \epsilon) \subset W$, also auch von $B(0, 1)$, einen Ball $B(0, \delta) \subset V$. Damit folgt $\|Av\| < 1$ aus $\|v\| < \delta$. Aus der Linearität folgt dann $\|Av\| = \frac{1}{\delta} \|A(\delta v)\| < \frac{1}{\delta}$ aus $\|v\| < 1$. Also ist (v) erfüllt.

(v) $\Rightarrow$ (iv): Wegen der Linearität folgt aus (v), dass für alle $v \in V \setminus \{0\}$ gilt

$$\|Av\| = \left\| A \left(2\|v\| \cdot \frac{v}{2\|v\|} \right) \right\| = \left\| 2\|v\| \cdot A \left(\frac{v}{2\|v\|} \right) \right\| = 2\|v\| \cdot \left\| A \left(\frac{v}{2\|v\|} \right) \right\| \leq 2\|v\|C.$$

(iv) $\Rightarrow$ (iii): Für $v, w \in V$ folgt $\|Av - Aw\| = \|A(v - w)\| \leq C\|v - w\|$ aus (iv). Also ist A sogar Lipschitzstetig mit Lipschitzkonstante C . Dann gilt auch (iii).

(iii) $\Rightarrow$ (ii) und (ii) $\Rightarrow$ (i): Sind offensichtlich.

q.e.d.

Satz 9.57. Jede lineare Abbildung A von $\mathbb{K}^n$ in einen normierten Vektorraum ist stetig.

Beweis: Wir benutzen wieder die Basis $e_1, \dots, e_n$ von $\mathbb{K}^n$. Dann gilt für alle $v \in \mathbb{K}^n$

$$\|Av\| \leq |v_1| \cdot \|Ae_1\| + \dots + |v_n| \cdot \|Ae_n\| \leq \|v\|_1 \max\{\|Ae_1\|, \dots, \|Ae_n\|\}.$$

Also folgt die Aussage aus Satz 9.37 und Satz 9.56.

q.e.d.

Definition 9.58. Seien V, W normierte Vektorräume. Dann sei $\mathcal{L}(V, W)$ die Menge aller linearen stetigen Abbildungen von V nach W zusammen mit den Abbildungen:

$$\begin{aligned} + : \mathcal{L}(V, W) \times \mathcal{L}(V, W) &\rightarrow \mathcal{L}(V, W), & (A, B) &\mapsto A + B : V \rightarrow W, & v &\mapsto Av + Bv \\ \cdot : \mathbb{K} \times \mathcal{L}(V, W) &\rightarrow \mathcal{L}(V, W), & (\lambda, A) &\mapsto \lambda A : V \rightarrow W, & v &\mapsto \lambda Av \\ \|\cdot\| : \mathcal{L}(V, W) &\rightarrow \mathbb{R}, & A &\mapsto \|A\| = \sup\{\|Av\| \mid v \in \overline{B(0, 1)}\} = \sup\{\|Av\| \mid v \in \overline{B(0, 1)}\}. \end{aligned}$$

Für $w \in \overline{B(0, 1)}$ gilt nämlich $\|Aw\| = \lim_{\lambda \uparrow 1} \lambda \|Aw\| = \lim_{\lambda \uparrow 1} \|A\lambda w\| \leq \sup\{\|Av\| \mid v \in \overline{B(0, 1)}\}$.

Satz 9.59. $\mathcal{L}(V, W)$ ist ein normierter Untervektorraum von $C_b(\overline{B(0, 1)}, W)$.

Beweis: Wegen Satz 9.56 $\mathcal{L}(V, W)$ ist die Menge aller linearen Abbildungen $A : V \rightarrow W$, deren Einschränkungen $A|_{\overline{B(0, 1)}}$ in $C_b(\overline{B(0, 1)}, W)$ liegen. Aus der Linearität zweier solcher Abbildungen A und B folgt die Linearität von $A + B$ und $\lambda \cdot A$ für $\lambda \in \mathbb{K}$. Für jeden linearen Operator $A \in \mathcal{L}(V, W)$ und $v \in V \setminus \{0\}$ gilt $Av = \|v\| \cdot A(\frac{v}{\|v\|})$. Also ist A durch seine Werte auf $\overline{B(0, 1)}$ eindeutig bestimmt und die Norm von $\mathcal{L}(V, W)$ ist einfach die Supremumsnorm der stetigen Abbildung von $\overline{B(0, 1)}$ nach W . Also ist $\mathcal{L}(V, W)$ ein normierter Untervektorraum von $C_b(\overline{B(0, 1)}, W)$. **q.e.d.**

Offenbar gilt $\|A\| = \sup\{\frac{\|Av\|}{\|v\|} \mid v \in V \setminus \{0\}\}$ für $A \in \mathcal{L}(V, W)$ und $\|Av\| \leq \|A\| \cdot \|v\|$ für alle $v \in V$. Aus der Konvergenz einer Folge $(A_n)_{n \in \mathbb{N}}$ in $\mathcal{L}(V, W)$ folgt also die gleichmäßige Konvergenz auf $\overline{B(0, 1)}$ und die punktweise Konvergenz auf V . Für $V = \mathbb{K}^n$ ist der Abschluss der Einheitskugel $\overline{B(0, 1)} = \{v \in \mathbb{K}^n \mid \|v\| \leq 1\}$ kompakt. Deshalb gibt es also für jedes $A \in \mathcal{L}(\mathbb{K}^n, W)$ ein $v \in \mathbb{K}^n \setminus \{0\}$ mit $\|A\| = \|A \frac{v}{\|v\|}\| = \frac{\|Av\|}{\|v\|}$.

Satz 9.60. Seien V ein normierter Vektorraum und W ein Banachraum. Dann ist $\mathcal{L}(V, W)$ ein Banachraum.

Beweis: Wegen Satz 9.44 (ii)-(iii) und Satz 9.59 müssen wir nur zeigen, dass die Grenzwerte $A \in C_b(\overline{B(0, 1)}, W)$ von Cauchyfolgen $(A_n)_{n \in \mathbb{N}}$ in $\mathcal{L}(V, W)$ linear sind. Für jedes $v \in V$ ist $(A_n v)_{n \in \mathbb{N}}$ wegen $\|(A_n - A_m)v\| \leq \|A_n - A_m\| \cdot \|v\|$ eine Cauchyfolge in W . Also setzt sich A als punktwiser Grenzwert zu folgender Abbildung fort:

$$A : V \rightarrow W, \quad v \mapsto Av = \lim_{n \rightarrow \infty} A_n v \quad \text{für alle } v \in V.$$

Wir müssen nur noch zeigen, dass A linear ist. Aus der Linearität von A_n folgt

$$\begin{aligned} \|A(v+w) - (Av+Aw)\| &\leq \\ &\leq \|(A-A_n)(v+w) - (A-A_n)v - (A-A_n)w)\| + \|A_n(v+w) - (A_nv + A_nw)\| \\ &\leq \|(A-A_n)(v+w)\| + \|(A-A_n)v\| + \|(A-A_n)w\|, \text{ und} \\ \| \lambda Av - A(\lambda v) \| &\leq \| \lambda(A-A_n)v - (A-A_n)(\lambda v) \| + \| \lambda A_nv - A_n(\lambda v) \| \\ &\leq |\lambda| \|(A-A_n)v\| + \|(A-A_n)(\lambda v)\|. \end{aligned}$$

für alle $n \in \mathbb{N}$ und alle $v, w \in V$. Im Grenzwert $n \rightarrow \infty$ konvergieren die rechten Seiten für alle $v, w \in V$ und $\lambda \in \mathbb{K}$ punktweise gegen Null, so dass A linear ist. **q.e.d.**

Satz 9.61. Seien U, V und W normierte Vektorräume und $A \in \mathcal{L}(U, V)$ und $B \in \mathcal{L}(V, W)$, dann ist $B \circ A \in \mathcal{L}(U, W)$ und es gilt $\|B \circ A\| \leq \|B\| \cdot \|A\|$. Insbesondere ist die Abbildung $\circ : \mathcal{L}(U, V) \times \mathcal{L}(V, W) \rightarrow \mathcal{L}(U, W)$, $(A, B) \mapsto B \circ A$ stetig.

Beweis: Für alle $u \in U$ gilt $\|(B \circ A)u\| = \|B(A(u))\| \leq \|B\| \cdot \|Au\| \leq \|B\| \cdot \|A\| \cdot \|u\|$. Also folgt die Ungleichung $\|B \circ A\| \leq \|B\| \cdot \|A\|$ aus der Definition 9.58. Für zwei normierte Vektorräume V, W mit Normen $\|\cdot\|_V$ und $\|\cdot\|_W$ ist

$$\|\cdot\|_{V \times W} : V \times W \rightarrow \mathbb{R}, \quad (v, w) \mapsto \|v\|_V + \|w\|_W$$

eine Norm auf $V \times W$ und induziert die Metrik des kartesischen Produktes der metrischen Räume V und W . Für $(A, B), (A', B') \in \mathcal{L}(U, V) \times \mathcal{L}(V, W)$ gilt dann

$$\begin{aligned} \|B \circ A - B' \circ A'\| &= \|B \circ A - B \circ A' + B \circ A' - B' \circ A'\| \\ &= \|B \circ (A - A') + (B - B') \circ A'\| \\ &\leq \|B\| \cdot \|A - A'\| + \|B - B'\| \cdot \|A'\| \\ &\leq (\|A - A'\| + \|B - B'\|)(\|B\| + \|A'\|) \\ &\leq (\|A - A'\| + \|B - B'\|)(\|B\| + \|A\| + \|A' - A\|) \\ &\leq (\|(A, B) - (A', B')\|)(\|B\| + \|A\| + \|(A, B) - (A', B')\|). \end{aligned}$$

Also ist diese Abbildung im Punkt $(A, B) \in \mathcal{L}(U, V) \times \mathcal{L}(V, W)$ stetig. **q.e.d.**

Wir bezeichnen die Komposition $B \circ A$ von linearen Operatoren auch mit BA .

Definition 9.62. Auf einem normierten Vektorraum V ist $\mathcal{L}(V) = \mathcal{L}(V, V)$ mit

$$\circ : \mathcal{L}(V) \times \mathcal{L}(V) \rightarrow \mathcal{L}(V), \quad (A, B) \mapsto AB \quad \text{und} \quad \|\cdot\| : \mathcal{L}(V) \rightarrow \mathbb{R}, \quad A \mapsto \|A\|$$

eine normierte Algebra mit Eins $\mathbb{1}_V$ und eine Banachalgebra für einen Banachraum V .

Satz 9.63. (Neumannsche Reihe) Sei $\mathcal{A}$ eine Banachalgebra mit Eins $\mathbb{1}$ und $A \in \mathcal{A}$ ein Operator mit $\|A\| < 1$. Dann ist $\mathbb{1} - A$ invertierbar und es gilt $(\mathbb{1} - A)^{-1} = \sum_{n=0}^{\infty} A^n$.

Beweis: Wegen $\|A^n\| \leq \|A\|^n$ ist $(\sum A^n)_{n \in \mathbb{N}_0}$ für $\|A\| < 1$ eine Cauchyfolge mit

$$\left\| \sum_{n=M}^N A^n \right\| \leq \sum_{n=M}^N \|A\|^n = \frac{\|A\|^M - \|A\|^{N+1}}{1 - \|A\|} \leq \frac{\|A\|^M}{1 - \|A\|} \text{ für } 0 \leq M \leq N.$$

Also konvergiert diese Reihe gegen ein $B \in \mathcal{A}$. Wie im Satz 9.61 ist wegen $\|B \cdot A\| \leq \|B\| \cdot \|A\|$ die Multiplikation $\cdot : \mathcal{A} \times \mathcal{A} \rightarrow \mathcal{A}$ stetig und es gilt

$$(\mathbb{1} - A)B = \sum_{n=0}^{\infty} A^n - \sum_{n=1}^{\infty} A^n = \mathbb{1} \quad \text{und} \quad B(\mathbb{1} - A) = \sum_{n=0}^{\infty} A^n - \sum_{n=1}^{\infty} A^n = \mathbb{1}.$$

Also ist $\sum_{n=0}^{\infty} A^n$ das Inverse von $\mathbb{1} - A$ mit $\|(\mathbb{1} - A)^{-1}\| \leq \frac{1}{1 - \|A\|}$. **q.e.d.**

Jede Potenzreihenfunktion $f(x) = \sum_{n \in \mathbb{N}_0} a_n x^n$ mit Konvergenzradius $R > 0$ definiert also auf jeder Banachalgebra $\mathcal{A}$ eine Abbildung

$$f : \{A \in \mathcal{A} \mid \|A\| < R\} \rightarrow \mathcal{A}, \quad A \mapsto f(A) = \sum_{n=0}^{\infty} a_n A^n \text{ mit } \|f(A)\| \leq \sum_{n=0}^{\infty} |a_n| \|A\|^n < \infty.$$

Viele der Aussagen, die wir für Potenzreihenfunktionen auf $\mathbb{K}$ gezeigt haben, lassen sich auf Potenzreihenfunktionen auf Banachalgebren $\mathcal{A}$ ausdehnen. Aber weil im allgemeinen $AB \neq BA$ für $A, B \in \mathcal{A}$, gilt im allgemeinen auch $\exp(A)\exp(B) \neq \exp(A+B)$.

Definition 9.64. Eine Derivation einer Algebra $\mathcal{A}$ ist eine lineare Abbildung $D : \mathcal{A} \rightarrow \mathcal{A}$, die $D(A \cdot B) = D(A) \cdot B + A \cdot D(B)$ für alle $A, B \in \mathcal{A}$ erfüllt.

Übungsaufgabe 9.65. (i) Zeige, dass jedes Element A einer Algebra $\mathcal{A}$ folgende Derivation definiert:

$$D_A : \mathcal{A} \rightarrow \mathcal{A}, \quad B \mapsto AB - BA$$

(ii) Sei $\mathcal{A}$ eine Banachalgebra und $D \in \mathcal{L}(\mathcal{A})$ eine Derivation von $\mathcal{A}$. Zeige dass $\exp(D)$ ein Algebrasomorphismus von $\mathcal{A}$ ist, d.h. ein invertierbares Element von

$$\{C \in \mathcal{L}(\mathcal{A}) \mid C(A \cdot B) = C(A) \cdot C(B) \text{ für alle } A, B \in \mathcal{A}\}.$$

(iii) Zeige $\exp(D_A)B = \exp(A) \cdot B \cdot \exp(-A)$ für alle A, B in einer Banachalgebra $\mathcal{A}$.

In der Vorlesung Analysis III wird gezeigt, dass die Derivationen der Algebra $C^\infty(\mathbb{R})$ von der Form $f \mapsto gf'$ für ein $g \in C^\infty(\mathbb{R})$ sind. Außerdem werden die entsprechenden Algebrasomorphismen bestimmt, soweit sie existieren.

In dem Buch L. Gillman, M. Jerison: "Rings of continuous functions" wird gezeigt, dass für jeden kompakten metrischen Raum X die einzigen Algebrasomorphismen von $C(X, \mathbb{R})$ von der Form $f \mapsto f \circ \Phi$ für einen Homöomorphismus $\Phi : X \rightarrow X$ sind. Daraus lässt sich folgern, dass $\mathcal{L}(C(X, \mathbb{R}))$ nur die triviale Derivation von $C(X, \mathbb{R})$ enthält.

Kapitel 10

Differentialrechnung von Funktionen mehrerer Veränderlicher

10.1 Ableitungen von $f : X \rightarrow Y$

Definition 10.1. (Ableitung) Seien X und Y normierte Vektorräume. Eine Abbildung f von einer offenen Menge $U \subset X$ nach Y heißt im Punkt $x_0 \in U$ differenzierbar, wenn es ein $A \in \mathcal{L}(X, Y)$ gibt, so dass die folgende Abbildung in x_0 stetig ist:

$$U \rightarrow \mathbb{R}, \quad x \mapsto \begin{cases} \frac{\|f(x) - f(x_0) - A(x - x_0)\|}{\|x - x_0\|} & \text{für } x \neq x_0 \\ 0 & \text{für } x = x_0. \end{cases}$$

A heißt Ableitung von f bei x_0 und wird mit $f'(x_0)$ oder $\frac{df}{dx}(x_0)$ bezeichnet. Wenn A und B beide diese Bedingung erfüllen, dann gibt es für alle $\epsilon > 0$ ein $\delta > 0$ mit

$$\frac{\|(A - B)(x - x_0)\|}{\|x - x_0\|} \leq \frac{\|A(x - x_0) - f(x) + f(x_0)\|}{\|x - x_0\|} + \frac{\|f(x) - f(x_0) - B(x - x_0)\|}{\|x - x_0\|} < 2\epsilon$$

für alle $x \in B(x_0, \delta) \setminus \{x_0\}$. Mit $x - x_0 = \delta y \in X$ für $0 < \|y\| < 1$ folgt $\|(A - B)y\| = \frac{\|(A - B)\delta y\|}{\delta} < 2\epsilon\|y\| < 2\epsilon$ und $\|A - B\| \leq 2\epsilon$. Weil das für alle $\epsilon > 0$ gilt, ist $A = B$.

Satz 10.2. Sei $f : U \subset X \rightarrow Y$ in $x_0 \in U$ differenzierbar. Dann ist f in x_0 stetig.

Beweis: Wegen $f(x) - f(x_0) = A(x - x_0) + (f(x) - f(x_0) - A(x - x_0))$ folgt aus der Differenzierbarkeit von f in x_0 , dass es ein $\delta > 0$ gibt, so dass $\|f(x) - f(x_0)\| \leq (\|A\| + 1)\|x - x_0\|$ gilt für alle $\|x - x_0\| < \delta$. Dann ist f in x_0 auch stetig. **q.e.d.**

Beispiel 10.3. (i) Sei f konstant. Dann ist
$$\frac{\|f(x) - f(x_0)\|}{\|x - x_0\|} = 0 \quad \text{für } x \neq x_0.$$
 Also ist f differenzierbar mit $f'(x_0) = 0$.

(ii) Für $f \in \mathcal{L}(X, Y)$ und $x \neq x_0 \in X$ gilt
$$\frac{\|f(x) - f(x_0) - f(x - x_0)\|}{\|x - x_0\|} = 0.$$

Also ist f in x_0 differenzierbar und $f'(x_0) = f$.

(iii) Die Abbildung $Y \rightarrow \mathcal{L}(\mathbb{R}, Y)$, $y \mapsto$ Multiplikation mit y besitzt offenbar die Umkehrabbildung $\mathcal{L}(\mathbb{R}, Y) \rightarrow Y$, $A \mapsto A(1)$ und ist ein isometrischer Isomorphismus von normierten Vektorräumen. Deshalb können wir die Ableitungen von differenzierbaren $f : (a, b) \rightarrow Y$ auf $(a, b) \subset \mathbb{R}$ auch durch $f' : (a, b) \rightarrow Y$ beschreiben. Für $x \neq x_0$ gilt

$$\frac{\|f(x) - f(x_0) - f'(x_0)(x - x_0)\|}{\|x - x_0\|} = \left\| \frac{f(x) - f(x_0)}{x - x_0} - f'(x_0) \right\|.$$

Dann sind für $f : (a, b) \rightarrow \mathbb{K}$ und $x_0 \in (a, b)$ die Definitionen 7.1 und 10.1 äquivalent.

Satz 10.4. (i) Sei $f, g : U \subset X \rightarrow Y$ in $x_0 \in U$ differenzierbar. Dann sind $f + g$ und $\lambda \cdot f$ für alle $\lambda \in \mathbb{K}$ in x_0 differenzierbar und es gilt

$$(f + g)'(x_0) = f'(x_0) + g'(x_0) \quad \text{bzw.} \quad (\lambda f)'(x_0) = \lambda f'(x_0).$$

(ii) Seien $f, g : U \subset X \rightarrow Y$ in x_0 differenzierbar und Y eine normierte Algebra. Dann ist $f \cdot g$ in x_0 differenzierbar und es gilt (Leibnizregel)

$$\begin{aligned} (f \cdot g)'(x_0) &= f'(x_0) \cdot g(x_0) + f(x_0) \cdot g'(x_0) \quad \text{mit} \\ (f'(x_0) \cdot g(x_0))(x) &= f'(x_0)(x) \cdot g(x_0) \quad \text{und} \quad (f(x_0) \cdot g'(x_0))(x) = f(x_0) \cdot g'(x_0)(x). \end{aligned}$$

(iii) Seien $f : U \subset X \rightarrow Y$ in $x_0 \in U$ und $g : V \subset Y \rightarrow Z$ in $f(x_0) \in V$ differenzierbar mit $f[U] \subset V$. Dann ist $g \circ f$ im Punkt x_0 differenzierbar mit

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

Beweis:(i) Mit den drei Eigenschaften von Normen folgt (i) für $x \neq x_0$ aus

$$\begin{aligned} & \frac{\|f(x) + g(x) - f(x_0) - g(x_0) - f'(x_0)(x - x_0) - g'(x_0)(x - x_0)\|}{\|x - x_0\|} \\ & \leq \frac{\|f(x) - f(x_0) - f'(x_0)(x - x_0)\|}{\|x - x_0\|} + \frac{\|g(x) - g(x_0) - g'(x_0)(x - x_0)\|}{\|x - x_0\|}. \\ & \frac{\|\lambda f(x) - \lambda f(x_0) - \lambda f'(x_0)(x - x_0)\|}{\|x - x_0\|} = |\lambda| \frac{\|f(x) - f(x_0) - f'(x_0)(x - x_0)\|}{\|x - x_0\|}. \end{aligned}$$

(ii) Weil in einer normierten Algebra $\|AB\| \leq \|A\| \cdot \|B\|$ gilt, folgt für $x \neq x_0$

$$\begin{aligned} & \frac{\|f(x) \cdot g(x) - f(x_0) \cdot g(x_0) - f'(x_0)(x - x_0)g(x_0) - f(x_0)g'(x_0)(x - x_0)\|}{\|x - x_0\|} \\ & \leq \frac{\|f(x) - f(x_0) - f'(x_0)(x - x_0)\|}{\|x - x_0\|} \|g(x)\| + \|f'(x_0)\| \cdot \|g(x) - g(x_0)\| + \\ & \quad + \|f(x_0)\| \frac{\|g(x) - g(x_0) - g'(x_0)(x - x_0)\|}{\|x - x_0\|}. \end{aligned}$$

Weil g in x_0 stetig ist, gibt es für jedes $\epsilon > 0$ ein $\delta > 0$, so dass aus $\|x - x_0\| < \delta$ folgt $\|g(x) - g(x_0)\| < \epsilon$ bzw. $\|g(x)\| \leq \|g(x_0)\| + \epsilon$. Dann folgt (ii).

$$\begin{aligned}
 \text{(iii)} \quad & \text{Aus Satz 9.61 folgt für } x \neq x_0 \quad \frac{\|(g \circ f)(x) - (g \circ f)(x_0) - g'(f(x_0))f'(x_0)(x - x_0)\|}{\|x - x_0\|} \\
 & \leq \frac{\|g(f(x)) - g(f(x_0)) - g'(f(x_0))(f(x) - f(x_0))\|}{\|f(x) - f(x_0)\|} \cdot \frac{\|f(x) - f(x_0)\|}{\|x - x_0\|} \\
 & \quad + \frac{\|g'(f(x_0))(f(x) - f(x_0) - f'(x_0)(x - x_0))\|}{\|x - x_0\|} \\
 & \leq \frac{\|g(f(x)) - g(f(x_0)) - g'(f(x_0))(f(x) - f(x_0))\|}{\|f(x) - f(x_0)\|} \\
 & \quad \cdot \left(\|f'(x_0)\| + \frac{\|f(x) - f(x_0) - f'(x_0)(x - x_0)\|}{\|x - x_0\|} \right) \\
 & \quad + \|g'(f(x_0))\| \frac{\|f(x) - f(x_0) - f'(x_0)(x - x_0)\|}{\|x - x_0\|}.
 \end{aligned}$$

Ersetze hierbei für $f(x) = f(x_0)$ die beiden Differenzenquotienten von g wie in der Definition von $g'(f(x_0))$ durch Null. Aus Satz 10.2 und Korollar 9.32 folgt (iii). **q.e.d.**

10.2 Schrankensatz

Wir verallgemeinern in diesem Abschnitt den Schrankensatz auf differenzierbare Abbildungen zwischen normierten Vektorräumen.

Lemma 10.5. *Seien f eine stetige Abbildung von einem kompakten Intervall $[a, b]$ mit $a < b$ in einen normierten Vektorraum Y und ϕ eine stetige reelle Funktion auf $[a, b]$. Wenn im Komplement einer abzählbaren Teilmenge von (a, b) sowohl f als auch ϕ differenzierbar sind und dort gilt $\|f'\| \leq \phi'$, dann gilt auch $\|f(b) - f(a)\| \leq \phi(b) - \phi(a)$.*

Beweis: Sei $(x_n)_{n \in \mathbb{N}}$ eine Abzählung der Punkte in $[a, b]$, an denen entweder f oder ϕ nicht differenzierbar ist oder $\|f'\| > \phi'$. Mit kontinuierlicher Induktion zeigen wir

$$\|f(x) - f(a)\| \leq \phi(x) - \phi(a) + \epsilon(x - a) + \epsilon \sum_{x_n < x} 2^{-n} \quad \text{für alle } \epsilon > 0 \text{ und } x \in [a, b].$$

Sei also $\epsilon > 0$ und A_ϵ die Menge aller $y \in [a, b]$, so dass diese Ungleichung für alle $x \in [a, y]$ gilt. Wegen der Stetigkeit von f und ϕ gilt auch für $y = \sup A_\epsilon$

$$\begin{aligned}
 \|f(y) - f(a)\| & \leq \phi(y) - \phi(a) + \epsilon(y - a) + \epsilon \sup \left\{ \sum_{x_n < x} 2^{-n} \mid x \in (a, y) \right\} \\
 & = \phi(y) - \phi(a) + \epsilon(y - a) + \epsilon \sum_{x_n < y} 2^{-n}.
 \end{aligned}$$

Deshalb ist A_ϵ ein Intervall von der Form $A_\epsilon = [a, y]$. Wenn $y \in (a, b)$ und f und ϕ in y differenzierbar sind und $\|f'(y)\| \leq \phi'(y)$ gilt, dann gibt es ein $\delta > 0$, so dass

$$\phi'(y) - \frac{\epsilon}{2} < \frac{\phi(x) - \phi(y)}{x - y} < \phi'(y) + \frac{\epsilon}{2} \quad \text{und} \quad \frac{\|f(x) - f(y) - f'(y)(x - y)\|}{|x - y|} < \frac{\epsilon}{2}$$

für alle $x \in (y - \delta, y) \cup (y, y + \delta)$ gilt. Dann folgt für $x \in (y, y + \delta)$

$$\begin{aligned} \|f(x) - f(y)\| &\leq \|f'(y)(x - y)\| + \|f(x) - f(y) - f'(y)(x - y)\| \\ &< \left(\|f'(y)\| + \frac{\epsilon}{2}\right) |x - y| \leq \left(\phi'(y) - \frac{\epsilon}{2} + \epsilon\right) (x - y) \\ &< \left(\frac{\phi(x) - \phi(y)}{x - y} + \epsilon\right) (x - y) = \phi(x) - \phi(y) + \epsilon(x - y) \quad \text{und} \\ \|f(x) - f(a)\| &\leq \|f(x) - f(y)\| + \|f(y) - f(a)\| \\ &< \phi(x) - \phi(y) + \epsilon(x - y) + \phi(y) - \phi(a) + \epsilon(y - a) + \epsilon \sum_{x_n < y} 2^{-n} \\ &\leq \phi(x) - \phi(a) + \epsilon(x - a) + \epsilon \sum_{x_n < x} 2^{-n}. \end{aligned}$$

Woraus $y + \delta \in A_\epsilon$ folgt, im Widerspruch zu $y = \sup A_\epsilon$. Wenn es andererseits ein $N \in \mathbb{N}$ gibt mit $x_N = y < b$, dann gibt es ein $\delta > 0$, so dass aus $x \in (y, y + \delta)$ folgt

$$\|f(x) - f(y)\| - (\phi(x) - \phi(y)) < \epsilon 2^{-N}.$$

Dann folgt für dieselben x wieder $\|f(x) - f(a)\| \leq \|f(x) - f(y)\| + \|f(y) - f(a)\|$

$$< \epsilon 2^{-N} + \phi(x) - \phi(a) + \epsilon(y - a) + \epsilon \sum_{x_n < y} 2^{-n} \leq \phi(x) - \phi(a) + \epsilon(x - a) + \epsilon \sum_{x_n < x} 2^{-n}.$$

Also gilt wieder $y + \delta \in A_\epsilon$, was $y = \sup A_\epsilon$ widerspricht. Dann muß aber $\sup A_\epsilon = b$ gelten. Weil das für alle $\epsilon > 0$ gilt, folgt auch $\|f(b) - f(a)\| \leq \phi(b) - \phi(a)$. **q.e.d.**

Korollar 10.6. (Schranksatz) Sei f eine stetige Abbildung von einer offenen Teilmenge U des normierten Vektorraumes X in den normierten Vektorraum Y . Wenn f im Komplement einer abzählbaren Teilmenge S von $D = \{(1 - t)a + tb \mid t \in [0, 1]\} \subset U$ differenzierbar ist mit $a, b \in U$, und die Ableitung auf $D \setminus S$ beschränkt ist, dann gilt

$$\begin{aligned} \|f(b) - f(a)\| &\leq \|b - a\| \sup\{\|f'(x)\| \mid x \in D \setminus S\} \quad \text{und} \\ \|f(b) - f(a) - A(b - a)\| &\leq \|b - a\| \sup\{\|f'(x) - A\| \mid x \in D \setminus S\} \quad \text{für } A \in \mathcal{L}(X, Y). \end{aligned}$$

Auf konvexen Mengen ist jede obere Schranke an $\|f'\|$ eine Lipschitzkonstante von f .

Beweis: Die Abbildung $x : \mathbb{R} \rightarrow X$, $t \mapsto x(t) = (t - 1)a + tb = a + t(b - a)$ ist wegen Beispiel 10.3 (i)-(ii) und Satz 10.4 (i) differenzierbar mit $x'(t) : s \mapsto s(b - a)$

als Element von $\mathcal{L}(\mathbb{R}, X)$ bzw. mithilfe von Beispiel 10.3 (iii) mit $x'(t) = (b - a) \in X$. Wegen Satz 10.4 (iii) ist dann auch die Abbildung $t \mapsto f(x(t))$ bei den $t \in [0, 1]$ differenzierbar, für die f bei $x(t) \in U$ differenzierbar ist, mit der Ableitung $(f \circ x)'(t) : s \mapsto sf'(x(t))(b - a)$ bzw. mithilfe von Beispiel 10.3 (iii) $(f \circ x)'(t) = f'(x(t))(b - a) \in Y$. Dann folgt die erste Behauptung aus Lemma 10.5 mit den beiden Funktionen

$$f \circ x : [0, 1] \rightarrow Y, \quad t \mapsto f(x(t)), \quad \phi : [0, 1] \rightarrow \mathbb{R}, \quad t \mapsto t\|b - a\| \sup\{\|f'(x)\| \mid x \in D \setminus S\}.$$

Die zweite Behauptung folgt aus diesem Lemma mit den beiden Funktionen

$$t \mapsto f(x(t)) - tA(b - a) \in Y, \quad t \mapsto t\|b - a\| \sup\{\|f'(x) - A\| \mid x \in D \setminus S\} \in \mathbb{R}. \quad \mathbf{q.e.d.}$$

10.3 Partielle Ableitungen

Definition 10.7. Eine Abbildung f von einer offenen Teilmenge U eines normierten Vektorraumes X in einen normierten Vektorraum Y heißt stetig differenzierbar, wenn

- (i) f in allen $x_0 \in U$ differenzierbar ist, und
- (ii) die Abbildung $f' : U \rightarrow \mathcal{L}(X, Y)$, $x \mapsto f'(x)$ stetig ist.

Wegen Beispiel 10.3 (iii) stimmt im Falle von $X = Y = \mathbb{R}$ diese Definition mit der Definition von stetig differenzierbaren reellen Funktionen auf Intervallen in $\mathbb{R}$ überein.

Definition 10.8. (partielle Ableitung) Sei f eine Funktion von einer offenen Teilmenge $U \subset X_1 \times X_2$ des kartesischen Produktes der normierten Vektorräume X_1 und X_2 in den normierten Vektorraum Y . Dann heißt f im Punkt $(x_1, x_2) \in U \subset X_1 \times X_2$ partiell differenzierbar, falls die Abbildung $x \mapsto f(x, x_2)$ im Punkt $x = x_1$, und die Abbildung $x \mapsto f(x_1, x)$ im Punkt $x = x_2$ differenzierbar ist. Die Ableitungen heißen partielle Ableitungen an der Stelle (x_1, x_2) und werden mit $\frac{\partial f}{\partial x_1}(x_1, x_2)$ und $\frac{\partial f}{\partial x_2}(x_1, x_2)$ bezeichnet. Allgemeiner heißt eine Abbildung von einer offenen Menge $U \subset X_1 \times \dots \times X_n$ eines n -fachen kartesischen Produktes von normierten Vektorräumen in einen normierten Vektorraum Y im Punkt $(x_1, \dots, x_n) \in U$ partiell differenzierbar, wenn für $i = 1, \dots, n$ die Abbildungen $x \mapsto f(x_1, \dots, x_{i-1}, x, x_{i+1}, \dots, x_n)$ bei $x = x_i$ differenzierbar sind.

Die wichtigsten Beispiele sind reelle Funktionen auf offenen Teilmengen des $\mathbb{R}^n$. Für jedes $A \in \mathcal{L}(X_1 \times X_2, Y)$ sind die beiden folgenden Abbildungen stetig und linear

$$A_1 : X_1 \rightarrow Y, \quad x_1 \mapsto A((x_1, 0)) \quad \text{und} \quad A_2 : X_2 \rightarrow Y, \quad x_2 \mapsto A((0, x_2)).$$

Weil für $x \in X_1 \setminus \{x_1\}$ mit $(x, x_2) \in U$ bzw. $x \in X_2 \setminus \{x_2\}$ mit $(x_1, x) \in U$

$$\begin{aligned} \frac{\|f(x, x_2) - f(x_1, x_2) - A((x, x_2) - (x_1, x_2))\|}{\|(x, x_2) - (x_1, x_2)\|} &= \frac{\|f(x, x_2) - f(x_1, x_2) - A_1(x - x_1)\|}{\|x - x_1\|} \\ \frac{\|f(x_1, x) - f(x_1, x_2) - A((x_1, x) - (x_1, x_2))\|}{\|(x_1, x) - (x_1, x_2)\|} &= \frac{\|f(x_1, x) - f(x_1, x_2) - A_2(x - x_2)\|}{\|x - x_2\|} \end{aligned}$$

gilt, folgt aus den Definitionen der folgende Satz:

Satz 10.9. Eine im Punkt $(x_1, x_2) \in U$ differenzierbare Funktion f von einer offenen Teilmenge $U \subset X_1 \times X_2$ des kartesischen Produktes zweier normierter Vektorräume in einen normierten Vektorraum Y ist in (x_1, x_2) auch partiell differenzierbar. **q.e.d.**

Beispiel 10.10. Die Umkehrung dieses Satzes gilt nicht:

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad (x, y) \mapsto \begin{cases} \frac{2xy}{x^2+y^2} & \text{für } (x, y) \neq 0 \\ 0 & \text{für } (x, y) = 0 \end{cases} \quad \text{besitzt die partiellen Ableitungen}$$

$$\frac{\partial f}{\partial x} = \begin{cases} \frac{2y}{x^2+y^2} - \frac{4x^2y}{(x^2+y^2)^2} & \text{für } (x, y) \neq 0 \\ 0 & \text{für } (x, y) = 0 \end{cases} \quad \frac{\partial f}{\partial y} = \begin{cases} \frac{2x}{x^2+y^2} - \frac{4y^2x}{(x^2+y^2)^2} & \text{für } (x, y) \neq 0 \\ 0 & \text{für } (x, y) = 0 \end{cases}$$

Für alle $r \in (0, \infty)$ und alle $\phi \in \mathbb{R}$ gilt $f(r \cos \phi, r \sin \phi) = 2 \sin \phi \cos \phi = \sin(2\phi)$, und deshalb $\lim_{r \rightarrow 0+} f(r \cos \phi, r \sin \phi) = \sin(2\phi)$. Also ist f im Punkt $(x, y) = 0$ nicht stetig.

Aber es gilt folgende Umkehrung.

Satz 10.11. Seien X_1, X_2, Y normierte Vektorräume, $U \subset X_1 \times X_2$ offen und $f : U \rightarrow Y$ bei $x = (x_1, x_2) \in U$ partiell differenzierbar. Wenn eine der partiellen Ableitungen

$$\frac{\partial f}{\partial x_1} : U \rightarrow \mathcal{L}(X_1, Y) \quad \text{oder} \quad \frac{\partial f}{\partial x_2} : U \rightarrow \mathcal{L}(X_2, Y)$$

auf U existiert und bei x stetig ist, dann ist f bei x differenzierbar. Also ist f genau dann auf U stetig differenzierbar, wenn f auf U stetig partiell differenzierbar ist.

Beweis: Wenn $A_1 \in \mathcal{L}(X_1, Y)$ und $A_2 \in \mathcal{L}(X_2, Y)$, dann sind auch die Abbildungen

$$X_1 \times X_2 \rightarrow Y, \quad (x_1, x_2) \mapsto A_1(x_1) \text{ bzw. } X_1 \times X_2 \rightarrow Y, \quad (x_1, x_2) \mapsto A_2(x_2)$$

lineare und stetige Abbildungen in $\mathcal{L}(X_1 \times X_2, Y)$. Also ist

$$A_1 \times A_2 : X_1 \times X_2 \rightarrow Y, \quad (x_1, x_2) \mapsto A_1(x_1) + A_2(x_2)$$

eine stetige lineare Abbildung in $\mathcal{L}(X_1 \times X_2, Y)$. Umgekehrt sind für jede stetige lineare Abbildung $A \in \mathcal{L}(X_1 \times X_2, Y)$ die Abbildungen $A_1 : X_1 \rightarrow Y, x_1 \mapsto A((x_1, 0))$ und $A_2 : X_2 \rightarrow Y, x_2 \mapsto A((0, x_2))$ stetig und linear. Und es gilt

$$\|A((x_1, x_2))\| = \|A((x_1, 0)) + A((0, x_2))\| \leq \|A((x_1, 0))\| + \|A((0, x_2))\|.$$

Aufgrund der Definition der Norm von $\mathcal{L}(X_1 \times X_2, Y)$ folgt dann

$$\|A_1\| \leq \|A\| \quad \|A_2\| \leq \|A\| \quad \|A\| \leq \|A_1\| + \|A_2\| \leq 2\|A\|.$$

Also ist die Abbildung $\mathcal{L}(X_1, Y) \times \mathcal{L}(X_2, Y) \rightarrow \mathcal{L}(X_1 \times X_2, Y)$, $(A_1, A_2) \mapsto A$ eine bijektive Abbildung von normierten Vektorräumen und die beiden Normen von $\mathcal{L}(X_1, Y) \times \mathcal{L}(X_2, Y)$ und $\mathcal{L}(X_1 \times X_2, Y)$ sind bezüglich dieser Identifikation äquivalent. Daraus folgt, dass für jede stetig differenzierbare Funktion $f : U \rightarrow Y$, die beiden partiellen Ableitungen $\frac{\partial f}{\partial x_1} : U \rightarrow \mathcal{L}(X_1, Y)$ und $\frac{\partial f}{\partial x_2} : U \rightarrow \mathcal{L}(X_2, Y)$ stetig sind.

Wenn umgekehrt f in (x_1, x_2) partiell differenzierbar ist, dann folgt

$$\begin{aligned} \left\| f(y_1, y_2) - f(x_1, x_2) - \frac{\partial f(x_1, x_2)}{\partial x_1}(y_1 - x_1) - \frac{\partial f(x_1, x_2)}{\partial x_2}(y_2 - x_2) \right\| &\leq \\ &\leq \left\| f(y_1, y_2) - f(x_1, y_2) - \frac{\partial f(x_1, x_2)}{\partial x_1}(y_1 - x_1) \right\| + \\ &\quad + \left\| f(x_1, y_2) - f(x_1, x_2) - \frac{\partial f(x_1, x_2)}{\partial x_2}(y_2 - x_2) \right\|. \end{aligned}$$

Wenn $\frac{\partial f}{\partial x_1}$ auf U existiert und bei (x_1, x_2) stetig ist, dann gibt es für jedes $\epsilon > 0$ ein $\delta > 0$, so dass

$$\left\| \frac{\partial f(z_1, z_2)}{\partial x_1} - \frac{\partial f(x_1, x_2)}{\partial x_1} \right\| < \epsilon$$

für $(z_1, z_2) \in B((x_1, x_2), \delta)$ gilt. Aus Korollar 10.6 folgt für $(y_1, y_2) \in B((x_1, x_2), \delta)$

$$\left\| f(y_1, y_2) - f(x_1, y_2) - \frac{\partial f(x_1, x_2)}{\partial x_1}(y_1 - x_1) \right\| < \epsilon \|y_1 - x_1\|.$$

Weil $\frac{\partial f}{\partial x_2}$ in (x_1, x_2) existiert, gibt es auch ein $\delta' > 0$, so dass für $y_2 \in B(x_2, \delta')$ folgt

$$\left\| f(x_1, y_2) - f(x_1, x_2) - \frac{\partial f(x_1, x_2)}{\partial x_2}(y_2 - x_2) \right\| < \epsilon \|y_2 - x_2\|.$$

Dann folgt für $(y_1, y_2) \in B((x_1, x_2), \min\{\delta, \delta'\})$ auch

$$\|f(y_1, y_2) - f(x_1, x_2) - f'(x_1, x_2)((y_1, y_2) - (x_1, x_2))\| < \epsilon(\|y_1 - x_1\| + \|y_2 - x_2\|),$$

wobei $f'(x_1, x_2) = \frac{\partial f(x_1, x_2)}{\partial x_1} \times \frac{\partial f(x_1, x_2)}{\partial x_2}$ durch die partiellen Ableitungen gegeben ist. Also ist f differenzierbar, und mit den partiellen Ableitungen stetig differenzierbar. **q.e.d.**

Durch mehrmaliges Anwenden erhalten wir dann auch die entsprechende Aussage für Abbildungen von offenen Teilmengen U des n -fachen kartesischen Produktes von normierten Vektorräumen in einen normierten Vektorraum. Unsere wichtigsten Beispiele sind wieder reelle Funktionen auf offenen Teilmengen des $\mathbb{R}^n$.

Beispiel 10.12. (i) $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $(x, y) \mapsto \begin{cases} \frac{2xy}{x^2+y^2} & \text{für } (x, y) \neq 0 \\ 0 & \text{für } (x, y) = 0. \end{cases}$

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

Für $y = 0$ ist $\frac{\partial f}{\partial x} = 0$ und für $x = 0$ ist $\frac{\partial f}{\partial y} = 0$, so dass diese partiellen Ableitungen für alle $(x, y) \in \mathbb{R}^2$ existieren. Allerdings sind sie in keiner Umgebung von $(x, y) = (0, 0)$ beschränkt, und deshalb auch nicht stetig. Wir hatten schon im Beispiel 10.10 gesehen, dass f bei $(0, 0)$ nicht stetig und deshalb auch nicht differenzierbar ist.

$$(ii) \quad f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad (x, y) \mapsto \begin{cases} \frac{x^3 - y^3}{x^2 + y^2} & \text{für } (x, y) \neq 0 \\ 0 & \text{für } (x, y) = 0. \end{cases}$$

Offenbar gilt $|f(x, y)| \leq \frac{|x|^3}{x^2 + y^2} + \frac{|y|^3}{x^2 + y^2} \leq |x| + |y|$. Also ist f stetig.

$$\frac{\partial f}{\partial x} = \begin{cases} \frac{3x^2(x^2 + y^2) - 2x(x^3 - y^3)}{(x^2 + y^2)^2} = \frac{x(x^3 + 3xy^2 + 2y^3)}{(x^2 + y^2)^2} & \text{für } (x, y) \neq 0 \\ 1 & \text{für } (x, y) = 0 \end{cases}$$

$$\frac{\partial f}{\partial y} = \begin{cases} -\frac{3y^2(x^2 + y^2) + 2y(x^3 - y^3)}{(x^2 + y^2)^2} = -\frac{y(y^3 + 3yx^2 + 2x^3)}{(x^2 + y^2)^2} & \text{für } (x, y) \neq 0 \\ -1 & \text{für } (x, y) = 0. \end{cases}$$

Wegen $\frac{\partial f(x, 0)}{\partial x} = 1$ und $\frac{\partial f(0, y)}{\partial y} = -1$ ist f partiell differenzierbar. In Beispiel 10.14 (iii) werden wir sehen, dass f in $(0, 0)$ nicht differenzierbar ist.

(iii) Alle Polynome in endlich vielen Variablen sind partiell unendlich oft stetig differenzierbar, und deshalb differenzierbar.

Definition 10.13. (Richtungsableitung) Für eine Funktion $f : U \rightarrow Y$ von einer offenen Teilmenge U eines normierten Vektorraumes X in einem normierten Vektorraum Y ist die Richtungsableitung in $x_0 \in U$ in Richtung $x_1 \in X$ die Ableitung bei $t = 0$ von

$$(-\epsilon, \epsilon) \rightarrow Y, \quad t \mapsto f(x_0 + tx_1).$$

Wie in Beispiel 10.3 (iii) identifizieren wir dabei $\mathcal{L}(\mathbb{R}, Y)$ durch $A \mapsto A(1)$ mit Y .

Beispiel 10.14. (i) Sei $U \subset X$ offen und $f : U \rightarrow Y$ in $x_0 \in U$ differenzierbar. Für $x_1 \in X$ ist die Abbildung $x : \mathbb{R} \rightarrow X$, $t \mapsto x(t) = x_0 + tx_1$ wegen Beispiel 10.3 (i)-(ii) und Satz 10.4 differenzierbar mit $x'(t) = x_1$ im Sinne von Beispiel 10.3 (iii). Dann gibt es ein Intervall $(-\epsilon, \epsilon)$ im Urbild $x^{-1}[U]$ von U unter x , und wegen Satz 10.4 (iii) ist $(-\epsilon, \epsilon) \rightarrow Y$, $t \mapsto f(x(t))$ bei $t = 0$ differenzierbar mit der Ableitung $f'(x_0)(x_1)$ im Sinne von Beispiel 10.3 (iii). Also existiert die Richtungsableitung und es gilt

$$\frac{d}{dt} f(x_0 + tx_1)|_{t=0} = f'(x_0)(x_1).$$

$$(ii) \quad f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad (x, y) \mapsto \begin{cases} \frac{2xy}{x^2 + y^2} & \text{für } (x, y) \neq 0 \\ 0 & \text{für } (x, y) = 0. \end{cases}$$

Für $t \neq 0$ ist dann $f(t \cos(\phi), t \sin(\phi)) = \sin(2\phi)$ und $f(t \cos(\phi), t \sin(\phi)) = 0$ für $t = 0$. Also ist f in $t = 0$ für $\phi \notin \frac{\pi}{2}\mathbb{Z}$ nicht stetig und auch nicht differenzierbar. Für $\phi \in \frac{\pi}{2}\mathbb{Z}$ existieren die Richtungsableitungen und verschwinden.

$$(iii) \quad f: \mathbb{R}^2 \rightarrow \mathbb{R}, \quad (x, y) \mapsto \begin{cases} \frac{x^3 - y^3}{x^2 + y^2} & \text{für } (x, y) \neq 0 \\ 0 & \text{für } (x, y) = 0. \end{cases}$$

Dann gilt $f(t(x, y)) = f(tx, ty) = tf(x, y)$ für $(x, y) \in \mathbb{R}^2$ und $x_0 = (0, 0)$. Also existieren alle Richtungsableitungen und setzen sich im Punkt $(0, 0)$ zu f zusammen. Wegen $2 = f(2, 0) \neq f(1, 1) + f(1, -1) = 0 + 1$ ist f im Punkt $(0, 0)$ nicht differenzierbar.

$$(iv) \quad f: \mathbb{R}^2 \rightarrow \mathbb{R}, \quad (x, y) \mapsto \begin{cases} \frac{2x^3y}{x^4 + y^2} & \text{für } (x, y) \neq 0 \\ 0 & \text{für } (x, y) = 0. \end{cases}$$

Die Richtungsableitungen von f in $(0, 0)$ in Richtung von (x, y) verschwinden alle:

$$\left. \frac{d}{dt} f(tx, ty) \right|_{t=0} = \begin{cases} \left. \frac{d}{dt} \frac{2t^2x^3y}{t^2x^4 + y^2} \right|_{t=0} = 0 & \text{für } y \neq 0 \\ 0 & \text{für } y = 0, \end{cases} \quad |f(x, y)| = 2|x| \sqrt{\frac{x^4}{x^4 + y^2} \frac{y^2}{x^4 + y^2}} \leq 2|x|.$$

Wegen $\frac{d}{dt}(t, t^2) = (1, 2t)$ und $\frac{d}{dt}f(t, t^2) = \frac{d}{dt}t = 1 \neq 0$ ist f in $(0, 0)$ nicht differenzierbar.

Wir wollen den wichtigsten Fall von Funktionen $\mathbb{R}^n \rightarrow \mathbb{R}^m$ genauer betrachten.

Definition 10.15. (Partielle Ableitungen in $\mathbb{R}^n$) Sei $f: U \rightarrow \mathbb{R}^m$ eine Funktion von einer offenen Teilmenge $U \subset \mathbb{R}^n$ in den $\mathbb{R}^m$. Dann sind die Komponenten $(f_1, \dots, f_m)$ von f offenbar reelle Funktionen auf U . Die Funktion f ist in $(x_1, \dots, x_n) \in U$ genau dann partiell differenzierbar, wenn für alle $i = 1, \dots, n$ und $j = 1, \dots, m$ die Funktionen

$$x \mapsto f_j(x_1, \dots, x_{i-1}, x, x_{i+1}, \dots, x_n)$$

bei $x = x_i$ differenzierbar sind. Die entsprechenden Ableitungen heißen partielle Ableitungen von f und werden mit $\frac{\partial f_j}{\partial x_i}(x_1, \dots, x_n)$ bezeichnet. Wenn diese partiellen Ableitungen für alle $x \in U$ existieren, heißt f auf U partiell differenzierbar.

Definition 10.16. (Vektorfeld, Gradient, Divergenz und Rotation) Eine Abbildung von einer offenen Menge $U \subset \mathbb{R}^n$ nach $\mathbb{R}^n$ wird Vektorfeld genannt. Das Vektorfeld

$$\text{grad } f = \nabla f = \left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n} \right)$$

der partiellen Ableitungen einer partiell differenzierbaren reellen Funktion f auf einer offenen Menge $U \subset \mathbb{R}^n$ heißt Gradient von f . Wenn f ein partiell differenzierbares Vektorfeld ist, dann ist die Divergenz von f folgende reelle Funktion

$$\text{div } f = \nabla \cdot f = \frac{\partial f_1}{\partial x_1} + \dots + \frac{\partial f_n}{\partial x_n}.$$

$$\text{Die lineare Abbildung } \Delta: f \mapsto \Delta f = \text{div grad } f = \frac{\partial^2 f}{\partial x_1^2} + \dots + \frac{\partial^2 f}{\partial x_n^2}$$

auf den zweimal partiell differenzierbaren reellen Funktionen heißt Laplaceoperator.

Im Fall von $n = 3$ ist die Rotation eines differenzierbaren Vektorfeldes f definiert durch

$$\text{rot } f = \left(\frac{\partial f_3}{\partial x_2} - \frac{\partial f_2}{\partial x_3}, \frac{\partial f_1}{\partial x_3} - \frac{\partial f_3}{\partial x_1}, \frac{\partial f_2}{\partial x_1} - \frac{\partial f_1}{\partial x_2} \right).$$

Wenn die reelle Funktion f in $x_0 \in U$ differenzierbar ist, dann ist f auch in x_0 partiell differenzierbar und die partiellen Ableitungen sind die Richtungsableitungen in Richtung der kanonischen Einheitsvektoren $e_1, \dots, e_n$ aus dem Beweis von Satz 9.37. Wegen der Linearität der Ableitung ist die Ableitung die lineare Abbildung:

$$f'(x_0) : \mathbb{R}^n \rightarrow \mathbb{R}, \quad (x_1, \dots, x_n) \mapsto x_1 \frac{\partial f}{\partial x_1}(x_0) + \dots + x_n \frac{\partial f}{\partial x_n}(x_0).$$

Wenn wir den $\mathbb{R}^n$ mit den Spaltenvektoren bezeichnen, können wir diese Abbildung durch das Matrixprodukt des Zeilenvektors ∇f mit dem Spaltenvektor x darstellen:

$$f'(x_0) : \mathbb{R}^n \rightarrow \mathbb{R}, \quad x \mapsto (\nabla f(x_0)) \cdot x.$$

Oder allgemeiner, für eine $\mathbb{R}^m$ -wertige Funktion können wir die Ableitung $f'(x_0)$ von f an der Stelle x_0 als lineare Abbildung mit der Jacobimatrix identifizieren:

$$\frac{\partial f(x_0)}{\partial x} = \begin{pmatrix} \frac{\partial f_1(x_0)}{\partial x_1} & \dots & \frac{\partial f_1(x_0)}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_m(x_0)}{\partial x_1} & \dots & \frac{\partial f_m(x_0)}{\partial x_n} \end{pmatrix} \quad f'(x_0) : \mathbb{R}^n \rightarrow \mathbb{R}^m, \quad x \mapsto \frac{\partial f(x_0)}{\partial x} \cdot x.$$

Die lineare Abbildung ist einfach die Matrixmultiplikation der Jacobimatrix, einer $m \times n$ -Matrix, mit dem Spaltenvektoren in $\mathbb{R}^n$. Insbesondere ist also die Richtungsableitung einer reellen Funktion f auf U an der Stelle $x_0 \in U$ in Richtung eines Vektors $x_1 \in \mathbb{R}^n$ das Skalarprodukt des Gradienten $\nabla f(x_0)$ von f an der Stelle x_0 mit dem Vektor x_1 :

$$\frac{d}{dt} f(x_0 + tx_1)|_{t=0} = x_1 \cdot \nabla f(x_0).$$

Satz 10.9 und Satz 10.11 zeigen insbesondere, dass folgendes gilt:

Korollar 10.17. *Sei $f : U \rightarrow \mathbb{R}^m$ eine Funktion auf einer offenen Teilmenge $U \subset \mathbb{R}^n$.*

- (i) *Wenn f in $x_0 \in U$ differenzierbar ist, dann existieren in x_0 alle partiellen Ableitungen $\frac{\partial f_j}{\partial x_i}(x_0)$ und setzen sich zu der Jacobimatrix zusammen.*
- (ii) *Wenn f auf U stetig differenzierbar ist, dann existieren auf U die partiellen Ableitungen $\frac{\partial f_j}{\partial x_i}$ und setzen sich zusammen zu einer stetigen Funktion von U in die $m \times n$ -Matrizen in $\mathbb{R}^{m \times n}$. Diese Matrizen heißen Jacobimatrizen von f .*
- (iii) *Wenn f auf U partiell differenzierbar ist, und alle partiellen Ableitungen $\frac{\partial f_j}{\partial x_i}$ auf U stetig sind, dann ist f auf U stetig differenzierbar. Die Ableitung bei x_0 ist die Multiplikation der Jacobimatrix $\frac{\partial f(x_0)}{\partial x}$ mit Spaltenvektoren in $\mathbb{R}^n$. **q.e.d.***

Definition 10.18. *Eine Nullstelle $x_0 \in U$ der Ableitung f' einer auf einer offenen Menge U reellen differenzierbaren Funktion f heißt kritischer Punkt.*

Satz 10.19. *Jedes lokale Maximum (bzw. Minimum) einer differenzierbaren reellen Funktion auf einer offenen Menge ist ein kritischer Punkt.*

Beweis: Sei x_0 ein solches lokales Maximum (bzw. Minimum). Dann ist für alle $x_1 \in X$ die entsprechende Abbildung $t \mapsto f(x_0 + tx_1)$ auf einer Umgebung von $t = 0$ differenzierbar und besitzt dort ein lokales Maximum (bzw. Minimum). Also verschwindet die entsprechende Richtungsableitung. Dann verschwindet auch $f'(x_0)$ auf allen x_1 . **q.e.d.**

10.4 Höhere Ableitungen

Sei f eine auf einer offenen Teilmenge U eines Banachraumes X differenzierbare Funktion in den Banachraum Y . Wenn f zweimal differenzierbar ist, dann ist f' stetig. Die Ableitung f' ist dann eine stetige Abbildung von U nach $\mathcal{L}(X, Y)$. Die zweite Ableitung $f''(x_0)$ ist an den Stellen $x_0 \in U$, wo sie existiert, ein Element von $\mathcal{L}(X, \mathcal{L}(X, Y))$.

Definition 10.20. *Eine Abbildung $A : V \times V \rightarrow W$ heißt bilinear, wenn für alle $v, v', v'' \in V$ und $\lambda \in \mathbb{K}$ gilt*

$$\begin{aligned} A(v + v'', v') &= A(v, v') + A(v'', v') & \text{und} & & A(v, v' + v'') &= A(v, v') + A(v, v'') \\ A(\lambda v, v') &= \lambda A(v, v') & \text{und} & & A(v, \lambda v') &= \lambda A(v, v'). \end{aligned}$$

Das kartesische Produkt $V \times V$ von (normierten) Vektorräumen ist wieder ein normierter Vektorraum. Die bilinearen Abbildungen von $V \times V$ nach W unterscheiden sich von den linearen Abbildungen von $V \times V$ nach W . Es gibt einen anderen Vektorraum $V \otimes V$, den man das Tensorprodukt von V mit V nennt, so dass die linearen Abbildungen von $V \otimes V$ nach W genau die bilinearen Abbildungen von $V \times V$ nach W sind. Allerdings besitzt $V \otimes V$ keine natürliche Norm. Für die Dimensionen gilt

$$\dim(V \times V) = \dim(V) + \dim(V) \qquad \dim(V \otimes V) = \dim(V) \cdot \dim(V).$$

Die bilinearen Abbildungen von $V \times V$ nach W lassen sich mit den linearen Abbildungen von V in die linearen Abbildungen von V nach W identifizieren:

Lemma 10.21. *Eine Abbildung $A : V \times V \rightarrow W$ ist genau dann bilinear, wenn*

$$B : V \rightarrow \{\text{Abbildungen } V \rightarrow W\}, \qquad v \mapsto B(v), \qquad B(v) : V \rightarrow W,$$

*eine lineare Abbildung von V in die linearen Abbildungen von V nach W ist, wobei B aus A als $B(v)(v') = A(v, v')$ und A aus B als $A(v, v') = B(v)(v')$ definiert ist. **q.e.d.***

Satz 10.22. (Satz von Schwarz) *Sei f eine differenzierbare Abbildung von einer offenen Teilmenge $U \subset X$ eines normierten Vektorraumes X in den normierten Vektorraum Y . Wenn f im Punkt x_0 zweimal differenzierbar ist, dann ist die der zweiten Ableitung entsprechende bilineare Abbildung $f''(x_0) : X \times X \rightarrow Y$ symmetrisch, d.h.*

$$(f''(x_0)x)y = (f''(x_0)y)x \quad \text{für alle } x, y \in X.$$

Beweis: Für $t \in [0, 1]$ und kleine $x, y \in X$ sei $g(t) = f(x_0 + tx + y) - f(x_0 + tx)$. Dann ist g differenzierbar mit der Ableitung mit Werten in $Y \simeq \mathcal{L}(\mathbb{R}, Y)$:

$$\begin{aligned} g'(t) &= f'(x_0 + tx + y)x - f'(x_0 + tx)x \\ &= ((f'(x_0 + tx + y) - f'(x_0)) - (f'(x_0 + tx) - f'(x_0)))x \end{aligned}$$

Weil f in x_0 zweimal differenzierbar ist, gibt es zu jedem $\epsilon > 0$ ein $\delta > 0$, so dass $B(x_0, 2\delta) \subset U$ und außerdem für $x, y \in B(0, \delta) \subset X$ und $t \in [0, 1]$ die Ungleichungen

$$\begin{aligned} \|f'(x_0 + tx + y) - f'(x_0) - f''(x_0)tx - f''(x_0)y\| &\leq \epsilon\|tx + y\| \leq \epsilon(t\|x\| + \|y\|) \\ \|f'(x_0 + tx) - f'(x_0) - f''(x_0)tx\| &\leq \epsilon\|tx\| = \epsilon t\|x\| \end{aligned}$$

gelten. Daraus folgt für $t \in [0, 1]$ $\|g'(t) - (f''(x_0)y)x\| \leq \epsilon\|x\|(2\|x\| + \|y\|)$.

Die Anwendung von Korollar 10.6 auf die Funktion $t \mapsto g(t) - t(f''(x_0)y)x$ ergibt dann

$$\|g(1) - g(0) - (f''(x_0)y)x\| \leq \sup\{\|g'(t) - (f''(x_0)y)x\| \mid t \in [0, 1]\} \leq \epsilon\|x\|(2\|x\| + \|y\|).$$

Weil $g(1) - g(0) = f(x_0 + x + y) - f(x_0 + x) - f(x_0 + y) + f(x_0)$ in x und y symmetrisch ist gilt dann auch $\|g(1) - g(0) - (f''(x_0)x)y\| \leq \epsilon\|y\|(2\|y\| + \|x\|)$. Daraus folgt

$$\|(f''(x_0)y)x - (f''(x_0)x)y\| \leq 2\epsilon(\|x\|^2 + \|x\| \cdot \|y\| + \|y\|^2) \quad \text{für alle } x, y \in B(0, \delta).$$

Diese Ungleichung gilt wegen der Linearität nicht nur für $x, y \in B(0, \delta)$, sondern für $x, y \in X$. Im Grenzwert $\epsilon \rightarrow 0$ folgt $(f''(x_0)y)x = (f''(x_0)x)y$ für alle $x, y \in X$. **q.e.d.**

Zusammen mit Satz 10.11 erhalten wir

Korollar 10.23. Sei $f : U \rightarrow \mathbb{R}^m$ eine zweimal stetig partiell differenzierbare Funktion von einer offenen Teilmenge $U \subset \mathbb{R}^n$ nach $\mathbb{R}^m$. Dann vertauschen die partiellen Ableitungen, d.h. für alle $i, j = 1, \dots, n$ und $k = 1, \dots, m$ gilt

$$\partial_i \partial_j f_k = \frac{\partial^2 f_k}{\partial x_i \partial x_j} = \frac{\partial^2 f_k}{\partial x_j \partial x_i} = \partial_j \partial_i f_k \quad \text{und} \quad \text{rot grad } f = 0 \quad \text{für } n = 3, m = 1. \quad \mathbf{q.e.d.}$$

Durch mehrfaches Anwenden und differenzieren erhalten wir dann auch

Korollar 10.24. Sei f eine Abbildung von einer offenen Teilmenge U eines normierten Vektorraumes X in den normierten Vektorraum Y , die in $x_0 \in U$ n -mal differenzierbar ist. Dann ist $f^{(n)}(x_0)$ eine multilineare symmetrische Abbildung von $X \times X \times \dots \times X$ nach Y . D.h. für jede Permutation $\sigma : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ und $x_1, \dots, x_n \in X$ gilt

$$(\dots ((f^{(n)}(x_0)x_1)x_2)\dots)x_n = (\dots ((f^{(n)}(x_0)x_{\sigma(1)})x_{\sigma(2)})\dots)x_{\sigma(n)}. \quad \mathbf{q.e.d.}$$

Beispiel 10.25. Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $(x, y) \mapsto f(x, y)$ mit $f(x, y) = xy \frac{x^2 - y^2}{x^2 + y^2}$ für $(x, y) \neq (0, 0)$ und $f(0, 0) = 0$. Dann ist f zweimal partiell differenzierbar.

$$\frac{\partial f(x, y)}{\partial x} = y \frac{(x^2 - y^2)(x^2 + y^2) + 2x^2(x^2 + y^2) - 2x^2(x^2 - y^2)}{(x^2 + y^2)^2} = y \frac{x^4 + 4x^2y^2 - y^4}{(x^2 + y^2)^2}$$

$$\frac{\partial f(x, y)}{\partial y} = x \frac{(x^2 - y^2)(x^2 + y^2) - 2y^2(x^2 + y^2) - 2y^2(x^2 - y^2)}{(x^2 + y^2)^2} = x \frac{x^4 - 4x^2y^2 - y^4}{(x^2 + y^2)^2}.$$

$$\frac{\partial^2 f(x, 0)}{\partial x^2} = 0, \quad \frac{\partial^2 f(0, y)}{\partial y \partial x} = -1, \quad \frac{\partial^2 f(x, 0)}{\partial x \partial y} = 1, \quad \frac{\partial^2 f(0, y)}{\partial y^2} = 0.$$

Also existieren auf $\mathbb{R}^2$ alle zweiten partiellen Ableitungen, mit $\frac{\partial^2 f}{\partial y \partial x}(0, 0) \neq \frac{\partial^2 f}{\partial x \partial y}(0, 0)$.

Definition 10.26. Sei $f : U \rightarrow \mathbb{R}$ eine reelle Funktion auf einer offenen Teilmenge U des normierten Vektorraumes X , die bei $x_0 \in U$ zweimal differenzierbar ist. Dann definiert die zweite Ableitung eine symmetrische Bilinearform auf X :

$$f''(x_0) : X \times X \rightarrow \mathbb{R}, \quad (x, y) \mapsto f''(x_0)(x, y).$$

Für $X = \mathbb{R}^n$ identifizieren wir die Elemente von X wieder mit den Spaltenvektoren. Dann ist diese Bilinearform durch die sogenannte Hessematrix gegeben:

$$f''(x_0)(x, y) = y^t \cdot \frac{\partial^2 f(x_0)}{\partial x^2} \cdot x = \sum_{i,j=1}^n x_j \frac{\partial^2 f(x_0)}{\partial x_j \partial x_i} y_i.$$

Satz 10.27. Sei $f : U \rightarrow \mathbb{R}$ eine auf einer offenen Menge zweimal differenzierbare reelle Funktion f . Dann ist die zweite Ableitung bei allen lokalen Minima (Maxima) eine nicht negative (nicht positive) Bilinearform: $f''(x_0)(x, x) \geq 0$ bzw. ≤ 0 für alle $x \in X$. Gibt es umgekehrt einen kritischen Punkt $x_0 \in U$ und ein $\epsilon > 0$ mit

$$f''(x_0)(x, x) \geq \epsilon \|x\|^2 \quad \text{bzw.} \quad f''(x_0)(x, x) \leq -\epsilon \|x\|^2 \quad \text{für alle } x \in X,$$

dann ist der kritische Punkt ein striktes lokales Minimum bzw. Maximum.

Beweis: Wenn x_0 ein lokales Maximum bzw. Minimum von f ist, dann für alle $x \in X$ auch $t = 0$ von $t \mapsto f(x_0 + tx)$. Deshalb folgt die erste Aussage aus Korollar 7.17.

Umgekehrt folgt aus der zweimaligen Differenzierbarkeit, dass für ein $\delta > 0$

$$-\frac{\epsilon}{2} \|x\|^2 < f'(x_0 + x)x - f'(x_0)x - f''(x_0)(x, x) < \frac{\epsilon}{2} \|x\|^2$$

für alle $x \in B(0, \delta)$ gilt. Daraus und den obigen Bedingungen folgt für die gleichen x

$$f'(x_0 + x)x = f'(x_0 + x)x - f'(x_0)x - f''(x_0)(x, x) + f''(x_0)(x, x) > \frac{\epsilon}{2} \|x\|^2 \quad \text{bzw.} \quad < -\frac{\epsilon}{2} \|x\|^2.$$

Also ist $f(x_0 + x) - f(x_0) = \int_0^1 f'(x_0 + tx)x dt > \frac{\epsilon}{4}\|x\|^2$ bzw. $< -\frac{\epsilon}{4}\|x\|^2$. **q.e.d.**

Auf endlichdimensionalen Räumen zeigt Beweis von Satz 9.37 wie $f''(x_0)(x, x) \geq \epsilon\|x\|^2$ aus $f''(x_0)(x, x) > 0$ für alle $x \in X \setminus \{0\}$ folgt. In unendlichdimensionalen Räumen gilt das nicht, und diese Bedingung ist nicht hinreichend für ein lokales Minimum.

Beispiel 10.28. Sei $f : C([0, 1]) \rightarrow \mathbb{R}$, $x \mapsto \int_0^1 x^2(t)(t - x(t))dt$. Dann ist $f''(0)(x, x) = 2 \int_0^1 x^2(t)tdt > 0$ für alle $x \in C([0, 1]) \setminus \{0\}$. Sei $x_\epsilon(t) = \begin{cases} \epsilon - t & \text{für } 0 \leq t < \epsilon \\ 0 & \text{für } \epsilon \leq t \leq 1 \end{cases}$ mit $\epsilon \in (0, 1)$. Dann gilt $\|x_\epsilon\|_\infty = \epsilon$ und $f(sx_\epsilon) = s^2 \int_0^\epsilon (\epsilon - t)^2tdt - s^3 \int_0^\epsilon (\epsilon - t)^3dt = -\frac{s^2}{3}(\epsilon - t)^3t \Big|_0^\epsilon - (\frac{s^2}{12} - \frac{s^3}{4})(\epsilon - t)^4 \Big|_0^\epsilon = (\frac{s^2}{12} - \frac{s^3}{4})\epsilon^4$. Also ist $x = 0$ kein lokales Minimum.

Zum Abschluss wollen wir das Taylorpolynom und die Taylorreihe einer Funktion $f : U \rightarrow Y$ auf einer offenen konvexen Teilmenge $U \subset X$ eines normierten Vektorraumes X in einem normierten Vektorraum Y betrachten. Für $x_0, x \in U$ sei

$$g : [0, 1] \rightarrow \mathbb{R}, \quad t \mapsto f(x_0 + t(x - x_0))$$

Wenn f auf U n -mal differenzierbar ist, dann ist auch g n -mal differenzierbar. Wegen der Kettenregel Satz 10.4 (iii) ist die m -te Ableitung von g gleich

$$g^{(m)}(t) = (\dots (f^{(m)}(x_0 + t(x - x_0))(x - x_0)) \dots (x - x_0)),$$

also die m -lineare symmetrische Form zu $f^{(m)}(x_0 + t(x - x_0))$ ausgewertet auf $((x - x_0), \dots, (x - x_0)) \in X^{\times m}$. Dann erhalten wir für n -mal differenzierbare bzw. glatte Funktionen das Taylorpolynom der Ordnung n von f bei x_0 bzw. die Taylorreihe:

$$x \mapsto \sum_{k=0}^n \frac{f^{(k)}(x_0)((x - x_0), \dots, (x - x_0))}{k!} \quad \text{bzw.} \quad x \mapsto \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)((x - x_0), \dots, (x - x_0))}{k!}.$$

Eine unendlich oft differenzierbare Funktion f heißt wieder reell analytisch in x_0 , wenn die entsprechende Taylorreihe auf einer Umgebung von x_0 gegen $f(x)$ konvergiert. Auf einem Banachraum X ist z.B. $\exp : \mathcal{L}(X) \rightarrow \mathcal{L}(X)$ eine analytische Funktion.

Für reelle Funktionen, also $Y = \mathbb{R}$ ergibt der Satz von Taylor:

Satz 10.29. (von Taylor in höheren Dimensionen) Sei $f : U \rightarrow \mathbb{R}$ eine auf einer offenen konvexen Teilmenge eines normierten Vektorraumes definierte $(n+1)$ -mal differenzierbare Funktion. Dann gibt es für jedes $x, x_0 \in U$ ein $\xi \in (0, 1)$, so dass

$$f(x) = \sum_{k=0}^n \frac{f^{(k)}(x_0)((x - x_0), \dots, (x - x_0))}{k!} + \frac{f^{(n+1)}(x_0 + \xi(x - x_0))((x - x_0), \dots, (x - x_0))}{(n+1)!}$$

gilt. Hierbei bezeichnen wir mit $f^{(k)}(x_0)$ bzw. $f^{(n+1)}(x_0 + \xi(x - x_0))$ die entsprechende multilineare Abbildung von $X^{\times k}$ bzw. $X^{\times (n+1)}$ nach $\mathbb{R}$. Der erste Term heißt wieder Taylorpolynom von f in x_0 der Ordnung n und der zweite Term Restglied **q.e.d.**

Kapitel 11

Nichtlineare Analysis

11.1 Der Banachsche Fixpunktsatz

In diesem Kapitel bieten wir eine kurze Einführung in die nichtlineare Analysis. Der bei weitem wichtigste Satz der nichtlinearen Analysis ist der sogenannte Banachsche Fixpunktsatz. Wir werden gleich mehrere Anwendungen kennenlernen.

Banachscher Fixpunktsatz 11.1. *Sei $f : X \rightarrow X$ eine lipschitzstetige Abbildung eines vollständigen metrischen Raumes X auf sich selber mit Lipschitzkonstante $L < 1$. Dann hat f genau einen Fixpunkt: $x \in X$ mit $f(x) = x$. Für jedes $x_0 \in X$ konvergiert die induktiv durch $x_{n+1} = f(x_n)$ definierte Folge $(x_n)_{n \in \mathbb{N}_0}$ gegen den Fixpunkt.*

Beweis: Sei $x_0 \in X$ und $(x_n)_{n \in \mathbb{N}_0}$ induktiv definiert durch $x_{n+1} = f(x_n)$. Dann gilt

$$d(x_n, x_{n+1}) = d(f(x_{n-1}), f(x_n)) \leq Ld(x_{n-1}, x_n) \leq \dots \leq L^n d(x_0, x_1).$$

Mit der Dreiecksungleichung folgt dann für $n < m$

$$\begin{aligned} d(x_n, x_m) &\leq d(x_n, x_{n+1}) + \dots + d(x_{m-1}, x_m) \\ &\leq (L^n + \dots + L^{m-1})d(x_0, x_1) \\ &= L^n \frac{1 - L^{m-n}}{1 - L} d(x_0, x_1) \leq \frac{L^n}{1 - L} d(x_0, x_1) \end{aligned}$$

Also ist $(x_n)_{n \in \mathbb{N}}$ eine Cauchyfolge und es existiert $x = \lim_{n \rightarrow \infty} x_n$. Aus der Stetigkeit von f folgt dann $f(x) = \lim_{n \rightarrow \infty} f(x_n) = \lim_{n \rightarrow \infty} x_{n+1} = x$. Also ist x ein Fixpunkt. Ist $y \in X$ ein zweiter Fixpunkt, so gilt $d(x, y) = d(f(x), f(y)) \leq L \cdot d(x, y)$. Es folgt $(1-L)d(x, y) \leq 0$ und wegen $L < 1$ auch $0 \leq d(x, y) \leq 0$. Also gilt $x = y$. **q.e.d.**

Eine Anwendung ist z.B. der Satz von Picard Lindelöf über die Existenz und Eindeutigkeit von Anfangswertproblemen von gewöhnlichen Differentialgleichungen. Eine gewöhnliche Differentialgleichung ist eine Gleichung, von der Form

$$\dot{u}(t) = f(t, u(t)) \quad \text{mit } u : I \rightarrow X \quad \text{und } f : I \times U \rightarrow X.$$

Hierbei ist $I \subset \mathbb{R}$ ein offenes Intervall, X ein normierter Vektorraum und $U \subset X$ eine offene Teilmenge. Der Punkt über u bezeichnet die Ableitung nach t . Diese Variable steht in vielen Anwendungen für die Zeit. Das Anfangswertproblem besteht aus der Suche nach einer differenzierbaren Funktion $u : I \rightarrow U$, die die Differentialgleichung erfüllt und an einem Punkt $t_0 \in I$ den Anfangswert $u(t_0) = u_0 \in U$ annimmt.

Definition 11.2. Eine Funktion f von einem metrischen Raum X in den metrischen Raum Y heißt lokal lipschitzstetig, wenn es für jedes $x_0 \in X$ eine Umgebung $U \subset X$ von x_0 gibt und eine Lipschitzkonstante $L > 0$, so dass für alle $x, x' \in U$ gilt

$$d(f(x), f(x')) \leq Ld(x, x').$$

Satz 11.3. (Lokale Existenz und Eindeutigkeit) Sei I ein offenes Intervall, $U \subset \mathbb{R}^n$ eine offene Teilmenge und $f : I \times U \rightarrow \mathbb{R}^n$ eine stetige Abbildung, die bezüglich der zweiten Variablen lokal lipschitzstetig ist, d.h. für jedes $(t_0, u_0) \in I \times U$ gibt es ein $\delta > 0$ und ein $L > 0$, so dass für alle $(t, u), (t, \tilde{u}) \in (t_0 - \delta, t_0 + \delta) \times B(u_0, \delta)$ gilt

$$\|f(t, u) - f(t, \tilde{u})\|_1 \leq L\|u - \tilde{u}\|_1.$$

Dann gibt es für jedes $(t_0, u_0) \in I \times U$ ein $\epsilon > 0$, so dass das Anfangswertproblem $\dot{u}(t) = f(t, u(t))$ mit $u(t_0) = u_0$ auf $(t_0 - \epsilon, t_0 + \epsilon)$ genau eine Lösung besitzt.

Beweis: Wir benutzen auf $\mathbb{R}^n$ die Norm $\|\cdot\|_1$. Wegen der lokalen Lipschitzstetigkeit gibt es $\delta > 0$ und $L > 0$, so dass für alle $(t, u), (t, \tilde{u}) \in [t_0 - \delta, t_0 + \delta] \times \overline{B(u_0, \delta)} \subset I \times U$ auch $\|f(t, u) - f(t, \tilde{u})\|_1 \leq L\|u - \tilde{u}\|_1$ gilt. Wegen der Stetigkeit von f ist die Abbildung

$$F : u \mapsto F(u) \text{ mit } F(u)(t) = u_0 + \left(\int_{t_0}^t f_1(s, u(s)) ds, \dots, \int_{t_0}^t f_n(s, u(s)) ds \right)$$

eine stetige Abbildung von $C([t_0 - \delta, t_0 + \delta], \overline{B(u_0, \delta)})$ nach $C([t_0 - \delta, t_0 + \delta], \mathbb{R}^n)$. Sei

$$\|f(\cdot, u_0)\|_\infty = \sup\{\|f(s, u_0)\|_1 \mid s \in [t_0 - \delta, t_0 + \delta]\}.$$

Wenn $\epsilon \leq \delta$ und $\epsilon(\|f(\cdot, u_0)\|_\infty + L\delta) \leq \delta$, dann gilt für alle $u \in C([t_0 - \epsilon, t_0 + \epsilon], \overline{B(u_0, \delta)})$ und alle $t \in [t_0 - \epsilon, t_0 + \epsilon]$ wegen Satz 8.3

$$\|F(u)(t) - u_0\|_1 = \left\| \int_{t_0}^t (f(s, u_0) + f(s, u(s)) - f(s, u_0)) ds \right\|_1 \leq \epsilon(\|f(\cdot, u_0)\|_\infty + L\delta) \leq \delta$$

Also bildet F den vollständigen metrischen Raum $C([t_0 - \epsilon, t_0 + \epsilon], \overline{B(u_0, \delta)})$ auf sich selber ab. Für $u, \tilde{u} \in C([t_0 - \epsilon, t_0 + \epsilon], \overline{B(u_0, \delta)})$ und $t \in [t_0 - \epsilon, t_0 + \epsilon]$ gilt

$$\|F(u)(t) - F(\tilde{u})(t)\|_1 \leq \left| \int_{t_0}^t \|f(s, u(s)) - f(s, \tilde{u}(s))\|_1 ds \right| \leq \epsilon L\|u - \tilde{u}\|_\infty.$$

Sei also ϵ kleiner als $\epsilon < \min \left\{ \delta, \frac{\delta}{\|f(\cdot, u_0)\|_\infty + L\delta}, \frac{1}{L} \right\} = \min \left\{ \delta, \frac{\delta}{\|f(\cdot, u_0)\|_\infty + L\delta} \right\}$.

Dann definiert die Abbildung F eine lipschitzstetige Abbildung mit Lipschitzkonstante $\epsilon \cdot L < 1$ von dem vollständigen metrischen Raum $C([t_0 - \epsilon, t_0 + \epsilon], \overline{B}(u_0, \delta))$ auf sich selber. Jeder Fixpunkt ist wegen dem Hauptsatz der Differential- und Integralrechnung stetig differenzierbar und es gilt $\dot{u}(t) = f(t, u)$ für alle $t \in (t_0 - \epsilon, t_0 + \epsilon)$ mit $u(t_0) = u_0$. Also löst u dieses Anfangswertproblem auf $(t_0 - \epsilon, t_0 + \epsilon)$. Wenn u umgekehrt auf $(t_0 - \epsilon, t_0 + \epsilon)$ dieses Anfangswertproblem löst, dann ist die Ableitung von $F(u) - u$ gleich Null, und beide Funktionen $F(u)$ und u sind bei $t = t_0$ gleich u_0 . Also stimmen beide Funktionen überein und jede Lösung des obigen Anfangswertproblems ist ein Fixpunkt von F . Also folgt die Existenz und Eindeutigkeit dieses Anfangswertproblems auf $(t_0 - \epsilon, t_0 + \epsilon)$ aus dem Banachschen Fixpunktsatz. **q.e.d.**

Satz 11.4* (Globale Existenz und Eindeutigkeit) Sei $O \subset \mathbb{R} \times \mathbb{R}^n$ eine offene Teilmenge und $f : O \rightarrow \mathbb{R}^n$ eine stetige Abbildung, die wie bei der lokalen Existenz und Eindeutigkeit lokal lipschitzstetig ist. Dann gibt es für jedes $(t_0, u_0) \in O$ genau ein maximales Intervall $(a, b) \subset \mathbb{R}$, das t_0 enthält, und auf dem das Anfangswertproblem

$$\dot{u}(t) = f(t, u) \quad \text{mit} \quad u(t_0) = u_0$$

genau eine Lösung u enthält. Das Intervall ist in dem Sinne maximal, dass an beiden Rändern, also bei a und b , eine der folgenden Bedingungen erfüllt ist:

- (i) $a = -\infty$ (bzw. $b = \infty$).
- (ii) $t \mapsto \|f(t, u(t))\|_1$ ist für alle $\epsilon > 0$ auf $(a, a + \epsilon)$ (bzw. $(b - \epsilon, b)$) unbeschränkt.
- (iii) Die Lösung u lässt sich stetig auf $[a, b)$ (bzw. $(a, b]$) fortsetzen, der Graph der Fortsetzung liegt aber nicht in O , d.h. $\lim_{t \downarrow a} (t, u(t)) \notin O$ (bzw. $\lim_{t \uparrow b} (t, u(t)) \notin O$).

Beweis*: Seien zuerst $q_1 : (a_1, b_1) \rightarrow \mathbb{R}^n$ und $q_2 : (a_2, b_2) \rightarrow \mathbb{R}^n$ Lösungen von $\dot{q}(t) = f(t, q(t))$ mit $q_1(t_1) = q_2(t_1)$ für ein $t_1 \in (a_1, b_1) \cap (a_2, b_2)$. Wenn das Infimum von

$$A = \{a \in (a_1, t_1] \cap (a_2, t_1] \mid q_1(t) = q_2(t) \text{ für alle } t \in [a, t_1]\}$$

in $(a_1, t_1] \cap (a_2, t_1]$ liegt, dann liegt es wegen der Stetigkeit von $q_1 - q_2$ in A . Dann würde A wegen dem Satz von Picard-Lindelöf eine Umgebung vom $\inf A$ enthalten, was der Definition des Infimums widerspricht. Also gilt $A = (a_1, t_1] \cap (a_2, t_1]$. Weil genauso auch

$$\{b \in [t_1, b_1) \cap [t_1, b_2) \mid q_1(t) = q_2(t) \text{ für alle } t \in [t_1, b]\} = [t_1, b_1) \cap [t_1, b_2)$$

gilt stimmen dann q_1 und q_2 auf $(a_1, b_1) \cap (a_2, b_2)$ überein. Insbesondere definieren auf der Vereinigung ihrer offenen Definitionsbereiche alle Lösungen $q : I \ni t_0 \rightarrow \mathbb{R}^n$ von

$$\dot{q}(t) = f(t, q) \quad \text{mit} \quad q(t_0) = q_0$$

auch eine Lösung q . Wir zeigen jetzt, dass diese Vereinigung der Definitionsbereiche an den Rändern die Bedingung (iii) erfüllt, wenn (i) und (ii) nicht gelten. Dann ist $\dot{q}$ auf einer Menge $(a, a + \epsilon)$ bzw. $(b - \epsilon, b)$ beschränkt und q ist dort lipschitzstetig. Für jede Folge $(t_n)_{n \in \mathbb{N}}$, die gegen a bzw. b konvergiert, konvergiert $((t_n, q(t_n))_{n \in \mathbb{N}}$ in $\mathbb{R} \times \mathbb{R}^n$ gegen den gleichen Grenzwert. Läge dieser in O , dann besäße das Anfangswertproblem

$$\dot{q}(t) = f(t, q(t)) \quad \text{mit} \quad q(a) = \lim_{t \rightarrow a+} q(t) \quad \text{bzw.} \quad q(b) = \lim_{t \rightarrow b-} q(t)$$

wegen dem Satz von Picard-Lindelöf eine Lösung auf einer Umgebung von a bzw. b , die in obiger Vereinigung der Definitionsbereiche enthalten wäre. Das zeigt (iii). **q.e.d.**

Bemerkung 11.5*. Wenn (ii) erfüllt ist, kann $t \mapsto f(t, u(t))$ nicht stetig auf $[a, a + \epsilon)$ bzw. $(b - \epsilon, b]$ fortgesetzt werden. Also können u und f nicht so stetig auf größere Definitionsbereiche fortgesetzt werden, dass a (bzw. b) im Definitionsbereich von u und $(a, u(a))$ (bzw. $(b, u(b))$) im Definitionsbereich von f liegt.

11.2 Das Lösen von nichtlinearen Gleichungen

Die Lösungen der Gleichungen von der Form

$$Ax = y, \quad A \in \mathcal{L}(V, W), \quad x \in V \quad \text{und} \quad y \in W$$

in einem (endlichdimensionalen) Vektorraum V sind in der linearen Algebra untersucht worden. Wenn A invertierbar ist, dann ist $x = A^{-1}y$ die eindeutige Lösung. In diesem Abschnitt nutzen wir das Verständnis dieser Gleichungen für Gleichungen von der Form

$$f(x) = y, \quad f : V \rightarrow W, \quad x \in V \quad \text{und} \quad y \in W$$

mit nichtlinearen Abbildungen f . Dabei nehmen wir an, dass f differenzierbar ist, und durch lineare Abbildungen angenähert werden kann. Ausgangspunkt ist die Beobachtung, dass kleine Störungen von invertierbaren linearen Abbildungen invertierbar sind.

Lemma 11.6. Seien V und W Banachräume und A ein invertierbares Element von $\mathcal{L}(V, W)$. D.h. es gibt ein Element $A^{-1} \in \mathcal{L}(W, V)$ mit $AA^{-1} = \mathbb{1}_W$ und $A^{-1}A = \mathbb{1}_V$. Dann sind alle Elemente des folgenden Balles um A invertierbar:

$$B \in B\left(A, \frac{1}{\|A^{-1}\|}\right) \subset \mathcal{L}(V, W) \quad \text{mit} \quad \|B^{-1} - A^{-1}\| \leq \frac{\|A^{-1}\|^2 \|A - B\|}{1 - \|A^{-1}\| \|A - B\|}$$

Beweis: Offenbar ist $B = A - (A - B) = A(\mathbb{1}_V - A^{-1}(A - B))$. Wegen Satz 9.61 gilt $\|A^{-1}(A - B)\| \leq \|A^{-1}\| \cdot \|A - B\| < 1$ für $B \in B\left(A, \frac{1}{\|A^{-1}\|}\right)$. Dann folgt aus der Neumannschen Reihe, dass $\mathbb{1}_V - A^{-1}(A - B)$ invertierbar ist in $\mathcal{L}(V)$ und der inverse

Operator beschränkt ist durch $\frac{1}{1-\|A^{-1}\|\|A-B\|}$. Also ist auch B invertierbar und es gilt

$$\begin{aligned} B^{-1} &= (\mathbb{1}_V - A^{-1}(A-B))^{-1} A^{-1} \quad \text{mit} \quad \|B^{-1}\| \leq \frac{\|A^{-1}\|}{1 - \|A^{-1}\|\|A-B\|}, \\ B^{-1} - A^{-1} &= ((\mathbb{1}_V - A^{-1}(A-B))^{-1} - \mathbb{1}_V) A^{-1} = A^{-1}(A-B)(\mathbb{1}_V - A^{-1}(A-B))^{-1} A^{-1} \\ \text{mit} \quad \|B^{-1} - A^{-1}\| &\leq \frac{\|A^{-1}\|^2 \|A-B\|}{1 - \|A^{-1}\| \cdot \|A-B\|}. \end{aligned} \quad \mathbf{q.e.d.}$$

Damit bilden die invertierbaren Elemente von $\mathcal{L}(V, W)$ eine offene Teilmenge.

Korollar 11.7. Für Banachräume V und W und invertierbare $A \in \mathcal{L}(V, W)$ ist

$$B \left(A, \frac{1}{\|A^{-1}\|} \right) \rightarrow \mathcal{L}(W, V), \quad B \mapsto B^{-1}$$

eine analytische Abbildung, also insbesondere unendlich oft stetig differenzierbar.

Beweis: Aus Lemma 11.6 und der Neumannschen Reihe folgt für alle $B \in B(0, \frac{1}{\|A^{-1}\|})$

$$\|(A+B)^{-1} - A^{-1} + A^{-1}BA^{-1}\| \leq \left\| A^{-1} \sum_{n=2}^{\infty} (-BA^{-1})^n \right\| \leq \frac{\|B\|^2 \|A^{-1}\|^3}{1 - \|B\| \|A^{-1}\|}.$$

Insbesondere ist $A \mapsto A^{-1}$ differenzierbar mit der Ableitung $B \mapsto -A^{-1}BA^{-1}$ bei A , und damit einmal mehr differenzierbar als $A \mapsto A^{-1}$, also unendlich oft. Für $B \in B(0, \frac{1}{\|A^{-1}\|}) \subset \mathcal{L}(V, W)$ und $t \in [0, 1]$ gilt $(A + tB)^{-1} = A^{-1} \sum_{n \in \mathbb{N}_0} t^n (-BA^{-1})^n$. **q.e.d.**

Lemma 11.8. Sei $f : U \rightarrow V$ eine bijektive Abbildung zwischen offenen Teilmengen der normierten Vektorräume $U \subset X$ und $V \subset Y$. Wenn f bei $x_0 \in U$ differenzierbar, $f'(x_0)$ in $\mathcal{L}(X, Y)$ invertierbar, und f^{-1} bei $y_0 = f(x_0)$ stetig ist, dann ist f^{-1} bei y_0 differenzierbar mit $(f^{-1})'(y_0) = (f'(x_0))^{-1}$.

Beweis: Sei $x_0 \in U$ wie im Lemma, $0 < \epsilon \leq 1$ und $\delta > 0$ so gewählt, dass

$$\|f(x) - f(x_0) - f'(x_0)(x - x_0)\| \leq \frac{\epsilon}{2} \|(f'(x_0))^{-1}\|^{-1} \cdot \|x - x_0\| \quad \text{für alle } x \in B(x_0, \delta) \subset U$$

gilt. Es folgt $\|f'(x_0)(x - x_0)\| \leq \|f(x) - f(x_0)\| + \frac{1}{2} \|(f'(x_0))^{-1}\|^{-1} \cdot \|x - x_0\|$ und damit

$$\|(f'(x_0))^{-1}\| \cdot \|f(x) - f(x_0)\| \geq \|(f'(x_0))^{-1}\| \cdot \|f'(x_0)(x_0 - x)\| - \frac{1}{2} \|x - x_0\| \geq \frac{1}{2} \cdot \|x - x_0\|.$$

Bei y_0 ist f^{-1} stetig, und $f[B(x_0, \delta)]$ eine Umgebung von $f(x_0)$. Für $x \in B(x_0, \delta)$ gilt

$$\begin{aligned} \|x - x_0 - (f'(x_0))^{-1}(f(x) - f(x_0))\| &\leq \|(f'(x_0))^{-1}\| \cdot \|f'(x_0)(x - x_0) - f(x) + f(x_0)\| \\ &\leq \frac{\epsilon}{2} \|x - x_0\| \leq \epsilon \|(f'(x_0))^{-1}\| \cdot \|f(x) - f(x_0)\|. \end{aligned}$$

Also ist f^{-1} bei $f(x_0)$ differenzierbar mit $(f^{-1})'(y_0) = (f'(x_0))^{-1}$.

q.e.d.

Satz 11.9. (Satz über die inverse Funktion) Seien X, Y Banachräume, $f : U \rightarrow Y$ eine differenzierbare Abbildung von einer offenen Teilmenge $U \subset X$ nach Y . Wenn f' bei $x_0 \in U$ stetig und $f'(x_0)$ in $\mathcal{L}(X, Y)$ invertierbar ist, dann gibt es offene Umgebungen $V \subset U$ und $W \subset Y$ von x_0 bzw. $f(x_0)$, so dass $f|_V : V \rightarrow W$ bijektiv ist mit differenzierbarer Umkehrabbildung $f^{-1} : W \rightarrow V$ mit Ableitung $y \mapsto (f'(f^{-1}(y)))^{-1}$.

Dabei ist f^{-1} auf W genauso oft (stetig) differenzierbar wie f auf V .

Beweis: Zuerst ersetzen wir f durch $h \circ f \circ g$ mit folgenden Abbildungen

$$g : X \rightarrow X, \quad x \mapsto x + x_0, \quad h : Y \rightarrow X, \quad y \mapsto (f'(x_0))^{-1}(y - f(x_0)).$$

Die Abbildungen g und h sind bijektiv und glatt und haben glatte Umkehrabbildungen. Dadurch wird $Y = X$, $x_0 = 0 = f(x_0)$, und $\underline{f'(x_0)} = \mathbb{1}_X$. Weil f' stetig bei 0 ist, gibt es ein $\delta > 0$, so dass $\|f'(x) - \mathbb{1}_X\| \leq \frac{1}{2}$ für $x \in \overline{B(0, \delta)} \subset U$ gilt. Wegen dem Schrankensatz ist für jedes $y \in X$ die Abbildung

$$F_y : \overline{B(0, \delta)} \rightarrow \overline{B(y, \frac{\delta}{2})}, \quad x \mapsto y + x - f(x)$$

lipschitzstetige mit Lipschitzkonstante $\frac{1}{2}$. Dann gilt auch für alle $x \in \overline{B(0, \delta)}$

$$\|F_y(x) - y\| = \|F_y(x) - F_y(0)\| \leq \frac{1}{2}\|x - 0\| \leq \frac{\delta}{2}.$$

Wenn y in $\overline{B(0, \frac{\delta}{2})}$ liegt, dann liegt $\overline{B(y, \frac{\delta}{2})}$ in $\overline{B(0, \delta)}$. Also definiert F_y dann eine Abbildung von $\overline{B(0, \delta)}$ auf sich selbst. Aus dem Banachschen Fixpunktsatz folgt, dass für jedes $y \in \overline{B(0, \frac{\delta}{2})}$ die Abbildung F_y auf $\overline{B(0, \delta)}$ genau einen Fixpunkt hat und der Fixpunkt in $\overline{B(y, \frac{\delta}{2})}$ liegt. Weil x genau dann ein Fixpunkt von F_y ist, wenn $f(x) = y$ ist, gibt es für alle $y \in \overline{B(0, \frac{\delta}{2})}$ auf $\overline{B(0, \delta)}$ genau eine Lösung von $f(x) = y$. Sei also $W = \overline{B(0, \frac{\delta}{2})}$ und $V = f^{-1}[W] \cap \overline{B(0, \delta)}$. Dann ist W und V als Urbild einer offenen Menge unter einer stetigen Abbildung offen und $f : V \rightarrow W$ bijektiv. Weil F_0 auf $\overline{B(0, \delta)}$ lipschitzstetig ist mit Lipschitzkonstante $\frac{1}{2}$, gilt für alle $x, x' \in \overline{B(0, \delta)}$ auch

$$\begin{aligned} \|x - x'\| &= \|F_0(x) - F_0(x') + f(x) - f(x')\| \leq \frac{1}{2}\|x - x'\| + \|f(x) - f(x')\| \\ &\text{oder auch} \quad \|x - x'\| \leq 2\|f(x) - f(x')\|. \end{aligned}$$

Also ist f^{-1} lipschitzstetig und wegen Lemma 11.6 und dem vorangehenden Lemma differenzierbar. Die letzte Aussage folgt aus dem Korollar 11.7 und Satz 10.4 (iii). **q.e.d.**

Beispiel 11.10. Die Voraussetzung der stetigen Differenzierbarkeit kann nicht abgeschwächt werden zu einfacher Differenzierbarkeit. Die Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$, $x \mapsto f(x)$

$$f(x) = \begin{cases} \frac{x}{2} + x^2 \sin(\frac{1}{x}) & \text{für } x \neq 0 \\ 0 & \text{für } x = 0 \end{cases} \quad f'(x) = \begin{cases} \frac{1}{2} + 2x \sin(\frac{1}{x}) - \cos(\frac{1}{x}) & \text{für } x \neq 0 \\ \frac{1}{2} & \text{für } x = 0 \end{cases}$$

ist differenzierbar mit $f'(0) = \frac{1}{2}$. Aber f ist in keiner Umgebung der 0 injektiv.

Korollar 11.11. (Satz über die implizite Funktion) Seien X und Y Banachräume, U eine offene Teilmenge von $X \times Y$ und $f : U \rightarrow Y$ differenzierbar. Wenn f' in $(x_0, y_0) \in U$ stetig und $\frac{\partial f}{\partial y}(x_0, y_0)$ in $\mathcal{L}(Y)$ invertierbar ist, dann gibt es offene Umgebungen V von x_0 in X , W von $f(x_0, y_0)$ in Y und O von (x_0, y_0) in U und eine differenzierbare Funktion $g : V \times W \rightarrow Y$, so dass $f^{-1}[\{z\}] \cap O = \text{Graph}(g(\cdot, z))$ für alle $z \in W$ gilt:

$$\{(x, y) \in O \mid f(x, y) = z\} = \{(x, g(x, z)) \mid x \in V\} \quad \text{für alle } z \in W.$$

Beweis: Für $F : U \rightarrow X \times Y$, $(x, y) \mapsto (x, f(x, y))$ gilt $F(x + v, y + w) - F(x, y) = (v, f(x + v, y + w) - f(x, y))$. Deshalb ist F bei allen $(x, y) \in U$ differenzierbar mit

$$F'(x, y) : X \times Y \rightarrow X \times Y, \quad (v, w) \mapsto \left(v, \frac{\partial f(x, y)}{\partial x} v + \frac{\partial f(x, y)}{\partial y} w \right).$$

Wenn $\frac{\partial f(x, y)}{\partial y}$ in $\mathcal{L}(Y)$ invertierbar ist, dann ist der inverse Operator gegeben durch

$$(F'(x, y))^{-1} : X \times Y \rightarrow X \times Y, \quad (v, w) \mapsto \left(v, \left(\frac{\partial f(x, y)}{\partial y} \right)^{-1} \left(w - \frac{\partial f(x, y)}{\partial x} v \right) \right)$$

Also erfüllt F in (x_0, y_0) die Voraussetzungen des Satzes über die inverse Funktion. Deshalb gibt es Umgebungen V von x_0 in X , W von $f(x_0, y_0)$ in W und O von (x_0, y_0) in U , so dass die Abbildung $O \rightarrow V \times W$, $(x, y) \mapsto (x, f(x, y))$ bijektiv ist und eine Umkehrabbildung besitzt. Diese Umkehrabbildung muss aber wegen der Gestalt von F von der Form $V \times W \rightarrow O$, $(x, y) \mapsto (x, g(x, y))$ sein, mit einer differenzierbaren Funktion $g : V \times W \rightarrow Y$. Insbesondere sind für alle $z \in W$ alle Lösungen $(x, y) \in O$ von $f(x, y) = z$ im Graphen von $V \rightarrow Y$, $x \mapsto g(x, z)$ enthalten. **q.e.d.**

Definition 11.12 (Untermannigfaltigkeit). Eine Teilmenge $A \subset \mathbb{R}^n$ heißt Untermannigfaltigkeit, wenn bei jedem $x \in A$ eine stetig differenzierbare Funktion $f : O \rightarrow \mathbb{R}^m$ auf einer in $\mathbb{R}^n$ offenen Menge existiert, so dass $f'(x)$ surjektiv und $f^{-1}[\{0\}] = A \cap O$ ist.

Nach einer geeigneten Permutation der Koordinaten erfüllt f bei $x \in A$ als Funktion auf $O \subset \mathbb{R}^{n-m} \times \mathbb{R}^m$ die Voraussetzungen des Satzes der impliziten Funktion. Also gilt lokal $A = \text{Graph}(g)$ für eine Funktion $g : V \rightarrow \mathbb{R}^m$ auf einer offenen Menge $V \subset \mathbb{R}^{n-m}$.

Beispiel 11.13. (i) Höhenlinien: Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ eine stetig differenzierbare Funktion, die z.B. in Abhängigkeit von Längen- und Breitengraden die Höhe über dem Meeresspiegel beschreibt. Wenn die partielle Ableitung $\frac{\partial f}{\partial y}$ in einem Punkt $(x_0, y_0) \in \mathbb{R}$ nicht verschwindet, dann gibt es eine stetig differenzierbare Funktion

$$g : (x_0 - \epsilon, x_0 + \epsilon) \times (f(x_0, y_0) - \delta, f(x_0, y_0) + \delta) \rightarrow \mathbb{R},$$

so dass für $z \in (f(x_0, y_0) - \delta, f(x_0, y_0) + \delta)$ die Höhenlinien zur Höhe z der Graph

$$\{(x, g(x, z)) \mid x \in (x_0 - \epsilon, x_0 + \epsilon)\} \quad \text{von} \quad (x_0 - \epsilon, x_0 + \epsilon) \rightarrow \mathbb{R}, \quad x \mapsto g(x, z)$$

ist. Für festes z zeigt also der Vektor $(1, \frac{\partial g}{\partial x}(x, z))$ in Richtung der Höhenlinie durch $(x, g(x, z))$ und steht senkrecht auf dem Gradienten von f bei $(x, g(x, z))$.

(ii) Hyperflächen: Sei $f : \mathbb{R}^n \rightarrow \mathbb{R}$ eine stetig differenzierbare Funktion, deren Ableitung in einem Punkt x_0 nicht verschwindet. Dann verschwindet auch mindestens eine partielle Ableitung nicht. Nach einer geeigneten Permutation der Koordinaten, können wir $\frac{\partial f}{\partial x_n}(x_0) \neq 0$ annehmen. Sei $y_0 \in \mathbb{R}^{n-1}$ der Vektor der ersten $n-1$ Koordinaten von x_0 . Dann lassen sich für $z \in (f(x_0) - \delta, f(x_0) + \delta)$ lokal die Niveaumengen $f^{-1}[\{z\}]$ durch den Graphen einer Funktion $g : B(y_0, \epsilon) \times (f(x_0) - \delta, f(x_0) + \delta) \rightarrow \mathbb{R}$ beschreiben als $\{(y, g(y, z)) \mid y \in B(y_0, \epsilon)\}$, also durch $y \in B(y_0, \epsilon) \subset \mathbb{R}^{n-1}$ parametrisiert. Für alle (y, z) in dieser Umgebung von $(y_0, f(x_0))$ ist das Bild der partiellen Ableitungen

$$\mathbb{1}_{\mathbb{R}^n} \times \frac{\partial g(y, z)}{\partial y} \in \mathcal{L}(\mathbb{R}^{n-1}, \mathbb{R}^{n-1}) \times \mathcal{L}(\mathbb{R}^{n-1}, \mathbb{R}) \simeq \mathcal{L}(\mathbb{R}^{n-1}, \mathbb{R}^{n-1} \times \mathbb{R})$$

der Kern von $f'(y, g(y, z)) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R})$, der Tangentialraum an die Niveauflächen heißt.

Definition 11.14. Eine unendlich oft (stetig) differenzierbare bijektive Abbildung mit unendlich oft (stetig) differenzierbarer Umkehrabbildung heißt Diffeomorphismus.

Beispiel 11.15. Polarkoordinaten Die Abbildung

$$(0, \infty) \times \mathbb{R}/2\pi\mathbb{Z} \rightarrow \mathbb{R}^2 \setminus \{(0, 0)\}, \quad (r, \varphi) \mapsto (r \cos \varphi, \sin \varphi)$$

heißt Polarkoordinaten von $\mathbb{R}^2$. Offenbar ist diese Abbildung unendlich oft stetig differenzierbar. Die Umkehrabbildung ist dann gegeben durch

$$(x, y) \mapsto (r, \varphi) \quad \text{mit} \quad r = \sqrt{x^2 + y^2} \quad \text{und} \\ \varphi = \begin{cases} \arccos\left(\frac{x}{r}\right) & \text{für } y \geq 0 \\ -\arccos\left(\frac{x}{r}\right) & \text{für } y < 0 \end{cases} \quad \text{mod } 2\pi \quad \begin{cases} \arcsin\left(\frac{y}{r}\right) & \text{für } x \geq 0 \\ \pi - \arcsin\left(\frac{y}{r}\right) & \text{für } x < 0. \end{cases}$$

Also ist diese Abbildung ein Diffeomorphismus von $(0, \infty) \times \mathbb{R}/2\pi\mathbb{Z}$ nach $\mathbb{R}^2 \setminus \{(0, 0)\}$. Hier beschreibt $\mathbb{R}/2\pi\mathbb{Z}$ den Raum aller Äquivalenzklassen von $\mathbb{R}$, wobei

$$x \sim y \Leftrightarrow x - y \in 2\pi\mathbb{Z}.$$

Dieser Raum ist offenbar lokal diffeomorph zu $\mathbb{R}$, weil in jedem Intervall dessen Länge kleiner ist als 2π , verschiedene Elemente verschiedene Äquivalenzklassen repräsentieren. Deshalb sind die Einschränkungen der Abbildung $\mathbb{R} \rightarrow \mathbb{R}/2\pi\mathbb{Z}$ auf beliebige offene Intervalle mit Längen nicht größer als 2π , die jedes Element auf die entsprechende Äquivalenzklasse abbilden, Diffeomorphismen.

11.3 Lagrangemultiplikatoren

Ziel dieses Abschnittes ist es die lokalen Extremwerte von Funktionen auf solchen Teilmengen eines normierten Vektorraumes X zu bestimmen, die die Nullstellen von endlich vielen reellen differenzierbaren Funktionen bilden. Wir sprechen dann von Zwangsbedingungen, wegen denen nur die Punkte in diesen Nullstellenmengen in Betracht kommen. Diese Situation ist recht allgemein und kommt in vielen Anwendungen der Wirtschaftswissenschaften vor. Dieses Verfahren ist die Grundlage für die nichtlineare Optimierung, in der man nach Extremwerten auf Teilmengen eines Banachraumes sucht. Darauf aufbauend wird in der konvexen Analysis nach Bedingungen gesucht, die die Existenz und Eindeutigkeit von solchen Extremwertproblemen garantieren.

Definition 11.16. Sei $g = (g_1, \dots, g_m) : U \rightarrow \mathbb{R}^m$ eine Abbildung auf einer offenen Teilmenge eines normierten Vektorraumes X . Für $x_0 \in U$ heißt $A = g^{-1}[\{g(x_0)\}] = \{x \in U \mid g(x) = g(x_0)\}$ Niveaumenge. Sie ist abgeschlossen in U , wenn g stetig ist.

Definition 11.17. Die Niveaumenge A heißt in einem Punkt $x_0 \in A$ stetig differenzierbar (glatt), wenn es eine offene Umgebung U von $x_0 \in X$ und eine stetig differenzierbare (glatte) bijektive Funktion Φ von einer offenen Teilmenge O eines normierten Vektorraumes Y auf $A \cap U$ gibt, so dass $\Phi'(\Phi^{-1}(x_0))$ ein linearer Isomorphismus von Y auf den Kern von $g'(x_0)$ d.h. $\{x \in X \mid g'(x_0)(x) = 0\}$ ist. Ein solcher Punkt x_0 heißt kritischer Punkt der Einschränkung $f|_{(A \cap U)}$ einer differenzierbaren Funktion $f : U \rightarrow \mathbb{R}$ auf die Niveaumenge, wenn $\Phi^{-1}(x_0)$ ein kritischer Punkt von $f \circ \Phi$ ist.

In Lemma 12.34 werden wir zeigen, dass im $\mathbb{R}^d$ eine Niveaumenge in allen Punkten eine Untermannigfaltigkeit (Definition 11.12) ist, in denen sie stetig differenzierbar ist.

Aufgrund der Definition ist ein Punkt $x_0 \in U$, an dem die Niveaumenge A stetig differenzierbar ist, höchstens dann ein lokaler Extremwert der Einschränkung $f|_A$ einer differenzierbaren Funktion $f : U \rightarrow \mathbb{R}$, wenn er ein kritischer Punkt im Sinne dieser Definition ist. Deshalb kommen als die lokalen Extremwerte von $f|_A$ neben diesen kritischen Punkten nur solche Punkte in Betracht, an denen die Niveaumenge nicht stetig differenzierbar ist. Sie werden Singularitäten genannt. Im folgenden werden also einerseits diese kritischen Punkte und andererseits die Singularitäten charakterisiert.

Satz 11.18. Sei $f : U \rightarrow \mathbb{R}$ eine differenzierbare und $g : U \rightarrow \mathbb{R}^m$ eine stetig differenzierbare Funktion auf einer offenen Teilmenge U eines normierten Vektorraumes X und die Niveaumenge $A = g^{-1}[\{g(x_0)\}]$ bei $x_0 \in U$ stetig differenzierbar. Dann ist x_0 genau dann ein kritischer Punkt von der Einschränkung $f|_A$ von f auf A , wenn es reelle Zahlen $\lambda_1, \dots, \lambda_m$ gibt, so dass $f'(x_0) = \lambda_1 g'_1(x_0) + \dots + \lambda_m g'_m(x_0)$ gilt.

Definition 11.19. Die Zahlen $\lambda_1, \dots, \lambda_m$ heißen Lagrangemultiplikatoren.

Beweis: Ein stetig differenzierbarer Punkt $x_0 \in U$ von $A = g^{-1}[\{g(x_0)\}]$ ist genau dann ein kritischer Punkt von $f|_A$, wenn $\Phi^{-1}(x_0)$ ein kritischer Punkt von $f \circ \Phi$ ist.

Hierbei ist $\Phi : O \rightarrow U \cap A$ wie in der Definition 11.17 bijektiv und stetig differenzierbar, und $\Phi'(\Phi^{-1}(x_0))$ ein linearer Isomorphismus auf den Kern von $g'(x_0)$. Das ist äquivalent dazu, dass $f'(x_0)\Phi'(\Phi^{-1}(x_0))$ verschwindet, oder dazu, dass $f'(x_0)$ auf dem Bild von $\Phi'(\Phi^{-1}(x_0))$, also auf dem Kern $Y = \{x \in X \mid g'(x_0)(x) = 0\}$ von $g'(x_0)$ verschwindet.

Für alle $r = 1, \dots, R = \dim g'(x_0)[X]$ gibt es dann einen kleinsten Index $l_r > l_{r-1}$, so dass $(g'_1(x_0), \dots, g'_{l_r}(x_0))$ den Rang r hat genauso wie $(g'_{l_1}(x_0), \dots, g'_{l_r}(x_0))$. Die Komponenten von $g'(x_0)$ sind also Linearkombinationen von denen von $\tilde{g}'(x_0) = (g'_{l_1}(x_0), \dots, g'_{l_R}(x_0))$, und Y ist auch der Kern von $\tilde{g}'(x_0)$. Diese lineare Abbildung bildet geeignet gewählte $z_1, \dots, z_R \in X$ auf die Standardbasis $e_1, \dots, e_R$ von $\mathbb{R}^R$ ab. Dann liegt $x - \sum_{r=1}^R g'_{l_r}(x_0)(x)z_r$ für alle $x \in X$ in Y . Aus $f'(x_0)|_Y = 0$ folgt also

$$f'(x_0)(x) = \sum_{r=1}^R \lambda_r g'_{l_r}(x_0)(x) \quad \text{für alle } x \in X \quad \text{mit} \quad \lambda_r = f'(x_0)(z_r).$$

Umgekehrt verschwindet $\lambda_1 g'_1(x_0) + \dots + \lambda_m g'_m(x_0)$ für $(\lambda_1, \dots, \lambda_m) \in \mathbb{R}^m$ auf Y . **q.e.d.**

Lemma 11.20. *Sei $g : U \rightarrow \mathbb{R}^m$ eine stetig differenzierbare Abbildung auf einer offenen Menge U eines normierten Vektorraumes X . Dann ist für alle $r \in \mathbb{N}_0$ die Menge $\{x \in U \mid \dim(g'(x)[X]) \geq r\}$ entweder leer oder offen.*

Beweis: Sei $x_0 \in \{x \in U \mid \dim(g'(x)[X]) \geq r\}$. Dann gibt es Indizes $1 \leq l_1 < \dots < l_r \leq m$, so dass $(g'_{l_1}(x_0), \dots, g'_{l_r}(x_0))$ eine surjektive Abbildung nach $\mathbb{R}^r$ ist. Sie bildet geeignet gewählte Elemente $z_1, \dots, z_r \in X$ auf die Standardbasis $e_1, \dots, e_r$ von $\mathbb{R}^r$ ab. Weil g stetig differenzierbar ist, sind die Funktionen $x \mapsto g'_{l_i}(x)(z_j)$ auf U stetig. Deshalb enthält $\{x \in U \mid \dim(g'(x)[X]) \geq r\}$ die offenen Umgebung von x_0 in U , auf der die Determinante der $r \times r$ Matrix $(g'_{l_i}(x)(z_j))_{1 \leq i, j \leq r}$ nicht verschwindet. **q.e.d.**

Typischerweise werden die Zwangsbedingungen glatte Funktionen sein. Aber selbst dann sind die Niveaumengen nicht immer glatt.

Satz 11.21. (Rangsatz) *Für eine stetig differenzierbare Abbildung $g : U \rightarrow \mathbb{R}^m$ auf einer offenen Menge U eines Banachraumes X ist die Niveaumenge $A = g^{-1}[\{g(x_0)\}]$ in allen lokalen Maxima x_0 von der Funktion $x \mapsto \dim(g'(x)[X])$ stetig differenzierbar.*

Beweis: Wenn $g'(x_0) \in \mathcal{L}(X, \mathbb{R}^m)$ surjektiv ist, dann gibt es Elemente $z_1, \dots, z_m \in X$, die durch $g'(x_0)$ auf die Standardbasis $e_1, \dots, e_m$ von $\mathbb{R}^m$ abgebildet werden. Sei Z der von $z_1, \dots, z_m$ aufgespannte Unterraum von X , und Y der Kern von $g'(x_0)$. Für alle $x \in X$ gilt $g'(x_0)(x - g'_1(x_0)(x)z_1 - \dots - g'_m(x_0)(x)z_m) = 0$. Deshalb ist die Abbildung

$$I : X \rightarrow Y \times Z, x \mapsto (x - g'_1(x_0)(x)z_1 - \dots - g'_m(x_0)(x)z_m, g'_1(x_0)(x)z_1 + \dots + g'_m(x_0)(x)z_m)$$

ein linearer stetiger Isomorphismus von X nach $Y \times Z$ mit $I^{-1}(y, z) = y + z$, der U auf eine offene Teilmenge von $O \subset Y \times Z$ abbildet. Dann ist $g \circ I^{-1} : O \rightarrow \mathbb{R}^m$ stetig

differenzierbar, und die partielle Ableitung nach z ist bei $I(x_0)$ in $\mathcal{L}(Z, \mathbb{R}^m)$ invertierbar. In diesem Fall folgt dann die Aussage aus dem Satz über die implizite Funktion.

Wenn die Dimension $\dim(g'(x)[X])$ bei x_0 gleich r und auf einer Umgebung nicht größer als r ist, dann gibt es Indizes $1 \leq l_1 < l_2 \dots < l_r \leq m$, so dass die linearen Abbildungen $g'_{l_1}(x_0), \dots, g'_{l_r}(x_0)$ der Komponenten von $\tilde{g}(x) = (g_{l_1}(x), \dots, g_{l_r}(x))$ bei $x = x_0$ linear unabhängig sind. Also ist schon gezeigt, dass die Niveaumenge $\tilde{A} = \tilde{g}^{-1}[\{\tilde{g}(x_0)\}]$ bei x_0 stetig differenzierbar ist. Wegen Lemma 11.20 sind $g'_{l_1}(x), \dots, g'_{l_r}(x)$ auf einer Umgebung von x_0 linear unabhängig. Weil $\dim(g'(x)[X])$ auf einer Umgebung von x_0 nicht größer als r ist, stimmen auf einer Umgebung von x_0 wie im Beweis von Satz 11.18 die Kerne von $g'(x)$ und $\tilde{g}'(x)$ überein. Wegen Satz 11.18 besteht $\tilde{A}$ nur aus kritischen Punkten von g , und $\tilde{A}$ stimmt in dieser Umgebung mit A überein. **q.e.d.**

Weil die Funktion $\dim(g'(x)[X])$ nur die endlich vielen Werte $0, \dots, m$ annehmen kann, gibt es in jeder offenen Menge U ein lokales Maximum. Die Menge aller solcher lokalen Maxima ist wegen Lemma 11.20 sogar offen und dicht in U . Mit Hilfe von dem Lemma 11.20 und dem Rangsatz können wir die Singularitäten der Niveaumengen von stetig differenzierbaren Funktionen $g : U \rightarrow \mathbb{R}^m$ dadurch bestimmen, dass wir

- (i) wie im Beweis des Rangsatzes maximal viele Komponenten $\tilde{g}$ von g auswählen, deren Ableitungen $\tilde{g}'(x)$ an möglichst vielen Punkten surjektiv sind,
- (ii) und dann die Punkte bestimmen, an denen diese Ableitung nicht surjektiv sind.
Das sind die Nullstellen der Determinante aus dem Beweis von Lemma 11.20.

Beispiel 11.22. (i) $g : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad (x, y) \mapsto g(x, y) = x^2 + y^2.$

Der Gradient $\nabla g(x, y) = (2x, 2y)$ von g verschwindet nur bei $(x, y) = 0$. Also sind alle Niveaumengen $g(x, y) = g_0$ mit $g_0 \neq g(0, 0) = 0$ glatte 1-dimensionale Teilmengen von $\mathbb{R}^2$. Es sind jeweils die Kreise mit Radius $\sqrt{g_0}$ um den Nullpunkt. Für $g_0 = 0$ besteht die Niveaumenge nur aus $\{0\}$. Der Nullpunkt ist eine Singularität der Niveaumenge.

(ii) $g : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad (x, y) \mapsto g(x, y) = x^2 - y^2.$

Der Gradient $\nabla g(x, y) = (2x, -2y)$ verschwindet wieder nur bei $(x, y) = (0, 0)$. Also sind alle Niveaumengen $g(x, y) = g_0$ mit $g_0 \neq g(0, 0) = 0$ glatte eindimensionale Teilmengen von $\mathbb{R}^2$. Es sind jeweils zwei Hyperebenen. Die Niveaumenge $g(x, y) = x^2 - y^2 = (x - y)(x + y) = 0$ besteht aber aus zwei Geraden $y = x$ und $y = -x$, die sich im Nullpunkt schneiden. Diese Niveaumenge hat also im Nullpunkt eine Singularität, weil sich dort zwei glatte Teilmengen schneiden. Man spricht von einem Doppelpunkt.

(iii) $g : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad (x, y) \mapsto g(x, y) = y^2 - x^3.$

Der Gradient $\nabla g(x, y) = (-3x^2, 2y)$ verschwindet wieder nur im Nullpunkt. Also sind wieder alle Niveaumengen $g(x, y) = g_0$ mit $g_0 \neq g(0, 0) = 0$ glatte eindimensionale Teilmengen von $\mathbb{R}^2$. Die Niveaumenge $g(x, y) = y^2 - x^3 = 0$ besteht aus zwei Lösungen $y = \pm\sqrt{x^3}$ mit $x \geq 0$, die sich bei $(x, y) = 0$ einer gemeinsamen Halbgeraden parallel zu der x -Achse annähern. Man nennt deshalb die Singularität im Nullpunkt eine Spitze.

Satz 11.23* (Whitney) Jede nicht leere abgeschlossene Teilmenge $A \subset \mathbb{R}^n$ ist die Nullstellenmenge einer glatten reellen Funktion auf $\mathbb{R}^n$.

Beweis*: Sei $A \subset \mathbb{R}^n$ eine nicht leere abgeschlossene Menge. Dann besitzt $\mathbb{Q}^n \cap O$ mit $O = \mathbb{R}^n \setminus A$ eine Abzählung durch eine Folge $(x_m)_{m \in \mathbb{N}}$. Für jedes $m \in \mathbb{N}$ ist $B(x_m, r_m) = \bigcup_{\{r > 0 \mid B(x_m, r) \subset O\}} B(x_m, r)$ ein offener Ball in O . Die Vereinigung dieser Bälle $(B(x_m, r_m))_{m \in \mathbb{N}}$ ist dann eine offene Teilmenge von O . Jeder Punkt $x \in O$ ist in einem Ball $B(x, r) \subset O$ enthalten, und $B(x, \frac{r}{2})$ enthält ein x_m mit $r_m > \frac{r}{2}$. Dann liegt x in $B(x_m, r_m)$ und die Vereinigung der Bälle $(B(x_m, r_m))_{m \in \mathbb{N}}$ ist O . Die Funktion

$$\psi : \mathbb{R}^n \rightarrow [0, \infty), \quad x \mapsto \begin{cases} \exp(\frac{1}{x^2-1}) & \text{für } \|x\| < 1 \\ 0 & \text{für } \|x\| > 1 \end{cases}$$

ist unendlich oft stetig differenzierbar, und für alle $m \in \mathbb{N}_0$ sind alle partiellen Ableitungen höchstens m -ter Ordnung beschränkt durch ein $C_m > 0$. Alle partiellen Ableitungen höchstens m -ter Ordnung von $\psi_m(x) = \frac{1}{C_m} \left(\frac{\min\{1, r_m\}}{2} \right)^m \psi\left(\frac{x-x_m}{r_m}\right)$ sind beschränkt durch 2^{-m} . Für jedes Monom D in $\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n}$ folgt aus Satz 8.38 induktiv im Grad von D , dass die Summe der partiellen Ableitungen $(\sum D\psi_m)_{m \in \mathbb{N}}$ auf ganz $\mathbb{R}^n$ gleichmäßig gegen die entsprechende partielle Ableitung Df von $f = \sum_{m \in \mathbb{N}} \psi_m$ konvergiert. Deshalb ist f glatt und die Nullstellenmenge von f ist gleich A . **q.e.d.**

Beispiel 11.24. (i) Die sogenannte Cantormenge ist definiert als das Komplement $A = [0, 1] \setminus I$ folgender offenen Teilmenge von $[0, 1]$:

$$I = \bigcup_{n \in \mathbb{N}_0} \bigcup_{(z_1, \dots, z_n) \in \{0, 1\}^n} \left(\frac{1}{3^{n+1}} + \sum_{l=1}^n \frac{2z_l}{3^l}, \frac{2}{3^{n+1}} + \sum_{l=1}^n \frac{2z_l}{3^l} \right) = \left(\frac{1}{3}, \frac{2}{3}\right) \cup \left(\frac{1}{9}, \frac{2}{9}\right) \cup \left(\frac{7}{9}, \frac{8}{9}\right) \dots$$

Sie ist offenbar eine abgeschlossene Teilmenge von $[0, 1]$ und wegen dem Satz von Whitney die Niveaumenge einer glatten reellen Funktion f auf $\mathbb{R}$. Weil in jeder Umgebung von jedem Punkt von A sowohl Elemente von A als auch Elemente von I enthalten sind, verschwinden alle Ableitungen von f auf A . Alle Punkte von A sind Singularitäten.

(ii) Der Sierpinski Teppich ist das Komplement $A = [0, 1]^2 \setminus I$ der offenen Teilmenge

$$I = \bigcup_{n \in \mathbb{N}_0} \left[\bigcup_{(z_1, \dots, z_n) \in \{0, 1, 2\}^n} \left(\frac{1}{3^{n+1}} + \sum_{l=1}^n \frac{z_l}{3^l}, \frac{2}{3^{n+1}} + \sum_{l=1}^n \frac{z_l}{3^l} \right) \right]^2 = \left(\frac{1}{3}, \frac{2}{3}\right)^2 \cup \left(\left(\frac{1}{9}, \frac{2}{9}\right) \cup \left(\frac{4}{9}, \frac{5}{9}\right) \cup \left(\frac{7}{9}, \frac{8}{9}\right)\right)^2 \dots$$

Die Menge I ist eine Teilmenge von dem kartesischen Produkt des Komplementes der Cantormenge in (i) mit sich selber. Wenn also x oder y zu der Cantormenge gehören, dann sind $\{x\} \times [0, 1]$ und $[0, 1] \times \{y\}$ Teilmengen von A . Deshalb ist A zusammenhängend. Wegen dem Satz von Whitney ist A die Nullstellenmenge einer glatten Funktion auf $\mathbb{R}^2$. Wieder enthält jede Umgebung von jedem Punkt von A sowohl Elemente von A als auch Elemente von I , so dass alle Ableitungen von f auf A verschwinden. Alle Punkte von A sind Singularitäten.

Kapitel 12

Das Lebesgueintegral auf dem $\mathbb{R}^d$

12.1 Treppenfunktionen

Zunächst führen wir die Quader im $\mathbb{R}^d$ ein.

Definition 12.1. Ein Quader ist ein d -faches kartesisches Produkt von Intervallen

$$Q = I_1 \times \dots \times I_d = \{x \in \mathbb{R}^d \mid x_1 \in I_1, \dots, x_d \in I_d\} \subset \mathbb{R}^d.$$

Die Intervalle $I_1, \dots, I_d \subset \mathbb{R}$ können den linken bzw. rechten Rand enthalten oder nicht und nur aus einem Punkt bestehen. Wenn $I_1, \dots, I_d$ beschränkt sind heißt Q endlich.

Für jeden solchen Quader definieren wir das Volumen als das Produkt der Längen von $I_1, \dots, I_d$. Wir bezeichnen es mit $\mu(Q)$. Endliche Quader haben endliches Volumen.

Definition 12.2. Eine Teilmenge $A \subset \mathbb{R}^d$ heißt Nullmenge, wenn es für jedes $\epsilon > 0$ eine Folge von endlichen Quadern $(Q_n)_{n \in \mathbb{N}}$ im $\mathbb{R}^d$ gibt, mit

$$A \subset \bigcup_{n=1}^{\infty} Q_n \quad \text{und} \quad \sum_{n=1}^{\infty} \mu(Q_n) \leq \epsilon.$$

Beispiel 12.3. Die Vereinigung von abzählbar vielen Quadern ohne Volumen ist eine Nullmenge. Insbesondere ist jede abzählbare Teilmenge von $\mathbb{R}^d$ eine Nullmenge.

Lemma 12.4. Eine höchstens abzählbare Vereinigung von Nullmengen ist wieder eine Nullmenge.

Beweis: Sei $\bigcup_{n=1}^{\infty} A_n$ eine höchstens abzählbare Vereinigung von Nullmengen. Dann besitzt für jedes $\epsilon > 0$ jedes A_n eine Überdeckung von Quadern, deren gesamtes Volumen nicht größer ist als $\epsilon \cdot 2^{-n}$. Die höchstens abzählbare Vereinigung dieser jeweils höchstens abzählbar vielen Quader ist wegen Satz 2.40 eine höchstens abzählbare Menge von Quader mit einem Volumen nicht größer als $\sum_{n=1}^{\infty} \epsilon 2^{-n} = \epsilon$. Also wird $\bigcup_{n=1}^{\infty} A_n$ von abzählbar vielen Quadern überdeckt, deren Volumen nicht größer ist als ϵ . **q.e.d.**

Definition 12.5. Eine Treppenfunktion ist eine endliche Linearkombination von charakteristischen Funktionen von endlichen Quadern, d.h. der Funktionen, die bei Punkten innerhalb eines Quaders gleich 1, und außerhalb des Quaders gleich 0 sind.

Proposition 12.6. Endlich viele Quader Q_i lassen sich so in endlich viele paarweise disjunkte Quader P_j zerlegen, so dass jeder Q_i eine Vereinigung von endlich vielen P_j 's ist. Jede Treppenfunktion ist eine endliche Linearkombination von charakteristischen Funktionen von paarweise disjunkten endlichen Quadern.

Beweis: Für endlich viele Intervalle $I_1, \dots, I_n$ ordnen wir alle endlichen Intervallgrenzen der Reihe nach an $-\infty < x_1 < \dots < x_m < \infty$. Dadurch erhalten wir eine Zerlegung

$$\mathbb{R} = (-\infty, x_1) \cup \{x_1\} \cup (x_1, x_2) \cup \{x_2\} \cup \dots \cup (x_{m-1}, x_m) \cup \{x_m\} \cup (x_m, \infty)$$

von $\mathbb{R}$ in endlich viele paarweise disjunkte Intervalle, so dass jedes der Intervalle $I_1, \dots, I_n$ eine Vereinigung von endlich vielen dieser Intervalle ist. Für jeden Faktor des kartesischen Produktes sind durch n Quader $Q_1, \dots, Q_n$ auch n Intervalle vorgegeben und damit auch eine solche Zerlegung von $\mathbb{R}$. Die kartesischen Produkte von jeweils einem dieser Intervalle aus den Zerlegungen aller d Faktoren des kartesischen Produktes bilden Quader, die paarweise disjunkt sind und deren Vereinigung gleich $\mathbb{R}^d$ ist. Dabei ist jeder der Quader $Q_1, \dots, Q_n$ eine Vereinigung von endlich vielen von diesen paarweise disjunkten Quadern. Eine endliche Linearkombination von $\chi_{Q_1}, \dots, \chi_{Q_n}$ nimmt auf jedem dieser paarweise disjunkten Quader der Zerlegung genau einen Wert an und ist eine endliche Linearkombination von charakteristischen Funktionen von paarweise disjunkten endlichen Quadern. **q.e.d.**

Als nächstes wollen wir das Integral von Treppenfunktionen definieren. Zunächst definieren wir für jede charakteristische Funktion χ_Q eines Quaders das Integral

$$\int \chi_Q d\mu = \mu(Q).$$

Proposition 12.7. Sei f eine Treppenfunktion und seien

$$f = \sum_i c_i \chi_{Q_i} = \sum_j d_j \chi_{R_j}$$

zwei Zerlegungen in endliche Linearkombinationen von charakteristischen Funktionen von endlichen Quadern. Dann gilt

$$\int f d\mu = \sum_i c_i \mu(Q_i) = \sum_j d_j \mu(R_j).$$

Beweis: Wir zerlegen alle diese endlichen Quader Q_i und R_j wie im vorangehenden Beweis beschrieben in endlich viele paarweise disjunkte endliche Quader P_k . Es folgt

$$f|_{P_k} = \sum_{\{i|Q_i \supset P_k\}} c_i = \sum_{\{j|R_j \supset P_k\}} d_j, \quad \chi_{Q_i} = \sum_{\{k|P_k \subset Q_i\}} \chi_{P_k}, \quad \chi_{R_j} = \sum_{\{k|P_k \subset R_j\}} \chi_{P_k}.$$

Hierbei sei die Summe über eine leere Menge gleich Null. Die Gesamtlänge einer disjunkten Vereinigung von Intervallen ist gleich der Summe der Intervalllängen. Wegen dem Distributivgesetz folgt dann

$$\begin{aligned}\mu(Q_i) &= \sum_{\{k|P_k \subset Q_i\}} \mu(P_k), \quad \mu(R_j) = \sum_{\{k|P_k \subset R_j\}} \mu(P_k), \\ \sum_i c_i \mu(Q_i) &= \sum_i c_i \sum_{\{k|P_k \subset Q_i\}} \mu(P_k) = \sum_k \sum_{\{i|Q_i \supset P_k\}} c_i \mu(P_k) \\ &= \sum_k \sum_{\{j|R_j \supset P_k\}} d_j \mu(P_k) = \sum_j d_j \sum_{\{k|P_k \subset R_j\}} \mu(P_k) = \sum_j d_j \mu(R_j). \quad \text{q.e.d.}\end{aligned}$$

Wegen dieser Proposition definiert das Integral $f \mapsto \int f d\mu$ eine lineare Abbildung von dem Raum aller Treppenfunktionen nach $\mathbb{R}$.

Proposition 12.8. *Seien f und g zwei Treppenfunktionen mit $f \geq g$. Dann gilt*

$$\int f d\mu \geq \int g d\mu.$$

Beweis: Wir zerlegen die beiden Vereinigungen von Quadern der Treppenfunktion f und der Treppenfunktion g in eine gemeinsame disjunkte Vereinigung von Quadern. Auf jedem der Quader ist f größer oder gleich g . Deshalb gilt das auch für die Summen, die die entsprechenden Integrale berechnen. q.e.d.

12.2 Lebesgueintegrale Funktionen auf dem $\mathbb{R}^d$

Satz 12.9. *Sei $(f_n)_{n \in \mathbb{N}}$ eine monoton wachsende Folge von Treppenfunktionen, deren Integrale $(\int f_n d\mu)_{n \in \mathbb{N}}$ beschränkt sind. Dann ist folgende Menge eine Nullmenge:*

$$\{x \in \mathbb{R}^d \mid (f_n(x))_{n \in \mathbb{N}} \text{ konvergiert nicht} \}.$$

Beweis: Sei $M > 0$ eine obere Schranke von $(\int (f_n - f_1) d\mu)_{n \in \mathbb{N}} = (\int f_n d\mu - \int f_1 d\mu)_{n \in \mathbb{N}}$:

$$\int (f_n - f_1) d\mu \leq M \quad \text{für alle } n \in \mathbb{N}.$$

Dann sind für alle $\epsilon > 0$ und alle $n \in \mathbb{N}$ die Mengen

$$S_{n,\epsilon} = \{x \in \mathbb{R}^d \mid f_n(x) - f_1(x) \geq \frac{M}{\epsilon}\} = \{x \in \mathbb{R}^d \mid f_n(x) \geq \frac{M}{\epsilon} + f_1(x)\}$$

eine endliche Vereinigungen von Quadern, die als Mengen jeweils im Nachfolger $S_{n+1,\epsilon}$ enthalten sind. Wegen Proposition 12.6 ist das relative Komplement eines Quaders in einem anderen Quader wieder eine disjunkte Vereinigung von Quadern ist. Dann ist

$$\bigcup_{n=1}^{\infty} S_{n,\epsilon} = S_{1,\epsilon} \cup \bigcup_{n=1}^{\infty} (S_{n+1,\epsilon} \setminus S_{n,\epsilon})$$

eine abzählbare Vereinigung von disjunkten Quadern. Weil $f_n - f_1$ nichtnegative Funktionen sind, ist $\frac{\epsilon}{M}(f_n - f_1)$ größer oder gleich $\chi_{S_{n,\epsilon}}$. Also erfüllt das Gesamtvolumen von $S_{n,\epsilon}$

$$\int \chi_{S_{n,\epsilon}} d\mu \leq \int \frac{\epsilon}{M}(f_n - f_1) d\mu = \frac{\epsilon}{M} \int (f_n - f_1) d\mu \leq \epsilon.$$

Wegen der Monotonie ist dann auch das Gesamtvolumen der abzählbaren Vereinigung $\bigcup_{n=1}^{\infty} S_{n,\epsilon}$ nicht größer als ϵ . Also ist die Schnittmenge S eine Nullmenge:

$$S = \bigcap_{\epsilon > 0} \left(\bigcup_{n=1}^{\infty} S_{n,\epsilon} \right) = \{x \in \mathbb{R}^d \mid (f_n(x))_{n \in \mathbb{N}} \text{ konvergiert nicht} \}. \quad \text{q.e.d.}$$

Die Komplemente von Nullmengen werden fast überall genannt (bzw. a.e. für almost everywhere). Der vorangehende Satz besagt also, dass jede monoton wachsende Folge von Treppenfunktionen mit beschränkten Integralen fast überall konvergiert.

Satz 12.10. *Für jede Nullmenge $A \subset \mathbb{R}^d$ gibt es eine monoton wachsende Folge $(f_n)_{n \in \mathbb{N}}$ von Treppenfunktionen mit beschränkten Integralen $(\int f_n d\mu)_{n \in \mathbb{N}}$, so dass A in der Menge enthalten ist, auf der die Folge $(f_n)_{n \in \mathbb{N}}$ nicht konvergiert.*

Beweis: Sei A eine Nullmenge. Dann gibt es für jedes $n \in \mathbb{N}$ eine Überdeckung von A mit abzählbar vielen Quadern, deren Gesamtvolumen nicht größer ist als 2^{-n} . Sei nun $(Q_n)_{n \in \mathbb{N}}$ eine Abzählung der Vereinigung aller dieser Quader. Dann gehört jeder Punkt von A zu unendlich vielen Quadern. Also definiert die Reihe $(\sum \chi_{Q_n})_{n \in \mathbb{N}}$ eine monoton wachsende Folge von Treppenfunktionen, die auf A nicht konvergiert. Die Integrale $(\sum \int \chi_{Q_n} d\mu)_{n \in \mathbb{N}}$ sind beschränkt durch $\sum_{n \in \mathbb{N}} 2^{-n} = 1$. q.e.d.

Für jede monoton wachsende Folge $(f_n)_{n \in \mathbb{N}}$ von Treppenfunktionen mit beschränkten Integralen $(\int f_n d\mu)_{n \in \mathbb{N}}$ können wir jetzt den Grenzwert fast überall definieren:

$$f(x) = \begin{cases} \lim_{n \rightarrow \infty} f_n(x) & \text{wenn } (f_n(x))_{n \in \mathbb{N}} \text{ beschränkt ist} \\ 0 & \text{wenn } (f_n(x))_{n \in \mathbb{N}} \text{ nicht beschränkt ist.} \end{cases}$$

Wir wollen $\int f d\mu$ als $\lim_{n \rightarrow \infty} \int f_n d\mu$ definieren. Die folgenden Lemmata zeigen, dass diese Definition nur von der fast überall definierten Funktion f abhängt.

Lemma 12.11. *Sei $(f_n)_{n \in \mathbb{N}}$ eine monoton fallende Folge von Treppenfunktionen, die fast überall gegen Null konvergiert. Dann ist $(\int f_n d\mu)_{n \in \mathbb{N}}$ eine Nullfolge.*

Beweis: Sei Q_0 ein kompakter Quader, außerhalb dessen f_1 verschwindet. Für jedes $n \in \mathbb{N}$ haben alle Quader Q , auf denen f_n negativ ist, das Maß $\mu(Q) = 0$, weil $(f_n)_{n \in \mathbb{N}}$ auf diesen Q gegen eine negative Zahl konvergiert. Insbesondere gilt $\int \chi_{\mathbb{R}^d \setminus Q_0} f_n d\mu = 0$ für alle $n \in \mathbb{N}$, und es genügt $\lim_{n \rightarrow \infty} \int \chi_{Q_0} f_n d\mu = 0$ zu zeigen. Für jedes $n \in \mathbb{N}$ sei A_n die Menge der Unstetigkeitsstellen von f_n in Q_0 , also der $x \in Q_0$, bei denen

f_n lokal nicht konstant ist. Dann ist A_n als eine endliche Vereinigung von Quadern ohne Volumen eine Nullmenge. Dann ist auch $A = \bigcup_{n=1}^{\infty} A_n$ eine Nullmenge. Sei B die Nullmenge aller $x \in Q_0$, an denen $(f_n)_{n \in \mathbb{N}}$ nicht gegen Null konvergiert. Für jedes $\epsilon > 0$ gibt es eine Überdeckung $\bigcup_{m=1}^{\infty} Q_m \supset (A \cup B)$ durch Quader, deren Gesamtvolumen nicht größer ist als $\frac{\epsilon}{3}$. Indem wir die Kanten der Quader mit positivem Volumen um den Faktor $\sqrt[d]{2}$ verlängern, dabei aber den Mittelpunkt festhalten, und die Quader mit verschwindendem Volumen durch größere offene Quader mit Volumen $\frac{\epsilon}{3} 2^{-m}$ ersetzen, erhalten wir auch eine solche Überdeckung $\bigcup_{m=1}^{\infty} Q_m \supset (A \cup B)$ durch offene Quader, deren Gesamtvolumen nicht größer ist als ϵ . Für jeden Punkt $x \in Q_0 \setminus (A \cup B)$ gibt es ein $N_x \in \mathbb{N}$ mit $f_{N_x}(x) \leq \epsilon$. Weil $(f_n)_{n \in \mathbb{N}}$ monoton fallend ist, gilt $f_n(x) \leq \epsilon$ für alle $n \geq N_x$. Weil alle f_{N_x} bei den Punkten von $Q_0 \setminus (A \cup B)$ lokal konstant sind, gibt es eine Überdeckung von offenen Quadern $(R_x)_{x \in Q_0 \setminus (A \cup B)}$ von $Q_0 \setminus (A \cup B)$, so dass $f_n \leq \epsilon$ auf R_x für $n \geq N_x$ gilt. Zusammen mit $(Q_m)_{m \in \mathbb{N}}$ bilden die $(R_x)_{x \in Q_0 \setminus (A \cup B)}$ eine offene Überdeckung von Q_0 . Sie besitzt eine endliche Teilüberdeckung, weil Q_0 kompakt ist. Sei N das Maximum der entsprechenden endlich vielen N_x 's. Dann gilt für alle $n \geq N$

$$0 \leq \int \chi_{Q_0} f_n d\mu \leq \epsilon (\max\{f_1(x) \mid x \in Q_0\} + \mu(Q_0)).$$

Auf den Quadern $(Q_m)_{m \in \mathbb{N}}$ schätzen wir dabei f durch $\max\{f_1(x) \mid x \in Q_0\}$ ab und auf den endlich vielen der $(R_x)_{x \in Q_0 \setminus (A \cup B)}$ durch ϵ . Es folgt $\lim_{n \rightarrow \infty} \int f_n d\mu = 0$. **q.e.d.**

Lemma 12.12. *Seien f und g fast überall definierte Grenzwerte von monoton wachsenden Folgen $(f_n)_{n \in \mathbb{N}}$ und $(g_n)_{n \in \mathbb{N}}$ von Treppenfunktionen mit beschränkten $(\int f_n d\mu)_{n \in \mathbb{N}}$ und $(\int g_n d\mu)_{n \in \mathbb{N}}$. Wenn fast überall $f \geq g$ gilt, dann gilt auch*

$$\lim_{n \rightarrow \infty} \int f_n d\mu \geq \lim_{n \rightarrow \infty} \int g_n d\mu.$$

Beweis: Für jedes feste $m \in \mathbb{N}$ erfüllen die Funktionenfolgen

$$((g_m - f_n)^+)_{n \in \mathbb{N}} = \left(\frac{1}{2}(g_m - f_n + |g_m - f_n|) \right)_{n \in \mathbb{N}}$$

die Voraussetzungen von dem vorangehenden Lemma, weil fast überall $g_m - f \leq g - f \leq 0$ gilt. Deshalb konvergieren die entsprechenden Integrale gegen Null. Wegen $g_m - f_n \leq (g_m - f_n)^+$ folgt aus Proposition 12.8 und Lemma 12.11

$$\int g_m d\mu - \lim_{n \rightarrow \infty} \int f_n d\mu \leq 0 \quad \text{und damit auch} \quad \lim_{m \rightarrow \infty} \int g_m d\mu \leq \lim_{n \rightarrow \infty} \int f_n d\mu. \text{q.e.d.}$$

Aus Lemma 12.12 folgt, dass wir das Integral auf die Grenzwerte von monoton wachsenden Folgen von Treppenfunktionen mit beschränkten Integralen konsistent fortsetzen

können. Seien nämlich $(f_n)_{n \in \mathbb{N}}$ und $(g_n)_{n \in \mathbb{N}}$ monoton wachsenden Folgen von Treppenfunktionen mit beschränkten Integralen $(\int f_n d\mu)_{n \in \mathbb{N}}$ und $(\int g_n d\mu)_{n \in \mathbb{N}}$, deren Grenzwerte fast überall übereinstimmen, dann können wir Lemma 12.12 sowohl auf diese Folge, als auch auf die vertauschten Folgen anwenden und erhalten

$$\lim_{n \rightarrow \infty} \int f_n d\mu \geq \lim_{n \rightarrow \infty} \int g_n d\mu \geq \lim_{n \rightarrow \infty} \int f_n d\mu.$$

Definition 12.13. Sei $L^1(\mathbb{R}^d)$ die Menge der Äquivalenzklassen von fast überall definierten Funktionen f , für die es monoton wachsende Folgen $(g_n)_{n \in \mathbb{N}}$ und $(h_n)_{n \in \mathbb{N}}$ von Treppenfunktionen mit beschränkten Integralen $(\int g_n d\mu)_{n \in \mathbb{N}}$ und $(\int h_n d\mu)_{n \in \mathbb{N}}$ gibt, und

$$f = \lim_{n \rightarrow \infty} g_n - \lim_{n \rightarrow \infty} h_n$$

fast überall gilt. Hierbei werden zwei Funktionen miteinander identifiziert, wenn sie fast überall miteinander übereinstimmen.

Satz 12.14. (Eigenschaften der lebesgueintegrierbaren Funktionen)

(i) $L^1(\mathbb{R}^d)$ ist ein Vektorraum über $\mathbb{R}$ und das Integral über Treppenfunktionen induziert eine lineare Abbildung

$$\int : L^1(\mathbb{R}^d) \rightarrow \mathbb{R}, \quad f \rightarrow \int f d\mu$$

(ii) Wenn $f \in L^1(\mathbb{R}^d)$ fast überall nicht negativ ist, dann gilt $\int f d\mu \geq 0$.

(iii) Wenn $f \in L^1(\mathbb{R}^d)$, dann ist auch $|f| \in L^1(\mathbb{R}^d)$ mit $\left| \int f d\mu \right| \leq \int |f| d\mu$.

Beweis: (i) Die monoton wachsende Folgen $(g_n)_{n \in \mathbb{N}}$ und $(h_n)_{n \in \mathbb{N}}$ von Treppenfunktionen mit beschränkten Integralen können wir addieren und mit $\lambda \in \mathbb{R}_0^+$ multiplizieren. Für die Multiplikation mit -1 vertauschen wir sie. Um die Wohldefiniertheit des Integrals zu zeigen, seien $(g_n)_{n \in \mathbb{N}}$, $(h_n)_{n \in \mathbb{N}}$, $(\tilde{g}_n)_{n \in \mathbb{N}}$ und $(\tilde{h}_n)_{n \in \mathbb{N}}$ solche Folgen, wobei fast überall

$$g(x) - h(x) = \tilde{g}(x) - \tilde{h}(x) \quad \text{mit}$$

$$g = \lim_{n \rightarrow \infty} g_n, \quad h = \lim_{n \rightarrow \infty} h_n, \quad \tilde{g} = \lim_{n \rightarrow \infty} \tilde{g}_n, \quad \tilde{h} = \lim_{n \rightarrow \infty} \tilde{h}_n$$

gilt. Weil dann fast überall auch $g(x) + \tilde{h}(x) = \tilde{g}(x) + h(x)$ gilt, folgt aus Lemma 12.12

$$\lim_{n \rightarrow \infty} \int (g_n + \tilde{h}_n) d\mu = \lim_{n \rightarrow \infty} \int (\tilde{g}_n + h_n) d\mu.$$

Daraus folgt wegen der Linearität des Integrals von Treppenfunktionen

$$\int (g-h)d\mu = \lim_{n \rightarrow \infty} \int g_n d\mu - \lim_{n \rightarrow \infty} \int h_n d\mu = \lim_{n \rightarrow \infty} \int \tilde{g}_n d\mu - \lim_{n \rightarrow \infty} \int \tilde{h}_n d\mu = \int (\tilde{g} - \tilde{h})d\mu.$$

Das definiert die Abbildung $\int$ von $L^1(\mathbb{R})$ nach $\mathbb{R}$. Die Linearität folgt aus den Rechenregeln für Folgen und der Linearität des Integrals auf Treppenfunktionen.

(ii) Seien g bzw. h die fast überall definierte Grenzwert der monoton wachsenden Folgen von Treppenfunktionen $(g_n)_{n \in \mathbb{N}}$ und $(h_n)_{n \in \mathbb{N}}$ mit beschränkten Integralen. Aus $g - h \geq 0 \iff g \geq h$ folgt mit Lemma 12.12 $\int g d\mu \geq \int h d\mu \iff \int (g - h) d\mu \geq 0$.

(iii) Sei f fast überall die Differenz $g - h$ der Grenzwerte der monoton wachsenden Folgen $(g_n)_{n \in \mathbb{N}}$ und $(h_n)_{n \in \mathbb{N}}$ von Treppenfunktionen mit beschränkten Integralen. Aus $\min\{g_n, h_n\} \leq g_n \leq g_{n+1}$ und $\min\{g_n, h_n\} \leq h_n \leq h_{n+1}$ folgt $\min\{g_n, h_n\} \leq \min\{g_{n+1}, h_{n+1}\}$, und aus $\max\{g_{n+1}, h_{n+1}\} \geq g_{n+1} \geq g_n$ und $\max\{g_{n+1}, h_{n+1}\} \geq h_{n+1} \geq h_n$ folgt $\max\{g_{n+1}, h_{n+1}\} \geq \max\{g_n, h_n\}$. Also sind $(\tilde{g}_n)_{n \in \mathbb{N}} = (\max\{g_n, h_n\})_{n \in \mathbb{N}}$ und $(\tilde{h}_n)_{n \in \mathbb{N}} = (\min\{g_n, h_n\})_{n \in \mathbb{N}}$ monoton wachsende Folgen von Treppenfunktionen. Für die nicht negativen $g_n - \tilde{h}_1$ und $h_n - \tilde{h}_1$ gilt $0 \leq \tilde{h}_n - \tilde{h}_1 \leq \tilde{g}_n - \tilde{h}_1 = \max\{g_n - \tilde{h}_1, h_n - \tilde{h}_1\} \leq g_n - \tilde{h}_1 + h_n - \tilde{h}_1$. Also haben $(\tilde{g}_n)_{n \in \mathbb{N}}$ und $(\tilde{h}_n)_{n \in \mathbb{N}}$ beschränkte Integrale. Dann ist $|f|$ fast überall die Differenz der entsprechenden Grenzwerte $\tilde{g} - \tilde{h}$. Deshalb ist $|f| \in L^1(\mathbb{R}^d)$. Wegen (ii) folgt dann aus $-|f| \leq f \leq |f|$

$$-\int |f| d\mu \leq \int f d\mu \leq \int |f| d\mu \quad \text{bzw.} \quad \left| \int f d\mu \right| \leq \int |f| d\mu. \quad \text{q.e.d.}$$

Satz 12.15. Eine beschränkte Funktion f , die außerhalb einer beschränkten Menge verschwindet und deren Unstetigkeitsstellen eine Nullmenge bilden, gehört zu $L^1(\mathbb{R}^d)$.

Beweis: Wir wählen einen Quader $Q_0 = [a_1, b_1] \times \dots \times [a_d, b_d]$, außerhalb dessen f verschwindet. Wir zerlegen für alle $n \in \mathbb{N}$ und alle $i = 1, \dots, d$ das Intervall $[a_i, b_i]$:

$$[a_i, b_i] = \left[a_i, a_i + \frac{b_i - a_i}{2^n} \right] \cup \left(a_i + \frac{b_i - a_i}{2^n}, a_i + 2 \frac{b_i - a_i}{2^n} \right] \cup \dots \cup \left(a_i + (2^n - 1) \frac{b_i - a_i}{2^n}, b_i \right].$$

Die kartesischen Produkte dieser Zerlegungen ergeben eine Zerlegung $\mathcal{P}_n$ von Q_0 in eine Vereinigung von 2^{nd} paarweise disjunkten Quadern. Dann sei f_n die Treppenfunktion, die auf jedem der 2^{nd} Quader gleich dem Infimum der entsprechenden Funktionswerte von f ist. Offenbar ist $(f_n)_{n \in \mathbb{N}}$ eine monoton wachsende Folge von Treppenfunktionen, deren Integrale durch $\|f\|_\infty \cdot \mu(Q_0)$ beschränkt sind. An allen Punkten $x_0 \in \mathbb{R}^d$, an denen f stetig ist, gibt es für jedes $\epsilon > 0$ ein $\delta > 0$, so dass $f(x) \in B(f(x_0), \epsilon)$ aus $x \in B(x_0, \delta)$ folgt. Dann gibt es auch ein $N \in \mathbb{N}$, so dass der Durchmesser von Q_0 kleiner ist als $2^N \delta$. Für alle $n \geq N$ ist dann der Teilquader Q_n der 2^{nd} Teilquader von Q , der x_0 enthält, in $B(x_0, \delta)$ enthalten. Deshalb gilt dann

$$f(x_0) - \epsilon \leq \inf\{f(x) \mid x \in Q_n\} = f_n(x_0) \leq f(x_0).$$

Also konvergiert $(f_n(x_0))$ gegen $f(x_0)$. Weil die Menge der Unstetigkeitsstellen von f eine Nullmenge ist, konvergiert $(f_n)_{n \in \mathbb{N}}$ dann fast überall gegen f . q.e.d.

Satz 12.16*: Sei $f \in L^1(\mathbb{R}^d)$ mit $\int |f| d\mu = 0$. Dann ist f fast überall gleich Null.

Beweis:* Seien $(\tilde{g}_n)_{n \in \mathbb{N}}$ und $(\tilde{h}_n)_{n \in \mathbb{N}}$ monoton wachsende Folgen von Treppenfunktionen mit beschränkten Integralen, so dass fast überall gilt

$$f(x) = \lim_{n \rightarrow \infty} \tilde{g}_n(x) - \lim_{n \rightarrow \infty} \tilde{h}_n(x).$$

Für alle $n \in \mathbb{N}$ seien $g_n = \max\{\tilde{g}_n, \tilde{h}_n\}$ und $h_n = \min\{\tilde{g}_n, \tilde{h}_n\}$.

Dann gilt fast überall $|f(x)| = \lim_{n \rightarrow \infty} g_n(x) - \lim_{n \rightarrow \infty} h_n(x)$.

Wegen $\int |f| d\mu = 0$ gilt also auch $\lim_{n \rightarrow \infty} \int g_n d\mu = \lim_{n \rightarrow \infty} \int h_n d\mu$.

Für $\epsilon, \delta > 0$ sei $N \in \mathbb{N}$ so gewählt, dass $\lim_{n \rightarrow \infty} \int g_n d\mu - \int h_m d\mu \leq \epsilon \delta$.

für alle $m \geq N$ gilt. Wir definieren g als den fast überall definierten Grenzwert $\lim_{n \rightarrow \infty} g_n$. Weil $(g_n)_{n \in \mathbb{N}}$ und $(h_n)_{n \in \mathbb{N}}$ monoton wachsend sind, ist die Menge

$$\{x \in \mathbb{R}^d \mid |f(x)| > \delta\} \quad \text{für ein festes } m \in \mathbb{N} \text{ in den Mengen}$$

$$\{x \in \mathbb{R}^d \mid g(x) - h_m(x) > \delta\} = \{x \in \mathbb{R}^d \mid g_n(x) - h_m(x) > \delta \text{ für ein } n \in \mathbb{N}\}$$

enthalten. Sei also $A_1 = \{x \in \mathbb{R}^d \mid g_1(x) - h_m(x) > \delta\}$ und für $n = 2, \dots$

$$A_n = \{x \in \mathbb{R}^d \mid g_n(x) - h_m(x) > \delta \text{ und } g_{n-1}(x) - h_m(x) \leq \delta\}. \quad \text{Dann gilt}$$

$$\{x \in \mathbb{R}^d \mid g(x) - h_m(x) > \delta\} = \bigcup_{n=1}^{\infty} A_n \quad \text{und} \quad \sum_{n=1}^{\infty} \mu(A_n) \leq \lim_{n \rightarrow \infty} \frac{1}{\delta} \int (g_n - h_m) d\mu \leq \epsilon$$

für $m \geq N$. Also ist für alle $\delta > 0$ die Menge $\{x \in \mathbb{R}^d \mid g(x) - \lim_{m \rightarrow \infty} h_m(x) > \delta\}$ eine Nullmenge. Weil die abzählbare Vereinigung von Nullmengen wieder eine Nullmenge ist, folgt, dass die Menge $\bigcup_{n \in \mathbb{N}} \{x \in \mathbb{R}^d \mid g(x) - \lim_{m \rightarrow \infty} h_m(x) > \frac{1}{n}\}$ eine Nullmenge ist. Also ist fast überall $g(x) \leq h(x) = \lim_{m \rightarrow \infty} h_m(x)$. Aufgrund der Konstruktion gilt aber fast überall $g(x) \geq h(x)$. Also ist fast überall $|f(x)| = g(x) - h(x) = 0$. **q.e.d.**

12.3 Das Riemann- und das Lebesgueintegral

In diesem Abschnitt setzen wir das Riemannintegral mit dem Lebesgueintegral in Beziehung setzen. Für $d > 1$ sind die riemannintegrierbaren Funktionen f auf einem kompakten Quader Q_0 dadurch charakterisiert, dass das Unterintegral, also das Supremum aller Integrale von Treppenfunktionen nicht größer als f , mit dem Oberintegral, also dem Infimum aller Integrale von Treppenfunktionen nicht kleiner als f , übereinstimmt. Damit sie nach oben und unten durch Treppenfunktionen beschränkt sind, müssen sie beschränkt sein. Zunächst charakterisieren wir die riemannintegrierbaren Funktionen.

Satz 12.17. (*Lebesguekriterium*) Eine beschränkte Funktion auf einem kompakten Quader $Q_0 = [a_1, b_1] \times \dots \times [a_d, b_d]$ ist genau dann riemannintegrabel, wenn ihre Unstetigkeitsstellen eine Nullmenge bilden. Insbesondere sind alle riemannintegrablen Funktionen auch lebesgueintegrabel und die beiden Integrale stimmen überein.

Beweis: Wir zerlegen für alle $n \in \mathbb{N}$ und alle $i = 1, \dots, d$ das Intervall $[a_i, b_i]$:

$$[a_i, b_i] = [a_i, a_i + \frac{b_i - a_i}{2^n}] \cup (a_i + \frac{b_i - a_i}{2^n}, a_i + 2\frac{b_i - a_i}{2^n}] \cup \dots \cup (a_i + (2^n - 1)\frac{b_i - a_i}{2^n}, b_i].$$

Das entspricht für $d = 1$ der Partition $p_n \in \mathcal{P}[a, b]$ aus dem Beweis von Korollar 8.24. Die kartesischen Produkte dieser Zerlegungen ergeben eine Zerlegung $\mathcal{P}_n$ von Q_0 in eine Vereinigung von 2^{nd} paarweise disjunkte Quadern. Für alle $f \in B(Q_0, \mathbb{R})$ seien

$$\begin{aligned} f_n(x) &= \inf \{f(y) \mid y \in Q \text{ mit } Q \in \mathcal{P}_n \text{ und } x \in Q\} \\ F_n(x) &= \sup \{f(y) \mid y \in Q \text{ mit } Q \in \mathcal{P}_n \text{ und } x \in Q\} \end{aligned}$$

Dann sind $(f_n)_{n \in \mathbb{N}}$ und $(F_n)_{n \in \mathbb{N}}$ monoton wachsende bzw. monoton fallende Folgen von Treppenfunktionen mit beschränkten Integralen. Im Beweis von Satz 12.15 haben wir gezeigt, dass $(f_n)_{n \in \mathbb{N}}$ an den Punkten, an denen f stetig ist, gegen f konvergiert. Dort konvergiert auch $(-F_n)_{n \in \mathbb{N}}$ gegen $-f$, und deshalb $(F_n - f_n)_{n \in \mathbb{N}}$ gegen Null. Wenn die Unstetigkeitsstellen von f also eine Nullmenge bilden, dann konvergiert wegen Lemma 12.11 $(\int (F_n - f_n) d\mu)_{n \in \mathbb{N}}$ gegen Null und f ist riemannintegrabel.

Wenn umgekehrt f riemannintegrabel ist, dann gibt es Folgen $\tilde{f}_n \leq f \leq \tilde{F}_n$ von Treppenfunktionen mit $\lim_{n \rightarrow \infty} \int (\tilde{F}_n - \tilde{f}_n) d\mu = 0$. Die Minima $F_n = \min\{\tilde{F}_1, \dots, \tilde{F}_n\} \geq f$ sind eine monoton fallende Folge und die Maxima $f_n = \max\{\tilde{f}_1, \dots, \tilde{f}_n\} \leq f$ eine monoton wachsende Folge von Treppenfunktionen mit $\lim_{n \rightarrow \infty} \int (F_n - f_n) d\mu = 0$. Die Differenz $(F_n - f_n)_{n \in \mathbb{N}}$ ist dann eine monoton fallende Folge von nichtnegativen Treppenfunktionen, deren Grenzwert ein verschwindendes Integral hat. Wegen Satz 12.16 (siehe auch Korollar 12.20) konvergieren dann $(F_n)_{n \in \mathbb{N}}$ und $(f_n)_{n \in \mathbb{N}}$ fast überall gegen f . Die Unstetigkeitsstellen aller Treppenfunktionen beider Folgen sind abzählbare Vereinigungen von Nullmengen und damit eine Nullmenge $A \subset Q_0$. Von jedem Punkt $x_0 \in Q_0 \setminus A$ sind alle Quadern der Treppenfunktionen, die x_0 enthalten, eine Umgebung. Wenn f bei $x_0 \in Q_0 \setminus A$ unstetig ist, dann gilt für ein $\epsilon > 0$ und alle $\delta > 0$

$$\sup\{f(x) \mid x \in B(x_0, \delta) \cap Q_0\} - \inf\{f(x) \mid x \in B(x_0, \delta) \cap Q_0\} \geq \epsilon.$$

Deshalb ist nur dann fast überall $\lim(F_n - f_n) = 0$, wenn die Unstetigkeitsstellen von f in $Q_0 \setminus A$ eine Nullmenge bilden, und damit auch die Unstetigkeitsstellen von f . **q.e.d.**

Der Beweis zeigt auch, dass $f \in B(\mathbb{R}^d, \mathbb{R})$ genau dann riemannintegrabel ist, wenn f fast überall gleich den Grenzwerten sowohl einer monoton fallenden als auch einer monoton wachsenden Folge von Treppenfunktionen mit beschränkten Integralen ist.

Beispiel 12.18. (i) Sei $(r_n)_{n \in \mathbb{N}}$ eine Abzählung aller rationalen Zahlen in $(0, 1)$. Dann ist für $0 < \epsilon < 1$ das Komplement $A = [0, 1] \setminus O$ in $[0, 1]$ der Teilmenge

$$O = \bigcup_{n=1}^{\infty} I_n \quad \text{mit} \quad I_n = (r_n - 2^{-(n+1)}\epsilon, r_n + 2^{-(n+1)}\epsilon) \cap (0, 1)$$

keine Nullmenge, weil alle I_n höchstens die Länge $\epsilon 2^{-n}$ haben und $\sum_{n=1}^{\infty} \epsilon 2^{-n} = \epsilon < 1$. Die monoton wachsende Folge von Treppenfunktionen mit beschränkten Integralen

$$(h_n)_{n \in \mathbb{N}} \quad \text{mit} \quad h_n = \chi_{I_1 \cup \dots \cup I_n} \quad \text{für alle } n \in \mathbb{N}$$

konvergiert punktweise gegen χ_O , und $\chi_A = \chi_{[0,1]} - \chi_O$ liegt in $L^1(\mathbb{R})$. Weil die $\mathbb{Q} \cap (0, 1)$ dicht in $[0, 1]$ liegen, ist χ_A an allen Punkten in A unstetig. Also ist χ_A eine lebesgueintegrierbare Funktion, die nicht riemannintegrierbar ist. Insbesondere ist O ein Beispiel für eine offene Menge, so dass das Komplement A von O im Abschluss $\bar{O} = [0, 1]$ von O , also der Rand von O , keine Nullmenge ist und positives Maß $\int \chi_A d\mu \geq 1 - \epsilon$ hat.

(ii) Sei $p \in \mathbb{N} \setminus \{1, 2\}$ und χ die Funktion $\chi : \mathbb{R} \rightarrow \mathbb{R}$, $x \mapsto \chi(x)$ mit

$$\chi(x) = \begin{cases} 0 & \text{wenn } x \in \bigcup_{m \in \mathbb{Z}} (mp + 1, mp + 2) \\ 1 & \text{sonst.} \end{cases}$$

Dann definiert die Folge $(x \mapsto \chi_{[0,1]}(x) \prod_{k=1}^n \chi(p^k x))_{n \in \mathbb{N}}$ eine monoton fallende Folge von Treppenfunktionen mit beschränkten Integralen. Der Grenzwert ist lebesguesintegrierbar und die charakteristische Funktion χ_A der abgeschlossenen Menge $A \subset [0, 1]$ mit

$$[0, 1] \setminus A = \bigcup_{n \in \mathbb{N}_0} \bigcup_{z_1, \dots, z_n \in \{0, 2, \dots, p-1\}} \left(\sum_{k=1}^n \frac{z_k}{p^k} + \frac{1}{p^{n+1}}, \sum_{k=1}^n \frac{z_k}{p^k} + \frac{2}{p^{n+1}} \right).$$

Für $p = 3$ ist A die Cantormenge aus Beispiel 11.24 (i). Alle Punkte, deren p -adische Darstellung nur Ziffern in $\{0, 2, \dots, p-1\}$ hat, liegen in A . Wenn man bei der p -adischen Darstellung fordert, dass die Ziffernfolgen nicht gegen $p-1$ konvergieren, dann sind das alle Punkte von A bis auf die abzählbar vielen linken Intervallgrenzen der offenen disjunkten Intervalle von $[0, 1] \setminus A$. Wegen der $(p-1)$ -adischen Darstellung ist A dann gleichmächtig zu $[0, 1]$ und damit nicht abzählbar. Indem man eine oder zwei der Ziffern der p -adischen Darstellung der Punkte von A in eine 1 umwandelt, erhält man in jeder Umgebung eines Punktes von A einen Punkt der offenen Menge $[0, 1] \setminus A$. Also ist A auch die Menge der Unstetigkeitsstellen von χ_A . Das Integral von χ_A ist

$$\int \chi_A d\mu = 1 - \sum_{n=0}^{\infty} \frac{1}{p} \left(\frac{p-1}{p} \right)^n = 1 - 1 = 0.$$

Also ist A eine nicht abzählbare Nullmenge, und χ_A riemannintegrierbar.

12.4 Konvergenzsätze

In diesem Abschnitt werden drei Aussagen darüber bewiesen, wann Grenzwertbildungen mit der Integration vertauschen. Als erstes beweisen wir die Konvergenz von monotonen Folgen mit beschränkten Integralen, die unserem Zugang zugrunde liegt.

Satz 12.19. (*Satz der monotonen Konvergenz von Beppo Levi*) Sei $(f_n)_{n \in \mathbb{N}}$ eine monotone Folge in $L^1(\mathbb{R}^d)$ deren Integrale $(\int f_n d\mu)_{n \in \mathbb{N}}$ beschränkt sind. Dann konvergiert $(f_n)_{n \in \mathbb{N}}$ fast überall gegen eine Funktion f in $L^1(\mathbb{R}^d)$ und es gilt

$$\lim_{n \rightarrow \infty} \int f_n d\mu = \int f d\mu.$$

Beweis: Sei $(f_n)_{n \in \mathbb{N}}$ eine monotone Folge in $L^1(\mathbb{R}^d)$ mit beschränkten Integralen. Durch Übergang zu $(\pm f_n)_{n \in \mathbb{N}}$ können wir annehmen, dass $(f_n)_{n \in \mathbb{N}}$ eine monoton wachsende Folge von Funktionen mit beschränkten Integralen ist. Für alle $n \in \mathbb{N}$ seien $(\tilde{g}_{nm})_{m \in \mathbb{N}}$ und $(\tilde{h}_{nm})_{m \in \mathbb{N}}$ monoton wachsende Folgen von Treppenfunktionen mit beschränkten Integralen, so dass fast überall gilt

$$f_n = \lim_{m \rightarrow \infty} \tilde{g}_{nm} - \lim_{m \rightarrow \infty} \tilde{h}_{nm}.$$

Die entsprechenden Folgen der Integrale $(\int \tilde{g}_{nm} d\mu)_{m \in \mathbb{N}}$ und $(\int \tilde{h}_{nm} d\mu)_{m \in \mathbb{N}}$ konvergieren. Für alle $n \in \mathbb{N}$ sei $M(n) \in \mathbb{N}$ so gewählt, dass für alle $m, m' \geq M(n)$ gilt

$$\left| \int \tilde{h}_{nm} d\mu - \int \tilde{h}_{nm'} d\mu \right| \leq 2^{-n}.$$

Wir definieren jetzt induktiv in $n \in \mathbb{N}$ neue Folgen $(h_{nm})_{m \in \mathbb{N}}$ und $(g_{nm})_{m \in \mathbb{N}}$ durch

$$h_{nm} = \begin{cases} h_{n-1m} & \text{für } m < M(n) \\ \tilde{h}_{nm} - \tilde{h}_{nM(n)} + h_{n-1m} & \text{für } m \geq M(n) \end{cases} \quad \text{und} \quad g_{nm} = \tilde{g}_{nm} - \tilde{h}_{nM(n)} + h_{n-1m}$$

mit $h_{0m} = 0$ für alle $m \in \mathbb{N}$. Weil die Folgen $(\tilde{h}_{nm})_{m \in \mathbb{N}}$ monoton wachsend sind, folgt

$$h_{n-1m} \leq h_{nm}, \quad 0 \leq h_{nm}, \quad h_{nm} \leq h_{nm+1} \quad \text{induktiv in } n \in \mathbb{N} \text{ für alle } m \in \mathbb{N}.$$

Mit den Folgen $(\tilde{g}_{nm})_{m \in \mathbb{N}}$ sind auch die Folgen $(g_{nm})_{m \in \mathbb{N}}$ monoton wachsend. Aufgrund der Wahl von $M(n)$ sind die Integrale $(\int h_{nm} d\mu - \int h_{n-1m} d\mu)_{m \in \mathbb{N}}$ beschränkt durch 2^{-n} , und $(\int h_{nm} d\mu)_{n, m \in \mathbb{N}}$ beschränkt durch $1 = \sum_{n=1}^{\infty} 2^{-n}$. Weil $(\int f_n d\mu)_{n \in \mathbb{N}}$ beschränkt sind, und fast überall $g_{nm} \leq \lim_{m \rightarrow \infty} \tilde{g}_{nm} - \tilde{h}_{nM(n)} + h_{n-1m} \leq f_n + \lim_{m \rightarrow \infty} h_{nm}$ gilt, sind auch die Integrale $(\int g_{nm} d\mu)_{n, m \in \mathbb{N}}$ beschränkt. Für alle $n \in \mathbb{N}$ bzw. $m \in \mathbb{N}$ seien

$$\begin{aligned} g_n &= \lim_{m \rightarrow \infty} g_{nm} & \text{und} & \quad h_n = \lim_{m \rightarrow \infty} h_{nm}, \\ \tilde{g}_m &= \max\{g_{1m}, \dots, g_{mm}\} \leq f_m + h_m & \text{und} & \quad \tilde{h}_m = \max\{h_{1m}, \dots, h_{mm}\} = h_{mm}. \end{aligned}$$

Dann gilt für alle $n \in \mathbb{N}$ auch fast überall $f_n = g_n - h_n$. Aufgrund $h_{n-1m} \leq h_{nm}$ ist $(h_n)_{n \in \mathbb{N}}$ fast überall monoton wachsend und $(g_n)_{n \in \mathbb{N}} = (f_n + h_n)_{n \in \mathbb{N}}$ auch. Dann sind $(\tilde{h}_m)_{m \in \mathbb{N}}$ und $(\tilde{g}_m)_{m \in \mathbb{N}}$ monoton wachsende Folgen von Treppenfunktionen, deren Integrale beschränkt sind durch $\int \tilde{h}_m d\mu \leq 1$ und $\int \tilde{g}_m d\mu \leq \int f_m d\mu + 1$. Seien $\tilde{g}$ und $\tilde{h}$ die entsprechenden fast überall definierten Grenzwerte. Für $n \leq m$ gilt fast überall

$$g_{nm} \leq \tilde{g}_m \leq \max\{g_1, \dots, g_m\} = g_m \quad \text{und} \quad h_{nm} \leq \tilde{h}_m \leq \max\{h_1, \dots, h_m\} = h_m.$$

Im Komplement der abzählbar vielen Nullmengen, auf denen eine dieser Ungleichungen nicht gilt, oder $\lim_{m \rightarrow \infty} \tilde{g}_m$ bzw. $\lim_{m \rightarrow \infty} \tilde{h}_m$ nicht existiert, folgt im Grenzwert $m \rightarrow \infty$

$$\tilde{g}_n \leq g_n \leq \tilde{g} \quad \text{und} \quad \tilde{h}_n \leq h_n \leq \tilde{h}.$$

Also konvergieren $(f_n)_{n \in \mathbb{N}} = (g_n - h_n)_{n \in \mathbb{N}}$ fast überall gegen $\tilde{g} - \tilde{h} \in L^1(\mathbb{R}^d)$, und $(\int f_n d\mu)_{n \in \mathbb{N}} = (\int (g_n - h_n) d\mu)_{n \in \mathbb{N}}$ wegen Satz 12.14 (ii) gegen $\int (\tilde{g} - \tilde{h}) d\mu$. Daraus folgt

$$\int f d\mu = \lim_{m \rightarrow \infty} \int (\tilde{g}_m - \tilde{h}_m) d\mu = \lim_{n \rightarrow \infty} \int (g_n - h_n) d\mu = \lim_{n \rightarrow \infty} \int f_n d\mu. \quad \mathbf{q.e.d.}$$

Korollar 12.20. (Norm $\|\cdot\|_1$)

Auf $L^1(\mathbb{R}^d)$ definiert $\|\cdot\|_1 : L^1(\mathbb{R}^d) \rightarrow \mathbb{R}, \quad f \mapsto \|f\|_1 = \int |f| d\mu$ eine Norm.

Beweis: Aus Satz 12.14 (i)-(ii) folgt die Dreiecksungleichung und die Eigenschaft

$$\|\lambda f\|_1 = \int |\lambda| \cdot |f| d\mu = |\lambda| \cdot \int |f| d\mu = |\lambda| \|f\|_1$$

Zuletzt zeigen wir, dass für $\|f\|_1 = 0$ fast überall $f = 0$ gilt. Das folgt aus dem Satz der monotonen Konvergenz, weil dann $(n|f|)_{n \in \mathbb{N}}$ fast überall konvergiert. $\mathbf{q.e.d.}$

Korollar 12.21. (Lebesgue's Satz der beschränkten Konvergenz) Eine Folge $(f_n)_{n \in \mathbb{N}}$ in $L^1(\mathbb{R}^d)$ konvergiere fast überall gegen f , und für ein $k \in L^1(\mathbb{R}^d)$ und alle $n \in \mathbb{N}$ gelte fast überall $|f_n| \leq k$. Dann ist $f \in L^1(\mathbb{R}^d)$ und $(\int f_n d\mu)_{n \in \mathbb{N}}$ konvergiert gegen $\int f d\mu$.

Beweis: Seien $(g_{nm})_{m \in \mathbb{N}}$ und $(h_{nm})_{m \in \mathbb{N}}$ definiert durch

$$g_{nm} = \min\{f_n, f_{n+1}, \dots, f_{n+m}\} \quad \text{und} \quad h_{nm} = \max\{f_n, f_{n+1}, \dots, f_{n+m}\}.$$

Für alle $n \in \mathbb{N}$ sind wegen den Eigenschaften der lebesgueintegrierbaren Funktionen $(g_{nm})_{m \in \mathbb{N}}$ monoton fallende Folgen in $L^1(\mathbb{R}^d)$ mit durch $\int k d\mu$ beschränkten Integralen, und $(h_{nm})_{m \in \mathbb{N}}$ monoton wachsende Folgen von Funktionen mit durch $\int k d\mu$ beschränkten Integralen. Wegen dem Satz der monotonen Konvergenz konvergieren diese Folgen fast überall gegen Funktionen $(g_n)_{n \in \mathbb{N}}$ und $(h_n)_{n \in \mathbb{N}}$ in $L^1(\mathbb{R}^d)$. Fast überall sind

$$g_n = \inf\{f_n, f_{n+1}, f_{n+2}, \dots\} \quad \text{und} \quad h_n = \sup\{f_n, f_{n+1}, f_{n+2}, \dots\}$$

monotone Folgen in $L^1(\mathbb{R}^d)$ mit beschränkten Integralen. Mit $(f_n)_{n \in \mathbb{N}}$ konvergieren $(g_n)_{n \in \mathbb{N}}$ und $(h_n)_{n \in \mathbb{N}}$ fast überall gegen f . Dann gilt auch

$$\int g_n d\mu \leq \int f_n d\mu \leq \int h_n d\mu \quad \text{und} \quad \int f d\mu \leq \lim_{n \rightarrow \infty} \int f_n d\mu \leq \int f d\mu. \quad \mathbf{q.e.d.}$$

Korollar 12.22. (Vollständigkeit von $L^1(\mathbb{R}^d)$, Satz von Riesz-Fischer)
 $L^1(\mathbb{R}^d)$ ist mit $\|\cdot\|_1$ ein Banachraum.

Beweis: Sei $(f_n)_{n \in \mathbb{N}}$ eine Cauchyfolge in $L^1(\mathbb{R}^d)$. Dann gibt es eine Teilfolge $(f_{n_m})_{m \in \mathbb{N}}$, so dass für alle $m \in \mathbb{N}$ gilt $\|f_{n_{m+1}} - f_{n_m}\|_1 \leq 2^{-m}$. Die Reihe $(\sum_{m=1}^n |f_{n_{m+1}} - f_{n_m}|)_{n \in \mathbb{N}}$ erfüllt dann die Voraussetzungen des Satzes der monotonen Konvergenz. Also konvergiert sie fast überall gegen eine Funktion $k \in L^1(\mathbb{R}^d)$. Dann konvergiert auch die Folge $(\sum_{m=1}^n f_{n_{m+1}} - f_{n_m})_{n \in \mathbb{N}} = (f_{n_{n+1}} - f_{n_1})_{n \in \mathbb{N}}$ fast überall und erfüllt mit $k \in L^1(\mathbb{R}^d)$ die Voraussetzungen von Lebesgue's Satz der beschränkten Konvergenz. Dann konvergiert auch die Teilfolge gegen einen Grenzwert $f \in L^1(\mathbb{R}^d)$. Weil $(f_n)_{n \in \mathbb{N}}$ eine Cauchyfolge ist, konvergiert $(\|f_n - f\|_1)_{n \in \mathbb{N}}$ gegen Null, und damit auch $(f_n)_{n \in \mathbb{N}}$ gegen f . $\mathbf{q.e.d.}$

Im Allgemeinen konvergieren in $L^1(\mathbb{R}^d)$ konvergente Folge $(f_n)_{n \in \mathbb{N}}$ nicht fast überall gegen den Grenzwert f . Eine Teilfolge konvergiert aber immer fast überall gegen f .

Beispiel 12.23. Wir betrachten für alle $n \in \mathbb{N}$ und alle $k \in \{1 - n^2, 2 - n^2, \dots, n^2\}$ die charakteristische Funktion $\chi_{n,k}$ der abgeschlossenen Intervalle $[\frac{k-1}{n}, \frac{k}{n}]$. Sei $m \mapsto (n, k)$ eine Abzählung solcher Paare (n, k) mit $n \in \mathbb{N}$ und $k \in \{1 - n^2, 2 - n^2, \dots, n^2\}$. Dann ist die entsprechende Folge $(\chi_m)_{m \in \mathbb{N}}$ in $L^1(\mathbb{R})$ eine Nullfolge, weil es für alle $n \in \mathbb{N}$ nur endlich viele $m \in \mathbb{N}$ gibt mit $\|\chi_m\|_1 \geq \frac{1}{n}$. Andererseits gibt es für alle $x \in \mathbb{R}$ unendlich viele $m \in \mathbb{N}$ mit $\chi_m(x) = 1$. Also konvergiert $(\chi_m(x))_{m \in \mathbb{N}}$ für kein $x \in \mathbb{R}$ gegen 0.

Satz 12.24. (Fatou's Lemma) Sei $(f_n)_{n \in \mathbb{N}}$ eine Folge in $L^1(\mathbb{R}^d)$ von fast überall nicht negativen Funktionen, die fast überall gegen f konvergieren. Wenn $(\int f_n d\mu)_{n \in \mathbb{N}}$ beschränkt ist, dann ist $f \in L^1(\mathbb{R}^d)$ und es gilt

$$\int f d\mu \leq \liminf_{n \rightarrow \infty} \left\{ \int f_n d\mu, \int f_{n+1} d\mu, \dots \right\}.$$

Beweis: Für alle $n \in \mathbb{N}$ erfüllt die Folge $(g_{nm})_{m \in \mathbb{N}}$ mit $g_{nm} = \min\{f_n, f_{n+1}, \dots, f_{n+m}\}$ wieder die Voraussetzungen des Satzes der monotonen Konvergenz und konvergiert gegen $g_n \in L^1(\mathbb{R}^d)$ mit $g_n = \inf\{f_n, f_{n+1}, \dots\}$. Die Folgen $(g_n)_{n \in \mathbb{N}}$ erfüllen wieder die Voraussetzungen des Satzes der monotonen Konvergenz und konvergieren fast überall gegen f . Also gilt auch $f \in L^1(\mathbb{R}^d)$ und fast überall $f_{n+m} \geq g_n$ für alle $m \in \mathbb{N}_0$. Es folgt $\int f d\mu = \lim_{n \rightarrow \infty} \int g_n d\mu$ und $\int f d\mu \leq \lim_{n \rightarrow \infty} \inf\{\int f_n d\mu, \int f_{n+1} d\mu, \dots\}$. $\mathbf{q.e.d.}$

12.5 Der Satz von Fubini

Für jeden Quader $Q = I_1 \times \dots \times I_d \times I_{d+1} \subset \mathbb{R}^d \times \mathbb{R}$ und jedes $x \in \mathbb{R}^d$ ist die Funktion $\mathbb{R} \rightarrow \mathbb{R}$, $y \mapsto \chi_Q(x, y)$ eine Treppenfunktion auf $\mathbb{R}$. Wenn wir diese Funktion integrieren erhalten wir eine Treppenfunktion auf dem $\mathbb{R}^d$:

$$\int \chi_Q(x, y) d\mu(y) = \begin{cases} \text{Länge von } I_{d+1} & \text{wenn } x \in I_1 \times \dots \times I_d, \\ 0 & \text{wenn } x \notin I_1 \times \dots \times I_d. \end{cases}$$

Also ist

$$\int \left(\int \chi_Q(x, y) d\mu(y) \right) d\mu(x) = \mu(Q).$$

Wegen der Linearität des Integrals definiert die Abbildung $\int d\mu(y)$ also eine lineare Abbildung von den Treppenfunktionen auf $\mathbb{R}^d \times \mathbb{R}$ in die Treppenfunktionen auf $\mathbb{R}^d$. Und für jede Treppenfunktion f auf $\mathbb{R}^d \times \mathbb{R}$ gilt

$$\int \left(\int f(x, y) d\mu(y) \right) d\mu(x) = \int f d\mu.$$

In diesem Abschnitt wollen wir zeigen, dass diese Abbildung eine Abbildung

$$\int d\mu(y) : L^1(\mathbb{R}^d \times \mathbb{R}) \rightarrow L^1(\mathbb{R}^d)$$

induziert, so dass für alle $f \in L^1(\mathbb{R}^d \times \mathbb{R})$ gilt

$$\int \left(\int f(x, y) d\mu(y) \right) d\mu(x) = \int f d\mu.$$

Wenn $f \geq g$ zwei Treppenfunktionen auf $\mathbb{R}^d \times \mathbb{R}$ sind, dann erfüllen für jedes $x \in \mathbb{R}^d$ die entsprechenden Treppenfunktionen $f_x : y \rightarrow f(x, y)$ bzw. $g_x : y \rightarrow g(x, y)$ auch $f_x \geq g_x$. Wegen Proposition 12.8 gilt für die Integrale auch

$$\int f(x, y) d\mu(y) \geq \int g(x, y) d\mu(y).$$

Also definiert $\int d\mu(y)$ eine lineare monotone Abbildung von den Treppenfunktionen auf $\mathbb{R}^d \times \mathbb{R}$ in die Treppenfunktionen auf $\mathbb{R}^d$. Damit diese Abbildungen eine Abbildung von $L^1(\mathbb{R}^d \times \mathbb{R})$ nach $L^1(\mathbb{R}^d)$ induziert, müssen zwei fast überall definierte Grenzwerte von monoton wachsenden Treppenfunktionen, die fast überall übereinstimmen, auch auf zwei fast überall definierte Grenzwerte von monoton wachsenden Treppenfunktionen abgebildet werden, die fast überall übereinstimmen.

Lemma 12.25. *Sei $S \subset \mathbb{R}^d \times \mathbb{R}$ eine Nullmenge. Dann ist fast überall in $x \in \mathbb{R}^d$, die Menge $S_x = \{y \in \mathbb{R} \mid (x, y) \in S\}$ eine Nullmenge in $\mathbb{R}$.*

Beweis: Wegen Satz 12.10 gibt es eine monoton wachsende Folge $(f_n)_{n \in \mathbb{N}}$ von Treppenfunktionen auf $\mathbb{R}^d \times \mathbb{R}$ mit beschränkten Integralen, die auf S divergiert. Dann sind auch die entsprechenden Integrale $(\int f_n d\mu(y))_{n \in \mathbb{N}}$ über $\mathbb{R}$ monoton wachsende Treppenfunktionen mit beschränkten Integralen auf $\mathbb{R}^d$. Wegen Satz 12.9 konvergieren sie fast überall auf $x \in \mathbb{R}^d$. Für alle $x \in \mathbb{R}^d$, für die $(\int f_n d\mu(y))_{n \in \mathbb{N}}$ konvergiert, sind die entsprechenden Einschränkungen von $(f_n)_{n \in \mathbb{N}}$ auf $\{x\} \times \mathbb{R}$ monoton wachsende Folgen von Treppenfunktionen auf $y \in \mathbb{R}$ mit beschränkten Integralen. Wegen Satz 12.9 sind also für alle $x \in \mathbb{R}^d$, so dass die Integrale über $y \in \mathbb{R}$ konvergieren, die Mengen der $y \in \mathbb{R}$, an denen $(f_n(x, y))_{n \in \mathbb{N}}$ nicht konvergiert, Nullmengen die S_x enthalten. **q.e.d.**

Proposition 12.26. *Die Integration über $y \in \mathbb{R}$ induziert eine lineare monotone Abbildung von $L^1(\mathbb{R}^d \times \mathbb{R})$ nach $L^1(\mathbb{R}^d)$, so dass für alle $f \in L^1(\mathbb{R}^d \times \mathbb{R})$ gilt*

$$\int \left(\int f(x, y) d\mu(y) \right) d\mu(x) = \int f d\mu.$$

Beweis: Die Integration über $y \in \mathbb{R}$ ist eine monotone lineare Abbildung von den Treppenfunktionen auf $(x, y) \in \mathbb{R}^d \times \mathbb{R}$ in die Treppenfunktionen auf $x \in \mathbb{R}^d$. Wegen Lemma 12.25 induziert sie eine Abbildung von den Äquivalenzklassen von den Grenzwerten von monoton wachsenden Folgen von Treppenfunktionen mit beschränkten Integralen auf $\mathbb{R}^d \times \mathbb{R}$ in die entsprechenden Äquivalenzklassen auf $\mathbb{R}^d$. Wegen der Definition der lebesgueintegrierbaren Funktionen induziert sie also auch eine Abbildung von $L^1(\mathbb{R}^d \times \mathbb{R})$ nach $L^1(\mathbb{R}^d)$. Folgende Gleichung gilt dann für alle Treppenfunktionen f auf $\mathbb{R}^d \times \mathbb{R}$ und damit auch für alle Grenzwerte $f \in L^1(\mathbb{R}^d \times \mathbb{R})$ solcher Treppenfunktionen:

$$\int \left(\int f(x, y) d\mu(y) \right) d\mu(x) = \int f d\mu. \quad \text{q.e.d.}$$

Die gleichen Argumente zeigen die analoge Aussage auch für die vertauschten Faktoren $\mathbb{R} \times \mathbb{R}^d$. Wenn wir die mehrfach anwenden erhalten wir also

Korollar 12.27. (Satz von Fubini) *Für alle Funktionen $f \in L^1(\mathbb{R}^d \times \mathbb{R}^{d'})$ gilt*

$$\int f d\mu = \int \left(\int f(x, y) d\mu(y) \right) d\mu(x) = \int \left(\int f(x, y) d\mu(x) \right) d\mu(y). \quad \text{q.e.d.}$$

Mit dem Satz von Fubini und dem Lebesguekriterium können wir jetzt auch Integrale auf dem $\mathbb{R}^d$ ausrechnen. Als erstes können wir für fast alle $(x_2, \dots, x_s) \in \mathbb{R}^d$ das Integral $\int_{-\infty}^{\infty} f(x_1, \dots, x_d) dx_1$ ausrechnen. Wenn f riemannintegrierbar ist können wir dabei die Methoden der eindimensionalen Integration, wie wir sie bei dem Riemannintegral kennen, benutzen. Wenn f außerhalb einer kompakten Menge verschwindet, benutzen wir das Riemannintegral und ansonsten das uneigentliche Riemannintegral. Danach integrieren wir genauso über $dx_2, \dots, dx_d$, und erhalten zuletzt

$$\int_{\mathbb{R}^d} f(x) d\mu = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, \dots, x_d) dx_1 \dots dx_d.$$

Wir können die Reihenfolge dieser eindimensionalen Integrale beliebig vertauschen.

12.6 Jacobis Transformation von Maßen

In diesem Abschnitt untersuchen wir, wie sich die Integration unter Koordinatentransformationen verhält und verallgemeinern die Substitutionsregel auf $\mathbb{R}^d$. Dafür benutzen wir $\|\cdot\|_\infty$ aus Definition 9.9 deren Bälle Quader mit gleichlangen Kanten sind.

Definition 12.28. Eine Teilmenge $A \subset \mathbb{R}^d$ heißt messbar, wenn für alle $f \in L^1(\mathbb{R}^d)$ das Produkt $f \cdot \chi_A$ mit der charakteristischen Funktion von A in $L^1(\mathbb{R}^d)$ liegt.

Satz 12.29. (i) Das Komplement einer messbaren Menge ist messbar.

(ii) Die Schnittmenge von abzählbar vielen messbaren Mengen ist messbar.

(iii) Jede offene Menge ist messbar.

Beweis: (i) Aus $f, f\chi_A \in L^1(\mathbb{R}^d)$ folgt $f\chi_{\mathbb{R}^d \setminus A} = f(1 - \chi_A) = f - f\chi_A \in L^1(\mathbb{R}^d)$.

(ii) Seien $(A_n)_{n \in \mathbb{N}}$ messbar und $A = \bigcap_n A_n$. Mit $f \in L^1(\mathbb{R}^d)$ liegt auch

$$f_n = f\chi_{A_1 \cap \dots \cap A_n} = \min\{f^+ \chi_{A_1}, \dots, f^+ \chi_{A_n}\} - \min\{f^- \chi_{A_1}, \dots, f^- \chi_{A_n}\} \in L^1(\mathbb{R}^d),$$

wobei $f^\pm = \frac{|f| \pm f}{2}$ ist. Die Folge $(f_n)_{n \in \mathbb{N}}$ konvergiert fast überall gegen $f\chi_A$ und erfüllt $|f_n| \leq |f|$. Aus Lebesgues Satz der beschränkten Konvergenz folgt $f\chi_A \in L^1(\mathbb{R}^d)$.

(iii) Jeder Quader ist messbar, weil die Multiplikation einer Treppenfunktion mit der charakteristischen Funktion eines Quaders eine Treppenfunktion ist. Für eine offene Menge $U \neq \mathbb{R}^d$ und $x \in U$ sei $r > 0$ mit $B(x, 3r) \subset U$ und y ein Element in $B(x, r) \cap \mathbb{Q}^d$. Dann ist $B(y, 2r) \subset B(x, 3r) \subset U$. Also ist $x \in B(y, \sup\{r > 0 \mid B(y, 2r) \subset U\}) \subset U$, und U folgende abzählbare Vereinigung von offenen bzw. kompakten $\|\cdot\|_\infty$ -Bällen

$$U = \bigcup_{y \in U \cap \mathbb{Q}^d} B(y, \sup\{r > 0 \mid B(y, 2r) \subset U\}) = \bigcup_{y \in U \cap \mathbb{Q}^d} \overline{B(y, \sup\{r > 0 \mid B(y, 2r) \subset U\})}.$$

Dann folgt (iii) aus (i) und (ii) und den de Morganschen Regeln.

q.e.d.

Definition 12.30. Für messbare Teilmengen A von $\mathbb{R}^d$ sei $L^1(A) \subset L^1(\mathbb{R}^d)$ der Teilraum aller lebesgueintegrierbaren Funktionen auf $\mathbb{R}^d$, die außerhalb von A verschwinden.

Der Beweis des Satzes von Riesz-Fischer gibt für jede Cauchyfolge in $L^1(\mathbb{R}^d)$ eine fast überall konvergente Teilfolge an. Also ist $L^1(A)$ abgeschlossen und ein Banachraum.

Lemma 12.31. (i) Die Multiplikation mit einer beschränkten stetigen Funktion f ist eine lineare Abbildung in $\mathcal{L}(L^1(\mathbb{R}^d))$, deren Norm beschränkt ist durch $\|f\|_\infty$.

(ii) Sei $\Phi : U \rightarrow O$ eine bijektive lipschitzstetige Abbildung zwischen den offenen Mengen $U, O \subset \mathbb{R}^d$ mit Lipschitzkonstante L . Dann ist $f \mapsto f \circ \Phi^{-1}$ eine lineare stetige Abbildung $L^1(U) \rightarrow L^1(O)$, deren Norm beschränkt ist durch L^d .

(iii) Seien $U, O \subset \mathbb{R}^d$ offenen Mengen und $\Phi : U \rightarrow O$ eine surjektive Abbildung mit $\|\Phi(x) - \Phi(y) - (x - y)\|_\infty \leq \epsilon \|x - y\|_\infty$ für $x, y \in U$ und ein $\epsilon \in (0, 1)$. Dann gilt

$$\left| \int_O f \circ \Phi^{-1} d\mu - \int_U f d\mu \right| \leq (2^d - 1)\epsilon \|f\|_1 \quad \text{für alle } f \in L^1(U).$$

Beweis: (i) Für eine beschränkte stetige Funktion f und eine Treppenfunktion g gilt $|fg| \leq \|f\|_\infty |g|$ und wegen Satz 12.15 $gf \in L^1(\mathbb{R}^d)$ und wegen Satz 12.14 auch $\|fg\|_1 \leq \|f\|_\infty \|g\|_1$. Für eine monoton wachsende Folge $(g_n)_{n \in \mathbb{N}}$ von Treppenfunktionen mit beschränkten Integralen gilt entweder $g_n - g_m \geq 0$ oder $g_n - g_m \leq 0$, also $\|g_n - g_m\|_1 \leq |\int g_n d\mu - \int g_m d\mu|$. Dann ist $(g_n)_{n \in \mathbb{N}}$ mit $(\int g_n d\mu)_{n \in \mathbb{N}}$ eine Cauchyfolge und konvergiert bezüglich $\|\cdot\|_1$. Also liegen die Treppenfunktionen dicht in $L^1(\mathbb{R}^d)$, woraus (i) folgt.

(ii) Wir zeigen zuerst, dass die Linearkombinationen von χ_B mit kompakten und bis auf Nullmengen paarweise disjunkten $\|\cdot\|_\infty$ -Bällen $B \subset U$ dicht in $L^1(U)$ liegen. Das folgt wenn $\chi_{Q \cap U}$ für jeden endlichen offenen Quader Q im Abschluss solcher Linearkombinationen liegt, weil die Treppenfunktionen dicht in $L^1(\mathbb{R}^d)$ liegen. Im Beweis von Satz 12.29 (iii) wurde $Q \cap U$ in eine abzählbare Vereinigung von kompakten $\|\cdot\|_\infty$ -Bällen $(B_n)_{n \in \mathbb{N}}$ in U zerlegt. Neben den Koordinaten ihrer Zentren können wir auch ihre Radien in $\mathbb{Q}$ wählen, indem wir sie um einen Faktor in $[1, \frac{3}{2})$ vergrößern. Weil solche $\|\cdot\|_\infty$ -Bälle Quader mit rationalen Koordinaten sind, können wir induktiv $B_{n+1} \setminus (B_1 \cup \dots \cup B_n)$ jeweils in solche bis auf Nullmengen paarweise disjunkte $\|\cdot\|_\infty$ -Bälle zerlegen. Dann konvergiert $(\sum \chi_{B_n})_{n \in \mathbb{N}}$ gegen $\chi_{Q \cap U}$ mit bis auf Nullmengen paarweise disjunkten B_n .

Für kompakte $\|\cdot\|_\infty$ -Bälle $B \subset U$ ist $\Phi[B]$ kompakt und messbar, also liegt $\chi_B \circ \Phi^{-1} = \chi_{\Phi[B]}$ in $L^1(O)$. Wegen der Lipschitzstetigkeit von Φ ist $\Phi[B]$ in einem $\|\cdot\|_\infty$ -Ball enthalten, dessen Radius L mal so groß ist wie der von B . Also gilt zunächst für $f = \chi_B$

$$\int |f \circ \Phi^{-1}| d\mu \leq L^d \int |f| d\mu, \quad \text{und} \quad \|f \circ \Phi^{-1}\|_{L^1(O)} \leq L^d \|f\|_{L^1(U)}.$$

Dann gilt das auch für Linearkombinationen von χ_B mit bis auf Nullmengen paarweise disjunkten kompakten $\|\cdot\|_\infty$ -Bällen in U , und deren Grenzwerte, also für alle $f \in L^1(U)$.

(iii) Für $x, y \in U$ folgt aus $\|\Phi(x) - \Phi(y) - (x - y)\|_\infty \leq \epsilon \|x - y\|_\infty$
 $\|y - x\|_\infty = \|\Phi(x) - \Phi(y) - (x - y) + \Phi(y) - \Phi(x)\|_\infty \leq \epsilon \|x - y\|_\infty + \|\Phi(y) - \Phi(x)\|_\infty,$
 $(1 - \epsilon)\|x - y\|_\infty \leq \|\Phi(x) - \Phi(y)\|_\infty \leq (1 + \epsilon)\|x - y\|_\infty.$

Also ist Φ injektiv und damit bijektiv, und Φ^{-1} lipschitzstetig mit Lipschitzkonstante $(1 - \epsilon)^{-1}$. Indem wir $f \in L^1(U)$ in $f = \frac{|f|+f}{2} - \frac{|f|-f}{2}$ zerlegen, genügt es die Ungleichung für $f \geq 0$ zu zeigen. Die Anwendung von (ii) auf Φ und Φ^{-1} ergibt für solche f

$$(1 - \epsilon)^d \int_U f d\mu = (1 - \epsilon)^d \int_U f \circ \Phi^{-1} \circ \Phi d\mu \leq \int_O f \circ \Phi^{-1} d\mu \leq (1 + \epsilon)^d \int_U f d\mu.$$

Aus $1 - (1 - \epsilon)^d \leq (1 + \epsilon)^d - 1 \iff 2 \leq (1 + \epsilon)^d + (1 - \epsilon)^d = \sum_{0 \leq 2l \leq d} \binom{d}{2l} 2\epsilon^{2l}$ folgt

$$\left| \int_O f \circ \Phi^{-1} d\mu - \int_U f d\mu \right| \leq ((1 + \epsilon)^d - 1) \|f\|_1 \leq \sum_{l=1}^d \binom{d}{l} \epsilon \|f\|_1 = (2^d - 1)\epsilon \|f\|_1. \text{ q.e.d.}$$

Satz 12.32. Sei $A \in \mathcal{L}(\mathbb{R}^d)$ eine invertierbare lineare Abbildung. Dann ist

$$\pi(A) : L^1(\mathbb{R}^d) \rightarrow L^1(\mathbb{R}^d), \quad f \mapsto \pi(A)f \quad \text{mit } (\pi(A)f)(x) = f(A^{-1}x)$$

eine stetige lineare Abbildung, deren Norm beschränkt ist durch $\|A\|^d$. Es gilt

$$\int \pi(A)f d\mu = |\det A| \int f d\mu \quad \text{für alle } f \in L^1(\mathbb{R}^d).$$

Beweis: Aus Lemma 12.31 (ii) folgt $\pi(A) \in \mathcal{L}(L^1(\mathbb{R}^d))$ und $\|\pi(A)\| \leq \|A\|^d$. Wie im Beweis von Lemma 12.31 (ii) genügt es die letzte Gleichung für $f = \chi_Q$ mit kompakten Quadern Q zu zeigen. Weil $A^{-1}x$ genau dann in Q liegt, wenn x in $A[Q]$ liegt, ist $\chi_Q \circ A^{-1} = \chi_{A[Q]}$. Wir betrachten zuerst solche A , die kompakte Quader auf kompakte Quader abbilden. Das sind einerseits Diagonalmatrizen und andererseits die Permutationen der Koordinaten. In beiden Fällen gilt die Formel offenbar für alle χ_Q mit kompakten Quadern Q und deshalb auch für alle $f \in L^1(\mathbb{R}^d)$. Wenn die elementare Zeilenumformung des Gaußschen Algorithmus, die für $i \neq j$ zu der i -ten Zeile das λ -fache der j -ten Zeile hinzuaddiert, der Linksmultiplikation mit A entspricht, dann berechnen wir $\int \chi_{A[Q]} d\mu$ indem wir mit dem Satz von Fubini zuerst über x_i integrieren. Für festes $x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_d$ ist $\{x_i \in \mathbb{R} \mid (x_1, \dots, x_d) \in A[Q]\}$ das um λx_j verschobene Intervall $\{x_i \in \mathbb{R} \mid (x_1, \dots, x_d) \in Q\}$, also ein Intervall gleicher Länge. Für solche A ist $\det A = 1$, und die Formel gilt also wieder zunächst für alle χ_Q und dann für alle $f \in L^1(\mathbb{R}^d)$. Für zwei invertierbare $A, B \in \mathcal{L}(\mathbb{R}^d)$ ist $\pi(A \cdot B) = \pi(A) \cdot \pi(B)$. Wegen $|\det(A \cdot B)| = |\det A| \cdot |\det B|$ gilt die Formel also auch für die Produkte solcher Matrizen. Mit dem Gaußschen Algorithmus können wir jede invertierbare Matrix A durch elementare Zeilenumformungen in eine diagonale Matrix umwandeln. Die inversen Zeilenumformungen verwandeln die Diagonalmatrix in A . Also ist A ein Produkt von Matrizen, für die die Formel gilt. Damit gilt die Formel auch für A . **q.e.d.**

Satz 12.33. (Jacobis Transformationsformel) Sei $\Phi : U \rightarrow O$ stetig differenzierbare und bijektiv mit stetig differenzierbarer Umkehrabbildung zwischen den offenen Teilmenge U und O von $\mathbb{R}^d$. Dann ist die Abbildung $f \mapsto |\det \Phi'| f \circ \Phi$ eine bijektive Isometrie von $L^1(O)$ nach $L^1(U)$, d.h. für $f \in L^1(O)$ gilt $|\det \Phi'| f \circ \Phi \in L^1(U)$ mit

$$\int_U |\det \Phi'| f \circ \Phi d\mu = \int_O f d\mu.$$

Beweis: Im Beweis von Lemma 12.31 (ii) wurde gezeigt, dass die Linearkombinationen von χ_Q mit kompakten Quadern $Q \subset O$ dicht in $L^1(O)$ liegen. Also genügt es

$$\int |\det \Phi'| \chi_Q \circ \Phi d\mu = \int \chi_Q d\mu$$

für kompakte Quader $Q \subset O$ zu zeigen. Wegen Korollar 9.35, und weil $(\Phi^{-1})'$ stetig ist, ist $\|(\Phi^{-1})'\|$ auf Q durch ein positives $L > 0$ beschränkt, und L wegen dem Schranken-satz Korollar 10.6 auf der konvexen Menge Q eine Lipschitzkonstante von Φ^{-1} . Weil

Φ' wegen Satz 9.39 auf der kompakten Menge $\Phi^{-1}[Q]$ gleichmäßig stetig ist, gibt es für jedes $0 < \epsilon < 1$ eine Zerlegung $Q = Q_1 \cup \dots \cup Q_N$ in paarweise disjunkte Quader, auf denen jeweils $\|A_n - \Phi'(\Phi^{-1}(x))\| \leq \frac{\epsilon}{L}$ oder auch $\|A_n \circ (\Phi^{-1})'(x) - \mathbb{1}\| \leq \frac{\epsilon}{L} \|(\Phi^{-1})'(x)\| \leq \epsilon$ gilt. Hierbei ist $A_n = \Phi'(\Phi^{-1}(x_n))$ am Zentrum x_n von Q_n . Aus dem Schrankensatz folgt $\|(A_n \circ \Phi^{-1})(x) - (A_n \circ \Phi^{-1})(y) - (x - y)\|_\infty \leq \epsilon \|x - y\|_\infty$, und mit Lemma 12.31 (iii)

$$\left| \int \chi_{Q_n} \circ \Phi \circ A_n^{-1} d\mu - \int \chi_{Q_n} d\mu \right| \leq (2^d - 1) \epsilon \|\chi_{Q_n}\|_1 \quad \text{für } n = 1, \dots, N.$$

Auf $x \in Q_n$ gilt $\|\mathbb{1} - \Phi'(\Phi^{-1}(x)) \circ A_n^{-1}\| \leq \|A_n - \Phi'(\Phi^{-1}(x))\| \cdot \|A_n^{-1}\| \leq \frac{\epsilon}{L} L = \epsilon$. Wir wenden Lemma 12.31 (iii) auf $\Phi'(\Phi^{-1}(x)) \circ A_n^{-1} \in \mathcal{L}(\mathbb{R}^d)$ an und erhalten mit Satz 12.32

$$\begin{aligned} & \left| \int |\det \Phi'| \chi_{Q_n} \circ \Phi d\mu - \int \chi_{Q_n} \circ \Phi \circ A_n^{-1} d\mu \right| \leq \int \left| |\det \Phi'| - |\det A_n| \right| \chi_{Q_n} \circ \Phi d\mu \\ &= \int \left| |\det \Phi' A_n^{-1}| - 1 \right| \cdot |\det A_n| \chi_{Q_n} \circ \Phi d\mu \leq (2^d - 1) \epsilon \int |\det A_n| \chi_{Q_n} \circ \Phi d\mu \leq (4^d - 2^d) \epsilon \|\chi_{Q_n}\|_1. \end{aligned}$$

Für $\sum_{n=1}^N$ folgt $\left| \int |\det \Phi'| \chi_Q \circ \Phi d\mu - \int \chi_Q d\mu \right| \leq 4^d \epsilon \|\chi_Q\|_1 \rightarrow 0$ für $\epsilon \rightarrow 0$. **q.e.d.**

12.7 Integration über Untermannigfaltigkeiten

In diesem Abschnitt definieren wir die Integration von reellen Funktionen auf stetig differenzierbare Untermannigfaltigkeiten $A \subset \mathbb{R}^d$. Zunächst charakterisieren wir Untermannigfaltigkeiten durch die Existenz von geeigneten Parametrisierungen.

Lemma 12.34. *Eine Teilmenge $A \subset \mathbb{R}^d$ ist genau dann eine k -dimensionale Untermannigfaltigkeit, wenn A überdeckt wird durch die Bilder $\Phi[U]$ von Homöomorphismen $\Phi : U \rightarrow \Phi[U]$ von offenen Teilmengen $U \subset \mathbb{R}^k$ auf offene Teilmengen von A , die als Abbildung nach $\mathbb{R}^d$ stetig differenzierbar sind und auf U eine injektive Ableitung haben.*

Insbesondere ist eine Niveaumenge $A \subset \mathbb{R}^d$ in allen $x \in A$ eine Untermannigfaltigkeit ist, in denen die Niveaumenge gemäß Definition 11.17 stetig differenzierbar ist.

Beweis: Wenn A eine Untermannigfaltigkeit von $\mathbb{R}^d$ ist, dann liegt gemäß Definition 11.12 jedes $x \in A$ im Definitionsbereich einer stetig differenzierbaren Abbildung $f : O \rightarrow \mathbb{R}^l$ mit surjektivem $f'(x)$. Nach einer geeigneten Permutation der Koordinaten von $\mathbb{R}^d$ bilden die hinteren l partiellen Ableitungen von f eine invertierbare $l \times l$ -Matrix, so dass f auf $O \subset \mathbb{R}^{d-l} \times \mathbb{R}^l$ die Voraussetzungen des Satzes der impliziten Funktion erfüllt. Deshalb gibt es offene Teilmengen $U \subset \mathbb{R}^{d-l}$ und $V \subset \mathbb{R}^l$ und eine stetig differenzierbare Abbildung $g : U \times V \rightarrow \mathbb{R}^l$, so dass $f(z, g(z, y)) = y$ für alle $(z, y) \in U \times V$ gilt, und $\{z, g(z, f(x)) \mid z \in U\}$ die Schnittmenge von A mit einer

offenen Umgebung von x im $\mathbb{R}^d$ ist. Weil die Ableitung von $\Phi(z) = (z, g(z, f(x)))$ auf ganz U injektiv ist, erfüllt A die Bedingung im Lemma mit $k = d - l$.

Wenn umgekehrt $A \subset \mathbb{R}^d$ die Bedingung im Lemma erfüllt und $x \in A$ im Bild einer entsprechenden Abbildung $\Phi : U \rightarrow \Phi[U]$ liegt, dann bilden nach einer geeigneten Permutation der Koordinaten von $\mathbb{R}^d$ die partiellen Ableitungen der ersten k Komponenten von Φ bei $\Phi^{-1}(x)$ eine invertierbare $k \times k$ -Matrix. Mit der Projektion auf die erste Komponente $P_1 : \mathbb{R}^k \times \mathbb{R}^{d-k} \rightarrow \mathbb{R}^k$ erfüllt $P_1 \circ \Phi$ die Voraussetzungen des Satzes der inversen Funktion. Dann ist die Einschränkung $P_1 \circ \Phi|_V$ auf eine Umgebung V von $\Phi^{-1}(x)$ eine bijektive Abbildung auf eine offene Teilmenge $W \subset \mathbb{R}^k$ mit stetig differenzierbare Umkehrfunktion $(P_1 \circ \Phi|_V)^{-1}$. Weil W in $\mathbb{R}^k$ und $\Phi[V]$ in A offen sind, gibt es eine in $\mathbb{R}^d$ offene Vereinigung O von offenen Bällen in $W \times \mathbb{R}^{d-k}$, so dass $A \cap O = \Phi[V]$ die Nullstellenmenge von $f : O \rightarrow \mathbb{R}^{d-k}, (z, y) \mapsto y - P_2 \circ \Phi \circ (P_1 \circ \Phi|_V)^{-1}(z)$ mit der Projektion $P_2 : \mathbb{R}^k \times \mathbb{R}^{d-k} \rightarrow \mathbb{R}^{d-k}$ auf die zweite Komponente ist. Dann ist $\frac{\partial f}{\partial y} = \mathbb{1}_{\mathbb{R}^{d-k}}$, und A erfüllt bei allen $x \in A$ die Bedingung in der Definition 11.12. **q.e.d.**

Für zwei solche Parametrisierungen $\Phi : U \rightarrow \Phi[U]$ und $\tilde{\Phi} : \tilde{U} \rightarrow \tilde{\Phi}[\tilde{U}]$, deren Bilder beide ein $x \in A$ enthalten, zeigt der Beweis, dass die Verkettungen $P \circ \Phi|_V$ und $\tilde{P} \circ \tilde{\Phi}|_{\tilde{V}}$ mit geeigneten Projektionen stetig differenzierbare Umkehrabbildungen haben. Dann sind $\Phi^{-1} \circ \tilde{\Phi} = (P \circ \Phi|_V)^{-1} P \circ \tilde{\Phi}$ und $\tilde{\Phi}^{-1} \circ \Phi = (\tilde{P} \circ \tilde{\Phi}|_{\tilde{V}})^{-1} \tilde{P} \circ \Phi$ stetig differenzierbar und Umkehrabbildungen voneinander. Also haben in einem $x \in A$ alle Parametrisierungen die gleiche Dimension k , und diese Dimension hängt nur von der Menge A ab.

Mit diesen Parametrisierungen Φ können wir Funktionen $f : A \rightarrow \mathbb{R}$ integrieren, wenn f außerhalb einer der Mengen $\Phi[U]$ verschwindet. Mithilfe einer sogenannten Zerlegung der Eins zerlegen wir jedes $f : A \rightarrow \mathbb{R}$ in eine Summe solcher Funktionen.

Definition 12.35. (Zerlegung der Eins) Eine glatte Zerlegung der Eins bezüglich einer offenen Überdeckung einer offenen Teilmenge $\Omega \subset \mathbb{R}^d$, ist eine abzählbare Familie $(h_n)_{n \in \mathbb{N}}$ von glatten Funktionen $h_n : \Omega \rightarrow [0, 1]$, so dass

- (i) für jedes $x \in \Omega$ auf einer Umgebung von x nur endlich viele h_n ungleich Null sind.
- (ii) Für alle $x \in \Omega$ gilt $\sum_{n=1}^{\infty} h_n(x) = 1$.
- (iii) Jedes h_n außerhalb einer kompakten Teilmenge eines Elementes der Überdeckung verschwindet.

Satz 12.36. (Existenz der Zerlegung der Eins) Jede offene Überdeckung $\mathcal{U}$ einer offenen Teilmenge $\Omega \subset \mathbb{R}^d$ besitzt eine glatte Zerlegung der Eins.

Beweis: Für $0 < a < b < \infty$ sei $h_{a,b} \in C^\infty(\mathbb{R})$ definiert durch

$$h_{a,b} : \mathbb{R} \rightarrow [0, 1], \quad x \mapsto h_{a,b}(x) = \begin{cases} 1 & \text{für } 0 \leq |x| \leq a \\ \exp\left(\frac{1}{|x|^2 - b^2} \exp\left(\frac{-1}{|x|^2 - a^2}\right)\right) & \text{für } a < |x| < b \\ 0 & \text{für } b \leq |x| \end{cases}$$

Der Beweis von Satz 12.29 (iii) stellt jede offene Menge $\Omega \subset \mathbb{R}^d$ als abzählbare Vereinigung von offenen $\|\cdot\|_\infty$ -Bällen dar, deren Abschlüsse kompakte Teilmengen von Ω sind. Der Abschluss der Vereinigung von endlich vielen dieser Bälle ist kompakt in Ω , und wird jeweils von endlich vielen solcher Bälle überdeckt. Sei also O_n induktiv für alle $n \in \mathbb{N}$ die offene Vereinigung endlich vieler solcher Bälle, so dass O_{n+1} jeweils die kompakte Menge $\bar{O}_n$ enthält. Dabei sei O_1 einer der Bälle und zusätzlich sei $O_0 = O_{-1} = \emptyset$. Für jedes $m \in \mathbb{N}$ und jedes $x \in \bar{O}_m \setminus O_{m-1}$ ist $B(x, 3r)$ (bezüglich $\|\cdot\|_2$) für ein $r > 0$ in $O_{m+1} \setminus \bar{O}_{m-2}$ und in einer der offenen Mengen von $\mathcal{U}$ enthalten. Also wird die kompakte Menge $\bar{O}_m \setminus O_{m-1}$ von endlich vielen solchen Bällen $B(x, r)$ überdeckt. Induktiv erhalten wir eine Folge von offenen Bällen $(B(x_n, r_n))_{n \in \mathbb{N}}$, die Ω überdecken, so dass jeder $B(x_n, 3r_n)$ jeweils in einer der offenen Mengen von $\mathcal{U}$ enthalten, und jeweils nur zu endlich vielen anderen der Bälle $(B(x_n, 3r_n))_{n \in \mathbb{N}}$ nicht schnittfremd ist. Dann ist

$$\begin{aligned} h_n : \Omega &\rightarrow [0, 1], \quad x \mapsto h_n(x) = h_{r_n, 2r_n}(\|x - x_n\|) \prod_{m=1}^{n-1} (1 - h_{r_m, 2r_m}(\|x - x_m\|)) \\ &= \prod_{m=1}^{n-1} (1 - h_{r_m, 2r_m}(\|x - x_m\|)) - \prod_{m=1}^n (1 - h_{r_m, 2r_m}(\|x - x_m\|)). \end{aligned}$$

eine glatte Zerlegung der Eins bezüglich der Überdeckung $\mathcal{U}$. Denn induktiv gilt

$$h_1(x) + \dots + h_n(x) + (1 - h_{r_1, 2r_1}(\|x - x_1\|)) \cdots (1 - h_{r_n, 2r_n}(\|x - x_n\|)) = 1 \quad \text{für jedes } n \in \mathbb{N},$$

und für alle $n \in \mathbb{N}$ verschwindet $(1 - h_{r_n, 2r_n}(\|x - x_n\|))$ auf $B(x_n, r_n)$. **q.e.d.**

Definition 12.37. Sei $A \subset \mathbb{R}^d$ eine k -dimensionale Untermannigfaltigkeit von $\mathbb{R}^d$ und $f \in C(A, \mathbb{R})$ verschwinde außerhalb einer kompakten Teilmenge $K \subset A$. Wir überdecken K durch endlich viele offene Mengen $O \subset \mathbb{R}^d$ mit $A \cap O = \Phi[U]$ für ein Φ wie im Lemma 12.34 und wählen eine entsprechende Zerlegung der Eins $(h_l)_{l \in \mathbb{N}}$. Dann sei

$$\int_A f d\sigma = \sum_{l \in \mathbb{N}} \int_U (h_l f) \circ \Phi \sqrt{\det((\Phi')^T \Phi')} d\mu.$$

Zur Motivation bemerken wir, dass das Volumen des k -dimensionalen Parallelotops, das von den Spaltenvektoren einer $n \times k$ Matrix A aufgespannt wird, gleich $\sqrt{\det(A^T A)}$ ist. Dabei ist $A^T A$ die Matrix aller Skalarprodukte zwischen diesen Vektoren.

Lemma 12.38. Das Integral $\int_A f d\sigma$ hängt weder von der Wahl der Parametrisierung noch von der Wahl der Zerlegung der Eins ab.

Beweis: Auf K ist wegen der Bedingung (i) der Zerlegung der Eins die Summe in der Definition von $\int_A f d\sigma$ endlich. Für zwei Überdeckungen von K durch Mengen $\Phi[U]$ und $\Psi[V]$ wie im Lemma 12.34 mit entsprechenden Zerlegungen der Eins, bilden die

Schnittmengen von zwei Mengen (aus beiden Überdeckungen jeweils eine) auch eine Überdeckung durch solche Mengen und die Produkte von zwei Elementen (aus beiden Zerlegungen der Eins jeweils eine) eine entsprechende Zerlegung der Eins. Wegen der Bedingung (ii) an die Zerlegungen der Eins und der Linearität des Integrals genügt es dann den Fall zu betrachten, dass K in den Bildern $\Phi[U]$ und $\Psi[V]$ von zwei stetig differenzierbaren Homöomorphismen wie im Lemma 12.34 enthalten ist. Dann sind die Einschränkungen von Φ auf $\Phi^{-1}[\Phi[U] \cap \Psi[V]]$ und von Ψ auf $\Psi^{-1}[\Phi[U] \cap \Psi[V]]$ beides Homöomorphismen auf die offene Teilmenge $\Phi[U] \cap \Psi[V]$ von A . Die Verkettung des zweiten mit dem inversen des ersten ergibt einen Homöomorphismus $\Upsilon : \Psi^{-1}[\Phi[U] \cap \Psi[V]] \rightarrow \Phi^{-1}[\Phi[U] \cap \Psi[V]]$ so dass $\Psi(x) = \Phi(\Upsilon(x))$ für alle $x \in \Psi^{-1}[\Phi[U] \cap \Psi[V]]$ gilt. Aus der Bemerkung direkt im Anschluss an den Beweis von Lemma 12.34 folgt, dass Υ und Υ^{-1} stetig differenzierbar sind. Es folgt

$$\begin{aligned} \int_{\Psi^{-1}[\Phi[U] \cap \Psi[V]]} f \circ \Psi \sqrt{\det((\Psi')^T \Psi')} d\sigma &= \int_{\Psi^{-1}[\Phi[U] \cap \Psi[V]]} f \circ \Phi \circ \Upsilon \sqrt{\det(((\Phi \circ \Upsilon)')^T (\Phi \circ \Upsilon)')} d\sigma = \\ &= \int_{\Psi^{-1}[\Phi[U] \cap \Psi[V]]} \left(f \circ \Phi \sqrt{\det((\Phi')^T \Phi')} \right) \circ \Upsilon |\det \Upsilon'| d\sigma = \int_{\Phi^{-1}[\Phi[U] \cap \Psi[V]]} f \circ \Phi \sqrt{\det((\Phi')^T \Phi')} d\sigma. \end{aligned}$$

Im letzten Schritt haben wir Jacobis Transformationsformel benutzt.

q.e.d.

12.8 Der Gaußsche Satz

Im Gaußschen Satz werden beschränkte offene Teilmengen $\Omega \subset \mathbb{R}^d$ betrachtet, deren Rand $\partial\Omega = \bar{\Omega} \setminus \Omega$ eine $d - 1$ -dimensionale Untermannigfaltigkeit ist. Wegen dem Satz der impliziten Funktion ist die Menge $\partial\Omega$ dann nach einer geeigneten Permutation der Koordinaten lokal bei jedem $y \in \partial\Omega$ der Graph einer stetig differenzierbaren Funktion $g : U \rightarrow (a, b)$ auf einer offenen Umgebung $U \subset \mathbb{R}^{d-1}$ von $(y_1, \dots, y_{d-1})$. Dann ist Ω lokal von der Gestalt $\{(x, z) \in U \times (a, b) \mid z > g(x)\}$ bzw. $\{(x, z) \in U \times (a, b) \mid z < g(x)\}$. Im Gaußschen Satz wird die äußere Normale N auf $\partial\Omega$ benutzt. Das ist der Vektor, der auf der Tangentialebene des Randes senkrecht steht, der die Länge Eins hat und aus Ω heraus zeigt. In den beiden Fällen ist der Rand die Nullstellenmenge von

$$U \times (a, b) \rightarrow \mathbb{R}, \quad (x, z) \mapsto z - g(x).$$

Dann ist die Tangentialebene im Punkt $(x, g(x))$ der Kern der Ableitung dieser Abbildung an dieser Stelle. Also steht der Gradient dieser Abbildung senkrecht auf der Tangentialebene. Ob der Gradient aus der Menge herauszeigt oder in die Menge hineinzeigt erkennt man an der letzten Komponente. Also ist die äußere Normale bei den Randpunkten $(x, g(x))$ der obigen Mengen Ω gegeben durch

$$N(x, g(x)) = \pm \frac{(\nabla^T g(x), -1)^T}{\sqrt{1 + (\nabla g(x))^2}} \quad \text{für} \quad \Omega = \{(x, z) \in U \times (a, b) \mid z \gtrless g(x)\}.$$

Andererseits gilt für die entsprechende Abbildung $\Phi(x) = (x, g(x))$

$$\det((\Phi'(x))^T \Phi'(x)) = \det \left(\begin{pmatrix} \mathbb{1} & \nabla g(x) \end{pmatrix} \begin{pmatrix} \mathbb{1} \\ \nabla^T g(x) \end{pmatrix} \right) = 1 + (\nabla g(x))^2.$$

Die lineare Abbildung dieser Matrix wirkt auf dem orthogonalen Komplement von $\nabla g(x)$ nämlich wie die Identität und multipliziert $\nabla g(x)$ mit $1 + (\nabla g(x))^2$.

Beispiel 12.39. (i): Sei $f : \mathbb{R}^d \rightarrow \mathbb{R}$ stetig differenzierbar, so dass der Gradient ∇f keine gemeinsamen Nullstellen mit f hat. Dann hat $\Omega = \{x \in \mathbb{R}^d \mid f(x) < 0\}$ einen stetig differenzierbaren Rand, und Der Rand $\partial\Omega$ ist die Hyperfläche, auf der f verschwindet. Auf dem Rand $\partial\Omega$ ist ∇f ein Vektor, der orthogonal auf dem Tangentialraum an den Rand steht, und nach außen zeigt. Deshalb ist $\frac{\nabla f}{\|\nabla f\|}$ die äußere Normale.
(ii): Für $y \in \mathbb{R}^d$ und $r > 0$ ist der Ball gleich $B(y, r) = \{x \in \mathbb{R}^d \mid (x - y)^2 - r^2 < 0\}$. Die Funktion $f(x) = (x - y)^2 - r^2$ ist offenbar unendlich oft differenzierbar, und der Gradient $\nabla f(x) = 2(x - y)$ hat nur eine Nullstelle bei $x = y$. Also hat der Ball $B(y, r)$ einen unendlich oft differenzierbaren Rand mit der äußeren Normalen $N(x) = \frac{x - y}{\|x - y\|}$.

Satz 12.40. (Gaußscher Satz oder Divergenzsatz) Sei $\Omega \subset \mathbb{R}^d$ offen und beschränkt, $\partial\Omega = \bar{\Omega} \setminus \Omega$ eine $d - 1$ dimensionale Untermannigfaltigkeit und $f : \bar{\Omega} \rightarrow \mathbb{R}^d$ stetig und auf Ω differenzierbar, mit sich stetig auf $\bar{\Omega}$ fortsetzenden ersten partiellen Ableitungen. Dann gilt:

$$\int_{\Omega} \nabla \cdot f d\mu = \int_{\partial\Omega} f \cdot N d\sigma$$

Hierbei ist N die äußere Normale und $N d\sigma$ das entsprechende Maß auf dem Rand $\partial\Omega$.

Beweis: Wir überdecken $\partial\Omega$ durch offenen Mengen $O \subset \mathbb{R}^d$ wie in Definition 12.37. In dem Beweis haben wir gesehen, dass wir nach einer geeigneten Permutation der Komponenten von $\mathbb{R}^d$ die entsprechende Funktion f so wählen können, dass sie von der Gestalt $U \times (a, b) \rightarrow \mathbb{R}$ mit $(x, z) \mapsto z - g(x)$ ist. O.B.d.A. sei $\partial\Omega$ durch endlich viele offenen Mengen $U \times (a, b)$ überdeckt, so dass $\Omega \cap U \times (a, b) = \{(x, z) \mid z < g(x)\}$ gilt. Zusammen mit Ω erhalten wir eine endlich Überdeckung des Abschlusses von Ω . Dann wählen wir eine entsprechende Zerlegung der Eins. Wegen der Kompaktheit von $\bar{\Omega}$ und der lokalen Endlichkeit hat diese Zerlegung der Eins nur endlich viele Elemente. Wegen der Linearität genügt es die Aussage für jeden der Summanden einzeln zu zeigen.

Als erstes betrachten wir stetig differenzierbare Funktionen $f : \Omega \rightarrow \mathbb{R}^d$, die außerhalb einer kompakten Teilmenge von Ω verschwinden. Durch den Wert Null außerhalb von Ω setzen sich diese Funktionen stetig differenzierbar auf $\mathbb{R}^d$ fort. Wir wählen einen endlichen Quader, der Ω enthält. Dann verschwindet f auf dem Rand des Quaders. Wir integrieren beim i -tem Summanden von $\nabla \cdot f = \partial_1 f_1 + \dots, + \partial_d f_d$ mit Fubini zuerst über die Koordinate x_i . Wegen dem Hauptsatz ergibt dieses Integral die Differenz der Funktionswerte von f_i an den Randpunkten der entsprechenden Intervalle, also Null. Damit verschwinden in diesem ersten Fall beide Seiten des Gaußschen Satzes.

Als zweites betrachten wir für $1 \leq i < d$ die Komponenten f_i auf $\{(x, z) \in U \times (a, b) \mid z \leq g(x)\}$, die außerhalb einer kompakten Teilmenge verschwinden. Dann ist

$$\{(x, z) \in U \times (a, b) \mid z \leq g(x)\} \rightarrow \mathbb{R}^d, \quad (x, z) \mapsto \int_a^z f_i(x, y) dy$$

auf $\{(x, z) \in U \times (a, b) \mid z < g(x)\}$ stetig differenzierbar mit den partiellen Ableitungen

$$\frac{\partial}{\partial x_j} \int_a^z f_i(x, y) dy = \int_a^z \frac{\partial f_i(x, z)}{\partial x_j} dz \quad \text{für } 1 \leq j < d, \quad \frac{\partial}{\partial z} \int_a^z f_i(x, y) dy = f_i(x, z),$$

die sich stetig auf $\{(x, z) \in U \times (a, b) \mid z \leq g(x)\}$ fortsetzen. Weil f dann mit den ersten partiellen Ableitungen beschränkt und gleichmäßig stetig ist konvergiert im Grenzwert $\epsilon \downarrow 0$ die folgende Funktion mit ihren partiellen Ableitungen gleichmäßig auf U :

$$x \mapsto F_\epsilon(x) = \int_a^{g(x)-\epsilon} f(x, z) dz \quad \text{mit} \quad \frac{\partial F_\epsilon}{\partial x_i}(x) = \int_a^{g(x)-\epsilon} \frac{\partial f(x, z)}{\partial x_i} dz + \frac{\partial g(x)}{\partial x_i} f(x, g(x) - \epsilon).$$

Weil F_ϵ außerhalb einer kompakten Teilmenge von U verschwindet, folgt mit den Argumenten des ersten Falles, dass das Integral von der rechten Seite über U verschwindet. Für $1 \leq i < d$ folgt im Grenzwert $\epsilon \downarrow 0$ auch in diesem Fall der Gaußsche Satz:

$$\int_U \int_a^{g(x)} \frac{\partial f_i(x, z)}{\partial x_i} dz d\mu_{\mathbb{R}^{d-1}} = - \int_U f_i(x, g(x)) \frac{\partial g(x)}{\partial x_i} d\mu_{\mathbb{R}^{d-1}} = \int_U f_i(x, g(x)) N_i(x, g(x)) d\sigma.$$

Weil f auf $U \times \{a\}$ verschwindet folgt auch im letzten Fall aus dem Hauptsatz:

$$\int_U \int_a^{g(x)} \frac{\partial f_d(x, z)}{\partial x_d} dz d\mu_{\mathbb{R}^{d-1}} = \int_U f_d(x, g(x)) d\mu_{\mathbb{R}^{d-1}} = \int_U f_d(x, g(x)) N_d(x, g(x)) d\sigma. \mathbf{q.e.d.}$$

Beispiel 12.41. Sei $\Omega = B(0, r) \subset \mathbb{R}^d$ und $f(x) = x$. Auf dem Rand $\partial B(0, r) = \{x \in \mathbb{R}^d \mid \|x\| = r\}$ ist die äußere Normale gleich $N(x) = \frac{x}{\|x\|}$. Dann folgt mit $\nabla \cdot f = d$

$$\int_{B(0, r)} d \cdot d\mu = d \int \chi_{B(0, r)} d\mu = \int_{\partial B(0, r)} x \cdot \frac{x}{\|x\|} d\sigma = r \int_{\partial B(0, r)} d\sigma.$$

Also ist die Oberfläche von $\partial B(0, r)$ gleich $\frac{d}{r}$ mal dem Volumen von $B(0, r)$.

12.9 Maßtheorie

In diesem Abschnitt untersuchen wir, wann wir eine lebesgueintegrierbare Funktion über eine Teilmenge integrieren können. Das führt dann zu einer allgemeineren Definition vom Volumen von messbaren Mengen. Dieses Volumen heißt Maß.

Definition 12.42. Eine Teilmenge $\mathcal{B}$ der Potenzmenge $\mathcal{P}(X)$ einer Menge X heißt σ -Algebra, wenn diese Teilmenge $\mathcal{B} \subset \mathcal{P}(X)$ unter Komplementbildung und dem Schnitt von abzählbar vielen Elementen von $\mathcal{B}$ abgeschlossen ist. Wenn X ein metrischer Raum ist, dann heißen die Elemente der kleinsten σ -Algebra, die alle offenen (und abgeschlossenen) Mengen enthält, Borelmengen.

Definition 12.43. Sei $\mathcal{B}$ eine σ -Algebra auf der Menge X . Dann heißt $\mu : \mathcal{B} \rightarrow \bar{\mathbb{R}}_0^+$ Maß, wenn für alle Folgen $(A_n)_{n \in \mathbb{N}}$ von paarweise disjunkten Mengen in $\mathcal{B}$ gilt

$$\mu \left(\bigcup_{n=1}^{\infty} A_n \right) = \sum_{n=1}^{\infty} \mu(A_n) \quad (\sigma\text{-Additivität}).$$

Satz 12.44. (Lebesguemaß) Seien $\mathcal{L}(\mathbb{R}^d)$ die messbaren Mengen von $\mathbb{R}^d$. Dann definiert

$$\mu : \mathcal{L}(\mathbb{R}^d) \rightarrow \bar{\mathbb{R}}_0^+, \quad A \mapsto \lim_{n \rightarrow \infty} \int \chi_{Q_n} \chi_A d\mu \quad \text{mit} \quad Q_n = [-n, n]^d \quad \text{für alle } n \in \mathbb{N}$$

ein Maß auf $\mathcal{L}(\mathbb{R}^d)$. Es heißt Lebesguemaß. Weil die Borelmengen $\mathcal{B}(\mathbb{R}^d)$ messbar sind, definiert es auch ein Borelmaß auf $\mathcal{B}(\mathbb{R}^d)$.

Beweis: Wegen Satz 12.29 bilden die messbaren Mengen eine σ -Algebra, die die Borelmengen enthält. Seien $(A_n)_{n \in \mathbb{N}}$ paarweise disjunkte messbare Mengen und

$$f_n = \sum_{k=1}^n \chi_{A_k} = \chi_{A_1 \cup \dots \cup A_n} \quad \text{für alle } n \in \mathbb{N}.$$

Dann konvergiert $(f_n)_{n \in \mathbb{N}}$ punktweise gegen $f = \chi_A$ mit $A = \bigcup_{n \in \mathbb{N}} A_n$. Für alle $g \in L^1(\mathbb{R}^d)$ sind $|gf_n|$ durch $|g|$ beschränkt. Wegen Lebesgues Satz der beschränkten Konvergenz konvergiert die Folge $(gf_n)_{n \in \mathbb{N}}$ gegen $gf \in L^1(\mathbb{R}^d)$, und es gilt

$$\lim_{n \rightarrow \infty} \int gf_n d\mu = \int gf d\mu.$$

Wenn wir für g die Folge $(\chi_{Q_n})_{n \in \mathbb{N}}$ der charakteristischen Funktionen von den Quadern $Q_n = [-n, n]^d$ einsetzen, dann konvergiert die entsprechende monotone Folge $(\chi_{Q_n} f)_{n \in \mathbb{N}}$ entweder gegen eine lebesgueintegrierbare Funktion, und $\mu(\bigcup_{n \in \mathbb{N}} A_n)$ ist endlich, oder das Maß der disjunkten Vereinigung der Borelmengen $(A_n)_{n \in \mathbb{N}}$ ist unendlich. In beiden Fällen folgt die σ -Additivität aus den Rechenregeln für Folgen. **q.e.d.**

Lemma 12.45. (Translationsinvarianz des Lebesguesmaßes) Für $x \in \mathbb{R}^d$ ist eine Menge $A \subset \mathbb{R}^d$ genau dann messbar bzw. eine Borelmenge, wenn $A+x = \{y \in \mathbb{R}^d \mid y-x \in A\}$ messbar bzw. eine Borelmenge ist. Wenn A messbar ist, dann gilt $\mu(A+x) = \mu(A)$.

Beweis: Für alle Quader $Q \subset \mathbb{R}^d$ und $x \in \mathbb{R}^d$ ist $Q+x$ ein Quader mit $\mu(Q+x) = \mu(Q)$. Deshalb folgt $f_x \in L^1(\mathbb{R}^d)$ und $\int f_x d\mu = \int f d\mu$ aus $f \in L^1(\mathbb{R}^d)$ und $x \in \mathbb{R}^d$ für $f_x(y) = f(y-x)$. Für jede Menge $A \subset \mathbb{R}^d$ gilt $\chi_{A+x} = (\chi_A)_x$. Das Lemma folgt. **q.e.d.**

Satz 12.46. (Vitali) Die Relation $x \sim y \iff x - y \in \mathbb{Q}$ definiert eine Äquivalenzrelation. Für jedes $x \in \mathbb{R}$ gibt es genau ein $n_x \in \mathbb{Z}$ mit $x \in [n_x, n_x + 1)$. Deshalb enthält jede Äquivalenzklasse $[x]$ einen Repräsentanten $x - n_x \in [0, 1)$. Eine Menge $V \subset [0, 1)$, die für jedes $x \in \mathbb{R}$ genau einen Repräsentanten in der entsprechenden Äquivalenzklasse $[x]$ enthält, heißt Vitalimenge. Solche Mengen sind nicht messbar.

Die Definition der Menge V benutzt das sogenannte Auswahlaxiom der Mengenlehre. Es besagt, dass aus jeder nicht leeren Menge ein Element ausgewählt werden kann. In der Mengenlehre unterscheidet man zwischen Aussagen, die das Auswahlaxiom voraussetzen, und solchen, die das nicht tun. Dieser Satz gehört zu den ersteren.

Beweis: Sei $(q_n)_{n \in \mathbb{N}}$ eine Abzählung von $\mathbb{Q} \cap [-1, 1]$ und $V_n = q_n + V$ für alle $n \in \mathbb{N}$. Aufgrund der Definition von V sind die Mengen $(V_n)_{n \in \mathbb{N}}$ paarweise disjunkt und es gilt

$$[0, 1] \subset \bigcup_{n \in \mathbb{N}} V_n \subset [-1, 2).$$

Wegen dem vorangehenden Lemma sind entweder alle Mengen $(V_n)_{n \in \mathbb{N}}$ messbar oder keine. Im ersten Fall haben alle das gleiche Maß und aus der σ -Additivität folgt $1 \leq \sum_{n=1}^{\infty} \mu(V_n) = \sum_{n=1}^{\infty} \mu(V) \leq 3$, was unmöglich ist. Also ist V nicht messbar. **q.e.d.**

Dieser Satz zeigt, dass es unter der Annahme des Auswahlaxiomes unmöglich ist ein Maß auf allen Teilmengen von $\mathbb{R}$ zu definieren, das translationsinvariant ist, also das vorhergehende Lemma erfüllt, und $[0, 1]$ ein positives endliches Maß zuordnet.

Definition 12.47. Eine Äquivalenzklasse von fast überall definierten reellen Funktionen auf $\mathbb{R}^d$ heißt messbar, wenn es eine Folge von Treppenfunktionen gibt, die fast überall gegen einen Repräsentanten der Äquivalenzklasse konvergiert.

Satz 12.48. (i) Die messbaren Funktionen bilden mit der punktweisen Addition und Multiplikation und der Skalarmultiplikation eine Algebra, die $L^1(\mathbb{R}^d)$ enthält.

(ii) Wenn f und g messbar sind, dann auch $|f|$, $\max\{f, g\}$ und $\min\{f, g\}$. Ist außerdem fast überall f ungleich Null, dann ist auch $\frac{1}{f}$ messbar.

(iii) Eine messbare Funktion ist genau dann lebesgueintegrierbar, wenn ihr Betrag durch eine lebesgueintegrierbare Funktion beschränkt ist.

(iv) Der Grenzwert einer fast überall konvergenten Folge von messbaren Funktionen ist messbar.

(v) Die Äquivalenzklassen von stetigen Funktionen sind messbar.

(vi) Eine Menge $A \subset \mathbb{R}^d$ ist genau dann messbar, wenn χ_A messbar ist.

Beweis (i): Mit den Treppenfunktionen bilden die messbaren Funktionen wegen den Rechenregeln für Grenzwerte von Folgen eine Algebra. Alle lebesgueintegrierbaren Funktionen sind als Äquivalenzklassen von Grenzwerten von Treppenfunktionen messbar.

(ii): Der Betrag, das Maximum und das Minimum von endlich vielen Treppenfunktionen sind wieder Treppenfunktionen. Wenn eine Folge $(f_n)_{n \in \mathbb{N}}$ von Treppenfunktionen fast überall gegen eine nicht verschwindende Funktion f konvergiert, dann konvergiert die folgende Folge von Treppenfunktionen fast überall gegen $\frac{1}{f}$:

$$(g_n)_{n \in \mathbb{N}} \quad \text{mit} \quad g_n : \mathbb{R}^d \rightarrow \mathbb{R}, \quad x \mapsto g_n(x) = \begin{cases} \frac{1}{f_n(x)} & \text{wenn } f_n(x) \neq 0 \\ 0 & \text{wenn } f_n(x) = 0 \end{cases}$$

(iii): Sei $(f_n)_{n \in \mathbb{N}}$ eine Folge von Treppenfunktionen, die fast überall gegen f konvergiert, das $|f| \leq k$ für ein $k \in L^1(\mathbb{R}^d)$ erfüllt. Dann konvergiert auch $(\max\{-k, \min\{k, f_n\}\})_{n \in \mathbb{N}}$ fast überall gegen f . Aus Lebesgue's Satz der beschränkten Konvergenz folgt (v).

(iv): Sei $(f_n)_{n \in \mathbb{N}}$ eine Folge von messbaren Funktionen, die fast überall gegen f konvergiert. Sei h folgende positive lebesgueintegrierbare Hilfsfunktion auf $\mathbb{R}^d$

$$h = \sum_{n=1}^{\infty} \frac{2^{-n-d}}{n^d - (n-1)^d} \chi_{(-n,n)^d \setminus (1-n,n-1)^d} = \sum_{n=1}^{\infty} \left(\frac{2^{-n-d}}{n^d - (n-1)^d} - \frac{2^{-n-1-d}}{(n+1)^d - n^d} \right) \chi_{(-n,n)^d}$$

mit dem Integral $\int h d\mu = \sum_{n=1}^{\infty} 2^{-n} = 1$. Dann konvergiert die Folge $(g_n)_{n \in \mathbb{N}}$

$$g_n : \mathbb{R}^d \rightarrow \mathbb{R}, \quad x \mapsto g_n(x) = \frac{h(x)f_n(x)}{h(x) + |f_n(x)|} \quad \text{und} \quad |g_n| \leq h \quad \text{für alle } n \in \mathbb{N},$$

die wegen (ii) und (iii) aus messbaren Funktionen in $L^1(\mathbb{R}^d)$ besteht, fast überall gegen $g = \frac{hf}{h+|f|}$. Wegen Lebesgue's Satz der beschränkten Konvergenz ist g lebesgueintegrierbar, also wegen (i) messbar. Dann ist wegen (i)–(iii) auch folgende Funktion messbar:

$$\frac{hg}{h - |g|} = \frac{h \frac{hf}{h+|f|}}{h - \frac{h|f|}{h+|f|}} = \frac{h^2 f}{h^2 + h|f| - h|f|} = f.$$

(v): Für jede stetige Funktion f ist $(\chi_{(-n,n)^d} f)_{n \in \mathbb{N}}$ wegen Satz 12.15 eine Folge in $L^1(\mathbb{R}^d)$. Sie konvergiert punktweise gegen f . Also folgt (iv) aus (i) und (iv).

(iv): Für eine messbare Menge $A \subset \mathbb{R}^d$ konvergiert $(\chi_A \chi_{(-n,n)^d})_{n \in \mathbb{N}}$ punktweise gegen χ_A . Wegen (i) und (iv) ist χ_A messbar. Die Umkehrung folgt aus (i) und (iii). **q.e.d.**

Satz 12.49. Eine Äquivalenzklasse f von fast überall definierten reellen Funktionen auf $\mathbb{R}^d$ ist genau dann messbar, wenn eine der folgenden äquivalenten Bedingungen gilt:

- (i) Für alle $a \in \mathbb{R}$ sind die Mengen $\{x \in \mathbb{R}^d \mid f(x) \leq a\}$ messbar.
- (ii) Für alle $a \in \mathbb{R}$ sind die Mengen $\{x \in \mathbb{R}^d \mid f(x) \geq a\}$ messbar.
- (iii) Für alle $a \in \mathbb{R}$ sind die Mengen $\{x \in \mathbb{R}^d \mid f(x) < a\}$ messbar.

(iv) Für alle $a \in \mathbb{R}$ sind die Mengen $\{x \in \mathbb{R}^d \mid f(x) > a\}$ messbar.

(v) Für alle $k \geq 0$ in $L^1(\mathbb{R}^d)$ liegt die Funktion $\max\{-k, \min\{k, f\}\}$ in $L^1(\mathbb{R}^d)$.

Beweis: Für eine messbare Funktion f konvergiert $(n(\max\{a + \frac{1}{n}, f\} - \max\{a, f\}))_{n \in \mathbb{N}}$ fast überall gegen $\chi_{\{x \in \mathbb{R}^d \mid f(x) \leq a\}}$. Aus Satz 12.48 (i), (ii), (iv) und (vi) folgt (i).

(i) $\Leftrightarrow$ (ii) $\Leftrightarrow$ (iii) $\Leftrightarrow$ (iv): Wegen Satz 12.29 sind (i) und (iv) bzw. (ii) und (iii) äquivalent. Wegen $\{x \in \mathbb{R}^d \mid f(x) < a\} = \bigcup_{n \in \mathbb{N}} \{x \in \mathbb{R}^d \mid f(x) \leq a - \frac{1}{n}\}$ und $\{x \in \mathbb{R}^d \mid f(x) \leq a\} = \bigcap_{n \in \mathbb{N}} \{x \in \mathbb{R}^d \mid f(x) < a + \frac{1}{n}\}$ sind auch (i) und (iii) äquivalent.

(iv) $\Rightarrow$ (v): Wenn eine Funktion f die äquivalenten Bedingungen (i)-(iv) erfüllt, dann

$$\begin{aligned} \text{sind folgende Funktionen für alle } n \in \mathbb{N} \text{ messbar: } f_n &= \sum_{l=1}^{n^2} \frac{\chi_{f^{-1}[[\frac{l}{n}, \infty))} - \chi_{f^{-1}[(-\infty, -\frac{l}{n}]]}}{n} \\ &= n\chi_{f^{-1}[[n, \infty))} - n\chi_{f^{-1}[(-\infty, -n]]} + \sum_{l=1}^{n^2-1} \left(\frac{l}{n} \chi_{f^{-1}[[\frac{l}{n}, \frac{l+1}{n}])} - \frac{l}{n} \chi_{f^{-1}[(-\frac{l+1}{n}, -\frac{l}{n}]]} \right). \end{aligned}$$

Bei allen $x \in \mathbb{R}^d$ gilt $|f_n(x) - f(x)| < \frac{1}{n}$ für $n \geq |f(x)|$. Wegen Satz 12.48 (iv) ist f messbar und mit $k \geq 0$ liegt $\max\{-k, \min\{k, f\}\}$ wegen Satz 12.48 (ii)-(iii) in $L^1(\mathbb{R}^d)$. Wenn (v) gilt, dann liegt die Folge $(\max\{-n\chi_{(-n,n)^d}, \min\{n\chi_{(-n,n)^d}, f\}\})_{n \in \mathbb{N}}$ in $L^1(\mathbb{R}^d)$ und konvergiert punktweise gegen f . Wegen Satz 12.48 (i) und (iv) ist f messbar. **q.e.d.**

12.10 Die Räume $L^p(\mathbb{R}^d)$

Eine komplexwertige Funktion ist genau dann lebesgueintegrierbar bzw. messbar, wenn sowohl der Realteil als auch der Imaginärteil lebesgueintegrierbar bzw. messbar sind. Eine messbare komplexe Funktion ist wegen Satz 12.48 (iii) genau dann lebesgueintegrierbar, wenn der Absolutbetrag lebesgueintegrierbar ist. Aus Satz 12.49 folgt, dass für jede messbare Funktion f und $p \in \mathbb{R}^+$ auch die Funktion $|f|^p$ messbar ist.

Definition 12.50. Für alle $1 \leq p < \infty$ sei $L^p(\mathbb{R}^d)$ die Menge aller Äquivalenzklassen von fast überall definierten messbaren Funktionen von $\mathbb{R}^d$ nach $\mathbb{K}$, für die die Funktion $|f|^p$ lebesgueintegrierbar ist. Der Raum $L^\infty(\mathbb{R}^d)$ ist definiert als die Menge aller Äquivalenzklassen von messbaren beschränkten Funktionen von $\mathbb{R}^d$ nach $\mathbb{K}$.

Wegen Satz 12.48 (iii) stimmt dieser $L^1(\mathbb{R}^d)$ mit dem aus Definition 12.13 überein.

Satz 12.51. Für alle $1 \leq p \leq \infty$ ist $L^p(\mathbb{R}^d)$ zusammen mit der Abbildung

$$\|\cdot\| : L^p(\mathbb{R}^d) \rightarrow \mathbb{R}, \quad f \mapsto \|f\|_p = \begin{cases} \left(\int |f|^p d\mu \right)^{\frac{1}{p}} & \text{für } 1 \leq p < \infty \\ \inf\{C \in [0, \infty) \mid |f| \leq C \text{ a.e.}\} & \text{für } p = \infty. \end{cases}$$

ein normierter Vektorraum. Für $1 \leq p, q, r \leq \infty$ mit $\frac{1}{p} + \frac{1}{q} = \frac{1}{r}$ und $f \in L^p(\mathbb{R}^d)$ und $g \in L^q(\mathbb{R}^d)$ gilt $fg \in L^r(\mathbb{R}^d)$ und die Höldersche Ungleichung: $\|fg\|_r \leq \|f\|_p \|g\|_q$.

Die Dreiecksungleichung der Norm $\|\cdot\|_p$ wird wieder Minkowski-Ungleichung genannt: für alle $f, g \in L^p(\mathbb{R}^d)$ gilt $f + g \in L^p(\mathbb{R}^d)$ und $\|f + g\|_p \leq \|f\|_p + \|g\|_p$.

Beweis: Wir beweisen zuerst die Höldersche Ungleichung. Im Fall $r = \infty$ sind beide anderen Exponenten auch gleich $p = \infty = q$, und die Hölderungleichung folgt aus den Eigenschaften des Betrags. Für $1 \leq r < \infty$ gehen wir anstatt der Funktionen f und g zu den Funktionen $|f|^r$ und $|g|^r$, und anstatt der Exponenten r , p und q zu den Exponenten 1 , $\frac{r}{p}$ und $\frac{r}{q}$ über. Dann genügt es den Fall $1 = \frac{1}{p} + \frac{1}{q}$ für nicht-negative reelle Funktionen zu betrachten. Es gilt nämlich $\|f\|_p^r = \| |f|^r \|_{\frac{p}{r}}$ bzw. $\|g\|_q^r = \| |g|^r \|_{\frac{q}{r}}$. Für $p = 1$, $q = \infty$ bzw. $p = \infty$, $q = 1$ folgt die Höldersche Ungleichung aus den Eigenschaften der lebesguesintegrierbaren Funktionen. Für $f = 0$ oder $g = 0$ ist die Aussage offensichtlich. Sei also $f \neq 0$, $g \neq 0$ und $1 < p, q < \infty$ mit $1 = \frac{1}{p} + \frac{1}{q}$. Die Youngsche Ungleichung ergibt fast überall für die Funktionen $\frac{|f|}{\|f\|_p}$ und $\frac{|g|}{\|g\|_q}$

$$\frac{|f|}{\|f\|_p} \frac{|g|}{\|g\|_q} \leq \frac{1}{p} \frac{|f|^p}{\|f\|_p^p} + \frac{1}{q} \frac{|g|^q}{\|g\|_q^q}$$

Wegen $f \in L^p(\mathbb{R}^d)$ und $g \in L^q(\mathbb{R}^d)$ ist die rechte Seite lebesguesintegrierbar und das Integral gleich $\frac{1}{p} + \frac{1}{q} = 1$. Also ist auch die linke Seite lebesguesintegrierbar und das Integral kleiner oder gleich 1. Daraus folgt $fg \in L^1(\mathbb{R}^d)$, $\|fg\|_1 \leq \|f\|_p \|g\|_q$.

Wegen Korollar 12.20 erfüllt die Abbildungen $f \mapsto \|f\|_p$ die erste Eigenschaft einer Norm. Die zweite Eigenschaft ist offensichtlich. Für $p = 1$ und $p = \infty$ ist die Dreiecksungleichung schon gezeigt. Sei also $1 < p < \infty$ und $f, g \in L^p(\mathbb{R}^d)$. Offenbar gilt für $f, g \in L^p(\mathbb{R}^d)$ fast überall $|f + g|^p \leq 2^p \max\{|f|, |g|\}^p \leq 2^p(|f|^p + |g|^p)$. Also liegt $f + g$ in $L^p(\mathbb{R}^d)$. Für $f, g \in L^p(\mathbb{R}^d)$ folgt aus $|f + g|^p = |f + g| |f + g|^{p-1} \leq (|f| + |g|) |f + g|^{p-1}$ mit $1 = \frac{1}{p} + \frac{1}{q} \Leftrightarrow (p-1)q = p \Leftrightarrow \frac{1}{q} = \frac{p-1}{p}$, also $\| |f + g|^{p-1} \|_q = \|f + g\|_p^{p-1}$, auch

$$\|f + g\|_p^p \leq \|f \cdot |f + g|^{p-1}\|_1 + \|g \cdot |f + g|^{p-1}\|_1 \leq (\|f\|_p + \|g\|_p) \|f + g\|_p^{p-1}.$$

aus der Hölderschen Ungleichung. Daraus folgt die Minkowski Ungleichung. **q.e.d.**

Satz 12.52. (Satz von Riesz Fischer) Für alle $1 \leq p \leq \infty$ ist $L^p(\mathbb{R}^d)$ ein Banachraum.

Beweis: Für eine Cauchyfolge $(f_n)_{n \in \mathbb{N}}$ in $L^\infty(\mathbb{R}^d)$ sei $(g_n)_{n \in \mathbb{N}}$ eine Folge von Repräsentanten in $B(\mathbb{R}^d, \mathbb{R})$ mit $\|g_n\|_\infty = \|f_n\|_\infty$. Für alle $(n, m) \in \mathbb{N}^2$ ist $\{x \in \mathbb{R}^d \mid |g_m(x) - g_n(x)| > \|f_m - f_n\|_\infty\}$ eine Nullmenge, also auch ihre Vereinigung S . Wenn wir auf S die Funktionswerte von allen g_n durch Null ersetzen, wird $\|g_m - g_n\|_\infty = \|f_m - f_n\|_\infty$. Dann folgt die Vollständigkeit von $L^\infty(\mathbb{R}^d)$ folgt aus den Sätzen 9.44 (iii) und 12.48 (iv).

Für eine Cauchyfolge $(f_n)_{n \in \mathbb{N}_0}$ in $L^p(\mathbb{R}^d)$ mit $1 \leq p < \infty$ sei $(f_{n_m})_{m \in \mathbb{N}_0}$ eine Teilfolge mit $\|f_{n_m} - f_{n_{m-1}}\|_p \leq 2^{-m}$ für alle $m \in \mathbb{N}$. Die monoton wachsende Folge $g_m = |f_{n_0}| + \sum_{l=1}^m |f_{n_l} - f_{n_{l-1}}|$ für alle $m \in \mathbb{N}$ erfüllt wegen der Minkowskiungleichung $\|g_m\|_p \leq \|f_{n_0}\|_p + 1$ für alle $m \in \mathbb{N}$. Wegen dem Satz der monotonen Konvergenz, konvergiert $(g_m^p)_{m \in \mathbb{N}}$ gegen eine Funktion $k \in L^1(\mathbb{R}^d)$. Wegen dem Monotonieprinzip konvergiert

$(g_m)_{m \in \mathbb{N}}$ fast überall gegen eine messbare Funktion g . Wegen $|f_{n_m} - f_{n_l}| \leq |g_m - g_l|$ konvergiert $(f_{n_m})_{m \in \mathbb{N}}$ fast überall gegen eine messbare Funktionen f . Wegen Lebesgue's Satz der beschränkten Konvergenz ist $|f|^p$ lebesgueintegabel und damit $f \in L^p(\mathbb{R}^d)$. Weil fast überall $|f_{n_m} - f| \leq |g_m - g|$ und $\|g_m - g\|_p \leq 2^{-m}$ gilt, konvergiert $(f_{n_m})_{m \in \mathbb{N}}$ in $L^p(\mathbb{R}^d)$ gegen f . Als Cauchyfolge konvergiert auch $(f_n)_{n \in \mathbb{N}_0}$ in $L^p(\mathbb{R}^d)$ gegen f . **q.e.d.**

Definition 12.53. Für jede messbare Menge $A \subset \mathbb{R}^d$ und alle $1 \leq p \leq \infty$ sei

$$L^p(A) = \{f \in L^p(\mathbb{R}^d) \mid f|_{\mathbb{R}^d \setminus A} = 0 \text{ a.e.}\}$$

Der vorangehende Beweis zeigt, dass $L^p(A)$ in $L^p(\mathbb{R}^d)$ abgeschlossen ist.

Korollar 12.54. Für A messbar und $1 \leq p \leq \infty$ ist $L^p(A)$ ein Banachraum. **q.e.d.**

Satz 12.55. Für alle messbaren Mengen $A \subset \mathbb{R}^d$ und alle $1 \leq p, q \leq \infty$ mit $\frac{1}{p} + \frac{1}{q} = 1$ ist folgende Abbildung eine Isometrie:

$$j : L^q(A) \rightarrow \mathcal{L}(L^p(A), \mathbb{K}), \quad g \mapsto j(g) \text{ mit } j(g) : L^p(A) \rightarrow \mathbb{K}, \quad f \mapsto \int_A f g d\mu.$$

Beweis: Wegen der Hölderschen Ungleichung ist j lipschitzstetig mit $L = 1$.

Für $1 \leq q < \infty$ und $g \in L^q(A) \setminus \{0\}$ definieren wir eine messbare Funktion f so, dass in der Youngschen Ungleichung im Beweis der Hölderschen Ungleichung Gleichheit gilt:

$$f(x) = \begin{cases} \frac{|g(x)|^{\frac{q}{p}+1}}{g(x)} = \frac{|g(x)|^{q(\frac{1}{p}+\frac{1}{q})}}{g(x)} = \frac{|g(x)|^q}{g(x)} & \text{für } g(x) \neq 0 \\ 0 & \text{für } g(x) = 0. \end{cases}$$

Wegen $\|f\|_p^p = \|g\|_q^q = \|fg\|_1$ ist $f \in L^p(A)$, und j tatsächlich eine Isometrie:

$$\|g\|_q = \|g\|_q^{q-q(1-\frac{1}{q})} = \frac{\|g\|_q^q}{\|g\|_q^{\frac{q}{p}}} = \frac{\int_A g f d\mu}{\|f\|_p} = \frac{j(g)(f)}{\|f\|_p} \leq \|j(g)\| \leq \|g\|_q$$

Für $p = 1, q = \infty$ gibt es für jedes $\epsilon > 0$ eine messbare Teilmenge B von der messbaren Menge $\{x \in A \mid |g(x)| \geq \|g\|_\infty - \epsilon\}$ mit $0 < \mu(B) < \infty$. Dann gilt für $f = \chi_B \bar{g} \in L^1(A)$

$$\|g\|_\infty - \epsilon \leq \frac{\int_B |g|^2 d\mu}{\int_B |g| d\mu} = \frac{j(g)(f)}{\|f\|_1} \leq \|j(g)\| \leq \|g\|_\infty.$$

Weil das für alle $\epsilon > 0$ gilt, ist j auch für $p = 1$ eine Isometrie.

q.e.d.

Mithilfe von weiteren Sätzen der Maßtheorie kann man zeigen, dass für $1 < q \leq \infty$ diese Abbildung j sogar ein Isomorphismus von Banachräumen ist.

Index

- Äquivalenzrelation, 9
 - $\sim$ von Normen, 111
- Abbildung, 9
 - bijektive $\sim$, 10
 - Bild einer $\sim$, 9
 - bilineare $\sim$, 139
 - Definitionsbereich einer $\sim$, 9
 - differenzierbare $\sim$, 129
 - stetig $\sim$, 133
 - identische $\sim$, 10
 - injektive $\sim$, 10
 - Komposition von $\sim$ en, 10
 - lineare $\sim$, 125
 - stetige $\sim$, 126
 - Nachfolger $\sim$, 18, 20
 - stetige $\sim$, 59, 114
 - surjektive $\sim$, 10
 - Umkehr $\sim$, 10
 - differenzierbare $\sim$, 147
 - Wertebereich einer $\sim$, 9
- Abel, Niels Henrik 1802-1829
 - $\sim$ scher Grenzwertsatz, 101
- Ableitung
 - $\sim$ der Umkehrfunktion, 73
 - $\sim$ einer
 - $\sim$ Abbildung, 129
 - $\sim$ Funktion, 71
 - $\sim$ Potenzreihenfunktion, 72
 - $\sim$ und Konvexität, 79, 80
 - $\sim$ und Monotonie, 76
 - höhere $\sim$, 82
 - $\sim$ und Konvexität, 80
 - partielle $\sim$, 133
 - stetige $\sim$, 134
- Abschluß einer Menge, 58, 111
- absolut
 - $\sim$ e Konvergenz, 44
- Abstand $\rightarrow$ Metrik, 16, 30
- Addition
 - $\sim$ stheorem, 55
 - Axiome der $\sim$, 11
- Algebra
 - σ - $\sim$, 179
 - Banach $\sim$, 119
 - $B(X, Y)$, 121
 - $C_b(X, Y)$, 121
 - $\mathcal{L}(V)$, 127
 - Fundamentalsatz der $\sim$, 70
- analytisch
 - reell $\sim$ e Funktion, 84
- Ankathete, 69
- Archimedes von Syrakus 287 a.C.–212 a.C.
 - $\sim$ ischer Körper, 23
 - Satz von $\sim$ –Eudoxos, 23
- Arcus
 - $\sim$ cosinus, 68
 - $\sim$ cotangens, 68
 - $\sim$ sinus, 68
 - $\sim$ tangens, 68
- Argument, 69
- arithmetisch
 - $\sim$ es Mittel, 15
- Arzelà, Cesare 1847-1912
 - Satz von Arzela Ascoli, 124
- Ascoli, Giulio 1843-1896
 - Satz von Arzela Ascoli, 124

- Axiome
 - $\sim$ der Addition, 11
 - $\sim$ der Multiplikation, 11
 - Auswahl $\sim$, 180
 - Distributivgesetz, 12
 - Ordnungs $\sim$, 14
 - Peano $\sim$, 27
 - Vollständigkeits $\sim$, 22, 38
- Ball, 57, 110
- Banach, Stefan 1892–1945
 - $\sim$ algebra, 119
 - $B(X, Y)$, 121
 - $C_b(X, Y)$, 121
 - $\mathcal{L}(V)$, 127
 - $\sim$ raum, 112
 - $B(X, Y)$, 121
 - $C_b(X, Y)$, 121
 - $L^1(A)$, 170
 - $L^1(\mathbb{R}^d)$, 167
 - $L^p(\mathbb{R}^d)$, 182
 - $\mathcal{L}(V)$, 127
 - $\mathcal{L}(V, W)$, 126
 - $\sim$ scher Fixpunktsatz, 143
- Beppo Levi 1875–1961
 - Satz von $\sim$, 165
- Bernoulli, Johann 1667–1748
 - $\sim$ Ungleichung, 19
 - $\sim$ Zahlen, 104
- beschränkt
 - $\sim$ e Folge, 35
 - $\sim$ e Funktion, 62, 120
 - $\sim$ e Menge, 17, 58, 114
- Betrag, 15, 30
- Beweis
 - $\sim$ durch vollständige Induktion, 19
- Bild, 9
 - Ur $\sim$, 10
- bilinear
 - $\sim$ e Abbildung, 139
- binomische Formel, 40
- Bolzano, Bernhard Placidus Johann Nepomuk 1781–1848
 - Auswahlprinzip von $\sim$ –Weierstraß, 37
- Borel, Felix Edouard Justin Emile 1871–1956, 59
 - $\sim$ maß, 179
 - $\sim$ menge, 179
 - Satz von $\sim$, 103
 - Satz von Heine $\sim$, 59
- Bruch
 - Partial $\sim$ zerlegung, 90
- Cantor, Georg Ferdinand Ludwig Philipp 1845–1918, 7
 - $\sim$ menge, 154, 164
- Cauchy, Augustin Louis 1789–1857
 - $\sim$'s Vedichtungssatz, 46
 - $\sim$ –Produkt von Reihen, 48
 - $\sim$ –Schwarzsche Ungleichung, 108
 - $\sim$ folge, 37, 112
 - $\sim$ kriterium, 37, 99
 - $\sim$ für Reihen, 44
- charakteristisch
 - $\sim$ e Funktion, 85, 156
- Cosinus, 55, 66–69
- Cotangens cot, 68
 - Partialbruchzerlegung von $\sim$, 106
- Darboux, Jean Gaston 1842–1917
 - Kriterium von $\sim$, 93
- Darstellung
 - Dezimabruch $\sim$, 47
 - Polar $\sim$, 69
- de L'Hôpital, Guillaume Francois Antoine Marquis 1661–1704
 - 1. Regel von $\sim$, 78
 - 2. Regel von $\sim$, 78
- de Morgan, Augustus 1806–1871
 - $\sim$ sche Regel, 8
- Dedekind, Julius Wilhelm Richard 1831–1916
 - $\sim$ sche Schnitte, 28

- Definitionsbereich, 9
- Derivation, 128
- Diffeomorphismus, 150
- Differential
 - $\sim$ einer Funktion, 71
 - $\sim$ - und Integralrechnung
 - Hauptsatz der $\sim$, 86, 97
 - gewöhnliche $\sim$ gleichung, 143
 - Existenz und Eindeutigkeit von $\sim$ en, 144, 145
- differenzierbar
 - $\sim$ e Abbildung, 129
 - $\sim$ e Funktion, 71
 - $\sim$ e Niveaumenge, 151
- Dini, Ulisse 1845–1918
 - Satz von $\sim$, 123
- Divergenz
 - $\sim$ eines Vektorfeldes, 137
 - $\sim$ satz, 177
- Doppelpunkt, 153
- Drehung, 69
- Dreiecksungleichung, 30, 107
- Eigenschaften
 - $\sim$ der Exponentialfunktion, 49, 65
 - $\sim$ der lebesgueintegrierbaren Funktionen, 160
 - $\sim$ des Abstandes, 16
 - $\sim$ des Betrags, 16, 30
 - $\sim$ des Logarithmus, 65
 - $\sim$ des Riemannintegrals, 95
 - $\sim$ von Potenzreihenfunktionen, 52
- Ein
 - $\sim$ s 1, 11
 - Zerlegung der $\sim$, 174
 - imaginäre $\sim$ heit, 29
- Euler, Leonard 1707–1783
 - $\sim$ sche Formel, 55
 - $\sim$ sche Konstante γ , 100
 - $\sim$ sche Zahl e , 42, 49
- Exponent
 - $\sim$ ialfunktion, 45, 49, 65
- Extremwert
 - lokaler $\sim$, 75
- fast überall, 158
- Fatou, Pierre Joseph Louis 1878–1929
 - $\sim$'s Lemma, 167
- Fischer, Ernst Sigismund 1875–1954
 - Satz von Riesz- $\sim$, 167, 183
- Fixpunkt
 - Banachscher $\sim$ satz, 143
- Fläche
 - Hyper $\sim$, 150
- Folge, 31
 - $\sim$ von Funktionen, 62, 119
 - $\sim$ nkompakt, 113
 - Cauchy $\sim$, 37, 112
 - Konvergenz einer $\sim$, 31
 - monoton fallende $\sim$, 35
 - monoton wachsende $\sim$, 35
 - monotone $\sim$, 35
 - Null $\sim$, 31
 - streng monoton fallende $\sim$, 35
 - streng monoton wachsende $\sim$, 35
 - Teil $\sim$, 36
 - Grenzwert einer $\sim$, 38
 - konvergente $\sim$, 37
 - monotone $\sim$, 37
 - Zahlen $\sim$, 31
- Formel
 - binomische $\sim$, 40
 - Eulersche $\sim$, 55
 - Taylor $\sim$, 83, 98
- Fubini, Guido 1879–1943
 - Satz von $\sim$, 169
- Fundamentalsatz
 - $\sim$ der Algebra, 70
- Funktion
 - Γ - $\sim$, 101
 - $\sim$ swert
 - Grenzwert von $\sim$ en, 77

- Arcus
 - $\sim$ cosinus, 68
 - $\sim$ cotangens, 68
 - $\sim$ sinus, 68
 - $\sim$ tangens, 68
- beschränkte $\sim$, 62, 120
- charakteristische $\sim$, 85, 156
- Cosinus, 55, 66–69
- Cotangens, 68
 - Partialbruchzerlegung der $\sim$, 106
- differenzierbare $\sim$, 71
- Exponential $\sim$, 45, 49, 65
- Folge von $\sim$ en, 62, 119
- Gradient einer $\sim$, 137
- Graph einer $\sim$, 71
- implizite $\sim$
 - Satz über die $\sim$, 149
- inverse $\sim$
 - Satz über die $\sim$, 148
- konkave $\sim$, 79
- konvexe $\sim$, 79
- lebesgueintegrabale $\sim$, 160
- Logarithmus, 65
 - $\sim$ zur Basis a , 66
 - natürlicher $\sim$, 65
- monoton fallende $\sim$, 63
- monoton wachsende $\sim$, 63
- Monotonie einer $\sim$, 63
- reellanalytische $\sim$, 84
- Regel $\sim$, 85
- riemannintegrabale $\sim$, 93, 98, 162
- Riemannsche ζ - $\sim$, 43, 100, 106
- Sinus, 55, 66–69
 - Produktzerlegung der $\sim$, 106
- Stamm $\sim$, 87
- streng monoton fallende $\sim$, 64
- streng monoton wachsende $\sim$, 63
- Tangens, 68
- Treppen $\sim$, 85, 156
- trigonometrische $\sim$ en, 66–68
- Gauß, Johann Carl Friedrich 1777–1855
 - $\sim$ scher Satz, 177
- Gegenkathete, 69
- Gleichung
 - Differential $\sim$, 143
 - gewöhnliche $\sim$, 143
- Un $\sim$
 - $\sim$ von Jensen, 80
 - Bernoulli $\sim$, 19
 - Cauchy–Schwarzsche $\sim$, 108
 - Dreiecks $\sim$, 30, 107
 - Hölder’sche $\sim$, 110, 182
 - Minkowski- $\sim$, 108, 110, 183
 - Youngsche $\sim$, 81
- Gradient einer Funktion, 137
- Graph, 71
- Grenzwert $\lim$, 31
 - $\sim$ einer Reihe, 43
 - $\sim$ einer Teilfolge, 38
 - $\sim$ einer Zahlenfolge, 31
 - $\sim$ von Funktionswerten, 77
- Abelscher $\sim$ satz, 101
- Häufungspunkt, 38
- Höhenlinien, 149
- Hölder, Otto Ludwig 1859–1937
 - $\sim$ sche Ungleichung, 110, 182
- Hauptsatz, 86, 97
- Heine, Heinrich Eduard 1821–1881
 - Satz von $\sim$ –Borel, 59
- Hesse, Ludwig Otto 1811–1874
 - $\sim$ matrix, 141
- Hypothenuse, 69
- Identität
 - $\sim$ ssatz für Potenzreihenfunktionen, 53
- imaginär
 - $\sim$ e Einheit, 29
- Induktion
 - kontinuierliche $\sim$, 131
 - vollständige $\sim$, 19
- Integral

- ~ von Regelfunktionen, 86
- ~ von Treppenfunktionen, 85
- ~kriterium für Reihen, 99
- ~rechnung
 - Hauptsatz der Differential- und ~, 86, 97
 - Mittelwertsatz der ~, 98
- Lebesgue~, 160
 - Eigenschaften des ~, 160
- Ober~, 92, 162
- Riemann~, 93, 162
 - Eigenschaften des ~, 95
 - uneigentliches ~, 98
- Unter~, 92, 162
- Integration
 - ~ durch Substitution, 88
 - partielle ~, 89
- Intervall, 17
 - ~schachtelungsprinzip, 25
- Jacobi, Karl Gustav Jacob 1804–1851
 - ~matrix, 138
 - ~s Transformation von Maßen, 172
- Jensen, Johan Ludwig William Valdemar 1859–1925
 - Ungleichung von ~, 80
- Körper, 12
 - angeordneter ~, 11, 14
 - archimedischer ~, 23
- Kettenregel, 73
- Komplement, 8
 - ~ einer Nullmenge, 158
- komplex
 - ~e Konjugation, 29
 - ~e Zahlen, 28
- Komposition, 10
- Konjugation
 - komplexe ~, 29
- konkav, 79
- Konvergenz
 - ~ einer Folge, 31
 - ~prinzipien, 35
 - ~radius einer Potenzreihe, 51
 - absolute ~, 44
 - gleichmäßige ~, 62, 119
 - Lebesgue's Satz der beschränkten ~, 166
 - punktweise ~, 62, 119
- konvex, 79
 - ~ität
 - 2. Ableitung und ~, 80
 - Ableitung und ~, 79, 80
- Koordinaten
 - Polar~, 150
- Kriterium
 - ~ von Darboux, 93
 - ~ von Riemann, 94
 - Cauchy~, 37, 99
 - für Reihen, 44
 - Integral~ für Reihen, 99
 - Lebesgue~, 163
 - Majoranten~, 44, 99
 - Monotonie~, 99
- Lagrange, Joseph-Louis 1736–1813
 - ~multiplikator, 151
- Lebesgue, Henri Léon 1875–1941
 - ~'s Satz der beschränkten Konvergenz, 166
 - ~Integral, 160
 - ~integrale Funktion, 160
 - ~en auf A , 170
 - ~kriterium, 163
 - ~maß, 179
- Leibniz, Gottfried Wilhelm von 1646–1716
 - ~regel, 72, 82, 130
 - alternierende Reihe von ~, 46
- Limes $\lim$, 31
 - ~inferior $\underline{\lim}$, 39
 - ~superior $\overline{\lim}$, 39
 - von Funktionswerten, 77
- Lindelöf, Ernst 1870–1946

- Satz von Picard $\sim$, 143
- linear
 - $\sim$ e Abbildung, 125
 - stetige $\sim$, 126
 - bi $\sim$ e Abbildung, 139
 - nicht $\sim$ e Analysis, 143, 146–152
- Lipschitz, Rudolf Otto Sigismund 1832–1903
 - $\sim$ Stetigkeit, 61, 117
 - lokale $\sim$, 144
- Logarithmus, 65
 - $\sim$ zur Basis $a \log_a$, 66
 - natürlicher $\sim \ln$, 65
- Mächtigkeit einer Menge, 20
- Majorantenkriterium, 44
- Matrix
 - Hesse $\sim$, 141
 - Jacobi $\sim$, 138
- Maximum, 17
 - lokales $\sim$, 75, 139
- Maß, 179
- Menge, 7
 - abgeschlossene $\sim$, 58, 111
 - Abschluß einer $\sim$, 58, 111
 - abzählbare $\sim$, 21
 - über $\sim$, 21, 26
 - beschränkte $\sim$, 17, 58, 114
 - Borel $\sim$, 179
 - Cantor $\sim$, 154, 164
 - endliche $\sim$, 21
 - höchstens abzählbare $\sim$, 21
 - induktive $\sim$, 18
 - kompakte $\sim$, 58, 112
 - Mächtigkeit einer $\sim$, 20
 - messbare $\sim$, 170
 - nicht $\sim$, 180
 - Niveau $\sim$, 151
 - Null $\sim$, 155
 - Komplement einer $\sim$, 158
 - offene $\sim$, 57, 110
 - Potenz $\sim$, 8
 - Teil $\sim$ von $\mathbb{R}$
 - $\sim$ mit nicht abzählbarem Rand, 164
 - undendliche $\sim$, 21
 - Vitali $\sim$, 180
- Metrik, 16, 30, 107
 - des kartesischen Produktes, 107
 - diskrete $\sim$, 107
 - euklidische $\sim$, 108
- Minimum, 17
 - lokales $\sim$, 75, 139
- Minkowski, Hermann 1864–1909
 - $\sim$ –Ungleichung, 108, 110, 183
- Mittel
 - $\sim$ wertsatz, 76
 - $\sim$ der Integralsrechnung, 98
 - verallgemeinerter $\sim$, 76
 - arithmetisches $\sim$, 15
- Monotonie
 - $\sim$ der Ordnung, 14
 - $\sim$ einer Folge, 35
 - $\sim$ einer Funktion, 63
 - $\sim$ kriterium, 99
 - $\sim$ prinzip, 35
 - $\sim$ für Reihen, 44
 - Ableitung und $\sim$, 76
 - Satz der $\sim$ en Konvergenz, 165
- Multiplikation
 - Axiome der $\sim$, 11
 - Skalar $\sim$, 108
- Nachfolgerabbildung, 18, 20
- Neumann, Carl Gottfried 1832–1925
 - $\sim$ sche Reihe, 127
- nichtlineare Analysis, 143, 146–152
- Niveaumenge, 151
 - differenzierbare $\sim$, 151
 - kritischer Punkt auf $\sim$ n, 151
 - Singularität einer $\sim$, 151
- Norm, 108, 109
 - $\|\cdot\|_1$, 166

- $\| \cdot \|_\infty$, 62, 109, 182
- $\| \cdot \|_p$, 109, 182
- Äquivalenzrelation von $\sim$ en, 111
- euklidische $\sim$, 108
- Supremums $\sim$, 62, 109, 182
- Null 0, 11
 - $\sim$ folge, 31
 - $\sim$ menge, 155
 - Komplement einer $\sim$, 158
 - nicht abzählbare $\sim$, 164
- Ordnung, 14
 - Totalität der $\sim$, 14
 - Transitivität der $\sim$, 14
 - Um $\sim$
 - $\sim$ von Reihen, 50
- Partialbruchzerlegung, 90
 - $\sim$ von cot, 106
- Partition, 93
 - Verfeinerung einer $\sim$, 93
- Peano, Giuseppe 1858-1932
 - $\sim$ Axiome, 27
- Pi π , 66
- Picard, Emile 1856-1941, 143
 - Satz von $\sim$ Lindelöf, 143
- Polar
 - $\sim$ darstellung
 - $\sim$ einer komplexen Zahl, 69
- Polynom
 - $\sim$ division, 91
 - Taylor $\sim$, 82
- Potenz
 - $\sim$ menge, 8
 - $\sim$ reihe, 51
 - $\sim$ von Cosinus, 56
 - $\sim$ von Sinus, 56
 - $\sim$ nfunktion, 51, 128
 - Identitätssatz von $\sim$ nfunktionen, 53
 - Konvergenzradius einer $\sim$, 51
- Prinzip
 - Auswahl $\sim$ von Bolzano–Weierstraß, 37
- Intervallschachtelungs $\sim$, 25
- Konvergenz $\sim$ ien, 35
- Monotonie $\sim$, 35
 - $\sim$ für Reihen, 44
- Wohlordnungs $\sim$, 20
- Produkt
 - $\sim$ von Reihen, 48
 - $\sim$ zerlegung von sin, 106
 - kartesisches $\sim$, 8
 - Metrik des $\sim$, 107
 - Skalar $\sim$, 109
 - Tensor $\sim$, 139
 - unendliches $\sim$, 106
- Punkt
 - Doppel $\sim$, 153
 - Häufungs $\sim$, 38
 - kritischer $\sim$, 76, 138
 - $\sim$ auf Niveaumengen, 151
 - isolierter $\sim$, 77
- Quader, 155
- Quotienten
 - $\sim$ regel, 74
 - $\sim$ test, 45
- Rand
 - $\sim$ einer Menge
 - $\sim$ mit positivem Lebesguemaß, 164
- Rangsatz, 152
- rational
 - $\sim$ e Zahlen, 20
- Raum
 - $B(X, Y)$, 120
 - $B(X, \mathbb{K})$, 62
 - $C(I)$, 82
 - $C(X, \mathbb{K})$, 62
 - $C^n(I)$, 82
 - $C_b(X, Y)$, 120
 - $L^1(A)$, 170
 - $L^1(\mathbb{R}^d)$, 160
 - $\mathcal{L}(V)$, 127
 - $\mathcal{L}(V, W)$, 126

- Banach~, 112
 - $B(X, Y)$, 121
 - $C_b(X, Y)$, 121
 - $L^1(A)$, 170
 - $L^1(\mathbb{R}^d)$, 167
 - $L^p(\mathbb{R}^d)$, 182
 - $\mathcal{L}(V)$, 127
 - $\mathcal{L}(V, W)$, 126
- metrischer ~, 107
 - Vervollständigung eines ~, 118
 - vollständiger ~, 58, 112
- Tangential~, 150
- Vektor~, 108
- reell
 - ~analytische Funktion, 84
 - ~e Zahlen, 11, 22
- Reflexivität einer Relation, 9
- Regel
 - ~funktion, 85
 - 1. ~ von de L'Hopital, 78
 - 2. ~ von de L'Hopital, 78
 - de Morgansche ~n, 8
 - Ketten~, 73
 - Leibniz~, 72, 82, 130
 - Quotienten~, 74
 - Rechen~n
 - ~ der Ableitung, 72, 82, 130
 - ~ für Folgen, 33
 - ~ für Reihen, 48
 - Substitutions~, 88
- Reihe, 43
 - geometrische ~, 43
 - absolut konvergente ~, 44
 - alternierende ~
 - ~ von Leibniz, 46
 - Cauchy Kriterium für ~n, 44
 - Integralkriterium für ~n, 99
 - konvergente ~, 43
 - absolut ~, 44
 - Neumannsche ~, 127
 - Potenz~, 51
 - ~nfunktion, 51, 128
 - Konvergenzradius einer ~, 51
 - Produkt von ~n, 48
 - Rechenregeln für ~n, 48
 - Taylor~, 83, 142
 - Umordnung von ~n, 50
- Relation, 9
 - Ordnungs~, 14
 - Reflexivität einer ~, 9
 - Symmetrie einer ~, 9
 - Transitivität einer ~, 9
- Riemann, Georg Friedrich Bernhard 1826–1866, 51
 - ~integrabel, 93, 98, 162
 - ~integral, 93, 162
 - Eigenschaften des ~, 95
 - uneigentliches ~, 98
 - ~sche ζ -Funktion, 43, 100, 106
 - ~summen, 94
 - Kriterium von ~, 94
- Riesz, Frigyes 1880–1956
 - Satz von ~–Fischer, 167, 183
- Rolle, Michel 1652–1719
 - Satz von ~, 76
- Rotation
 - ~ eines Vektorfeldes, 137
- Russell, Bertrand Arthur William 1872–1970
 - ~sche Antinomie, 7
- Satz
 - ~ über die implizite Funktion, 149
 - ~ über die inverse Funktion, 148
 - ~ der monotonen Konvergenz, 165
 - ~ von Arzela Ascoli, 124
 - ~ von Beppo Levi, 165
 - ~ von Borel, 103
 - ~ von Dini, 123
 - ~ von Fubini, 169
 - ~ von Heine–Borel, 59
 - ~ von Picard Lindelöf, 143

- ~ von Riesz–Fischer, 167, 183
- ~ von Rolle, 76
- ~ von Schwarz, 139
- ~ von Stone–Weierstraß, 121
- ~ von Vitali, 180
- ~ von Whitney, 154
- Abelscher Grenzwert~, 101
- Banachscher Fixpunkt~, 143
- Cauchy’s Verdichtungs~, 46
- Divergenz~, 177
- Fundamental~ der Algebra, 70
- Gaußscher ~, 177
- Haupt~, 86, 97
- Identitäts~
 - ~ für Potenzreihenfunktionen, 53
- Lebesgue’s ~ der beschränkten Konvergenz, 166
- Mittelwert~, 76
 - ~ der Integralrechnung, 98
 - verallgemeinerter ~, 76
- Rang~, 152
- Schranken~, 76, 132
- Zwischenwert~, 63
- Schranke
 - ~nsatz, 76, 132
 - obere ~, 16
 - untere ~, 16
- Schwarz, Karl Herman Amandus 1843–1921
 - Cauchy–~sche Ungleichung, 108
 - Satz von ~, 139
- Sekante, 71
- Sierpinski, Waclaw Franciszek 1882–1962
 - ~teppich, 154
- Singularität, 151
 - ~ einer Niveaumenge, 151
- Doppelpunkt, 153
- Spitze, 153
- Sinus, 55, 66–69
 - Produktzerlegung von ~, 106
- Skalar
 - ~multiplikation, 108
 - ~produkt, 109
- Spitze, 153
- Stammfunktion, 87
- stetig
 - ~ differenzierbar, 134
 - ~ partiell differenzierbar, 134
 - ~e Abbildung, 114
 - ~keit, 59
 - gleichmäßige ~, 61, 117
 - Lipschitz~, 61, 117
- Stone, Marshall Harvey 1903–1989
 - Satz von ~–Weierstraß, 121
- Substitutionsregel, 88
- Summe
 - Ober~en, 93
 - Riemann~en, 94
 - Unter~en, 93
- Supremum, 17
 - ~snorm, 62, 109, 182
- Symmetrie einer Relation, 9
- Tangens tan, 68
- Tangente, 71
- Tangentialraum, 150
- Taylor, Brook 1685–1731
 - ~formel, 83, 98, 142
 - ~polynom, 82
 - ~reihe, 83, 142
- Teilfolge, 36
 - konvergente ~, 37
 - monotone ~, 37
- Tensorprodukt, 139
- Test
 - Quotienten~, 45
 - Wurzel~, 45
- Totalität der Ordnung, 14
- Transformation
 - ~ der Variablen, 88
 - Jacobis ~ von Maßen, 172
- Transitivität
 - ~ einer Relation, 9

- $\sim$ der Ordnung, 14
- Treppenfunktion, 85, 156
- trigonometrische Funktionen, 66–68
- Umgebung, 57, 110
- Umkehrabbildung, 10
 - Ableitung der $\sim$, 73
 - Stetigkeit der $\sim$, 64
- Umordnung von Reihen, 50
- Ungleichung
 - $\sim$ von Jensen, 80
 - Bernoulli $\sim$, 19
 - Cauchy–Schwarzsche $\sim$, 108
 - Dreiecks $\sim$, 30, 107
 - Hölder’sche $\sim$, 110, 182
 - Minkowski– $\sim$, 108, 110, 183
 - Youngsche $\sim$, 81
- Untermannigfaltigkeit, 149
- Urbild, 10
- Variable
 - Transformation der $\sim_n$, 88
- Vektor
 - $\sim$ feld, 137
 - Divergenz eines $\sim$ es, 137
 - Rotation eines $\sim$ es, 137
 - $\sim$ raum, 108
- Verdichtungssatz
 - Cauchy’s $\sim$, 46
- Verfeinerung, 93
- Vitali, Giuseppe 1875–1932
 - $\sim$ menge, 180
 - Satz von $\sim$, 180
- vollständig
 - $\sim$ e Induktion, 19
 - $\sim$ keit
 - $\sim$ eines metrischen Raumes, 58, 112
 - $\sim$ saxiom, 22, 38
- Ver $\sim$ ung
 - $\sim$ eines metrischen Raumes, 118
- Volumen, 155
- Weierstraß, Karl Theodor Wilhelm 1815–1897
 - Auswahlprinzip von Bolzano– $\sim$, 37
 - Satz von Stone– $\sim$, 121
- Wertebereich, 9
- Whitney, Hassler 1907–1989
 - Satz von $\sim$, 154
- Winkel, 69
- Wohlordnungsprinzip, 20
- Wurzel
 - $\sim$ test, 45
 - k -te $\sim$, 35
 - Quadrat $\sim$, 24
- Young, Grace Chisholm 1886–1944
 - $\sim$ ’sche Ungleichung, 81
- Zahl
 - $\sim$ en $\mathbb{K}$, 31
 - $\sim$ enfolge, 31
 - konvergente $\sim$, 31
 - Bernoulli $\sim$ en, 104
 - erweiterte $\sim$ engerade, 26
 - Eulersche $\sim$ e , 42, 49
 - ganze $\sim$ en $\mathbb{Z}$, 20
 - komplexe $\sim$ en $\mathbb{C}$, 28
 - $\sim$ enebene, 30
 - Polardarstellung der $\sim$, 69
 - natürliche $\sim$ en $\mathbb{N}$, 18
 - Pi π , 66
 - rationale $\sim$ en $\mathbb{Q}$, 20
 - reelle $\sim$ en $\mathbb{R}$, 11, 22
- Zerlegung
 - $\sim$ der Eins, 174
 - Partialbruch $\sim$, 90
 - $\sim$ von $\cot$, 106
 - Produkt $\sim$ von $\sin$, 106
- Zwischen
 - $\sim$ punkte, 94
 - $\sim$ wertsatz, 63