

Analysis I, II

Universität Duisburg-Essen

2021/2022

Frank Müller

4. April 2022

0

Inhaltsverzeichnis

I. Allgemeine Grundlagen	3
1. Logik und Mengenlehre	5
1.1. Mathematische Logik	5
1.2. Mengen und ihre Teilmengen	11
1.3. Mengenoperationen	14
1.4. Kartesisches Produkt und Relationen	17
1.5. Abbildungen zwischen Mengen	20
2. Reelle und komplexe Zahlen	27
2.1. Zahlenbereiche	27
2.2. Die Körperaxiome	30
2.3. Angeordnete Körper	33
2.4. Vollständige Induktion	36
2.5. Die komplexen Zahlen	41
2.6. Rechnen in bewerteten Körpern	45
II. Das Grenzwertkonzept	49
3. Zahlenfolgen	51
3.1. Folgen und Konvergenz	51
3.2. CAUCHYfolgen und Vollständigkeit	56
3.3. Rechenregeln für Folgengrenzwerte	59
3.4. Das Intervallschachtelungsprinzip in \mathbb{R}	61
3.5. Supremum und Infimum von Mengen	66
3.6. Folgen in \mathbb{R} und ihre Häufungswerte	69
3.7. \mathbb{Q} als Teilmenge von \mathbb{R}	73
4. Reihen	79
4.1. Definition und Rechenregeln	79

4.2.	Konvergenzkriterien für Reihen bei absoluter Konvergenz	81
4.3.	Reelle Reihen, CAUCHYSche Produktformel und Umordnungen	86
4.4.	Potenzreihen	91
5.	Topologische Grundlagen im \mathbb{R}^d	95
5.1.	Der Euklidische Vektorraum \mathbb{R}^d	95
5.2.	Folgen im \mathbb{R}^d	99
5.3.	Offene, abgeschlossene und kompakte Mengen	101
5.4.	Zur Klassifizierung von Punkten	105
6.	Stetige Funktionen	109
6.1.	Funktionsgrenzwerte	109
6.2.	Der Stetigkeitsbegriff	112
6.3.	Funktionen einer reellen Veränderlichen	116
6.4.	Stetige Funktionen auf kompakten Mengen	120
6.5.	Der Fundamentalsatz der Algebra	123
6.6.	Funktionenfolgen und gleichmäßige Konvergenz	127
III.	Differential- und Integralrechnung in einer Veränderlichen	131
7.	Die gewöhnliche Ableitung	133
7.1.	Differenzierbarkeit, der Ableitungsbegriff	133
7.2.	Rechenregeln für Ableitungen	135
7.3.	Lokale Extrema und der Mittelwertsatz	138
7.4.	Höhere Ableitungen und lokale Extrema	143
8.	Die elementaren Funktionen	149
8.1.	Funktionenfolgen und Differenzierbarkeit	149
8.2.	Exponentialfunktion, Logarithmus und allgemeine Potenz	151
8.3.	Die Kreisfunktionen cos und sin	156
8.4.	Polarkoordinatendarstellung komplexer Zahlen	160
8.5.	Weitere elementare Funktionen	163
9.	Das eindimensionale RIEMANNsche Integral	169
9.1.	Der Integralbegriff	169
9.2.	Kriterien und Rechenregeln	173
9.3.	Die Additivität des Integrals	179
9.4.	Der Fundamentalsatz der Differential- und Integralrechnung	183

9.5. Drei Anwendungen des Fundamentalsatzes	186
9.6. Uneigentliche Integrale	191
9.7. Die TAYLORSche Formel	196
9.8. TAYLORreihen	200
 IV. Differential- und Integralrechnung in mehreren Veränderlichen	205
 10. Differenzierbare Funktionen mehrerer Veränderlicher	207
10.1. Partielle Ableitungen	207
10.2. Mittelwertsatz und Richtungsableitung	211
10.3. Die totale Ableitung	215
10.4. Partielle Ableitungen höherer Ordnung, der Satz von Schwarz	218
10.5. Die TAYLORSche Formel im \mathbb{R}^n	223
 11. Anwendungen der mehrdimensionalen Differentialrechnung	227
11.1. Lokale Extrema: Bedingung erster Ordnung	227
11.2. Lokale Extrema: Bedingungen zweiter Ordnung	230
11.3. Diffeomorphismen	233
11.4. Der Umkehrsatz	236
11.5. Implizite Funktionen	239
11.6. Mannigfaltigkeiten und lokale Extrema unter Nebenbedingungen	242
 12. Das RIEMANNsche Integral im \mathbb{R}^n	247
12.1. Das Integral auf Quadern	247
12.2. Rechenregeln und iterierte Integration auf Quadern	252
12.3. Der Satz von HEINE-BOREL, Nullmengen und Mengen vom Inhalt Null	256
12.4. Unstetigkeitsstellen integrierbarer Funktionen auf Quadern	260
12.5. Integration über quadrierbare Mengen	262
12.6. Normalbereiche und Additivität des Integrals	267
12.7. Die Transformationsformel für Testfunktionen - Kurzfassung	271
12.8. Die allgemeine Transformationsformel - Kurzfassung	274
 A. Die Transformationsformel	279
A.1. Testfunktionen	279
A.2. Quadrierbarkeit und Diffeomorphismen	282
A.3. Die Transformationsformel für Testfunktionen	285
A.4. Uneigentliche Integrale im \mathbb{R}^n	290

A.5. Die allgemeine Transformationsformel	295
Literaturverzeichnis	301
Index	303

Teil I.

Allgemeine Grundlagen

In Teil I sollen Grundlagen der Mathematik im Allgemeinen gelegt werden, wobei wir die späteren Anwendungen in der Analysis natürlich vor Augen haben. Kapitel 1 ist den grundlegendsten Begriffen der mathematischen Logik und der Mengenlehre gewidmet. In Kapitel 2 diskutieren wir die Struktur und das Wesen der uns aus der Schule bekannten natürlichen, rationalen und reellen Zahlen und erweitern diese anschließend auf die komplexe Ebene.

Kapitel 1.

Logik und Mengenlehre

Die Mathematik basiert auf den Regeln der Logik. Wir wollen diese hier anführen, ohne zu sehr ins Detail zu gehen. Unsere Hauptziele sind zum einen das Nachvollziehen von logischen Regeln, die einige der grundlegenden Beweistechniken der Mathematik begründen. Zum anderen werden wir die Grundlagen der Mengenlehre ableiten. Abschließend betrachten wir noch Relationen auf und Abbildungen zwischen Mengen.

1.1. Mathematische Logik

Mathematische Theorien entstehen, indem aus bekannten Aussagen neue Aussagen gefolgert werden. Die hier auftretenden Termini „Aussage“ und „Folgerung“ sind Begriffe der Logik. Wir beginnen mit der

Definition 1.1.1

Eine *Aussage* ist ein umgangssprachlicher Satz, der entweder wahr oder falsch ist. Jeder Aussage lässt sich also eindeutig genau ein *Wahrheitswert* zuordnen:

w (für *wahr*) oder f (für *falsch*).

Definition 1.1.1 geht auf ARISTOTELES zurück, sie wird auch *Prinzip vom ausgeschlossenen Dritten* genannt.

Beispiel 1.1.2. (a) Beispiele von Aussagen im Sinne von Definition 1.1.1 sind etwa

- ▶ „4 ist eine gerade Zahl.“
- ▶ „5 ist eine gerade Zahl.“
- ▶ „Jede gerade Zahl größer 2 ist Summe zweier Primzahlen.“

Die erste Aussage ist sicher wahr, die zweite sicher falsch. Von der dritten Aussage, der GOLDBACH-Vermutung, weiß man zwar nicht, ob sie wahr oder falsch ist, sie ist aber mit Sicherheit genau eines von beiden.

(b) Beispiele von umgangssprachlichen Sätzen, die *keine Aussagen* sind:

- ▶ „Seid endlich ruhig!“
- ▶ „Ist die Vorlesung bald zu Ende?“
- ▶ „Jede meiner Aussagen ist falsch.“

Aufforderungen und Fragen sind also keine Aussagen. Beim dritten Satz findet eine (ungünstige) Vermischung mit der sogenannten *Metaebene* statt (Aussagen über Aussagen).

Wir bezeichnen Aussagen mit a, b, c, \dots . Zwei Aussagen a, b lassen sich zu neuen Aussagen verknüpfen, entsprechend unserer Alltagssprache. Die Wahrheitswerte lassen sich am besten in Form von *Wahrheitstafeln* darstellen:

Definition 1.1.3: Logische Verknüpfungen

Für zwei Aussagen a und b erklären wir:

- ▶ die *UND-Verknüpfung* $a \wedge b$ (lies: „ a und b “) und die *ODER-Verknüpfung* $a \vee b$ (lies: „ a oder b “) mit den Wahrheitswerten

a	b	$a \wedge b$	$a \vee b$
w	w	w	w
w	f	f	w
f	w	f	w
f	f	f	f

- ▶ die *Folgerung* oder *Implikation* $a \rightarrow b$ (lies: „wenn a , dann b “ oder „aus a folgt b “) und die *Äquivalenz* $a \leftrightarrow b$ (lies: „genau dann a , wenn b “) mit

a	b	$a \rightarrow b$	$a \leftrightarrow b$
w	w	w	w
w	f	f	f
f	w	w	f
f	f	w	w

auch die Schreibweise $b \leftarrow a$ für $a \rightarrow b$ ist zulässig,

- ▶ die *Negation* $\neg a$ (lies: „nicht a “) mit

a	$\neg a$
w	f
f	w

- Beispiel 1.1.4.** (a) „4 ist eine gerade Zahl oder 5 ist eine gerade Zahl“ ist also eine wahre Aussage. Umgangssprachlich sagt man natürlich besser „4 oder 5 ist eine gerade Zahl“. *Übrigens:* Bei der ODER-Verknüpfung ist kein „Entweder-Oder“ gemeint; dazu erklärt man das *ausschließende ODER* oder *Kontravalenz* (\rightarrow Übungen).
- (b) „Es gilt nicht, dass jede gerade Zahl größer 2 Summe zweier Primzahlen ist“ ist genau dann wahr, wenn die GOLDBACH-Vermutung falsch ist. Eine schönere Formulierung ist hier: „Es gibt eine gerade Zahl größer 2, die nicht Summe zweier Primzahlen ist“. Wir werden das unten mit Hilfe von Quantoren präzisieren.
- (c) Die Definition der Folgerung $a \rightarrow b$ ist etwas gewöhnungsbedürftig, aber durchaus praktisch und sinnvoll. Wir halten fest:
- ▶ Ist die *Prämisse* a wahr, so stimmt der Wahrheitswert von $a \rightarrow b$ mit dem von b überein.
 - ▶ Ist a falsch, so ist der Wahrheitswert von $a \rightarrow b$ immer wahr („aus Falschem kann alles folgen“).

Zur Erklärung ein Beispiel: „Wenn x gerade ist, so ist $x + 1$ ungerade“ ist immer wahr.

Denn: Ist „ x ist gerade“ falsch, so ist die Folgerung nach Definition wahr; in diesem Fall ist nichts zu zeigen, wir wollen dann ja auch gar keine Aussage treffen. Ist „ x ist gerade“ wahr, so ist $x + 1$ sicher ungerade, also die Folgerung wahr. \square

Mit den Verknüpfungen aus Definition 1.1.3 lassen sich nun auch kompliziertere Aussagen zusammensetzen, etwa

$$((a \wedge b) \rightarrow (\neg c)) \vee ((b \wedge c) \leftrightarrow ((\neg b) \wedge a)). \quad (1.1.1)$$

Um solche Zusammensetzungen übersichtlich zu halten, legen wir fest:

Vereinbarung. Die Verknüpfungen aus Definition 1.1.3 binden in folgender Reihenfolge:

$$\neg \text{ vor } \wedge, \vee \text{ vor } \rightarrow, \leftrightarrow.$$

Aussage (1.1.1) lässt sich also schreiben als

$$(a \wedge b \rightarrow \neg c) \vee (b \wedge c \leftrightarrow \neg b \wedge a).$$

Definition 1.1.5

Eine Aussage heißt *Tautologie*, wenn sie immer wahr ist, und *Kontradiktion*, wenn sie immer falsch ist.

Satz 1.1.6

Zu beliebigen Aussagen a, b, c , sind folgende Aussagen Tautologien:

- (a) *Regel vom ausgeschlossenen Dritten*: $a \vee \neg a$
- (b) *Widerspruchsregel*: $\neg(a \wedge \neg a)$
- (c) *Doppelte Verneinung*: $\neg\neg a \leftrightarrow a$
- (d) *Kontraposition*: $(a \rightarrow b) \leftrightarrow (\neg b \rightarrow \neg a)$
- (e) *Modus ponens*: $(a \rightarrow b) \wedge a \rightarrow b$
- (f) *Modus tollens*: $(a \rightarrow b) \wedge \neg b \rightarrow \neg a$
- (g) *Modus barbara*: $(a \rightarrow b) \wedge (b \rightarrow c) \rightarrow (a \rightarrow c)$

Beweis. Zum Beweis kann man Wahrheitstabeln verwenden, wir zeigen das hier am Beispiel des Modus barbara; die anderen Aussagen belassen wir zur Übung (siehe auch *Ergänzungen*). Es gilt

a	b	c	$a \rightarrow b$	$b \rightarrow c$	$(a \rightarrow b) \wedge (b \rightarrow c)$	$a \rightarrow c$	$(a \rightarrow b) \wedge (b \rightarrow c) \rightarrow (a \rightarrow c)$
w	w	w	w	w	w	w	w
w	w	f	w	f	f	f	w
w	f	w	f	w	f	w	w
w	f	f	f	w	f	f	w
f	w	w	w	w	w	w	w
f	w	f	w	f	f	w	w
f	f	w	w	w	w	w	w
f	f	f	w	w	w	w	w

wie behauptet. □

Vereinbarung. Ist eine Folgerung $a \rightarrow b$ eine Tautologie, also immer wahr, so schreiben wir

$$a \Rightarrow b \quad (\text{lies: „}a \text{ impliziert } b\text{“})$$

und nennen dies einen (*allgemein gültigen*) *logischen Schluss*. Ist eine Äquivalenz $a \leftrightarrow b$ eine Tautologie, also immer wahr, so schreiben wir

$$a \Leftrightarrow b \quad (\text{lies: „}a \text{ äquivalent } b\text{“})$$

und nennen dies eine (*allgemein gültige*) *äquivalente Umformung*. Die Zeichen \Rightarrow und \Leftrightarrow binden *schwächer* als alle anderen logischen Verknüpfungen.

Bemerkung 1.1.7. In der Mathematik arbeitet man mit genau solchen logischen Schlüssen $a \Rightarrow b$, um von einer wahren Aussage a auf eine neue wahre Aussage b zu schließen. Und man arbeitet mit äquivalenten Umformungen $a \Leftrightarrow b$, um eine bekannte Aussage a in eine Aussage b „umzuschreiben“ und so z.B. neue Eigenschaften des betrachteten mathematischen Objekts ablesen zu können.

Aus Satz 1.1.6 können wir einige der wichtigsten logischen Schlüsse und äquivalenten Umformungen ablesen, die wir für unsere Beweise benötigen werden.

Folgerung 1.1.8: Beweistechniken

Für beliebige Aussagen a, b, c gelten

- ▶ **Direkter Beweis, Abtrennungsregel:** $(a \rightarrow b) \wedge a \Rightarrow b$.
(Wenn a wahr ist und b aus a folgt, so ist b wahr.)
- ▶ **Kontrapositionsregel:** $a \rightarrow b \Leftrightarrow \neg b \rightarrow \neg a$.
(b folgt genau dann aus a , wenn $\neg a$ aus $\neg b$ folgt.)
- ▶ **Widerspruchsbeweis:** $\neg a \rightarrow b \wedge \neg b \Rightarrow a$.
(Wenn aus „ a ist falsch“ eine Kontradiktion folgt, muss a wahr sein.)
- ▶ **Indirekter Beweis:** $(\neg a \rightarrow b) \wedge \neg b \Rightarrow a$.
(Wenn man „ b ist falsch“ weiß und b aus $\neg a$ folgt, so muss a wahr sein.)
- ▶ **Kettenschluss:** $(a \rightarrow b) \wedge (b \rightarrow c) \Rightarrow (a \rightarrow c)$.
(Wenn b aus a und c aus b folgen, so folgt auch c aus a .)

Beweis. Abgesehen vom Widerspruchsbeweis entsprechen die Regeln den Tautologien aus Satz 1.1.6 (e), (d), (f) und (g). Wir zeigen den Widerspruchsbeweis (alternativ mit Wahrheitstafel):

Nach Satz 1.1.6 ist $\neg(b \wedge \neg b)$ immer wahr, also $b \wedge \neg b$ eine Kontradiktion, d.h. immer falsch. Damit also $\neg a \rightarrow b \wedge \neg b$ wahr ist (im anderen Fall ist nichts zu zeigen), muss auch $\neg a$ falsch bzw. a wahr sein. Das war die Behauptung. \square

Bemerkung 1.1.9. Weitere wichtige Schlussregeln und äquivalente Umformungen sind z.B. (→ Übungen + Ergänzungen):

- ▶ **Äquivalenzregel:** $a \Leftrightarrow b \Leftrightarrow (a \rightarrow b) \wedge (b \rightarrow a)$.
- ▶ **Kreisschluss:** $(a \Leftrightarrow b) \wedge (b \Leftrightarrow c) \Leftrightarrow (a \rightarrow b) \wedge (b \rightarrow c) \wedge (c \rightarrow a)$.
- ▶ **Fallunterscheidung:** $(a \rightarrow b) \wedge (\neg a \rightarrow b) \Rightarrow b$

Beim Kreisschluss wird auf der rechten Seite die *Assoziativität* der UND-Verknüpfung verwendet: $(a \wedge b) \wedge c \Leftrightarrow a \wedge (b \wedge c)$, d.h. die Klammerung kann weggelassen werden; dasselbe gilt für \vee und \Leftrightarrow , nicht aber für \rightarrow (*Nachrechnen*).

Definition 1.1.10: Aussageformen und Quantoren

- ▶ Eine *Aussageform* ist ein (mathematischer) Ausdruck p , der durch Einsetzen einer Variablen x zu einer Aussage $p(x)$ wird.
- ▶ Ist p eine Aussageform, so schreiben wir mit dem *Existenzquantor* \exists :

$$\exists x p(x) \quad \text{oder} \quad \exists x : p(x)$$

für die Aussage „Es existiert ein x , so dass $p(x)$ wahr ist“.

- ▶ Ist p eine Aussageform, so schreiben wir mit dem *Allquantor* \forall :

$$\forall x p(x) \quad \text{oder} \quad \forall x : p(x)$$

für die Aussage „Für alle x ist $p(x)$ wahr“.

Beispiel 1.1.11. (a) Aussageformen sind z.B.

- ▶ $x^2 + 7 = 5$
- ▶ $2x3$, wobei x eine Relation darstellt (z.B. $>$ oder $=$, siehe Abschnitt A.1.4)

Zur Präzisierung der Definition einer Aussageform bedarf es der Festlegung, aus welchem Bereich x zu wählen ist; wir führen dazu in Abschnitt A.1.2 den Begriff der *Menge* ein.

(b) Was ist der Wahrheitswert der Aussage

$$\exists x : x^2 + 7 = 5?$$

Auch hier muss die Menge für x präzisiert werden: Lassen wir für x nur *reelle* Zahlen zu, ist die Aussage falsch; erlauben wir für x *komplexe* Zahlen, so ist die Aussage aber wahr.

Bemerkung 1.1.12. Wichtig sind noch die Negationen der in Definition 1.1.10 erklärten Aussagen; z.B. „Für alle x ist $p(x)$ wahr“ negieren wir umgangssprachlich mit „Es gibt ein x , so dass $p(x)$ falsch ist“. Wir setzen daher

$$\neg(\forall x : p(x)) :\Leftrightarrow \exists x : \neg p(x) \tag{1.1.2}$$

und analog

$$\neg(\exists x : p(x)) :\Leftrightarrow \forall x : \neg p(x). \tag{1.1.3}$$

Dabei bedeutet der Doppelpunkt *links* von \Leftrightarrow , dass die *linke* Seite der Äquivalenz durch die rechte Seite definiert ist.

Beispiel 1.1.13. Die Formalisierung der Aussage „Zu jeder Zahl x lässt sich eine Zahl y so finden, dass $xy = 2$ gilt“ ist

$$\forall x \exists y : xy = 2.$$

Als Negation erhalten wir

$$\neg(\forall x \exists y : xy = 2) \Leftrightarrow \exists x : \neg(\exists y : xy = 2) \Leftrightarrow \exists x \forall y : xy \neq 2,$$

also „Es gibt ein x , so dass für alle y gilt $xy \neq 2$ “.

1.2. Mengen und ihre Teilmengen

Definition 1.2.1: Mengendefinition nach CANTOR

Eine *Menge* M ist eine Zusammenfassung bestimmter wohlunterschiedener Objekte unserer Anschauung oder unseres Denkens, welche die *Elemente von M* genannt werden. Für ein beliebiges Objekt m ist also eindeutig festgelegt, ob m Element von M ist - wir schreiben dann $m \in M$ - oder nicht Element von M ist - dann schreiben wir $m \notin M$ für $\neg(m \in M)$.

Diese Definition ist nicht exakt, es lassen sich recht leicht Paradoxien (also logische Widersprüche) finden; man spricht daher auch vom *naiven Mengenbegriff*. Es gibt eine axiomatische Mengendefinition von ZERMELO, FRAENKEL u.a., die aber deutlich komplizierter ist. Alle Mengen, die wir benötigen werden, sind auch Mengen im axiomatischen Sinn, daher begnügen wir uns (wie allgemein üblich) hier mit dem naiven Mengenbegriff.

Definition 1.2.2

- ▶ Zwei Mengen M_1, M_2 heißen *gleich*, i.Z. $M_1 = M_2$, wenn sie dieselben Elemente enthalten.
- ▶ Die Menge ohne Elemente heißt *leere Menge*, wir schreiben \emptyset . Für jedes Objekt m gilt also $m \notin \emptyset$.

Bemerkung 1.2.3. Sind zwei Mengen M_1, M_2 gleich, so handelt es sich im Sinne von Definition 1.2.1 um dieselbe Menge (mit zwei verschiedenen Bezeichnungen). In diesem Sinne gilt: *Es gibt nur eine leere Menge.*

Denn: Sind \emptyset_1 und \emptyset_2 leere Mengen, so ist jedes Element von \emptyset_1 auch in \emptyset_2 und jedes Element von \emptyset_2 auch in \emptyset_1 enthalten. Es gilt also $\emptyset_1 = \emptyset_2$. □

Definition 1.2.4: Darstellungen von Mengen

Grundsätzlich können Mengen auf zwei Arten geschrieben werden:

- Durch Aufzählung ihrer Elemente m_1, m_2, m_3, \dots , also

$$M = \{m_1, m_2, m_3, \dots\},$$

wobei Mehrfachnennungen möglich sind (siehe Beispiel 1.2.5 unten).

- Durch Angabe der Eigenschaften ihrer Elemente

$$M = \{m : p(m)\},$$

wobei p eine Aussageform ist.

Beispiel 1.2.5. (a) Beispiele für Aufzählungen sind

$$M_1 = \{1, 2, 5, 3, 4\}, \quad M_2 = \{R, O, M, A\}, \quad M_3 = \{1, 1, 1\}.$$

Alle Mengen haben endlich viele Elemente. Auf die Reihenfolge der Elemente kommt es nicht an, andere Darstellungen von M_1, M_2 sind z.B.

$$M_1 = \{1, 2, 3, 4, 5\}, \quad M_2 = \{A, M, O, R\}.$$

Bei M_3 gibt es eine Mehrfachnennung, es gilt

$$M_3 = \{1, 1, 1\} = \{1, 1\} = \{1\}.$$

Den Sinn dieser Regelung werden wir später erkennen.

- (b) Hat eine Menge unendlich viele Elemente, so nutzen wir die „Punktchen-Schreibweise“, um ein „und so weiter“ zu kennzeichnen; z.B. ist

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

die *Menge der natürlichen Zahlen*. Übrigens lässt sich nicht jede Menge mit unendlich vielen Elementen als Aufzählung schreiben, siehe Kapitel 3.

- (c) Z.B. lässt sich M_1 auch durch die Eigenschaften ihrer Elemente beschreiben:

$$M_1 = \{m : m \in \mathbb{N} \wedge m < 6\} =: \{m \in \mathbb{N} : m < 6\}.$$

Dabei bedeutet der Doppelpunkt *rechts* von $=$, dass die *rechte* Seite der Gleichung durch die linke Seite definiert ist. Die kürzere Schreibweise rechts ist üblich, wenn eine *Grundmenge* (hier \mathbb{N}) bekannt ist.

Bemerkung 1.2.6. Ist M eine Menge mit endlich vielen Elementen m_1, \dots, m_n , so bezeichnen wir mit

$$\#M = |M| := n$$

die *Mächtigkeit* von M , also die Anzahl ihrer Elemente. Für die Mengen aus Beispiel 1.2.5 (a) gilt also $|M_1| = 5$, $|M_2| = 4$, $|M_3| = 1$.

Definition 1.2.7

Eine Menge A heißt *Teilmenge* einer Menge M , wenn jedes Element von A auch in M enthalten ist. Wir schreiben dann

$$A \subset M \quad (\text{lies auch: „} A \text{ enthalten in } M \text{“}).$$

Weiter heißt A *echte Teilmenge* von M , wenn gilt

$$A \subset M \wedge A \neq M.$$

Wir schreiben dann

$$A \subsetneq M \quad (\text{lies auch: „} A \text{ echt enthalten in } M \text{“}).$$

Beispiel 1.2.8. ▶ $\{O, M, A\} \subset \{A, M, O, R\}$

- ▶ $\{O, P, A\} \not\subset \{A, M, O, R\}$, lies: „ $\{O, P, A\}$ ist nicht enthalten in $\{A, M, O, R\}$ “.
- ▶ $M_1 \subsetneq \mathbb{N}$ mit M_1 aus Beispiel 1.2.5.

Definition 1.2.9: Potenzmenge

Die *Potenzmenge* einer Menge M ist die Menge ihrer Teilmengen; wir schreiben

$$\mathcal{P}(M) := \{A : A \subset M\}.$$

Beispiel 1.2.10. (a) $\mathcal{P}(M)$ enthält also immer \emptyset und M als Elemente (*nicht* als Teilmengen). Einelementige Mengen $M = \{m\}$ sind die einzigen, deren Potenzmengen auch keine weiteren Elemente enthalten: $\mathcal{P}(\{m\}) = \{\emptyset, \{m\}\}$. Die leere Menge ist die einzige, deren Potenzmenge nur ein Element enthält: $\mathcal{P}(\emptyset) = \{\emptyset\}$.

(b) Man kann natürlich auch die Potenzmenge von der Potenzmenge von der Potenzmen-

ge. . . bilden. Ist z.B. $M = \emptyset$, so gilt

$$\begin{aligned}\mathcal{P}(\emptyset) &= \{\emptyset\}, \\ \mathcal{P}(\mathcal{P}(\emptyset)) &= \mathcal{P}(\{\emptyset\}) = \{\emptyset, \{\emptyset\}\}, \\ \mathcal{P}(\mathcal{P}(\mathcal{P}(\emptyset))) &= \mathcal{P}(\{\emptyset, \{\emptyset\}\}) = \{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \{\emptyset, \{\emptyset\}\}\}.\end{aligned}$$

Zum besseren Verständnis stelle man sich \emptyset als leere Kiste vor. Dann ist z.B. $\{\{\emptyset\}\}$ eine Kiste, die eine Kiste enthält, die wiederum eine leere Kiste enthält.

1.3. Mengenoperationen

Aus bekannten Mengen kann man neue Mengen erklären m.H. der folgenden Setzungen:

Definition 1.3.1: Mengenoperationen

Zu zwei Mengen M, N erklären wir deren

- ▶ **Vereinigung** $M \cup N := \{m : m \in M \vee m \in N\}$
- ▶ **Durchschnitt** $M \cap N := \{m : m \in M \wedge m \in N\}$
- ▶ **Differenz** $M \setminus N := \{m : m \in M \wedge m \notin N\}$

$M \setminus N$ heißt auch **relatives Komplement** von N in M . Gilt zusätzlich $N \subset M$, so heißt

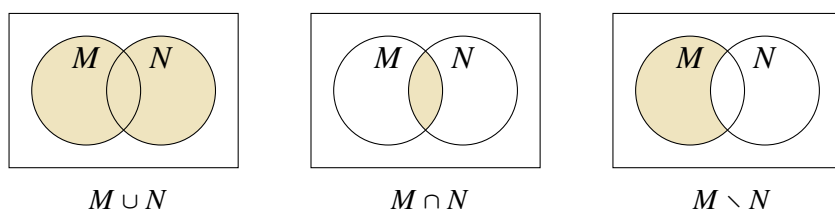
$$C_M N = M \setminus N$$

das **Komplement** von N bezüglich M ; ist M aus dem Kontext bekannt, schreibt man meist

$$N^c = C_M N.$$

Schließlich heißen zwei Mengen M, N **disjunkt**, wenn $M \cap N = \emptyset$ gilt.

Bemerkung 1.3.2. Zur Veranschaulichung der Mengenoperationen kann man sogenannte **VENN-Diagramme** nutzen. Sie dienen als graphische Hilfsmittel zur Plausibilitätsprüfung, stellen aber



keine Beweise dar, da die tatsächliche Art und Lage der Mengen zueinander nicht berücksichtigt wird (siehe Bemerkung 1.3.5 unten).

Beispiel 1.3.3. ▶ $\{O, M, A\} \cap \{O, P, A\} = \{O, A\}$

▶ $\{O, P, A\} \cup \{A, M, O, R\} = \{R, A, M, P, O\}$

▶ $\mathbb{N} \setminus M_1 = \{m \in \mathbb{N} : m \geq 6\} = C_{\mathbb{N}}M_1 = M_1^c$ mit M_1 aus Beispiel 1.2.5 (a).

▶ $\mathbb{N} \cap \{O, M, A\} = \emptyset$

Satz 1.3.4: Rechenregeln für Mengenoperationen

(a) Sind M, N, O Mengen, so gelten die *Distributivgesetze*

$$(M \cap N) \cup O = (M \cup O) \cap (N \cup O),$$

$$(M \cup N) \cap O = (M \cap O) \cup (N \cap O).$$

(b) Für zwei Mengen M, N gelten die *Verschmelzungsgesetze*

$$(M \cap N) \cup M = M = (M \cup N) \cap M.$$

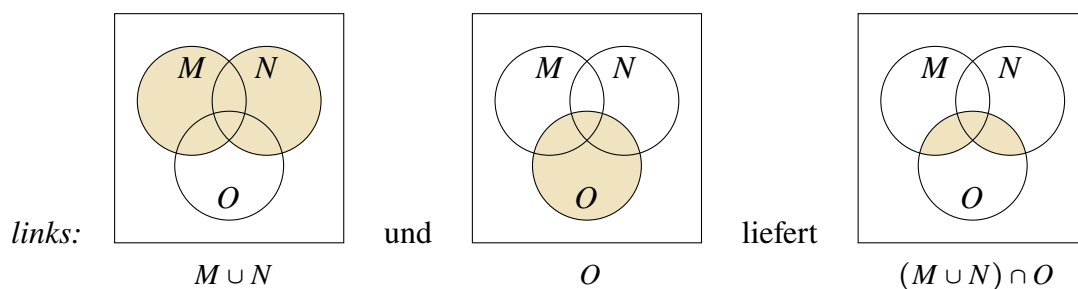
(c) Ist $A \subset M$ eine Teilmenge von M , so gilt für das Komplement von A bez. M das *Doppelkomplementgesetz*

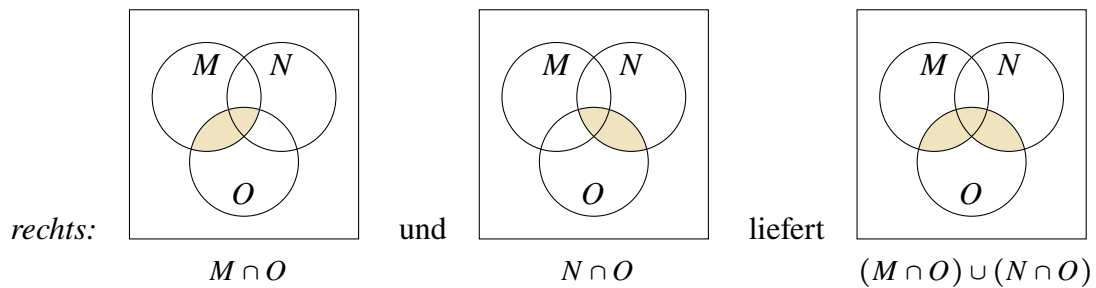
$$(A^c)^c = A.$$

(d) Sind $A, B \subset M$ zwei Teilmengen von M , so gelten für deren Komplemente bez. M die *Regeln von de Morgan*

$$(A \cap B)^c = A^c \cup B^c, \quad (A \cup B)^c = A^c \cap B^c.$$

Bemerkung 1.3.5. Wir zeigen am Beispiel des 2. Distributivgesetzes wie Venn-Diagramme zur Illustration angewendet werden können:





Beweis von Satz 1.3.4. Wir diskutieren hier (a) und (d), Aussagen (b) und (c) werden in den Ergänzungen bewiesen.

- (a) Wir zeigen $(M \cup N) \cap O = (M \cap O) \cup (N \cap O)$, die erste Relation folgt analog (\rightarrow Nachrechnen). Zu beliebigem m und beliebiger Menge X ist die Aussage $m \in X$ entweder wahr oder falsch. Die zugehörige Wahrheitstafel wird *Mengentafel* genannt:

M	N	O	$M \cup N$	$(M \cup N) \cap O$	$M \cap O$	$N \cap O$	$(M \cap O) \cup (N \cap O)$
w	w	w	w	w	w	w	w
w	w	f	w	f	f	f	f
w	f	w	w	w	w	f	w
w	f	f	w	f	f	f	f
f	w	w	w	w	f	w	w
f	w	f	w	f	f	f	f
f	f	w	f	f	f	f	f
f	f	f	f	f	f	f	f

Da die blau markierten Spalten übereinstimmen, ist $m \in (M \cap N) \cup O$ genau dann wahr, wenn $m \in (M \cup O) \cap (N \cup O)$ wahr ist, wie behauptet.

- (d) $(A \cap B)^c = A^c \cup B^c$. Wir nutzen die logischen Äquivalenzen

$$\neg(a \wedge b) \Leftrightarrow \neg a \vee \neg b, \quad (a \vee b) \wedge c \Leftrightarrow (a \wedge c) \vee (b \wedge c);$$

die erste Relation zeigen wir in den Übungen, die Wahrheitstafel der zweiten ist analog zur Mengentafel in (a). Damit folgt

$$\begin{aligned}
 m \in (A \cap B)^c &\stackrel{\text{Def. 1.3.1}}{\Leftrightarrow} m \in M \wedge m \notin A \cap B \\
 &\stackrel{\text{Def. 1.3.1}}{\Leftrightarrow} m \in M \wedge \neg(m \in A \wedge m \in B) \\
 &\Leftrightarrow m \in M \wedge (m \notin A \vee m \notin B) \\
 &\Leftrightarrow (m \in M \wedge m \notin A) \vee (m \in M \wedge m \notin B) \\
 &\stackrel{\text{Def. 1.3.1}}{\Leftrightarrow} m \in A^c \vee m \in B^c \\
 &\stackrel{\text{Def. 1.3.1}}{\Leftrightarrow} m \in A^c \cup B^c.
 \end{aligned}$$

$(A \cup B)^c = A^c \cap B^c$. Unter Verwendung von (c) und der ersten de Morganschen Regel folgt sofort

$$\begin{aligned} m \in (A \cup B)^c &\stackrel{(c)}{\Leftrightarrow} m \in ((A^c)^c \cup (B^c)^c)^c \\ &\Leftrightarrow m \in ((A^c \cap B^c)^c)^c \\ &\stackrel{(c)}{\Leftrightarrow} m \in A^c \cap B^c, \end{aligned}$$

wie behauptet. □

Bemerkung 1.3.6. Sei J eine beliebige nichtleere Menge, und zu jedem $j \in J$ sei M_j wiederum eine Menge; J wird dann *Indexmenge* genannt und wir schreiben $\{M_j\}_{j \in J}$ für die *Familie* der Mengen M_j . Analog zu Definition 1.3.1 erklären wir dann die *Vereinigung* bzw. den *Durchschnitt* der M_j als

$$\begin{aligned} \bigcup_{j \in J} M_j &:= \{m : \exists j (m \in M_j)\}, \\ \bigcap_{j \in J} M_j &:= \{m : \forall j (m \in M_j)\}. \end{aligned}$$

Gilt $M_j \subset X$ für alle $j \in J$ mit einer Menge X , so folgen für deren Komplemente in X (\rightarrow Übung):

$$\left(\bigcup_{j \in J} M_j\right)^c = \bigcap_{j \in J} M_j^c, \quad \left(\bigcap_{j \in J} M_j\right)^c = \bigcup_{j \in J} M_j^c.$$

Aufgrund ihrer Analogie zu Satz 1.3.4 (d) sprechen wir auch hier von den *Regeln von de Morgan*.

1.4. Kartesisches Produkt und Relationen

Relationen auf einer Menge M werden wir als Teilmengen des kartesischen Produkts von M mit sich selbst definieren. Dazu benötigen wir zunächst letzteren Begriff, der auch von unabhängigem Interesse ist:

Definition 1.4.1: Kartesisches Produkt

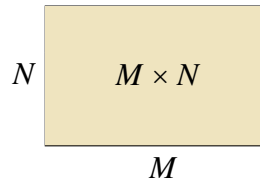
Zu beliebigen Mengen M, N erklären wir das *kartesische Produkt*

$$M \times N := \{(m, n) : m \in M, n \in N\}.$$

Dabei heißt (m, n) *geordnetes Paar* mit den Komponenten $m \in M, n \in N$.

Bemerkung 1.4.2. (a) Der Name „geordnetes Paar“ rührt daher, dass (m, n) und (n, m) i.A. verschieden sind; ein willkürliches Vertauschen der Reihenfolge ist also nicht zulässig. Insbesondere gilt daher i.A. auch $M \times N \neq N \times M$.

Man kann sich $M \times N$ schematisch als Rechteck mit den Seiten M und N veranschaulichen.



(b) Analog erklärt man das kartesische Produkt von 2, 3, ... oder allgemein $d \in \mathbb{N}$ Mengen M_1, \dots, M_d , also

$$M_1 \times \dots \times M_d := \{(m_1, \dots, m_d) : m_j \in M_j, j = 1, \dots, d\}.$$

Die Elemente (m_1, \dots, m_d) heißen dann *d-Tupel* mit den Komponenten m_1, \dots, m_d ; geordnete Paare sind also 2-Tupel.

Beispiel 1.4.3. Der dreidimensionale Raum unserer Anschauung lässt sich durch die Menge aller 3-Tupel (oder *geordneten Tripel*)

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

beschreiben, wobei \mathbb{R} die *reellen Zahlen* bezeichnet (siehe Kapitel 2); x, y, z sind dann die Koordinaten der Raumpunkte.

In Verallgemeinerung werden wir später auch den *d-dimensionalen EUKLIDischen Raum* betrachten:

$$\mathbb{R}^d := \underbrace{\mathbb{R} \times \dots \times \mathbb{R}}_{d\text{-mal}}, \quad d \in \mathbb{N}.$$

Beziehungen zwischen Elementen von Mengen lassen sich allgemein als Relationen definieren:

Definition 1.4.4: Relationen

Ist M eine beliebige Menge, so nennen wir eine Teilmenge $R \subset M \times M$ (*zweistellige Relation*) auf M . Zwei Elemente $x, y \in M$ stehen in (dieser) Relation zueinander, wenn $(x, y) \in R$ gilt; wir schreiben dann kürzer xRy .

Wir betrachten hier näher *Äquivalenzrelationen*. Eine andere interessante Klasse sind *Ordnungsrelationen* (wie „ \leq “ oder „ \subset “), die wir aber nur exemplarisch und nach Bedarf betrachten werden.

Definition 1.4.5: Äquivalenzrelationen

Eine Relation R auf M heißt *Äquivalenzrelation*, wenn folgendes für alle $x, y, z \in M$ gilt:

(R) *Reflexivität*: xRx

(S) *Symmetrie*: $xRy \Rightarrow yRx$

(T) *Transitivität*: $xRy \wedge yRz \Rightarrow xRz$

Eine Äquivalenzrelation wird meist allgemein mit \sim bezeichnet, und zwei Elemente $x, y \in M$ mit $x \sim y$ heißen *äquivalent*. Schließlich nennt man die Teilmenge

$$[x] := \{y \in M : y \sim x\}$$

Äquivalenzklasse zu $x \in M$.

Satz 1.4.6

Ist \sim eine Äquivalenzrelation auf der Menge M , so bildet die sogenannte *Restklassenmenge modulo \sim* ,

$$M/\sim := \{[x] : x \in M\},$$

eine *Zerlegung* von M : Die Vereinigung aller Äquivalenzklassen ist gleich M , und je zwei Äquivalenzklassen sind gleich oder disjunkt.

Beweis. ▶ Nach Definition 1.4.5 (R) gilt $x \in [x]$ für alle $x \in M$ und somit $M \subset \bigcup_{x \in M} [x]$. Andererseits gilt natürlich $[x] \subset M$ für jedes $x \in M$ und folglich $\bigcup_{x \in M} [x] \subset M$. Somit sind beide Mengen gleich.

▶ Gilt $[x] \cap [y] \neq \emptyset$, so existiert ein $z \in [x] \cap [y]$. Ist nun $a \in [x]$ beliebig, d.h. $a \sim x$, so folgt wegen $z \sim x$ aus Definition 1.4.5 (S) und (T) zunächst $a \sim z$. Da andererseits $z \sim y$ gilt, folgt wiederum aus Definition 1.4.5 (T) auch $a \sim y$. Folglich haben wir $[x] \subset [y]$, und analog zeigt man $[y] \subset [x]$; insgesamt folgt $[x] = [y]$, wie behauptet. \square

Beispiel 1.4.7. (a) Die Gleichheitsrelation „ $=$ “ ist auf jeder Menge M eine Äquivalenzrelation mit den Äquivalenzklassen $[x] = \{x\}$, $x \in M$.

(b) Zu $p \in \mathbb{N} \setminus \{1\}$ definiert

$$n \sim m :\Leftrightarrow p \text{ teilt } n - m$$

eine Äquivalenzrelation auf den ganzen Zahlen \mathbb{Z} , die mit $n \equiv m \pmod{p}$ bezeichnet wird. Ihre Äquivalenzklassen sind die p Restklassen bezüglich Division durch p .

(c) Auf der Menge G aller Geraden in der anschaulichen Ebene $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ ist durch

$$g_1 \parallel g_2 :\Leftrightarrow g_1 \text{ ist parallel zu } g_2$$

eine Äquivalenzrelation definiert. Die Äquivalenzklasse $[g]$ von $g \in G$ enthält alle zu g parallelen Geraden.

(d) Die \subset -Relation ist keine Äquivalenzrelation, sie ist nicht symmetrisch.

1.5. Abbildungen zwischen Mengen

Abschließend besprechen wir noch den zentralen Begriff der Abbildung oder Funktion:

Definition 1.5.1: Abbildungen

Sind M, N zwei Mengen, so ordnet eine *Abbildung* oder *Funktion* f von M nach N jedem Element $x \in M$ genau ein Element $y \in N$ zu. Wir schreiben dann

$$f : M \rightarrow N, \quad x \mapsto y = f(x).$$

Bemerkung 1.5.2. Der etwas unklare Begriff „zuordnen“ lässt sich mit einer mengentheoretischen, etwas unanschaulichen Definition umgehen: $f : M \rightarrow N$ ist dabei definiert als Teilmenge von $M \times N$, in der jedes $x \in M$ in genau einem Paar aus $M \times N$ an erster Stelle vorkommt.

Beispiel 1.5.3. (a) *Lineare Funktion:* $f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto f(x) = mx + b$ mit Steigung $m \in \mathbb{R}$ und Ordinatenabschnitt $b \in \mathbb{R}$.

(b) $\pi : \mathbb{N} \rightarrow \mathbb{N}, n \mapsto \pi(n) := \#\{p \in \mathbb{N} : p \text{ ist Primzahl} \wedge p \leq n\}$. Eine einfache Darstellung für die Zuordnungsvorschrift π gibt es hier nicht.

(c) *Identische Abbildung:* $\text{id}_M : M \rightarrow M, x \mapsto x$ mit beliebiger Menge M . Wenn diese Menge M klar ist, schreiben wir auch $\text{id} = \text{id}_M$.

(d) *Charakteristische Funktion:* Ist $A \subset M$ eine Teilmenge einer Menge M , so setzt man

$$\chi_A : M \rightarrow \{0, 1\}, \quad x \mapsto \chi_A(x) = \begin{cases} 1, & x \in A \\ 0, & x \in M \setminus A \end{cases}.$$

(e) *Einschränkung:* Sind M, N Mengen, $A \subset M$ eine Teilmenge und $f : M \rightarrow N$ eine Funktion, so setzt man

$$f|_A : A \rightarrow N, \quad x \mapsto f(x).$$

Die Funktion f wird durch $f|_A$ also auf die Punkte aus A *eingeschränkt*.

- (f) *d-Tupel* (vgl. Beispiel 1.4.2 (b)): Zu $d \in \mathbb{N}$ setzt man $A_d := \{1, \dots, d\}$ und $f : A_d \rightarrow M$, $k \mapsto y_k := f(k)$. Dann wird f durch das d -Tupel (y_1, \dots, y_d) mit den Koordinaten $y_k = f(k)$ eindeutig beschrieben.
- (g) *Folgen*: Eine Abbildung $f : \mathbb{N} \rightarrow M$, $k \mapsto y_k = f(k)$ wird als Folge in M bezeichnet; wir schreiben dann

$$\{y_k\}_{k=1,2,\dots} = \{y_k\}_k.$$

Folgen werden wir in Kapitel 3 genau untersuchen.

Die folgende Klasse von Abbildungen treffen wir überall in der Mathematik an:

Definition 1.5.4: Operationen

Ist M eine Menge, so wird eine Abbildung

$$\star : M \times M \rightarrow M, \quad (x, y) \mapsto \star(x, y) =: x \star y$$

als (*binäre*) *Operation* auf M bezeichnet. Eine Operation heißt *kommutativ*, wenn gilt

$$x \star y = y \star x \quad \text{für alle } x, y \in M$$

und *assoziativ*, wenn gilt

$$(x \star y) \star z = x \star (y \star z) \quad \text{für alle } x, y, z \in M.$$

Beispiel 1.5.5. Für eine Menge M ist z.B. der Durchschnitt

$$\cap : \mathcal{P}(M) \times \mathcal{P}(M) \rightarrow \mathcal{P}(M), \quad (A, B) \mapsto A \cap B$$

eine Operation, die sowohl kommutativ als auch assoziativ ist (\rightarrow *Nachrechnen*, z.B. mit Mengentafeln). Analog sind Vereinigung \cup und Differenz \setminus Operationen, letztere ist jedoch weder kommutativ noch assoziativ.

Weitere Standardbeispiele sind die Rechenoperationen $+$ und \cdot auf Zahlenbereichen (siehe Abschnitt A.2.2).

Definition 1.5.6

Seien M, N Mengen und $f : M \rightarrow N$ eine Abbildung. Dann heißt M *Definitionsbereich* und N *Wertebereich* von f . Zu $x \in M$ heißt $y = f(x)$ *Bildpunkt* oder *Funktionswert* von x

unter f . Die Menge aller Bildpunkte,

$$f(M) := \{f(x) : x \in M\} = \{y \in N : \exists x \in M : f(x) = y\} \subset N,$$

wird als **Bild** oder **Bildmenge** von M unter f bezeichnet. Zu einem Punkt $y \in f(M)$ heißt dann $x \in M$ mit $f(x) = y$ **Urbild** von y unter f .

Schließlich nennen wir

$$\Gamma(f) := \{(x, y) \in M \times N : y = f(x)\} = \{(x, f(x)) : x \in M\}$$

den **Graph** von f .

Vereinbarung. Eine Abbildung $f : M \rightarrow N$ bleibt unverändert, wenn wir den Wertebereich N durch eine andere Menge \hat{N} mit $f(M) \subset \hat{N}$ ersetzen. Insbesondere schreiben wir gelegentlich $f : M \rightarrow f(M)$.

Beispiel 1.5.7. ► Wir bestimmen das Bild der Abbildungen aus Beispiel 1.5.3:

(a) Wir benötigen eine Fallunterscheidung: $f(\mathbb{R}) = \begin{cases} \mathbb{R}, & \text{falls } m \neq 0 \\ \{b\}, & \text{falls } m = 0 \end{cases}$.

(b) $\pi(\mathbb{N}) = \{0, 1, 2, \dots\} = \mathbb{N} \cup \{0\}$, denn es gibt unendlich viele Primzahlen.

(c) $\text{id}_M(M) = M$.

(d) Sei $M \neq \emptyset$, dann: $\chi_A(M) = \begin{cases} \{0\}, & \text{falls } A = \emptyset \\ \{0, 1\}, & \text{falls } A \notin \{\emptyset, M\} \\ \{1\}, & \text{falls } A = M \end{cases}$.

(e) Es gilt immer $f|_A(A) \subset f(M)$ für beliebiges $A \subset M$. Z.B. für die charakteristische Funktion χ_A gilt $\chi_A|_A(A) = \{1\} \subsetneq \{0, 1\} = \chi_A(M)$, falls $A \notin \{\emptyset, M\}$. Anders als beim Wertebereich (siehe obige Vereinbarung), führt eine Abänderung des Definitionsbereichs also i.A. auch zu einer anderen Funktion.

► Sind $M, N \subset \mathbb{R}$, so lässt sich der Graph $\Gamma(f)$ einer Funktion $f : M \rightarrow N$ wie aus der Schule gewohnt im x, y -Koordinatensystem veranschaulichen.

Definition 1.5.8

Sind M, N, O drei Mengen und $f : M \rightarrow N, g : N \rightarrow O$ zwei Abbildungen, so setzen wir

$$g \circ f : M \rightarrow O, \quad x \mapsto (g \circ f)(x) := g(f(x))$$

für die **Komposition** von f mit g (lies: „ g nach f “ oder „ g kringel f “).

Bemerkung 1.5.9. (a) Die Komposition $g \circ f$ wird immer *von rechts nach links* ausgeführt. Man schreibt manchmal auch zur Veranschaulichung

$$M \xrightarrow{f} N \xrightarrow{g} O.$$

- (b) Bezeichnet $F(M)$ die Menge der Abbildungen $f : M \rightarrow M$, so ist durch die Komposition eine Operation $\circ : F(M) \times F(M) \rightarrow F(M)$, $(g, f) \mapsto g \circ f$ im Sinne von Definition 1.5.4 erklärt. Die Komposition ist immer assoziativ, aber i.A. *nicht* kommutativ (\rightarrow Nachrechnen).
- (c) In Definition 1.5.8 stimmen der Wertebereich von f und der Definitionsbereich von g überein. Wegen obiger Vereinbarung kann diese Bedingung abgeschwächt werden zu $f(M) \subset N$.

Definition 1.5.10

Eine Abbildung $f : M \rightarrow N$ heißt

- ▶ *injektiv*, wenn jeder Bildpunkt $y \in f(M)$ genau ein Urbild hat,
- ▶ *surjektiv*, wenn jeder Punkt $y \in N$ Bildpunkt ist, d.h. $f(M) = N$,
- ▶ *bijektiv*, wenn f sowohl injektiv als auch surjektiv ist.

Bemerkung 1.5.11. ▶ Eine Abbildung $f : M \rightarrow f(M)$ ist immer surjektiv, insbesondere sind injektive Abbildungen $f : M \rightarrow f(M)$ immer bijektiv. Sprechweise: *Injektive Abbildungen sind bijektiv auf ihr Bild.*

- ▶ Eine Abbildung $f : M \rightarrow N$ ist genau dann injektiv, wenn für alle $x_1, x_2 \in M$ gilt

$$f(x_1) = f(x_2) \Rightarrow x_1 = x_2.$$

- ▶ $f : M \rightarrow N$ ist schließlich genau dann bijektiv, wenn jeder Punkt $y \in N$ genau ein Urbild $x \in M$ besitzt.

Beispiel 1.5.12. Wir beziehen uns wieder auf Beispiel 1.5.3:

- (a) Für $m \neq 0$ ist $f : \mathbb{R} \rightarrow \mathbb{R}$ bijektiv, für $m = 0$ weder injektiv noch surjektiv.
- (b) $\pi : \mathbb{N} \rightarrow \mathbb{N} \cup \{0\}$ ist surjektiv, aber nicht injektiv; z.B. gilt $\pi(3) = \pi(4) = 2$.
- (c) $\text{id}_M : M \rightarrow M$ ist natürlich bijektiv.
- (d) $\chi_A : M \rightarrow \{0, 1\}$ ist für $M \neq \emptyset$ genau dann surjektiv, wenn $A \notin \{\emptyset, M\}$ gilt. Und χ_A ist injektiv genau dann, wenn $\#M = 2$, $\#A = 1$ gilt.

Satz 1.5.13

Eine Abbildung $f : M \rightarrow N$ ist genau dann bijektiv, wenn eine Abbildung $\varphi : N \rightarrow M$ so existiert, dass gilt

$$\varphi \circ f = \text{id}_M, \quad f \circ \varphi = \text{id}_N. \quad (1.5.1)$$

In diesem Fall ist φ eindeutig bestimmt.

Beweis. ▶ *Existenz:* Wir beweisen die behauptete Äquivalenz mit der Äquivalenzregel, siehe Bemerkung 1.1.9, also durch getrennte Beweise der Hin- und Rückrichtung.

▷ „ \Rightarrow “: Sei f bijektiv. Zu beliebigem $y \in N$ existiert dann genau ein Urbild $x \in M$ mit $y = f(x)$. Wir definieren dann $\varphi : N \rightarrow M$ durch die Zuordnung $y \mapsto x = \varphi(y)$. Dann ist (1.5.1) offenbar erfüllt.

▷ „ \Leftarrow “: Sei $\varphi : N \rightarrow M$ mit (1.5.1) gegeben. Für beliebiges $y \in N$ folgt dann

$$y = \text{id}_N(y) = (f \circ \varphi)(y) = f(\varphi(y)),$$

d.h. $\varphi(y) \in M$ ist Urbild von y , folglich ist f surjektiv. Sind $x_1, x_2 \in M$ mit $f(x_1) = f(x_2)$ gewählt, so folgt weiter

$$x_1 = \text{id}_M(x_1) \stackrel{(1.5.1)}{=} \varphi(f(x_1)) = \varphi(f(x_2)) \stackrel{(1.5.1)}{=} \text{id}_M(x_2) = x_2,$$

d.h. f ist auch injektiv, insgesamt also bijektiv.

▶ *Eindeutigkeit:* Gäbe es zwei Funktionen $\varphi_1, \varphi_2 : N \rightarrow M$ mit Eigenschaft (1.5.1), so folgte

$$\varphi_1 = \varphi_1 \circ \text{id}_N \stackrel{(1.5.1)}{=} \varphi_1 \circ (f \circ \varphi_2) \stackrel{\text{Bem. 1.5.9(b)}}{=} (\varphi_1 \circ f) \circ \varphi_2 \stackrel{(1.5.1)}{=} \text{id}_M \circ \varphi_2 = \varphi_2,$$

d.h. φ ist eindeutig bestimmt. □

Definition 1.5.14

Ist $f : M \rightarrow N$ bijektiv, so heißt die gemäß Satz 1.5.13 eindeutig bestimmte Funktion $f^{-1} := \varphi : N \rightarrow M$ mit der Eigenschaft

$$f^{-1} \circ f = \text{id}_M, \quad f \circ f^{-1} = \text{id}_N$$

die *Umkehrabbildung* oder *Umkehrfunktion* zu f . Bijektive Funktionen werden daher auch *umkehrbar* genannt.

Hilfssatz 1.5.15

Sind $f : M \rightarrow N$, $g : N \rightarrow O$ umkehrbar, so ist auch $g \circ f : M \rightarrow O$ umkehrbar mit der Umkehrfunktion

$$(g \circ f)^{-1} = f^{-1} \circ g^{-1} : O \rightarrow M.$$

Beweis. Man zeigt, dass $\varphi = f^{-1} \circ g^{-1}$ die Eigenschaften aus Definition 1.5.14 bzw. (1.5.1) erfüllt (\rightarrow Ergänzungen). \square

Kapitel 2.

Reelle und komplexe Zahlen

Mathematik ohne Zahlen ist undenkbar. Wir beginnen mit einer Diskussion der aus der Schule bekannten Zahlenbereiche und geben eine axiomatische Definition der reellen Zahlen als vollständig archimedisch angeordneter Körper. Die nächsten Paragraphen sind den Rechenregeln in angeordneten Körpern gewidmet, und wir führen das Beweisprinzip der vollständigen Induktion ein. Abschließend betrachten wir den (vollständig) bewerteten Körper der komplexen Zahlen.

2.1. Zahlenbereiche

Wir erinnern zunächst an die „Typen“ von Zahlen, die wir aus der Schule kennen, hier aber kritisch betrachten wollen:

Natürliche und ganze Zahlen

Die *natürlichen Zahlen* wollen wir als bekannt voraussetzen:

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

Eine Definition mit Hilfe der *Peanoschen Axiomen* entnimmt man z.B. [DeRa] Abschnitt 2.1, eine Definition als *kleinste induktive Teilmenge der reellen Zahlen* findet man z.B. in [Hil1] Abschnitt 1.4. Oft nutzen wir auch die Erweiterung

$$\mathbb{N}_0 := \mathbb{N} \cup \{0\}.$$

Ebenfalls voraussetzen wollen wir die *ganzen Zahlen*

$$\mathbb{Z} := \{\dots, -2, -1, 0, 1, 2, \dots\} = \{x : x \in \mathbb{N}_0 \vee -x \in \mathbb{N}\}.$$

Definieren kann man eine *negative* ganze Zahl $x \in \mathbb{Z} \setminus \mathbb{N}_0$ als Lösung der Gleichung

$$x + m = 0$$

zu vorgegebenem $m \in \mathbb{N}$. Dieses *ideale Element* $x =: -m$ wurde sozusagen erdacht, um die Lösbarkeit dieser in \mathbb{N} nicht lösbaren Gleichung zu erzwingen.

Bemerkung 2.1.1. Es gilt $\mathbb{N} \subset \mathbb{Z}$. Und wir vereinbaren, dass wir auch die Rechenregeln der Operationen

$$\begin{aligned} + : \mathbb{Z} \times \mathbb{Z} &\rightarrow \mathbb{Z}, & (x, y) &\mapsto x + y && (\text{Addition}), \\ \cdot : \mathbb{Z} \times \mathbb{Z} &\rightarrow \mathbb{Z}, & (x, y) &\mapsto x \cdot y && (\text{Multiplikation}) \end{aligned}$$

als bekannt voraussetzen; wir schreiben wie üblich auch $xy = x \cdot y$ für die Multiplikation.

Rationale Zahlen

Die *rationalen Zahlen* werden als (ideale) Lösungen der Gleichung

$$q \cdot x = p \tag{2.1.1}$$

zu gegebenen $p \in \mathbb{Z}$, $q \in \mathbb{N}$ erklärt. Wir schreiben zunächst $x = (p, q)$ für diese Lösung und betrachten folgende Äquivalenzrelation: Auf $\mathbb{Z} \times \mathbb{N}$ erklären wir

$$(\tilde{p}, \tilde{q}) \sim (p, q) :\Leftrightarrow \tilde{p} \cdot q = p \cdot \tilde{q}.$$

Man rechnet leicht nach, dass diese Relation die Eigenschaften (R), (S) und (T) aus Definition 1.4.5 erfüllt. Ist $x = (p, q)$ dann die Lösung von (2.1.1), so folgt für $(\tilde{p}, \tilde{q}) \sim (p, q)$:

$$\tilde{p}q = p\tilde{q} \stackrel{(2.1.1)}{=} (qx)\tilde{q} = q(\tilde{q}x),$$

woraus $\tilde{p} = \tilde{q} \cdot x$ folgt. Also ist x durch (2.1.1) nur modulo \sim festgelegt (eine ähnliche Rechnung zeigt noch, dass x keine weitere Gleichung (2.1.1) löst); wir schreiben dann für die Äquivalenzklassen

$$\frac{p}{q} := [(p, q)]$$

und nennen $\frac{p}{q}$ die rationale Zahl oder den *Bruch* mit *Zähler* $p \in \mathbb{Z}$ und *Nenner* $q \in \mathbb{N}$. Für die Menge der rationalen Zahlen schreiben wir

$$\mathbb{Q} := \left\{ \frac{p}{q} : p \in \mathbb{Z}, q \in \mathbb{N} \right\}.$$

Bemerkung 2.1.2. ► Wir halten noch einmal fest: verschiedene Paare $(p, q) \sim (\tilde{p}, \tilde{q})$ definieren denselben Bruch; z.B. ist also $\frac{1}{2} = \frac{3}{6}$. Insbesondere hat jeder Bruch ungleich 0 eine Darstellung $\frac{p}{q}$ mit *teilerfremden* $p \in \mathbb{Z}$, $q \in \mathbb{N}$. Hier kommt uns die Konvention zugute, dass Mehrfachnennungen von Elementen in einer Menge zulässig sind; siehe Definition 1.2.4.

- Wir können offenbar die Zahlen $p \in \mathbb{Z}$ und $\frac{p}{1} \in \mathbb{Q}$ *identifizieren*, also gemäß $p = \frac{p}{1}$ gleichsetzen, denn es gilt

$$\frac{p}{1} = 1 \cdot \frac{p}{1} \stackrel{(2.1.1)}{=} p.$$

In diesem Sinne haben wir also $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q}$.

Wir erinnern schließlich an die Definition von Addition und Multiplikation für Brüche $\frac{p_1}{q_1}, \frac{p_2}{q_2} \in \mathbb{Q}$:

$$\frac{p_1}{q_1} + \frac{p_2}{q_2} := \frac{p_1 q_2 + p_2 q_1}{q_1 q_2} \in \mathbb{Q}, \quad \frac{p_1}{q_1} \cdot \frac{p_2}{q_2} := \frac{p_1 p_2}{q_1 q_2} \in \mathbb{Q}. \quad (2.1.2)$$

Als *Übungsaufgabe* rechnet man nach, dass diese Setzungen unabhängig vom Repräsentanten der Brüche sind.

Reelle Zahlen

Die Menge der rationalen Zahlen weist schon eine wichtige Struktur auf, sie ist nämlich ein *Körper*; siehe Definition 2.2.1 und Satz 2.2.3 unten. Aber sie hat ein großes Manko, es gibt einfach nicht genug rationale Zahlen:

Beispiel 2.1.3. Die Lösung $x = \sqrt{2}$ der Gleichung $x^2 = 2$ ist nicht rational.

Beweis. Wäre $x = \frac{p}{q}$ richtig mit $p \in \mathbb{Z}, q \in \mathbb{N}$ teilerfremd, so müsste $\frac{p^2}{q^2} = \left(\frac{p}{q}\right)^2 = 2$ bzw. $p^2 = 2q^2$ gelten. Damit wäre aber p^2 und daher auch p durch 2 teilbar, d.h. $p = 2l$ mit einem $l \in \mathbb{Z}$ und folglich

$$q^2 = \frac{p^2}{2} = 2l^2.$$

Also wäre auch q^2 und somit q durch 2 teilbar, im Widerspruch zur Annahme, dass p und q teilerfremd sind. \square

Wir erweitern \mathbb{Q} daher zur Menge der *reellen Zahlen* $\mathbb{R} \supset \mathbb{Q}$, die wir uns auf der Zahlengerade angeordnet denken:

Definition 2.1.4: Reelle Zahlen

Es gibt eine Menge $\mathbb{R} \supset \mathbb{Q}$ von Zahlen, genannt *reelle Zahlen*, die drei Gruppen von Grundgesetzen (Axiomen) erfüllen:

- ▶ die Körperaxiome (A1)-(A4), (M1)-(M4), (D)
- ▶ die Anordnungsaxiome (O1), (O2) und das Archimedische Axiom (O3)
- ▶ das Vollständigkeitsaxiom (V)

Die reellen Zahlen bilden also einen *vollständig archimedisch angeordneten Körper*. Die Elemente der (nichtleeren) Menge $\mathbb{R} \setminus \mathbb{Q}$ heißen *irrationale Zahlen*.

Die erwähnten Axiome werden wir in den Abschnitten A.2.2 (Körperaxiome) und A.2.3 (Anordnungsaxiome und Archimedisches Axiom) diskutieren; das Vollständigkeitsaxiom erfordert in unserer Formulierung die Betrachtung von Zahlenfolgen und wird erst in Kapitel 3 erklärt.

- Bemerkung 2.1.5.** (a) Weitere Beispiele irrationaler Zahlen sind die Kreiszahl π und die Eulersche Zahl e . Doch was sind das genau für Zahlen? $\pi = 3,14159\dots$. Aber für welches „und so weiter“ stehen hier die Punkte? Neben dem in Definition 2.1.4 angegebenen *axiomatischen Ansatz*, kann man \mathbb{R} auch *konstruktiv* durch *Vervollständigung* der rationalen Zahlen gewinnen (\rightarrow *Ergänzungen*, Kapitel 3).
- (b) Wir werden später in Kapitel 3 sehen, dass *überabzählbar* unendlich viele irrationale Zahlen existieren, und in diesem Sinne sogar mehr als es rationale Zahlen gibt.
- (c) Man kann zeigen, dass \mathbb{R} durch die Axiome in Definition 2.1.4 (bis auf Isomorphie) eindeutig bestimmt ist. Wir verzichten hier auf nähere Erklärungen und verweisen auf den Paragraphen „Eine algebraische Charakterisierung“ in [Dei] Abschnitt 1.3.

2.2. Die Körperaxiome

Definition 2.2.1: Körper

Ein Tripel $(\mathbb{K}, +, \cdot)$ heißt *Körper* mit der nichtleeren *Grundmenge* \mathbb{K} und den *Rechenoperationen* $+, \cdot : \mathbb{K} \times \mathbb{K} \rightarrow \mathbb{K}$, wenn die folgenden *Körperaxiome* erfüllt sind:

► **Axiome der Addition.**

- (A1) *Kommutativität*: Für alle $a, b \in \mathbb{K}$ gilt $a + b = b + a$.
- (A2) *Assoziativität*: Für alle $a, b, c \in \mathbb{K}$ gilt $(a + b) + c = a + (b + c)$.
- (A3) *Existenz des Nullelements 0*: Es existiert ein neutrales Element $0 \in \mathbb{K}$, d.h. für alle $a \in \mathbb{K}$ gilt $a + 0 = a$.
- (A4) *Existenz des Negativen*: Für alle $a \in \mathbb{K}$ existiert ein $b \in \mathbb{K}$, so dass $a + b = 0$ richtig ist; wir schreiben $b = -a$.

► **Axiome der Multiplikation.**

- (M1) *Kommutativität*: Für alle $a, b \in \mathbb{K}$ gilt $a \cdot b = b \cdot a$.
- (M2) *Assoziativität*: Für alle $a, b, c \in \mathbb{K}$ gilt $(a \cdot b) \cdot c = a \cdot (b \cdot c)$.
- (M3) *Existenz des Einselements 1*: Es existiert ein neutrales Element $1 \in \mathbb{K} \setminus \{0\}$, d.h. für alle $a \in \mathbb{K}$ gilt $a \cdot 1 = a$.
- (M4) *Existenz des Inversen*: Für alle $a \in \mathbb{K} \setminus \{0\}$ existiert ein $b \in \mathbb{K}$, so dass $a \cdot b = 1$ richtig ist; wir schreiben $b = a^{-1}$.

► **Distributivgesetz.**

- (D) Für alle $a, b, c \in \mathbb{K}$ gilt $a \cdot (b + c) = a \cdot b + a \cdot c$.

Die Elemente eines Körpers nennen wir *Zahlen*. Wenn klar ist, welche Rechenoperationen

benutzt werden, schreiben wir auch einfach \mathbb{K} für den betrachteten Körper.

Bemerkung 2.2.2. ▶ In einem Körper können wir noch *Subtraktion* und *Division* erklären:

$$a - b := a + (-b) \in \mathbb{K} \quad \text{für } a, b \in \mathbb{K},$$

$$\frac{a}{b} := a \cdot b^{-1} \in \mathbb{K} \quad \text{für } a \in \mathbb{K}, b \in \mathbb{K} \setminus \{0\}.$$

- ▶ Nach Definition 2.1.4 ist \mathbb{R} ein Körper. \mathbb{N} und \mathbb{Z} sind keine Körper - beiden fehlt das Inverse, d.h. (M4) ist verletzt, den natürlichen Zahlen fehlt auch das Negative und sogar das Nullelement 0. Für die rationalen Zahlen \mathbb{Q} gilt der

Satz 2.2.3

Die Menge \mathbb{Q} der rationalen Zahlen ist - zusammen mit den Rechenoperationen aus (2.1.2) - ein Körper.

Beweis. → Ergänzungen. □

Bemerkung 2.2.4. Ein Körper muss nach (A3) und (M3) mindestens zwei Elemente enthalten, nämlich das Nullelement 0 und das Einselement $1 \neq 0$. Umgekehrt kann man jede zweielementige Menge $M = \{x, y\}$ durch geeignete Definition der Verknüpfungen zu einem Körper machen:

+	x	y
x	x	y
y	y	x

und

·	x	y
x	x	x
y	x	y

Hierbei wird x als Nullelement und y als Einselement interpretiert.

Bei unserer axiomatischen Definition des Körpers ist - wie bei jeder anderen Axiomatik auch - ein Grundprinzip das Vermeiden von Redundanz. Viele Ihnen aus der Schule geläufige Rechenregeln lassen sich aus den Körperaxiomen ableiten, die wichtigsten fassen wir zusammen zum

Satz 2.2.5: Rechnen in Körpern

In einem Körper \mathbb{K} (z.B. $\mathbb{K} = \mathbb{Q}$ oder $\mathbb{K} = \mathbb{R}$) gelten folgende Rechenregeln:

- Die Gleichung $a + x = b$ besitzt für beliebig vorgegebene $a, b \in \mathbb{K}$ genau eine Lösung $x \in \mathbb{K}$. Insbesondere sind das Nullelement 0 und das negative Element eindeutig bestimmt.
- Die Gleichung $ax = b$ besitzt für beliebig vorgegebene $a \in \mathbb{K} \setminus \{0\}$, $b \in \mathbb{K}$ genau

eine Lösung $x \in \mathbb{K}$. Insbesondere sind das Einselement 1 und das inverse Element eindeutig bestimmt.

- (c) Für alle $x \in \mathbb{K}$ gilt $x \cdot 0 = 0$ und $(-1) \cdot x = -x$.
- (d) Für alle $x \in \mathbb{K}$ gilt $-(-x) = x$ und, falls zusätzlich $x \neq 0$, auch $(x^{-1})^{-1} = x$.
- (e) Für alle $x, y \in \mathbb{K} \setminus \{0\}$ ist $xy \neq 0$ richtig.
- (f) Für alle $x, y \in \mathbb{K}$ ist $-(x + y) = -x - y$ richtig und, falls zusätzlich $x, y \neq 0$ gilt, auch $(xy)^{-1} = x^{-1}y^{-1}$.

Beweis. Hier beweisen wir (a), (c) und (e), die Aussagen (b), (d) und (f) in den Ergänzungen.

- (a) Wir zeigen zunächst, dass $x := b + (-a) = b - a$ die Gleichung $a + x = b$ löst, d.h. wir beweisen die *Existenz einer Lösung*:

$$\begin{aligned} a + x &= a + (b - a) \stackrel{(A1)}{=} a + ((-a) + b) \stackrel{(A2)}{=} (a + (-a)) + b \\ &\stackrel{(A4)}{=} 0 + b \stackrel{(A1)}{=} b + 0 \stackrel{(A3)}{=} b. \end{aligned}$$

Die *Eindeutigkeit der Lösung* ergibt sich wie folgt: Gäbe es zwei Lösungen $x_1, x_2 \in \mathbb{K}$ mit $x_1 \neq x_2$, d.h. es gälte

$$a + x_1 = b = a + x_2.$$

Addieren wir von rechts auf beiden Seiten $-a$, so folgte

$$\begin{aligned} (a + x_1) + (-a) &= (a + x_2) + (-a) \\ \stackrel{(A1), (A2)}{\Rightarrow} x_1 + (a + (-a)) &= x_2 + (a + (-a)) \\ \stackrel{(A4)}{\Rightarrow} x_1 + 0 &= x_2 + 0 \\ \stackrel{(A3)}{\Rightarrow} x_1 &= x_2, \end{aligned}$$

also ein Widerspruch. Folglich gibt es genau eine Lösung.

- (c) $x \cdot 0 = 0$: Nach (A3) gilt $0 + 0 = 0$ und folglich

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

Addition von $-(x \cdot 0)$ auf beiden Seiten von rechts (und Ausnutzen der Axiome (A2), (A4) und (A3)) liefert die Behauptung.

$(-1) \cdot x = -x$: Es gilt

$$x + (-1) \cdot x \stackrel{(M3), (M1)}{=} x \cdot 1 + x \cdot (-1) \stackrel{(D)}{=} x \cdot (1 + (-1)) \stackrel{(A4)}{=} x \cdot 0 = 0.$$

Die *Eindeutigkeit des Negativen* – siehe (a) – liefert die Behauptung.

- (e) Angenommen, es gibt $x, y \in \mathbb{K} \setminus \{0\}$ mit $xy = 0$. Nach Multiplikation mit y^{-1} (beachte $y \neq 0$) von rechts folgt

$$x \stackrel{(M3)}{=} x \cdot 1 \stackrel{(M4)}{=} x(yy^{-1}) \stackrel{(M2)}{=} (xy)y^{-1} = 0 \cdot y^{-1} \stackrel{(c), (M1)}{=} 0,$$

im Widerspruch zur Voraussetzung $x \neq 0$. Somit ist die Behauptung richtig. \square

Bemerkung 2.2.6. In Mehrfachsummen und Mehrfachprodukten lassen wir i.F. die Klammern meist weg, also $a + b + c + \dots$ und $a \cdot b \cdot c \cdot \dots$ für $a, b, c, \dots \in \mathbb{K}$, denn wegen (A2) und (M2) spielt die Reihenfolge der Summation bzw. Multiplikation keine Rolle. Ebenso können wir in Mehrfachsummen und Mehrfachprodukten die Reihenfolge der Summanden bzw. Faktoren beliebig vertauschen.

2.3. Angeordnete Körper

Wir wenden uns nun den Anordnungs- und dem Archimedischen Axiom zu: Für das in der Analysis wesentliche Rechnen mit Ungleichungen müssen wir entscheiden können, ob ein Element eines Körpers „größer“ oder „kleiner“ als ein anderes Element ist. Hierzu verwenden wir die

Definition 2.3.1: Angeordnete Körper

Wir nennen einen Körper $(\mathbb{K}, +, \cdot)$ *angeordnet*, wenn eine Teilmenge $P \subset \mathbb{K}$ *positiver* Elemente existiert - wir schreiben $x > 0$ für $x \in P$ -, so dass folgende *Anordnungsaxiome* erfüllt sind:

- (O1) *Trichotomiegesetz*. Für jedes $x \in \mathbb{K}$ gilt genau eine der drei Beziehungen

$$x > 0, \quad x = 0, \quad -x > 0.$$

Die $x \in \mathbb{K}$ mit $-x > 0$ heißen die *negativen Elemente*.

- (O2) *Abgeschlossenheit bez. +, ·*. Für alle $x, y \in \mathbb{K}$ mit $x > 0$ und $y > 0$ gilt

$$x + y > 0 \quad \text{und} \quad xy > 0.$$

Beispiel 2.3.2. (a) Nach unserer Festlegung in Definition 2.1.4 bilden die reellen Zahlen einen angeordneten Körper.

- (b) Auch der Körper \mathbb{Q} der rationalen Zahlen ist angeordnet mittels

$$x > 0 \iff x = \frac{p}{q} \text{ mit } p, q \in \mathbb{N}.$$

Dann sind die $x = \frac{p}{q}$ mit $-p, q \in \mathbb{N}$ gerade die negativen Zahlen, (O1) ist also erfüllt (wegen $\frac{0}{q} = 0$). Und (O2) folgt sofort aus der Definition (2.1.2) von Addition und Multiplikation rationaler Zahlen. Diese Anordnung entspricht unserer Vorstellung der Anordnung von \mathbb{Q} auf der Zahlengeraden (vgl. auch Definition 2.3.3 unten).

(c) Der Körper $(\{x, y\}, +, \cdot)$ aus Bemerkung 2.2.4 kann nicht angeordnet werden.

Denn: Da x das Nullelement ist, müsste anderenfalls für y entweder $y > x$ oder $-y > x$ gelten. Per Definition ist aber $-y \in \{x, y\}$, also $-y = y$, Widerspruch! \square

Definition 2.3.3: Größer- und Kleinerrelation

In einem angeordneten Körper definieren wir:

$$x > y \quad :\Leftrightarrow \quad x - y > 0,$$

$$x < y \quad :\Leftrightarrow \quad y > x \Leftrightarrow y - x > 0,$$

$$x \geq y \quad :\Leftrightarrow \quad (x > y) \vee (x = y),$$

$$x \leq y \quad :\Leftrightarrow \quad y \geq x \Leftrightarrow (x < y) \vee (x = y).$$

Satz 2.3.4: Rechnen in angeordneten Körpern

In einem angeordneten Körper \mathbb{K} gelten folgende Aussagen:

(a) Für je zwei Elemente $x, y \in \mathbb{K}$ gilt genau eine der Relationen

$$x < y, \quad x = y, \quad x > y.$$

(b) *Transitivität:* Für alle $x, y, z \in \mathbb{K}$ gilt: $x < y$ und $y < z$ implizieren $x < z$.

(c) *Translationsinvarianz:* Für alle $x, y, a \in \mathbb{K}$ gilt: Aus $x < y$ folgt $x + a < y + a$.

(d) *Skalierungsinvarianz:* Für alle $x, y, a \in \mathbb{K}$ mit $x < y$ und $a > 0$ gilt $xa < ya$.

(e) *Spiegelung:* Für alle $x, y \in \mathbb{K}$ mit $x < y$ haben wir $-x > -y$.

(f) Für alle $x \in \mathbb{K} \setminus \{0\}$ ist $x^2 > 0$ richtig; insbesondere gilt $1 > 0$.

(g) Für jedes $x \in \mathbb{K}$ mit $x > 0$ ist $x^{-1} > 0$ erfüllt.

(h) Für alle $x, y \in \mathbb{K}$ mit $0 < x < y$ gilt $x^{-1} > y^{-1}$.

Bemerkung 2.3.5. Wegen Satz 2.3.4 (a) sind in einem angeordneten Körper für je zwei Elemente

$x, y \in \mathbb{K}$ das *Minimum* $\min\{x, y\}$ und das *Maximum* $\max\{x, y\}$ wohl definiert:

$$\min\{x, y\} := \begin{cases} x, & \text{falls } x \leq y \\ y, & \text{sonst} \end{cases}, \quad \max\{x, y\} := \begin{cases} x, & \text{falls } x \geq y \\ y, & \text{sonst} \end{cases}.$$

Beweis von Satz 2.3.4. Wir werden die Körperaxiome und deren Folgerungen aus Satz 2.2.5 ohne Kommentar benutzen. Aussagen (d), (e), (g) und (h) beweisen wir in den Ergänzungen.

(a) Wende (O1) auf $x - y$ an und benutze Definition 2.3.3.

(b) Per Voraussetzung ist $y - x > 0$ und $z - y > 0$, so dass (O2) liefert

$$z - x = (z - y) + (y - x) > 0 \quad \text{bzw.} \quad x < z.$$

(c) Aus der Voraussetzung $y - x > 0$ folgt sofort

$$(y + a) - (x + a) = y - x > 0 \quad \text{bzw.} \quad x + a < y + a.$$

(f) Wegen $x \neq 0$ muss nach (O1) entweder $x > 0$ oder $-x > 0$ gelten. (O2) liefert dann

$$x^2 = x \cdot x > 0$$

bzw.

$$x^2 = x \cdot (-(-x)) = x \cdot ((-1)(-x)) = ((-1)x) \cdot (-x) = (-x)(-x) > 0.$$

Schließlich beachten wir noch $1 = 1 \cdot 1 > 0$. □

Definition 2.3.6

Ein angeordneter Körper $\mathbb{K} \supset \mathbb{N}$ heißt *archimedisch angeordnet*, wenn das *Archimedische Axiom* gilt:

(O3) Zu jedem $c \in \mathbb{K}$ existiert ein $n \in \mathbb{N}$ mit $n > c$.

Nach unserer obigen Festlegung ist \mathbb{R} also ein archimedisch angeordneter Körper. Es gilt auch der

Hilfssatz 2.3.7

Der Körper \mathbb{Q} der rationalen Zahlen ist archimedisch angeordnet.

Beweis. Gemäß Beispiel 2.3.2 (b) haben wir nur noch (O3) nachzurechnen: Sei dazu $c = \frac{p}{q} \in \mathbb{Q}$ beliebig. Für $c \leq 0$ wählen wir $n = 1 > 0 \geq c$ gemäß Satz 2.3.4 (b), (f) und sind fertig. Für $c > 0$

bzw. $p \in \mathbb{N}$ setzen wir $n := p + 1$ und erhalten

$$\begin{aligned} n &= p + 1 = \frac{p}{q} \cdot q + 1 \\ &\stackrel{\text{Satz 2.3.4 (c), (d)}}{\geq} \frac{p}{q} + 1 = c + 1 \\ &\stackrel{\text{Satz 2.3.4 (c), (f)}}{>} c, \end{aligned}$$

wie behauptet. □

Bemerkung 2.3.8. ▶ Archimedes hat (O3) geometrisch formuliert: Hat man zwei Strecken auf einer Geraden, so kann man, in dem man die kürzere hinreichend oft abträgt, die längere übertreffen.

- ▶ Gemäß Hilfssatz 2.3.7 erfüllt \mathbb{Q} alle Axiome der reellen Zahlen, abgesehen vom Vollständigkeitsaxiom. Genau diese zusätzliche Eigenschaft der reellen Zahlen wird uns in Kapitel 3 das Lösen von Gleichungen wie $x^2 = 2$ ermöglichen.

In einem angeordneten Körper kann man immer einen Abstandsbegriff erklären; wir werden in Abschnitt A.2.6 sehen, dass damit jeder angeordnete auch ein *bewerteter Körper* ist.

Definition 2.3.9

Sei \mathbb{K} ein angeordneter Körper (z.B. $\mathbb{K} = \mathbb{R}$). Zu beliebigem $x \in \mathbb{K}$ erklären wir den *Betrag* oder *Absolutbetrag* von x als

$$|x| := \begin{cases} x, & \text{falls } x \geq 0 \\ -x, & \text{sonst} \end{cases}.$$

Bemerkung 2.3.10. Für den so erklärten Absolutbetrag gilt

$$|x| = \max\{x, -x\} \quad \text{für alle } x \in \mathbb{K}$$

und folglich

$$-|x| \leq x \leq |x| \quad \text{für alle } x \in \mathbb{K}.$$

2.4. Vollständige Induktion

Wir lernen nun ein wichtiges Beweisprinzip kennen und anwenden, welches darauf beruht, dass jede Zahl $n \in \mathbb{N}_0 = \mathbb{N} \cup \{0\}$ einen eindeutig definierten Nachfolger, nämlich $n + 1 \in \mathbb{N}$, besitzt:

Hilfssatz 2.4.1: Beweisprinzip der vollständigen Induktion

Sei p eine Aussageform (siehe Definition 1.1.10) und $n_0 \in \mathbb{N}_0$. Dann ist die Aussage $p(n)$ für alle $n \in \mathbb{N}_0$ mit $n \geq n_0$ wahr, falls folgendes gilt:

(IA) *Induktionsanfang*: Die Aussage $p(n_0)$ ist wahr.

(IS) *Induktionsschritt*: Für alle $n \geq n_0$ gilt: Wenn $p(n)$ wahr ist, so ist auch $p(n+1)$ wahr.

Beweis. Sind (IA) und (IS) erfüllt und angenommen, es existiert ein $n > n_0$, für das $p(n)$ nicht wahr ist. Wegen (IS) ist dann auch $p(n-1)$ falsch und dann $p(n-2)$, $p(n-3)$ usw. Nach $n - n_0$ Schritten würde also folgen, dass auch $p(n_0)$ falsch ist, im Widerspruch zu (IA). \square

Für unsere erste Anwendung benötigen wir noch die

Definition 2.4.2

Sei K ein Körper und $a \in K$ gewählt. Dann erklären wir die *n -te Potenz* induktiv als

$$a^0 := 1, \quad a^{n+1} := a^n \cdot a \quad \text{für } n \in \mathbb{N}_0.$$

Gilt zusätzlich $a \neq 0$, so erklären wir auch negative Potenzen:

$$a^{-n} := (a^{-1})^n \quad \text{für } n \in \mathbb{N}.$$

Bemerkung 2.4.3. ► Für beliebige $a, b \in K \setminus \{0\}$ und $n, m \in \mathbb{Z}$ folgen aus Definition 2.4.2 die bekannten Potenzgesetze (\rightarrow Übungsaufgaben):

$$a^n a^m = a^{n+m}, \quad (a^n)^m = a^{nm}, \quad a^n b^n = (ab)^n. \quad (2.4.1)$$

► Analog zu Definition 2.4.2 erklärt man den Ausdruck na für $a \in K$ als

$$0a := 0, \quad (n+1)a := na + a \quad \text{für } n \in \mathbb{N}_0,$$

sowie

$$(-n)a := -(na), \quad n \in \mathbb{N},$$

und es gelten die bekannten, zu (2.4.1) analogen Rechenregeln. (*Nachrechnen!*)

Satz 2.4.4: BERNOULLISCHE UNGLEICHUNG

Sei K ein angeordneter Körper (z.B. $K = \mathbb{R}$) und $x \in K$ mit $x \geq -1$ sei gewählt. Dann gilt

für alle $n \in \mathbb{N}_0$ die Ungleichung

$$(1+x)^n \geq 1+nx.$$

Beweis. Mittels vollständiger Induktion:

(IA) $n = 0$: Wir haben zu zeigen, dass $p(0)$ wahr ist, also in unserem Fall

$$(1+x)^0 \geq 1+0 \cdot x.$$

Das ist offenbar richtig.

(IS) $n \rightarrow n+1$: Es sei $p(n)$, also $(1+x)^n \geq 1+nx$, für ein $n \in \mathbb{N}_0$ wahr (genannt *Induktionsvoraussetzung* (IV)). Zu zeigen ist, dass $p(n+1)$ mit diesem n wahr ist, d.h.

$$(1+x)^{n+1} \geq 1+(n+1)x.$$

Hierzu berechnen wir m.H. der Induktionsvoraussetzung (beachte $1+x \geq 0$):

$$\begin{aligned} (1+x)^{n+1} &= (1+x)^n(1+x) \stackrel{(IV)}{\geq} (1+nx)(1+x) \\ &= 1+nx+x+nx^2 \geq 1+(n+1)x, \end{aligned}$$

d.h. $p(n+1)$ ist wahr. Also ist auch der Induktionsschritt (IS) bewiesen und nach dem Prinzip der vollständigen Induktion gilt die Behauptung für alle $n \in \mathbb{N}_0$. \square

Bevor wir die nächste Anwendung der vollständigen Induktion angeben, schieben wir eine wichtige Folgerung von Satz 2.4.4 ein, die in beliebigen *archimedisch* angeordneten Körpern gilt.

Folgerung 2.4.5

Sei \mathbb{K} ein archimedisch angeordneter Körper und sei $b \in \mathbb{K}$ positiv. Dann gilt:

- (a) Ist $b > 1$, so existiert zu jedem positiven $K \in \mathbb{K}$ ein $N \in \mathbb{N}$ mit der Eigenschaft $b^N > K$.
- (b) Ist $0 < b < 1$, so existiert zu jedem positiven $\delta \in \mathbb{K}$ ein $N \in \mathbb{N}$ mit der Eigenschaft $b^N < \delta$.

Beweis. (a) Wegen $b > 1$ ist $x := b - 1 > 0$ richtig. Also ist die Bernoullische Ungleichung, Satz 2.4.4, anwendbar: Für alle $n \in \mathbb{N}$ gilt

$$b^n = (1+x)^n \geq 1+nx.$$

Nach dem Archimedischen Axiom (O3) aus Definition 2.3.6 können wir $N \in \mathbb{N}$ so wählen, dass $N > \frac{K}{x}$ ausfällt. Dann folgt für $n = N$:

$$b^N \geq 1+Nx > 1+K > K.$$

(b) Wegen $0 < b < 1$ ist $\hat{b} := \frac{1}{b} > 1$ richtig. Nach (a) existiert zu $\hat{K} := \frac{1}{\delta}$ ein $N \in \mathbb{N}$ mit

$$\hat{b}^N > \hat{K} \quad \Leftrightarrow \quad b^N = (\hat{b}^N)^{-1} < \hat{K}^{-1} = \delta,$$

wie behauptet. □

Satz 2.4.6: GAUSSsche Summenformel

Für jede natürliche Zahl $n \in \mathbb{N}$ gilt die Relation

$$1 + 2 + 3 + \dots + n = \frac{n(n+1)}{2}.$$

Bemerkung 2.4.7. Die Punkte deuten an, dass die Summation in der gleichen Weise fortgesetzt wird. Dies kann man m.H. von *Summen- und Produktzeichen* wie folgt kompakter (und präziser) schreiben:

Hat man viele, eventuell unendlich viele Variablen, so benutzt man statt a, b, c, \dots sinnvoller die Bezeichnungen a_1, a_2, a_3, \dots . Die natürlichen Zahlen $1, 2, 3, \dots$ heißen hierbei *Indizes* und dienen der Unterscheidung der Variablen $a_k, k \in \mathbb{N}$.

Für $n \in \mathbb{N}$ Variablen a_1, a_2, \dots, a_n setzen wir

$$\begin{aligned} \sum_{k=1}^n a_k &:= a_1 + a_2 + \dots + a_n && \text{(Summe)}, \\ \prod_{k=1}^n a_k &:= a_1 \cdot a_2 \cdot \dots \cdot a_n && \text{(Produkt)}. \end{aligned}$$

Hierbei kann man die Indexmenge $1, \dots, n$ natürlich auch durch andere (endliche) Teilmengen von \mathbb{N} oder allgemeiner von \mathbb{Z} ersetzen. Mit dieser Notation erscheint die in Satz 2.4.6 behauptete Formel als ($a_k = k$ für $k = 1, \dots, n$):

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

Wir weisen noch auf die sogenannte *Indexverschiebung* hin: Man kann den Summationsbereich beliebig verschieben, indem man im Index entsprechend korrigiert; es gilt also z.B.

$$\sum_{k=1}^n a_k = \sum_{k=0}^{n-1} a_{k+1} = \sum_{k=k_0+1}^{n+k_0} a_{k-k_0} \quad \text{für beliebiges } k_0 \in \mathbb{Z}.$$

Beweis von Satz 2.4.6. (Vollständige Induktion)

(IA) $n = 1$: Offenbar gilt $\sum_{k=1}^1 k = 1$ und $\frac{1 \cdot (1+1)}{2} = 1$, d.h. $p(1)$ ist wahr.

(IS) $n \rightarrow n + 1$: Die Induktionsvoraussetzung $\sum_{k=1}^n k = \frac{n(n+1)}{2}$ gelte für ein $n \in \mathbb{N}$. Zum Beweis von (IS) berechnen wir mit der Induktionsvoraussetzung (IV):

$$\begin{aligned} \sum_{k=1}^{n+1} k &= \sum_{k=1}^n k + (n+1) \stackrel{(IV)}{=} \frac{n(n+1)}{2} + (n+1) \\ &= \frac{n(n+1) + 2(n+1)}{2} = \frac{(n+1)(n+2)}{2}, \end{aligned}$$

d.h. auch $p(n+1)$ ist wahr. Somit gilt die Behauptung für alle $n \in \mathbb{N}$. \square

Satz 2.4.8: Summenformel der geometrischen Reihe

Für alle $x \in \mathbb{K} \setminus \{1\}$ im Körper \mathbb{K} und alle $n \in \mathbb{N}$ gilt

$$\sum_{k=0}^{n-1} x^k = \frac{1-x^n}{1-x}.$$

Beweis. (Vollständige Induktion)

(IA) $n = 1$: Es gilt $\sum_{k=0}^0 x^k = x^0 = 1$ und $\frac{1-x^1}{1-x} = 1$, also $p(1)$ ist wahr.

(IS) $n \rightarrow n + 1$: Die zu beweisende Relation gelte für festes $n \in \mathbb{N}$ (IV). Wir berechnen dann

$$\begin{aligned} \sum_{k=0}^{(n+1)-1} x^k &= \sum_{k=0}^n x^k = \sum_{k=0}^{n-1} x^k + x^n \stackrel{(IV)}{=} \frac{1-x^n}{1-x} + x^n \\ &= \frac{1-x^n + x^n(1-x)}{1-x} = \frac{1-x^{n+1}}{1-x}, \end{aligned}$$

wie behauptet. \square

Definition 2.4.9

Wir erklären für $n \in \mathbb{N}_0$ die *Binomialkoeffizienten*

$$\binom{n}{0} := 1, \quad \binom{n}{k} := \prod_{j=1}^k \frac{n-j+1}{j} = \frac{n(n-1) \cdot \dots \cdot (n-k+1)}{1 \cdot 2 \cdot \dots \cdot k} \quad \text{für } k \in \mathbb{N}.$$

Bemerkung 2.4.10. Offenbar gilt

$$\binom{n}{k} = \begin{cases} \frac{n!}{k!(n-k)!}, & \text{falls } k \leq n \\ 0, & \text{falls } k > n \end{cases}$$

mit der bekannten *Fakultät*:

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

Insbesondere halten wir noch $\binom{n}{1} = n$ und $\binom{n}{k} = \binom{n}{n-k}$ für $k \leq n$ fest.

Hilfssatz 2.4.11

Für alle natürlichen Zahlen $k, n \in \mathbb{N}$ gilt die Relation

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}.$$

(Veranschaulichung: *Pascalsches Dreieck*).

Beweis. Für $k \geq n$ ist nach obiger Bemerkung nichts zu zeigen. Sei also $k < n$. Dann finden wir

$$\begin{aligned} \binom{n-1}{k-1} + \binom{n-1}{k} &= \frac{(n-1)!}{(k-1)!(n-k)!} + \frac{(n-1)!}{k!(n-k-1)!} \\ &= \frac{k(n-1)! + (n-1)!(n-k)}{k!(n-k)!} \\ &= \frac{n!}{k!(n-k)!} = \binom{n}{k}, \end{aligned}$$

wie behauptet. □

Satz 2.4.12: BINOMISCHER LEHRSATZ

Sei \mathbb{K} Körper und $n \in \mathbb{N}_0$ beliebig. Dann gilt für alle $a, b \in \mathbb{K}$ die Identität

$$(a+b)^n = \sum_{k=0}^n \binom{n}{k} a^k b^{n-k}. \quad (2.4.2)$$

Beweis. Mittels vollständiger Induktion \rightarrow *Ergänzungen*. □

2.5. Die komplexen Zahlen

Die Punkte der reellen Ebene lassen sich durch die geordneten Paare $z = (x, y)$ des $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ darstellen. Es ist sicher etwas überraschend, dass wir den \mathbb{R}^2 auch mit einer Körperstruktur

ausstatten können. Dieser Körper der *komplexen Zahlen* \mathbb{C} liefert eine Erweiterung der reellen Zahlengerade auf die reelle Ebene, in dem wir auch Gleichungen der Form $z^2 + 1 = 0$ lösen können.

Wir erklären Addition und Multiplikation für zwei geordnete Paare des \mathbb{R}^2 wie folgt: Für $z_1 = (x_1, y_1)$, $z_2 = (x_2, y_2)$ setzen wir

$$\begin{aligned} z_1 + z_2 &:= (x_1 + x_2, y_1 + y_2) && \text{(komplexe Addition),} \\ z_1 \cdot z_2 &:= (x_1 x_2 - y_1 y_2, x_1 y_2 + x_2 y_1) && \text{(komplexe Multiplikation).} \end{aligned} \tag{2.5.1}$$

Geometrisch entspricht die komplexe Addition der Vektoraddition in \mathbb{R}^2 . Eine geometrische Deutung der komplexen Multiplikation als *Drehstreckung* folgt erst in Kapitel 7.

Satz 2.5.1

$(\mathbb{R}^2, +, \cdot)$ den Rechenoperationen aus (2.5.1) ist ein Körper mit:

Nullelement: $0 := (0, 0)$,

Einselement: $1 := (1, 0)$,

negativem Element: zu $z = (x, y)$ setzen wir $-z := (-x, -y)$,

inverses Element: zu $z = (x, y) \neq 0$ setzen wir $z^{-1} := \left(\frac{x}{x^2 + y^2}, \frac{-y}{x^2 + y^2} \right)$.

Wir schreiben $(\mathbb{C}, +, \cdot)$ oder einfach \mathbb{C} für den *Körper der komplexen Zahlen*.

Beweis. 1. Die Axiome (A1)-(A4) folgen durch komponentenweises Anwenden der entsprechenden Axiome in \mathbb{R} ; z.B. haben wir für $z = (x, y)$:

$$z + (-z) \stackrel{(2.5.1)}{=} (x + (-x), y + (-y)) = (0, 0) = 0,$$

also (A4).

2. (M1) ist klar, da die komplexe Multiplikation symmetrisch bezüglich der Vertauschung $(x_1, y_1) \leftrightarrow (x_2, y_2)$ ist. Zum Beweis von (M2) betrachten wir $z_1 = (x_1, y_1)$, $z_2 = (x_2, y_2)$, $z_3 = (x_3, y_3)$ und berechnen

$$\begin{aligned} (z_1 z_2) z_3 &= (x_1 x_2 - y_1 y_2, x_1 y_2 + x_2 y_1) \cdot (x_3, y_3) \\ &= ((x_1 x_2 - y_1 y_2) x_3 - (x_1 y_2 + x_2 y_1) y_3, (x_1 x_2 - y_1 y_2) y_3 + (x_1 y_2 + x_2 y_1) x_3) \end{aligned}$$

und

$$\begin{aligned} z_1 (z_2 z_3) &= (x_1, y_1) \cdot (x_2 x_3 - y_2 y_3, x_2 y_3 + x_3 y_2) \\ &= (x_1 (x_2 x_3 - y_2 y_3) - y_1 (x_2 y_3 + x_3 y_2), x_1 (x_2 y_3 + x_3 y_2) + y_1 (x_2 x_3 - y_2 y_3)). \end{aligned}$$

Ein Vergleich der rechten Seiten zeigt (M2). Mit dem oben erklärten Einselement $(1, 0)$ haben wir für beliebiges $z = (x, y) \in \mathbb{C}$:

$$z \cdot 1 = (x, y) \cdot (1, 0) = (x \cdot 1 - y \cdot 0, x \cdot 0 + y \cdot 1) = (x, y) = z,$$

also (M3). Schließlich gilt auch (M4), denn wir berechnen mit der Inversen zu $z = (x, y) \neq 0$:

$$\begin{aligned} zz^{-1} &= (x, y) \cdot \left(\frac{x}{x^2 + y^2}, \frac{-y}{x^2 + y^2} \right) \\ &= \left(x \frac{x}{x^2 + y^2} - y \frac{-y}{x^2 + y^2}, x \frac{-y}{x^2 + y^2} + y \frac{x}{x^2 + y^2} \right) \\ &= (1, 0) = 1. \end{aligned}$$

3. Das Distributivgesetz lässt sich zur Übung leicht nachrechnen. \square

Bemerkung 2.5.2. Es gelten alle für beliebige Körper abgeleiteten Rechenregeln. Insbesondere sind Null- und Einselement sowie negatives und inverses Element eindeutig bestimmt. Und natürlich können wir auch Subtraktion und Division für $z_1, z_2 \in \mathbb{C}$ erklären:

$$z_1 - z_2 := z_1 + (-z_2), \quad \frac{z_1}{z_2} := z_1 \cdot z_2^{-1}, \text{ falls } z_2 \neq 0.$$

\mathbb{R} als Unterkörper von \mathbb{C}

Wir wollen uns überlegen, in welchem Sinn \mathbb{R} in \mathbb{C} enthalten ist. Dazu betrachten wir die Teilmenge

$$\mathbb{C}_{\mathbb{R}} := \{z = (x, y) \in \mathbb{C} : y = 0\}.$$

Für beliebige $z_1 = (x_1, 0), z_2 = (x_2, 0) \in \mathbb{C}_{\mathbb{R}}$ erhalten wir dann aus (2.5.1):

$$(x_1, 0) + (x_2, 0) = (x_1 + x_2, 0),$$

$$(x_1, 0) \cdot (x_2, 0) = (x_1 \cdot x_2, 0).$$

Also sind mit $z_1, z_2 \in \mathbb{C}_{\mathbb{R}}$ auch $z_1 + z_2, z_1 \cdot z_2 \in \mathbb{C}_{\mathbb{R}}$. Ferner gilt $0, 1 \in \mathbb{C}_{\mathbb{R}}$. Und mit $z = (x, 0) \in \mathbb{C}_{\mathbb{R}}$ ist auch $-z = (-x, 0) \in \mathbb{C}_{\mathbb{R}}$ und für $z \neq 0$ auch $z^{-1} = (x^{-1}, 0) \in \mathbb{C}_{\mathbb{R}}$ richtig. Also ist $\mathbb{C}_{\mathbb{R}}$ ein *Unterkörper von \mathbb{C}* , d.h. eine Teilmenge von \mathbb{C} , die bez. der Rechenoperationen in \mathbb{C} einen Körper bildet.

Da die komplexe Addition und Multiplikation von Elementen aus $\mathbb{C}_{\mathbb{R}}$ in der ersten Komponente den reellen Operationen entsprechen, können wir $\mathbb{C}_{\mathbb{R}}$ mit \mathbb{R} identifizieren durch den *Körperisomorphismus*:

$$I : \mathbb{R} \rightarrow \mathbb{C}_{\mathbb{R}}, \quad x \mapsto (x, 0).$$

In diesem Sinne gilt also

$$\mathbb{R} \subset \mathbb{C}.$$

Dementsprechend werden in der komplexen Ebene \mathbb{C} Zahlen auf der x -Achse als reelle Zahlen aufgefasst; man spricht daher von der *reellen Achse*. Die y -Achse heißt *imaginäre Achse*.

Bemerkung 2.5.3. Im oben beschriebenen Sinne ist auch \mathbb{Q} ein Unterkörper von \mathbb{R} .

Die Standardschreibweise $z = x + iy$

Die „wichtigste“ komplexe Zahl ist die *imaginäre Einheit*

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

Sie hat die Eigenschaft

$$i^2 = (0, 1) \cdot (0, 1) = (-1, 0) = -1, \quad (2.5.2)$$

d.h. $z = i$ ist Lösung der Gleichung

$$z^2 + 1 = 0.$$

Mit i berechnen wir für beliebige $z = (x, y) \in \mathbb{C}$:

$$x + iy = (x, 0) + (0, 1)(y, 0) = (x, 0) + (0, y) = (x, y) = z.$$

Die linke Seite dieser Gleichung werden wir i.F. als Schreibweise für die komplexe Zahl z verwenden, also

$$z = x + iy, \quad x, y \in \mathbb{R}.$$

Dabei heißt x *Realteil* und y *Imaginärteil* von z , und wir schreiben

$$x = \operatorname{Re} z, \quad y = \operatorname{Im} z.$$

Zwei Zahlen $z_1, z_2 \in \mathbb{C}$ stimmen genau dann überein, wenn sowohl ihre Real- als auch ihre Imaginärteile übereinstimmen.

Bemerkung 2.5.4. (a) *Faustregel:* Mit komplexen Zahlen kann man „wie mit reellen Zahlen rechnen“, wenn man (2.5.2) beachtet. Z.B. ist

$$\begin{aligned} z^2 &= (x + iy)^2 = x^2 + 2ixy + (iy)^2 \\ &= x^2 + 2ixy + i^2 y^2 \stackrel{(2.5.2)}{=} x^2 + 2ixy - y^2 \end{aligned}$$

richtig, also

$$\operatorname{Re}(z^2) = x^2 - y^2, \quad \operatorname{Im}(z^2) = 2xy.$$

(b) \mathbb{C} kann *nicht* angeordnet werden.

Denn: Gäbe es den Begriff der Positivität, so dass (O1) und (O2) aus Definition 2.3.1 erfüllt sind. Wegen $i \neq 0, 1 \neq 0$ folgte dann aus Satz 2.3.4 (f):

$$1 = 1^2 > 0, \quad -1 \stackrel{(2.5.2)}{=} i^2 > 0,$$

also ein Widerspruch zu (O1)!

□

Definition 2.5.5

Sei $z = x + iy \in \mathbb{C}$, so heißt

$$\bar{z} := x - iy$$

die *konjugiert komplexe Zahl* zu z (lies: „ z konjugiert“ oder „ z quer“). Mit

$$|z| := \sqrt{x^2 + y^2}$$

bezeichnen wir den *Betrag* von z .

Bemerkung 2.5.6. (a) Die Konjugation $z \mapsto \bar{z}$ entspricht geometrisch einer Spiegelung an der reellen Achse. Es gelten die Rechenregeln

$$\operatorname{Re} z = \frac{1}{2}(z + \bar{z}), \quad \operatorname{Im} z = \frac{1}{2i}(z - \bar{z}) \quad (2.5.3)$$

sowie

$$\overline{\bar{z}} = z, \quad \overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2, \quad \overline{z_1 \cdot z_2} = \bar{z}_1 \cdot \bar{z}_2. \quad (2.5.4)$$

(b) Der Betrag $|z|$ von $z \in \mathbb{C}$ entspricht geometrisch dem Abstand zum Nullpunkt (gemessen in der euklidischen Metrik). Es gilt

$$|z|^2 = z \cdot \bar{z}, \quad |z| = |\bar{z}|. \quad (2.5.5)$$

Und wir haben die Abschätzungen

$$|\operatorname{Re} z| \leq |z|, \quad |\operatorname{Im} z| \leq |z|, \quad |z| \leq |\operatorname{Re} z| + |\operatorname{Im} z|, \quad (2.5.6)$$

die man leicht als Übungsaufgabe nachrechnet. Die exakte Definition der Wurzel $\sqrt{x^2 + y^2}$ im Betrag von z liefern wir in Kapitel 3 nach.

2.6. Rechnen in bewerteten Körpern

In Bemerkung 2.5.4 (b) haben wir gesehen, dass \mathbb{C} nicht angeordnet werden kann. Der Betrag aus Definition 2.5.5 liefert immerhin einen Abstandsbegriff, eine sogenannte *Bewertung*:

Definition 2.6.1

Ein Körper \mathbb{K} heißt *bewerteter Körper*, wenn es einen angeordneten Körper \mathbb{K}_0 und eine Abbildung $|\cdot| : \mathbb{K} \rightarrow \mathbb{K}_0$ gibt, so dass folgende Regeln gelten:

(a) *Positivität*: Für jedes $x \in \mathbb{K}$ gilt $|x| \geq 0$ und $(|x| = 0 \Leftrightarrow x = 0)$.

(b) *Multiplikativität*: Für alle $x, y \in \mathbb{K}$ gilt $|x \cdot y| = |x| \cdot |y|$.

(c) *Dreiecksungleichung*: Für beliebige $x, y \in \mathbb{K}$ haben wir $|x + y| \leq |x| + |y|$.

Die Abbildung $|\cdot| : \mathbb{K} \rightarrow \mathbb{K}_0$ nennt man dann *Betragsfunktion*.

Satz 2.6.2

Der Betrag $|\cdot| : \mathbb{C} \rightarrow \mathbb{R}$ aus Definition 2.5.5 ist eine Betragsfunktion, \mathbb{C} ist also ein bewerteter Körper.

Beweis. Wir weisen die Eigenschaften aus Definition 2.6.1 (a)-(c) nach.

(a) $|z| \geq 0$ für alle $z \in \mathbb{C}$ ist per Definition klar. Und wir bemerken $|z| = \sqrt{x^2 + y^2} = 0$ gdw. $x^2 + y^2 = 0$ gdw. $x = y = 0$ gdw. $z = 0$.

(b) Für $z_1, z_2 \in \mathbb{C}$ berechnen wir

$$\begin{aligned} |z_1 z_2| &= \sqrt{|z_1 z_2|^2} \stackrel{(2.5.5)}{=} \sqrt{(z_1 z_2)(\overline{z_1 z_2})} \stackrel{(2.5.4)}{=} \sqrt{(z_1 z_2)(\overline{z_1} \overline{z_2})} \\ &= \sqrt{(z_1 \overline{z_1})(z_2 \overline{z_2})} \stackrel{(2.5.5)}{=} \sqrt{|z_1|^2 |z_2|^2} = |z_1| |z_2|, \end{aligned}$$

wie behauptet.

(c) Für $z_1, z_2 \in \mathbb{C}$ schätzen wir ab

$$\begin{aligned} |z_1 + z_2|^2 &\stackrel{(2.5.5)}{=} (z_1 + z_2)(\overline{z_1 + z_2}) \stackrel{(2.5.4)}{=} z_1 \overline{z_1} + (z_1 \overline{z_2} + z_2 \overline{z_1}) + z_2 \overline{z_2} \\ &\stackrel{(2.5.4), (2.5.5)}{=} |z_1|^2 + (z_1 \overline{z_2} + \overline{z_1} z_2) + |z_2|^2 \\ &\stackrel{(2.5.3)}{=} |z_1|^2 + 2 \operatorname{Re}(z_1 \overline{z_2}) + |z_2|^2 \\ &\stackrel{(b), (2.5.5), (2.5.6)}{\leq} |z_1|^2 + 2|z_1| |z_2| + |z_2|^2 = (|z_1| + |z_2|)^2, \end{aligned}$$

und die Behauptung folgt. \square

Bemerkung 2.6.3. (a) Der Begriff *Dreiecksungleichung* leitet sich aus den geometrischen Interpretationen der komplexen Addition als Vektoraddition und des komplexen Betrages als Abstand zum Ursprung ab.

(b) Für zwei Punkte $x, y \in \mathbb{K}$ stellt $|x - y|$ den *Abstand* zueinander dar; in \mathbb{C} - und damit auch in \mathbb{R} - ist das der Euklidische Abstand. Dieser Abstandsbegriff wird ab Kapitel 3 zentrale Bedeutung haben, z.B. bei der Konvergenz von Folgen.

(c) Auf dem Unterkörper \mathbb{R} von \mathbb{C} wird durch $|\cdot| : \mathbb{C} \rightarrow \mathbb{R}$ ein Betrag *induziert*. Für beliebiges $z = (x, 0) =: x \in \mathbb{R}$ gilt dabei

$$|z| = |x| = \begin{cases} x, & \text{falls } x \geq 0 \\ -x, & \text{sonst} \end{cases}$$

mit dem Betrag auf \mathbb{R} aus Definition 2.3.9. (Nachrechnen!) Allgemeiner erfüllt letztere Betragsdefinition die Eigenschaften aus Definition 2.6.1 für jeden angeordneten Körper, es gilt also der

Satz 2.6.4

Jeder angeordnete Körper \mathbb{K} wird mit der Betragsfunktion $|\cdot| : \mathbb{K} \rightarrow \mathbb{K}$ aus Definition 2.3.9 zu einem bewerteten Körper.

Beweis. → Ergänzungen □

Wir geben noch einige Rechenregeln an, die in beliebigen bewerteten Körpern und damit insbesondere in \mathbb{R} und \mathbb{C} gelten.

Satz 2.6.5: Rechnen in bewerteten Körpern

In einem bewerteten Körper \mathbb{K} mit Betragsfunktion $|\cdot| : \mathbb{K} \rightarrow \mathbb{K}_0$ gelten folgende Aussagen:

- (a) Für jedes $x \in \mathbb{K}$ ist $|-x| = |x|$ richtig.
- (b) Für jedes $x \in \mathbb{K} \setminus \{0\}$ ist $|x^{-1}| = |x|^{-1}$ erfüllt.
- (c) Für beliebige $x, y \in \mathbb{K}$ gilt $|x - y| \geq ||x| - |y||$.
- (d) Für alle $x, y \in \mathbb{K}$ mit $y \neq 0$ gilt $\left|\frac{x}{y}\right| = \frac{|x|}{|y|}$.

Beweis. Die Aussagen (b) und (d) lassen wir für die Ergänzungen.

- (a) Die Multiplikativität liefert für $x = y = 1$ zunächst $|1| = |1 \cdot 1| = |1| \cdot |1|$ bzw. $1 = |1|$. Setzen wir $x = y = -1$ ein, so folgt $1 = |1| = |(-1)(-1)| = |-1|^2$. Wegen der Positivität ist somit $|-1| > 0$ richtig. Aufgrund von

$$0 = |-1|^2 - 1^2 = (|-1| - 1)(|-1| + 1),$$

muss also $|-1| = 1$ gelten. Für beliebige $x \in \mathbb{K}$ finden wir nun

$$|-x| = |(-1)x| = |-1| |x| = 1 \cdot |x| = |x|,$$

wie behauptet.

- (c) Mit der Dreiecksungleichung berechnen wir

$$|x| = |(x - y) + y| \leq |x - y| + |y| \quad \text{bzw.} \quad |x| - |y| \leq |x - y|$$

und

$$|y| = |(y - x) + x| \stackrel{(a)}{\leq} |x - y| + |x| \quad \text{bzw.} \quad -(|x| - |y|) \leq |x - y|,$$

also (siehe Bemerkung 2.3.10)

$$|x - y| \geq \max \{ |x| - |y|, -(|x| - |y|) \} = ||x| - |y||,$$

wie behauptet. □

Teil II.

Das Grenzwertkonzept

In Teil II legen wir mit der Betrachtung des Grenzwertbegriffs die Grundlage der Analysis. In Kapitel 3 werden Zahlenfolgen diskutiert und der Konvergenzbegriff eingeführt. Kapitel 4 ist Reihen und deren Konvergenzkriterien gewidmet. In Kapitel 5 betrachten wir dann zunächst den d -dimensionalen Euklidischen Raum \mathbb{R}^d , $d \in \mathbb{N}$, bevor für Funktionen $f : \Omega \rightarrow \mathbb{R}^d$, $\Omega \subset \mathbb{R}^n$, die Begriffe Funktionsgrenzwert und Stetigkeit eingeführt und diskutiert werden.

Kapitel 3.

Zahlenfolgen

Die Arbeit mit Zahlenfolgen bildet ein zentrales Element der Analysis. Der Konvergenzbegriff wird in Abschnitt A.3.1 eingeführt. In Abschnitt A.3.2 erklären wir das Vollständigkeitsaxiom, der Begriff der CAUCHYfolge wird definiert. Nachdem wir in Abschnitt A.3.3 Rechenregeln für Grenzwerte bereitgestellt haben, lernen wir in Abschnitt A.3.4 das Intervallschachtelungsprinzip kennen. Supremum und Infimum diskutieren wir in Abschnitt A.3.5 und wenden diese in Abschnitt 3.6 auf die Menge der Häufungswerte einer Folge an. Die Einbettung von \mathbb{Q} in \mathbb{R} betrachten wir schließlich in Abschnitt 3.7 genauer.

3.1. Folgen und Konvergenz

Bereits in Abschnitt A.1.5 haben wir Folgen als Beispiele von Abbildungen erwähnt, wir erinnern uns:

Definition 3.1.1

Sei M eine beliebige Menge. Eine Abbildung $f : \mathbb{N} \rightarrow M$, $n \mapsto x_n = f(n)$ heißt *Folge* in M ; wir schreiben

$$\{x_n\}_n = \{x_n\}_{n \in \mathbb{N}} = \{x_n\}_{n=1,2,\dots} = \{x_1, x_2, \dots\} \subset M.$$

Ist \mathbb{K} ein Körper, so nennen wir eine Folge $\{x_n\}_n \subset \mathbb{K}$ auch *Zahlenfolge* in \mathbb{K} ; für $\mathbb{K} = \mathbb{Q}, \mathbb{R}, \mathbb{C}$ sprechen wir auch von *rationalen*, *reellen* bzw. *komplexen* Zahlenfolgen. Das Element x_n heißt schließlich *n-tes Glied* der Folge.

Bemerkung 3.1.2. Vorsicht, auch wenn es die Notation $\{x_n\}_n \subset M$ suggeriert, handelt es sich bei Folgen *nicht* um Mengen; die Stelle, an der das n -te Glied steht, ist nämlich wesentlich. Vielmehr verallgemeinern Folgen also den Begriff des d -Tupels (x_1, \dots, x_d) , weshalb in der Literatur inzwischen auch meist $(x_n)_n$ geschrieben wird.

Für uns interessant sind Zahlenfolgen in \mathbb{Q} , \mathbb{R} und \mathbb{C} , Folgen im \mathbb{R}^d , $d \in \mathbb{N}$, sowie *Funktionsfolgen*, d.h. M stellt eine Menge von Funktionen dar.

Beispiel 3.1.3. ▶ $\{n\}_n = \{1, 2, 3, \dots\}$ ist die Folge der natürlichen Zahlen.

- ▶ $\{a\}_n = \{a, a, a, \dots\}$ mit $a \in M$ ist eine *konstante Folge* in M .
- ▶ $\{x^n\}_n = \{x, x^2, x^3, \dots\}$ mit $x \in \mathbb{K}$ ist eine *Produktfolge*.
- ▶ $\{\frac{1}{n}\}_n = \{1, \frac{1}{2}, \frac{1}{3}, \dots\}$ ist eine wichtige rationale Folge.
- ▶ Mit $a_n \in \{0, 1, \dots, 9\}$, $n \in \mathbb{N}$, und $a_0 \in \mathbb{Z}$ erklären wir die *endlichen Dezimalbrüche*

$$x_n := \sum_{k=0}^n \frac{a_k}{10^k} = a_0, a_1 a_2 \dots a_n, \quad n \in \mathbb{N}.$$

Die Folge $\{x_n\}_n \subset \mathbb{Q}$ kann dann als *unendlicher Dezimalbruch* interpretiert werden.

Sowohl für die folgende zentrale Definition als auch die weiteren Überlegungen benötigen wir einen Abstandsbegriff, wir formulieren sie deshalb in bewerteten Körpern (vgl. Definition 2.6.3 (b)).

Vereinbarung. Wenn nichts anderes gesagt wird, ist in diesem Abschnitt \mathbb{K} immer ein *bewerteter Körper* mit Betragsfunktion $|\cdot| : \mathbb{K} \rightarrow \mathbb{K}_0$ und dem angeordneten Körper \mathbb{K}_0 . Unsere Standardbeispiele sind $\mathbb{K} = \mathbb{R}, \mathbb{C}$ mit $\mathbb{K}_0 = \mathbb{R}$ sowie $\mathbb{K} = \mathbb{Q}$ mit $\mathbb{K}_0 = \mathbb{Q}$.

Definition 3.1.4: Konvergenz und Grenzwert

Sei \mathbb{K} ein bewerteter Körper. Eine Zahlenfolge $\{x_n\}_n \subset \mathbb{K}$ heißt *konvergent (gegen $\alpha \in \mathbb{K}$)*, wenn für jedes $\varepsilon \in \mathbb{K}_0$ mit $\varepsilon > 0$ ein $N = N(\varepsilon) \in \mathbb{N}$ existiert, so dass gilt

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

Wir schreiben dann

$$\lim_{n \rightarrow \infty} x_n = \alpha \quad \text{oder} \quad x_n \rightarrow \alpha \quad (n \rightarrow \infty).$$

α heißt der *Grenzwert* oder *Limes* der Folge $\{x_n\}_n$. Schließlich nennen wir eine Folge *divergent*, wenn sie nicht konvergiert.

Bemerkung 3.1.5. ∞ ist das Symbol für den *unendlich fernen Punkt* oder einfach *Unendlich*. Wir stellen uns unendlich ferne Punkte wie folgt vor:

- ▶ $\mathbb{K} = \mathbb{R}$; auf der reellen Zahlengerade unterscheiden wir $-\infty$ („ganz links“) und $+\infty$ („ganz rechts“). Gelegentlich schreibt man auch kurz ∞ für $+\infty$. Mit den Symbolen $+\infty, -\infty$ wird \mathbb{R} zu der *erweiterten Zahlengeraden* $\overline{\mathbb{R}} := \{-\infty\} \cup \mathbb{R} \cup \{+\infty\}$, die wir

gemäß $-\infty < x < +\infty$ für alle $x \in \mathbb{R}$ anordnen können. Allerdings ist $\overline{\mathbb{R}}$ kein Körper, wie auch immer Addition und Multiplikation in $\overline{\mathbb{R}}$ erklärt werden.

- $\mathbb{K} = \mathbb{C}$; in der komplexen Ebene gibt es keine ausgezeichneten Richtungen, es gibt nur einen unendlich fernen Punkt ∞ , egal in welche Richtung man „immer weiter“ läuft. Diese Vorstellung lässt sich auch topologisch erklären, in dem man $\overline{\mathbb{C}} := \mathbb{C} \cup \{\infty\}$ als *Riemannsche Zahlenkugel* interpretiert.

Die Schreibweise $n \rightarrow \infty$ suggeriert, dass die natürlichen Zahlen „über alle Grenzen wachsen“; diesen Eindruck werden wir später mit Hilfe des Archimedischen Axioms bestätigen.

Geometrische Deutung der Konvergenz

Für $r \in \mathbb{K}_0$ mit $r > 0$ enthält die *r-Umgebung*

$$U_r(\alpha) := \{x \in \mathbb{K} : |x - \alpha| < r\}$$

von $\alpha \in \mathbb{K}$ alle Zahlen $x \in \mathbb{K}$, die von α einen Abstand kleiner r haben. Wir bemerken:

- $\mathbb{K} = \mathbb{R}$; dann ist $U_r(\alpha) = \{x \in \mathbb{R} : \alpha - r < x < \alpha + r\} =: (\alpha - r, \alpha + r)$ ein um α zentriertes Intervall der Länge $2r$.
- $\mathbb{K} = \mathbb{C}$; dann ist $U_r(\alpha) = \{z = x + iy \in \mathbb{C} : (x - \beta)^2 + (y - \gamma)^2 < r^2\}$ eine Kreisscheibe vom Radius r um den Mittelpunkt $\alpha = \beta + i\gamma$.

Eine Folge $\{x_n\}_n \subset \mathbb{K}$ konvergiert genau dann gegen $\alpha \in \mathbb{K}$, wenn in *jeder* ε -Umgebung von α alle bis auf endlich viele Glieder der Folge liegen; dabei kann die Zahl dieser Ausnahmen von ε abhängen.

Hilfssatz 3.1.6

Der Grenzwert einer konvergenten Folge $\{x_n\}_n \subset \mathbb{K}$ ist eindeutig bestimmt.

Beweis. → Ergänzungen. □

Definition 3.1.7

Sei \mathbb{K} ein bewerteter Körper.

- Eine Folge $\{x_n\}_n \subset \mathbb{K}$ heißt *Nullfolge*, falls $x_n \rightarrow 0$ ($n \rightarrow \infty$) gilt.
- Eine Folge $\{x_n\}_n \subset \mathbb{K}$ heißt *beschränkt*, wenn ein $c \in \mathbb{K}_0$ mit $c > 0$ so existiert, dass

gilt

$$|x_n| \leq c \quad \text{für alle } n \in \mathbb{N}.$$

Anderenfalls heißt die Folge *unbeschränkt*.

Bemerkung 3.1.8. ▶ Offenbar gilt $x_n \rightarrow \alpha$ ($n \rightarrow \infty$) genau dann, wenn $|x_n - \alpha| \rightarrow 0$ ($n \rightarrow \infty$) richtig ist, d.h. wenn $\{x_n - \alpha\}_n$ Nullfolge ist.

- ▶ Geometrisch heißt Beschränktheit einer Folge: Alle Folgenglieder liegen in *einer gemeinsamen* c -Umgebung von 0.
- ▶ In einem *angeordneten* Körper \mathbb{K} heißt eine Folge $\{x_n\}_n \subset \mathbb{K}$ *von oben beschränkt* (bzw. *von unten beschränkt*), wenn mit einem $c > 0$ gilt

$$x_n \leq c \quad (\text{bzw. } x_n \geq -c) \quad \text{für alle } n \in \mathbb{N}.$$

Offenbar ist eine Folge $\{x_n\}_n$ genau dann beschränkt, wenn sie von oben *und* unten beschränkt ist.

Beispiel 3.1.9. (a) Die konstante Folge $\{a\}_n \subset \mathbb{K}$ konvergiert trivialerweise gegen a .

- (b) Die Folge $\{\frac{1}{n}\}_n \subset \mathbb{Q}$ ist Nullfolge. Zu beliebigem $\varepsilon > 0$ existiert nämlich nach dem Archimedischen Axiom (O3) ein $N = N(\varepsilon) \in \mathbb{N}$ mit $N > \frac{1}{\varepsilon}$. Es folgt

$$\left| \frac{1}{n} \right| \leq \frac{1}{N} < \varepsilon \quad \text{für alle } n \geq N.$$

- (c) $\lim_{n \rightarrow \infty} \frac{n}{2^n} = 0$. Mit vollständiger Induktion zeigt man nämlich $2^n \geq n^2$ für alle $n \geq 4$ (*Übungsaufgabe*). Es folgt $\frac{1}{2^n} \leq \frac{1}{n^2}$ bzw. $\frac{n}{2^n} \leq \frac{1}{n}$ und somit für beliebiges $\varepsilon > 0$:

$$\left| \frac{n}{2^n} - 0 \right| = \frac{n}{2^n} \leq \frac{1}{n} < \varepsilon \quad \text{für } n \geq N > \max \left\{ 3, \frac{1}{\varepsilon} \right\}.$$

- (d) Die Folge $\{x_n\}_n = \{(-1)^n\}_n$ ist divergent. Sonst müsste z.B. für $\varepsilon = 1$ ein $N \in \mathbb{N}$ existieren mit $|x_n - \alpha| < 1$ für alle $n \geq N$, wobei $\alpha \in \mathbb{R}$ der Grenzwert der Folge sei. Insbesondere für x_n und x_{n+1} mit $n \geq N$ hätten wir dann den Widerspruch

$$2 = |x_n - x_{n+1}| \leq |x_n - \alpha| + |x_{n+1} - \alpha| < 2.$$

Also divergiert $\{(-1)^n\}_n$. Wegen $|(-1)^n| = 1$, $n \in \mathbb{N}$, ist die Folge aber beschränkt. Beschränktheit einer Folge impliziert also nicht deren Konvergenz (die Umkehrung gilt jedoch, siehe Hilfssatz 3.2.3).

In *angeordneten Körpern* weisen einige divergente Folgen ein „kontrolliertes“ divergentes Verhalten auf:

Definition 3.1.10: Bestimmte Divergenz

Sei \mathbb{K} ein angeordneter Körper. Eine Folge $\{x_n\}_n \subset \mathbb{K}$ heißt *bestimmt divergent gegen $+\infty$* (bzw. *gegen $-\infty$*), wenn zu jedem $c \in \mathbb{K}$ ein $N \in \mathbb{N}$ existiert, so dass

$$x_n > c \quad (\text{bzw. } x_n < c) \quad \text{für alle } n \geq N$$

richtig ist. Wir schreiben dann

$$\lim_{n \rightarrow \infty} x_n = +\infty \quad (\text{bzw. } \lim_{n \rightarrow \infty} x_n = -\infty).$$

Bemerkung 3.1.11. Offensichtlich divergiert $\{x_n\}_n$ genau dann bestimmt gegen $+\infty$, wenn $\{-x_n\}_n$ bestimmt gegen $-\infty$ divergiert.

Beispiel 3.1.12. (a) Die Folge $\{n\}_n$ divergiert bestimmt gegen $+\infty$, denn nach dem Archimedischen Axiom (O3) existiert zu jedem $c > 0$ ein $N \in \mathbb{N}$ mit $n \geq N > c$ für alle $n \geq N$. Das erklärt auch die Schreibweise $\lim_{n \rightarrow \infty}$, die präzise $\lim_{n \rightarrow +\infty}$ lauten müsste.

(b) Zu $x \in \mathbb{R}$ betrachten wir die Produktfolge $\{x^n\}_n$. Das Konvergenzverhalten hängt von x ab, wir unterscheiden fünf Fälle.

- Für $|x| < 1$ gilt $\lim_{n \rightarrow \infty} x^n = 0$. Nach Folgerung 2.4.5 (b) gibt es nämlich zu jedem $\varepsilon > 0$ ein $N = N(\varepsilon) \in \mathbb{N}$ mit $|x|^N < \varepsilon$, so dass folgt

$$|x^n - 0| = |x|^n \leq |x|^N < \varepsilon \quad \text{für alle } n \geq N.$$

- Für $x = 1$ haben wir die konstante Folge $\{1^n\}_n = \{1\}_n$ mit $\lim_{n \rightarrow \infty} 1^n = 1$.
- Für $x = -1$ haben wir die divergente Folge $\{(-1)^n\}_n$, siehe Beispiel 3.1.9 (d).
- Für $x > 1$ ist $\{x^n\}_n$ bestimmt divergent gegen $+\infty$ nach Folgerung 2.4.5 (a).
- Für $x < -1$ ist $\{x^n\}_n$ divergent (wiederum nach Folgerung 2.4.5 (a)), aber nicht bestimmt divergent. Ist nämlich $c > 0$ und gilt $x^n > c$ für ein $n \in \mathbb{N}$, so ist n gerade und für den Nachfolger ergibt sich

$$x^{n+1} = -|x|^{n+1} < -|x|^n < -c < c.$$

Satz 3.1.13

Es sei \mathbb{K} ein angeordneter Körper.

- (a) Sei $\{x_n\}_n \subset \mathbb{K}$ bestimmt divergent gegen $+\infty$ oder $-\infty$. Dann gilt $x_n \neq 0$ für alle $n \geq n_0$ mit einem $n_0 \in \mathbb{N}$, und $\{x_n^{-1}\}_{n \geq n_0}$ ist eine Nullfolge.

- (b) Sei $\{x_n\}_n \subset \mathbb{K}$ Nullfolge mit $x_n > 0$ (bzw. $x_n < 0$) für alle $n \geq n_0$ mit einem $n_0 \in \mathbb{N}$.
Dann ist $\{x_n^{-1}\}_{n \geq n_0}$ bestimmt divergent gegen $+\infty$ (bzw. gegen $-\infty$).

Beweis. → Ergänzungen. □

3.2. CAUCHYfolgen und Vollständigkeit

Wir beginnen mit einem weiteren zentralen Begriff für Folgen:

Definition 3.2.1: CAUCHYfolgen

Sei \mathbb{K} ein bewerteter Körper mit Betragsfunktion $|\cdot| : \mathbb{K} \rightarrow \mathbb{K}_0$ und dem angeordneten Körper \mathbb{K}_0 . Eine Folge $\{x_n\}_n \subset \mathbb{K}$ heißt **CAUCHYfolge**, falls zu jedem $\varepsilon \in \mathbb{K}_0$ mit $\varepsilon > 0$ ein $N = N(\varepsilon) \in \mathbb{N}$ existiert mit

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

Noch einmal erinnern wir an unsere zentralen Beispiele $\mathbb{K} = \mathbb{R}, \mathbb{C}$ (mit $\mathbb{K}_0 = \mathbb{R}$) sowie $\mathbb{K} = \mathbb{Q}$ (mit $\mathbb{K}_0 = \mathbb{Q}$).

- Bemerkung 3.2.2.** ▶ Geometrisch bedeutet diese Definition, dass die „Streuung“ der Folgenglieder einer CAUCHYfolge, also deren Abstand zueinander, mit hinreichend großem Index beliebig klein gemacht, also unter jede positive Schranke gebracht werden kann.
- ▶ Wir weisen auf einen häufigen Fehler hin: Es genügt *nicht*, dass der Abstand $|x_{n+1} - x_n|$ zweier aufeinander folgender Folgenglieder mit hinreichend großem n beliebig klein wird! Beispiel:

$$\left\{ \underbrace{1, 2}_{=1}, \underbrace{2\frac{1}{2}, 3}_{=\frac{1}{2}}, \underbrace{3\frac{1}{3}, 3\frac{2}{3}}_{=\frac{1}{3}}, \underbrace{4, 4\frac{1}{4}, 4\frac{2}{4}, 4\frac{3}{4}}_{=\frac{1}{4}}, 5, \dots \right\}.$$

Offenbar ist die Folge $\{x_n\}_n$ unbeschränkt und somit nach dem folgenden Hilfssatz 3.2.3 keine CAUCHYfolge. Der Abstand $|x_{n+1} - x_n|$ geht aber gegen 0.

Hilfssatz 3.2.3

Für Zahlenfolgen $\{x_n\}_n \subset \mathbb{K}$ in einem bewerteten Körper \mathbb{K} gelten die Implikationen

$$\{x_n\}_n \text{ ist konvergent} \Rightarrow \{x_n\}_n \text{ ist CAUCHYfolge} \Rightarrow \{x_n\}_n \text{ ist beschränkt.}$$

Beweis. Sei $|\cdot| : \mathbb{K} \rightarrow \mathbb{K}_0$ mit dem angeordneten Körper \mathbb{K}_0 die Betragsfunktion in \mathbb{K} .

1. Sei $\{x_n\}_n$ konvergent mit Grenzwert $\alpha \in \mathbb{K}$. Dann existiert zu jedem $\varepsilon \in \mathbb{K}_0$ mit $\varepsilon > 0$ ein $N = N(\varepsilon) \in \mathbb{N}$ mit $|x_n - \alpha| < \frac{\varepsilon}{2}$ für alle $n \geq N$. Folglich haben wir

$$|x_n - x_m| = |(x_n - \alpha) + (\alpha - x_m)| \leq |x_n - \alpha| + |x_m - \alpha| < \varepsilon \quad \text{für alle } m, n \geq N(\varepsilon),$$

d.h. $\{x_n\}_n$ ist CAUCHYfolge.

2. Sei $\{x_n\}_n$ CAUCHYfolge. Gemäß Definition gibt es speziell zu $\varepsilon = 1 \in \mathbb{K}_0$ ein $N \in \mathbb{N}$ mit $|x_n - x_m| < 1$ für alle $n, m \geq N$. Damit folgt insbesondere für $m = N$:

$$|x_n| = |(x_n - x_N) + x_N| \leq |x_n - x_N| + |x_N| < 1 + |x_N| \quad \text{für alle } n \geq N.$$

Setzen wir $c := \max\{|x_1|, \dots, |x_{N-1}|, 1 + |x_N|\}$, so folgt die Behauptung. \square

Wir kommen nun zur

Definition 3.2.4

Ein bewerteter Körper \mathbb{K} heißt *vollständig bewertet*, wenn folgendes *Vollständigkeitsaxiom* erfüllt ist:

(V) Jede CAUCHYfolge in \mathbb{K} konvergiert.

Ist \mathbb{K} zusätzlich (archimedisch) angeordnet, so heißt \mathbb{K} dann *vollständig (archimedisch) angeordnet*.

Satz 3.2.5: CAUCHYSches Konvergenzkriterium

In einem vollständig bewerteten Körper \mathbb{K} ist eine Folge $\{x_n\}_n \subset \mathbb{K}$ genau dann konvergent, wenn $\{x_n\}_n$ eine CAUCHYfolge ist.

Beweis. Sofort aus Hilfssatz 3.2.3 und Definition 3.2.4. \square

Bemerkung 3.2.6. Um Konvergenz einer Folge in einem vollständig bewerteten Körper zu beweisen, genügt es also, deren CAUCHYfolgen-Eigenschaft nachzurechnen. Der Vorteil ist, dass keine Kenntnis (oder Raten) des Grenzwertes notwendig ist; dieser muss dann im Nachhinein bestimmt werden. Wir werden dieses Prinzip auch immer wieder theoretisch ausnutzen.

In Definition 2.1.4 hatten wir \mathbb{R} als vollständig archimedisch angeordneten Körper definiert (zu \mathbb{C} kommen wir gleich). Hingegen ist der angeordnete Körper \mathbb{Q} *nicht* vollständig:

Beispiel 3.2.7. Wir zeigen später in Satz 3.7.3, dass sich jede reelle Zahl $x \in \mathbb{R}$ als unendlicher Dezimalbruch $\{x_n\}_n \subset \mathbb{Q}$ schreiben lässt. Z.B. für $x = \sqrt{2}$ haben wir also

$$x = \lim_{n \rightarrow \infty} x_n, \quad x_n := \sum_{k=0}^n \frac{a_k}{10^k} = a_0, a_1 a_2 \dots a_n, \quad n \in \mathbb{N}, \quad (3.2.1)$$

für gewisse $a_k \in \{0, \dots, 9\}$, $k \in \mathbb{N}$, und $a_0 \in \mathbb{Z}$; siehe Beispiel 3.1.3. Für beliebiges $N \in \mathbb{N}$ und für $m, n \geq N$ (o.B.d.A. gelte $n > m$) schätzen wir ab

$$\begin{aligned} |x_n - x_m| &= \left| \sum_{k=0}^n \frac{a_k}{10^k} - \sum_{k=0}^m \frac{a_k}{10^k} \right| = \left| \sum_{k=m+1}^n \frac{a_k}{10^k} \right| \\ &\stackrel{\text{Def. 2.6.1 (c)}}{\leq} \sum_{k=m+1}^n \left| \frac{a_k}{10^k} \right| \leq \sum_{k=m+1}^n \frac{9}{10^k} = 9 \sum_{k=m+1}^n \left(\frac{1}{10} \right)^k \\ &\stackrel{k \rightarrow k-m-1}{=} \frac{9}{10} \sum_{k=0}^{n-m-1} \left(\frac{1}{10} \right)^{k+m} = \frac{9}{10} \left(\frac{1}{10} \right)^m \sum_{k=0}^{n-m-1} \left(\frac{1}{10} \right)^k \\ &\stackrel{\text{Satz 2.4.7}}{=} \frac{9}{10} \left(\frac{1}{10} \right)^m \frac{1 - \left(\frac{1}{10} \right)^{n-m}}{1 - \frac{1}{10}} \leq \frac{9}{10} \left(\frac{1}{10} \right)^m \frac{1}{1 - \frac{1}{10}} \\ &= \left(\frac{1}{10} \right)^m \leq \left(\frac{1}{10} \right)^N. \end{aligned}$$

Wenden wir nun Folgerung 2.4.5 (b) im archimedisch angeordneten Körper \mathbb{Q} an (Hilfssatz 2.3.7) mit $b = \frac{1}{10} < 1$ und $\delta = \varepsilon > 0$ für beliebig gewähltes positives $\varepsilon \in \mathbb{Q}$, so existiert also ein $N = N(\varepsilon) \in \mathbb{N}$ mit

$$|x_n - x_m| \leq \left(\frac{1}{10} \right)^N < \varepsilon \quad \text{für alle } m, n \geq N(\varepsilon). \quad (3.2.2)$$

Folglich ist $\{x_n\}_n$ eine CAUCHYfolge. Wäre \mathbb{Q} vollständig, müsste auch $\sqrt{2} = x = \lim_{n \rightarrow \infty} x_n \in \mathbb{Q}$ gelten, im Widerspruch zu $\sqrt{2} \in \mathbb{R} \setminus \mathbb{Q}$.

Abschließend zeigen wir, dass \mathbb{C} vollständig bewertet ist. Dazu benötigen wir noch den

Hilfssatz 3.2.8

Eine Folge $\{z_n\}_n \subset \mathbb{C}$ ist genau dann konvergent (bzw. CAUCHYfolge, Nullfolge, beschränkt), wenn die reellen Folgen $\{\operatorname{Re}(z_n)\}_n$ und $\{\operatorname{Im}(z_n)\}_n$ konvergent (bzw. CAUCHYfolgen, Nullfolgen, beschränkt) sind. Für konvergente Folgen $\{z_n\}_n$ gilt

$$\lim_{n \rightarrow \infty} z_n = \lim_{n \rightarrow \infty} \operatorname{Re}(z_n) + i \lim_{n \rightarrow \infty} \operatorname{Im}(z_n).$$

Beweis. Die Aussagen ergeben sich sofort aus den Ungleichungen (2.5.6). □

Satz 3.2.9: Vollständigkeit von \mathbb{C}

Eine Folge $\{z_n\}_n \subset \mathbb{C}$ ist genau dann konvergent, wenn sie CAUCHYfolge ist. Insbesondere ist \mathbb{C} ein vollständig bewerteter Körper.

Beweis. Wir argumentieren mit dem CAUCHYSchen Konvergenzkriterium in \mathbb{R} :

$$\begin{array}{lll}
 \{z_n\}_n \text{ ist konvergent} & \stackrel{\text{Hilfssatz 3.2.8}}{\Leftrightarrow} & \{\operatorname{Re}(z_n)\}_n, \{\operatorname{Im}(z_n)\}_n \subset \mathbb{R} \text{ sind konvergent} \\
 & \stackrel{\text{Satz 3.2.5}}{\Leftrightarrow} & \{\operatorname{Re}(z_n)\}_n, \{\operatorname{Im}(z_n)\}_n \subset \mathbb{R} \text{ sind CAUCHYfolgen} \\
 & \stackrel{\text{Hilfssatz 3.2.8}}{\Leftrightarrow} & \{z_n\}_n \text{ ist CAUCHYfolge,}
 \end{array}$$

wie behauptet. □

3.3. Rechenregeln für Folgengrenzwerte

Die Grenzwerte vieler Folgen lassen sich auf bekannte Grenzwerte einfacherer Folgen zurückführen. Dazu geben wir verschiedene Rechenregeln an und beginnen mit dem:

Satz 3.3.1

Seien $\{x_n\}_n, \{y_n\}_n \subset \mathbb{K}$ zwei konvergente Folgen in einem bewerteten Körper \mathbb{K} mit $x_n \rightarrow \alpha, y_n \rightarrow \beta$ ($n \rightarrow \infty$). Dann gelten:

- (a) Es konvergieren auch $\{x_n + y_n\}_n$ und $\{x_n y_n\}_n$ mit

$$\lim_{n \rightarrow \infty} (x_n + y_n) = \alpha + \beta, \quad \lim_{n \rightarrow \infty} (x_n y_n) = \alpha \beta.$$

- (b) Falls zusätzlich $\beta \neq 0$ und $y_n \neq 0$ für alle $n \in \mathbb{N}$ richtig ist, so konvergiert auch $\{\frac{x_n}{y_n}\}_n$ mit

$$\lim_{n \rightarrow \infty} \frac{x_n}{y_n} = \frac{\alpha}{\beta}.$$

- (c) Es konvergiert auch $\{|x_n|\}_n$ mit

$$\lim_{n \rightarrow \infty} |x_n| = |\alpha|.$$

Wir notieren noch die folgende direkte Konsequenz aus Satz 3.3.1 (a):

Folgerung 3.3.2

Ist \mathbb{K} bewerteter Körper, konvergieren $\{x_n\}_n, \{y_n\}_n \subset \mathbb{K}$ und sind $a, b \in \mathbb{K}$ beliebig, so konvergiert auch $\{ax_n + by_n\}_n$ mit

$$\lim_{n \rightarrow \infty} (ax_n + by_n) = a \lim_{n \rightarrow \infty} x_n + b \lim_{n \rightarrow \infty} y_n.$$

Beweis von Satz 3.3.1. (a) Wir nutzen die üblichen Notationen $|\cdot|$ und \mathbb{K}_0 . Zu beliebig vorgegebenem $\varepsilon \in \mathbb{K}_0$ mit $\varepsilon > 0$ existiert ein $N = N(\varepsilon) \in \mathbb{N}$, so dass

$$|x_n - \alpha| < \varepsilon, |y_n - \beta| < \varepsilon \quad \text{für alle } n \geq N.$$

Es folgt

$$|(x_n + y_n) - (\alpha + \beta)| \leq |x_n - \alpha| + |y_n - \beta| < 2\varepsilon \quad \text{für alle } n \geq N.$$

Da $\varepsilon > 0$ beliebig war, folgt $x_n + y_n \rightarrow \alpha + \beta$ ($n \rightarrow \infty$).

Nach Hilfssatz 3.2.3 existiert ferner ein $c > 0$ mit $|x_n| \leq c$ für alle $n \in \mathbb{N}$. Somit folgt noch

$$\begin{aligned} |x_n y_n - \alpha \beta| &= |x_n(y_n - \beta) + \beta(x_n - \alpha)| \\ &\leq |x_n| |y_n - \beta| + |\beta| |x_n - \alpha| \\ &< (c + |\beta|)\varepsilon \quad \text{für alle } n \geq N, \end{aligned}$$

also $x_n y_n \rightarrow \alpha \beta$ ($n \rightarrow \infty$).

(b) Wegen $|\beta| > 0$ gibt es ein $N \in \mathbb{N}$ mit $|y_n - \beta| < \frac{|\beta|}{2}$ für alle $n \geq N$. Wir folgern mit Satz 2.6.5

$$|y_n| = |\beta + (y_n - \beta)| \geq |\beta| - |y_n - \beta| > \frac{|\beta|}{2} > 0$$

und somit

$$\left| \frac{x_n}{y_n} - \frac{\alpha}{\beta} \right| = \frac{1}{|y_n| |\beta|} |\beta x_n - \alpha y_n| \leq \frac{2}{|\beta|^2} |\beta x_n - \alpha y_n| \quad \text{für } n \geq N.$$

Nach (a) bzw. Folgerung 3.3.2 konvergiert $\beta x_n - \alpha y_n \rightarrow \beta \alpha - \alpha \beta = 0$ ($n \rightarrow \infty$), also ist $\left\{ \frac{x_n}{y_n} - \frac{\alpha}{\beta} \right\}_n$ Nullfolge, wie behauptet.

(c) Aus Satz 2.6.5 (c) folgt sofort

$$||x_n| - |\alpha|| \leq |x_n - \alpha| \rightarrow 0 \quad (n \rightarrow \infty),$$

wie behauptet. □

Beispiel 3.3.3. Wir zeigen $\lim_{n \rightarrow \infty} \frac{5n^3 + 8n^2}{n^3 - 4} = 5$. Dazu kürzen wir

$$\frac{5n^3 + 8n^2}{n^3 - 4} = \frac{n^3(5 + 8\frac{1}{n})}{n^3(1 - \frac{4}{n^3})} = \frac{5 + 8\frac{1}{n}}{1 - \frac{4}{n^3}}.$$

Da $\{\frac{1}{n}\}_n$ Nullfolge ist, konvergieren nach Satz 3.3.1 (a) und Folgerung 3.3.2 auch $\{5 + 8\frac{1}{n}\}_n$ und $\{1 - \frac{4}{n^3}\}_n$, nämlich gegen 5 bzw. 1. Satz 3.3.1 (b) liefert dann

$$\lim_{n \rightarrow \infty} \frac{5n^3 + 8n^2}{n^3 - 4} = \frac{\lim_{n \rightarrow \infty} (5 + 8\frac{1}{n})}{\lim_{n \rightarrow \infty} (1 - \frac{4}{n^3})} = \frac{5}{1} = 5.$$

Bemerkung 3.3.4. Im bewerteten Körper \mathbb{C} folgern wir noch: Ist $\{z_n\}_n \subset \mathbb{C}$ konvergent, so konvergiert auch $\{\overline{z_n}\}_n$ und es gilt

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

Denn: Schreiben wir $z_n = x_n + iy_n$ und $\lim_{n \rightarrow \infty} z_n = \alpha = \beta + i\gamma$, so gilt nach Hilfssatz 3.2.8 $x_n \rightarrow \beta$, $y_n \rightarrow \gamma$ ($n \rightarrow \infty$). Folgerung 3.3.2 entnehmen wir dann

$$\overline{z_n} = x_n - iy_n \rightarrow \beta - i\gamma = \overline{\alpha} \quad (n \rightarrow \infty),$$

wie behauptet. □

Ist der Körper angeordnet, so haben wir noch zusätzlich den

Satz 3.3.5

Seien \mathbb{K} angeordnet und $\{x_n\}_n, \{y_n\}_n \subset \mathbb{K}$ konvergent mit $x_n \rightarrow \alpha$, $y_n \rightarrow \beta$ ($n \rightarrow \infty$). Dann folgen:

- (a) Gilt $x_n \leq y_n$ für alle $n \in \mathbb{N}$, so folgt $\alpha \leq \beta$.
- (b) *Sandwich-Lemma*. Gilt $\alpha = \beta$ und ist $\{z_n\}_n \subset \mathbb{K}$ eine weitere Folge mit $x_n \leq z_n \leq y_n$ für alle $n \in \mathbb{N}$, so konvergiert auch $\{z_n\}_n$ gegen α .

Beweis. → Übungsaufgabe. □

3.4. Das Intervallschachtelungsprinzip in \mathbb{R}

In diesem Abschnitt benötigen wir neben der Vollständigkeit die archimedische Anordnung des betrachteten Körpers \mathbb{K} , weshalb wir uns auf $\mathbb{K} = \mathbb{R}$ beschränken können und werden; siehe Bemerkung 2.1.5 (c).

Das Vollständigkeitsaxiom lässt sich zwar kompakt formulieren, ist aber etwas unanschaulich. Genau umgekehrt verhält es sich mit dem unten stehenden praktischen *Intervallschachtelungs-Prinzip*, welches aus der Vollständigkeit folgt. Zunächst benötigen wir die

Definition 3.4.1: Intervalle reeller Zahlen

Zu $a, b \in \mathbb{R}$ erklären wir folgende Mengen:

- ▶ *offenes Intervall:* $(a, b) := \{x \in \mathbb{R} : a < x < b\}$.
- ▶ *abgeschlossenes Intervall:* $[a, b] := \{x \in \mathbb{R} : a \leq x \leq b\}$.
- ▶ *halboffene Intervalle:*

$$(a, b] := \{x \in \mathbb{R} : a < x \leq b\}, \quad [a, b) := \{x \in \mathbb{R} : a \leq x < b\}.$$

Die Punkte a und b heißen *Endpunkte des Intervalls* oder *Intervallgrenzen*. Gilt $a \leq b$, so setzen wir

$$|I| = \text{diam}(I) := b - a \geq 0$$

für die *Länge* oder den *Durchmesser* von I .

Bemerkung 3.4.2. (a) Für $a > b$ sind alle angegebenen Intervalle leer. Für $a = b$ sind offene und halboffene Intervalle leer, während das abgeschlossene Intervall $[a, a] = \{a\}$ nur den Punkt a enthält.

(b) Es kann auch $a = -\infty$ und $b = +\infty$ gewählt werden, wenn der jeweilige Endpunkt nicht zum Intervall gehört, z.B. $[a, +\infty) = \{x \in \mathbb{R} : a \leq x\}$.

(c) Ist I ein Intervall mit den Endpunkten $a \leq b$, so gilt offenbar für beliebige $x, x' \in I$ dann $|x - x'| \leq |I|$.

Satz 3.4.3: Intervallschachtelungsprinzip

Es sei $I_1 \supset I_2 \supset \dots \supset I_n \supset I_{n+1} \supset \dots$ eine absteigende Folge von nichtleeren abgeschlossenen Intervallen in \mathbb{R} , genannt *Intervallschachtelung*. Es gelte

$$\lim_{n \rightarrow \infty} |I_n| = 0. \quad (3.4.1)$$

Dann gibt es genau eine reelle Zahl x mit $x \in I_n$ für alle $n \in \mathbb{N}$, d.h. $\{x\} = \bigcap_{n \in \mathbb{N}} I_n$.

Beweis. ▶ *Existenz.* Wir schreiben $I_n = [a_n, b_n]$. Formel (3.4.1) besagt, dass zu jedem $\varepsilon > 0$ ein $N = N(\varepsilon) \in \mathbb{N}$ existiert mit $0 \leq |I_n| < \varepsilon$ für alle $n \geq N(\varepsilon)$. Sind nun $m, n \geq N$, so folgt $a_m, a_n \in I_N$ und somit gemäß Bemerkung 3.4.2 (c):

$$|a_n - a_m| \leq |I_N| < \varepsilon \quad \text{für alle } m, n \geq N(\varepsilon),$$

d.h. $\{a_n\}_n$ ist eine CAUCHYfolge. Nach dem Vollständigkeitsaxiom existiert daher ein Punkt $x \in \mathbb{R}$ mit $\lim_{n \rightarrow \infty} a_n = x$.

Nun gilt $a_m \leq a_n \leq b_m$ für beliebiges $m \in \mathbb{N}$ und alle $n \geq m$. Gemäß Satz 3.3.5 (a) liefert der Grenzübergang $n \rightarrow \infty$ nun $a_m \leq x \leq b_m$ bzw. $x \in I_m$ für alle $m \in \mathbb{N}$.

- *Eindeutigkeit.* Gäbe es ein weiteres Element $x' \in \bigcap_{n \in \mathbb{N}} I_n \setminus \{x\}$, so hätten wir für beliebiges $\varepsilon > 0$ und das oben bestimmte $N = N(\varepsilon) \in \mathbb{N}$:

$$|x - x'| \leq |I_N| < \varepsilon.$$

Also folgte $x = x'$, Widerspruch! Folglich ist $x \in \bigcap_{n \in \mathbb{N}} I_n$ eindeutig bestimmt. \square

Bemerkung 3.4.4. (a) Die Aussage scheint offensichtlich: Eine Folge von ineinander liegenden, abgeschlossenen Intervallen, deren Durchmesser gegen 0 geht, zieht sich auf einen Punkt zusammen. Jedoch wird die Aussage in \mathbb{Q} falsch, da der gemeinsame Punkt dann keine rationale Zahl sein muss.

- (b) Die Konstruktion zeigt, dass für eine Intervallschachtelung mit $I_n = [a_n, b_n]$ und $\{x\} = \bigcap_{n \in \mathbb{N}} I_n$ gilt

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

Für eine erste wichtige Folgerung des Intervallschachtelungsprinzips benötigen wir noch den folgenden zentralen Begriff:

Definition 3.4.5

Seien M eine Menge und $\{x_n\}_{n \in \mathbb{N}} \subset M$ eine Folge. Weiter seien unendlich viele natürliche Zahlen $1 \leq n_1 < n_2 < n_3 < \dots$ gewählt (also $\{n_k\}_k \subset \mathbb{N}$ mit $n_k < n_{k+1}$ für alle $k \in \mathbb{N}$). Dann heißt

$$\{x_{n_k}\}_{k \in \mathbb{N}} = \{x_{n_1}, x_{n_2}, x_{n_3}, \dots\}$$

Teilfolge von $\{x_n\}_n$; wir schreiben oft zur Verdeutlichung $\{x_{n_k}\}_k \subset \{x_n\}_n$.

Satz 3.4.6: BOLZANO-WEIERSTRASS

Jede beschränkte Folge $\{x_n\}_n \subset \mathbb{R}$ besitzt eine konvergente Teilfolge.

Beweis. Da $\{x_n\}_n$ beschränkt ist, existiert ein $c > 0$ mit $-c \leq x_n \leq c$ für alle $n \in \mathbb{N}$. Wir konstruieren eine Intervallschachtelung $I_1 \supset I_2 \supset \dots$ mit folgenden Eigenschaften:

- I_k enthält unendlich viele Glieder der Folge $\{x_n\}_n$ für alle $k \in \mathbb{N}$,

- $|I_k| = 2c \cdot \left(\frac{1}{2}\right)^{k-1}$ für alle $k \in \mathbb{N}$; gemäß Beispiel 3.1.12 (b) folgt dann $|I_k| \rightarrow 0$ ($k \rightarrow \infty$).

Wir starten dazu mit $I_1 := [-c, c]$ und definieren im k -ten Schritt I_{k+1} wie folgt: Ist $I_k = [a_k, b_k]$ wie oben gegeben, so setzen wir $y_k := \frac{1}{2}(a_k + b_k)$ und erklären

$$I_{k+1} = [a_{k+1}, b_{k+1}] := \begin{cases} [a_k, y_k], & \text{falls } [a_k, y_k] \text{ unendlich viele Glieder} \\ & \text{von } \{x_n\}_n \text{ enthält} \\ [y_k, b_k], & \text{sonst} \end{cases}.$$

Damit hat offenbar auch I_{k+1} die gewünschten Eigenschaften. Nach dem Prinzip der vollständigen Induktion haben wir so unsere Intervallschachtelung $\{I_k\}_k$ gefunden.

Wir definieren nun eine Teilfolge $\{x_{n_k}\}_k$ mit $x_{n_k} \in I_k$ für alle $k \in \mathbb{N}$, und zwar wiederum induktiv:

- Für $k = 1$ setzen wir $n_1 = 1$, also $x_{n_1} = x_1 \in I_1$.
- Ist $x_{n_k} \in I_k$ für ein $k \in \mathbb{N}$, so existiert per Konstruktion ein $n_{k+1} > n_k$ mit $x_{n_{k+1}} \in I_{k+1}$ (da in I_{k+1} wieder *unendlich viele* Glieder von $\{x_n\}_n$ liegen).

Wir haben also $a_k \leq x_{n_k} \leq b_k$ für alle $k \in \mathbb{N}$. Gemäß Satz 3.4.3 und Bemerkung 3.4.4 (b) existiert ein $x \in \mathbb{R}$ mit

$$\lim_{k \rightarrow \infty} a_k = x = \lim_{k \rightarrow \infty} b_k.$$

Wegen des Sandwich-Lemmas, Satz 3.3.5 (b), konvergiert also auch $\{x_{n_k}\}_k$ gegen x , wie behauptet. \square

Wir geben noch ein weiteres hilfreiches Konvergenzkriterium an; dafür benötigen wir die

Definition 3.4.7

Eine Folge $\{x_n\}_n \subset \mathbb{R}$ heißt

- (i) *monoton wachsend* (bzw. *fallend*), falls $x_n \leq x_{n+1}$ (bzw. $x_n \geq x_{n+1}$) für alle $n \in \mathbb{N}$ gilt.
- (ii) *streng monoton wachsend* (bzw. *fallend*), falls $x_n < x_{n+1}$ (bzw. $x_n > x_{n+1}$) für alle $n \in \mathbb{N}$ richtig ist.

Bemerkung 3.4.8. Die Sprechweise ist leider nicht eindeutig. In der Literatur wird häufig z.B. für eine monoton wachsende Folge auch „schwach monoton wachsende“ oder „monoton nicht fallende“ Folge geschrieben.

Satz 3.4.9: Monotone Konvergenz

Jede beschränkte monotone Folge $\{x_n\}_n \subset \mathbb{R}$ ist konvergent.

Beweis. Sei $\{x_n\}_n$ monoton fallend. Da $\{x_n\}_n$ beschränkt ist, existiert nach dem Satz von BOLZANO-WEIERSTRASS eine konvergente Teilfolge $\{x_{n_k}\}_k$. Aus der Relation $x_{n_k} \geq x_{n_l}$ für alle $k \leq l$ erhalten wir nach Grenzübergang $l \rightarrow \infty$ die Ungleichung

$$x_{n_k} \geq x := \lim_{l \rightarrow \infty} x_{n_l} \quad \text{für alle } k \in \mathbb{N}.$$

Zu beliebigem $\varepsilon > 0$ gibt es nun ein $k_0 = k_0(\varepsilon) \in \mathbb{N}$ mit $|x_{n_k} - x| < \varepsilon$ für alle $k \geq k_0(\varepsilon)$. Wir setzen $N = N(\varepsilon) := n_{k_0}(\varepsilon)$. Für jedes $n \geq N$ existiert dann ein (eindeutiges) $k \geq k_0$ mit $n_k \leq n < n_{k+1}$. Die Monotonie liefert nun

$$x_{n_k} \geq x_n \geq x_{n_{k+1}} \geq x,$$

bzw. $x_{n_k} - x \geq x_n - x \geq 0$. Zusammen mit der Konvergenz der Teilfolge $\{x_{n_k}\}_k$ folgt

$$|x_n - x| \leq |x_{n_k} - x| < \varepsilon \quad \text{für alle } n \geq N(\varepsilon),$$

wie behauptet. Der Fall einer monoton wachsenden Folge $\{x_n\}_n$ folgt analog oder durch Übergang zur monoton fallenden Folge $\{-x_n\}_n$. \square

Bemerkung 3.4.10. Da eine monoton fallende (bzw. wachsende) Folge $\{x_n\}_n$ trivial nach oben (bzw. unten) beschränkt ist - nämlich durch x_1 -, genügt es für die Anwendung von Satz 3.4.9, die *richtige* einseitige Beschränkung der betrachteten Folge zu zeigen.

Beispiel 3.4.11. Wir betrachten eine *rekursiv definierte Folge*, nämlich

$$a_1 = 1, \quad a_{n+1} = \frac{3 + 5a_n}{20}, \quad n \in \mathbb{N}.$$

Man zeigt zunächst leicht mit vollständiger Induktion, dass $a_n > \frac{1}{5}$ für alle $n \in \mathbb{N}$ gilt, d.h. die Folge ist nach unten beschränkt. Es folgt dann

$$a_n - a_{n+1} = a_n - \frac{3 + 5a_n}{20} = \frac{15a_n - 3}{20} > \frac{3 - 3}{20} = 0,$$

d.h. $\{a_n\}_n$ ist streng monoton fallend (und somit beschränkt gemäß $1 = a_1 \geq a_n > \frac{1}{5}$ für alle $n \in \mathbb{N}$). Gemäß Satz 3.4.9 konvergiert die Folge; für den Grenzwert $a = \lim_{n \rightarrow \infty} a_n$ folgt:

$$a = \lim_{n \rightarrow \infty} a_{n+1} = \frac{3 + 5 \lim_{n \rightarrow \infty} a_n}{20} = \frac{3 + 5a}{20} \quad \text{bzw.} \quad a = \frac{1}{5}.$$

Abschließend halten wir die Beobachtung fest, dass uns die formale Festlegung von $\mathbb{R} \supset \mathbb{Q}$ als vollständiger Körper erlaubt, u.a. die in \mathbb{Q} nicht lösbare Gleichung $x^2 = 2$ zu lösen:

Satz 3.4.12

Sei $c > 0$ eine gegebene positive reelle Zahl. Dann besitzt die Gleichung $x^s = c$ für jedes $s \in \mathbb{N}$ genau eine positive Lösung $x \in \mathbb{R}$.

Beweis. Mit Intervallschachtelung \rightarrow Ergänzungen. □

Definition 3.4.13

- Die in Satz 3.4.12 konstruierte eindeutige Lösung $x > 0$ der Gleichung $x^s = c$ heißt *s-te Wurzel* von $c > 0$, und wir schreiben $x = \sqrt[s]{c}$.
- Ist $q = \frac{r}{s} \in \mathbb{Q}$ ($r \in \mathbb{Z}$, $s \in \mathbb{N}$) beliebig, so setzen wir für die *q-te Potenz* von $c > 0$:

$$c^q = c^{\frac{r}{s}} := (\sqrt[s]{c})^r.$$

Bemerkung 3.4.14. Für alle $x, y > 0$ und $p, q \in \mathbb{Q}$ gelten die bekannten Rechenregeln (\rightarrow Nachrechnen):

$$x^p y^p = (xy)^p, \quad x^p x^q = x^{p+q}, \quad (x^p)^q = x^{pq}.$$

3.5. Supremum und Infimum von Mengen

Sei \mathbb{K} ein angeordneter Körper (z.B. $\mathbb{K} = \mathbb{R}$) und $M \subset \mathbb{K}$ eine nichtleere Menge mit endlich vielen Elementen, d.h. $\#M \in \mathbb{N}$. Dann gibt es immer ein kleinstes und ein größtes Element von M . Hat M unendlich viele Elemente, so muss das nicht mehr gelten. Wir können aber eine Art Ersatz definieren, wenn M im folgenden Sinne beschränkt ist:

Definition 3.5.1

Sei \mathbb{K} angeordnet. Eine Menge $M \subset \mathbb{K}$ heißt *nach oben* (bzw. *nach unten*) *beschränkt*, wenn ein $c \in \mathbb{K}$ existiert mit

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

Wir nennen dann c *obere* (bzw. *untere*) *Schranke von M*.

Die Menge M heißt *beschränkt*, wenn sie sowohl nach oben als auch nach unten beschränkt ist.

Bemerkung 3.5.2. (a) M ist genau dann beschränkt, wenn ein $c > 0$ existiert mit $|x| \leq c$ für alle $x \in M$.

(b) Eine Folge $\{x_n\}_n \subset \mathbb{K}$ ist genau dann beschränkt, wenn die zugehörige Menge $\{x_n : n \in \mathbb{N}\}$ beschränkt ist; vgl. Definition 3.1.7.

Zum Beispiel ist das Intervall $(0, 1) \subset \mathbb{R}$ eine beschränkte Menge, ein „größtes“ Element lässt sich aber nicht finden. Stattdessen führen wir folgende Begriffe ein:

Definition 3.5.3

- ▶ Ist \mathbb{K} angeordnet und $M \subset \mathbb{K}$ nach oben beschränkt, so heißt $\sigma \in \mathbb{K}$ *Supremum* oder *kleinste obere Schranke* von M , i.Z. $\sigma = \sup M$, falls folgendes gilt:
 - (i) σ ist obere Schranke von M .
 - (ii) Für jede weitere obere Schranke $\hat{\sigma}$ von M gilt $\sigma \leq \hat{\sigma}$.
- ▶ Ist $M \subset \mathbb{K}$ nach unten beschränkt, so heißt $\tau \in \mathbb{K}$ *Infimum* oder *größte untere Schranke* von M , i.Z. $\tau = \inf M$, falls folgendes gilt:
 - (i) τ ist untere Schranke von M .
 - (ii) Für jede weitere untere Schranke $\hat{\tau}$ von M gilt $\tau \geq \hat{\tau}$.

Aus der Definition ist sofort klar, dass Infimum und Supremum, wenn sie existieren, eindeutig bestimmt sind. Außerdem haben wir folgende praktische Charakterisierung von Infimum und Supremum:

Hilfssatz 3.5.4

- ▶ Ist \mathbb{K} angeordnet und $M \subset \mathbb{K}$ nach oben beschränkt und ist $\sigma \in \mathbb{K}$ eine obere Schranke von M , so gilt

$$\sigma = \sup M \Leftrightarrow \text{Für alle } \varepsilon \in \mathbb{K} \text{ mit } \varepsilon > 0 \text{ existiert ein } x \in M \text{ mit } x > \sigma - \varepsilon.$$
- ▶ Ist $M \subset \mathbb{K}$ nach unten beschränkt und ist $\tau \in \mathbb{K}$ untere Schranke von M , so gilt

$$\tau = \inf M \Leftrightarrow \text{Für alle } \varepsilon \in \mathbb{K} \text{ mit } \varepsilon > 0 \text{ existiert ein } x \in M \text{ mit } x < \tau + \varepsilon.$$

Beweis. Wir zeigen nur die Aussage für das Supremum.

- ▶ „ \Rightarrow “: Ist $\sigma = \sup M$ und existierte ein $\varepsilon > 0$ mit $x \leq \sigma - \varepsilon$ für alle $x \in M$, so wäre auch $\sigma - \varepsilon < \sigma$ obere Schranke von M , im Widerspruch zur Definition des Supremums.
- ▶ „ \Leftarrow “: Sei umgekehrt σ obere Schranke von M , und zu jedem $\varepsilon > 0$ existiere ein $x \in M$

mit $x > \sigma - \varepsilon$. Gäbe es eine obere Schranke $\hat{\sigma} < \sigma$ von M , so folgte für $\varepsilon := \sigma - \hat{\sigma} > 0$:

$$x \leq \hat{\sigma} = \sigma - \varepsilon \quad \text{für alle } x \in M,$$

Widerspruch! Also gilt doch $\hat{\sigma} \geq \sigma$ für alle oberen Schranken $\hat{\sigma}$ von M , d.h. wir haben $\sigma = \sup M$. \square

Für das folgende wichtige Resultat benötigen wir die vollständig archimedische Anordnung von \mathbb{K} , weshalb wir es wieder in \mathbb{R} formulieren.

Satz 3.5.5

Jede nichtleere, nach oben (bzw. unten) beschränkte Menge $M \subset \mathbb{R}$ besitzt ein Supremum (bzw. Infimum).

Beweis. Wir zeigen die Existenz des Supremums. Die des Infimums folgt dann durch Übergang zur Menge $-M := \{x \in \mathbb{R} : -x \in M\}$, für die gilt $\inf M = -\sup(-M)$.

Zum Beweis konstruieren wir induktiv eine Intervallschachtelung $I_1 \supset I_2 \supset \dots$ mit $I_n = [x_n, c_n]$, so dass für alle $n \in \mathbb{N}$ gilt:

$$|I_n| \leq \left(\frac{1}{2}\right)^{n-1} (c_1 - x_1), \quad x_n \in M, \quad c_n \text{ ist obere Schranke an } M. \quad (3.5.1)$$

- $n = 1$: Da M nichtleer und nach oben beschränkt ist, existiert ein $x_1 \in M$ und ein c_1 mit $x \leq c_1$ für alle $x \in M$.
- $n \rightarrow n + 1$: Sei $I_n = [x_n, c_n]$ mit (3.5.1). Wir setzen $y_n := \frac{1}{2}(x_n + c_n)$ und erklären I_{n+1} wie folgt:

- (i) Falls $M \cap [y_n, c_n] = \emptyset$, dann ist y_n obere Schranke an M und wir setzen $x_{n+1} := x_n \in M$, $c_{n+1} := y_n$.
- (ii) Falls $M \cap [y_n, c_n] \neq \emptyset$, so existiert ein $x_{n+1} \in M$ mit $x_{n+1} \geq y_n$. Dann setzen wir $c_{n+1} := c_n$.

Offenbar ist dann (3.5.1) erfüllt für I_{n+1} und wir haben $I_{n+1} \subset I_n$.

Wegen Folgerung 2.4.5 folgt $|I_n| \rightarrow 0$ ($n \rightarrow \infty$). Nach Satz 3.4.3 existiert also ein eindeutiges $\sigma \in \bigcap_{n \in \mathbb{N}} I_n$ und gemäß Bemerkung 3.4.4 (b) gilt $x_n \rightarrow \sigma, c_n \rightarrow \sigma$ ($n \rightarrow \infty$).

Wir zeigen noch $\sigma = \sup M$: Da $x \leq c_n$ für alle $x \in M$ und $n \in \mathbb{N}$ gilt, liefert Grenzübergang $x \leq \sigma$ für alle $x \in M$, d.h. σ ist obere Schranke. Ist ferner $\varepsilon > 0$ beliebig, so existiert wegen $x_n \rightarrow \sigma$ ($n \rightarrow \infty$) ein $N \in \mathbb{N}$ mit $|x_N - \sigma| < \varepsilon$ und insbesondere $x_N > \sigma - \varepsilon$. Da gemäß (3.5.1) $x_N \in M$ gilt, folgt $\sigma = \sup M$ aus Hilfssatz 3.5.4, wie behauptet. \square

Bemerkung 3.5.6. Es gilt auch die Umkehrung von Satz 3.5.5:

Ist \mathbb{K} ein angeordneter Körper und besitzt jede nach oben (bzw. unten) beschränkte Menge $M \subset \mathbb{K}$ ein Supremum (bzw. Infimum), so ist \mathbb{K} vollständig archimedisch angeordnet (und damit i. W. $\mathbb{K} = \mathbb{R}$).

Zum Beweis zeigt man, dass das Intervallschachtelungsprinzip auch unter diesem *DEDEKINDSchen Kriterium* gilt; daraus folgt wie bei uns der Satz von BOLZANO-WEIERSTRASS und mit diesem lässt sich die Vollständigkeit beweisen; siehe z.B. [Hil1] Abschnitt 1.7, Satz 5 und Abschnitt 1.11, Satz 2.

Übrigens dient das DEDEKINDSche Kriterium auch als alternative Definition der Vollständigkeit und ersetzt dabei sowohl (V) als auch (O3); letzteres folgt dann als *Satz von ARCHIMEDES*, siehe [Hil1] Abschnitt 1.4, Satz 5.

Beispiel 3.5.7. ▶ Ist $[a, b)$ nichtleer, so gilt $\inf[a, b) = a$ und $\sup[a, b) = b$.

Denn: a ist offenbar untere Schranke von $[a, b) := \{x \in \mathbb{R} : a \leq x < b\}$. Und wegen $a \in [a, b)$ muss jede weitere untere Schranke a' insbesondere $a' \leq a$ erfüllen, d.h. $a = \inf[a, b)$.

Andererseits ist offenbar b obere Schranke von $[a, b)$. Und ist $\varepsilon > 0$ beliebig gewählt, so setzen wir $x := \max\{a, b - \frac{\varepsilon}{2}\}$. Dann ist $x \in [a, b)$ und $x \geq b - \frac{\varepsilon}{2} > b - \varepsilon$ richtig, nach Hilfssatz 3.5.4 gilt somit $b = \sup[a, b)$. □

▶ Für $A := \{\frac{n+1}{n} : n \in \mathbb{N}\}$ ist $\inf A = 1$, $\sup A = 2$ erfüllt.

Denn: Es gelten $\frac{n+1}{n} > 1$ für alle $n \in \mathbb{N}$ und $\lim_{n \rightarrow \infty} \frac{n+1}{n} = 1$, also auch $\frac{n+1}{n} < 1 + \varepsilon$ für beliebiges $\varepsilon > 0$ und hinreichend großes $n \in \mathbb{N}$, so dass $\inf A = 1$ folgt. Weiter gilt offenbar $2 \geq \frac{n+1}{n}$ für alle $n \in \mathbb{N}$ sowie $2 \in A$, woraus sich sofort $\sup A = 2$ ergibt. □

Bemerkung 3.5.8. (a) Obige Beispiele zeigen, dass $\inf M$ zur Menge M dazu gehören kann oder nicht. Wenn $\inf M \in M$ gilt, so schreiben wir auch $\min M := \inf M$ für das *Minimum von M* . Ebenso sprechen wir vom *Maximum von M* , falls $\sup M \in M$ gilt und schreiben dann $\max M := \sup M$. Für Mengen mit endlich vielen Elementen ist immer $\min M = \inf M$ und $\max M = \sup M$ erfüllt.

(b) Für nach oben bzw. nach unten unbeschränkte Mengen M schreiben wir auch

$$\sup M = +\infty \quad \text{bzw.} \quad \inf M = -\infty.$$

3.6. Folgen in \mathbb{R} und ihre Häufungswerte

Wir beschränken uns wieder auf den Körper der reellen Zahlen. Wir wissen bereits, dass jede konvergente Folge $\{x_n\}_n \subset \mathbb{R}$ beschränkt ist (Hilfssatz 3.2.3). Umgekehrt sind beschränkte

Folgen in \mathbb{R} zwar nicht immer konvergent, besitzen aber immer eine konvergente Teilfolge (Satz 3.4.6).

Definition 3.6.1

$x \in \mathbb{R}$ heißt *Häufungswert* einer Folge $\{x_n\}_n \subset \mathbb{R}$, wenn es eine Teilfolge $\{x_{n_k}\}_k \subset \{x_n\}_n$ gibt mit $\lim_{k \rightarrow \infty} x_{n_k} = x$.

Satz 3.4.6 von BOLZANO-WEIERSTRASS besagt also: Jede beschränkte Folge hat einen Häufungswert.

Beispiel 3.6.2. ▶ $\{(-1)^n\}_n$ besitzt die Häufungswerte -1 und $+1$.

- ▶ Die Folge $\{n\}_n$ der natürlichen Zahlen besitzt keine Häufungswerte.

Denn: Gäbe es eine Teilfolge $\{n_k\}_k$ mit $n_k \rightarrow x$ ($k \rightarrow \infty$), so existierte insbesondere ein $N \in \mathbb{N}$ mit $|n_k - x| < \frac{1}{2}$ für alle $k \geq N$. Für $k, l \geq N$ mit $k \neq l$ folgte also

$$1 \leq |n_k - n_l| \leq |n_k - x| + |n_l - x| < \frac{1}{2} + \frac{1}{2} = 1,$$

Widerspruch. □

- ▶ Es gibt natürlich auch unbeschränkte Folgen mit Häufungswerten, z.B. hat $\{x_n\}_n = \{[1 + (-1)^n]n\}_n$ den Häufungswert 0, da gilt $x_{2k-1} = 0$ für alle $k \in \mathbb{N}$.

Will man alle Häufungswerte einer beschränkten Folge qualitativ kontrollieren, so wird man auf die folgenden Begriffe geführt:

Definition 3.6.3

Sei $\{x_n\}_n \subset \mathbb{R}$ eine beschränkte Folge und bezeichne H die Menge ihrer Häufungswerte. Wir setzen dann

$$\liminf_{n \rightarrow \infty} x_n := \inf H \quad (\text{Limes inferior}),$$

$$\limsup_{n \rightarrow \infty} x_n := \sup H \quad (\text{Limes superior}).$$

Bemerkung 3.6.4. Offenbar ist H beschränkt, wenn $\{x_n\}_n$ beschränkt ist. Man beachte noch

$$\liminf_{n \rightarrow \infty} (-x_n) = -\limsup_{n \rightarrow \infty} x_n, \quad \limsup_{n \rightarrow \infty} (-x_n) = -\liminf_{n \rightarrow \infty} x_n.$$

Beispiel 3.6.5. Die Folge $\{x_n\}_n = \{(-1)^n + \frac{1}{n}\}_n$ ist beschränkt und wir haben die konvergenten Teilfolgen

$$x_{2k} = 1 + \frac{1}{2k} \rightarrow 1 \quad (k \rightarrow \infty), \quad x_{2k-1} = -1 + \frac{1}{2k-1} \rightarrow -1 \quad (k \rightarrow \infty).$$

Also gilt $H = \{-1, 1\}$ und $\liminf_{n \rightarrow \infty} x_n = -1$, $\limsup_{n \rightarrow \infty} x_n = 1$. Man beachte, dass $\liminf_{n \rightarrow \infty} x_n$ und $\limsup_{n \rightarrow \infty} x_n$ zu H gehören. Dies ist immer so:

Satz 3.6.6

Ist $\{x_n\}_n \subset \mathbb{R}$ beschränkt, so ist $\liminf_{n \rightarrow \infty} x_n$ der kleinste, $\limsup_{n \rightarrow \infty} x_n$ der größte Häufungswert von $\{x_n\}_n$, d.h. es gilt

$$\liminf_{n \rightarrow \infty} x_n = \min H, \quad \limsup_{n \rightarrow \infty} x_n = \max H.$$

Beweis. \rightarrow Ergänzungen. □

Satz 3.6.7: Charakterisierung von \limsup

Sei $\{x_n\}_n \subset \mathbb{R}$ eine beschränkte Folge. Dann ist $\eta = \limsup_{n \rightarrow \infty} x_n$ genau dann erfüllt, wenn gilt:

- (i) η ist Häufungswert von $\{x_n\}_n$.
- (ii) Für alle $\varepsilon > 0$ existiert ein $N = N(\varepsilon) \in \mathbb{N}$, so dass gilt

$$x_n < \eta + \varepsilon \quad \text{für alle } n \geq N(\varepsilon).$$

Beweis. \blacktriangleright „ \Rightarrow “: Sei $\eta = \limsup_{n \rightarrow \infty} x_n$. Nach Satz 3.6.6 ist dann $\eta \in H$, also (i) erfüllt.

Wäre (ii) falsch, so gäbe es ein $\varepsilon > 0$ und eine Teilfolge $\{x'_{k_l}\}_l = \{x_{n_k}\}_k$ mit

$$x'_{k_l} \geq \eta + \varepsilon \quad \text{für alle } l \in \mathbb{N}. \quad (3.6.1)$$

Da $\{x'_{k_l}\}_l$ beschränkt ist, existiert nach dem Satz von Bolzano–Weierstraß eine weitere Teilfolge $\{x'_{k_{l_i}}\}_i \subset \{x'_{k_l}\}_l \subset \{x_n\}_n$ und ein $\zeta \in \mathbb{R}$ mit $x'_{k_{l_i}} \rightarrow \zeta$ ($i \rightarrow \infty$). Also ist $\zeta \in H$ und somit $\zeta \leq \eta$. Andererseits liefert aber (3.6.1) angewendet auf $\{x'_{k_{l_i}}\}_i$ nach Grenzübergang $i \rightarrow \infty$: $\zeta \geq \eta + \varepsilon$, Widerspruch!

- \blacktriangleright „ \Leftarrow “: Seien nun (i) und (ii) erfüllt. Wäre $\eta \neq \limsup_{n \rightarrow \infty} x_n = \max H$, so gäbe es ein $\zeta \in H$ mit $\zeta > \eta$. Wir setzen $\varepsilon := \frac{\zeta - \eta}{2} > 0$. Für die Teilfolge $\{x_{n_k}\}_k$ mit $\lim_{k \rightarrow \infty} x_{n_k} = \zeta$ existiert dann ein $\hat{k} \in \mathbb{N}$, so dass $n_{\hat{k}} \geq N(\varepsilon)$ und $|x_{n_{\hat{k}}} - \zeta| < \varepsilon$ richtig ist. Aus (ii) für $n = n_{\hat{k}}$ erhalten wir dann

$$x_{n_{\hat{k}}} < \eta + \varepsilon = \eta + \frac{\zeta - \eta}{2} = \zeta - \frac{\zeta - \eta}{2} = \zeta - \varepsilon < x_{n_{\hat{k}}},$$

Widerspruch! Also ist $\eta = \max H$, wie behauptet. □

Wir halten noch die entsprechende Aussage für den Limes inferior fest, die sich analog oder mit Bemerkung 3.6.4 beweisen lässt:

Satz 3.6.8: Charakterisierung von \liminf

Sei $\{x_n\}_n$ eine beschränkte Folge. Dann ist $\xi = \liminf_{n \rightarrow \infty} x_n$ genau dann erfüllt, wenn gilt:

- (i) ξ ist Häufungswert von $\{x_n\}_n$.
- (ii) Für alle $\varepsilon > 0$ existiert ein $N = N(\varepsilon) \in \mathbb{N}$, so dass gilt

$$x_n > \xi - \varepsilon \quad \text{für alle } n \geq N(\varepsilon).$$

Schließlich notieren wir die folgenden äquivalenten Darstellungen:

Satz 3.6.9

Ist $\{x_n\}_n$ eine beschränkte Folge, so gilt

$$\limsup_{n \rightarrow \infty} x_n = \lim_{n \rightarrow \infty} \left(\sup \{x_k : k \geq n\} \right),$$

$$\liminf_{n \rightarrow \infty} x_n = \lim_{n \rightarrow \infty} \left(\inf \{x_k : k \geq n\} \right).$$

Beweis. → Ergänzungen. □

Bemerkung 3.6.10. (a) Die Darstellungen aus Satz 3.6.9 werden häufig auch als Definition von \limsup und \liminf verwendet. Sie sind zwar etwas unanschaulich, haben aber den Vorteil, dass sie auch für *unbeschränkte* Folgen Sinn machen: Ist $\{x_n\}_n$ etwa *nach oben unbeschränkt*, d.h. es existiert zu jedem $c \in \mathbb{R}$ ein $n \in \mathbb{N}$ mit $x_n > c$, so ist $\sup \{x_k : k \geq n\} = +\infty$ für alle $n \in \mathbb{N}$. Dann setzen wir

$$\limsup_{n \rightarrow \infty} x_n = +\infty.$$

Entsprechend ist für eine *nach unten unbeschränkte* Folge, d.h. es existiert zu jedem $c \in \mathbb{R}$ ein $n \in \mathbb{N}$ mit $x_n < c$,

$$\liminf_{n \rightarrow \infty} x_n = -\infty.$$

(b) Als *Übungsaufgabe* zeige man: Eine Folge $\{x_n\}_n \subset \mathbb{R}$ ist genau dann konvergent gegen $\alpha \in \mathbb{R}$, wenn sie beschränkt ist und wenn gilt

$$\liminf_{n \rightarrow \infty} x_n = \alpha = \limsup_{n \rightarrow \infty} x_n.$$

3.7. \mathbb{Q} als Teilmenge von \mathbb{R}

Wir wissen zwar (nach unserer Definition) dass $\mathbb{Q} \subset \mathbb{R}$ gilt, aber wo genau liegen die rationalen Zahlen auf der reellen Geraden? Die (stark vereinfachte) Antwort auf diesen Frage ist: Überall. Zur genaueren Klärung des Sachverhalts beginnen wir mit der

Definition 3.7.1

Eine Teilmenge $M \subset \mathbb{K}$ eines bewerteten Körpers \mathbb{K} heißt *dicht* in \mathbb{K} , falls es zu jedem $x \in \mathbb{K}$ eine Folge $\{x_n\}_n \subset M$ so gibt, dass $\lim_{n \rightarrow \infty} x_n = x$ gilt.

Bemerkung 3.7.2. Ist M dicht in \mathbb{K} , so lässt sich jede Zahl $x \in \mathbb{K}$ also beliebig gut durch Elemente aus M approximieren: In jeder ε -Umgebung $U_\varepsilon(x)$ findet man ein Element aus M . Natürlich liegt \mathbb{K} selbst dicht in \mathbb{K} .

Satz 3.7.3

\mathbb{Q} liegt dicht in \mathbb{R} .

Für den Beweis benötigen wir noch die nachstehende einfache Konsequenz aus dem Archimedischen Axiom (O3):

Hilfssatz 3.7.4

Zu jeder Zahl $x \in \mathbb{R}$ existiert genau ein $\nu \in \mathbb{Z}$, so dass $\nu \leq x < \nu + 1$ richtig ist.

Beweis. \rightarrow Übungsaufgabe. □

Beweis von Satz 3.7.3. Zu gegebenem $x \in \mathbb{R}$ konstruieren wir mittels vollständiger Induktion ein $a_0 \in \mathbb{Z}$ und eine Folge $\{a_k\}_{k \in \mathbb{N}} \subset \{0, 1, \dots, 9\}$, so dass für den unendlichen Dezimalbruch $\{x_n\}_{n \in \mathbb{N}} \subset \mathbb{Q}$ mit

$$x_n = \sum_{k=0}^n a_k \cdot 10^{-k}, \quad n \in \mathbb{N},$$

gilt $x_n \rightarrow x$ ($n \rightarrow \infty$).

1. Sei zunächst $0 \leq x < 1$ richtig. Wir behaupten, dass dann eine Folge $\{a_k\}_k \subset \{0, 1, \dots, 9\}$ und eine Nullfolge $\{\xi_n\}_n \subset \mathbb{R}$ so existieren, dass für alle $n \in \mathbb{N}$ gilt

$$x = \sum_{k=1}^n a_k \cdot 10^{-k} + \xi_n \quad \text{und} \quad 0 \leq \xi_n < 10^{-n}. \quad (3.7.1)$$

Offenbar folgt mit $a_0 = 0$ daraus $|x - x_n| = \xi_n \rightarrow 0$ ($n \rightarrow \infty$) und somit die Behauptung.

(IA) $n = 1$. Wegen $0 \leq x \cdot 10 < 10$ und Hilfssatz 3.7.4 existiert ein $a_1 \in \{0, 1, 2, \dots, 9\}$, so dass $a_1 \leq x \cdot 10 < a_1 + 1$ richtig ist. Mit $\xi_1 := x - a_1 \cdot 10^{-1}$ haben wir dann

$$x = a_1 \cdot 10^{-1} + \xi_1 \quad \text{und} \quad 0 \leq \xi_1 < 10^{-1}.$$

(IS) $n \rightarrow n+1$: Angenommen wir haben die Darstellung (3.7.1) für ein $n \in \mathbb{N}$. Dann ist $0 \leq \xi_n \cdot 10^{n+1} < 10$ richtig und wieder nach Hilfssatz 3.7.4 existiert ein $a_{n+1} \in \{0, 1, \dots, 9\}$ mit $a_{n+1} \leq \xi_n \cdot 10^{n+1} < a_{n+1} + 1$. Setzen wir noch $\xi_{n+1} := \xi_n - a_{n+1} \cdot 10^{-(n+1)}$, so finden wir

$$x \stackrel{(IV)}{=} \sum_{k=1}^n a_k \cdot 10^{-k} + \xi_n = \sum_{k=1}^{n+1} a_k \cdot 10^{-k} + \xi_{n+1}$$

und $0 \leq \xi_{n+1} < 10^{-(n+1)}$, wie behauptet.

2. Sei nun $x \in \mathbb{R}$ beliebig. Nach Hilfssatz 3.7.4 existiert ein $a_0 \in \mathbb{Z}$ mit $a_0 \leq x < a_0 + 1$. Dann gilt für $y := x - a_0$ natürlich $0 \leq y < 1$, und nach Teil 1 existiert eine Folge $\{a_k\}_k \subset \{0, \dots, 9\}$, so dass $y_n = \sum_{k=1}^n a_k \cdot 10^{-k} \rightarrow y$ ($n \rightarrow \infty$) richtig ist. Also hat $x_n := a_0 + y_n = \sum_{k=0}^n a_k \cdot 10^{-k}$ die gesuchte Form, und es gilt

$$x_n = a_0 + y_n \rightarrow a_0 + y = x \quad (n \rightarrow \infty),$$

wie behauptet. □

Bemerkung 3.7.5. ▶ Der soeben gegebene Beweis von Satz 3.7.3 bestätigt unsere Vorstellung, dass sich jede reelle Zahl als unendlicher Dezimalbruch darstellen lässt. Wir haben hier nämlich

$$x = \lim_{n \rightarrow \infty} \sum_{k=0}^n a_k \cdot 10^{-k} = a_0, a_1 a_2 a_3 \dots$$

gezeigt. Übrigens ist diese Darstellung nicht eindeutig, denn z.B. lässt sich die Zahl 1 sowohl als $1,00000\dots$ schreiben, als auch als

$$\begin{aligned} 0,99999\dots &= \lim_{n \rightarrow \infty} \sum_{k=1}^n 9 \cdot 10^{-k} = 9 \cdot \lim_{n \rightarrow \infty} \sum_{k=0}^n \left(\frac{1}{10}\right)^k - 9 \\ &\stackrel{\text{Satz 2.4.8}}{=} 9 \cdot \lim_{n \rightarrow \infty} \frac{1 - \left(\frac{1}{10}\right)^{n+1}}{1 - \frac{1}{10}} - 9 = 1 - \lim_{n \rightarrow \infty} \left(\frac{1}{10}\right)^n \\ &\stackrel{\text{Beispiel 3.1.12 (b)}}{=} 1. \end{aligned}$$

- ▶ Auch die irrationalen Zahlen $\mathbb{R} \setminus \mathbb{Q}$ liegen übrigens dicht in \mathbb{R} (\rightarrow Übungsaufgabe).

Eine ebenso naheliegende Frage wie die nach der Verteilung der rationalen und irrationalen Zahlen in \mathbb{R} ist die nach deren Anzahl. Hier ist die Antwort: Jeweils unendlich viele, aber mehr irrationale als rationale. Das klingt komisch und bedarf ebenfalls einer Präzisierung, wir beginnen mit der

Definition 3.7.6

Eine nichtleere Menge M heit:

- ▶ *endlich*, wenn sie endlich viele Elemente hat, d.h. $\#M \in \mathbb{N}$.
- ▶ *abzhlbar unendlich*, wenn eine Bijektion $f : \mathbb{N} \rightarrow M$ existiert.
- ▶ *abzhlbar*, wenn sie endlich oder abzhlbar unendlich ist.
- ▶ *berabzhlbar*, wenn sie nicht abzhlbar ist.

Bemerkung 3.7.7. ▶ Ist M abzhlbar unendlich mit Bijektion $f : \mathbb{N} \rightarrow M$, so setzen wir $x_n := f(n)$, $n \in \mathbb{N}$, und finden eine Folge $\{x_n\}_n$ mit $M = f(\mathbb{N}) = \{x_n : n \in \mathbb{N}\}$. Wir schreiben dafr auch kurz (und nicht ganz korrekt) $M = \{x_n\}_n$.

▶ Zwei Mengen M, N , fr die eine Bijektion $f : N \rightarrow M$ existiert, heien *gleichmchtig*. Eine abzhlbar unendliche Menge ist also gleichmchtig zu den natrlichen Zahlen. Und zwei endliche Mengen M, N sind genau dann gleichmchtig, wenn $\#M = \#N$ gilt.

Beispiel 3.7.8. (a) Die Menge \mathbb{N} der natrlichen Zahlen ist abzhlbar unendlich mit der identischen Abbildung $f = \text{id}_{\mathbb{N}} : \mathbb{N} \rightarrow \mathbb{N}$.

(b) Die ganzen Zahlen \mathbb{Z} sind abzhlbar unendlich mit der Bijektion

$$f(n) := \begin{cases} \frac{1}{2}(1-n), & \text{falls } n \text{ ungerade ist} \\ \frac{1}{2}n, & \text{falls } n \text{ gerade ist} \end{cases}, \quad n \in \mathbb{N}.$$

(c) Sind M und N abzhlbar unendlich, so ist auch $M \cup N$ abzhlbar. Deutlich allgemeiner gilt der folgende

Satz 3.7.9

Die Vereinigung abzhlbar vieler abzhlbarer Mengen ist abzhlbar.

Beweis. O.B.d.A. seien M_n , $n \in \mathbb{N}$, abzhlbar unendlich viele abzhlbar unendliche Mengen (siehe Bemerkung 3.7.10). Wir schreiben $M_n = \{x_{nm} : m \in \mathbb{N}\}$ fr $n \in \mathbb{N}$. Die Elemente der Vereinigungsmenge

$$\bigcup_{n \in \mathbb{N}} M_n = \{x_{nm} : m, n \in \mathbb{N}\}$$

können wir z.B. wie folgt abzählen:

$$\begin{array}{ccccccc}
 M_1 : & x_{11} & \rightarrow & x_{12} & & x_{13} & \rightarrow & x_{14} & \dots \\
 & & \swarrow & & \nearrow & & \swarrow & & \nearrow \\
 M_2 : & x_{21} & & x_{22} & & x_{23} & & x_{24} & \dots \\
 & \downarrow & \nearrow & & \swarrow & & \nearrow & & \swarrow \\
 M_3 : & x_{31} & & x_{32} & & x_{33} & & x_{34} & \dots \\
 & & \swarrow & & \nearrow & & \swarrow & & \nearrow \\
 M_4 : & x_{41} & & x_{42} & & x_{43} & & x_{44} & \dots \\
 \vdots & \downarrow & \nearrow & \vdots & \swarrow & \vdots & \nearrow & \vdots & \\
 \end{array}$$

also mit der Abzählung $y_1 := x_{11}, y_2 := x_{12}, y_3 := x_{21}, y_4 := x_{31}, \dots$ □

Bemerkung 3.7.10. Im Beweis von Satz 3.7.9 wurde nur der „interessanteste“ Fall untersucht:

- Haben wir nur endlich viele abzählbar unendliche Mengen M_1, \dots, M_N , so kann man z.B. folgendes einfachere Schema für das Abzählen nutzen:

$$\begin{array}{cccc}
 M_1 : & x_{11} & & x_{12} \rightarrow x_{13} \\
 & \downarrow & & \uparrow \quad \downarrow \\
 M_2 : & x_{21} & & x_{22} \quad x_{23} \\
 & \downarrow & & \uparrow \quad \downarrow \\
 M_3 : & x_{31} & & x_{32} \quad x_{33} \\
 \vdots & \vdots & & \vdots \quad \vdots \\
 M_N : & x_{N1} & \rightarrow & x_{N2} \quad x_{N3} \rightarrow \dots
 \end{array}$$

- Sind eine oder mehrere der Mengen M_n endlich, so werden die Lücken, die im ersten oder zweiten Schema dadurch in der jeweiligen Zeile entstehen, einfach übergangen.

Folgerung 3.7.11

Die Menge der rationalen Zahlen ist abzählbar.

Beweis. Wir setzen $M_n := \left\{ \frac{m}{n} : m \in \mathbb{Z} \right\}$ für $n \in \mathbb{N}$. Da \mathbb{Z} abzählbar ist, ist auch M_n abzählbar für jedes $n \in \mathbb{N}$. Und nach Satz 3.7.9 gilt dies auch für die Vereinigung

$$\bigcup_{n \in \mathbb{N}} M_n = \left\{ \frac{m}{n} : m \in \mathbb{Z}, n \in \mathbb{N} \right\} = \mathbb{Q},$$

wie behauptet. □

Folgerung 3.7.11 enthält die überraschende Beobachtung, dass die rationalen Zahlen gleichmächtig zu den natürlichen Zahlen sind. Für die reellen Zahlen wird diese Aussage falsch:

Satz 3.7.12

Die Menge der reellen Zahlen ist überabzählbar.

Beweis. Mittels Intervallschachtelung \rightarrow Ergänzungen. □

Folgerung 3.7.13

Die Menge der irrationalen Zahlen $\mathbb{R} \setminus \mathbb{Q}$ ist überabzählbar.

Beweis. Anderenfalls wäre nach Folgerung 3.7.11 und Satz 3.7.9 auch $(\mathbb{R} \setminus \mathbb{Q}) \cup \mathbb{Q} = \mathbb{R}$ abzählbar, im Widerspruch zu Satz 3.7.12. □

Kapitel 4.

Reihen

Reihen sind unendliche Summen. Die bei dieser unendlichen Summationen auftretenden Besonderheiten werden im vorliegenden Kapitel diskutiert. In Abschnitt A.4.1 geben wir die Definition und das fundamentale Cauchysche Konvergenzkriterium an, aus dem in Abschnitt A.4.2 mehrere Konvergenzkriterien abgeleitet werden. Abschnitt A.4.3 enthält verschiedene Resultate, insbesondere das Leibnizkriterium und den Cauchyschen Produktsatz; letzterer wird dann auf die in Abschnitt A.4.4 diskutierten Potenzreihen angewendet und uns später als Grundlage für die Betrachtung der *elementaren Funktionen* dienen.

4.1. Definition und Rechenregeln

Wir wollen Summen von abzählbar unendlich vielen Zahlen betrachten, wobei wir uns auf den Körper der komplexe Zahlen beschränken wollen. Wir werden später feststellen, dass dabei die Reihenfolge der Summation wichtig ist. Wir beginnen mit der

Definition 4.1.1: Reihen

- Zu einer Folge $\{z_k\}_k \subset \mathbb{C}$ komplexer Zahlen erklären wir die (*n-ten*) *Partialsummen*

$$s_n := \sum_{k=1}^n z_k = z_1 + z_2 + \dots + z_n, \quad n \in \mathbb{N}.$$

Die Folge $\{s_n\}_n \subset \mathbb{C}$ der Partialsummen heißt dann (*unendliche*) *Reihe*, die z_k werden auch als *Reihenglieder* bezeichnet.

- Konvergiert $\{s_n\}_n$, so sagen wir, dass die Reihe *konvergiert* und schreiben für den Grenzwert, genannt *Wert der Reihe*,

$$\sum_{k=1}^{\infty} z_k := \lim_{n \rightarrow \infty} s_n = \lim_{n \rightarrow \infty} \left(\sum_{k=1}^n z_k \right).$$

- Wenn die Reihe $\{s_n\}_n$ nicht konvergiert, heißt sie *divergent*. Wenn $z_k = x_k \in \mathbb{R}$ für alle $k \in \mathbb{N}$ und

$$\lim_{n \rightarrow \infty} s_n = \pm \infty$$

gilt, nennen wir die Reihe *bestimmt divergent (gegen $\pm \infty$)*.

Bemerkung 4.1.2. (a) Es ist üblich, auch die Folge der Partialsummen mit $\sum_{k=1}^{\infty} z_k$ zu bezeichnen; wir verwenden dieses Symbol also sowohl für die Reihe selbst (unabhängig von deren Konvergenz), als auch im Konvergenzfall für ihren Wert.

- (b) Wir können (und werden) auch Reihen der Form

$$\sum_{k=k_0}^{\infty} z_k = \lim_{n \rightarrow \infty} \sum_{k=k_0}^n z_k$$

mit einem $k_0 \in \mathbb{N}_0$ (oder sogar $k_0 \in \mathbb{Z}$) betrachten. Falls klar ist, über welche k summiert wird, schreiben wir auch kurz $\sum_k z_k$ für die Reihe bzw. ihren Wert.

- (c) Eine Reihe ist eine spezielle Folge, nämlich die von Partialsummen. Umgekehrt kann man jede Folge $\{\zeta_n\}_n \subset \mathbb{C}$ auch durch eine Reihe darstellen, denn es gilt

$$\zeta_n = \zeta_1 + \sum_{k=2}^n (\zeta_k - \zeta_{k-1}), \quad n \in \mathbb{N} \setminus \{1\} \quad (\text{Teleskopsumme}).$$

Satz 4.1.3: Rechenregeln für Reihen

Sind $\sum_{k=1}^{\infty} z_k$ und $\sum_{k=1}^{\infty} \zeta_k$ konvergente Reihen, so gelten:

- (a) Mit beliebigen $a, b \in \mathbb{C}$ konvergiert auch die Reihe $\sum_{k=1}^{\infty} (az_k + b\zeta_k)$, und es gilt

$$\sum_{k=1}^{\infty} (az_k + b\zeta_k) = a \sum_{k=1}^{\infty} z_k + b \sum_{k=1}^{\infty} \zeta_k.$$

- (b) Auch die Reihe $\sum_{k=1}^{\infty} \overline{z_k}$ ist konvergent, und es gilt

$$\sum_{k=1}^{\infty} \overline{z_k} = \overline{\sum_{k=1}^{\infty} z_k}.$$

Beweis. Sofort durch Anwenden von Folgerung 3.3.2 bzw. Bemerkung 3.3.4 auf die Folge der Partialsummen (\rightarrow Nachrechnen). \square

Beispiel 4.1.4. (a) *Geometrische Reihe* $\sum_{k=0}^{\infty} z^k$. Für $|z| < 1$ gilt $|z^n| = |z|^n \rightarrow 0$ ($n \rightarrow \infty$) nach Beispiel 3.1.12 (b). Satz 2.4.8 liefert also

$$\sum_{k=0}^{\infty} z^k = \lim_{n \rightarrow \infty} \sum_{k=0}^n z^k = \lim_{n \rightarrow \infty} \frac{1 - z^{n+1}}{1 - z} = \frac{1}{1 - z} (1 - z \lim_{n \rightarrow \infty} z^n) = \frac{1}{1 - z}. \quad (4.1.1)$$

(b) Als Beispiel einer Teleskopsumme halten wir fest:

$$\sum_{k=1}^{\infty} \frac{1}{k(k+1)} = \lim_{n \rightarrow \infty} \sum_{k=1}^n \frac{1}{k(k+1)} = \lim_{n \rightarrow \infty} \frac{n}{n+1} = 1.$$

Denn: Setzen wir $y_k = \frac{k}{k+1}$ für $k \in \mathbb{N}$, so folgt

$$y_k - y_{k-1} = \frac{k}{k+1} - \frac{k-1}{k} = \frac{k^2 - (k^2 - 1)}{k(k+1)} = \frac{1}{k(k+1)} \quad \text{für } k \in \mathbb{N} \setminus \{1\}$$

und somit

$$\frac{n}{n+1} = y_n = y_1 + \sum_{k=2}^n (y_k - y_{k-1}) = \frac{1}{2} + \sum_{k=2}^n \frac{1}{k(k+1)} = \sum_{k=1}^n \frac{1}{k(k+1)}, \quad n \in \mathbb{N},$$

woraus die Behauptung sofort folgt. \square

Im folgenden Abschnitt A.4.2 werden wir wichtige Konvergenzkriterien für Reihen kennenlernen, die alle auf folgender Konsequenz des Vollständigkeitsaxioms beruhen wir:

Satz 4.1.5: CAUCHYSches Konvergenzkriterium für Reihen

Eine Reihe $\sum_{k=1}^{\infty} z_k$ konvergiert genau dann, wenn für beliebige $\varepsilon > 0$ ein $N = N(\varepsilon) \in \mathbb{N}$ so existiert, dass gilt

$$\left| \sum_{k=m+1}^n z_k \right| < \varepsilon \quad \text{für alle } n > m \geq N(\varepsilon). \quad (4.1.2)$$

Beweis. Wir beachten

$$|s_n - s_m| = \left| \sum_{k=1}^n z_k - \sum_{k=1}^m z_k \right| = \left| \sum_{k=m+1}^n z_k \right| \quad \text{für alle } n > m.$$

Also ist (4.1.2) äquivalent dazu, dass $\{s_n\}_n$ eine CAUCHYfolge bildet und somit nach Satz 3.2.9 auch äquivalent zur Konvergenz der Folge $\{s_n\}_n$ bzw. der Reihe $\sum_k z_k$. \square

Bemerkung 4.1.6. Satz 4.1.5 zeigt übrigens auch: Die Reihe $\sum_{k=1}^{\infty} z_k$ konvergiert (bzw. divergiert) genau dann, wenn für beliebiges $k_0 \in \mathbb{N}$ die Reihe $\sum_{k=k_0}^{\infty} z_k$ konvergiert (bzw. divergiert). Die ersten endlich vielen Glieder beeinflussen das Konvergenzverhalten der Reihe also nicht (aber natürlich ihren Wert).

4.2. Konvergenzkriterien für Reihen bei absoluter Konvergenz

Das Untersuchen einer Reihe auf Konvergenz ist meist einfacher, als den Wert der Reihe tatsächlich zu bestimmen (letzterer lässt sich dann aber bei Bedarf durch die Partialsummen

beliebig gut annähern). In diesem Abschnitt werden wir eine Reihe praktischer *Kriterien* für die Reihenkonvergenz kennenlernen. Allgemein unterscheidet man bei Kriterien k für eine Aussage p :

- ▶ **Hinreichendes Kriterium:** Wenn das Kriterium erfüllt ist, dann gilt die jeweilige Aussage; d.h. $k \rightarrow p$.
- ▶ **Notwendiges Kriterium:** Wenn die Aussage erfüllt ist, dann gilt das Kriterium; d.h. $p \rightarrow k$.
- ▶ **Äquivalentes Kriterium:** Die Aussage gilt genau dann, wenn das Kriterium gilt; d.h. $p \leftrightarrow k$.

Die Cauchyschen Konvergenzkriterien aus den Sätzen 3.2.5, 3.2.9, 4.1.5 sind Beispiele für äquivalente Kriterien, ebenso wie die Umkehrbarkeit für bijektive Abbildungen aus Satz 1.5.13. Beschränktheit einer Folge in \mathbb{R} ist nach Hilfssatz 3.2.3 *notwendig* für deren Konvergenz und nach Satz 3.4.6 *hinreichend* für die Konvergenz einer Teilfolge.

Bei unseren Konvergenzkriterien für Reihen beginnen wir mit einem einfachen notwendigen Kriterium, das sich auch - durch Bilden der Kontraposition - zum Beweis der Divergenz einer Reihe eignet:

Satz 4.2.1

Sei $\{z_k\}_k \subset \mathbb{C}$ eine Folge. Wenn $\sum_{k=1}^{\infty} z_k$ konvergiert, so muss gelten

$$\lim_{k \rightarrow \infty} z_k = 0.$$

Beweis. Sei $\sum_k z_k$ konvergent und $\varepsilon > 0$ beliebig gewählt. Nach Satz 4.1.5 existiert dann ein $N(\varepsilon) \in \mathbb{N}$ mit $|z_n| < \varepsilon$ für alle $n > N(\varepsilon)$ (wende (4.1.2) mit $m = n - 1$ an). Also ist $\{z_n\}_n$ eine Nullfolge. \square

Beispiel 4.2.2. (a) $\sum_{k=1}^{\infty} (-1)^k$ ist divergent, da $\{(-1)^k\}_k$ keine Nullfolge ist.

(b) Für $|z| \geq 1$ ist die geometrische Reihe $\sum_{k=0}^{\infty} z^k$ divergent, da dann $|z^k| = |z|^k \geq 1$ für alle $k \in \mathbb{N}$ gilt, d.h. $\{z^k\}_k$ keine Nullfolge ist.

(c) Das Kriterium in Satz 4.2.1 ist *nicht* hinreichend. Ein Gegenbeispiel ist die divergente *harmonische Reihe* $\sum_{k=1}^{\infty} \frac{1}{k}$.

Denn: Für beliebiges $m \in \mathbb{N}$ gilt:

$$\left| \sum_{k=m+1}^{2m} \frac{1}{k} \right| \geq \sum_{k=m+1}^{2m} \frac{1}{2m} = \frac{m}{2m} = \frac{1}{2}.$$

Also ist das CAUCHYSche Konvergenzkriterium (4.1.2) für $\varepsilon < \frac{1}{2}$ nicht erfüllbar. \square

Definition 4.2.3

Eine Reihe $\sum_{k=1}^{\infty} z_k$, $\{z_k\}_k \subset \mathbb{C}$, heißt *absolut konvergent*, wenn die Reihe $\sum_{k=1}^{\infty} |z_k|$ konvergiert.

Absolute Konvergenz ist eine stärkere Bedingung als Konvergenz im Sinne von Definition 4.2.1 (die wir im Folgenden auch *Konvergenz im gewöhnlichen Sinne* nennen), denn es gilt

Hilfssatz 4.2.4

Ist die Reihe $\sum_{k=1}^{\infty} z_k$ absolut konvergent, so konvergiert sie auch im gewöhnlichen Sinne.

Beweis. Nach der Dreiecksungleichung in \mathbb{C} gilt

$$\left| \sum_{k=m+1}^n z_k \right| \leq \sum_{k=m+1}^n |z_k| \quad \text{für alle } n > m.$$

Satz 4.1.5 liefert die Behauptung. \square

Bemerkung 4.2.5. Wir werden in Beispiel 4.3.4 anhand der *alternierenden harmonischen Reihe* sehen, dass die Umkehrung von Hilfssatz 4.2.4 *nicht* gilt. Im Falle einer Reihe $\sum_{k=1}^{\infty} x_k$ mit $\{x_k\}_k \subset \mathbb{R}$ und $x_k \geq 0$ für alle $k \in \mathbb{N}$ fallen absolute und gewöhnliche Konvergenz natürlich zusammen.

Eines der wichtigsten (hinreichenden) Konvergenzkriterien enthält der folgende

Satz 4.2.6: Majorantenkriterium

Zwei Folgen $\{z_k\}_k \subset \mathbb{C}$ und $\{\mu_k\}_k \subset \mathbb{R}$ mit

$$|z_k| \leq \mu_k \quad \text{für alle } k \in \mathbb{N}$$

seien gegeben. Dann gilt: Konvergiert die Reihe $\sum_{k=1}^{\infty} \mu_k$, so konvergiert $\sum_{k=1}^{\infty} z_k$ absolut. Die Reihe $\sum_k \mu_k$ heißt *Majorante* von $\sum_k z_k$.

Beweis. Zu beliebigem $\varepsilon > 0$ existiert nach Satz 4.1.5 ein $N(\varepsilon) \in \mathbb{N}$ mit

$$\sum_{k=m+1}^n |z_k| \leq \sum_{k=m+1}^n \mu_k < \varepsilon \quad \text{für alle } n > m \geq N(\varepsilon).$$

Wiederum nach Satz 4.1.5 konvergiert also auch $\sum_k |z_k|$, d.h. $\sum_k z_k$ konvergiert absolut. \square

Beispiel 4.2.7. Die Reihe $\sum_{k=1}^{\infty} \frac{1}{k^\alpha}$ konvergiert (absolut) für rationales $\alpha \geq 2$. Wir haben nämlich

$$0 < \frac{1}{k^\alpha} \leq \frac{1}{k^2} \leq \frac{2}{k(k+1)} \quad \text{für alle } k \in \mathbb{N},$$

also ist $\sum_{k=1}^{\infty} \frac{2}{k(k+1)}$ eine Majorante, die gemäß Beispiel 4.1.4 (b) konvergiert.

Folgerung 4.2.8: Minorantenkriterium

Sind $\{x_k\}_k, \{\mu_k\}_k \subset \mathbb{R}$ gegeben mit

$$x_k \geq \mu_k \geq 0 \quad \text{für alle } k \in \mathbb{N}$$

und divergiert $\sum_{k=1}^{\infty} \mu_k$, so divergiert auch $\sum_{k=1}^{\infty} x_k$. Die Reihe $\sum_k \mu_k$ heißt *Minorante* von $\sum_k x_k$.

Beweis. Wäre $\sum_k x_k$ konvergent, so wäre $\sum_k \mu_k$ nach Satz 4.2.6 ebenfalls konvergent, Widerspruch! \square

Folgerung 4.2.9

Sind $\sum_{k=1}^{\infty} z_k$ und $\sum_{k=1}^{\infty} \zeta_k$ absolut konvergent und $a, b \in \mathbb{C}$ beliebig, so konvergiert auch die Linearkombination $\sum_{k=1}^{\infty} (az_k + b\zeta_k)$ absolut.

Beweis. Wir beachten

$$|az_k + b\zeta_k| \leq |a||z_k| + |b||\zeta_k| =: \mu_k, \quad k \in \mathbb{N}.$$

Nach Satz 4.1.3 konvergiert $\sum_k \mu_k$, da $\sum_k |z_k|$ und $\sum_k |\zeta_k|$ konvergieren. Satz 4.2.6 liefert die Behauptung. \square

Aus dem Majorantenkriterium lassen sich zwei weitere Sätze ableiten, die sowohl hinreichende als auch notwendige Kriterien enthalten und häufig zur Anwendung kommen.

Satz 4.2.10: Quotientenkriterium

Es sei $\sum_{k=1}^{\infty} z_k$ eine Reihe mit $z_k \neq 0$ für alle $k \in \mathbb{N}$. Dann gilt:

(a) Existiert ein $q \in (0, 1)$ und ein $k_0 \in \mathbb{N}$ mit

$$\left| \frac{z_{k+1}}{z_k} \right| \leq q < 1 \quad \text{für alle } k \geq k_0,$$

so konvergiert die Reihe $\sum_{k=1}^{\infty} z_k$ absolut.

(b) Existiert ein $k_0 \in \mathbb{N}$, so dass gilt

$$\left| \frac{z_{k+1}}{z_k} \right| \geq 1 \quad \text{für alle } k \geq k_0,$$

dann divergiert die Reihe.

Beweis. (a) Es gilt $|z_{k+1}| \leq q|z_k|$ und folglich

$$|z_k| \leq (|z_{k_0}|q^{-k_0})q^k \quad \text{für alle } k \geq k_0,$$

wie man leicht mit vollständiger Induktion zeigt. Nach Satz 4.2.6 konvergiert also die Reihe $\sum_k z_k$ absolut, da sie (ab dem k_0 -ten Glied) die gemäß Beispiel 4.1.4 (a) konvergente Majorante

$$\sum_{k=k_0}^{\infty} (|z_{k_0}|q^{-k_0})q^k = (|z_{k_0}|q^{-k_0}) \sum_{k=k_0}^{\infty} q^k$$

besitzt.

(b) Nach Voraussetzung gilt $|z_k| \geq |z_{k_0}| > 0$ für alle $k \geq k_0$. Folglich ist $\{z_k\}_k$ keine Nullfolge, nach Satz 4.2.1 ist $\sum_k z_k$ somit divergent. \square

Bemerkung 4.2.11. Wir können in Satz 4.2.10 (a) die Voraussetzung nicht durch die schwächere Bedingung

$$\left| \frac{z_{k+1}}{z_k} \right| < 1 \quad \text{für alle } k \geq k_0$$

ersetzen, wie das Beispiel der divergenten harmonischen Reihe $\sum_{k=1}^{\infty} \frac{1}{k}$ zeigt.

Umgekehrt ist die dort angegebene Bedingung aber auch keine notwendige Bedingung, denn z.B. $\sum_{k=1}^{\infty} \frac{1}{k^2}$ konvergiert gemäß Beispiel 4.2.7, aber es gilt

$$\left| \frac{z_{k+1}}{z_k} \right| = \frac{k^2}{(k+1)^2} \rightarrow 1 \quad (k \rightarrow \infty).$$

Beispiel 4.2.12. (a) Die Reihe $\sum_{k=1}^{\infty} \frac{k^2}{2^k}$ konvergiert, denn mit $x_k := \frac{k^2}{2^k}$ haben wir

$$\left| \frac{x_{k+1}}{x_k} \right| = \frac{(k+1)^2}{2^{k+1}} \frac{2^k}{k^2} = \frac{(k+1)^2}{2k^2} = \frac{1}{2} \left(1 + \frac{1}{k} \right)^2 \leq \frac{8}{9} < 1 \quad \text{für alle } k \geq 3.$$

Satz 4.2.10 (a) liefert die Behauptung.

(b) Die Reihe $\sum_{k=1}^{\infty} \frac{k^k}{k!}$ divergiert, denn mit $x_k = \frac{k^k}{k!}$ gilt

$$\left| \frac{x_{k+1}}{x_k} \right| = \frac{(k+1)^{k+1}}{(k+1)!} \frac{k!}{k^k} = \left(\frac{k+1}{k} \right)^k \geq 1 \quad \text{für alle } k \in \mathbb{N},$$

wir können also Satz 4.2.10 (b) anwenden.

- (c) Die (*komplexe*) *Exponentialreihe* $\sum_{k=0}^{\infty} \frac{z^k}{k!}$ konvergiert absolut für beliebiges $z \in \mathbb{C}$. Mit $z_k := \frac{z^k}{k!}$ gilt nämlich

$$\left| \frac{z_{k+1}}{z_k} \right| = \frac{|z|^{k+1}}{(k+1)!} \frac{k!}{|z|^k} = \frac{|z|}{k+1} \leq \frac{1}{2} \quad \text{für alle } k \geq 2|z| - 1.$$

Satz 4.2.13: Wurzelkriterium

Sei $\sum_{k=1}^{\infty} z_k$ eine beliebige Reihe. Dann gilt:

- (a) Existiert ein $q \in (0, 1)$ und ein $k_0 \in \mathbb{N}$ mit

$$\sqrt[k]{|z_k|} \leq q < 1 \quad \text{für alle } k \geq k_0,$$

so konvergiert $\sum_{k=1}^{\infty} z_k$ absolut.

- (b) Gilt

$$\limsup_{k \rightarrow \infty} \sqrt[k]{|z_k|} > 1,$$

so divergiert die Reihe.

Beweis. → Übungsaufgabe

□

Beispiel 4.2.14. Für die Reihe $\sum_{k=1}^{\infty} k^{\alpha k}$ mit $\alpha \in \mathbb{Z} \setminus \{0\}$ gilt

$$\sqrt[k]{|k^{\alpha k}|} = k^{\alpha} \rightarrow \begin{cases} 0, & \text{falls } \alpha < 0 \\ +\infty, & \text{falls } \alpha > 0 \end{cases} \quad (k \rightarrow \infty).$$

Satz 4.2.13 liefert also Konvergenz für $\alpha \in \mathbb{N}$ und Divergenz für $\alpha \in \mathbb{Z} \setminus \mathbb{N}$.

4.3. Reelle Reihen, Cauchysche Produktformel und Umordnungen

In diesem Abschnitt geben wir wichtige spezifische Resultate an, die - aus unterschiedlichen Gründen - nicht in das Raster des letzten Abschnitts passen. Wir beginnen mit zwei Konvergenzkriterien für den Spezialfall *reeller Reihen*, d.h. $\sum_k x_k$ mit $\{x_k\}_k \subset \mathbb{R}$.

Für reelle Reihen mit nichtnegativen Einträgen gilt das folgende äquivalente Konvergenzkriterium:

Satz 4.3.1

Die Reihe $\sum_{k=1}^{\infty} x_k$ mit $x_k \in \mathbb{R}$ und $x_k \geq 0$ für alle $k \in \mathbb{N}$ konvergiert genau dann, wenn die zugehörige Folge der Partialsummen beschränkt ist.

Bemerkung 4.3.2. Eine Reihe $\sum_{k=1}^{\infty} z_k$ nennen wir *beschränkt*, wenn die zugehörige Folge der Partialsummen beschränkt ist; wir schreiben dann auch

$$\left| \sum_{k=1}^{\infty} z_k \right| < +\infty.$$

Jede konvergente Reihe ist beschränkt gemäß Hilfssatz 3.2.3. Satz 4.3.1 behauptet für den Fall reeller $x_k \geq 0$, $k \in \mathbb{N}$, die Äquivalenz

$$\sum_k x_k \text{ konvergent} \quad \Leftrightarrow \quad \sum_k x_k < +\infty.$$

Die Reihe $\sum_{k=1}^{\infty} (-1)^k$ ist ein Beispiel einer beschränkten, aber nicht konvergenten Reihe.

Beweis von Satz 4.3.1. Wegen $x_k \geq 0$ ist die Folge der Partialsummen $s_n = \sum_{k=1}^n x_k$, $n \in \mathbb{N}$, monoton wachsend. Satz 3.4.9 über die monotone Konvergenz liefert also die Konvergenz der Reihe, wenn wir ihre Beschränktheit voraussetzen. Umgekehrt ist, wie in Bemerkung 4.3.2 erwähnt, jede konvergente Reihe auch beschränkt. \square

Im Falle sogenannter alternierender Reihen haben wir die folgende Aussage:

Satz 4.3.3: LEIBNIZ-Kriterium

Ist $\{x_k\}_k \subset \mathbb{R}$ eine monoton fallende Nullfolge, so konvergiert die Reihe $\sum_{k=1}^{\infty} (-1)^k x_k$.

Beispiel 4.3.4. Die Reihen

$$\sum_{k=1}^{\infty} \frac{(-1)^k}{k} \quad (\text{alternierende harmonische Reihe}),$$

$$\sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} \quad (\text{LEIBNIZreihe})$$

konvergieren nach Satz 4.3.3. Wir werden später für deren Werte berechnen

$$\sum_{k=1}^{\infty} \frac{(-1)^k}{k} = -\log 2, \quad \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} = \frac{\pi}{4}.$$

Der Beweis von Satz 4.3.3 benötigt etwas feinere Analysis, nämlich den folgenden

Hilfssatz 4.3.5

Sei $\{z_k\}_{k \in \mathbb{N}} \subset \mathbb{C}$ eine Folge, so dass $\sum_{k=1}^{\infty} z_k$ beschränkt ist, und sei $\{a_k\}_k \subset \mathbb{R}$ eine monoton fallende Nullfolge. Dann konvergiert die Reihe $\sum_{k=1}^{\infty} a_k z_k$.

Beweis. Wir wollen Satz 4.1.5 anwenden, also den Ausdruck $|\sum_{k=m+1}^n a_k z_k|$ für hinreichend große $m, n \in \mathbb{N}$ mit $n > m$ beliebig klein bekommen. Dazu setzen wir $s_n := \sum_{k=m+1}^n z_k$ für festes $m \in \mathbb{N}$ und beliebiges $n > m$. Mit vollständiger Induktion über n zeigt man dann leicht die Relation

$$\sum_{k=m+1}^n a_k z_k = s_n a_n + \sum_{k=m+1}^{n-1} s_k (a_k - a_{k+1}) \quad \text{für alle } n > m.$$

Da $\{a_k\}_k$ eine monoton fallende Nullfolge ist, gilt $a_k \geq a_{k+1} \geq 0$ für alle $k \in \mathbb{N}$. Daher können wir abschätzen

$$\begin{aligned} \left| \sum_{k=m+1}^n a_k z_k \right| &\leq |s_n| a_n + \sum_{k=m+1}^{n-1} |s_k| (a_k - a_{k+1}) \\ &\leq \max \{|s_{m+1}|, \dots, |s_n|\} \left[a_n + \sum_{k=m+1}^{n-1} (a_k - a_{k+1}) \right] \\ &= a_{m+1} \cdot \max \{|s_{m+1}|, \dots, |s_n|\} \quad \text{für } n > m. \end{aligned}$$

Sei nun $\varepsilon > 0$ beliebig. Wegen $a_k \rightarrow 0$ ($k \rightarrow \infty$) existiert ein $N(\varepsilon) \in \mathbb{N}$, so dass $0 \leq a_k < \varepsilon$ für alle $k \geq N(\varepsilon)$ richtig ist. Ferner gibt es ein $c > 0$ mit $|s_n| \leq c$ für alle $n > m$, denn $\{s_n\}_{n>m}$ ist beschränkt nach Voraussetzung. Insgesamt folgt

$$\left| \sum_{k=m+1}^n a_k z_k \right| < c\varepsilon \quad \text{für alle } n > m \geq N(\varepsilon),$$

also nach Satz 4.1.5 die Konvergenz der Reihe. □

Beweis von Satz 4.3.3. Sofort aus Hilfssatz 4.3.5 mit $z_k = (-1)^k$ (denn $|\sum_k (-1)^k| < +\infty$) und $a_k = x_k$, $k \in \mathbb{N}$. □

Bisher haben wir Produkte von Reihen bei unseren Betrachtungen ausgelassen. Zwar konvergiert das Produkt zweier konvergenter Reihen nach Satz 3.3.1 (a), aber die Möglichkeit der Darstellung als *eine* Reihe ist nicht so offensichtlich wie bei Summen (siehe Satz 4.1.3 (a)). Hier hilft der

Satz 4.3.6: CAUCHYSche Produktformel

Es seien $\sum_{k=1}^{\infty} z_k, \sum_{l=1}^{\infty} \zeta_l$ absolut konvergente Reihen. Setzen wir

$$c_j := \sum_{k=1}^j z_k \zeta_{j-k+1}, \quad j \in \mathbb{N},$$

so konvergiert auch die Reihe $\sum_{j=1}^{\infty} c_j$ absolut, und es gilt die *CAUCHYSche Produktformel*

$$\left(\sum_{k=1}^{\infty} z_k \right) \left(\sum_{l=1}^{\infty} \zeta_l \right) = \sum_{j=1}^{\infty} c_j.$$

Beweis. Wir illustrieren hier nur die Idee und schieben den Beweis in die *Ergänzungen*.

Das Produkt zweier endlicher Summen $\sum_{k=1}^n z_k, \sum_{l=1}^m \zeta_l$ lässt sich als Summe mit Summanden der Form $z_k \zeta_l$ schreiben. Diese erwartet man auch beim Produkt zweier Reihen, denn sie sind Grenzwerte ihrer Partialsummen. Das CAUCHY-Produkt zählt diese Summanden entlang endlicher Diagonalen ab:

	ζ_1	ζ_2	ζ_3	ζ_4	\dots
z_1	$z_1 \zeta_1$	$z_1 \zeta_2$	$z_1 \zeta_3$	$z_1 \zeta_4$	
z_2	$z_2 \zeta_1$	$z_2 \zeta_2$	$z_2 \zeta_3$		\ddots
z_3	$z_3 \zeta_1$	$z_3 \zeta_2$		\ddots	
z_4	$z_4 \zeta_1$		\ddots		
\vdots		\ddots			

Dabei sind die eingefärbten Produkte die Summanden von $c_1, c_2, c_3, c_4, \dots$ □

Um schließlich das Verhalten von Reihen unter Umordnungen zu erläutern, beginnen wir mit der

Definition 4.3.7

Zu einer Reihe $\sum_{k=1}^{\infty} z_k$ heißt $\sum_{l=1}^{\infty} \zeta_l$ eine *Umordnung*, wenn es eine bijektive Abbildung $\sigma : \mathbb{N} \rightarrow \mathbb{N}$ so gibt, dass gilt

$$\zeta_l = z_{\sigma(l)} \quad \text{für alle } l \in \mathbb{N}.$$

σ heißt (*unendliche*) *Permutation* der Reihenglieder.

Für endliche Summen in \mathbb{C} ist die Reihenfolge der Summation irrelevant für das Ergebnis;

jede Umordnung einer endlichen Summe liefert also den gleichen Wert. Wie verhält es sich aber bei unendlichen Reihen?

Definition 4.3.8

Wir nennen eine konvergente Reihe $\sum_{k=1}^{\infty} z_k$ *unbedingt konvergent*, wenn jede ihrer Umordnungen ebenfalls konvergiert und den gleichen Wert wie die ursprüngliche Reihe besitzt. Anderenfalls heißt $\sum_{k=1}^{\infty} z_k$ *bedingt konvergent*.

Satz 4.3.9: RIEMANNscher Umordnungssatz

Ist die reelle Reihe $\sum_{k=1}^{\infty} x_k$ konvergent, aber nicht absolut konvergent, so gibt es zu jedem $t \in \mathbb{R}$ eine Umordnung $\sum_{l=1}^{\infty} \xi_l$ der Reihe, so dass $t = \sum_{l=1}^{\infty} \xi_l$ gilt. Folglich ist $\sum_k x_k$ bedingt konvergent.

Beweis. → Ergänzungen

□

Bemerkung 4.3.10. Man kann sogar Umordnungen $\sum_l \xi_l$ einer beliebigen reellen konvergenten, aber nicht absolut konvergenten Reihe konstruieren, so dass $\sum_l \xi_l = \pm\infty$ gilt.

Satz 4.3.11: DIRICHLETScher Umordnungssatz

Eine komplexe konvergente Reihe ist genau dann unbedingt konvergent, wenn sie absolut konvergent ist.

Beweis. ▶ „ \Rightarrow “: Ist $\sum_k z_k$ eine komplexe, unbedingt konvergente Reihe, so sind auch $\sum_k \operatorname{Re}(z_k)$ und $\sum_k \operatorname{Im}(z_k)$ unbedingt konvergent. Nach Satz 4.3.9 müssen daher die reellen Reihen absolut konvergent sein, also ist auch $\sum_k z_k$ absolut konvergent.

▶ „ \Leftarrow “: Sei $\sum_{k=1}^{\infty} z_k$ absolut konvergent, d.h. $\sum_{k=1}^{\infty} |z_k| < +\infty$. Dann existiert nach Satz 4.1.5 zu jedem $\varepsilon > 0$ ein $N = N(\varepsilon) \in \mathbb{N}$, so dass

$$\sum_{k=N+1}^{N+p} |z_k| < \frac{\varepsilon}{2} \quad \text{für alle } p \in \mathbb{N} \quad (4.3.1)$$

richtig ist. Ist $\sum_l \zeta_l$ eine beliebige Umordnung von $\sum_k z_k$, so existiert ein $K \in \mathbb{N}$ mit $K \geq N$, so dass gilt

$$\{z_1, \dots, z_N\} \subset \{\zeta_1, \dots, \zeta_K\}. \quad (4.3.2)$$

Es bezeichne $s_n := \sum_{k=1}^n z_k$ bzw. $t_n := \sum_{l=1}^n \zeta_l$ die n -ten Partialsummen der beiden Reihen. Aus (4.3.1) und (4.3.2) folgt nun

$$|t_n - s_n| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \quad \text{für alle } n > K \geq N,$$

da sich die Terme z_1, \dots, z_N aufheben und die übrigen Terme in den Summen s_n bzw. t_n jeweils durch (4.3.1) abgeschätzt werden können. Es folgt also $\lim_{n \rightarrow \infty} |t_n - s_n| = 0$. Bezeichnet nun s den Wert von $\sum_k z_k$, so folgt

$$|t_n - s| \leq |t_n - s_n| + |s_n - s| \rightarrow 0 \quad (n \rightarrow \infty),$$

d.h. auch $\sum_k \zeta_k$ konvergiert gegen s , wie behauptet. \square

4.4. Potenzreihen

Eine wichtige Klasse von Reihen sind die *Potenzreihen*, denen wir deshalb einen eigenen Abschnitt widmen:

Definition 4.4.1

Zu einer Folge $\{a_k\}_{k \in \mathbb{N}_0} \subset \mathbb{C}$ und einem $z \in \mathbb{C}$ erklären wir die *Potenzreihe*

$$\mathcal{P}(z) := \sum_{k=0}^{\infty} a_k z^k.$$

Die a_k werden auch *Koeffizienten* von \mathcal{P} genannt.

Bemerkung 4.4.2. ▶ Gibt es eine Menge $M \subset \mathbb{C}$, so dass $\mathcal{P}(z)$ für alle $z \in M$ konvergiert, so stellt $\mathcal{P} : M \rightarrow \mathbb{C}$ eine Abbildung dar.

▶ Die Partialsummen

$$P_n(z) := \sum_{k=0}^n a_k z^k, \quad n \in \mathbb{N}_0,$$

einer Potenzreihe sind für alle $n \in \mathbb{N}_0$ *komplexe Polynome* $\mathcal{P}_n : \mathbb{C} \rightarrow \mathbb{C}$. Die Folge $\{\mathcal{P}_n\}_{n \in \mathbb{N}_0}$ ist also eine *Funktionenfolge* in \mathbb{C} . Wir werden Eigenschaften von Funktionenfolgen ab Kapitel 6 allgemein diskutieren.

Beispiel 4.4.3. Wir kennen bereits zwei Beispiele von Potenzreihen, welche sicher auch die beiden wichtigsten sind:

▶ *Geometrische Reihe*: $\sum_{k=0}^{\infty} z^k$, d.h. $a_k = 1$ für alle $k \in \mathbb{N}_0$.

- **Exponentialreihe:** $\sum_{k=0}^{\infty} \frac{z^k}{k!}$, d.h. $a_k = \frac{1}{k!}$ für alle $k \in \mathbb{N}_0$.

Während letztere nach Beispiel 4.2.12 (c) für alle $z \in \mathbb{C}$ absolut konvergiert, konvergiert die geometrische Reihe absolut für $|z| < 1$ und divergiert für $|z| \geq 1$, siehe Beispiele 4.1.4 (a) und 4.2.2 (b). I.A. hängt also das Konvergenzverhalten einer gegebenen Potenzreihe von der Wahl der Variablen z ab. Genauer haben wir den folgenden

Satz 4.4.4: CAUCHY-HADAMARD

Für eine Potenzreihe $\mathcal{P}(z) = \sum_{k=0}^{\infty} a_k z^k$ setzen wir $\alpha := \limsup_{k \rightarrow \infty} \sqrt[k]{|a_k|} \in [0, +\infty) \cup \{+\infty\}$. Erklären wir dann

$$R := \begin{cases} +\infty, & \text{falls } \alpha = 0 \\ \alpha^{-1}, & \text{falls } \alpha \in (0, +\infty) \\ 0, & \text{falls } \alpha = +\infty \end{cases}, \quad (4.4.1)$$

so konvergiert $\mathcal{P}(z)$ für $|z| < R$ absolut und divergiert für $|z| > R$.

Bemerkung 4.4.5. (a) Zur Erinnerung: Für nach oben unbeschränkte Folgen haben wir

$\limsup_{k \rightarrow \infty} x_k = +\infty$ gesetzt; siehe Bemerkung 3.6.10 (a).

(b) Die in (4.4.1) erklärte Größe $R \in [0, +\infty) \cup \{+\infty\}$ heißt **Konvergenzradius** der Reihe $\mathcal{P}(z)$. Die Kreisscheibe $K_R(0) := \{z \in \mathbb{C} : |z| < R\}$ nennen wir das **Konvergenzgebiet**.

(c) Wir setzen folgende Regeln für das Rechnen mit $+\infty$ fest:

$$\frac{1}{0} = +\infty, \quad \frac{1}{+\infty} = 0, \quad (+\infty) \cdot x = +\infty \quad \text{für } x \in \mathbb{R} \text{ mit } x > 0.$$

Dann liest sich (4.4.1) als $R = \frac{1}{\alpha}$.

Beweis von Satz 4.4.4. Offenbar gilt $|z| < R$ (bzw. $|z| > R$) genau dann, wenn

$$|z|\alpha = |z| \limsup_{k \rightarrow \infty} \sqrt[k]{|a_k|} = \limsup_{k \rightarrow \infty} \sqrt[k]{|a_k z^k|} < 1 \quad (\text{bzw. } > 1)$$

richtig ist. Wegen Satz 3.6.7 ist für $\limsup_{k \rightarrow \infty} \sqrt[k]{|a_k z^k|} < 1$ der Fall (a) aus Satz 4.2.13 gültig, d.h. $\sum_k a_k z^k$ konvergiert absolut. Im Fall $\limsup_{k \rightarrow \infty} \sqrt[k]{|a_k z^k|} > 1$ divergiert $\mathcal{P}(z) = \sum_k a_k z^k$ nach Satz 4.2.13 (b). \square

Als sehr praktisch erweist sich der folgende

Satz 4.4.6

Ist $\mathcal{P}(z) = \sum_{k=0}^{\infty} a_k z^k$ im Punkt $z_0 \in \mathbb{C} \setminus \{0\}$ konvergent, so konvergiert $\mathcal{P}(z)$ absolut für alle $z \in \mathbb{C}$ mit $|z| < |z_0|$.

Beweis. Da $\sum_k a_k z_0^k$ konvergiert, ist $\{a_k z_0^k\}_k$ Nullfolge und somit beschränkt; es gibt also ein $c > 0$ mit $|a_k z_0^k| \leq c$ für alle $k \in \mathbb{N}_0$. Sei nun $z \in \mathbb{C}$ mit $|z| < |z_0|$ bzw. $q := \left|\frac{z}{z_0}\right| \in [0, 1)$ beliebig gewählt. Dann folgt

$$|a_k z^k| = |a_k z_0^k| \left|\frac{z}{z_0}\right|^k \leq c q^k \quad \text{für alle } k \in \mathbb{N}_0.$$

Wegen $q \in [0, 1)$ konvergiert die Reihe $\sum_k c q^k = c \sum_k q^k = \frac{c}{1-q}$ und nach dem Majorantenkriterium, Satz 4.2.6, konvergiert $\mathcal{P}(z)$ absolut für $|z| < |z_0|$, wie behauptet. \square

Der nächste Satz macht u.a. Aussagen über das Konvergenzverhalten auf dem *Rand* des Konvergenzgebietes und enthält das Leibniz-Kriterium, Satz 4.3.3, als Spezialfall:

Satz 4.4.7

Ist $\{a_k\}_{k \in \mathbb{N}_0} \subset \mathbb{R}$ eine monoton fallende Nullfolge, so konvergiert die Potenzreihe $\mathcal{P}(z) = \sum_{k=0}^{\infty} a_k z^k$ für alle $z \in \mathbb{C} \setminus \{1\}$ mit $|z| \leq 1$.

Beweis. Aus der Summenformel der geometrischen Reihe, Satz 2.4.8, erhalten wir für beliebige $z \in \mathbb{C} \setminus \{1\}$ mit $|z| \leq 1$:

$$\left| \sum_{k=0}^n z^k \right| = \left| \frac{1 - z^{n+1}}{1 - z} \right| \leq \frac{1 + |z|^{n+1}}{|1 - z|} \leq \frac{2}{|1 - z|} \quad \text{für alle } n \in \mathbb{N}.$$

Somit ist $\sum_k z^k$ für solche z beschränkt, Hilfssatz 4.3.5 liefert also die Konvergenz von $\mathcal{P}(z)$. \square

Abschließend geben wir noch die zu Satz 4.3.6 analoge Aussage für Potenzreihen an:

Satz 4.4.8: CAUCHYSche Produktformel für Potenzreihen

Sind die Potenzreihen $\sum_{k=0}^{\infty} a_k z^k$ und $\sum_{k=0}^{\infty} b_k z^k$ für $|z| < R$ absolut konvergent, so gilt dies auch für die Potenzreihe $\sum_{k=0}^{\infty} c_k z^k$ mit den Koeffizienten

$$c_k := \sum_{l=0}^k a_l b_{k-l}, \quad k \in \mathbb{N}_0,$$

und wir haben die Identität

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

Beweis. Dies ergibt sich nach Indexverschiebung $k \rightarrow k + 1$ aus Satz 4.3.6, wenn man noch

$$\sum_{l=0}^k (a_l z^l) (b_{k-l} z^{k-l}) = \left(\sum_{l=0}^k a_l b_{k-l} \right) z^k = c_k z^k$$

beachtet. □

Bemerkung 4.4.9. Alle Resultate lassen sich direkt auf komplexe Potenzreihen der Form

$$\mathcal{P}_{z_0}(z) := \sum_{k=0}^{\infty} a_k (z - z_0)^k$$

übertragen. Das Konvergenzgebiet von $\mathcal{P}_{z_0}(z)$ ist dann eine Kreisscheibe $K_R(z_0) = \{z \in \mathbb{C} : |z - z_0| < R\}$ vom Radius $R \in [0, +\infty) \cup \{+\infty\}$ um den *Entwicklungspunkt* $z_0 \in \mathbb{C}$. Die bisher betrachteten Potenzreihen $\mathcal{P}(z)$ sind also Spezialfälle von $\mathcal{P}_{z_0}(z)$ mit $z_0 = 0$.

Kapitel 5.

Topologische Grundlagen im \mathbb{R}^d

Was man in der Mathematik unter „Topologie“ zu verstehen hat, ist nicht ganz so leicht zu fassen (schauen Sie einmal in den Wikipedia-Artikel dazu); es hat etwas mit Form und Invarianz zu tun. Für uns handelt es sich um spezielle Eigenschaften von Mengen und deren Punkten, die wir i.F. häufig verwenden werden. Obgleich wir viele Begriffe und Aussagen wortgleich auch für Teilmengen beliebiger normierter Vektorräume definieren könnten, beschränken wir uns weitestgehend auf den Euklidischen Raum \mathbb{R}^d ; in unseren Vorlesungen werden wir nämlich Abbildungen zwischen Teilmengen des \mathbb{R}^d untersuchen.

In Abschnitt A.5.1 führen wir die grundlegenden Begriffe ein, in Abschnitt A.5.2 diskutieren wir Folgen und deren Grenzwerte im \mathbb{R}^d . In Abschnitt A.5.3 werden wir spezielle Klassen von Mengen kennenlernen. Schließlich klassifizieren wir in Abschnitt A.5.4 Punkte von (bzw. in Bezug zu) Mengen.

5.1. Der Euklidische Vektorraum \mathbb{R}^d

In Beispiel 1.4.3 hatten wir zu einem $d \in \mathbb{N}$ den $\mathbb{R}^d = \mathbb{R} \times \dots \times \mathbb{R}$ als d -faches kartesisches Produkt mit sich selbst definiert. Zwar besitzt der \mathbb{R}^d für $d \geq 3$ keine Körperstruktur mehr, er lässt sich aber zu einem *Vektorraum* machen:

Definition 5.1.1

Für Elemente $x = (x_1, \dots, x_d)$, $y = (y_1, \dots, y_d)$ des \mathbb{R}^d und $\lambda \in \mathbb{R}$ erklären wir die Rechenoperationen:

- ▶ *Addition* $+: \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ durch $x + y := (x_1 + y_1, \dots, x_d + y_d)$,
- ▶ *skalare Multiplikation* $\cdot: \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ durch $\lambda \cdot x = \lambda x := (\lambda x_1, \dots, \lambda x_d)$.

Das Tripel $(\mathbb{R}^d, +, \cdot)$ heißt (*d-dimensionaler*) *Koordinatenraum* oder *Standardvektorraum*; wir schreiben auch kurz \mathbb{R}^d für diesen Raum.

Bemerkung 5.1.2. Der \mathbb{R}^d ist ein d -dimensionaler Vektorraum, die folgenden *Vektorraumaxiome* sind offenbar erfüllt:

- ▶ (A1)-(A4) aus Definition 2.2.1 mit dem neutralen Element $0 = (0, \dots, 0) \in \mathbb{R}^d$ und dem negativen Element $-x := (-x_1, \dots, -x_d) \in \mathbb{R}^d$.
- ▶ $\lambda(\mu x) = (\lambda\mu)x$ für alle $\lambda, \mu \in \mathbb{R}, x \in \mathbb{R}^d$.
- ▶ $\lambda(x + y) = \lambda x + \lambda y, (\lambda + \mu)x = \lambda x + \mu x$ für alle $\lambda, \mu \in \mathbb{R}, x, y \in \mathbb{R}^d$.
- ▶ $1x = x$ für alle $x \in \mathbb{R}^d$.

Wir sprechen bei Elementen $x \in \mathbb{R}^d$ gleichwertig von *Punkten* oder *Vektoren*.

Im \mathbb{R}^d kann auch ein Produkt zweier Vektoren erklärt werden:

Definition 5.1.3

Seien $x, y \in \mathbb{R}^d$, so erklären wir deren (*euklidisches*) *Skalarprodukt* oder auch *inneres Produkt* als

$$\langle x, y \rangle = \sum_{j=1}^d x_j y_j. \quad (5.1.1)$$

$(\mathbb{R}^d, \langle \cdot, \cdot \rangle)$ heißt *euklidischer Vektorraum*; wir schreiben auch hier wieder kurz \mathbb{R}^d .

Bemerkung 5.1.4. Allgemein bezeichnet man ein Paar $(V, \langle \cdot, \cdot \rangle)$ aus einem reellen Vektorraum und einer Abbildung $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{R}$ als *euklidischen Vektorraum* oder *Skalarproduktraum*, wenn folgende Eigenschaften erfüllt sind:

$$\langle x, x \rangle \geq 0, \quad \langle x, x \rangle = 0 \Leftrightarrow x = 0 \quad (\text{Positivität}) \quad (5.1.2)$$

$$\langle x, y \rangle = \langle y, x \rangle \quad (\text{Symmetrie}) \quad (5.1.3)$$

$$\langle \lambda x + \mu y, z \rangle = \lambda \langle x, z \rangle + \mu \langle y, z \rangle \quad (\text{Bilinearität}) \quad (5.1.4)$$

für alle $x, y, z \in V$ und $\lambda, \mu \in \mathbb{R}$. Dabei implizieren (5.1.3) und (5.1.4) auch Linearität im zweiten Argument:

$$\langle x, \lambda y + \mu z \rangle = \lambda \langle x, y \rangle + \mu \langle x, z \rangle. \quad (5.1.5)$$

Offenbar gelten (5.1.2)-(5.1.4) für das in (5.1.1) erklärte euklidische Skalarprodukt.

Alternative Bezeichnungen: Ein Skalarprodukt ist eine *symmetrische, positiv definite Bilinearform*. Im Fall unendlich-dimensionaler Vektorräume V spricht man statt von euklidischen Vektorräumen eher von *Prähilberträumen*.

Satz 5.1.5: CAUCHY-SCHWARZsche Ungleichung

Sind $x, y \in \mathbb{R}^d$ beliebig, so gilt

$$\langle x, y \rangle^2 \leq \langle x, x \rangle \langle y, y \rangle. \quad (5.1.6)$$

Gleichheit tritt genau dann ein, wenn x, y linear abhängig sind.

Bemerkung 5.1.6. Wir verwenden im Beweis ausschließlich die Eigenschaften (5.1.2)-(5.1.5), Satz 5.1.5 gilt also in beliebigen Skalarprodukträumen.

Beweis von Satz 5.1.5. ▶ Für $y = 0$ verschwinden beide Seiten von (5.1.6); die rechte Seite gemäß (5.1.2) und für die linke Seite rechnen wir

$$\langle x, 0 \rangle = \langle x, 0 \cdot x \rangle \stackrel{(5.1.5)}{=} 0 \cdot \langle x, x \rangle = 0.$$

▶ Für $y \neq 0$ beachten wir mit beliebigem $t \in \mathbb{R}$:

$$\begin{aligned} 0 & \stackrel{(5.1.2)}{\leq} \langle x - ty, x - ty \rangle \\ & \stackrel{(5.1.3)-(5.1.5)}{=} \langle x, x \rangle - 2t \langle x, y \rangle + t^2 \langle y, y \rangle. \end{aligned} \quad (5.1.7)$$

Wir beachten $\langle y, y \rangle > 0$ und setzen speziell $t = \frac{\langle x, y \rangle}{\langle y, y \rangle}$ ein. Dann folgt

$$0 \leq \langle x, x \rangle - 2 \frac{\langle x, y \rangle^2}{\langle y, y \rangle} + \frac{\langle x, y \rangle^2}{\langle y, y \rangle} = \langle x, x \rangle - \frac{\langle x, y \rangle^2}{\langle y, y \rangle},$$

also nach Umstellen die Behauptung.

▶ **Gleichheitsfall.** Für $y = 0$ haben wir dies schon gesehen (denn jedes $x \in \mathbb{R}^d$ ist linear abhängig von $y = 0$). Für $y \neq 0$ bemerken wir:

- ▷ „ \Rightarrow “: Gilt die Gleichheit in (5.1.6), so folgt mit $t = \frac{\langle x, y \rangle}{\langle y, y \rangle}$ in (5.1.7) die Gleichheit. Die Positivität des Skalarprodukts liefert dann $x - ty = 0$.
- ▷ „ \Leftarrow “: Sind umgekehrt $x = ty$ linear abhängig, so liefern (5.1.4) und (5.1.5) sofort die Gleichheit in (5.1.6).

Damit ist alles gezeigt. □

Wir erklären nun einen „natürlichen“ Längen- und Abstandsbegriff im \mathbb{R}^d :

Definition 5.1.7

Die (*euklidische*) *Norm* von $x \in \mathbb{R}^d$ definieren wir als

$$\|x\| := \sqrt{\langle x, x \rangle} = \left(\sum_{j=1}^d x_j^2 \right)^{\frac{1}{2}}.$$

Bemerkung 5.1.8. ▶ Die euklidische Norm ist die durch das Skalarprodukt (5.1.1) induzierte Norm. Zu zwei Punkten $x, y \in \mathbb{R}^d$ nennt man $\|x - y\|$ ihren (*euklidischen*) *Abstand*. Wir werden i.F. einfach von *Norm* und *Abstand* sprechen, erwähnen aber, dass es im \mathbb{R}^d viele weitere Normen gibt.

▶ In \mathbb{R}^1 bzw. \mathbb{R}^2 entspricht die Norm den in \mathbb{R} bzw. \mathbb{C} erklärten Beträgen.

Die Norm $\|\cdot\|$ hat dem Betrag in bewerteten Körpern sehr ähnliche Eigenschaften (vgl. Definition 2.6.1):

Satz 5.1.9

Für alle $x, y \in \mathbb{R}^d$ und $\lambda \in \mathbb{R}$ gilt

- (a) $\|x\| \geq 0$ und $\|x\| = 0 \Leftrightarrow x = 0$.
- (b) $\|\lambda x\| = |\lambda| \|x\|$.
- (c) $\|x + y\| \leq \|x\| + \|y\|$ (*Dreiecksungleichung*).

Beweis. Wir nutzen die Beziehung $\|\cdot\| = \sqrt{\langle \cdot, \cdot \rangle}$ und die Eigenschaften (5.1.2)-(5.1.6); siehe auch Bemerkung 5.1.10 (c) unten.

(a) Entspricht der Positivität des Skalarprodukts (5.1.2).

(b) Wir rechnen

$$\|\lambda x\| = \sqrt{\langle \lambda x, \lambda x \rangle} \stackrel{(5.1.4), (5.1.5)}{=} \sqrt{\lambda^2 \langle x, x \rangle} = |\lambda| \|x\|.$$

(c) Hier rechnen wir

$$\begin{aligned} \|x + y\|^2 &= \langle x + y, x + y \rangle \stackrel{(5.1.3) - (5.1.5)}{=} \langle x, x \rangle + 2\langle x, y \rangle + \langle y, y \rangle \\ &\stackrel{(5.1.6)}{\leq} \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2, \end{aligned}$$

also nach Wurzelziehen die behauptete Dreiecksungleichung. \square

Bemerkung 5.1.10. (a) Aus der Dreiecksungleichung folgt wie in Satz 2.6.5 (c) noch die *umgekehrte Dreiecksungleichung*

$$\|x - y\| \geq \left| \|x\| - \|y\| \right| \quad \text{für alle } x, y \in \mathbb{R}^d.$$

- (b) Allgemein heit eine Abbildung $\|\cdot\| : V \rightarrow \mathbb{R}$ auf einem reellen Vektorraum V *Norm*, wenn die Eigenschaften (a)-(c) aus Satz 5.1.9 fr alle $x, y \in V$ und $\lambda \in \mathbb{R}$ erfllt sind, das Tupel $(V, \|\cdot\|)$ wird *normierter Vektorraum* genannt. Satz 5.1.9 besagt demnach, dass \mathbb{R}^d mit der euklidische Norm einen normierten Vektorraum darstellt.
- (c) Allgemein lsst sich auf jedem Skalarproduktraum $(V, \langle \cdot, \cdot \rangle)$ durch $\|\cdot\| := \sqrt{\langle \cdot, \cdot \rangle}$ die *induzierte Norm* erklren. Der Beweis von Satz 5.1.2 zeigt dann: Jeder Skalarproduktraum ist auch normiert.

5.2. Folgen im \mathbb{R}^d

Die Verallgemeinerung des Betrages durch die Norm erlaubt uns, die bekannten Begriffe fr Folgen in bewerteten Krpern auf den \mathbb{R}^d zu bertragen:

Definition 5.2.1

Eine Folge $\{x_n\}_{n \in \mathbb{N}} \subset \mathbb{R}^d$ mit den Gliedern $x_n = (x_{n1}, \dots, x_{nd}) \in \mathbb{R}^d$, $n \in \mathbb{N}$, heit:

- *konvergent* gegen den *Grenzwert* oder *Limes* $x \in \mathbb{R}^d$, wenn es fr alle $\varepsilon > 0$ ein $N(\varepsilon) \in \mathbb{N}$ gibt mit

$$\|x_n - x\| < \varepsilon \quad \text{fr alle } n \geq N(\varepsilon).$$

Schreibweisen: $\lim_{n \rightarrow \infty} x_n = x$ oder $x_n \rightarrow x$ ($n \rightarrow \infty$). Wenn $\{x_n\}_n$ nicht konvergiert, nennen wir die Folge *divergent*.

- *CAUCHYfolge*, wenn fr alle (reellen) $\varepsilon > 0$ ein $N(\varepsilon) \in \mathbb{N}$ existiert mit

$$\|x_n - x_m\| < \varepsilon \quad \text{fr alle } m, n \geq N(\varepsilon).$$

- *Nullfolge*, wenn $\{x_n\}_n$ gegen $0 \in \mathbb{R}^d$ konvergiert.
- *beschrnkt*, falls ein reelles $c > 0$ existiert mit $\|x_n\| \leq c$ fr alle $n \in \mathbb{N}$.

Ferner heit $y \in \mathbb{R}^d$ *Hufungswert* von $\{x_n\}_n$, wenn eine Teilfolge $\{x_{n_k}\}_k \subset \{x_n\}_n$ existiert mit $\lim_{k \rightarrow \infty} x_{n_k} = y$.

Geometrische Interpretation

Z.B. ist $\{x_n\}_n \subset \mathbb{R}^d$ gegen $x \in \mathbb{R}^d$ konvergent, wenn in jeder ε -Umgebung

$$U_\varepsilon(x) := \{y \in \mathbb{R}^d : \|y - x\| < \varepsilon\}$$

alle bis auf endliche viele Glieder der Folge liegen. Fr $d = 3$ ist $U_\varepsilon(x)$ eine Kugel (englisch: *Ball*). Wir schreiben daher auch allgemein $B_\varepsilon(x) := U_\varepsilon(x)$ und nennen dies die (*d-dimensionale*)

Kugel um x vom Radius $\varepsilon > 0$.

Satz 5.2.2: Rechenregeln für Grenzwerte im \mathbb{R}^d

Sind $\{x_n\}_n, \{y_n\}_n \subset \mathbb{R}^d$ konvergente Folgen mit $\lim_{n \rightarrow \infty} x_n = x$ und $\lim_{n \rightarrow \infty} y_n = y$, so folgt:

- (a) Sind $\alpha, \beta \in \mathbb{R}$ beliebig, so konvergiert auch $\{\alpha x_n + \beta y_n\}_n$ mit Grenzwert $\alpha x + \beta y$.
- (b) Es gilt $\langle x_n, y_n \rangle \rightarrow \langle x, y \rangle$ und $\|x_n\| \rightarrow \|x\|$ für $n \rightarrow \infty$.
- (c) Ist $\{\alpha_n\}_n \subset \mathbb{R}$ eine Folge mit $\alpha_n \rightarrow \alpha$ ($n \rightarrow \infty$), so gilt $\alpha_n x_n \rightarrow \alpha x$ ($n \rightarrow \infty$).

Beweis. → Zur Übung. □

Satz 5.2.3: CAUCHYSches Konvergenzkriterium in \mathbb{R}^d

Eine Folge $\{x_n\}_n \subset \mathbb{R}^d$ ist genau dann konvergent, wenn $\{x_n\}_n$ CAUCHYfolge ist.

Der Beweis erfolgt genau wie im Beweis der Vollständigkeit von \mathbb{C} , Satz 3.2.9, indem man die Aussage auf die Koordinatenfolgen $\{x_{nj}\}_n, j = 1, \dots, d$, zurückführt mittels des folgenden

Hilfssatz 5.2.4

Eine Folge $\{x_n\}_n \subset \mathbb{R}^d$ ist genau dann beschränkt (bzw. konvergent, CAUCHYfolge, Nullfolge), wenn alle Koordinatenfolgen $\{x_{nj}\}_n \subset \mathbb{R}, j = 1, \dots, d$, beschränkt (bzw. konvergent, CAUCHYfolgen, Nullfolgen) sind. Für konvergente Folgen $\{x_n\}_n$ gilt

$$\lim_{n \rightarrow \infty} x_n = \left(\lim_{n \rightarrow \infty} x_{n1}, \dots, \lim_{n \rightarrow \infty} x_{nd} \right).$$

Beweis. Der Norm-Definition 5.1.7 entnimmt man sofort: Ist $y = (y_1, \dots, y_d) \in \mathbb{R}^d$ beliebig, so gelten die Ungleichungen

$$|y_j| \leq \|y\| \quad \text{für } j = 1, \dots, d, \quad \|y\| \leq \sum_{k=1}^d |y_k|.$$

Hieraus ergeben sich sofort die Behauptungen. □

Bemerkung 5.2.5. Satz 5.2.3 besagt, dass \mathbb{R}^d vollständig ist: Jede CAUCHYfolge konvergiert im \mathbb{R}^d (die Umkehrung gilt in jedem normierten Vektorraum). Folglich ist der \mathbb{R}^d ein vollständiger normierter Vektorraum (siehe Bemerkung 5.1.10 (b)); solche Räume werden *BANACHräume*

genannt. Genauer ist der \mathbb{R}^d auch ein sogenannter *HILBERTRAUM*, d.h. ein vollständiger Skalarproduktraum (was nun das Synonym „Prähilbertraum“ für Skalarprodukträume erklärt).

Satz 5.2.6: BOLZANO-WEIERSTRASS im \mathbb{R}^d

Jede beschränkte Folge $\{x_n\}_n \subset \mathbb{R}^d$ besitzt eine konvergente Teilfolge.

Beweis. Vollständige Induktion über die Raumdimension $d \in \mathbb{N}$.

- $d = 1$: Das ist die Aussage von Satz 3.4.6.
- $d \rightarrow d + 1$: Die Aussage sei für beschränkte Folgen $\{x_n\}_n \subset \mathbb{R}^d$ mit einem $d \in \mathbb{N}$ erfüllt. Sei $\{\tilde{x}_n\}_n \subset \mathbb{R}^{d+1}$ beschränkt mit den Folgengliedern

$$\tilde{x}_n = (\tilde{x}_{n1}, \dots, \tilde{x}_{nd}, \tilde{x}_{n,d+1}) = (x_n, \xi_n) \quad \text{mit } x_n := (\tilde{x}_{n1}, \dots, \tilde{x}_{nd}), \xi_n := \tilde{x}_{n,d+1}.$$

Damit sind nach Hilfssatz 5.2.4 auch die Folgen $\{x_n\}_n \subset \mathbb{R}^d$ und $\{\xi_n\}_n \subset \mathbb{R}$ beschränkt. Nach Induktionsvoraussetzung existiert also eine konvergente Teilfolge $\{x'_k\}_k = \{x_{n_k}\}_k$ von $\{x_n\}_n$ mit $\lim_{k \rightarrow \infty} x'_k = x \in \mathbb{R}^d$. Die entsprechende Teilfolge $\{\xi'_k\}_k = \{\xi_{n_k}\}_k \subset \mathbb{R}$ von $\{\xi_n\}_n$ muss zwar nicht konvergieren, ist aber sicher beschränkt. Also gibt es nach Satz 3.4.6 eine weitere Teilfolge $\{\xi'_{k_l}\}_l \subset \{\xi'_k\}_k$ mit $\lim_{l \rightarrow \infty} \xi'_{k_l} = \xi \in \mathbb{R}$. Die entsprechende Teilfolge $\{x'_{k_l}\}_l \subset \{x'_k\}_k$ konvergiert auch gegen x , so dass wiederum nach Hilfssatz 5.2.4 $\{\tilde{x}'_{k_l}\}_l = \{(x'_{k_l}, \xi'_{k_l})\}_l$ gegen (x, ξ) konvergiert. \square

Bemerkung 5.2.7. Da der Betrag in \mathbb{C} und die Norm in \mathbb{R}^2 übereinstimmen, liefert Satz 5.2.6 für $d = 2$: Jede beschränkte Folge $\{x_n\}_n \subset \mathbb{C}$ besitzt eine konvergente Teilfolge.

5.3. Offene, abgeschlossene und kompakte Mengen

Wir beginnen mit der

Definition 5.3.1

Eine Teilmenge $M \subset \mathbb{R}^d$ heißt:

- *offen*, wenn zu jedem $x_0 \in M$ ein $r > 0$ so existiert, dass gilt

$$B_r(x_0) = \{x \in \mathbb{R}^d : \|x - x_0\| < r\} \subset M.$$

- *abgeschlossen*, wenn für jede konvergente Folge $\{x_n\}_n \subset M$ gilt

$$x_0 := \lim_{n \rightarrow \infty} x_n \in M.$$

Beispiel 5.3.2. (a) *Intervalle in \mathbb{R} :*

- Das offene Intervall $(a, b) = \{x \in \mathbb{R} : a < x < b\}$ ist im Sinne von Definition 5.3.1 offen (\rightarrow Übungsaufgabe).
- Das abgeschlossene Intervall $[a, b] = \{x \in \mathbb{R} : a \leq x \leq b\}$ ist im Sinne von Definition 5.3.1 abgeschlossen (\rightarrow Übungsaufgabe).
- Das halboffene Intervall $[a, b)$ ist für $a < b$ weder offen noch abgeschlossen.

Denn: Die konvergente Folge $\{b - \frac{1}{n}\}_{n \geq N} \subset [a, b)$ mit hinreichend großem $N \in \mathbb{N}$ hat den Grenzwert $\lim_{n \rightarrow \infty} (b - \frac{1}{n}) = b \notin [a, b)$. Und für $a \in [a, b)$ gilt offenbar $B_r(a) \not\subset [a, b)$ für alle $r > 0$. \square

(b) *Kugeln in \mathbb{R}^d :*

- $B_R(\xi) \subset \mathbb{R}^d$ ist offen für beliebige $R > 0$, $\xi \in \mathbb{R}^d$.

Denn: Ist $x_0 \in B_R(\xi)$ beliebig, so ist $r := R - \|x_0 - \xi\| > 0$. Für beliebiges $x \in B_r(x_0)$ haben wir dann die Abschätzung

$$\|x - \xi\| \leq \|x - x_0\| + \|x_0 - \xi\| < r + \|x_0 - \xi\| = R,$$

also $x \in B_R(\xi)$ und somit $B_r(x_0) \subset B_R(\xi)$. \square

Wir bezeichnen daher $B_R(\xi)$ auch als *offene Kugel* im \mathbb{R}^d .

- Im Gegensatz dazu ist $\hat{B}_R(\xi) := \{x \in \mathbb{R}^d : \|x - \xi\| \leq R\}$ abgeschlossen und heißt *abgeschlossene Kugel* im \mathbb{R}^d .

Denn: Ist $\{x_n\}_n \subset \hat{B}_R(\xi)$ mit $x_n \rightarrow x_0$ ($n \rightarrow \infty$) beliebig, so gilt $\|x_n - \xi\| \leq R$ für alle $n \in \mathbb{N}$. Grenzübergang $n \rightarrow \infty$ liefert dann:

$$R \geq \lim_{n \rightarrow \infty} \|x_n - \xi\| \stackrel{\text{Satz 5.2.2 (b)}}{=} \left\| \lim_{n \rightarrow \infty} (x_n - \xi) \right\| = \|x_0 - \xi\|,$$

also $x_0 \in \hat{B}_R(\xi)$. \square

- Die Kugelschale $S_{\varrho, R}(\xi) := B_R(\xi) \setminus B_{\varrho}(\xi) = \{x \in \mathbb{R}^d : \varrho \leq \|x - \xi\| < R\}$ mit $0 < \varrho < R$ ist weder offen noch abgeschlossen (\rightarrow Übungsaufgabe).

(c) $\mathbb{Q} \subset \mathbb{R}$ ist weder offen noch abgeschlossen (\rightarrow Übungsaufgabe).

(d) \mathbb{R}^d und \emptyset sind die einzigen Teilmengen von \mathbb{R}^d , die sowohl offen als auch abgeschlossen sind.

Wir erinnern an Definition 1.3.1, das *Komplement* einer Menge $M \subset \mathbb{R}^d$:

$$M^c := \mathbb{R}^d \setminus M = \{x \in \mathbb{R}^d : x \notin M\}.$$

Satz 5.3.3

Eine Menge $M \subset \mathbb{R}^d$ ist genau dann offen, wenn ihr Komplement M^c abgeschlossen ist. Weiter ist M genau dann abgeschlossen, wenn M^c offen ist.

Beweis. Es genügt, die erste Aussage zu beweisen. Die zweite folgt dann unmittelbar aus dem Doppelkomplementgesetz $(M^c)^c = M$; siehe Satz 1.3.4 (c).

- „ \Rightarrow “: Sei $M \subset \mathbb{R}^d$ offen. Wäre M^c nicht abgeschlossen, so gäbe es eine konvergente Folge $\{x_n\}_n \subset M^c$ mit $x_n \rightarrow x_0 \notin M^c$ ($n \rightarrow \infty$). Da M offen ist, gäbe es dann zu $x_0 \in M$ ein $r > 0$ mit $B_r(x_0) \subset M$. Dann würde wegen $\|x_n - x_0\| \rightarrow 0$ ($n \rightarrow \infty$) aber $x_n \in B_r(x_0) \subset M$ für hinreichend großes $n \in \mathbb{N}$ gelten, im Widerspruch zu $\{x_n\}_n \subset M^c$. Also ist M^c abgeschlossen.
- „ \Leftarrow “: Sei nun M^c abgeschlossen. Wäre M nicht offen, so gäbe es ein $x_0 \in M$ mit $B_r(x_0) \not\subset M$ für alle $r > 0$. Insbesondere gäbe es zu jedem $n \in \mathbb{N}$ ein $x_n \in B_{\frac{1}{n}}(x_0) \cap M^c$. Für die so bestimmte Folge $\{x_n\}_n \subset M^c$ folgte $x_n \rightarrow x_0$ ($n \rightarrow \infty$). Da M^c abgeschlossen ist, würde $x_0 \in M^c$ folgen, Widerspruch! Also ist M offen. \square

Vereinbarung. Meist werden wir offene Mengen mit dem (ggf. indizierten) Symbol Ω und abgeschlossene Mengen mit A bezeichnen.

Der folgende Satz zeigt, wie man aus gegebenen offenen bzw. abgeschlossenen Mengen neue Mengen diesen Typs konstruieren kann.

Satz 5.3.4

- (a) Sind $\Omega_1, \dots, \Omega_n \subset \mathbb{R}^d$ endlich viele offene Mengen, so ist auch für $\bigcap_{j=1}^n \Omega_j$ offen.
- (b) Sind $A_1, \dots, A_n \subset \mathbb{R}^d$ endlich viele abgeschlossene Mengen, so ist auch $\bigcup_{j=1}^n A_j$ abgeschlossen.
- (c) Ist $\{\Omega_j\}_{j \in J}$ eine Familie offener Mengen mit beliebiger Indexmenge J , so ist auch die Vereinigung $\bigcup_{j \in J} \Omega_j = \{x \in \mathbb{R}^d : x \in \Omega_j \text{ für ein } j \in J\}$ offen.
- (d) Ist $\{A_j\}_{j \in J}$ eine Familie abgeschlossener Mengen mit beliebiger Indexmenge J , so ist auch der Durchschnitt $\bigcap_{j \in J} A_j = \{x \in \mathbb{R}^d : x \in A_j \text{ für alle } j \in J\}$ abgeschlossen.

Beweis. Wegen Satz 5.3.3 und den Regeln von de Morgan (siehe Bemerkung 1.3.6) genügt es die Aussagen (a) und (c) zu beweisen (\rightarrow Übungsaufgabe). \square

Abgeschlossene Mengen sind besonders interessant, wenn sie zusätzlich *beschränkt* sind im folgenden Sinne:

Definition 5.3.5

Eine Teilmenge $M \subset \mathbb{R}^d$ heißt

- ▶ *beschränkt*, falls ein $R > 0$ existiert mit $M \subset \hat{B}_R(0)$; anderenfalls nennen wir M *unbeschränkt*.
- ▶ *kompakt* oder *Kompaktum*, falls M beschränkt und abgeschlossen ist.

Bemerkung 5.3.6. Ist M nichtleer und beschränkt, so ist der *Durchmesser*

$$\text{diam } M := \sup \{ \|x - y\| : x, y \in M \}$$

wohldefiniert, d.h. $\text{diam } M$ ist endlich und eindeutig bestimmt.

Folgerung 5.3.7

Sei $\{K_j\}_{j \in J}$ eine Familie kompakter Teilmengen des \mathbb{R}^d mit Indexmenge J . Dann ist auch der Durchschnitt $\bigcap_{j \in J} K_j$ kompakt; ist zusätzlich $J = \{1, \dots, n\}$ endlich, so ist auch die Vereinigung $\bigcup_{j=1}^n K_j$ kompakt.

Beweis. Die Abgeschlossenheit von Durchschnitt und endlicher Vereinigung folgt aus Satz 5.3.4. Die Beschränktheit folgt für den Durchschnitt sofort aus der Relation $\bigcap_{j \in J} K_j \subset K_k$ für beliebiges $k \in J$. Und für die Vereinigung gilt offenbar: Haben wir $K_j \subset \hat{B}_{R_j}(0)$, $j = 1, \dots, n$, und setzen $R := \max\{R_1, \dots, R_n\}$, so folgt $\bigcup_{j=1}^n K_j \subset \hat{B}_R(0)$. \square

Wir schließen mit einer wichtigen Charakterisierung der kompakten Mengen:

Satz 5.3.8

Eine Teilmenge $K \subset \mathbb{R}^d$ ist genau dann kompakt, wenn sie *folgenkompakt* ist, d.h. wenn jede Folge $\{x_n\}_n \subset K$ eine konvergente Teilfolge $\{x_{n_l}\}_l$ enthält mit $\lim_{l \rightarrow \infty} x_{n_l} =: x_0 \in K$.

Beweis. ▶ „ \Rightarrow “: Sei K beschränkt und abgeschlossen. Eine beliebige Folge $\{x_n\}_n \subset K$ ist dann beschränkt und nach Satz 5.2.6 existiert eine konvergente Teilfolge $\{x_{n_l}\}_l \subset K$. Da K abgeschlossen ist, gilt $\lim_{l \rightarrow \infty} x_{n_l} =: x_0 \in K$.

- „ \Leftarrow “: Nun sei K folgenkompakt. Dann ist K offenbar abgeschlossen. Wäre K nicht beschränkt, so gäbe es insbesondere zu jedem $n \in \mathbb{N}$ ein $x_n \in K$ mit $x_n \notin \hat{B}_n(0)$. Somit gälte $\|x_n\| > n$ für alle $n \in \mathbb{N}$, d.h. aus $\{x_n\}_n$ könnten wir keine konvergente Teilfolge auswählen, Widerspruch! Also ist K auch beschränkt und folglich kompakt. \square

Bemerkung 5.3.9. ► Alle Begriffe dieses Abschnitts lassen sich wortgleich für beliebige normierte Vektorräume definieren. Allerdings wird Kompaktheit in solchen Räumen, abweichend von Definition 5.3.5, durch die *Heine-Borel-Eigenschaft* erklärt; Folgerung 5.3.7 bleibt dabei richtig. Im \mathbb{R}^d ist auch die Heine-Borel-Eigenschaft äquivalent zu unserer Definition der Kompaktheit; vgl. Analysis II.

- In beliebigen (unendlich-dimensionalen) normierten Vektorräumen sind nur Folgenkompaktheit und Heine-Borel-Eigenschaft äquivalent. Von Satz 5.3.8 gilt i.A. nur die Richtung „Folgenkompaktheit \Rightarrow Beschränktheit und Abgeschlossenheit“ (\rightarrow *Ergänzungen*).

5.4. Zur Klassifizierung von Punkten

Die Punkte von Mengen lassen sich in verschiedene Kategorien einteilen, mit deren Hilfe man wiederum offene und abgeschlossene Mengen charakterisieren kann.

Definition 5.4.1

Sei $M \subset \mathbb{R}^d$ eine beliebige Menge. Dann heißt ein Punkt $x_0 \in \mathbb{R}^d$:

- *innerer Punkt* von M , wenn ein $r > 0$ mit $B_r(x_0) \subset M$ existiert.
- *Randpunkt* von M , wenn zu jedem $r > 0$ Punkte $y \in M \cap B_r(x_0)$ und $z \in M^c \cap B_r(x_0)$ existieren.
- *Häufungspunkt* von M , wenn zu jedem $r > 0$ ein $x \in M \setminus \{x_0\}$ existiert mit $x \in B_r(x_0)$.
- *isolierter Punkt* von M , wenn $x_0 \in M$ gilt und x_0 kein Häufungspunkt von M ist, d.h. es existiert ein $r > 0$ mit $B_r(x_0) \cap M = \{x_0\}$.

Die Menge der inneren Punkte von $M \subset \mathbb{R}^d$ heißt das *Innere* von M ; wir schreiben \mathring{M} oder $\overset{\circ}{M}$. Die Menge der Randpunkte heißt *Rand* von M und wird mit ∂M bezeichnet. Schließlich heißt $\overline{M} := M \cup \partial M$ der (*topologischen*) *Abschluss* von M .

Wir haben folgende Charakterisierungen dieser Punkte durch Folgen:

Hilfssatz 5.4.2

Sei $M \subset \mathbb{R}^d$ eine beliebige Menge.

- (a) $x_0 \in \mathbb{R}^d$ ist genau dann innerer Punkt von M , wenn für jede Folge $\{x_n\}_n \subset \mathbb{R}^d$ mit $x_n \rightarrow x_0$ ($n \rightarrow \infty$) ein $N \in \mathbb{N}$ existiert mit $x_n \in M$ für alle $n \geq N$.
- (b) $x_0 \in \mathbb{R}^d$ ist genau dann Randpunkt von M , wenn zwei Folgen $\{y_n\}_n \subset M$ und $\{z_n\}_n \subset M^c$ existieren mit $y_n \rightarrow x_0, z_n \rightarrow x_0$ ($n \rightarrow \infty$).
- (c) $x_0 \in \mathbb{R}^d$ ist genau dann Häufungspunkt von M , wenn eine Folge $\{x_n\}_n \subset M \setminus \{x_0\}$ existiert mit $x_n \rightarrow x_0$ ($n \rightarrow \infty$).
- (d) $x_0 \in M$ ist genau dann isolierter Punkt von M , wenn für jede Folge $\{x_n\}_n \subset \mathbb{R}^d \setminus \{x_0\}$ mit $x_n \rightarrow x_0$ ($n \rightarrow \infty$) ein $N \in \mathbb{N}$ existiert mit $x_n \in M^c$ für alle $n \geq N$.

Beweis. Wir beweisen exemplarisch (a); (b) und (c) folgen auf analoge Weise, (d) ergibt sich aus (c) durch Negation (\rightarrow Nachrechnen).

- „ \Rightarrow “: Sei x_0 innerer Punkt von M und $r > 0$ mit $B_r(x_0) \subset M$ sei gewählt. Ist dann $\{x_n\}_n$ eine beliebige Folge mit $x_n \rightarrow x_0$ ($n \rightarrow \infty$), so wählen wir $N \in \mathbb{N}$ mit $\|x_n - x_0\| < r$ und somit $x_n \in B_r(x_0) \subset M$ für alle $n \geq N$.
- „ \Leftarrow “: Es existiere für jede Folge $\{x_n\}_n$ mit $x_n \rightarrow x_0$ ($n \rightarrow \infty$) ein N mit $x_n \in M$ für alle $n \geq N$. Wäre x_0 kein innerer Punkt, so gälte $B_r(x_0) \cap M^c \neq \emptyset$ für alle $r > 0$. Speziell existierte also eine Folge $\{y_n\}_n$ mit $y_n \in B_{\frac{1}{n}}(x_0) \cap M^c$ für alle $n \in \mathbb{N}$ und somit $y_n \rightarrow x_0$ ($n \rightarrow \infty$), Widerspruch! \square

Satz 5.4.3

Für eine beliebige Menge $M \subset \mathbb{R}^d$ gelten die folgenden Aussagen:

- (i) $\partial M = \partial(M^c)$.
- (ii) M ist genau dann offen, wenn $M = \text{int } M$ gilt.
- (iii) $\overline{M} = \text{int } M \cup \partial M, \partial M = \overline{M} \setminus \text{int } M$.
- (iv) Ist $\{x_n\}_n \subset M$ konvergent, so gilt $\lim_{n \rightarrow \infty} x_n =: x_0 \in \overline{M}$.
- (v) M ist abgeschlossen $\Leftrightarrow \partial M \subset M \Leftrightarrow M = \overline{M}$.

Beweis. (i) und (ii) sind aus den Definitionen sofort klar, (iii) und (iv) zeigen wir in den *Ergänzungen*. Wir beweisen (v): Zunächst ist $\partial M \subset M \Leftrightarrow M = \overline{M}$ wieder per Definition klar. Wir beweisen die erste Äquivalenz:

- „ \Rightarrow “: Seien M abgeschlossen und $x_0 \in \partial M$ gewählt. Dann existiert nach Hilfssatz 5.4.2 (b) eine Folge $\{y_n\}_n \subset M$ mit $y_n \rightarrow x_0$ ($n \rightarrow \infty$). Und es folgt $x_0 \in M$ wegen der Abgeschlossenheit von M , also $\partial M \subset M$.
- „ \Leftarrow “: Seien umgekehrt $\partial M \subset M$ und eine konvergente Folge $\{x_n\}_n \subset M$ gewählt. Wegen (iv) gilt dann $x_0 := \lim_{n \rightarrow \infty} x_n \in \overline{M} = M \cup \partial M = M$, also ist M abgeschlossen. \square

Beispiel 5.4.4. Für die offene Kugel $B_R(\xi)$ im \mathbb{R}^d gilt:

$$\begin{aligned} \text{int } B_R(\xi) &= B_R(\xi), \\ \partial B_R(\xi) &= \{x \in \mathbb{R}^d : \|x - \xi\| = R\} =: S_R(\xi), \\ \overline{B_R(\xi)} &= \{x \in \mathbb{R}^d : \|x - \xi\| \leq R\} = \hat{B}_R(\xi). \end{aligned}$$

Daher schreiben wir für die abgeschlossene Kugel in Zukunft $\overline{B_R(\xi)}$. Mit $\mathbb{S}^{d-1} := \{x \in \mathbb{R}^d : \|x\| = 1\} = S_1(0)$ bezeichnen wir noch die *Einheitssphäre* im \mathbb{R}^d .

Kapitel 6.

Stetige Funktionen

In diesem Kapitel lernen wir Funktionsgrenzwerte und den Stetigkeitsbegriff kennen. Diese Begriffe werden in den Abschnitten 6.1 bzw. 6.2 allgemein für Abbildungen $f : D \rightarrow \mathbb{R}^d$ mit $D \subset \mathbb{R}^n$ diskutiert. In Abschnitt 6.3 werden einige spezielle Begriffe und Aussagen für Funktionen einer reellen Veränderlichen zusammengestellt. In Abschnitt 6.4 werden drei zentrale Aussagen für stetige Funktionen auf kompakten Mengen bewiesen. Als Anwendung beweisen wir in Abschnitt 6.5 den berühmten Fundamentalsatz der Algebra. In Abschnitt 6.5 betrachten wir schließlich Folgen stetiger Funktionen und deren Grenzwerte.

6.1. Funktionsgrenzwerte

Wir untersuchen Abbildungen oder Funktionen $f : D \rightarrow \mathbb{R}^d$ mit Definitionsbereich $D \subset \mathbb{R}^n$ und beliebigen Dimensionen $n, d \in \mathbb{N}$. Die Abbildungsvorschrift $x \mapsto y = f(x)$ lässt sich in Koordinaten schreiben als

$$(y_1, \dots, y_d) = y = f(x) = (f_1(x_1, \dots, x_n), \dots, f_d(x_1, \dots, x_n)).$$

Die Funktionen $f_j : D \rightarrow \mathbb{R}$, $j = 1, \dots, d$, heißen auch *Komponenten(funktionen)* von f . Wir erinnern noch an das *Bild* $f(D)$ von D unter f aus Definition 1.5.6 und erklären allgemeiner das *Bild* einer beliebigen Teilmenge $M \subset D$ unter f gemäß

$$f(M) := \{f(x) : x \in M\} \subset f(D) \subset \mathbb{R}^d.$$

Gilt $n = 2$ oder/und $d = 2$, so denken wir uns - bei Bedarf - den jeweiligen Raum mit einer komplexen Struktur ausgestattet, also z.B. $f : D \rightarrow \mathbb{C}$ für $d = 2$.

Wir geben nun den zentralen Grenzwertbegriff für Funktionen:

Definition 6.1.1: Funktionsgrenzwerte
Sei $D \subset \mathbb{R}^n$ und x_0 ein Häufungspunkt von D . Zu der Funktion $f : D \rightarrow \mathbb{R}^d$ gäbe es ein

$a \in \mathbb{R}^d$, so dass für alle $\varepsilon > 0$ ein $\delta = \delta(\varepsilon) > 0$ existiert mit der Eigenschaft

$$\|f(x) - a\| < \varepsilon \quad \text{für alle } x \in D \text{ mit } 0 < \|x - x_0\| < \delta. \quad (6.1.1)$$

Dann heißt a der **Grenzwert** oder **Limes** der Funktion $f = f(x)$ im Punkt x_0 und wir schreiben

$$\lim_{x \rightarrow x_0} f(x) = a \quad \text{oder} \quad f(x) \rightarrow a \quad (x \rightarrow x_0).$$

Man sagt auch: $f(x)$ **konvergiert** (gegen a) in x_0 .

Existiert zu x_0 kein solches a , so heißt f **divergent** in x_0 .

Geometrische Interpretation

Es gilt $\lim_{x \rightarrow x_0} f(x) = a$ genau dann, wenn für jedes $\varepsilon > 0$ ein $\delta = \delta(\varepsilon) > 0$ so existiert, dass f die Menge $B'_\delta(x_0) \cap D$ in die Kugel $B_\varepsilon(a)$ abbildet. Dabei bezeichnet

$$B'_\delta(x_0) := B_\delta(x_0) \setminus \{x_0\}$$

die **punktierte Kugel**.

Satz 6.1.2: Charakterisierung durch Folgengrenzwerte

Seien $D \subset \mathbb{R}^n$, x_0 Häufungspunkt von D und $f : D \rightarrow \mathbb{R}^d$ gegeben. Dann konvergiert f genau dann in x_0 , wenn für *jede* Folge $\{x_p\}_p \subset D \setminus \{x_0\}$ mit $x_p \rightarrow x_0$ ($p \rightarrow \infty$) die Folge $\{f(x_p)\}_p$ konvergiert. In diesem Fall gilt für *alle* solche Folgen

$$\lim_{p \rightarrow \infty} f(x_p) = \lim_{x \rightarrow x_0} f(x).$$

Beweis. ▶ „ \Rightarrow “: Sei $\lim_{x \rightarrow x_0} f(x) = a$ erfüllt und $\{x_p\}_p \subset D \setminus \{x_0\}$ eine Folge mit $x_p \rightarrow x_0$ ($p \rightarrow \infty$). Zu beliebig vorgegebenem $\varepsilon > 0$ wählen wir $\delta = \delta(\varepsilon) > 0$ wie in Definition 6.1.1 und $N = N(\varepsilon) \in \mathbb{N}$ so, dass gilt

$$0 < \|x_p - x_0\| < \delta(\varepsilon) \quad \text{für alle } p \geq N(\varepsilon).$$

Dann folgt aus Formel (6.1.1)

$$\|f(x_p) - a\| < \varepsilon \quad \text{für alle } p \geq N(\varepsilon),$$

also $\lim_{p \rightarrow \infty} f(x_p) = a$.

▶ „ \Leftarrow “: Sei nun $\{f(x_p)\}_p$ für jede gegen x_0 konvergente Folge $\{x_p\}_p \subset D \setminus \{x_0\}$ konvergent. Gäbe es zwei solche Folgen $\{x_p\}_p, \{x'_p\}_p$ mit $\lim_{p \rightarrow \infty} f(x_p) \neq \lim_{p \rightarrow \infty} f(x'_p)$, so betrachten wir die *gemischte Folge* $\{\xi_p\}_p := \{x_1, x'_1, x_2, x'_2, \dots\}$. Dann konvergierte

$\{\xi_p\}_p \subset D \setminus \{x_0\}$ ebenfalls gegen x_0 , aber $\{f(\xi_p)\}_p$ hätte zwei Häufungswerte, wäre also divergent, Widerspruch.

Folglich haben *alle* Folgen $\{f(x_p)\}$ denselben Grenzwert $a \in \mathbb{R}^d$. Angenommen es gilt nicht $\lim_{x \rightarrow x_0} f(x) = a$. Dann gäbe ein $\varepsilon > 0$, so dass für alle $\delta > 0$ ein $x \in D$ existiert mit $0 < \|x - x_0\| < \delta$ und $\|f(x) - a\| \geq \varepsilon$. Wählen wir speziell $\delta = \frac{1}{p}$, so finden wir Punkte $x_p \in D$ mit $0 < \|x_p - x_0\| < \frac{1}{p}$ und $\|f(x_p) - a\| \geq \varepsilon > 0$ für alle $p \in \mathbb{N}$. Für diese Folge $\{x_p\}_p \subset D \setminus \{x_0\}$ gilt $\lim_{p \rightarrow \infty} x_p = x_0$, nach Voraussetzung müsste daher $\|f(x_p) - a\| \rightarrow 0$ ($p \rightarrow \infty$) erfüllt sein, Widerspruch! Also gilt doch $\lim_{x \rightarrow x_0} f(x) = a$. \square

Folgerung 6.1.3

Eine Funktion $f = (f_1, \dots, f_d) : D \rightarrow \mathbb{R}^d$ konvergiert genau dann in einem Häufungspunkt x_0 von $D \subset \mathbb{R}^n$, wenn dort jede Komponente $f_1, \dots, f_d : D \rightarrow \mathbb{R}$ konvergiert. Es gilt dann

$$\lim_{x \rightarrow x_0} f(x) = \left(\lim_{x \rightarrow x_0} f_1(x), \dots, \lim_{x \rightarrow x_0} f_d(x) \right).$$

Beweis. Sofort aus Satz 6.1.2 und Hilfssatz 5.2.4. \square

Satz 6.1.4: Rechenregeln für Funktionsgrenzwerte

Seien Funktionen $f, g : D \rightarrow \mathbb{R}^d$ gegeben mit $\lim_{x \rightarrow x_0} f(x) = a$, $\lim_{x \rightarrow x_0} g(x) = b$, wobei x_0 Häufungspunkt von $D \subset \mathbb{R}^n$ sei. Dann gelten die Rechenregeln:

$$\lim_{x \rightarrow x_0} [\lambda f(x) + \mu g(x)] = \lambda a + \mu b \quad \text{für alle } \lambda, \mu \in \mathbb{R},$$

$$\lim_{x \rightarrow x_0} \langle f(x), g(x) \rangle = \langle a, b \rangle,$$

$$\lim_{x \rightarrow x_0} \|f(x)\| = \|a\|$$

und für $d = 2$, also $f, g : D \rightarrow \mathbb{C}$, auch

$$\lim_{x \rightarrow x_0} [\lambda f(x) + \mu g(x)] = \lambda a + \mu b \quad \text{für alle } \lambda, \mu \in \mathbb{C},$$

$$\lim_{x \rightarrow x_0} f(x)g(x) = ab,$$

$$\lim_{x \rightarrow x_0} \overline{f(x)} = \overline{a},$$

$$\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = \frac{a}{b}, \quad \text{falls } g \neq 0 \text{ in } D \setminus \{x_0\} \text{ und } b \neq 0 \text{ ist.}$$

Beweis. Mit Satz 6.1.2 ergeben sich die Aussagen sofort aus den entsprechenden Rechenregeln für Folgengrenzwerte (Sätze 5.2.2, 3.3.1 und Bemerkung 3.3.4). Zur Übung kann man die Aussagen auch direkt mit dem ε - δ -Kriterium aus Definition 6.1.1 beweisen. \square

Bemerkung 6.1.5. Für reellwertige Funktion $f : D \rightarrow \mathbb{R}$ lassen sich auch sofort die Anordnungsregeln für Grenzwerte aus Satz 3.3.5 übertragen (\rightarrow Nachrechnen). Ferner können wir wieder den Begriff der bestimmten Divergenz erklären, vgl. Definition 3.1.10:

Definition 6.1.6

Sei $D \subset \mathbb{R}^n$ und x_0 Häufungspunkt von D . Wir sagen, eine Funktion $f : D \rightarrow \mathbb{R}$ *divergiert bestimmt* oder *konvergiert gegen* $+\infty$ (bzw. $-\infty$) für $x \rightarrow x_0$, wenn zu jedem $c > 0$ ein $\delta > 0$ existiert mit

$$f(x) > c \quad (\text{bzw. } f(x) < -c) \quad \text{für alle } x \in B'_\delta(x_0) \cap D.$$

Wir schreiben

$$\lim_{x \rightarrow x_0} f(x) = \pm\infty \quad \text{oder} \quad f(x) \rightarrow \pm\infty \quad (x \rightarrow x_0).$$

Bemerkung 6.1.7. Wie in Satz 3.1.13 folgt: Sei $f : D \rightarrow \mathbb{R}$, $D \subset \mathbb{R}^n$, x_0 Häufungspunkt von D , mit $f(x) > 0$ „nahe“ x_0 . Dann gilt

$$\lim_{x \rightarrow x_0} f(x) = +\infty \quad \Leftrightarrow \quad \lim_{x \rightarrow x_0} \frac{1}{f(x)} = 0.$$

6.2. Der Stetigkeitsbegriff

In der Schule bekommt man zu stetigen Funktionen vielleicht gesagt, dass man deren Graph „durchzeichnen kann, ohne abzusetzen“. Diese Vorstellung ist sicher hilfreich, greift aber natürlich zu kurz, spätestens wenn man stetige Funktionen mehrerer Veränderlicher betrachtet. Wir lehnen unsere Definition stattdessen an des Konzept des Funktionsgrenzwertes an:

Definition 6.2.1: Stetige Funktionen

Seien $D \subset \mathbb{R}^n$, $x_0 \in D$ und eine Funktion $f : D \rightarrow \mathbb{R}^d$ gegeben. Dann heißt f in x_0 *stetig*, wenn zu jedem $\varepsilon > 0$ ein $\delta = \delta(\varepsilon) > 0$ so existiert, dass gilt

$$\|f(x) - f(x_0)\| < \varepsilon \quad \text{für alle } x \in D \text{ mit } \|x - x_0\| < \delta. \quad (6.2.1)$$

Anderenfalls heißt f in x_0 *unstetig*.

Geometrische Interpretation

$f : D \rightarrow \mathbb{R}^d$ ist in $x_0 \in D$ genau dann stetig, wenn für jeden Radius $\varepsilon > 0$ ein $\delta > 0$ so existiert, dass f die Menge $B_\delta(x_0) \cap D$ in die Kugel $B_\varepsilon(f(x_0))$ abbildet, d.h.

$$f(B_\delta(x_0) \cap D) \subset B_\varepsilon(f(x_0)). \quad (6.2.2)$$

- Bemerkung 6.2.2.** (a) Ist $x_0 \in D$ isolierter Punkt von D , so ist offenbar jede Funktion $f : D \rightarrow \mathbb{R}^d$ in x_0 stetig.
- (b) Die Stetigkeit ist eine *lokale Eigenschaft*, d.h.: Ist f in x_0 stetig (bzw. unstetig) und ändern wir f in $D \setminus B_r(x_0)$ für ein $r > 0$ beliebig ab, so bleibt die resultierende Funktion in x_0 stetig (bzw. unstetig).

Satz 6.2.3: Charakterisierung der Stetigkeit

Sei $f : D \rightarrow \mathbb{R}^d$ auf $D \subset \mathbb{R}^n$ erklärt und sei $x_0 \in D$ Häufungspunkt. Dann sind folgende Aussagen äquivalent:

- (i) f ist stetig in x_0 .
- (ii) Es gilt $\lim_{x \rightarrow x_0} f(x) = f(x_0)$.
- (iii) Für jede Folge $\{x_p\}_p \subset D$ mit $x_p \rightarrow x_0$ ($p \rightarrow \infty$) gilt $\lim_{p \rightarrow \infty} f(x_p) = f(x_0)$.

Beweis. Sofort aus den Definitionen 6.1.1 und 6.2.1 sowie Satz 6.1.2. In (iii) kann die Forderung $x_p \neq x_0$ aus Satz 6.1.2 offenbar fallen gelassen werden. \square

Satz 6.2.4: Rechenregeln stetiger Funktionen

- (a) Sind $f, g : D \rightarrow \mathbb{R}^d$ stetig in $x_0 \in D$, so gilt dies auch für jede Linearkombination $\lambda f + \mu g$ mit $\lambda, \mu \in \mathbb{R}$, das Skalarprodukt $\langle f, g \rangle$ und die Norm $\|f\|$.
- (b) Sind $f, g : D \rightarrow \mathbb{C}$ stetig in $x_0 \in D$, so gilt dies auch für jede Linearkombination $\lambda f + \mu g$ mit $\lambda, \mu \in \mathbb{C}$, das Produkt fg und, falls $g \neq 0$ in D , auch für den Quotienten $\frac{f}{g}$.

Beweis. Sofort aus den Sätzen 6.1.4 und 6.2.3. (Falls x_0 kein Häufungspunkt, also isoliert ist, sind die Aussagen wegen Bemerkung 6.2.2 (a) trivial.) \square

Beispiel 6.2.5. (a) Die konstante Funktion $f(x) = c \in \mathbb{R}^d$ und die identische Abbildung $f(x) = x, x \in D \subset \mathbb{R}^n$, sind offenbar in allen Punkten $x_0 \in D$ stetig.

(b) Polynomfunktionen $p(z) = \sum_{k=0}^n a_k z^k, n \in \mathbb{N}$, mit Koeffizienten $a_0, \dots, a_n \in \mathbb{C}$ sind in jedem Punkt $z_0 \in \mathbb{C}$ stetig wegen (a) und Satz 6.2.4 (b).

(c) Die Funktion $f(x) := \|x\|^2, x \in \mathbb{R}^n$, ist in allen Punkten $x_0 \in \mathbb{R}^n$ stetig.

Denn: Ist $x_0 \in \mathbb{R}^d$ fixiert, so schätzen wir ab:

$$\begin{aligned} |f(x) - f(x_0)| &= \left| \|x\|^2 - \|x_0\|^2 \right| = (\|x\| + \|x_0\|) \|x - x_0\| \\ &\stackrel{\text{Bem. 5.1.10}}{\leq} (\|x - x_0\| + 2\|x_0\|) \|x - x_0\|, \quad x \in \mathbb{R}^d. \end{aligned}$$

Ist $\varepsilon > 0$ beliebig, so wähle $\delta = \delta(\varepsilon) > 0$ mit $(\delta + 2\|x_0\|)\delta < \varepsilon$. Dann folgt offenbar (6.2.1). \square

(d) Die *DIRICHLETSche Sprungfunktion*

$$f(x) := \chi_{\mathbb{Q}}(x) = \begin{cases} 1, & x \in \mathbb{Q} \\ 0, & x \in \mathbb{R} \setminus \mathbb{Q} \end{cases}$$

ist in keinem Punkt aus \mathbb{R} stetig. Die Funktion

$$g(x) := (\text{Id} \cdot \chi_{\mathbb{Q}})(x) = \begin{cases} x, & x \in \mathbb{Q} \\ 0, & x \in \mathbb{R} \setminus \mathbb{Q} \end{cases}$$

ist in $x_0 = 0$ und nur dort stetig (\rightarrow Übungsaufgaben).

(e) Die *Signumfunktion* $\text{sgn} : \mathbb{R} \rightarrow \mathbb{R}$ mit

$$\text{sgn}(x) := \begin{cases} 1, & x > 0 \\ 0, & x = 0 \\ -1, & x < 0 \end{cases}$$

ist in allen Punkten $x_0 \in \mathbb{R} \setminus \{0\}$ stetig und in $x_0 = 0$ unstetig.

Denn: Die Stetigkeit in allen $x_0 \in \mathbb{R} \setminus \{0\}$ ist klar wegen Beispiel (a) und des lokalen Charakters der Stetigkeit; siehe Bemerkung 6.2.2 (b). Die Unstetigkeit in 0 zeigen wir in Bemerkung 6.3.3 (a). \square

(f) Die Funktion

$$f(x, y) = \begin{cases} \frac{xy}{x^2 + y^2}, & \text{für } (x, y) \in \mathbb{R}^2 \setminus \{(0, 0)\} \\ a, & \text{für } (x, y) = (0, 0) \end{cases}$$

ist für jede Wahl von $a \in \mathbb{R}$ in allen Punkten aus $\mathbb{R}^2 \setminus \{(0, 0)\}$ stetig und in $(0, 0)$ unstetig.

Denn: Die Stetigkeit in $\mathbb{R}^2 \setminus \{(0, 0)\}$ folgt sofort aus Satz 6.2.4, da die Funktionen $f_1(x, y) = x, f_2(x, y) = y$ offenbar stetig in allen Punkten des \mathbb{R}^2 sind. Zum Beweis der

Unstetigkeit in $(0, 0)$ betrachten wir eine Nullfolge $\{x_p\}_p \subset \mathbb{R} \setminus \{0\}$; es gilt dann

$$\lim_{p \rightarrow \infty} f(x_p, 0) = \lim_{p \rightarrow \infty} \frac{0}{x_p^2} = 0 \neq \frac{1}{2} = \lim_{p \rightarrow \infty} \frac{x_p^2}{x_p^2 + x_p^2} = \lim_{p \rightarrow \infty} f(x_p, x_p).$$

Nach Satz 6.2.3 ist f also unstetig in $(0, 0)$. □

Satz 6.2.6: Komposition stetiger Funktionen

Seien Funktionen $f : D \rightarrow \mathbb{R}^d$ und $g : E \rightarrow \mathbb{R}^m$ gegeben mit $D \subset \mathbb{R}^n$, $E \subset \mathbb{R}^d$ und $f(D) \subset E$. Weiter seien f in $x_0 \in D$ und g in $y_0 = f(x_0) \in E$ stetig. Dann ist auch die Komposition $h := g \circ f : D \rightarrow \mathbb{R}^m$ in x_0 stetig.

Beweis. Da für isolierte Punkte nichts zu zeigen ist, können wir annehmen, dass x_0 Häufungspunkt von D und y_0 Häufungspunkt von E ist. Sei nun $\{x_p\}_p \subset D$ mit $x_p \rightarrow x_0$ ($p \rightarrow \infty$) eine beliebige Folge. Nach Satz 6.2.3 gilt dann

$$\lim_{p \rightarrow \infty} f(x_p) = f\left(\lim_{p \rightarrow \infty} x_p\right) = f(x_0) = y_0.$$

Somit folgt wiederum aus Satz 6.2.3

$$\lim_{p \rightarrow \infty} h(x_p) = \lim_{p \rightarrow \infty} g(f(x_p)) = g\left(\lim_{p \rightarrow \infty} f(x_p)\right) = g(y_0) = h(x_0),$$

also die behauptete Stetigkeit von $h = g \circ f$ in x_0 . □

Von besonderem Interesse sind Funktionen, die in *allen* Punkten ihres Definitionsbereichs stetig sind:

Definition 6.2.7

Eine Funktion $f : D \rightarrow \mathbb{R}^d$, $D \subset \mathbb{R}^n$, nennen wir *stetig (auf D)*, wenn f in allen Punkten $x \in D$ stetig ist. Mit $C^0(D, \mathbb{R}^d)$ bezeichnen wir die Menge aller stetigen Funktionen $f : D \rightarrow \mathbb{R}^d$. Für $d = 1$ schreiben wir auch kurz $C^0(D) := C^0(D, \mathbb{R})$ und für $d = 2$ auch $C^0(D, \mathbb{C}) := C^0(D, \mathbb{R}^2)$.

Bemerkung 6.2.8. ▶ Gemäß Satz 6.2.4 wird $C^0(D, \mathbb{R}^d)$ durch die Rechenoperationen

$$(f + g)(x) := f(x) + g(x), \quad (\lambda f)(x) := \lambda f(x) \quad \text{für } x \in D$$

zu einem (unendlich dimensionalen) Vektorraum; siehe Definition 5.1.1.

- Gemäß Satz 6.2.6 gilt für $f \in C^0(D, \mathbb{R}^d)$ und $g \in C^0(E, \mathbb{R}^m)$ mit $f(D) \subset E$ auch $g \circ f \in C^0(D, \mathbb{R}^m)$.

Wir schließen mit einer Stetigkeitsaussage für die Umkehrfunktion einer injektiven Abbildung mit *kompaktem* Definitionsbereich:

Satz 6.2.9: Stetigkeit der Umkehrfunktion

Sei $K \subset \mathbb{R}^n$ kompakt und $f : K \rightarrow \mathbb{R}^d$ sei stetig und injektiv mit Bild $W := f(K)$. Dann ist auch die Umkehrfunktion $f^{-1} : W \rightarrow \mathbb{R}^n$ von f stetig auf W .

Beweis. Sei $y_0 \in W$ ein beliebiger Häufungspunkt und sei $\{y_p\}_p \subset W$ mit $y_p \rightarrow y_0$ ($p \rightarrow \infty$). Gemäß Satz 6.2.3 ist dann zu zeigen

$$x_p := f^{-1}(y_p) \rightarrow f^{-1}(y_0) =: x_0 \quad (p \rightarrow \infty).$$

Die Folge $\{x_p\}_p \subset K$ ist beschränkt, da K beschränkt ist. Sei nun $\xi \in \mathbb{R}^n$ ein beliebiger Häufungswert von $\{x_p\}_p$ und $\{x_{p_k}\}_k$ eine Teilfolge mit $x_{p_k} \rightarrow \xi$ ($k \rightarrow \infty$). Da K abgeschlossen ist, gilt $\xi \in K$. Die Stetigkeit von f liefert also $f(x_{p_k}) \rightarrow f(\xi)$ ($k \rightarrow \infty$). Andererseits wissen wir

$$f(x_{p_k}) = f(f^{-1}(y_{p_k})) = y_{p_k} \rightarrow y_0 \quad (k \rightarrow \infty)$$

und somit $f(\xi) = y_0 = f(x_0)$. Die Injektivität von f liefert daher $\xi = x_0$ für alle Häufungswerte von $\{x_p\}_p$. Das bedeutet $\lim_{p \rightarrow \infty} x_p = x_0$, wie behauptet. \square

Bemerkung 6.2.10. Für Funktionen *einer* reellen Veränderlichen werden wir Satz 6.2.9 im folgenden Abschnitt noch auf nicht-kompakte Definitionsbereiche verallgemeinern; siehe Satz 6.3.8. Weitere wichtige Aussagen zu stetigen Funktionen mit kompakten Definitionsbereichen $K \subset \mathbb{R}^n$ werden wir in Abschnitt 6.4 beweisen.

6.3. Funktionen einer reellen Veränderlichen

Wir wollen einige Begriffe und Aussagen zusammentragen, die speziell für Funktionen einer reellen Veränderlichen gelten, d.h. $f : D \rightarrow \mathbb{R}^d$ mit $D \subset \mathbb{R}$. Wir beginnen mit der

Definition 6.3.1: Einseitige Grenzwerte

Es seien $D \subset \mathbb{R}$ und $f : D \rightarrow \mathbb{R}^d$ gegeben.

- (a) Gilt $(x_0, x_0 + \alpha) \subset D$ und gibt es ein $a \in \mathbb{R}^d$ so, dass für alle $\varepsilon > 0$ ein $\delta = \delta(\varepsilon) \in (0, \alpha)$

existiert mit

$$\|f(x) - a\| < \varepsilon \quad \text{für alle } x_0 < x < x_0 + \delta,$$

so heißt a der *rechtsseitige Limes* oder *Grenzwert* von f an der Stelle x_0 ; wir schreiben dann

$$f(x_0+) := \lim_{x \rightarrow x_0+} f(x) = a \quad \text{oder} \quad f(x) \rightarrow a \quad (x \rightarrow x_0+).$$

- (b) Gilt $(x_0 - \alpha, x_0) \subset D$ und gibt es ein $a \in \mathbb{R}^d$ so, dass für alle $\varepsilon > 0$ ein $\delta = \delta(\varepsilon) \in (0, \alpha)$ existiert mit

$$\|f(x) - a\| < \varepsilon \quad \text{für alle } x_0 - \delta < x < x_0,$$

so heißt a der *linksseitige Limes* oder *Grenzwerte* von f an der Stelle x_0 ; wir schreiben dann

$$f(x_0-) := \lim_{x \rightarrow x_0-} f(x) = a \quad \text{oder} \quad f(x) \rightarrow a \quad (x \rightarrow x_0-).$$

- (c) Gilt $(\beta, +\infty) \subset D$, so sagen wir, f *konvergiert gegen* $b \in \mathbb{R}^d$ *für* $x \rightarrow +\infty$, wenn $f(\frac{1}{t}) \rightarrow b$ ($t \rightarrow 0+$) gilt; wir schreiben dann

$$\lim_{x \rightarrow +\infty} f(x) = b \quad \text{oder} \quad f(x) \rightarrow b \quad (x \rightarrow +\infty).$$

- (d) Ist schließlich $(-\infty, \beta) \subset D$, so sagen wir, f *konvergiert gegen* $b \in \mathbb{R}^d$ *für* $x \rightarrow -\infty$, wenn $f(\frac{1}{t}) \rightarrow b$ ($t \rightarrow 0-$) richtig ist; wir schreiben dann

$$\lim_{x \rightarrow -\infty} f(x) = b \quad \text{oder} \quad f(x) \rightarrow b \quad (x \rightarrow -\infty).$$

Bemerkung 6.3.2. (a) Ist $f : D \rightarrow \mathbb{R}^d$, $D \subset \mathbb{R}$ und $(x_0 - \alpha, x_0 + \alpha) \setminus \{x_0\} \subset D$, so besitzt f in x_0 offenbar genau dann den Grenzwert $\lim_{x \rightarrow x_0} f(x) =: a$, wenn gilt

$$\lim_{x \rightarrow x_0-} f(x) = a = \lim_{x \rightarrow x_0+} f(x).$$

- (b) Die Charakterisierung aus Satz 6.1.2 und die Rechenregeln aus Satz 6.1.4 lassen sich sofort wörtlich auf die einseitigen Grenzwerte aus Definition 6.3.1 übertragen.
- (c) Analog zu Definition 6.1.6 erweitert man für reellwertige Funktionen $f : D \rightarrow \mathbb{R}$, $D \subset \mathbb{R}$, die einseitigen Grenzwerte auch auf Werte $\pm\infty$. Dann bleibt natürlich auch Bemerkung 6.1.7 für einseitige Grenzwerte richtig.

Beispiel 6.3.3. (a) Für die Signumfunktion $\text{sgn} : \mathbb{R} \rightarrow \mathbb{R}$ aus Bemerkung 6.2.5 (e) gilt in $x_0 = 0$:

$$\lim_{x \rightarrow 0-} \text{sgn}(x) = -1, \quad \lim_{x \rightarrow 0+} \text{sgn}(x) = +1.$$

Also besitzt sgn in $x_0 = 0$ keinen Grenzwert und ist dort somit auch unstetig.

(b) Für die Funktion $f(x) := \frac{1}{x} : (0, +\infty) \rightarrow \mathbb{R}$ gilt $\lim_{x \rightarrow +\infty} f(x) = 0$, denn wir haben $f(\frac{1}{t}) = t \rightarrow 0 \ (t \rightarrow 0+)$.

(c) Beispiel (b) und die Rechenregeln für Funktionsgrenzwerte aus Satz 6.2.4 liefern

$$\lim_{x \rightarrow +\infty} \frac{7x - 2}{3x + 1} = \lim_{x \rightarrow +\infty} \frac{7 - \frac{2}{x}}{3 + \frac{1}{x}} = \frac{7}{3}.$$

Wir konzentrieren uns nun auf stetige, reellwertige Funktionen $f : I \rightarrow \mathbb{R}$ an, wobei $I \subset \mathbb{R}$ immer ein Intervall ist. Wir geben für solche Funktionen einige wichtige Ergebnisse an und beginnen mit dem

Satz 6.3.4: Zwischenwertsatz

Sei $f : [a, b] \rightarrow \mathbb{R}$ stetig mit $f(a) \neq f(b)$. Dann existiert zu jedem Wert c zwischen $f(a)$ und $f(b)$ mindestens ein $\xi \in (a, b)$ mit $f(\xi) = c$.

Beweis. Wir können $f(a) < c < f(b)$ annehmen; anderenfalls gehen wir zu $-f$ und $-c$ über. Wir betrachten die Menge

$$M := \{x \in [a, b] : f(x) < c\},$$

die offenbar nichtleer und nach oben beschränkt ist. Setzen wir $\xi := \sup M$, so gibt es eine Folge $\{x_p\}_p \subset M$ mit $x_p \rightarrow \xi \ (p \rightarrow \infty)$; vgl. Hilfssatz 3.5.4. Die Stetigkeit von f liefert dann $f(\xi) = \lim_{p \rightarrow \infty} f(x_p) \leq c$, und nach Voraussetzung folgt noch $\xi < b$.

Wäre nun $f(\xi) < c$, so gäbe es wegen der Stetigkeit von f ein $\delta \in (0, b - \xi)$, so dass gilt

$$f(x) < c \quad \text{für alle } x \in [\xi, \xi + \delta), \quad \text{d.h. } [\xi, \xi + \delta) \subset M,$$

im Widerspruch zur Wahl von $\xi = \sup M$. Also folgt $f(\xi) = c$. □

Die nachstehenden Aussagen sind nicht zentral für unsere Vorlesung, so dass wir es bei einer Erwähnung belassen und die Beweise in die Ergänzungen verlagern.

Folgerung 6.3.5

Sei $I \subset \mathbb{R}$ ein beliebiges, nichtleeres Intervall und $f : I \rightarrow \mathbb{R}$ eine stetige Funktion. Dann ist auch $f(I) \subset \mathbb{R}$ ein Intervall.

Beweis. Direkt aus Satz 6.3.4 \rightarrow Ergänzungen. □

Definition 6.3.6

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

$$f(x) \leq f(x') \quad (\text{bzw. } f(x) \geq f(x')) \quad \text{für alle } x, x' \in D \text{ mit } x < x'$$

erfüllt ist. f heißt *streng monoton wachsend (bzw. fallend)*, wenn gilt

$$f(x) < f(x') \quad (\text{bzw. } f(x) > f(x')) \quad \text{für alle } x, x' \in D \text{ mit } x < x'.$$

Folgerung 6.3.7

Eine stetige Funktion $f : I \rightarrow \mathbb{R}$ auf einem Intervall $I \subset \mathbb{R}$ ist genau dann injektiv, wenn sie streng monoton (wachsend oder fallend) ist.

Beweis. → Übungsaufgabe. □

Wir schließen mit einer Verallgemeinerung von Satz 6.2.9 auf beliebige Intervalle:

Satz 6.3.8

Sei $I \subset \mathbb{R}$ ein beliebiges, nichtleeres Intervall. Dann besitzt jede stetige, injektive Funktion $f : I \rightarrow \mathbb{R}$ eine stetige Umkehrfunktion $f^{-1} : I^* \rightarrow \mathbb{R}$ auf dem Intervall $I^* := f(I)$.

Beweis. → Ergänzungen. □

Beispiel 6.3.9. Sei $q = \frac{r}{s} \in \mathbb{Q}$ mit $q > 0$ gewählt (d.h. $r \in \mathbb{N}$). Wir setzen $0^q := 0$ und erklären die *Potenzfunktion* $f : [0, +\infty) \rightarrow [0, +\infty)$, $x \mapsto x^q$. Dann ist f stetig, streng monoton wachsend und bijektiv mit Umkehrfunktion $f^{-1}(y) = y^{\frac{1}{q}}$, $y \in [0, +\infty)$.

Denn: Gemäß Definition 3.4.13 gilt $f([0, +\infty)) \subset [0, +\infty)$. Für $\varphi(y) := y^{\frac{1}{q}}$, $y \in [0, +\infty)$, ist analog $\varphi([0, \infty)) \subset [0, \infty)$. Folglich sind $\varphi \circ f$ und $f \circ \varphi$ wohl definiert, und wir finden mit Bemerkung 3.4.14:

$$(\varphi \circ f)(x) = (x^q)^{\frac{1}{q}} = x, \quad f \circ \varphi(y) = (y^{\frac{1}{q}})^q = y$$

für beliebige $x, y \in [0, +\infty)$. Gemäß Satz 1.5.13 ist $f : [0, +\infty) \rightarrow [0, +\infty)$ also bijektiv mit Umkehrabbildung $f^{-1} = \varphi$.

Weiter ist $f_1(x) := x^r$, $x \in [0, +\infty)$, gemäß Beispiel 6.2.5 (b) für beliebiges $r \in \mathbb{N}$ stetig. Nach Satz 6.3.8 sind dann auch die Umkehrfunktionen $f_2(x) := \sqrt[s]{x}$, $x \in [0, +\infty)$, für jedes $s \in \mathbb{N}$

stetig. Satz 6.2.6 liefert daher die Stetigkeit der Komposition

$$f_1 \circ f_2(x) = \left(\sqrt[r]{x}\right)^r \stackrel{\text{Def. 3.4.13}}{=} f(x), \quad x \in [0, +\infty).$$

Schließlich ist f nach Folgerung 6.3.7 und wegen $f(0) = 0$ auch streng monoton wachsend. \square

6.4. Stetige Funktionen auf kompakten Mengen

Wir widmen uns nun wieder dem allgemeinen Fall von Funktionen auf Teilmengen des \mathbb{R}^n .

Nach Satz 6.2.9 besitzen stetige, injektive Funktionen auf kompakten Mengen eine stetige Umkehrfunktion. In diesem Abschnitt lernen wir weitere wichtige Eigenschaften kennen, die Kompakta als Definitionsbereiche stetiger Funktionen auszeichnen. Wir beginnen mit dem

Satz 6.4.1

Ist $K \subset \mathbb{R}^n$ kompakt und $f \in C^0(K, \mathbb{R}^d)$, dann ist auch $f(K) \subset \mathbb{R}^d$ kompakt.

Beweis. Sei $\{y_p\}_p \subset f(K)$ eine beliebige Folge. Zu jedem y_p gibt es (mindestens) ein $x_p \in K$ mit $f(x_p) = y_p$. Da K kompakt ist, können wir nach Satz 5.3.8 aus $\{x_p\}_p \subset K$ eine konvergente Teilfolge $\{x_{p_l}\}_l$ auswählen mit $\lim_{l \rightarrow \infty} x_{p_l} =: x_0 \in K$. Die Stetigkeit von f liefert somit

$$y_{p_l} = f(x_{p_l}) \rightarrow f(x_0) =: y_0 \in f(K) \quad \text{für } l \rightarrow \infty,$$

d.h. $f(K)$ ist folgenkompakt. Wiederum Satz 5.3.8 entnehmen wir daher die behauptete Kompaktheit von $f(K)$. \square

Definition 6.4.2

Eine Funktion $f : D \rightarrow \mathbb{R}^d$ auf $D \subset \mathbb{R}^n$ heißt *beschränkt*, wenn ein $c > 0$ existiert mit

$$\|f(x)\| \leq c \quad \text{für alle } x \in D.$$

Bemerkung 6.4.3. \blacktriangleright $f : D \rightarrow \mathbb{R}^d$ ist also genau dann beschränkt, wenn ihr Bild $f(D)$ beschränkt ist. Satz 6.4.1 liefert, dass Kompaktheit von D und Stetigkeit von f dafür hinreichend sind.

- \blacktriangleright Für Folgen $\{x_n\}_n \subset \mathbb{R}^d$, also Abbildungen $f : \mathbb{N} \rightarrow \mathbb{R}^d, n \mapsto x_n := f(n)$, stimmt Definition 6.4.2 mit der ursprünglichen Definition 5.2.1 überein.

Eines der wichtigsten Hilfsmittel der Analysis enthält der folgende

Satz 6.4.4: WEIERSTRASScher Hauptsatz

Sei $K \subset \mathbb{R}^n$ kompakt und nichtleer, und sei $f \in C^0(K, \mathbb{R})$. Dann gibt es Punkte $\underline{x}, \bar{x} \in K$, so dass gilt

$$f(\underline{x}) \leq f(x) \leq f(\bar{x}) \quad \text{für alle } x \in K. \quad (6.4.1)$$

Bemerkung 6.4.5. Relation (6.4.1) können wir auch schreiben als

$$\begin{aligned} f(\underline{x}) &= \inf f(K) =: \inf_{x \in K} f(x) = \inf_K f, \\ f(\bar{x}) &= \sup f(K) =: \sup_{x \in K} f(x) = \sup_K f. \end{aligned}$$

Das heißt: Jede auf einem Kompaktum stetige Funktion nimmt dort ihr Infimum (= Minimum) bzw. Supremum (= Maximum) an. Die Aussage wird falsch, wenn man eine der Voraussetzungen fallen lässt.

Beweis von Satz 6.4.4. Nach Satz 6.4.1 ist $f(K) \subset \mathbb{R}$ beschränkt und abgeschlossen. Wegen der Beschränktheit existieren

$$\underline{m} := \inf f(K) \in \mathbb{R}, \quad \bar{m} := \sup f(K) \in \mathbb{R}.$$

Nach Hilfssatz 3.5.4 gibt es zwei Folgen $\{\underline{y}_p\}_p, \{\bar{y}_p\}_p \subset f(K)$ mit

$$\underline{y}_p \rightarrow \underline{m}, \quad \bar{y}_p \rightarrow \bar{m} \quad (p \rightarrow \infty).$$

Die Abgeschlossenheit von $f(K)$ liefert somit $\underline{m}, \bar{m} \in f(K)$; d.h. es gibt Punkte $\underline{x}, \bar{x} \in K$ mit

$$f(\underline{x}) = \underline{m} \leq f(x) \leq \bar{m} = f(\bar{x}) \quad \text{für alle } x \in K,$$

wie behauptet. □

Für die Formulierung des dritten grundlegenden Resultats benötigen wir noch die folgende Verschärfung des Stetigkeitsbegriffs:

Definition 6.4.6

Sei $D \subset \mathbb{R}^n$ und $f : D \rightarrow \mathbb{R}^d$ gegeben. Dann heißt f *gleichmäßig stetig* auf D , wenn zu jedem $\varepsilon > 0$ ein $\delta = \delta(\varepsilon) > 0$ existiert, so dass gilt

$$\|f(x) - f(x')\| < \varepsilon \quad \text{für alle } x, x' \in D \text{ mit } \|x - x'\| < \delta. \quad (6.4.2)$$

Bemerkung 6.4.7. Für eine in allen Punkten $x' \in D$ stetige Funktion $f \in C^0(D, \mathbb{R}^d)$ gilt (6.4.2) ebenfalls, jedoch mit einem i.A. von x' abhängigen $\delta = \delta(\varepsilon, x')$. Jede gleichmäßig stetige Funktion ist also stetig. Die Umkehrung gilt jedoch i.A. nicht, wie etwa das Beispiel $f(x) := \frac{1}{x}$, $x \in (0, 1]$, zeigt.

Denn: Angenommen es gäbe z.B. für $\varepsilon = 1$ ein $\delta > 0$, so dass $|f(x) - f(x')| < 1$ für alle $x, x' \in (0, 1]$ mit $|x - x'| < \delta$ richtig ist. Speziell für $0 < x < \min\{\delta, \frac{1}{2}\}$ und $x' = 2x$ folgte dann aber $|x - x'| = x < \delta$ und

$$|f(x) - f(x')| = \left| \frac{1}{x} - \frac{1}{2x} \right| = \frac{1}{2x} > 1,$$

Widerspruch! □

Satz 6.4.8: HEINE

Ist $K \subset \mathbb{R}^n$ kompakt und $f \in C^0(K, \mathbb{R}^d)$, so ist f gleichmäßig stetig.

Beweis. Angenommen, f ist nicht gleichmäßig stetig. Dann gäbe es ein $\varepsilon > 0$, so dass für alle $\delta > 0$ Punkte $x, x' \in K$ mit $\|x - x'\| < \delta$ existieren, für die gilt $\|f(x) - f(x')\| \geq \varepsilon$. Wählen wir insbesondere $\delta = \frac{1}{p}$, $p \in \mathbb{N}$, so finden wir Folgen $\{x_p\}_p, \{x'_p\}_p \subset K$ mit

$$\|x_p - x'_p\| < \frac{1}{p} \quad \text{für alle } p \in \mathbb{N} \quad (6.4.3)$$

und

$$\|f(x_p) - f(x'_p)\| \geq \varepsilon \quad \text{für alle } p \in \mathbb{N}. \quad (6.4.4)$$

Da K kompakt ist, existiert nach Satz 5.3.8 eine konvergente Teilfolge $\{x_{p_l}\}_l \subset \{x_p\}_p$ mit $\lim_{l \rightarrow \infty} x_{p_l} = x_0 \in K$. Für die entsprechende Teilfolge $\{x'_{p_l}\}_l \subset \{x'_p\}_p$ liefert (6.4.3) ebenfalls $\lim_{l \rightarrow \infty} x'_{p_l} = x_0$. Aus der Stetigkeit von f und (6.4.4) folgern wir schließlich

$$0 = \|f(x_0) - f(x_0)\| = \left\| \lim_{l \rightarrow \infty} f(x_{p_l}) - \lim_{l \rightarrow \infty} f(x'_{p_l}) \right\| = \lim_{l \rightarrow \infty} \|f(x_{p_l}) - f(x'_{p_l})\| \geq \varepsilon > 0,$$

also einen Widerspruch! □

Bemerkung 6.4.9. Im Beispiel $f(x) = \frac{1}{x}$, $x \in (0, 1]$, aus Bemerkung 6.4.7 ist zwar f stetig, aber $(0, 1]$ nicht kompakt.

6.5. Der Fundamentalsatz der Algebra

Hauptziel dieses Abschnitts ist ein Beweis der folgenden berühmten Aussage:

Satz 6.5.1: Fundamentalsatz der Algebra

Jedes nicht-konstante komplexe Polynom

$$p(z) = \sum_{k=0}^n a_k z^k = a_n z^n + \dots + a_1 z + a_0, \quad z \in \mathbb{C},$$

mit Koeffizienten $a_0, \dots, a_n \in \mathbb{C}$ besitzt eine Nullstelle in \mathbb{C} .

Bemerkung 6.5.2. (a) Der Satz wird für *reelle* Polynome $p : \mathbb{R} \rightarrow \mathbb{R}$ falsch, wie schon das einfache Beispiel $p(x) = x^2 + 1$, $x \in \mathbb{R}$, zeigt.

(b) Aus dem Satz folgt: Zu jedem $c \in \mathbb{C}$ besitzt ein nicht-konstantes komplexes Polynom p eine c -Stelle, d.h. ein $z \in \mathbb{C}$ mit $p(z) = c$; man betrachtet dazu einfach das Polynom $q := p - c$. Die Abbildung $p : \mathbb{C} \rightarrow \mathbb{C}$ ist also surjektiv.

(c) Für ein beliebiges Polynom p lässt sich mit einem eindeutig bestimmten $n \in \mathbb{N}_0$ annehmen

$$p(z) = \sum_{k=0}^n a_k z^k, \quad z \in \mathbb{C}, \quad \text{mit } a_n \neq 0. \quad (6.5.1)$$

Die Zahl $n =: \deg(p)$ heißt dann *Grad* von p . Offenbar gilt $\deg(p) \geq 1$ genau dann, wenn p nicht konstant ist.

(d) Es gibt viele verschiedene Beweise des Fundamentalsatzes; den ersten vollständigen verdankt man C. F. GAUSS (1799). Wir geben hier den *rein analytischen* Beweis von J.-R. ARGAND (1806) an, der wesentlich auf dem WEIERSTRASSschen Hauptsatz, Satz 6.4.4, beruht (siehe Hilfssatz 6.5.3).

Für den Beweis benötigen wir etwas Vorbereitung in Form von drei Hilfssätzen.

Hilfssatz 6.5.3: CAUCHYScher Minimumsatz

Ist $p : \mathbb{C} \rightarrow \mathbb{C}$ ein nicht-konstantes Polynom, so existiert ein $z_0 \in \mathbb{C}$ mit

$$|p(z_0)| = \inf_{z \in \mathbb{C}} |p(z)|.$$

Beweis. Sei p durch (6.5.1) gegeben mit einem $n \in \mathbb{N}$. Mit $\sigma := |a_0| + \dots + |a_{n-1}|$ können wir

für $|z| \geq 1$ abschätzen:

$$\begin{aligned} |p(z)| &\geq |a_n||z|^n - \sum_{k=0}^{n-1} |a_k||z|^k = |z|^n \left(|a_n| - \sum_{k=0}^{n-1} |a_k||z|^{k-n} \right) \\ &\geq |z|^n (|a_n| - \sigma|z|^{-1}) \quad \text{für alle } z \in \mathbb{C} : |z| \geq 1. \end{aligned}$$

Setzen wir $R := \max \left\{ 1, \frac{2\sigma}{|a_n|} \right\} > 0$, so folgt also

$$|p(z)| \geq |z|^n \left(|a_n| - \frac{\sigma}{R} \right) \geq \frac{|a_n|}{2} |z|^n \geq \frac{|a_n|}{2} |z| \quad \text{für alle } z \in \mathbb{C} : |z| \geq R.$$

Schließlich sei noch $R_0 := \max \left\{ R, \frac{2|a_0|}{|a_n|} \right\}$, dann haben wir

$$|p(z)| \geq \frac{|a_n|}{2} R_0 \geq |a_0| \quad \text{für alle } z \in \mathbb{C} \setminus B_{R_0}(0). \quad (6.5.2)$$

Nun gilt $|p| \in C^0(\mathbb{C})$; siehe Beispiel 6.2.5 (b) und Satz 6.2.4. Nach dem WEIERSTRASSschen Hauptsatz, Satz 6.4.4, nimmt also $|p|$ auf der kompakten Menge $\overline{B_{R_0}(0)}$ ihr Infimum in einem Punkt $z_0 \in \overline{B_{R_0}(0)}$ an. Dann muss aber insbesondere $|p(z_0)| \leq |p(0)| = |a_0|$ gelten, so dass (6.5.2) liefert

$$|p(z_0)| = \inf_{z \in \overline{B_{R_0}(0)}} |p(z)| = \inf_{z \in \mathbb{C}} |p(z)|,$$

wie behauptet. □

Hilfssatz 6.5.4

Zu beliebigen $n \in \mathbb{N}$ und $c \in \mathbb{C}$ hat die Gleichung $z^n = c$ eine Lösung $z \in \mathbb{C}$, genannt *n-te Wurzel* von c .

Beweis. Diese Aussage zeigen wir in Kapitel 8 m.H. der *Polarkordinatendarstellung* komplexer Zahlen. □

Hilfssatz 6.5.5: ARGANDS Ungleichung

Ist p ein nicht-konstantes Polynom und $z_0 \in \mathbb{C}$ ein Punkt mit $p(z_0) \neq 0$, so gibt es ein $\zeta \in \mathbb{C}$ mit

$$|p(z_0)| > |p(z_0 + \zeta)|.$$

Beweis. Wir entwickeln p aus (6.5.1) um z_0 mit dem Binomischen Satz 2.4.12:

$$\begin{aligned} p(z_0 + \zeta) &= \sum_{k=0}^n a_k (z_0 + \zeta)^k = \sum_{k=0}^n a_k \sum_{l=0}^k \binom{k}{l} z_0^{k-l} \zeta^l \\ &=: \sum_{k=0}^n b_k \zeta^k = b_0 + b_\nu \zeta^\nu + b_{\nu+1} \zeta^{\nu+1} + \dots + b_n \zeta^n, \quad \zeta \in \mathbb{C}. \end{aligned}$$

Dabei gilt für die Koeffizienten $b_0, b_\nu, \dots, b_n \in \mathbb{C}$ mit einem $\nu \in \{1, \dots, n\}$:

$$b_0 = p(z_0) \neq 0, \quad b_\nu \neq 0.$$

Mit Hilfssatz 6.5.4 wählen wir nun ein $a \in \mathbb{C} \setminus \{0\}$, so dass

$$a^\nu = -b_0 b_\nu^{-1} \neq 0$$

gilt. Dann können wir schreiben

$$\begin{aligned} p(z_0 + aw) &= b_0(1 + b_0^{-1} b_\nu a^\nu w^\nu) + \sum_{k=\nu+1}^n b_k a^k w^k \\ &= b_0(1 - w^\nu) + w^{\nu+1} \left(\sum_{k=\nu+1}^n b_k a^k w^{k-\nu-1} \right), \quad w \in \mathbb{C}. \end{aligned}$$

Wir setzen nun (beachte Satz 6.4.1)

$$C := \sup_{w \in B_1(0)} \left| \frac{1}{b_0} \sum_{k=\nu+1}^n b_k a^k w^{k-\nu-1} \right| \in [0, +\infty), \quad r := \min\{1, C^{-1}\}$$

und können wie folgt abschätzen:

$$\left| \frac{p(z_0 + aw)}{b_0} \right| \leq |1 - w^\nu| + |w|^{\nu+1} C < |1 - w^\nu| + |w|^\nu \quad \text{für alle } w \in \mathbb{C} : 0 < |w| < r.$$

Speziell für ein *reelles* $w \in (0, r)$ erhalten wir daher

$$\left| \frac{p(z_0 + aw)}{b_0} \right| < 1 - w^\nu + w^\nu = 1,$$

bzw. $|p(z_0 + aw)| < |b_0| = |p(z_0)|$, also die Behauptung mit $\zeta = aw$. □

Beweis von Satz 6.5.1. Sei also p ein nicht-konstantes Polynom. Nach Hilfssatz 6.5.3 existiert dann ein $z_0 \in \mathbb{C}$ mit $|p(z_0)| \leq |p(z)|$ für alle $z \in \mathbb{C}$. Wäre $p(z_0) \neq 0$, so gäbe es nach Hilfssatz 6.5.5 ein $\zeta \in \mathbb{C}$ mit $|p(z_0)| > |p(z_0 + \zeta)|$, Widerspruch! Also muss $p(z_0) = 0$ gelten. □

Hilfssatz 6.5.6

Ist $p(z) = \sum_{k=0}^n a_k z^k$ ein Polynom vom Grad $n \in \mathbb{N}$, so besitzt p eine Darstellung

$$p(z) = (z - z_0)\hat{p}(z), \quad z \in \mathbb{C}.$$

Dabei ist z_0 Nullstelle von p , und $\hat{p}(z) = \sum_{k=0}^{n-1} b_k z^k$ ist ein Polynom mit $b_{n-1} = a_n \neq 0$.

Beweis. Da p nach Voraussetzung nicht konstant ist (siehe Bemerkung 6.5.2 (c)), existiert gemäß Satz 6.5.1 eine Nullstelle $z_0 \in \mathbb{C}$ von p ; wir können dann schreiben

$$p(z) = p(z) - p(z_0) = \sum_{k=1}^n a_k (z^k - z_0^k), \quad z \in \mathbb{C}.$$

Wir beachten die Relation

$$z^k - z_0^k = (z - z_0) \sum_{l=0}^{k-1} z_0^{k-1-l} z^l, \quad z \in \mathbb{C}, \quad k \in \mathbb{N},$$

die man leicht mit vollständiger Induktion oder mit Satz 2.4.8 beweist (\rightarrow Nachrechnen). Einsetzen liefert

$$p(z) = (z - z_0) \left(\sum_{k=1}^n \sum_{l=0}^{k-1} a_k z_0^{k-1-l} z^l \right) = (z - z_0) (b_0 + \dots + b_{n-1} z^{n-1}), \quad z \in \mathbb{C},$$

mit Zahlen $b_0, \dots, b_{n-1} \in \mathbb{C}$. Insbesondere sehen wir noch $b_{n-1} = a_n$. \square

Mit Hilfssatz 6.5.6 lässt sich von jedem Polynom mit Grad $n \geq 1$ ein Faktor $z - z_0$ abspalten; der verbleibende Faktor ist wieder ein Polynom, nun mit Grad $n - 1$; offenbar lässt sich dieses Argument iterieren, bis der Grad 0 erreicht ist. Wir erhalten damit sofort den

Satz 6.5.7: Faktorisierung von Polynomen

Jedes Polynom $p(z) = \sum_{k=0}^n a_k z^k$ vom Grad $n \in \mathbb{N}$ besitzt $\nu \in \{1, \dots, n\}$ paarweise verschiedene Nullstellen $z_1, \dots, z_\nu \in \mathbb{C}$, und wir haben die Darstellung

$$p(z) = a_n (z - z_1)^{\alpha_1} (z - z_2)^{\alpha_2} \cdots (z - z_\nu)^{\alpha_\nu}, \quad z \in \mathbb{C}.$$

Dabei heißt $\alpha_j \in \mathbb{N}$ *Vielfachheit* der Nullstelle z_j , $j = 1, \dots, \nu$, und es gilt $\alpha_1 + \dots + \alpha_\nu = n$.

Bemerkung 6.5.8. Zählt man jede Nullstelle entsprechend ihrer Vielfachheit, so erhält man die Aussage: *Jedes komplexe Polynom vom Grad $n \in \mathbb{N}$ hat genau n (nicht notwendig verschiedene) Nullstellen.*

Folgerung 6.5.9

Ein Polynom vom Grad $n \in \mathbb{N}_0$, das in $n + 1$ paarweise verschiedenen Punkten verschwindet, ist identisch Null.

Beweis. Nach Satz 6.5.7 verschwindet ein Polynom vom Grad $n \in \mathbb{N}$ in höchstens n paarweise verschiedenen Punkten. Also muss $n = 0$ gelten, d.h. $p = \text{const.}$ Da aber p nach Voraussetzung eine Nullstelle besitzen soll, folgt $p = 0$ auf \mathbb{C} . \square

6.6. Funktionenfolgen und gleichmäßige Konvergenz

Wir betrachten nun Folgen $\{f_n\}_n$ von Funktionen $f_n : D \rightarrow \mathbb{R}^d$, die alle auf derselben nichtleeren Menge $D \subset \mathbb{R}^m$ erklärt sind; wir sprechen dann von *Funktionenfolgen*.

Definition 6.6.1

Eine Funktionenfolge $\{f_n\}_n$ mit $f_n : D \rightarrow \mathbb{R}^d$, $n \in \mathbb{N}$, heißt *punktweise konvergent* auf $D \subset \mathbb{R}^m$, wenn die Punktfolge $\{f_n(x)\}_n \subset \mathbb{R}^d$ für jedes $x \in D$ konvergiert. Die Grenzwerte

$$f(x) := \lim_{n \rightarrow \infty} f_n(x), \quad x \in D,$$

definieren dann eine Funktion $f : D \rightarrow \mathbb{R}^d$, den sogenannten *punktweisen Grenzwert* oder *Limes* der Funktionenfolge $\{f_n\}_n$. Schreibweise: $f_n \rightarrow f (n \rightarrow \infty)$ auf D .

Beispiel 6.6.2. Wir betrachten $f_n(x) := x^n$, $x \in D = [0, 1]$. $\{f_n\}_n$ konvergiert punktweise gegen die Funktion

$$f(x) := \begin{cases} 0, & x \in [0, 1) \\ 1, & x = 1 \end{cases}.$$

Das Beispiel zeigt, dass der punktweise Limes einer Folge stetiger Funktionen nicht wieder stetig sein muss. Um beim Grenzübergang in der Klasse der stetigen Funktionen zu bleiben, benötigen wir einen stärkeren Konvergenzbegriff, der auf WEIERSTRASS zurückgeht:

Definition 6.6.3: Gleichmäßige Konvergenz

Eine Folge $\{f_n\}_n$ von Funktionen $f_n : D \rightarrow \mathbb{R}^d$ mit $D \subset \mathbb{R}^m$ heißt *gleichmäßig konvergent* gegen $f : D \rightarrow \mathbb{R}^d$, in Zeichen $f_n \rightrightarrows f (n \rightarrow \infty)$ auf D , wenn zu jedem $\varepsilon > 0$ ein

$N = N(\varepsilon) \in \mathbb{N}$ existiert mit

$$\|f_n(x) - f(x)\| < \varepsilon \quad \text{für alle } x \in D \text{ und } n \geq N(\varepsilon). \quad (6.6.1)$$

Bemerkung 6.6.4. Formel (6.6.1) gilt in jedem Punkt $x \in D$ auch für den punktweisen Limes einer Funktionenfolge, allerdings mit einem i.A. von x abhängigen $N = N(\varepsilon, x) \in \mathbb{N}$. Jede gleichmäßig konvergente Funktionenfolge konvergiert auch punktweise.

Satz 6.6.5: WEIERSTRASSSCHER KONVERGENZSATZ

Die Folge $\{f_n\}_n$ stetiger Funktionen $f_n : D \rightarrow \mathbb{R}^d$ konvergiere auf $D \subset \mathbb{R}^m$ gleichmäßig gegen $f : D \rightarrow \mathbb{R}^d$. Dann ist f stetig auf D .

Beweis. Nach Definition 6.6.3 gibt es zu beliebig gewähltem $\varepsilon > 0$ ein $N = N(\varepsilon) \in \mathbb{N}$ mit

$$\|f_N(x) - f(x)\| < \frac{\varepsilon}{3} \quad \text{für alle } x \in D. \quad (6.6.2)$$

Sei $x_0 \in D$ gewählt. Da f_N stetig ist, finden wir ein $\delta = \delta(\varepsilon) > 0$, so dass gilt

$$\|f_N(x) - f_N(x_0)\| < \frac{\varepsilon}{3} \quad \text{für alle } x \in D \text{ mit } \|x - x_0\| < \delta. \quad (6.6.3)$$

Mit der Dreiecksungleichung erhalten wir aus (6.6.2) und (6.6.3)

$$\begin{aligned} \|f(x) - f(x_0)\| &\leq \|f(x) - f_N(x)\| + \|f_N(x) - f_N(x_0)\| + \|f_N(x_0) - f(x_0)\| \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon \quad \text{für alle } x \in D \text{ mit } \|x - x_0\| < \delta, \end{aligned}$$

wie behauptet. □

Analog zu Punktfolgen im \mathbb{R}^d haben wir für Funktionenfolgen ein zur gleichmäßigen Konvergenz äquivalentes gleichmäßiges CAUCHY-Kriterium:

Satz 6.6.6: CAUCHYS KONVERGENZKRITERIUM FÜR FUNKTIONENFOLGEN

Sei $\{f_n\}_n$ eine Folge von Funktionen $f_n : D \rightarrow \mathbb{R}^d$, $D \subset \mathbb{R}^m$. Dann konvergiert $\{f_n\}_n$ genau dann gleichmäßig (gegen ein $f : D \rightarrow \mathbb{R}^d$), wenn zu jedem $\varepsilon > 0$ ein $N = N(\varepsilon) \in \mathbb{N}$ existiert mit

$$\|f_n(x) - f_k(x)\| < \varepsilon \quad \text{für alle } x \in D \text{ und } n, k \geq N(\varepsilon). \quad (6.6.4)$$

Beweis. Punktweise Anwendung von Satz 5.2.3 (\rightarrow Ergänzungen). □

Definition 6.6.7

Ist $\{f_k\}_k$ eine Folge von Funktionen $f_k : D \rightarrow \mathbb{C}$, $D \subset \mathbb{R}^m$, so heißt die zugehörige **Funktionenreihe** $\sum_{k=1}^{\infty} f_k$ *punktweise* bzw. *gleichmäßig konvergent*, wenn die Folge $\{s_n\}_n$ der Partialsummen,

$$s_n(x) := \sum_{k=1}^n f_k(x), \quad x \in D, \quad n \in \mathbb{N},$$

punktweise bzw. gleichmäßig auf D konvergiert.

Beispiel 6.6.8. Die in Abschnitt A.4.4 untersuchten Potenzreihen $\mathcal{P}(z) = \sum_{k=0}^{\infty} a_k z^k$ sind Beispiele von Funktionenreihen; die Partialsummen $\mathcal{P}_n(z) = \sum_{k=0}^n a_k z^k$, $n \in \mathbb{N}$, sind dann komplexe Polynome. Ist $R \in (0, +\infty) \cup \{+\infty\}$ der Konvergenzradius einer Potenzreihe \mathcal{P} , so konvergiert die Reihe im Konvergenzgebiet $D = K_R(0) = \{z \in \mathbb{C} : |z| < R\}$ (und manchmal auch in Randpunkten von D) punktweise; siehe Satz 4.4.4 und Bemerkung 4.4.5.

Satz 6.6.9: Majorantenkriterium für Funktionenreihen

Sei $D \subset \mathbb{R}^m$ und $\{f_k\}_k$ eine Folge von Funktionen $f_k : D \rightarrow \mathbb{C}$. Ferner sei $\{c_k\}_k \subset \mathbb{R}$ eine Punktfolge mit der Eigenschaft

$$|f_k(x)| \leq c_k \quad \text{für alle } x \in D \text{ und alle } k \in \mathbb{N}. \quad (6.6.5)$$

Falls dann $\sum_{k=1}^{\infty} c_k$ konvergiert, so konvergiert $\sum_{k=1}^{\infty} f_k$ gleichmäßig auf D . Die Reihe $\sum_k c_k$ heißt *Majorante* von $\sum_k f_k$.

Beweis. Analog zu Satz 4.2.6 mit Hilfe von Satz 6.6.6 (\rightarrow Ergänzungen). \square

Als Konsequenz erhalten wir das folgende wichtige Resultat:

Satz 6.6.10

Es seien $\{a_k\}_k \subset \mathbb{C}$, $R \in (0, +\infty) \cup \{+\infty\}$ und $\mathcal{P}(z) := \sum_{k=0}^{\infty} a_k z^k$ eine in $K_R(0)$ konvergente Potenzreihe. Dann ist $\mathcal{P} : K_R(0) \rightarrow \mathbb{C}$ stetig.

Beweis. Ist $z_0 \in K_R(0)$ beliebig, so folgt $z_0 \in K_{R_0}(0)$ mit $R_0 := \frac{1}{2}(|z_0| + R) < R$. Nun ist für $D := K_{R_0}(0)$ die Folge $\{a_k z^k\}_k \subset C^0(D, \mathbb{C})$ durch $\{|a_k| R_0^k\}_k \subset \mathbb{R}$ majorisiert im Sinne von (6.6.5), und nach Satz 4.4.6 konvergiert $\sum_k |a_k| R_0^k$. Satz 6.6.9 liefert also die gleichmäßige Konvergenz der Potenzreihe \mathcal{P} auf D . Folglich ist \mathcal{P} stetig auf D gemäß Satz 6.6.5, also

insbesondere auch in $K_r(z_0) \subset D$ mit $r = \frac{1}{2}(R - |z_0|) > 0$. Da $z_0 \in K_R(0)$ beliebig war und die Stetigkeit eine lokale Eigenschaft ist (siehe Bemerkung 6.2.2 (b)), folgt $\mathcal{P} \in C^0(K_R(0), \mathbb{C})$, wie behauptet. \square

Folgerung 6.6.11

Die *komplexe Exponentialfunktion* oder *Exponentialreihe*

$$\exp z = e^z := \sum_{k=0}^{\infty} \frac{z^k}{k!}$$

ist auf ganz \mathbb{C} stetig.

Beweis. Sofort aus Satz 6.6.10 und Beispiel 4.2.12 (c). \square

Wir werden \exp in Kapitel 8 genauer untersuchen und hieraus auch die weiteren *elementaren Funktionen* wie Sinus, Cosinus, Hyperbelfunktionen, Logarithmus und allgemeine Potenz ableiten.

Teil III.

Differential- und Integralrechnung in einer Veränderlichen

In Teil III wenden wir die Idee des Grenzwerts an, um einen Ableitungs- und einen Integralbegriff für Funktionen *einer* reellen Veränderlichen zu etablieren. In Kapitel 7 behandeln wir zunächst Ableitungen und differenzierbare Funktionen. Kapitel 8 ist dann den elementaren Funktionen gewidmet (*wird fortgesetzt*).

Kapitel 7.

Die gewöhnliche Ableitung

Ableitungen von Funktionen sind ein zentrales Hilfsmittel in der Analysis, mit ihnen lassen sich Richtungen, Krümmungen, Geschwindigkeiten, ... der mathematisch modellierten Objekte beschreiben. Im vorliegenden Kapitel führen wir Ableitungen für Funktionen einer reellen Veränderlichen mittels eines Grenzwerts ein.

In Abschnitt 7.1 werden die grundlegenden Definitionen gelegt, Rechenregeln für Ableitungen stellen wir in Abschnitt 7.2 bereit. Die wichtige Aufgabe der Suche nach lokalen Extrema ist zentrales Thema des übrigen Kapitels: In Abschnitt 7.3 geben wir die Bedingungen 1. Ordnung an und thematisieren den Zusammenhang zum Mittelwertsatz. In Abschnitt 7.4 definieren wir dann höhere Ableitungen und wenden diese insbesondere zur Herleitung der Extremalbedingungen 2. Ordnung an.

7.1. Differenzierbarkeit, der Ableitungsbegriff

Wir untersuchen Funktionen einer reellen Veränderlichen $f : I \rightarrow \mathbb{R}^d$ für $d \in \mathbb{N}$; hier und im gesamten Kapitel seien $I, J, \dots \subset \mathbb{R}$ immer Intervalle beliebigen Typs. Wir beginnen mit einem der wichtigsten Begriffsbildungen der Analysis überhaupt:

Definition 7.1.1: Ableitung

Eine Funktion $f : I \rightarrow \mathbb{R}^d$ heißt *differenzierbar* an der Stelle $x_0 \in I$, falls der Grenzwert

$$f'(x_0) := \lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h} = \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} \quad (7.1.1)$$

existiert. $f'(x_0)$ heißt (*erste*) *Ableitung* oder *Differentialquotient* von f an der Stelle x_0 . Alternativ schreiben wir auch

$$f'(x_0) = \frac{df}{dx}(x_0) = Df(x_0).$$

Die Funktion $f : I \rightarrow \mathbb{R}^d$ heißt *differenzierbar (auf I)*, wenn f in jedem Punkt $x_0 \in I$ differenzierbar ist.

Bemerkung 7.1.2. (a) Um von den *partiellen Ableitungen* für Funktionen mehrerer Veränderlicher zu unterscheiden, sprechen wir bei f' auch von der *gewöhnlichen Ableitung*.

(b) Zum Beispiel sind die Funktionen $f(x) := x$, $g(x) := c$ mit einer Konstanten $c \in \mathbb{R}$ für alle $x \in \mathbb{R}$ differenzierbar und es gilt

$$f'(x) = 1, \quad g'(x) = 0 \quad \text{für alle } x \in \mathbb{R}.$$

(c) Eine Funktion $f = (f_1, \dots, f_d) : I \rightarrow \mathbb{R}^d$ ist genau dann in $x_0 \in I$ differenzierbar, wenn alle Komponentenfunktionen f_1, \dots, f_d in x_0 differenzierbar sind; siehe Folgerung 6.1.3. In diesem Fall gilt

$$f'(x_0) = (f'_1(x_0), \dots, f'_d(x_0)).$$

(d) Wie die Stetigkeit ist auch die Differenzierbarkeit (in einem Punkt) eine lokale Eigenschaft.

Geometrische Interpretation

► Der *Differenzenquotient*

$$\Delta_h f(x_0) := \frac{f(x_0 + h) - f(x_0)}{h}$$

einer Funktion $f : I \rightarrow \mathbb{R}$ gibt die Steigung der *Sekante* an den Graph $\Gamma(f)$ durch $(x_0, f(x_0))$ und $(x_0 + h, f(x_0 + h))$ an. Bei Grenzübergang $h \rightarrow 0$ geht die Sekante in die *Tangente*

$$T := \{(x, y) \in \mathbb{R}^2 : y = f'(x_0)(x - x_0) + f(x_0)\}$$

an $\Gamma(f)$ im Punkt $(x_0, f(x_0))$ über; $f'(x_0)$ ist die Steigung der Tangente.

► Eine Abbildung $f : I \rightarrow \mathbb{R}^d$ veranschaulicht man sich als *Kurve* im \mathbb{R}^d . Dabei sind $\Delta_h f(t_0)$ *Sekantenvektoren* in \mathbb{R}^d und $f'(t_0)$ wird als *Tangentenvektor* an die Kurve im Punkt $f(t_0)$ interpretiert (und abgetragen).

Satz 7.1.3

Ist $f : I \rightarrow \mathbb{R}^d$ gegeben, so sind die folgenden Aussagen äquivalent:

- (i) f ist in $x_0 \in I$ differenzierbar.
- (ii) Es existiert eine in x_0 stetige Funktion $\psi : I \rightarrow \mathbb{R}^d$, so dass gilt

$$f(x) = f(x_0) + (x - x_0)\psi(x) \quad \text{für alle } x \in I. \quad (7.1.2)$$

ψ ist dann eindeutig durch f festgelegt und es gilt $\psi(x_0) = f'(x_0)$.

Beweis. ▶ „(i)⇒(ii)“: Sei f in x_0 differenzierbar. Wir setzen dann

$$\psi(x) := \begin{cases} \frac{f(x) - f(x_0)}{x - x_0}, & \text{falls } x \in I \setminus \{x_0\} \\ f'(x_0), & \text{für } x = x_0 \end{cases}.$$

Offenbar ist ψ in x_0 stetig (vgl. Satz 6.2.3) mit $\psi(x_0) = f'(x_0)$, und Umstellen liefert die gesuchte Darstellung (7.1.2).

▶ „(i)⇐(ii)“: Haben wir umgekehrt (7.1.2), so liefert Umstellen für $x \neq x_0$:

$$\frac{f(x) - f(x_0)}{x - x_0} = \psi(x) \rightarrow \psi(x_0) \quad (x \rightarrow x_0),$$

also die Differenzierbarkeit von f in x_0 mit $f'(x_0) = \psi(x_0)$. □

Bemerkung 7.1.4. Die Darstellung (7.1.2) liefert eine *lineare Approximation* von f durch

$$L(x) := f(x_0) + (x - x_0)f'(x_0), \quad x \in \mathbb{R};$$

der Graph $\Gamma(L)$ ist dabei die Tangente an $\Gamma(f)$. Mit L finden wir nämlich die zu (7.1.2) äquivalente Darstellung

$$f(x) = L(x) + (x - x_0)\varphi(x) \quad \text{für alle } x \in I. \quad (7.1.3)$$

Dabei ist $\varphi := \psi - \psi(x_0) : I \rightarrow \mathbb{R}^d$ in x_0 stetig mit $\varphi(x_0) = 0$.

Folgerung 7.1.5

Eine in $x_0 \in I$ differenzierbare Funktion $f : I \rightarrow \mathbb{R}^d$ ist in x_0 auch stetig.

Beweis. Sofort aus Darstellung (7.1.2) und Satz 6.2.4. □

Bemerkung 7.1.6. Die Umkehrung von Folgerung 7.1.5 gilt nicht, wie etwa das Beispiel $f(x) := |x|$ im Punkt $x_0 = 0$ zeigt. Es gibt sogar stetige, nirgends differenzierbare Funktionen; siehe [Hil1] Abschnitt 3.1, Beispiel [6].

7.2. Rechenregeln für Ableitungen

Wir geben eine Reihe wichtiger Rechenregeln für differenzierbare Funktionen an. Den ersten Satz formulieren wir für komplexwertige Funktionen; siehe aber auch Bemerkung 7.2.2.

Satz 7.2.1: Rechenregeln differenzierbarer Funktionen

Seien $f, g : I \rightarrow \mathbb{C}$ in einem Punkt $x_0 \in I$ differenzierbare Funktionen. Dann gelten:

- (a) **Linearität:** Für beliebige $\lambda, \mu \in \mathbb{C}$ ist auch die Funktion $\lambda f + \mu g : I \rightarrow \mathbb{C}$ differenzierbar in x_0 mit der Ableitung

$$(\lambda f + \mu g)'(x_0) = \lambda f'(x_0) + \mu g'(x_0).$$

- (b) **Produktregel:** Auch die Funktion $fg : I \rightarrow \mathbb{C}$ ist differenzierbar in x_0 mit der Ableitung

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

- (c) **Quotientenregel:** Gilt $g'(x_0) \neq 0$, so ist auch $\frac{f}{g}$ differenzierbar in x_0 mit der Ableitung

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

Beweis. Nach Satz 7.1.3 haben wir die Darstellungen

$$f(x) = f(x_0) + (x - x_0)\psi(x), \quad g(x) = g(x_0) + (x - x_0)\chi(x),$$

mit in x_0 stetigen Funktionen $\psi, \chi : I \rightarrow \mathbb{C}$, die $\psi(x_0) = f'(x_0)$, $\chi(x_0) = g'(x_0)$ erfüllen. Damit folgen für $x \in I$:

$$\lambda f(x) + \mu g(x) = [\lambda f(x_0) + \mu g(x_0)] + (x - x_0)[\lambda \psi(x) + \mu \chi(x)],$$

$$f(x)g(x) = [f(x_0)g(x_0)] + (x - x_0)[\psi(x)g(x_0) + f(x_0)\chi(x) + (x - x_0)\psi(x)\chi(x)],$$

$$\frac{f(x)}{g(x)} = \frac{f(x_0)}{g(x_0)} + (x - x_0) \frac{\psi(x)g(x_0) - f(x_0)\chi(x)}{g(x)g(x_0)},$$

letzteres unter der Bedingung von (c) und nur für $x \in (x_0 - r, x_0 + r) \cap I$, wobei $r > 0$ so klein gewählt sei, dass $g \neq 0$ in $(x_0 - r, x_0 + r) \cap I$ gilt (siehe Folgerung 7.1.5). Wiederum Satz 7.1.3 und Satz 6.2.4 liefern die Behauptung. \square

Bemerkung 7.2.2. Mit dem gleichen Beweis erhält man die Linearität, Satz 7.2.1 (a), auch für Funktionen $f, g : I \rightarrow \mathbb{R}^d$, dann mit $\lambda, \mu \in \mathbb{R}$. Die Produktregel, Satz 7.2.1 (b), nimmt für solche Funktionen die Form

$$\langle f, g \rangle'(x_0) = \langle f'(x_0), g(x_0) \rangle + \langle f(x_0), g'(x_0) \rangle \quad (7.2.1)$$

an (\rightarrow Nachrechnen).

Beispiel 7.2.3. Es gilt (\rightarrow Ergänzungen):

$$\frac{d}{dx}(x^\nu) = \nu x^{\nu-1} \quad \text{für alle } \nu \in \mathbb{Z} \quad \text{und } x \in \mathbb{R} \setminus \{0\}.$$

Wir untersuchen nun die Komposition zweier differenzierbarer Funktionen:

Satz 7.2.4: Kettenregel

Seien $I, J \subset \mathbb{R}$ Intervalle und $f : I \rightarrow \mathbb{R}$, $g : J \rightarrow \mathbb{R}^d$ zwei Funktionen mit $f(I) \subset J$. Falls f in $x_0 \in I$ und g in $y_0 := f(x_0) \in J$ differenzierbar sind, so ist auch die Komposition $h := g \circ f : I \rightarrow \mathbb{R}^d$ in x_0 differenzierbar, und es gilt

$$h'(x_0) = g'(f(x_0))f'(x_0).$$

Beweis. Satz 7.1.3 entnehmen wir Darstellungen

$$f(x) = f(x_0) + (x - x_0)\psi(x), \quad g(y) = g(y_0) + (y - y_0)\chi(y)$$

mit in x_0 bzw. $y_0 = f(x_0)$ stetigen Funktionen $\psi : I \rightarrow \mathbb{R}$, $\chi : J \rightarrow \mathbb{R}^d$, für die $\psi(x_0) = f'(x_0)$ bzw. $\chi(y_0) = g'(y_0)$ gilt. Es folgt also

$$\begin{aligned} h(x) &= g(f(x)) = g(f(x_0)) + (f(x) - f(x_0))\chi(f(x)) \\ &= h(x_0) + (x - x_0)[\chi(f(x))\psi(x)]. \end{aligned}$$

Da die Funktion $\chi(f(x))\psi(x)$ nach Satz 6.2.6 und Folgerung 7.1.5 wieder stetig ist in x_0 , ist $h = g \circ f$ nach Satz 7.1.3 differenzierbar in x_0 und es gilt

$$h'(x_0) = \chi(f(x))\psi(x)|_{x=x_0} = g'(f(x_0))f'(x_0),$$

wie behauptet. □

Auch für die Umkehrfunktion einer differenzierbaren, injektiven Funktion $f : I \rightarrow \mathbb{R}$ erhalten wir eine schöne Aussage:

Satz 7.2.5: Ableitung der Umkehrfunktion

Sei $f : I \rightarrow \mathbb{R}$ eine stetige Funktion, die das Intervall $I \subset \mathbb{R}$ bijektiv auf $I^* := f(I)$ abbilde. Ist dann f in $x_0 \in I$ differenzierbar und gilt $f'(x_0) \neq 0$, so ist auch die Umkehrfunktion

$f^{-1} : I^* \rightarrow \mathbb{R}$ in $y_0 := f(x_0) \in I^*$ differenzierbar und es gilt

$$(f^{-1})'(y_0) = \frac{1}{f'(x_0)}.$$

Beweis. → Ergänzungen. □

Beispiel 7.2.6. Zu beliebigem $n \in \mathbb{N}$ ist die Umkehrfunktion zu $f(x) := x^n, x \in [0, +\infty)$, nach Beispiel 6.3.9 gegeben durch $f^{-1}(y) = \sqrt[n]{y}, y \in [0, +\infty)$. Für $x > 0$ gilt $f'(x) = nx^{n-1} > 0$ gemäß Beispiel 7.2.3, so dass Satz 7.2.5 liefert

$$(\sqrt[n]{y})' = \frac{1}{n(\sqrt[n]{y})^{n-1}} = \frac{1}{n}y^{\frac{1}{n}-1}, \quad y > 0.$$

Für die Potenzfunktion $g(x) := x^q = (\sqrt[s]{x})^r, x > 0$, mit einem $q = \frac{r}{s} \in \mathbb{Q} (r \in \mathbb{Z}, s \in \mathbb{N})$ folgt somit nach der Kettenregel, Satz 7.2.4:

$$\frac{d}{dx}(x^q) = \frac{d}{dx}(\sqrt[s]{x})^r = [r(\sqrt[s]{x})^{r-1}] \left[\frac{1}{s}x^{\frac{1}{s}-1} \right] = qx^{q-1}, \quad x > 0.$$

7.3. Lokale Extrema und der Mittelwertsatz

Ein wichtiges Teilgebiet der Analysis ist die Behandlung von Extremwertaufgaben. Extrema sind dabei allgemein wie folgt erklärt:

Definition 7.3.1

Sei $f : D \rightarrow \mathbb{R}$ auf einer Menge $D \subset \mathbb{R}^n$ mit $n \in \mathbb{N}$ erklärt. Wir sagen:

- ▶ f hat in $x_0 \in D$ ein *lokales Minimum* (bzw. *lokales Maximum*), wenn ein $r > 0$ so existiert, dass gilt

$$f(x) \geq f(x_0) \quad (\text{bzw. } f(x) \leq f(x_0)) \quad \text{für alle } x \in D \cap B_r(x_0). \quad (7.3.1)$$

Gilt in (7.3.1) die strikte Ungleichung für $x \neq x_0$, so hat f in x_0 ein *striktes lokales Minimum* (bzw. *striktes lokales Maximum*).

- ▶ f hat in x_0 ein *globales Minimum* (bzw. *globales Maximum*), wenn (7.3.1) für alle $x \in D$ gilt. Entsprechend werden die Begriffe *striktes globales Minimum* und *striktes globales Maximum* erklärt.
- ▶ Der Punkt x_0 heißt *lokale, strikte lokale, ... Minimalstelle* bzw. *lokale, strikte lokale ... Maximalstelle*.

Bemerkung 7.3.2. Zusammenfassend heißen Minima und Maxima auch *Extrema* und x_0 wird auch *Extremalstelle* genannt. Als Synonym für *lokal* wird auch *relativ* verwendet, statt *global* sagen wir auch *absolut*.

Wir wollen Extrema von Funktionen $f : I \rightarrow \mathbb{R}$ auf Intervallen $I \subset \mathbb{R}$ mit den Methoden der Differentialrechnung bestimmen. Der folgende fundamentale Satz sagt uns, welche Punkte dafür überhaupt infrage kommen.

Satz 7.3.3: FERMAT, notwendige Extremalbedingung 1. Ordnung

Besitzt $f : I \rightarrow \mathbb{R}$ in einem inneren Punkt $x_0 \in \text{int } I$ des Intervalls $I \subset \mathbb{R}$ ein lokales Extremum und ist f in x_0 differenzierbar, so gilt $f'(x_0) = 0$.

Beweis. O.B.d.A. sei x_0 Minimalstelle (sonst gehen wir zu $-f$ über). Da x_0 innerer Punkt ist, gibt es ein $\varepsilon > 0$, so dass $(x_0 - \varepsilon, x_0 + \varepsilon) \subset I$ gilt. Somit folgt (siehe Bemerkung 6.3.2 (a)):

$$0 \geq \lim_{x \rightarrow x_0^-} \frac{f(x) - f(x_0)}{x - x_0} = f'(x_0) = \lim_{x \rightarrow x_0^+} \frac{f(x) - f(x_0)}{x - x_0} \geq 0,$$

also $f'(x_0) = 0$. □

Bemerkung 7.3.4. ▶ Betrachte $f(x) := x$, $I = [0, 1]$. Dann ist $x_0 = 0$ (sogar globales) Minimum, aber es gilt $f'(0) = 1$. Für $x_0 \in \partial I$ wird Satz 7.3.3 also i.A. falsch.

- ▶ Die Bedingung $f'(x_0) = 0$ ist nicht hinreichend für ein Extremum, wie etwa das Beispiel $f(x) := x^3$, $x \in \mathbb{R}$, mit $f'(0) = 0$ zeigt.

Definition 7.3.5

Ist $f : I \rightarrow \mathbb{R}$ im inneren Punkt $x_0 \in \text{int } I$ differenzierbar und gilt $f'(x_0) = 0$, so heißt x_0 *kritischer* oder *stationärer Punkt* von f .

Bemerkung 7.3.6. Satz 7.3.3 besagt also: *Jede innere lokale Extremalstelle von f ist stationär.*

Geometrisch bedeutet dies, dass die Tangente $T = \{(x, y) : y = f(x_0) + f'(x_0)(x - x_0)\}$ an $\Gamma(f)$ im Punkt $(x_0, f(x_0))$ parallel zur x -Achse verläuft.

Der nächste Satz gibt eine hinreichende Bedingung für die Existenz eines kritischen Punktes an:

Satz 7.3.7: Satz von ROLLE

Sei $f : [a, b] \rightarrow \mathbb{R}$ stetig in $[a, b]$ und differenzierbar in (a, b) . Gilt zusätzlich $f(a) = f(b)$, so existiert ein $\xi \in (a, b)$ mit der Eigenschaft $f'(\xi) = 0$.

Beweis. Ist f konstant auf $[a, b]$, so gilt $f'(\xi) = 0$ für jedes $\xi \in (a, b)$. Sei also f nicht konstant. Dann existiert ein $x_0 \in (a, b)$ mit $f(x_0) \neq f(a)$. O.B.d.A. sei $f(x_0) > f(a)$, so dass folgt $\sup_{[a,b]} f > f(a) = f(b)$. Nach dem WEIERSTRASSschen Hauptsatz, Satz 6.4.4, nimmt also f ihr (globales) Maximum in einem inneren Punkt $\xi \in (a, b)$ an und nach Satz 7.3.3 gilt $f'(\xi) = 0$. \square

Die Bedingung $f(a) = f(b)$ kann im Satz von ROLLE natürlich nicht fallengelassen werden, wie die lineare Funktion $f(x) = mx + c$, $x \in [a, b]$, mit $m \neq 0$ zeigt. Der Satz besitzt aber eine Verallgemeinerung, die eines der meistgebrauchten Hilfsmittel der Differential- und Integralrechnung darstellt:

Satz 7.3.8: Mittelwertsatz

Es sei $f : [a, b] \rightarrow \mathbb{R}$ stetig in $[a, b]$ und differenzierbar in (a, b) . Dann gibt es ein $\xi \in (a, b)$, so dass gilt

$$f(b) - f(a) = f'(\xi)(b - a). \quad (7.3.2)$$

Bemerkung 7.3.9. Geometrisch heißt das, dass ein $\xi \in (a, b)$ so existiert, dass die Tangente an $(\xi, f(\xi))$ parallel zur Sekante durch $(a, f(a))$ und $(b, f(b))$ verläuft.

Satz 7.3.8 ergibt sich sofort als Spezialfall (mit $g = \text{Id}$) des folgenden

Satz 7.3.10: Allgemeiner Mittelwertsatz

Gegeben seien zwei stetige Funktionen $f, g : [a, b] \rightarrow \mathbb{R}$, die differenzierbar auf (a, b) seien. Weiter gelte $g' \neq 0$ auf (a, b) . Dann existiert ein $\xi \in (a, b)$, so dass gilt

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

Beweis. Nach dem ROLLESchen Satz gilt $g(a) \neq g(b)$. Wir betrachten die Hilfsfunktion

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

Offenbar ist φ stetig in $[a, b]$, differenzierbar in (a, b) und es gilt $\varphi(a) = \varphi(b) = f(a)$. Wieder nach dem ROLLESchen Satz existiert somit ein $\xi \in (a, b)$ mit

$$0 = \varphi'(\xi) = f'(\xi) - \frac{f(b) - f(a)}{g(b) - g(a)} g'(\xi),$$

also nach Umstellen die Behauptung. \square

In engem Zusammenhang mit der Suche nach Extrema steht das Monotonieverhalten einer Funktion, siehe unten stehenden Satz 7.3.13. Zunächst halten wir fest:

Folgerung 7.3.11: Monotonieverhalten

Ist $f : [a, b] \rightarrow \mathbb{R}$ stetig in $[a, b]$ und differenzierbar in (a, b) , so haben wir:

- (a) Gilt $f' > 0$ (bzw. $f' \geq 0$, $f' < 0$, $f' \leq 0$) auf (a, b) , so ist f streng monoton wachsend (bzw. monoton wachsend, streng monoton fallend, monoton fallend) auf $[a, b]$.
- (b) Ist umgekehrt f monoton wachsend (bzw. monoton fallend) auf $[a, b]$, so gilt $f' \geq 0$ (bzw. $f' \leq 0$) auf (a, b) .
- (c) Es gilt $f' \equiv 0$ in (a, b) genau dann, wenn $f \equiv \text{const}$ auf $[a, b]$ richtig ist.

Bemerkung 7.3.12. Strenge Monotonie impliziert *nicht* $f' > 0$ bzw. $f' < 0$ auf (a, b) , wie etwa das Beispiel $f(x) = x^3$, $x \in \mathbb{R}$, zeigt.

Beweis von Folgerung 7.3.11. (a) Wir betrachten nur den Fall $f' > 0$ auf (a, b) ; die anderen Aussagen folgen analog. Seien $x_1, x_2 \in [a, b]$ mit $x_1 < x_2$ gewählt. Nach Satz 7.3.8 existiert dann ein $\xi \in (x_1, x_2)$ mit

$$f(x_2) - f(x_1) = f'(\xi)(x_2 - x_1) > 0,$$

also $f(x_1) < f(x_2)$, wie behauptet.

- (b) Sei f monoton wachsend (bzw. fallend). Dann gilt für beliebiges $x_0 \in (a, b)$ und hinreichend kleines $h \neq 0$:

$$\frac{f(x_0 + h) - f(x_0)}{h} \geq 0 \quad (\text{bzw. } \leq 0).$$

Grenzübergang $h \rightarrow 0$ liefert die Behauptung.

- (c) Ist f konstant, so verschwindet die Ableitung bekanntlich identisch. Ist umgekehrt $f' \equiv 0$ auf (a, b) , so ist f nach (a) sowohl monoton wachsend als auch fallend auf $[a, b]$ und somit konstant. \square

Mit Folgerung 7.3.11 können wir einfache hinreichende Kriterien für das Vorliegen oder Nicht-Vorliegen eines Extremums angeben:

Satz 7.3.13

Sei $f \in C^0(I)$ in einer Umgebung $(x_0 - r, x_0 + r) \subset I$ eines kritischen Punktes $x_0 \in \text{int } I$ differenzierbar. Dann gelten:

- (a) Falls $f' < 0$ in $(x_0 - r, x_0)$ und $f' > 0$ in $(x_0, x_0 + r)$ richtig ist, so hat f in x_0 ein striktes lokales Minimum.
- (b) Falls $f' > 0$ in $(x_0 - r, x_0)$ und $f' < 0$ in $(x_0, x_0 + r)$ richtig ist, so hat f in x_0 ein striktes lokales Maximum.
- (c) Falls $f' < 0$ oder $f' > 0$ für alle $x \in (x_0 - r, x_0 + r) \setminus \{x_0\}$ gilt, so hat f in x_0 *kein* lokales Extremum.

Beweis. Folgerung 7.3.11 (i) entnehmen wir

$$\begin{aligned} f'(x) \leq 0 \text{ für } x_0 - r < x < x_0 &\Rightarrow f(x_0) \leq f(x) \text{ für } x_0 - r < x < x_0, \\ f'(x) \leq 0 \text{ für } x_0 < x < x_0 + r &\Rightarrow f(x_0) \geq f(x) \text{ für } x_0 < x < x_0 + r. \end{aligned}$$

Das liefert unmittelbar die Behauptungen. \square

Bemerkung 7.3.14. ▶ Zu Satz 7.3.13 (a), (b) analoge Aussagen gelten offenbar auch im „nicht-strikten“ Fall.

- ▶ Die Kriterien aus Satz 7.3.13 haben den Nachteil, dass die Ableitung von f nicht nur im kritischen Punkt x_0 , sondern auch in dessen Umgebung zu untersuchen ist. Für eine hinreichende Bedingung, die nur den kritischen Punkt selbst berücksichtigt, benötigen wir *höhere Ableitungen* von f , welche wir im nächsten Abschnitt 7.4 thematisieren.

Wir beschließen diesen Abschnitt mit einer Anwendung des (allgemeinen) Mittelwertsatzes, die oft sehr hilfreich bei der Berechnung von Grenzwerten ist.

Satz 7.3.15: L'HOSPITALSche Regel

Es seien $f, g : (a, b) \rightarrow \mathbb{R}$ zwei differenzierbare Funktionen. Es gelte $g' \neq 0$ auf (a, b) , und es existiere der Limes

$$\lim_{x \rightarrow a+} \frac{f'(x)}{g'(x)} =: c \in \mathbb{R}.$$

Dann gelten folgende Aussagen:

- (a) Falls $\lim_{x \rightarrow a+} f(x) = \lim_{x \rightarrow a+} g(x) = 0$ gilt, so folgt $g' \neq 0$ auf (a, b) und

$$\lim_{x \rightarrow a+} \frac{f(x)}{g(x)} = c.$$

(b) Falls $\lim_{x \rightarrow a+} f(x) = \pm\infty$, $\lim_{x \rightarrow a+} g(x) = \pm\infty$ gilt, so existiert ein $x_0 \in (a, b)$ mit $g \neq 0$ auf $(a, x_0]$ und es folgt

$$\lim_{x \rightarrow a+} \frac{f(x)}{g(x)} = c.$$

Analoge Aussagen sind für die Grenzwerte $x \rightarrow b-$ und $x \rightarrow \xi$ mit beliebigem $\xi \in (a, b)$ richtig.

Beweis. Wir zeigen nur (a) und belassen (b) für die *Ergänzungen*.

Nach Voraussetzung können wir zunächst f und g stetig (zu 0) in den Punkt $x = a$ fortsetzen. Der Satz von Rolle, Satz 7.3.7, liefert dann $g \neq 0$ auf (a, b) . Nach dem allgemeinen Mittelwertsatz, Satz 7.3.10, gibt es daher zu jedem hinreichend kleinen $h > 0$ ein $\vartheta = \vartheta(h) \in (0, 1)$ mit der Eigenschaft

$$\frac{f(a+h)}{g(a+h)} = \frac{f(a+h) - f(a)}{g(a+h) - g(a)} = \frac{f'(a+\vartheta h)}{g'(a+\vartheta h)}.$$

Für $h \rightarrow 0+$ (und somit $a + \vartheta h \rightarrow a+$) erhalten wir die Existenz des Grenzwertes $\lim_{x \rightarrow a+} \frac{f(x)}{g(x)}$ und die Relation

$$\lim_{x \rightarrow a+} \frac{f(x)}{g(x)} = \lim_{h \rightarrow 0+} \frac{f(a+h)}{g(a+h)} = \lim_{h \rightarrow 0+} \frac{f'(a+\vartheta h)}{g'(a+\vartheta h)} = c,$$

wie behauptet. □

Bemerkung 7.3.16. Satz 7.3.15 lässt sich durch geeignete Anpassungen noch erweitern: Einerseits gelten die Aussagen auch für uneigentliche Grenzwerte $c = \pm\infty$, andererseits auch für $a = -\infty$ bzw. $b = +\infty$.

7.4. Höhere Ableitungen und lokale Extrema

Zu einer beliebigen Funktion $f : I \rightarrow \mathbb{R}^d$ wollen wir nun induktiv „Ableitungen von Ableitungen“ erklären:

- $k = 1$: Ist f in $x_0 \in I$ differenzierbar, so setzen wir $f^{(1)}(x_0) := f'(x_0)$.
- $k \rightarrow k + 1$: Ist $f^{(k)}(x)$ für alle $x \in I$ erklärt und ist die Funktion $f^{(k)} : I \rightarrow \mathbb{R}^d$ in $x_0 \in I$ differenzierbar, so setzen wir $f^{(k+1)}(x_0) := (f^{(k)})'(x_0)$.

Wenn existent, ist so für $k \in \mathbb{N}$ und $x_0 \in I$ die Größe $f^{(k)}(x_0)$ definiert.

Definition 7.4.1

Ist $f : I \rightarrow \mathbb{R}^d$ gegeben und existiert für ein $k \in \mathbb{N}$ und ein $x_0 \in I$ die *k-te Ableitung* oder *Ableitung k-ter Ordnung* $f^{(k)}(x_0)$, so heißt f in x_0 *k-mal differenzierbar*. Alternative

Schreibweisen für die k -te Ableitung sind

$$f^{(k)}(x_0) = \frac{d^k f}{dx^k}(x_0) = D^k f(x_0).$$

Ist f in allen Punkten aus I k -mal differenzierbar, so sagen wir, f ist *k -mal differenzierbar (auf I)*; die k -te Ableitung fassen wir dann als Funktion $f^{(k)} : I \rightarrow \mathbb{R}^d$ auf.

Bemerkung 7.4.2. ▶ Nach unserer obigen induktiven Definition setzt die Existenz der k -ten Ableitung (in einem beliebigen Punkt) immer voraus, dass alle Ableitungen $f^{(1)}, \dots, f^{(k-1)}$ auf I existieren. Wendet man die Definition auf die Einschränkung $f|_{I \cap B_r(x_0)}$ an, so sieht man, dass es genügt, die Existenz der Ableitungen bis zur $(k-1)$ -ten Ordnung in einer Umgebung von x_0 vorauszusetzen.

- ▶ Für die Funktion $f : I \rightarrow \mathbb{R}^d$ schreibt man formal auch $f^{(0)} := f$; jede beliebige solche Funktion können wir dann auch als *0-mal differenzierbar (auf I)* bezeichnen und f selbst ist ihre *0-te Ableitung*. Übrigens schreibt man für die ersten drei Ableitungen auch

$$f' = f^{(1)}, \quad f'' = f^{(2)} = (f')', \quad f''' = f^{(3)} = (f'')'.$$

Eine kleine, aber in der Analysis oft entscheidende Verschärfung der k -maligen Differenzierbarkeit führt zu den folgenden Begriffen:

Definition 7.4.3: C^k -Räume

Für beliebiges $k \in \mathbb{N}_0$ erklären wir den Vektorraum $C^k(I, \mathbb{R}^d)$ der *k -mal stetig differenzierbaren Funktionen* $f : I \rightarrow \mathbb{R}^d$, die k -mal differenzierbar auf I sind und für die $f^{(k)} : I \rightarrow \mathbb{R}^d$ stetig ist. Weiter erklären wir

$$C^\infty(I, \mathbb{R}^d) := \bigcap_{k \in \mathbb{N}_0} C^k(I, \mathbb{R}^d),$$

den Vektorraum der *unendlich oft differenzierbaren Funktionen*.

Schließlich schreiben wir mit $k \in \mathbb{N}_0 \cup \{\infty\}$ auch $C^k(I)$ bzw. $C^k(I, \mathbb{C})$ für die reell- bzw. komplexwertigen k -mal stetig differenzierbaren Funktionen auf I .

Bemerkung 7.4.4. (a) Nach Folgerung 7.1.5 sind alle Ableitungen $f, f', \dots, f^{(k)}$ einer Funktion $f \in C^k(I, \mathbb{R}^d)$ stetig auf I . Insbesondere folgt

$$C^k(I, \mathbb{R}^d) \subset C^l(I, \mathbb{R}^d) \quad \text{für } l \leq k.$$

- (b) Die in Definition 7.4.3 enthaltene Behauptung, $C^k(I, \mathbb{R}^d)$ seien für alle $k \in \mathbb{N}_0 \cup \{\infty\}$ Vektorräume, ergibt sich für $k = 0, 1$ aus Satz 6.2.4 (a) bzw. Satz 7.2.1 (a) und folgt dann

induktiv für alle $k \in \mathbb{N}_0$ und somit auch für den Durchschnitt $C^\infty(I, \mathbb{R}^d)$. Auch die Produktregel aus Satz 7.2.1 (b) bzw. (7.2.1) lässt sich verallgemeinern; wir formulieren diese wieder für komplexwertige Funktionen:

Folgerung 7.4.5: LEIBNIZ-Regel

Gilt $f, g \in C^k(I, \mathbb{C})$ für ein $k \in \mathbb{N}$, so folgt $fg \in C^k(I, \mathbb{C})$ und es gilt

$$(fg)^{(k)} = \sum_{j=0}^k \binom{k}{j} f^{(j)} g^{(k-j)} \quad \text{auf } I.$$

Beweis. Analog zum Beweis des BINOMISCHEN Lehrsatzes mittels vollständiger Induktion (\rightarrow Ergänzungen). \square

Schließlich verallgemeinern wir noch die Sätze 7.2.4 und 7.2.5:

Folgerung 7.4.6

Sei $f \in C^k(I)$ für beliebiges $k \in \mathbb{N}$ gegeben. Dann gelten:

- (a) Für beliebiges $g \in C^k(J, \mathbb{R}^d)$ mit $J \supset f(I)$ folgt $g \circ f \in C^k(I, \mathbb{R}^d)$.
- (b) Ist f injektiv mit $f' \neq 0$ auf I , so gilt auch $f^{-1} \in C^k(I^*)$ mit $I^* := f(I)$.

Beweis. Durch vollständige Induktion über k :

- (a) Der Fall $k = 1$ ist in Satz 7.2.4 enthalten. Die Aussage gelte für ein $k \in \mathbb{N}$ und es seien $f \in C^{k+1}(I)$, $g \in C^{k+1}(J, \mathbb{R}^d)$. Die Kettenregel liefert dann

$$(g \circ f)' = (g' \circ f)f' \quad \text{auf } I.$$

Nach Voraussetzung gelten $g' \in C^k(J, \mathbb{R}^d)$, $f, f' \in C^k(I)$ - beachte Bemerkung 7.4.4 (a) -, und somit nach Induktionsvoraussetzung $g' \circ f \in C^k(I, \mathbb{R}^d)$ und wegen Folgerung 7.4.5 dann auch $(g \circ f)' \in C^k(I, \mathbb{R}^d)$ bzw. $g \circ f \in C^{k+1}(I, \mathbb{R}^d)$, wie behauptet.

- (b) Der Fall $k = 1$ ist in Satz 7.2.5 enthalten. Gilt die Aussage für ein $k \in \mathbb{N}$ und ist $f \in C^{k+1}(I)$, so entnehmen wir Satz 7.2.5 weiter

$$(f^{-1})' = \frac{1}{f' \circ f^{-1}} = h \circ f' \circ f^{-1} \quad \text{auf } I^*,$$

wobei wir noch $h(t) := \frac{1}{t}$ erklärt haben. Da $f' > 0$ oder $f' < 0$ auf I gilt (nach dem Zwischenwertsatz) und h sowohl auf dem Intervall $(0, +\infty)$, als auch auf $(-\infty, 0)$

unendlich oft differenzierbar ist (folgt induktiv aus Beispiel 7.2.3), liefern (a) und die Induktionsvoraussetzung $(f^{-1})' \in C^k(I^*)$ bzw. $f^{-1} \in C^{k+1}(I^*)$, wie behauptet. \square

Bemerkung 7.4.7. Offenbar bleiben die Aussagen der Folgerungen 7.4.5, 7.4.6 auch für $k = \infty$ richtig.

Wir wenden uns nun wieder der Suche nach lokalen Extrema zu und wollen Kriterien 2. Ordnung angeben.

Satz 7.4.8: Hinreichende Extremalbedingung

Sei $f \in C^1(I)$ gegeben, und in $x_0 \in \text{int } I$ sei f zweimal differenzierbar mit

$$f'(x_0) = 0 \quad \text{und} \quad f''(x_0) > 0 \quad (\text{bzw. } f''(x_0) < 0).$$

Dann besitzt f in x_0 ein striktes lokales Minimum (bzw. Maximum).

Beweis. O.B.d.A. gelte $f''(x_0) = \lim_{x \rightarrow x_0} \frac{f'(x) - f'(x_0)}{x - x_0} > 0$; der Fall $f''(x_0) < 0$ ergibt sich nach Übergang zu $-f$. Dann existiert ein $\varepsilon > 0$, so dass $(x_0 - \varepsilon, x_0 + \varepsilon) \subset I$ und

$$\frac{f'(x) - f'(x_0)}{x - x_0} > 0 \quad \text{für alle } x \in (x_0 - \varepsilon, x_0 + \varepsilon) \setminus \{x_0\}$$

erfüllt sind. Wegen $f'(x_0) = 0$ bedeutet dies

$$f'(x) < 0 \quad \text{für alle } x \in (x_0 - \varepsilon, x_0),$$

$$f'(x) > 0 \quad \text{für alle } x \in (x_0, x_0 + \varepsilon).$$

Nach Folgerung 7.3.13 (a) hat f in x_0 daher ein striktes lokales Minimum, wie behauptet. \square

Folgerung 7.4.9: Notwendige Extremalbedingung 2. Ordnung

Sei $f \in C^1(I)$ auf dem Intervall $I \subset \mathbb{R}$ gegeben. Sei $x_0 \in \text{int } I$ eine lokale Minimalstelle (bzw. Maximalstelle) von f , und sei f in x_0 zweimal differenzierbar. Dann gilt $f''(x_0) \geq 0$ (bzw. $f''(x_0) \leq 0$).

Beweis. Ist x_0 lokale Minimalstelle und gälte $f''(x_0) < 0$, so wäre x_0 nach Satz 7.4.8 auch strikte lokale Maximalstelle, Widerspruch! Also muss doch $f''(x_0) \geq 0$ gelten. Entsprechend folgt die Aussage für Maximalstellen. \square

Bemerkung 7.4.10. Das in Satz 7.4.8 angegebene hinreichende Kriterium ist nicht notwendig für ein Extremum, wie das Beispiel $f(x) = x^4$, $x \in \mathbb{R}$, mit dem strikten Minimum $x_0 = 0$ zeigt. Andererseits ist das notwendige Kriterium aus Folgerung 7.4.9 nicht hinreichend; hier betrachte man $g(x) = x^3$, $x \in \mathbb{R}$, im Punkt $x_0 = 0$.

Wir beschließen den Abschnitt mit einer Anwendung von Satz 7.4.8 auf die Untersuchung des Krümmungsverhaltens von Funktionen.

Definition 7.4.11

Eine Funktion $f : I \rightarrow \mathbb{R}$ heißt *konvex* auf dem Intervall I , wenn für alle $x_1, x_2 \in I$ und alle $\lambda \in (0, 1)$ gilt

$$f(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f(x_1) + (1 - \lambda)f(x_2). \quad (7.4.1)$$

Die Funktion f heißt *konkav*, wenn $-f$ konvex ist. Gilt schließlich in (7.4.1) die strikte Ungleichung für $x_1 \neq x_2$, so heißt f *strikt konvex*; gilt dies für $-f$, so nennen wir f *strikt konkav*.

Folgerung 7.4.12: Krümmungsverhalten

Sei $f \in C^2(I)$ gegeben.

- (a) f ist genau dann konvex (bzw. konkav), wenn $f'' \geq 0$ (bzw. $f'' \leq 0$) auf I gilt.
- (b) Gilt $f'' > 0$ (bzw. $f'' < 0$) auf I , so ist f strikt konvex (bzw. strikt konkav).

Bemerkung 7.4.13. Das Kriterium in (b) ist nicht notwendig, wie wieder das Beispiel $f(x) = x^4$, $x \in \mathbb{R}$, mit $f''(0) = 0$ zeigt.

Beweis von Satz 7.4.12. Teil (b) und die Richtung „ \Leftarrow “ von Teil (a) lassen sich auf gleiche Weise mit dem Mittelwertsatz beweisen (\rightarrow Ergänzungen). Wir zeigen hier die Richtung „ \Rightarrow “ von (a) für den konvexen Fall.

Sei also f konvex. Wir nehmen an, dass *nicht* $f'' \geq 0$ auf I gilt. Dann existierte aus Stetigkeitsgründen ein $x_0 \in \text{int } I$ mit $f''(x_0) < 0$. Wir erklären die Hilfsfunktion

$$\varphi(x) := f(x) - f'(x_0)(x - x_0), \quad x \in I,$$

und bemerken $\varphi \in C^2(I)$ und $\varphi'(x_0) = 0$, $\varphi''(x_0) = f''(x_0) < 0$. Nach Satz 7.4.8 besitzt also φ in x_0 ein striktes lokales Maximum. Insbesondere finden wir ein $h > 0$, so dass $[x_0 - h, x_0 + h] \subset I$ sowie

$$\varphi(x_0 - h) < \varphi(x_0), \quad \varphi(x_0 + h) < \varphi(x_0)$$

erfüllt sind. Hieraus erhalten wir

$$f(x_0) = \varphi(x_0) > \frac{1}{2}(\varphi(x_0 - h) + \varphi(x_0 + h)) = \frac{1}{2}(f(x_0 - h) + f(x_0 + h)). \quad (7.4.2)$$

Setzen wir schließlich $x_1 := x_0 - h$, $x_2 := x_0 + h$ und $\lambda = \frac{1}{2}$, so haben wir $x_0 = \lambda x_1 + (1 - \lambda)x_2$ und (7.4.2) besagt

$$f(\lambda x_1 + (1 - \lambda)x_2) > \lambda f(x_1) + (1 - \lambda)f(x_2),$$

im Widerspruch zur vorausgesetzten Konvexität (7.4.1) von f . Also gilt doch $f'' \geq 0$ auf I . \square

Kapitel 8.

Die elementaren Funktionen

In diesem Kapitel untersuchen wir die Exponentialfunktion genauer und leiten aus ihr die *elementaren Funktionen* ab. Abschnitt 8.1 ist zunächst allgemein der Frage gewidmet, wann der Grenzwert von Folgen differenzierbarer Funktionen wieder differenzierbar ist. In Abschnitt 8.2 wenden wir dies auf die Einschränkung von \exp auf die *reelle* Achse an und erklären damit die Logarithmusfunktion und die allgemeine Potenz. Abschnitt 8.3 ist den Kreisfunktionen \sin und \cos gewidmet, die sich durch Einschränkung der komplexen Exponentialfunktion auf die *imaginäre* Achse ergeben. In Abschnitt 8.4 führen wir mit Hilfe von \cos und \sin Polarkoordinaten ein. Der abschließende Abschnitt 8.5 gibt einen (unvollständigen) Überblick über weitere elementare Funktionen.

8.1. Funktionenfolgen und Differenzierbarkeit

In Abschnitt 6.6 haben wir mit dem WEIERSTRASSschen Konvergenzsatz, Satz 6.6.5, den gleichmäßigen Limes stetiger Funktionen als stetig erkannt. Insbesondere folgte daraus die Stetigkeit der komplexen Exponentialfunktion $\exp : \mathbb{C} \rightarrow \mathbb{C}$, die grundlegend für die Untersuchungen des aktuellen Kapitels ist.

Wir wollen zunächst angeben, unter welchen Voraussetzungen der Limes einer Folge differenzierbarer Funktionen selbst wieder differenzierbar ist, und dies auf Potenzreihen (und damit insbesondere die Exponentialfunktion) anwenden.

Satz 8.1.1

Sei $I = [a, b]$ und $\{f_n\}_n$ eine Folge von Funktionen $f_n \in C^1(I, \mathbb{R}^d)$ für alle $n \in \mathbb{N}$. Falls dann gilt

$$f_n \rightarrow f \ (n \rightarrow \infty), \quad f'_n \rightrightarrows g \ (n \rightarrow \infty) \quad \text{auf } I,$$

so folgt für den punktweisen Limes $f \in C^1(I, \mathbb{R}^d)$, und es gilt

$$\frac{d}{dx} \left(\lim_{n \rightarrow \infty} f_n(x) \right) = f' = g = \lim_{n \rightarrow \infty} \left(\frac{d}{dx} f_n(x) \right) \quad \text{auf } I.$$

Beweis. Wir werden den Beweis mit Hilfe der Integration in Kapitel 9 nachholen. \square

Wir wenden Satz 8.1.1 auf Potenzreihen an, wobei wir uns aber auf die reelle Achse einschränken:

Satz 8.1.2

Es sei $\mathcal{P}(x) := \sum_{k=0}^{\infty} a_k x^k$, $a_k \in \mathbb{C}$, eine Potenzreihe mit Konvergenzradius $R \in (0, +\infty) \cup \{+\infty\}$. Dann gehört $\mathcal{P} : (-R, R) \rightarrow \mathbb{C}$ zur Klasse $C^1((-R, R), \mathbb{C})$ und es gilt

$$\mathcal{P}'(x) = \sum_{k=1}^{\infty} k a_k x^{k-1}, \quad x \in (-R, R). \quad (8.1.1)$$

Bemerkung 8.1.3. Die formal durch gliedweises Differenzieren der Reihe erhaltene Potenzreihe hat also den gleichen Konvergenzradius (siehe Beweis) und stimmt mit der tatsächlichen Ableitung der Reihe überein.

Für den Beweis von Satz 8.1.2 benötigen wir noch den folgenden kleinen

Hilfssatz 8.1.4

Es gelten $\lim_{k \rightarrow \infty} \sqrt[k]{k} = 1$ und $\lim_{k \rightarrow \infty} \sqrt[k]{a} = 1$ für jedes $a \in (0, +\infty)$.

Beweis. ▶ Mit $x_k := \sqrt[k]{k} - 1 \geq 0$, $k \in \mathbb{N}$, liefert der BINOMISCHE Satz

$$k = (x_k + 1)^k = \sum_{l=0}^k \binom{k}{l} x_k^l \geq 1 + \frac{k(k-1)}{2} x_k^2, \quad k \in \mathbb{N},$$

und somit

$$0 \leq x_k^2 \leq \frac{2}{k} \quad \text{für alle } k \in \mathbb{N}.$$

Wir folgern $x_k \rightarrow 0$ bzw. $\sqrt[k]{k} \rightarrow 1$ für $k \rightarrow \infty$.

▶ Für $a \geq 1$ und $y_k := \sqrt[k]{a} - 1 \geq 0$ liefert die BERNOULLISCHE Ungleichung

$$a = (1 + y_k)^k \geq 1 + k y_k \quad \text{bzw.} \quad 0 \leq y_k \leq \frac{a-1}{k}, \quad k \in \mathbb{N}.$$

Wir folgern $y_k \rightarrow 0$ bzw. $\sqrt[k]{a} \rightarrow 1$ für $k \rightarrow \infty$. Falls $0 < a < 1$ gilt, wenden wir diese Argumentation auf $a^{-1} > 1$ an und finden

$$\sqrt[k]{a} = \frac{1}{\sqrt[k]{a^{-1}}} \rightarrow \frac{1}{1} = 1 \quad (k \rightarrow \infty),$$

also wieder die Behauptung. \square

Beweis von Satz 8.1.2. 1. Wir zeigen zunächst, dass die formal differenzierte Reihe

$$g(x) := \sum_{k=1}^{\infty} k a_k x^{k-1}$$

für jedes $R_0 \in (0, R)$ gleichmäßig auf $[-R_0, R_0]$ konvergiert.

Denn: Offenbar wird g auf $[-R_0, R_0]$ durch die Reihe $\sum_k k |a_k| R_0^{k-1}$ majorisiert. Letztere ist nach dem Wurzelkriterium, Satz 4.2.13, und Hilfssatz 8.1.4 konvergent:

$$\limsup_{k \rightarrow \infty} \sqrt[k]{k |a_k| R_0^{k-1}} = R_0 \cdot \limsup_{k \rightarrow \infty} \left(\frac{\sqrt[k]{k}}{\sqrt[k]{R_0}} \sqrt[k]{|a_k|} \right) = R_0 \cdot \limsup_{k \rightarrow \infty} \sqrt[k]{|a_k|} = \frac{R_0}{R} < 1.$$

Satz 6.6.9 liefert also die gleichmäßige Konvergenz von g auf $[-R_0, R_0]$.

2. Wir schreiben

$$\mathcal{P}_n(x) := \sum_{k=0}^n a_k x^k, \quad g_n(x) := \sum_{k=1}^n k a_k x^{k-1}, \quad n \in \mathbb{N},$$

für die Partialsummen von \mathcal{P} und g . Dann haben wir nach Voraussetzung und Teil 1:

$$\mathcal{P}_n \rightarrow \mathcal{P}, \quad g_n \rightrightarrows g \quad (n \rightarrow \infty) \quad \text{auf } [-R_0, R_0].$$

Wegen $\mathcal{P}'_n = g_n$ auf $[-R_0, R_0]$ für alle $n \in \mathbb{N}$ liefert somit Satz 8.1.1 $\mathcal{P} \in C^1([-R_0, R_0], \mathbb{C})$ sowie

$$\mathcal{P}'(x) = g(x) = \sum_{k=1}^{\infty} k a_k x^{k-1}, \quad x \in [-R_0, R_0].$$

Da schließlich $R_0 \in (0, R)$ beliebig war, folgt die Behauptung. \square

Folgerung 8.1.5

Es sei $\mathcal{P}(x) := \sum_{k=0}^{\infty} a_k x^k$, $a_k \in \mathbb{C}$, eine Potenzreihe mit Konvergenzradius $R \in (0, +\infty) \cup \{+\infty\}$. Dann gilt $\mathcal{P} \in C^\infty((-R, R), \mathbb{C})$ und für die n -te Ableitung gilt

$$\mathcal{P}^{(n)}(x) = \sum_{k=n}^{\infty} k(k-1)(k-2) \dots (k-n+1) a_k x^{k-n} \quad \text{auf } (-R, R), \quad n \in \mathbb{N}. \quad (8.1.2)$$

Beweis. Vollständige Induktion über $n \in \mathbb{N}$ (\rightarrow Ergänzungen). \square

8.2. Exponentialfunktion, Logarithmus und allgemeine Potenz

Bereits in Abschnitt 6.6, genauer in Folgerung 6.6.11, haben wir die komplexe Exponentialfunktion definiert,

$$\exp z = e^z := \sum_{k=0}^{\infty} \frac{z^k}{k!}, \quad z \in \mathbb{C}, \quad (8.2.1)$$

und $\exp \in C^0(\mathbb{C}, \mathbb{C})$ gefolgt. Bevor wir die Resultate von Abschnitt 8.1 auf $\exp|_{\mathbb{R}}$ anwenden, geben wir noch zwei Resultate für die komplexe Exponentialfunktion an.

Folgerung 8.2.1

Für beliebige $z \in \mathbb{C}$ gilt die Darstellung

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

Beweis. → Ergänzungen. □

Satz 8.2.2: Funktionalgleichung der Exponentialfunktion

Für beliebige $z_1, z_2 \in \mathbb{C}$ gilt die Identität

$$\exp(z_1 + z_2) = \exp z_1 \cdot \exp z_2.$$

Beweis. Da die Exponentialreihe für beliebige $z \in \mathbb{C}$ absolut konvergiert, liefern die CAUCHYSche Produktformel, Satz 4.3.6, und der BINOMISCHE Satz 2.4.12:

$$\begin{aligned} \exp z_1 \cdot \exp z_2 &= \left(\sum_{k=0}^{\infty} \frac{z_1^k}{k!} \right) \left(\sum_{k=0}^{\infty} \frac{z_2^k}{k!} \right) = \sum_{k=0}^{\infty} \left(\sum_{l=0}^k \frac{z_1^l}{l!} \frac{z_2^{k-l}}{(k-l)!} \right) \\ &= \sum_{k=0}^{\infty} \frac{1}{k!} \left(\sum_{l=0}^k \binom{k}{l} z_1^l z_2^{k-l} \right) = \sum_{k=0}^{\infty} \frac{(z_1 + z_2)^k}{k!} = \exp(z_1 + z_2), \end{aligned}$$

wie behauptet. □

Definition 8.2.3

Die Zahl

$$e := \exp 1 = \sum_{k=0}^{\infty} \frac{1}{k!} \in \mathbb{R}$$

wird *EULERSche Zahl* genannt.

Bemerkung 8.2.4. ▶ In den *Ergänzungen* zeigen wir, dass e eine irrationale Zahl ist.

▶ Mit der Funktionalgleichung zeigt man leicht

$$e^{\frac{p}{q}} = \exp\left(\frac{p}{q}\right) \quad \text{für alle } p \in \mathbb{Z}, q \in \mathbb{N}. \quad (8.2.2)$$

Dies erklärt sowohl die gebräuchliche Bezeichnung *e-Funktion* als auch die Schreibweise der Exponentialfunktion als Potenz:

$$e^z = \exp z, \quad z \in \mathbb{C}.$$

Wir konzentrieren uns nun auf die Einschränkung

$$e^x = \exp(x) = \exp|_{\mathbb{R}}(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!}, \quad x \in \mathbb{R},$$

der komplexen *e*-Funktion auf die *reelle* Achse; in Abschnitt 8.3 wenden wir uns dann der Einschränkung auf die imaginäre Achse zu.

Satz 8.2.5

Die *reelle Exponentialfunktion* $\exp : \mathbb{R} \rightarrow \mathbb{R}$ gehört zur Klasse $C^\infty(\mathbb{R})$ und es gilt

$$\exp' x = \frac{d}{dx} \exp x = \exp x, \quad x \in \mathbb{R}. \quad (8.2.3)$$

Beweis. Gemäß Beispiel 4.2.12 (c) ist die Exponentialreihe für alle $x \in \mathbb{R}$ konvergent. Nach Folgerung 8.1.5 gilt also $\exp \in C^\infty(\mathbb{R})$ und wir haben

$$\exp' x \stackrel{(8.1.1)}{=} \sum_{k=1}^{\infty} k \frac{1}{k!} x^{k-1} = \sum_{k=1}^{\infty} \frac{1}{(k-1)!} x^{k-1} = \sum_{k=0}^{\infty} \frac{x^k}{k!} = \exp x,$$

wie behauptet. □

Der nächste Satz bestätigt die Vorstellung, die wir aus der Schule von der reellen *e*-Funktion mitbringen.

Satz 8.2.6

Die reelle Exponentialfunktion \exp bildet \mathbb{R} bijektiv auf $(0, +\infty)$ ab, ist streng monoton wachsend, strikt konvex und erfüllt

$$\lim_{x \rightarrow -\infty} \exp x = 0, \quad \exp 0 = 1, \quad \exp 1 = e, \quad \lim_{x \rightarrow +\infty} \exp x = +\infty. \quad (8.2.4)$$

Beweis. Per Definition ist $e^0 = 1$ und $e^1 = e$ klar. Ferner haben wir

$$\exp x = 1 + \sum_{k=1}^{\infty} \frac{x^k}{k!} > 0 \quad \text{für alle } x \in [0, +\infty)$$

und nach Satz 8.2.2 auch

$$\exp x = \frac{1}{\exp(-x)} > 0 \quad \text{für alle } x \in (-\infty, 0),$$

also insgesamt $\exp(\mathbb{R}) \subset (0, +\infty)$. Zum Beweis von (8.2.4) beachten wir

$$\lim_{x \rightarrow +\infty} \exp x \geq \lim_{x \rightarrow +\infty} (1 + x) = +\infty$$

und

$$\lim_{x \rightarrow -\infty} \exp x = \lim_{x \rightarrow -\infty} \frac{1}{\exp(-x)} \stackrel{\xi := -x}{=} \lim_{\xi \rightarrow +\infty} \frac{1}{\exp \xi} = 0.$$

Ist nun $y \in (0, +\infty)$ beliebig, so existieren also $x_1 < x_2$ mit $e^{x_1} < y < e^{x_2}$. Nach dem Zwischenwertsatz, Satz 6.3.4, gibt es ein $x \in (x_1, x_2)$ mit $e^x = y$, d.h. $y \in \exp(\mathbb{R})$ und insgesamt $\exp(\mathbb{R}) = (0, +\infty)$.

Schließlich gilt nach Satz 8.2.5: $\exp' x = \exp x > 0$ für alle $x \in \mathbb{R}$, also ist \exp nach Folgerung 7.3.11 (a) streng monoton wachsend. Und wegen $\exp'' x = \exp' x = \exp x > 0$ liefert Satz 7.4.12 (b) die strikte Konvexität von \exp . \square

Definition 8.2.7

Die Umkehrfunktion von $\exp : \mathbb{R} \rightarrow (0, +\infty)$ nennen wir (*natürliche*) *Logarithmusfunktion* $\log : (0, +\infty) \rightarrow \mathbb{R}$. Für $x > 0$ heißt $y = \log x$ (*natürlicher*) *Logarithmus* von x .

Satz 8.2.8

Die Funktion $\log : (0, +\infty) \rightarrow \mathbb{R}$ ist streng monoton wachsend und strikt konkav, gehört zu $C^\infty((0, +\infty))$ und es gilt

$$\log' x = \frac{d}{dx} \log x = \frac{1}{x} \quad \text{für alle } x > 0. \quad (8.2.5)$$

Ferner ist die *Funktionalgleichung* der Logarithmusfunktion

$$\log(x_1 x_2) = \log x_1 + \log x_2 \quad \text{für alle } x_1, x_2 > 0 \quad (8.2.6)$$

erfüllt, und wir haben

$$\lim_{x \rightarrow 0+} \log x = -\infty, \quad \log 1 = 0, \quad \log e = 1, \quad \lim_{x \rightarrow +\infty} \log x = +\infty. \quad (8.2.7)$$

Beweis. Zunächst gehört \log nach Folgerung 7.4.6 als Umkehrfunktion von \exp zur Klasse $C^\infty((0, +\infty))$, und gemäß Satz 7.2.5 gilt

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

Außerdem ist \log offenbar streng monoton wachsend, und (8.2.5) liefert $\log'' x = -\frac{1}{x^2} < 0$ für $x > 0$, nach Satz 7.4.12 (b) ist \log also strikt konkav.

Zum Beweis von (8.2.6) seien $x_1, x_2 > 0$ beliebig gewählt. Wir erhalten dann aus Satz 8.2.2

$$\exp(\log x_1 + \log x_2) = \exp(\log x_1) \cdot \exp(\log x_2) = x_1 x_2.$$

Nehmen wir auf beiden Seiten den Logarithmus, so folgt die Behauptung (8.2.6).

Schließlich ist $\log 1 = \log(e^0) = 0$ und $\log e = \log(e^1) = 1$ richtig. Die Grenzwerte in (8.2.7) ergeben sich direkt aus der Monotonie und der Relation $\log((0, +\infty)) = \mathbb{R}$. Damit ist alles gezeigt. \square

Definition 8.2.9

Für beliebiges $\alpha \in \mathbb{R}$ erklären wir die (*allgemeine*) *Potenz* oder *Potenzfunktion* $x \mapsto x^\alpha$, $x \in (0, +\infty)$, durch die Formel

$$x^\alpha := e^{\alpha \log x} = \exp(\alpha \log x).$$

Bemerkung 8.2.10. Mit der Funktionalgleichung von \log rechnet man leicht nach, dass x^α für $\alpha \in \mathbb{Q}$ mit der ursprünglichen Definition rationaler Potenzen übereinstimmt; siehe Definition 3.4.13.

Satz 8.2.11

Die allgemeine Potenzfunktion $f(x) := x^\alpha$, $x \in (0, +\infty)$, erfüllt $f \in C^\infty((0, +\infty))$ und

$$\frac{d}{dx}(x^\alpha) = \alpha x^{\alpha-1}, \quad x > 0. \quad (8.2.8)$$

Weiter gelten für beliebige $x, y > 0$ und $\alpha, \beta \in \mathbb{R}$ die Relationen

$$x^\alpha y^\alpha = (xy)^\alpha, \quad x^\alpha x^\beta = x^{\alpha+\beta}, \quad (x^\alpha)^\beta = x^{\alpha\beta}, \quad (8.2.9)$$

$$\log(x^\alpha) = \alpha \log x. \quad (8.2.10)$$

Beweis. \rightarrow Ergänzungen. \square

Bemerkung 8.2.12. ▶ Für festes $c > 0$ können wir auch die *allgemeine Exponentialfunktion*

$$f(x) := c^x = e^{x \log c}, \quad x \in \mathbb{R},$$

betrachten. Es gilt $f \in C^\infty(\mathbb{R})$ und

$$f'(x) = c^x \cdot \log c, \quad x \in \mathbb{R}.$$

Für $c > 1$ ist also $f' > 0$ und $f : \mathbb{R} \rightarrow (0, +\infty)$ bijektiv. Die zugehörige Umkehrfunktion heißt *Logarithmus zur Basis $c > 1$* und wird mit $\log_c : (0, +\infty) \rightarrow \mathbb{R}$ bezeichnet. Der Logarithmus zur Basis $e > 1$ ist der natürliche Logarithmus (\rightarrow *Übungen*).

- ▶ Ganz allgemein kann man mit Funktionen $f : I \rightarrow (0, +\infty)$ und $g : I \rightarrow \mathbb{R}$ auf einem Intervall $I \subset \mathbb{R}$ die Funktion

$$h(x) := f(x)^{g(x)} = \exp[g(x) \log f(x)], \quad x \in I,$$

erklären. Sind $f, g \in C^k(I)$, $k \in \mathbb{N}_0 \cup \{\infty\}$, so folgt auch $h \in C^k(I)$.

8.3. Die Kreisfunktionen \cos und \sin

Wir schränken nun die komplexe Exponentialfunktion auf die *imaginäre* Achse ein, betrachten also die komplexwertige Funktion

$$\exp(ix) = e^{ix} = \sum_{k=0}^{\infty} \frac{(ix)^k}{k!}, \quad x \in \mathbb{R}.$$

Wegen

$$e^{-ix} = \sum_{k=0}^{\infty} \frac{(-ix)^k}{k!} = \sum_{k=0}^{\infty} \frac{(ix)^k}{k!} \stackrel{\text{Satz 4.1.3 (b)}}{=} \sum_{k=0}^{\infty} \frac{(ix)^k}{k!} = \overline{e^{ix}}, \quad x \in \mathbb{R},$$

erhalten wir mit (2.5.3) für den Real- bzw. Imaginärteil von e^{ix} :

Definition 8.3.1

Wir erklären die *Cosinusfunktion* $\cos : \mathbb{R} \rightarrow \mathbb{R}$ und die *Sinusfunktion* $\sin : \mathbb{R} \rightarrow \mathbb{R}$ gemäß

$$\begin{aligned} \cos x &:= \operatorname{Re}(e^{ix}) = \frac{1}{2}(e^{ix} + e^{-ix}), \\ \sin x &:= \operatorname{Im}(e^{ix}) = \frac{1}{2i}(e^{ix} - e^{-ix}), \quad x \in \mathbb{R}. \end{aligned}$$

Satz 8.3.2

Die Funktionen cos und sin gehören zur Klasse $C^\infty(\mathbb{R})$ mit den Ableitungen

$$\begin{aligned}\cos' x &= \frac{d}{dx} \cos x = -\sin x, \\ \sin' x &= \frac{d}{dx} \sin x = \cos x, \quad x \in \mathbb{R}.\end{aligned}\tag{8.3.1}$$

Es gelten die *EULERSche Formel*

$$e^{ix} = \cos x + i \sin x, \quad x \in \mathbb{R},\tag{8.3.2}$$

und die *Additionstheoreme*

$$\begin{aligned}\cos(x_1 + x_2) &= \cos x_1 \cos x_2 - \sin x_1 \sin x_2, \\ \sin(x_1 + x_2) &= \cos x_1 \sin x_2 + \sin x_1 \cos x_2, \quad x_1, x_2 \in \mathbb{R}.\end{aligned}\tag{8.3.3}$$

Die Cosinusfunktion ist *gerade*, die Sinusfunktion ist *ungerade*, d.h.

$$\cos(-x) = \cos x, \quad \sin(-x) = -\sin x, \quad x \in \mathbb{R}.\tag{8.3.4}$$

Schließlich haben wir die Potenzreihendarstellungen

$$\cos x = \sum_{l=0}^{\infty} \frac{(-1)^l}{(2l)!} x^{2l}, \quad \sin x = \sum_{l=0}^{\infty} \frac{(-1)^l}{(2l+1)!} x^{2l+1}, \quad x \in \mathbb{R},\tag{8.3.5}$$

wobei beide Reihen für alle $x \in \mathbb{R}$ absolut konvergieren.

Beweis. $\cos, \sin \in C^\infty(\mathbb{R})$ ist per Definition klar, da $\exp(\pm ix) \in C^\infty(\mathbb{R}, \mathbb{C})$ gilt gemäß Folgerung 8.1.5. Aus der Reihendarstellung erhalten wir auch

$$\frac{d}{dx}(e^{\pm ix}) \stackrel{(8.1.1)}{=} \sum_{k=1}^{\infty} k \frac{(\pm i)^k}{k!} x^{k-1} = \sum_{k=0}^{\infty} \frac{(\pm i)^{k+1}}{k!} x^k = \pm i e^{\pm ix}$$

und folgern daraus

$$\begin{aligned}\cos' x &= \frac{1}{2}(ie^{ix} - ie^{-ix}) = -\frac{1}{2i}(e^{ix} - e^{-ix}) = -\sin x, \\ \sin' x &= \frac{1}{2i}(ie^{ix} + ie^{-ix}) = \frac{1}{2}(e^{ix} + e^{-ix}) = \cos x,\end{aligned}$$

also (8.3.1). Die Eulersche Formel (8.3.2) entspricht der Definition von cos und sin. Und die Funktionalgleichung der Exponentialfunktion, Satz 8.2.2, liefert in Verbindung mit der

Eulerschen Formel:

$$\begin{aligned}
 \cos(x_1 + x_2) + i \sin(x_1 + x_2) &= e^{i(x_1 + x_2)} = e^{ix_1} e^{ix_2} \\
 &= (\cos x_1 + i \sin x_1)(\cos x_2 + i \sin x_2) \\
 &= (\cos x_1 \cos x_2 - \sin x_1 \sin x_2) + i(\cos x_1 \sin x_2 + \sin x_1 \cos x_2).
 \end{aligned}$$

Real- und Imaginärteil dieser Gleichung entsprechen den Formeln (8.3.3). Die Symmetrien (8.3.4) entnimmt man wieder direkt der Definition von \cos und \sin . Zum Beweis von (8.3.5) berechnen wir schließlich (beachte Satz 4.3.11):

$$\begin{aligned}
 \cos x + i \sin x &= e^{ix} = \sum_{k=0}^{\infty} \frac{1}{k!} i^k x^k \stackrel{\text{abs. Konv.}}{=} \sum_{k \text{ gerade}} \frac{1}{k!} i^k x^k + \sum_{k \text{ ungerade}} \frac{1}{k!} i^k x^k \\
 &= \sum_{l=0}^{\infty} \frac{1}{(2l)!} i^{2l} x^{2l} + \sum_{l=0}^{\infty} \frac{1}{(2l+1)!} i^{2l+1} x^{2l+1} \\
 &= \sum_{l=0}^{\infty} \frac{(-1)^l}{(2l)!} x^{2l} + i \sum_{l=0}^{\infty} \frac{(-1)^l}{(2l+1)!} x^{2l+1}.
 \end{aligned}$$

Vergleich von Real- und Imaginärteil dieser Identität liefert (8.3.5). Die absolute Konvergenz der beiden Reihen folgt nach dem Majorantenkriterium aus der absoluten Konvergenz der Exponentialreihe. Damit ist alles gezeigt. \square

Bemerkung 8.3.3. Wegen $\overline{e^{ix}} = e^{-ix}$ gilt $|e^{ix}|^2 = e^{ix} e^{-ix} = 1$ für alle $x \in \mathbb{R}$. Der Eulerschen Formel entnehmen wir daher die berühmte Relation

$$1 = \cos^2 x + \sin^2 x \quad \text{für alle } x \in \mathbb{R}.$$

Geometrisch stellt $f(x) := e^{ix}$, $x \in \mathbb{R}$, eine gleichförmige Bewegung mit Geschwindigkeit 1 auf der Einheitskreislinie dar, denn es gilt

$$|f(x)| = 1, \quad |f'(x)| = |ie^{ix}| = 1, \quad x \in \mathbb{R}.$$

Cosinus- und Sinusfunktion sind nach Definition die Projektionen dieser Kreisbewegung auf die reelle bzw. imaginäre Achse, weshalb man sie auch als *Kreisfunktionen* bezeichnet.

Wir wollen nun die Nullstellen der Kreisfunktionen untersuchen und beginnen mit dem

Satz 8.3.4: Die Kreiszahl π

Die Gleichung $\cos x = 0$ besitzt im Intervall $[0, 2]$ genau eine Lösung. Diese *kleinste*

positive Nullstelle von cos bezeichnen wir mit $\frac{\pi}{2}$. Es gilt dann

$$\cos x > 0 \quad \text{für alle } x \in \left[0, \frac{\pi}{2}\right), \quad \cos \frac{\pi}{2} = 0.$$

Beweis. Zunächst gilt per Definition $\cos 0 = \operatorname{Re}(e^0) = 1$. Und aus der Reihendarstellung von cos ermitteln wir

$$\begin{aligned} \cos x &= 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \frac{x^8}{8!} - \frac{x^{10}}{10!} + \frac{x^{12}}{12!} - + \dots \\ &= \left(1 - \frac{x^2}{2!} + \frac{x^4}{4!}\right) - \frac{x^6}{6!} \left(1 - \frac{x^2}{7 \cdot 8}\right) - \frac{x^{10}}{10!} \left(1 - \frac{x^2}{11 \cdot 12}\right) - \dots \end{aligned}$$

Für $x = 2$ erhalten wir

$$\cos 2 = -\frac{1}{3} - \frac{2^6}{6!} \left(1 - \frac{4}{7 \cdot 8}\right) - \frac{2^{10}}{10!} \left(1 - \frac{4}{11 \cdot 12}\right) - \dots < -\frac{1}{3}.$$

Nach dem Zwischenwertsatz, Satz 6.3.4, existiert somit ein $\xi \in (0, 2)$ mit $\cos \xi = 0$. Weiter entnehmen wir der Reihendarstellung von sin:

$$\begin{aligned} \cos' x &= -\sin x = -x + \frac{x^3}{3!} - \frac{x^5}{5!} + \frac{x^7}{7!} - \frac{x^9}{9!} + \frac{x^{11}}{11!} - + \dots \\ &= -x \left(1 - \frac{x^2}{2 \cdot 3}\right) - \frac{x^5}{5!} \left(1 - \frac{x^2}{6 \cdot 7}\right) - \frac{x^9}{9!} \left(1 - \frac{x^2}{10 \cdot 11}\right) - \dots < 0 \end{aligned}$$

für $x \in (0, 2)$. Nach Folgerung 7.3.11 (a) ist also cos in $[0, 2]$ streng monoton fallend und somit injektiv. Insbesondere ist die Nullstelle $\xi =: \frac{\pi}{2}$ eindeutig bestimmt und der Satz bewiesen. \square

Folgerung 8.3.5

Die Sinusfunktion ist im Intervall $[-\frac{\pi}{2}, \frac{\pi}{2}]$ streng monoton wachsend und es gilt

$$\sin\left(-\frac{\pi}{2}\right) = -1, \quad \sin 0 = 0, \quad \sin \frac{\pi}{2} = 1.$$

Die Cosinusfunktion ist im Intervall $[0, \pi]$ streng monoton fallend und es gilt

$$\cos 0 = 1, \quad \cos \frac{\pi}{2} = 0, \quad \cos \pi = -1.$$

Beweis. Da cos gerade ist, gilt gemäß (8.3.1): $\sin' x = \cos x > 0$ in $(-\frac{\pi}{2}, \frac{\pi}{2})$, d.h. sin ist in $[-\frac{\pi}{2}, \frac{\pi}{2}]$ streng monoton wachsend nach Folgerung 7.3.11 (a). Ferner gilt $\sin 0 = \operatorname{Im}(e^0) = 0$ und

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

also wegen der Monotonie $\sin(-\frac{\pi}{2}) = -1$, $\sin \frac{\pi}{2} = 1$. Schließlich erhalten wir die Aussagen über den Cosinus aus den Regeln der Phasenverschiebung (8.3.8) unten, die man mit den Additionstheoremen beweist. \square

Satz 8.3.6

Die Funktionen \cos und \sin sind 2π -periodisch, d.h. es gilt

$$\cos(x + 2\pi) = \cos x, \quad \sin(x + 2\pi) = \sin x \quad \text{für alle } x \in \mathbb{R}. \quad (8.3.6)$$

Ferner haben wir

$$\cos(x + \pi) = -\cos x, \quad \sin(x + \pi) = -\sin x \quad \text{für alle } x \in \mathbb{R} \quad (8.3.7)$$

und die Regeln der *Phasenverschiebung*

$$\cos\left(\frac{\pi}{2} - x\right) = \sin x, \quad \sin\left(\frac{\pi}{2} - x\right) = \cos x, \quad x \in \mathbb{R}, \quad (8.3.8)$$

Schließlich gilt für die Nullstellenmengen der Funktionen

$$\begin{aligned} \{x \in \mathbb{R} : \cos x = 0\} &= \left\{x = \frac{\pi}{2} + k\pi : k \in \mathbb{Z}\right\}, \\ \{x \in \mathbb{R} : \sin x = 0\} &= \{x = k\pi : k \in \mathbb{Z}\}. \end{aligned} \quad (8.3.9)$$

Beweis. Wir bemerken zunächst $e^{i\frac{\pi}{2}} = \cos \frac{\pi}{2} + i \sin \frac{\pi}{2} = i$. Damit folgt

$$e^{i\pi} = (e^{i\frac{\pi}{2}})^2 = i^2 = -1, \quad e^{2i\pi} = (e^{i\pi})^2 = (-1)^2 = 1,$$

also

$$\cos \pi = -1, \quad \sin \pi = 0; \quad \cos(2\pi) = 1, \quad \sin(2\pi) = 0.$$

Die Aussagen (8.3.6)-(8.3.8) folgen nun unmittelbar aus den Additionstheoremen (8.3.3). Ferner wissen wir bereits $\cos x > 0$ für alle $x \in (-\frac{\pi}{2}, \frac{\pi}{2})$ und $\cos \frac{\pi}{2} = 0$. Also folgt die Aussage (8.3.9) für den Cosinus aus Formel (8.3.7). Die Nullstellenmenge des Sinus lässt sich daraus m.H. der Phasenverschiebung (8.3.8) ablesen. \square

8.4. Polarkoordinatendarstellung komplexer Zahlen

Bei der Einführung der komplexen Zahlen in Abschnitt A.2.5 hatten wir eine geometrische Deutung der Multiplikation versprochen, die wir nun mit Hilfe der Polarkoordinatendarstellung nachholen wollen. Die Idee für letztere ist folgende: Als Punkt in der Ebene hat $z \in \mathbb{C} \setminus \{0\}$ einen Abstand von $|z| > 0$ zum Nullpunkt. Als Punkt auf der Einheitskreislinie lässt sich dann $\frac{z}{|z|}$ durch den Winkel (z.B.) zur positiven x -Achse beschreiben; dieser wiederum lässt sich mit \cos und \sin bzw. e^{ix} ausdrücken, vgl. Bemerkung 8.3.3. Genauer haben wir den

Satz 8.4.1: Polarkoordinaten

Jede komplexe Zahl $z \in \mathbb{C}$ besitzt eine Darstellung

$$z = r e^{i\varphi} = r(\cos \varphi + i \sin \varphi) \quad (8.4.1)$$

mit einem $\varphi \in \mathbb{R}$ und $r = |z|$. Für $z \neq 0$ ist die Darstellung (8.4.1) eindeutig, wenn wir $\varphi \in [0, 2\pi)$ fordern.

Für den Beweis benötigen wir noch den

Hilfssatz 8.4.2

Alle Lösungen $x \in \mathbb{R}$ der Gleichung $e^{ix} = 1$ haben die Form $x = 2k\pi$ mit einem $k \in \mathbb{Z}$.

Beweis. Wir beachten

$$\sin \frac{x}{2} = \frac{1}{2i} (e^{i\frac{x}{2}} - e^{-i\frac{x}{2}}) = \frac{e^{-i\frac{x}{2}}}{2i} (e^{ix} - 1), \quad x \in \mathbb{R}.$$

Also gilt $e^{ix} = 1 \Leftrightarrow \sin \frac{x}{2} = 0$. Die Behauptung ergibt sich nun aus (8.3.9). \square

Bemerkung 8.4.3. Aus Hilfssatz 8.4.2 und Satz 8.2.6 folgt noch

$$\{z \in \mathbb{C} : \exp z = 1\} = \{2ik\pi : k \in \mathbb{Z}\}.$$

Denn: Mit $z = x + iy \in \mathbb{C}$ haben wir $\exp z = e^x e^{iy}$ nach Satz 8.2.2. Wegen $|e^{iy}| = 1$ folgt aus $\exp z = 1$ zunächst $x = 0$ und damit $\exp z = e^{iy}$, also die Behauptung. \square

Beweis von Satz 8.4.1. 1. Für $z = 0$ ist $r = |z| = 0$ und (8.4.1) gilt mit beliebigem $\varphi \in \mathbb{R}$. Sei also $z = x + iy \neq 0$. Dann folgt $r := |z| > 0$ und wir können $\xi := \frac{x}{r}$, $\eta := \frac{y}{r}$ setzen. Es gilt somit

$$z = r(\xi + i\eta), \quad \xi^2 + \eta^2 = 1. \quad (8.4.2)$$

Insbesondere ist $\xi \in [-1, 1] = [\cos \pi, \cos 0]$ erfüllt; vgl. Folgerung 8.3.5. Nach dem Zwischenwertsatz existiert also ein $\alpha \in [0, \pi]$ mit $\cos \alpha = \xi$. Hieraus folgt noch

$$\eta = \pm \sqrt{1 - \xi^2} = \pm \sqrt{1 - \cos^2 \alpha} = \pm \sin \alpha.$$

Man beachte $\sin \alpha \geq 0$ wegen $\sin \frac{\pi}{2} = 1$ und (8.3.9).

- 1. Fall: Für $y \geq 0$ ist $\eta \geq 0$, also $\eta = \sin \alpha$. Dann setzen wir $\varphi := \alpha \in [0, \pi]$ und erhalten $\xi = \cos \varphi$, $\eta = \sin \varphi$, also aus (8.4.2) die gesuchte Darstellung (8.4.1).

- 2. Fall: Für $y < 0$ folgt $\alpha \in (0, \pi)$ und $\eta = -\sin \alpha$. Mit $\varphi := 2\pi - \alpha \in (\pi, 2\pi)$ erhalten wir dann aus den Symmetrieeigenschaften (8.3.4) und der Periodizität (8.3.6):

$$\begin{aligned}\xi &= \cos \alpha = \cos(2\pi - \varphi) = \cos \varphi, \\ \eta &= -\sin \alpha = -\sin(2\pi - \varphi) = \sin \varphi,\end{aligned}$$

also wieder (8.4.1).

2. Man beachte, dass der in Teil 1 des Beweises erklärte Winkel φ in $[0, 2\pi)$ liegt. Gibt es ein weiteres $\psi \in [0, 2\pi)$ mit $z = re^{i\psi}$, so folgt $e^{i\varphi} = e^{i\psi}$ bzw. $e^{i(\varphi-\psi)} = 1$. Hilfssatz 8.4.2 liefert also $\varphi - \psi = 2k\pi$ mit einem $k \in \mathbb{Z}$. Aus $|\varphi - \psi| < 2\pi$ erhalten wir dann $k = 0$ bzw. $\varphi = \psi$, wie behauptet. \square

Bemerkung 8.4.4. (a) Der Satz besagt im Speziellen, dass die Abbildung $\varphi \mapsto e^{i\varphi}$ das Intervall $[0, 2\pi)$ bijektiv auf die Einheitskreislinie S^1 abbildet. Die Periodizität der Kreisfunktionen liefert noch, dass dabei $[0, 2\pi)$ durch ein beliebiges halboffenes Intervall der Länge 2π ersetzt werden kann.

- (b) Für beliebiges $z = |z|e^{i\varphi} \neq 0$ misst $\varphi \in [0, 2\pi)$ den Winkel zwischen der positiven reellen Achse und dem Vektor z , gemessen in mathematisch positivem Sinn. Er wird *Argument von z* genannt und mit $\varphi = \arg z$ bezeichnet. Seine Berechnung gelingt mit Hilfe der Arcus-Funktionen; siehe Bemerkung 8.5.5 unten.
- (c) Mit einer Forderung $\varphi \in [\varphi_0, \varphi_0 + 2\pi)$ oder $\varphi \in (\varphi_0, \varphi_0 + 2\pi]$ für beliebiges $\varphi_0 \in \mathbb{R}$ folgt $\varphi = \arg z + 2k\pi$ mit einem (eindeutigen) $k = k(\varphi_0) \in \mathbb{Z}$; vgl. (a). φ misst also wieder den Winkel zur positiven x -Achse, wobei nun zusätzlich k -mal um den Ursprung gelaufen wird.

Geometrische Interpretation der komplexen Multiplikation

Für zwei komplexe Zahlen $z_1 = |z_1|e^{i\varphi_1}$ und $z_2 = |z_2|e^{i\varphi_2}$ mit $\varphi_1, \varphi_2 \in [0, 2\pi)$ erhalten wir mit Satz 8.2.2:

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

Bei der Multiplikation werden also die Beträge multipliziert und die Argumente (= Winkel) addiert.

Wir nutzen die Polarkoordinatendarstellung zum ausgelassen Beweis des Hilfssatzes 6.5.4, den wir hier noch ergänzen zur

Folgerung 8.4.5: MOIVRE, n -te Wurzel

Zu beliebigen $n \in \mathbb{N}$ und $c \in \mathbb{C} \setminus \{0\}$ hat die Gleichung $z^n = c$ genau n verschiedene

Lösungen, nämlich

$$z_k = \sqrt[n]{|c|} \exp\left(i \frac{\varphi + 2k\pi}{n}\right), \quad k = 0, \dots, n-1.$$

Dabei ist $\varphi = \arg c \in [0, 2\pi)$ gesetzt worden. Die Zahlen z_0, \dots, z_{n-1} heißen *n-te Wurzeln* von c .

Beweis. Dass z_0, \dots, z_{n-1} die Gleichung $z^n = c$ lösen, ist sofort klar:

$$z_k^n \stackrel{\text{Satz 8.2.2}}{=} |c| e^{i\varphi} e^{2ik\pi} \stackrel{\text{HS 8.4.2}}{=} |c| e^{i\varphi} \stackrel{\text{Satz 8.4.1}}{=} c, \quad k = 0, \dots, n-1.$$

Weiter gilt $z_k \neq z_l$ für $k \neq l$, denn $z_k = z_l$ impliziert nach Anwendung von Satz 8.2.2

$$\exp\left(i \frac{2k\pi}{n}\right) = \exp\left(i \frac{2l\pi}{n}\right),$$

also nach Multiplikation mit $\exp\left(-i \frac{2l\pi}{n}\right)$ und abermaliger Anwendung der Funktionalgleichung

$$\exp\left(i \frac{2(k-l)\pi}{n}\right) = 1.$$

Hilfssatz 8.4.2 entnehmen wir also $\frac{k-l}{n} \in \mathbb{Z}$, was wegen $|k-l| \leq n-1$ nur für $k=l$ möglich ist.

Schließlich gibt es auch keine weiteren Nullstellen des Polynoms $p(z) = z^n - c$, $z \in \mathbb{C}$, nach Folgerung 6.5.9 des Fundamentalsatzes der Algebra. \square

Beispiel 8.4.6. Für $c = 1$ erhalten wir als Lösungen der Gleichung $z^n = 1$ die *n-ten Einheitswurzeln*

$$z_k = \exp\left(i \frac{2k\pi}{n}\right) = \cos\left(\frac{2k\pi}{n}\right) + i \sin\left(\frac{2k\pi}{n}\right), \quad k = 0, \dots, n-1.$$

8.5. Weitere elementare Funktionen

Wir beschließen dieses Kapitel mit einer losen, aber keineswegs vollständigen Sammlung weiterer Funktionen, die sich aus der komplexen Exponentialfunktion ableiten. Wir beginnen mit der

Definition 8.5.1

$$\begin{aligned} \tan x &:= \frac{\sin x}{\cos x}, & x \in \mathbb{R} \setminus \left\{ \frac{\pi}{2} + k\pi : k \in \mathbb{Z} \right\} &=: D_{\tan} & \text{ (Tangens),} \\ \cot x &:= \frac{\cos x}{\sin x}, & x \in \mathbb{R} \setminus \{ k\pi : k \in \mathbb{Z} \} &=: D_{\cot} & \text{ (Cotangens).} \end{aligned}$$

Satz 8.5.2

Tangens und Cotangens sind in ihren Definitionsbereichen unendlich oft differenzierbar und es gelten

$$\begin{aligned}\tan' x &= \frac{d}{dx} \tan x = 1 + \tan^2 x = \frac{1}{\cos^2 x}, \quad x \in D_{\tan}, \\ \cot' x &= \frac{d}{dx} \cot x = -(1 + \cot^2 x) = -\frac{1}{\sin^2 x}, \quad x \in D_{\cot}.\end{aligned}\tag{8.5.1}$$

Ferner haben wir für x aus den jeweiligen Definitionsbereichen D_{\tan}, D_{\cot} :

$$\tan(x + \pi) = \tan x, \quad \cot(x + \pi) = \cot x$$

sowie

$$\tan\left(\frac{\pi}{2} - x\right) = \cot x, \quad \cot\left(\frac{\pi}{2} - x\right) = \tan x,$$

und es gelten die *Additionstheoreme*

$$\begin{aligned}\tan(x_1 + x_2) &= \frac{\tan x_1 + \tan x_2}{1 - \tan x_1 \tan x_2}, \quad x_1, x_2, x_1 + x_2 \in D_{\tan}, \\ \cot(x_1 + x_2) &= \frac{-1 + \cot x_1 \cot x_2}{\cot x_1 + \cot x_2}, \quad x_1, x_2, x_1 + x_2 \in D_{\cot}.\end{aligned}$$

Schließlich ist \tan in $(-\frac{\pi}{2}, \frac{\pi}{2})$ streng monoton wachsend mit

$$\lim_{x \rightarrow -\frac{\pi}{2}+} \tan x = -\infty, \quad \tan 0 = 0, \quad \lim_{x \rightarrow \frac{\pi}{2}-} \tan x = +\infty,$$

und \cot ist in $(0, \pi)$ streng monoton fallend mit

$$\lim_{x \rightarrow 0+} \cot x = +\infty, \quad \cot\left(\frac{\pi}{2}\right) = 0, \quad \lim_{x \rightarrow \pi-} \cot x = -\infty.$$

Beweis. Direkt aus den Aussagen über die Cosinus- und Sinusfunktion (\rightarrow Nachrechnen). \square

Aufgrund des Monotonieverhaltens von \sin , \cos , \tan und \cot können wir die entsprechenden Umkehrfunktionen erklären, wenn wir uns auf geeignete Monotonieintervalle beschränken: Wir

wählen die Bereiche

$$\begin{aligned}
 y = \sin x, \quad -\frac{\pi}{2} \leq x \leq \frac{\pi}{2} &\Rightarrow -1 \leq y \leq 1, \\
 y = \cos x, \quad 0 \leq x \leq \pi &\Rightarrow -1 \leq y \leq 1, \\
 y = \tan x, \quad -\frac{\pi}{2} < x < \frac{\pi}{2} &\Rightarrow -\infty < y < +\infty, \\
 y = \cot x, \quad 0 < x < \pi &\Rightarrow -\infty < y < +\infty.
 \end{aligned} \tag{8.5.2}$$

Definition 8.5.3

Die Umkehrfunktionen zu \sin , \cos , \tan und \cot auf den in (8.5.2) angegebenen Monotonieintervallen heißen *Arcus Sinus*, *Arcus Cosinus*, *Arcus Tangens* bzw. *Arcus Cotangens* und werden mit

$$\begin{aligned}
 \arcsin &:= \sin^{-1} : [-1, 1] \rightarrow \left[-\frac{\pi}{2}, \frac{\pi}{2}\right], & \arccos &:= \cos^{-1} : [-1, 1] \rightarrow [0, \pi], \\
 \arctan &:= \tan^{-1} : \mathbb{R} \rightarrow \left(-\frac{\pi}{2}, \frac{\pi}{2}\right), & \operatorname{arccot} &:= \cot^{-1} : \mathbb{R} \rightarrow (0, \pi)
 \end{aligned}$$

bezeichnet.

Satz 8.5.4

Es gelten $\arcsin, \arccos \in C^\infty((-1, 1))$ und $\arctan, \operatorname{arccot} \in C^\infty(\mathbb{R})$, und wir haben

$$\begin{aligned}
 \arcsin' y &= \frac{1}{\sqrt{1-y^2}}, & \arccos' y &= -\frac{1}{\sqrt{1-y^2}}, & y &\in (-1, 1), \\
 \arctan' y &= \frac{1}{1+y^2}, & \operatorname{arccot}' y &= -\frac{1}{1+y^2}, & y &\in \mathbb{R}.
 \end{aligned} \tag{8.5.3}$$

Ferner gelten die Relationen

$$\begin{aligned}
 \arcsin y + \arccos y &= \frac{\pi}{2} \quad \text{für alle } y \in [-1, 1], \\
 \arctan y + \operatorname{arccot} y &= \frac{\pi}{2} \quad \text{für alle } y \in \mathbb{R}.
 \end{aligned} \tag{8.5.4}$$

Beweis. Da die ersten Ableitungen der unendlich oft differenzierbaren Funktionen \sin, \tan auf $(-\frac{\pi}{2}, \frac{\pi}{2})$ und \cos, \cot auf $(0, \pi)$ nicht verschwinden, sind die Umkehrfunktionen in den angegebenen Bereichen ebenfalls unendlich oft differenzierbar nach Folgerung 7.4.6 (b). Gemäß

Satz 7.2.5 gelten weiter

$$\begin{aligned}\arcsin' y &= \frac{1}{\sin'(\arcsin y)} = \frac{1}{\sqrt{1 - \sin^2(\arcsin y)}} = \frac{1}{\sqrt{1 - y^2}}, \quad |y| < 1, \\ \arctan' y &= \frac{1}{\tan'(\arctan y)} = \frac{1}{1 + \tan^2(\arctan y)} = \frac{1}{1 + y^2}, \quad y \in \mathbb{R},\end{aligned}$$

und entsprechend erhalten wir die ersten Ableitungen für arccos und arccot.

Zum Beweis der ersten Relation in (8.5.4) wenden wir arccos auf die Identität $y = \sin x = \cos(\frac{\pi}{2} - x)$, $x \in [-\frac{\pi}{2}, \frac{\pi}{2}]$, an:

$$\arccos y = \frac{\pi}{2} - x = \frac{\pi}{2} - \arcsin y, \quad y \in [-1, 1].$$

Entsprechend wenden wir arccot auf $y = \tan x = \cot(\frac{\pi}{2} - x)$, $x \in \mathbb{R}$, an und erhalten die zweite Relation in (8.5.4). \square

Bemerkung 8.5.5. Die Konstruktion im Beweis von Satz 8.4.1 und die obige Definition von arccos zeigen, dass sich das Argument $\arg z$ einer Zahl $z = x + iy \neq 0$ wie folgt bestimmen lässt:

$$\arg z = \begin{cases} \arccos\left(\frac{x}{|z|}\right), & \text{falls } y \geq 0 \\ 2\pi - \arccos\left(\frac{x}{|z|}\right), & \text{falls } y < 0 \end{cases} \in [0, 2\pi).$$

Ausgehend von der komplexen Exponentialfunktion können wir auch die *komplexe Cosinus-* bzw. *Sinusfunktion* erklären:

$$\cos z := \frac{1}{2}(e^{iz} + e^{-iz}), \quad \sin z := \frac{1}{2i}(e^{iz} - e^{-iz}), \quad z \in \mathbb{C}.$$

Für $z = x \in \mathbb{R}$ erhalten wir dann die reellen Kreisfunktionen. Für $z = ix$, $x \in \mathbb{R}$, finden wir die *Hyperbelfunktionen*:

Definition 8.5.6

$$\begin{aligned}\cosh x &:= \cos(ix) = \frac{1}{2}(e^x + e^{-x}), \quad x \in \mathbb{R} && (\text{Cosinus hyperbolicus}), \\ \sinh x &:= -i \sin(ix) = \frac{1}{2}(e^x - e^{-x}), \quad x \in \mathbb{R} && (\text{Sinus hyperbolicus}).\end{aligned}$$

Bemerkung 8.5.7. Während $(\cos x, \sin x)$, $x \in \mathbb{R}$, eine Parametrisierung der Einheitskreislinie liefert, ergibt $(\cosh x, \sinh x)$, $x \in \mathbb{R}$, eine Parametrisierung des rechten Astes der Hyperbel $\{(x, y) \in \mathbb{R}^2 : x^2 - y^2 = 1\}$. Wir verzichten auf eine detaillierte Diskussion der Hyperbelfunktionen (verweisen dazu auf die Literatur und die Übungen) und halten hier nur fest

Satz 8.5.8

Es gilt $\cosh, \sinh \in C^\infty(\mathbb{R})$ sowie

$$\cosh' x = \sinh x, \quad \sinh' x = \cosh x, \quad x \in \mathbb{R}.$$

Beweis. Sofort klar aus der Definition und den Eigenschaften der reellen e -Funktion. □

Kapitel 9.

Das eindimensionale RIEMANNsche Integral

In diesem Kapitel führen wir den klassischen Integralbegriff nach B. RIEMANN ein. Abschnitt 9.1 enthält die Begriffsbildung, in Abschnitt 9.2 geben wir die drei zentralen Integrationskriterien an und folgern daraus wichtige Eigenschaften des Integrals. Der Additivität des Integrals ist Abschnitt 9.3 gewidmet. Den zentralen Fundamentalsatz der Differential- und Integralrechnung leiten wir in Abschnitt 9.4 her; diesen Zusammenhang zwischen Integral und Ableitungsbegriff nutzen wir dann zur Herleitung drei wichtiger Resultate in Abschnitt 9.5. In Abschnitt 9.6 führen wir den Begriff des uneigentlichen Integrals ein, der uns u.a. die Flächenbestimmung unter unbeschränkten Funktionsgraphen ermöglicht. Schließlich betrachten wir die Approximation von Funktionen mittels TAYLORpolynomen bzw. deren Darstellung durch TAYLORreihen in den abschließenden Abschnitten 9.7 und 9.8.

9.1. Der Integralbegriff

Ist $f : [a, b] \rightarrow \mathbb{R}$ eine positive Funktion, so wollen wir das *bestimmte Integral*

$$\int_a^b f(x) dx$$

als Flächeninhalt des Stückes des \mathbb{R}^2 erklären, das von der x -Achse und (dem Graphen) der Funktion f einerseits und den senkrechten Geraden durch $(a, 0)$ bzw. $(b, 0)$ andererseits berandet wird. Die grundlegende Idee hierbei ist, den Wert $\int_a^b f(x) dx$ durch den elementargeometrischen Flächeninhalt von einbeschriebenen Rechtecken geeigneter Höhe zu approximieren. Dieser Ansatz soll nun präzise gefasst werden. Wir beginnen mit der

Definition 9.1.1: Zerlegungen
Sei $I = [a, b]$ kompakt, $a < b$. Eine Menge $\mathcal{Z} = \{x_0, \dots, x_N\} \subset I$ mit $N \in \mathbb{N}$ und
$a = x_0 < x_1 < \dots < x_N = b$

nennen wir *Zerlegung* des Intervalls I . Die Punkte x_0, \dots, x_N heißen *Teilpunkte* von \mathcal{Z} . Wir schreiben auch

$$I_j := [x_{j-1}, x_j] \quad \text{und} \quad \Delta x_j := x_j - x_{j-1} = |I_j|, \quad j = 1, \dots, N,$$

für die von \mathcal{Z} induzierten Teilintervalle von I und deren Längen. Die Länge des größten Teilintervalls

$$\Delta(\mathcal{Z}) := \max\{\Delta x_1, \dots, \Delta x_N\}$$

wird als *Feinheit* von \mathcal{Z} bezeichnet.

Definition 9.1.2: Unter-, Ober- und Zwischensumme

Seien $I \subset \mathbb{R}$ ein kompaktes Intervall und $\mathcal{Z} = \{x_0, \dots, x_N\}$ eine Zerlegung von I gemäß Definition 9.1.1. Weiter sei $f : I \rightarrow \mathbb{R}$ beschränkt.

(a) Mit den Abkürzungen

$$\underline{m}_j := \inf_{I_j} f = \inf\{f(x) : x \in I_j\},$$

$$\overline{m}_j := \sup_{I_j} f = \sup\{f(x) : x \in I_j\}, \quad j = 1, \dots, N,$$

bilden wir die *Untersumme*

$$\underline{S}_{\mathcal{Z}}(f) := \sum_{j=1}^N \underline{m}_j \Delta x_j$$

und die *Obersumme*

$$\overline{S}_{\mathcal{Z}}(f) := \sum_{j=1}^N \overline{m}_j \Delta x_j$$

von f zu \mathcal{Z} .

(b) Aus jedem Teilintervall I_j wählen wir ein $\xi_j \in I_j$ und setzen $\xi := (\xi_1, \dots, \xi_N)$. Dann nennen wir

$$S_{\mathcal{Z}}(f) = S_{\mathcal{Z}}(f, \xi) := \sum_{j=1}^N f(\xi_j) \Delta x_j$$

eine *RIEMANNSche Zwischensumme* von f zu \mathcal{Z} und ξ .

Bemerkung 9.1.3. Wegen $\underline{m}_j \leq f(\xi_j) \leq \overline{m}_j$ für $\xi_j \in I_j$ gilt offenbar

$$\underline{S}_{\mathcal{Z}}(f) \leq S_{\mathcal{Z}}(f, \xi) \leq \overline{S}_{\mathcal{Z}}(f) \tag{9.1.1}$$

für jede Zerlegung von I und jede RIEMANNSche Zwischensumme. Ferner können wir wegen

Hilfssatz 3.5.4 zu jedem $\varepsilon > 0$ und jeder Zerlegung \mathcal{Z} Zwischenwerte $\underline{\xi}_j, \bar{\xi}_j \in I_j, j = 1, \dots, N$, so angeben, dass gilt

$$S_{\mathcal{Z}}(f, \underline{\xi}) < \underline{S}_{\mathcal{Z}}(f) + \varepsilon, \quad S_{\mathcal{Z}}(f, \bar{\xi}) > \bar{S}_{\mathcal{Z}}(f) - \varepsilon \quad (9.1.2)$$

(\rightarrow Nachrechnen).

Definition 9.1.4

- Eine Zerlegung \mathcal{Z}^* von I heißt *Verfeinerung* der Zerlegung \mathcal{Z} von I , wenn alle Teilpunkte von \mathcal{Z} auch Teilpunkte von \mathcal{Z}^* sind.
- Eine *gemeinsame Verfeinerung* $\mathcal{Z}_1 \vee \mathcal{Z}_2$ zweier Zerlegungen $\mathcal{Z}_1, \mathcal{Z}_2$ von I ist die Zerlegung von I , deren Teilpunkte zu \mathcal{Z}_1 oder \mathcal{Z}_2 gehören (doppelt auftretende Teilpunkte werden nur einmal gezählt).

Bemerkung 9.1.5. $\mathcal{Z}_1 \vee \mathcal{Z}_2$ ist also sowohl Verfeinerung von \mathcal{Z}_1 als auch von \mathcal{Z}_2 .

Hilfssatz 9.1.6

(a) Ist \mathcal{Z}^* Verfeinerung von \mathcal{Z} , so gilt

$$\underline{S}_{\mathcal{Z}}(f) \leq \underline{S}_{\mathcal{Z}^*}(f) \leq \bar{S}_{\mathcal{Z}^*}(f) \leq \bar{S}_{\mathcal{Z}}(f).$$

(b) Sind $\mathcal{Z}_1, \mathcal{Z}_2$ zwei beliebige Zerlegungen von I , so gilt

$$\underline{S}_{\mathcal{Z}_1}(f) \leq \bar{S}_{\mathcal{Z}_2}(f).$$

Beweis. (a) Seien I_l^* ein Teilintervall von \mathcal{Z}^* und I_j ein Teilintervall von \mathcal{Z} mit $I_l^* \subset I_j$.

Dann folgt

$$\underline{m}_l^* := \inf_{I_l^*} f \geq \inf_{I_j} f =: \underline{m}_j, \quad \bar{m}_l^* := \sup_{I_l^*} f \leq \sup_{I_j} f =: \bar{m}_j$$

und somit

$$\underline{m}_j \Delta x_j = \underline{m}_j \sum_{l: I_l^* \subset I_j} \Delta x_l^* \leq \sum_{l: I_l^* \subset I_j} \underline{m}_l^* \Delta x_l^*, \quad \bar{m}_j \Delta x_j \geq \sum_{l: I_l^* \subset I_j} \bar{m}_l^* \Delta x_l^*.$$

Durch Summierung über j erhalten wir also

$$\underline{S}_{\mathcal{Z}}(f) \leq \underline{S}_{\mathcal{Z}^*}(f) \stackrel{(9.1.1)}{\leq} \bar{S}_{\mathcal{Z}^*}(f) \leq \bar{S}_{\mathcal{Z}}(f).$$

(b) Wir wenden (a) auf die gemeinsame Verfeinerung $\mathcal{Z}_1 \vee \mathcal{Z}_2 =: \mathcal{Z}$ an und erhalten

$$\underline{S}_{\mathcal{Z}_1}(f) \leq \underline{S}_{\mathcal{Z}}(f) \leq \bar{S}_{\mathcal{Z}}(f) \leq \bar{S}_{\mathcal{Z}_2}(f),$$

wie behauptet. \square

Definition 9.1.7: Unter- und Oberintegral

Ist $f : I \rightarrow \mathbb{R}$ beschränkt, $I = [a, b]$, so erklären wir das *Unterintegral* $\underline{\mathcal{I}}(f)$ bzw. *Oberintegral* $\bar{\mathcal{I}}(f)$ von f als

$$\underline{\mathcal{I}}(f) := \sup \{ \underline{S}_{\mathcal{Z}}(f) : \mathcal{Z} \text{ ist Zerlegung von } I \},$$

$$\bar{\mathcal{I}}(f) := \inf \{ \bar{S}_{\mathcal{Z}}(f) : \mathcal{Z} \text{ ist Zerlegung von } I \}.$$

Bemerkung 9.1.8. Ist \mathcal{Z} eine beliebige Zerlegung von $I = [a, b]$ und $f : I \rightarrow \mathbb{R}$ beschränkt, so gilt nach Hilfssatz 9.1.6 (a):

$$-\infty < |I| \inf_I f \leq \underline{S}_{\mathcal{Z}}(f) \leq \bar{S}_{\mathcal{Z}}(f) \leq |I| \sup_I f < +\infty.$$

Also sind $\underline{\mathcal{I}}(f), \bar{\mathcal{I}}(f) \in \mathbb{R}$ wohldefiniert. Hilfssatz 9.1.6 (b) entnehmen wir dann durch sup- bzw. inf-Bildung:

$$\underline{S}_{\mathcal{Z}}(f) \leq \underline{\mathcal{I}}(f) \leq \bar{\mathcal{I}}(f) \leq \bar{S}_{\mathcal{Z}}(f) \quad \text{mit beliebiger Zerlegung } \mathcal{Z} \text{ von } I. \quad (9.1.3)$$

Definition 9.1.9: Integrierbarkeit und Integral

Eine beschränkte Funktion $f : I \rightarrow \mathbb{R}$ über dem Intervall $I = [a, b]$ heißt *RIEMANN-integrierbar*, wenn gilt $\underline{\mathcal{I}}(f) = \bar{\mathcal{I}}(f)$. Wir setzen dann

$$\mathcal{I}(f) := \underline{\mathcal{I}}(f) = \bar{\mathcal{I}}(f)$$

für das (*bestimmte*) *RIEMANNSche Integral* von f über $[a, b]$. Alternative Symbole sind

$$\mathcal{I}(f) = \int_a^b f(x) dx = \int_I f(x) dx.$$

Die Klasse aller RIEMANN-integrierbaren Funktionen auf I wird mit $\mathcal{R}(I)$ bezeichnet.

Bemerkung 9.1.10. ▶ Da wir weitere Integralbegriffe erst in der Analysis III kennenlernen werden, sagen wir i.F. kurz *integrierbar* für RIEMANN-integrierbar und *Integral* für RIEMANNSches Integral.

- ▶ Nach Definition 9.1.9 ist jede integrierbare Funktion beschränkt. In Abschnitt 9.6 werden wir (unter Zusatzvoraussetzungen) auch für unbeschränkte Funktionen einen Integralbegriff erklären.

Für komplex- und vektorwertige Funktionen erklären wir das Integral komponentenweise:

Definition 9.1.11

- (a) Eine beschränkte Funktion $f : I \rightarrow \mathbb{C}$ heißt *integrierbar* auf $I = [a, b]$, wenn $\operatorname{Re} f, \operatorname{Im} f \in \mathcal{R}(I)$ gilt. Wir setzen dann

$$\int_a^b f(x) dx := \int_a^b \operatorname{Re} f(x) dx + i \int_a^b \operatorname{Im} f(x) dx$$

und schreiben $f \in \mathcal{R}(I, \mathbb{C})$.

- (b) Entsprechend heißt $f = (f_1, \dots, f_d) : I \rightarrow \mathbb{R}^d$ *integrierbar* auf $I = [a, b]$, wenn $f_l \in \mathcal{R}(I)$ für alle $l = 1, \dots, d$ gilt. Wir schreiben dann $f \in \mathcal{R}(I, \mathbb{R}^d)$ und setzen

$$\int_a^b f(x) dx := \left(\int_a^b f_1(x) dx, \dots, \int_a^b f_d(x) dx \right).$$

9.2. Kriterien und Rechenregeln

Wir geben zunächst drei äquivalente Kriterien für die Integrierbarkeit reellwertiger Funktionen an, die uns z.B. beim Ableiten von Rechenregeln für das Integral helfen werden.

Satz 9.2.1: Integrabilitätskriterium I

Für eine beschränkte Funktion $f : I \rightarrow \mathbb{R}$, $I = [a, b]$, gilt:

$$f \in \mathcal{R}(I) \quad \Leftrightarrow \quad \text{Für alle } \varepsilon > 0 \text{ existiert eine Zerlegung } \mathcal{Z} \text{ von } I \text{ mit } \bar{S}_{\mathcal{Z}}(f) - \underline{S}_{\mathcal{Z}}(f) < \varepsilon.$$

Beweis. ▶ „ \Leftarrow “: Aus (9.1.3) erhalten wir

$$0 \leq \overline{\mathcal{I}}(f) - \underline{\mathcal{I}}(f) \leq \overline{\mathcal{S}}_{\mathcal{Z}}(f) - \underline{\mathcal{S}}_{\mathcal{Z}}(f) < \varepsilon$$

für beliebiges $\varepsilon > 0$ und geeignete Zerlegung \mathcal{Z} . Also folgt $\overline{\mathcal{I}}(f) = \underline{\mathcal{I}}(f)$ bzw. $f \in \mathcal{R}(I)$.

▶ „ \Rightarrow “: Nach Definition 9.1.7 und Hilfssatz 3.5.4 existieren zu beliebig gewähltem $\varepsilon > 0$ Zerlegungen $\underline{\mathcal{Z}}$ und $\overline{\mathcal{Z}}$ von I mit

$$\underline{\mathcal{S}}_{\underline{\mathcal{Z}}}(f) > \underline{\mathcal{I}}(f) - \frac{\varepsilon}{2}, \quad \overline{\mathcal{S}}_{\overline{\mathcal{Z}}}(f) < \overline{\mathcal{I}}(f) + \frac{\varepsilon}{2}.$$

Setzen wir $\mathcal{Z} := \underline{\mathcal{Z}} \vee \overline{\mathcal{Z}}$, so liefert Hilfssatz 9.1.6 (a):

$$\overline{\mathcal{S}}_{\mathcal{Z}}(f) - \underline{\mathcal{S}}_{\mathcal{Z}}(f) \leq \overline{\mathcal{S}}_{\overline{\mathcal{Z}}}(f) - \underline{\mathcal{S}}_{\underline{\mathcal{Z}}}(f) < \overline{\mathcal{I}}(f) - \underline{\mathcal{I}}(f) + \varepsilon \stackrel{f \in \mathcal{R}(I)}{=} \varepsilon,$$

wie behauptet. □

Beispiel 9.2.2. Die Dirichletsche Sprungfunktion $\chi_{\mathbb{Q}}$ ist über kein Intervall $I = [a, b]$, $a < b$, integrierbar.

Denn: Da sowohl die rationalen als auch die irrationalen Zahlen in I dicht liegen, gilt für jede Zerlegung $\mathcal{Z} = \{x_0, \dots, x_N\}$ von I :

$$\underline{m}_j = \inf_{I_j} \chi_{\mathbb{Q}} = 0, \quad \overline{m}_j = \sup_{I_j} \chi_{\mathbb{Q}} = 1, \quad j = 1, \dots, N.$$

Wir erhalten also

$$\overline{\mathcal{S}}_{\mathcal{Z}}(\chi_{\mathbb{Q}}) - \underline{\mathcal{S}}_{\mathcal{Z}}(\chi_{\mathbb{Q}}) = \sum_{j=1}^N (1 - 0) \Delta x_j = |I| > 0.$$

Folglich ist für $0 < \varepsilon \leq |I|$ das Kriterium aus Satz 9.2.1 nicht erfüllbar, es folgt $\chi_{\mathbb{Q}} \notin \mathcal{R}(I)$. □

Satz 9.2.3: Integrabilitätskriterium II

Für eine beschränkte Funktion $f : I \rightarrow \mathbb{R}$, $I = [a, b]$, gilt:

$$f \in \mathcal{R}(I) \quad \Leftrightarrow \quad \begin{array}{l} \text{Für alle } \varepsilon > 0 \text{ existiert ein } \delta = \delta(\varepsilon) > 0, \text{ so dass gilt:} \\ \overline{\mathcal{S}}_{\mathcal{Z}}(f) - \underline{\mathcal{S}}_{\mathcal{Z}}(f) < \varepsilon \text{ für alle Zerlegungen } \mathcal{Z} \text{ mit } \Delta(\mathcal{Z}) < \delta. \end{array}$$

Beweis. \rightarrow Ergänzungen. □

Zur Formulierung des dritten Kriteriums benötigen wir noch den folgenden Begriff:

Definition 9.2.4

Eine Folge $\{Z_n\}_n$ von Zerlegungen eines kompakten Intervalls I nennen wir *Zerlegungsfolge*. Eine Zerlegungsfolge heißt *ausgezeichnet*, wenn für deren Feinheiten gilt $\Delta(Z_n) \rightarrow 0$ ($n \rightarrow \infty$).

Satz 9.2.5: Integrabilitätskriterium III

Für eine beschränkte Funktion $f : I \rightarrow \mathbb{R}$, $I = [a, b]$, gilt:

$$f \in \mathcal{R}(I) \quad \Leftrightarrow \quad \begin{array}{l} \text{Für jede ausgezeichnete Zerlegungsfolge } \{Z_n\}_n \text{ von } I \\ \text{und jede Wahl von zugehörigen Zwischenwerten } \{\xi_n\}_n \\ \text{konvergiert die Folge } \{S_{Z_n}(f, \xi_n)\}_n. \end{array}$$

Ist f integrierbar, so gilt unabhängig von Zerlegungsfolge und Zwischenwerten:

$$\int_a^b f(x) dx = \lim_{n \rightarrow \infty} S_{Z_n}(f, \xi_n). \quad (9.2.1)$$

Beweis. ▶ „ \Rightarrow “: Seien $\{Z_n\}_n$ und $\{\xi_n\}_n$ wie vorausgesetzt fixiert; insbesondere gilt also $\Delta(Z_n) \rightarrow 0$ ($n \rightarrow \infty$). Zu beliebigem $\varepsilon > 0$ existiert dann nach Satz 9.2.3 ein $N = N(\varepsilon) \in \mathbb{N}$ mit

$$\overline{S}_{Z_n}(f) - \underline{S}_{Z_n}(f) < \varepsilon \quad \text{für alle } n \geq N.$$

Wegen $\underline{S}_{Z_n}(f) \leq \mathcal{I}(f) \leq \overline{S}_{Z_n}(f)$ und $\underline{S}_{Z_n}(f) \leq S_{Z_n}(f, \xi_n) \leq \overline{S}_{Z_n}(f)$ für alle $n \in \mathbb{N}$ folgt sofort

$$|\mathcal{I}(f) - S_{Z_n}(f, \xi_n)| \leq \overline{S}_{Z_n}(f) - \underline{S}_{Z_n}(f) < \varepsilon \quad \text{für alle } n \geq N,$$

also die Behauptung inklusive (9.2.1).

▶ „ \Leftarrow “: In den Ergänzungen. □

Wir fassen nun einige wichtige Eigenschaften reellwertiger integrierbarer Funktionen zusammen; für komplex- und vektorwertiger Funktionen siehe Bemerkung 9.2.7 (a) unten.

Satz 9.2.6: Rechenregeln

(a) Aus $f, g \in \mathcal{R}(I)$ folgt $\alpha f + \beta g \in \mathcal{R}(I)$ für alle $\alpha, \beta \in \mathbb{R}$, und es gilt

$$\mathcal{I}(\alpha f + \beta g) = \alpha \mathcal{I}(f) + \beta \mathcal{I}(g).$$

D.h. $\mathcal{R}(I)$ ist ein reeller Vektorraum.

(b) Sind $f, g \in \mathcal{R}(I)$ und gilt $f \leq g$ auf I , so folgt

$$\mathcal{I}(f) \leq \mathcal{I}(g).$$

(c) Mit $f \in \mathcal{R}(I)$ gilt auch $|f| \in \mathcal{R}(I)$, und wir haben

$$|\mathcal{I}(f)| \leq \mathcal{I}(|f|).$$

(d) Aus $f, g \in \mathcal{R}(I)$ folgt $f \cdot g \in \mathcal{R}(I)$, und es gilt

$$\mathcal{I}(|fg|) \leq \left(\sup_I |g| \right) \mathcal{I}(|f|).$$

(e) Gilt $f, g \in \mathcal{R}(I)$ sowie $|g| \geq c > 0$ auf I mit einer Konstante $c > 0$, so folgt auch $\frac{f}{g} \in \mathcal{R}(I)$, und wir haben

$$\mathcal{I}\left(\left|\frac{f}{g}\right|\right) \leq \frac{1}{c} \mathcal{I}(|f|).$$

Bemerkung 9.2.7. (a) Die Aussagen (a), (c) und (d) bleiben für komplex- und vektorwertige Funktionen richtig, wobei für $f, g \in \mathcal{R}(I, \mathbb{R}^d)$ in (c) der Betrag durch die Norm $\|\cdot\|$ und in (d) das Produkt durch das Skalarprodukt $\langle \cdot, \cdot \rangle$ zu ersetzen ist. Aussage (e) bleibt noch für komplexwertige Funktionen richtig (\rightarrow Übungsaufgaben).

(b) Zum Beweis einiger Aussagen benutzen wir folgende allgemeine Beobachtung (\rightarrow Übungsaufgabe): Ist $h : D \rightarrow \mathbb{R}$ beschränkt, $D \subset \mathbb{R}^m$, so gilt

$$\sup_{x, x' \in D} |h(x) - h(x')| = \sup_D h - \inf_D h. \quad (9.2.2)$$

Man nennt diesen Wert die *Oszillation* von h .

Beweis von Satz 9.2.6. (a) Ist $\{\mathcal{Z}_n\}_n$ eine beliebige ausgezeichnete Zerlegungsfolge und $\{\xi_n\}_n$ eine Folge zugehöriger Zwischenwerte, so gilt für die RIEMANNschen Zwischensummen:

$$S_{\mathcal{Z}_n}(\alpha f + \beta g, \xi_n) = \alpha S_{\mathcal{Z}_n}(f, \xi_n) + \beta S_{\mathcal{Z}_n}(g, \xi_n), \quad n \in \mathbb{N}.$$

Wenden wir Satz 9.2.5 auf f und g an, so folgt die Existenz des Grenzwerts

$$\lim_{n \rightarrow \infty} S_{\mathcal{Z}_n}(\alpha f + \beta g, \xi_n) = \alpha \lim_{n \rightarrow \infty} S_{\mathcal{Z}_n}(f, \xi_n) + \beta \lim_{n \rightarrow \infty} S_{\mathcal{Z}_n}(g, \xi_n) = \alpha \mathcal{I}(f) + \beta \mathcal{I}(g).$$

Wiederum Satz 9.2.5, nun angewendet auf $\alpha f + \beta g$, liefert somit $\alpha f + \beta g \in \mathcal{R}(I)$ und die behauptete Linearität des Integrals.

- (b) Nach (a) gilt $h := g - f \in \mathcal{R}(I)$. Ist \mathcal{Z} eine beliebige Zerlegung von I , so liefert Formel (9.1.3) wegen $h \geq 0$

$$0 \leq \underline{S}_{\mathcal{Z}}(h) \leq \mathcal{I}(h) \stackrel{(a)}{=} \mathcal{I}(g) - \mathcal{I}(f),$$

also die Behauptung.

- (c) Die umgekehrte Dreiecksungleichung liefert

$$||f(x)| - |f(x')|| \leq |f(x) - f(x')| \quad \text{für alle } x, x' \in I.$$

Mit Formel (9.2.2) und Satz 9.2.1 schätzen wir ab

$$\begin{aligned} \overline{S}_{\mathcal{Z}}(|f|) - \underline{S}_{\mathcal{Z}}(|f|) &= \sum_j \left(\sup_{I_j} |f| - \inf_{I_j} |f| \right) \Delta x_j \stackrel{(9.2.2)}{=} \sum_j \left(\sup_{x, x' \in I_j} ||f(x)| - |f(x')|| \right) \Delta x_j \\ &\leq \sum_j \left(\sup_{x, x' \in I_j} |f(x) - f(x')| \right) \Delta x_j \stackrel{(9.2.2)}{=} \sum_j \left(\sup_{I_j} f - \inf_{I_j} f \right) \Delta x_j \\ &= \overline{S}_{\mathcal{Z}}(f) - \underline{S}_{\mathcal{Z}}(f) < \varepsilon \end{aligned}$$

für eine geeignete Zerlegung \mathcal{Z} von I . Wiederum Satz 9.2.1 liefert also $|f| \in \mathcal{R}(I)$. Ferner entnehmen wir (b):

$$\mathcal{I}(f) \leq \mathcal{I}(|f|), \quad -\mathcal{I}(f) \stackrel{(a)}{=} \mathcal{I}(-f) \leq \mathcal{I}(|f|)$$

bzw. $|\mathcal{I}(f)| \leq \mathcal{I}(|f|)$.

- (d), (e) \rightarrow Ergänzungen. □

Bemerkung 9.2.8. Zu einem kompakten Intervall $I \subset \mathbb{R}$ setzen wir $V := \mathcal{R}(I)$ und

$$\langle f, g \rangle := \int_I f(x)g(x) dx, \quad f, g \in V.$$

Gemäß Satz 9.2.6 (d) ist dann $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{R}$ wohl definiert, und wegen Satz 9.2.6 (a) ist V ein Vektorraum und $\langle \cdot, \cdot \rangle$ eine symmetrische Bilinearform. Wegen Satz 9.2.6 (b) ist $\langle \cdot, \cdot \rangle$ positiv semidefinit, d.h. $\langle f, f \rangle \geq 0$ für alle $f \in \mathcal{R}(I)$; auf dem Unterraum $C^0(I) \subset \mathcal{R}(I)$ - siehe Satz 9.2.10 - ist $\langle \cdot, \cdot \rangle$ sogar positiv definit (\rightarrow Übungen). Da die Cauchy-Schwarzsche Ungleichung für positiv semidefinite Bilinearformen richtig bleibt (vgl. Satz 5.1.5 und Bemerkung 5.1.6), erhalten wir den

Satz 9.2.9: SCHWARZsche Ungleichung

Ist I ein kompaktes Intervall, so gilt für beliebige $f, g \in \mathcal{R}(I)$:

$$\left| \int_I f(x)g(x) dx \right| \leq \left(\int_I |f(x)|^2 dx \right)^{\frac{1}{2}} \left(\int_I |g(x)|^2 dx \right)^{\frac{1}{2}}.$$

Eine ganze Klasse integrierbarer Funktionen erhalten wir mit dem

Satz 9.2.10

Für beliebige kompakte Intervalle $I = [a, b]$ gilt $C^0(I) \subset \mathcal{R}(I)$.

Beweis. Nach dem Satz von Heine, Satz 6.4.8, ist jede Funktion $f \in C^0([a, b])$ gleichmäßig stetig. Zu beliebigem $\varepsilon > 0$ existiert also ein $\delta = \delta(\varepsilon) > 0$, so dass gilt

$$|f(x) - f(x')| < \frac{\varepsilon}{b-a} \quad \text{für alle } x, x' \in I \text{ mit } |x - x'| < \delta.$$

Ist \mathcal{Z} eine beliebige Zerlegung von I mit $\Delta(\mathcal{Z}) < \delta$, so folgt

$$\begin{aligned} \bar{S}_{\mathcal{Z}}(f) - \underline{S}_{\mathcal{Z}}(f) &= \sum_j \left[\sup_{I_j} f - \inf_{I_j} f \right] \Delta x_j \\ &\stackrel{(9.2.2)}{=} \left(\sum_j \sup_{x, x' \in I_j} |f(x) - f(x')| \right) \Delta x_j \\ &\leq \frac{\varepsilon}{b-a} \sum_j \Delta x_j = \varepsilon \end{aligned}$$

und somit $f \in \mathcal{R}(I)$ nach Satz 9.2.3. □

Mit dem Integrabilitätskriterium III kann man nun direkt Integrale (einfacher) stetiger Funktionen ausrechnen.

Beispiel 9.2.11. Seien $I = [a, b]$, $f = \text{Id} : I \rightarrow \mathbb{R}$. Nach Satz 9.2.9 ist f integrierbar. Zu $n \in \mathbb{N}$ wählen wir die äquidistante Zerlegung $\mathcal{Z}_n = \{x_{0n}, \dots, x_{nn}\}$ mit

$$x_{jn} := a + \frac{j}{n}(b-a), \quad j = 0, \dots, n.$$

Mit den Zwischenwerten $\xi_n = (\xi_{1n}, \dots, \xi_{nn}) = (x_{1n}, \dots, x_{nn})$ liefert dann Satz 9.2.5:

$$\begin{aligned} \int_a^b x \, dx &= \lim_{n \rightarrow \infty} S_{\mathcal{Z}_n}(\text{Id}, \xi_n) = \lim_{n \rightarrow \infty} \left(\sum_{j=1}^n \xi_{jn} \cdot \Delta x_{jn} \right) \\ &= \lim_{n \rightarrow \infty} \sum_{j=1}^n \left(a + \frac{j}{n}(b-a) \right) \frac{b-a}{n} \\ &= \lim_{n \rightarrow \infty} \left[a(b-a) + \frac{(b-a)^2}{n^2} \frac{n(n+1)}{2} \right] \\ &= a(b-a) + \frac{(b-a)^2}{2} = \frac{1}{2}(b^2 - a^2). \end{aligned}$$

Wir schließen mit dem

Satz 9.2.12: Mittelwertsatz der Integralrechnung

Es sei $I = [a, b]$ und $f \in C^0(I)$ sowie $p \in \mathcal{R}(I)$ mit $p \geq 0$ auf I seien gegeben. Dann existiert ein $\xi \in (a, b)$, so dass gilt

$$\int_a^b f(x)p(x) \, dx = f(\xi) \int_a^b p(x) \, dx.$$

Beweis. → Ergänzungen. □

Bemerkung 9.2.13. Speziell für $p(x) := 1, x \in [a, b]$, haben wir $p \in \mathcal{R}([a, b])$ und z.B. nach Satz 9.2.5:

$$\int_a^b 1 \, dx = \lim_{n \rightarrow \infty} S_{\mathcal{Z}_n}(1) = b - a$$

für eine ausgezeichnete Zerlegungsfolge $\{\mathcal{Z}_n\}_n$ und beliebige Zwischenwerte. Satz 9.2.12 liefert also

$$\int_a^b f(x) \, dx := \frac{1}{b-a} \int_a^b f(x) \, dx = f(\xi).$$

Die Größe $\int_a^b f(x) \, dx$ heißt *Mittelwert* von f über I und gibt die „mittlere Höhe“ von f an.

9.3. Die Additivität des Integrals

Eine wichtige, geometrisch naheliegende Eigenschaft des Integrals ist, dass man sich bei der Berechnung auf Integrale über Teilintervalle zurückziehen kann, die man im Anschluss aufsummiert; siehe Satz 9.3.2 unten. Zum Beweis starten wir mit dem einfachen

Hilfssatz 9.3.1

Ist $I = [a, b]$ ein kompaktes Intervall und $f \in \mathcal{R}(I)$, so gilt auch $f|_{I'} \in \mathcal{R}(I')$ für jedes kompakte Teilintervall $I' \subset I$; wir schreiben dann $f \in \mathcal{R}(I')$ und

$$\int_{I'} f \, dx := \int_{I'} f|_{I'} \, dx.$$

Beweis. Wegen $f \in \mathcal{R}(I)$ existiert nach Satz 9.2.1 zu jedem $\varepsilon > 0$ eine Zerlegung \mathcal{Z} von I mit $\bar{S}_{\mathcal{Z}}(f) - \underline{S}_{\mathcal{Z}}(f) < \varepsilon$. Für die Verfeinerung \mathcal{Z}^* von \mathcal{Z} , die zusätzlich die beiden Endpunkte von I' enthält, gilt dann nach Hilfssatz 9.1.6(a):

$$\bar{S}_{\mathcal{Z}^*}(f) - \underline{S}_{\mathcal{Z}^*}(f) \leq \bar{S}_{\mathcal{Z}}(f) - \underline{S}_{\mathcal{Z}}(f) < \varepsilon.$$

Die Teilzerlegung \mathcal{Z}' von \mathcal{Z}^* , die nur die in I' liegenden Teilpunkte enthält, ist dann offenbar Zerlegung von I' und es gilt

$$\bar{S}_{\mathcal{Z}'}(f) - \underline{S}_{\mathcal{Z}'}(f) \leq \bar{S}_{\mathcal{Z}^*}(f) - \underline{S}_{\mathcal{Z}^*}(f) < \varepsilon,$$

so dass $f \in \mathcal{R}(I')$ wiederum aus Satz 9.2.1 folgt. \square

Satz 9.3.2: Additivität des Integrals

Es sei $I = [a, b]$ in endlich viele abgeschlossene Teilintervalle I_1, \dots, I_μ zerlegt, die höchstens Endpunkte gemein haben, d.h.

$$I = I_1 \cup \dots \cup I_\mu, \quad \text{int } I_j \cap \text{int } I_k = \emptyset \text{ für } j \neq k.$$

(a) Dann gilt für beliebige $f \in \mathcal{R}(I)$ die Relation

$$\int_I f(x) \, dx = \sum_{j=1}^{\mu} \int_{I_j} f(x) \, dx. \quad (9.3.1)$$

(b) Ist $f \in \mathcal{R}(I_j)$ für alle $j = 1, \dots, \mu$, so folgt $f \in \mathcal{R}(I)$ und es gilt wieder (9.3.1).

Beweis. (a) Nach Hilfssatz 9.3.1 gilt zunächst $f \in \mathcal{R}(I_j)$ für $j = 1, \dots, \mu$, so dass alle Integrale in (9.3.1) erklärt sind. Es sei nun $\{\mathcal{Z}_n\}_n$ eine ausgezeichnete Zerlegungsfolge von I , deren Glieder die Endpunkte aller Teilintervalle I_1, \dots, I_μ als Teilpunkte enthalten, und $S_{\mathcal{Z}_n}(f)$ seien zugehörige RIEMANNSche Zwischensummen. Schreiben wir dann

$S_{\mathcal{Z}_n}^{(j)}(f)$ für die Zwischensummen, die nur die in I_j liegenden Teilpunkte von \mathcal{Z}_n enthalten, so gilt offenbar

$$S_{\mathcal{Z}_n}(f) = \sum_{j=1}^{\mu} S_{\mathcal{Z}_n}^{(j)}(f) \quad \text{für alle } n \in \mathbb{N}.$$

Satz 9.2.5 liefert nach Grenzübergang $n \rightarrow \infty$ die Behauptung.

(b) Wegen $f \in \mathcal{R}(I_j)$, $j = 1, \dots, \mu$, gibt es nach Satz 9.2.1 Zerlegungen \mathcal{Z}_j von I_j mit

$$\bar{S}_{\mathcal{Z}_j}(f) - \underline{S}_{\mathcal{Z}_j}(f) < \frac{\varepsilon}{\mu}, \quad j = 1, \dots, \mu.$$

Bezeichnet \mathcal{Z} die Zerlegung von I , die alle Teilpunkte von $\mathcal{Z}_1, \dots, \mathcal{Z}_\mu$ enthält, so folgt demnach

$$\bar{S}_{\mathcal{Z}}(f) - \underline{S}_{\mathcal{Z}}(f) = \sum_{j=1}^{\mu} [\bar{S}_{\mathcal{Z}_j}(f) - \underline{S}_{\mathcal{Z}_j}(f)] < \varepsilon,$$

also $f \in \mathcal{R}(I)$, wiederum nach Satz 9.2.1. Die Relation (9.3.1) folgt nun aus (a). \square

Wir wollen die Aussage von Satz 9.3.2 noch in eine etwas andere Fassung bringen; dazu benötigen wir die praktische

Definition 9.3.3

Ist $f \in \mathcal{R}(I)$, $I = [a, b]$, und seien $\alpha, \beta \in I$ mit $\alpha < \beta$ gewählt. Dann setzen wir

$$\int_{\alpha}^{\alpha} f(x) dx = 0 \quad \text{und} \quad \int_{\beta}^{\alpha} f(x) dx := - \int_{\alpha}^{\beta} f(x) dx.$$

Satz 9.3.4

Ist $f \in \mathcal{R}(I)$ und sind $\alpha, \beta, \gamma \in I$ beliebig gewählt, so folgt

$$\int_{\alpha}^{\beta} f(x) dx + \int_{\beta}^{\gamma} f(x) dx = \int_{\alpha}^{\gamma} f(x) dx. \quad (9.3.2)$$

Beweis. \rightarrow Ergänzungen \square

Wir schließen mit der Betrachtung einer naheliegenden Verallgemeinerung der Klasse stetiger Funktionen:

Definition 9.3.5

Eine Funktion $f : I \rightarrow \mathbb{R}$ heißt *stückweise stetig* auf dem kompakten Intervall I , wenn es eine Zerlegung $\mathcal{Z} = \{\xi_0, \dots, \xi_\mu\}$, $\mu \in \mathbb{N}$, von I so gibt, dass gilt:

- ▶ Die Einschränkungen $f|_{(\xi_{j-1}, \xi_j)}$ sind stetig für alle $j = 1, \dots, \mu$.
- ▶ Alle einseitigen Grenzwerte in den Punkten ξ_0, \dots, ξ_μ existieren.

Die Klasse aller stückweise stetigen Funktionen auf $[a, b]$ bezeichnen wir mit $PC([a, b])$.

Offenbar ist $C^0(I) \subset PC(I)$ richtig. Der folgende Satz liefert also eine Verschärfung von Satz 9.2.10:

Satz 9.3.6

Für jedes kompakte Intervall $I = [a, b]$ gilt $PC(I) \subset \mathcal{R}(I)$.

Für den Beweis benötigen wir noch die folgende Beobachtung:

Hilfssatz 9.3.7

Seien $I \subset \mathbb{R}$ ein kompaktes Intervall, $\Pi \subset I$ eine endliche Punktmenge und $f \in \mathcal{R}(I)$ gegeben. Ist $g : I \rightarrow \mathbb{R}$ eine weitere Funktion und gilt

$$f = g \quad \text{auf } I \setminus \Pi,$$

so folgt $g \in \mathcal{R}(I)$ sowie $\int_I g \, dx = \int_I f \, dx$.

Beweis. → Ergänzungen

□

Beweis von Satz 9.3.6. Sei $f \in PC(I)$ wie in Definition 9.3.5 gegeben. Wir setzen $I_j := [\xi_{j-1}, \xi_j]$ für $j = 1, \dots, \mu$. Dann stimmt $f|_{I_j}$ in $(\xi_{j-1}, \xi_j) = I_j \setminus \{\xi_{j-1}, \xi_j\}$ mit der auf I_j stetigen Funktion

$$g_j(x) := \begin{cases} \lim_{x \rightarrow \xi_{j-1}^+} f(x), & x = \xi_{j-1} \\ f(x), & \xi_{j-1} < x < \xi_j \\ \lim_{x \rightarrow \xi_j^-} f(x), & x = \xi_j \end{cases}, \quad j = 1, \dots, \mu,$$

überein. Nach Satz 9.2.10 gilt $g_j \in \mathcal{R}(I_j)$ und Hilfssatz 9.3.7 liefert $f \in \mathcal{R}(I_j)$ für $j = 1, \dots, \mu$. Satz 9.3.2 (b) entnehmen wir schließlich die Behauptung $f \in \mathcal{R}(I)$. □

9.4. Der Fundamentalsatz der Differential- und Integralrechnung

Abgesehen von der Linearität liefern die Rechenregeln für das Integral, Satz 9.2.6, keine Möglichkeiten zur Bestimmung von Integralen von Verknüpfungen von Funktionen; das unterscheidet die Integration grundlegend von der Differentiation und macht die Bestimmung von Integralen in der Praxis oft komplizierter als das Bestimmen von Ableitungen. Auch Satz 9.2.5 ist nur sehr eingeschränkt zur expliziten Berechnung von Integralen geeignet; siehe Beispiel 9.2.11.

Als zentrales Hilfsmittel wird sich stattdessen der sogenannte *Fundamentalsatz* oder *Hauptsatz der Differential- und Integralrechnung* und seine Anwendungen erweisen, die wir in den folgenden Abschnitten 9.4, 9.5 behandeln wollen. Sie stellen - wie der Name bereits andeutet - einen Zusammenhang zwischen den Ableitungen und Integralen von Funktionen her. Wir beginnen mit der

Definition 9.4.1

Seien ein *beliebiges* Intervall $I \subset \mathbb{R}$ und eine Funktion $f : I \rightarrow \mathbb{R}$ gegeben. Dann heißt eine differenzierbare Funktion $F : I \rightarrow \mathbb{R}$ *Stammfunktion* zu f , falls gilt

$$F'(x) = f(x) \quad \text{für alle } x \in I.$$

Satz 9.4.2

Seien $I \subset \mathbb{R}$ ein Intervall, $p \in I$ beliebig und $f \in C^0(I)$. Dann ist

$$F(x) := \int_p^x f(t) dt, \quad x \in I,$$

eine Stammfunktion zu f , und es gilt $F \in C^1(I)$.

Beweis. Für $x \in I$ und $h \neq 0$ mit $x + h \in I$ gilt

$$\begin{aligned}
 \frac{F(x+h) - F(x)}{h} &= \frac{1}{h} \left(\int_p^{x+h} f(t) dt - \int_p^x f(t) dt \right) \\
 &\stackrel{\text{Def. 9.3.3}}{=} \frac{1}{h} \left(\int_p^{x+h} f(t) dt + \int_x^p f(t) dt \right) \\
 &\stackrel{(9.3.2)}{=} \frac{1}{h} \int_x^{x+h} f(t) dt.
 \end{aligned} \tag{9.4.1}$$

Nach dem Mittelwertsatz der Integralrechnung existiert ein $\xi_h \in I$ zwischen x und $x + h$ mit

$$\frac{1}{h} \int_x^{x+h} f(t) dt = f(\xi_h). \tag{9.4.2}$$

Wegen $|\xi_h - x| \leq |h| \rightarrow 0$ ($h \rightarrow 0$) und der Stetigkeit von f haben wir

$$\lim_{h \rightarrow 0} f(\xi_h) = f\left(\lim_{h \rightarrow 0} \xi_h\right) = f(x).$$

Zusammen mit (9.4.1) und (9.4.2) finden wir also

$$\lim_{h \rightarrow 0} \frac{F(x+h) - F(x)}{h} = \lim_{h \rightarrow 0} f(\xi_h) = f(x), \quad x \in I,$$

und somit $F' \equiv f \in C^0(I)$. □

Bemerkung 9.4.3. Lässt man in Satz 9.4.2 die Forderung der Stetigkeit für f fallen, so ist F i.A. keine Stammfunktion von f . Z.B. erhält man für $f = \text{sgn}$ und $p = 0$ die Funktion

$$F(x) = \int_0^x \text{sgn}(t) dt = |x|, \quad x \in \mathbb{R},$$

die in $x = 0$ nicht differenzierbar ist. Wir beschränken unsere Betrachtungen daher i.F. auf stetige Funktionen $f : I \rightarrow \mathbb{R}$.

Satz 9.4.4

Seien $I \subset \mathbb{R}$ ein Intervall, $f \in C^0(I)$ eine Funktion und $F \in C^1(I)$ eine beliebige Stammfunktion von f . Dann ist eine Funktion $G \in C^1(I)$ genau dann Stammfunktion von f , wenn $G - F = \text{const}$ auf I gilt.

Beweis. Wir beachten

$$\begin{aligned} G - F \equiv \text{const auf } I &\stackrel{\text{Folg. 7.3.11 (c)}}{\Leftrightarrow} G' - F' = (G - F)' \equiv 0 \text{ auf } I \\ &\Leftrightarrow G' = F' = f \text{ auf } I, \end{aligned}$$

wie behauptet. □

Bemerkung 9.4.5. Ist $F \in C^1(I)$ eine Stammfunktion von $f \in C^0(I)$ (z.B. die in Satz 9.4.2 erklärte), so ist also die Menge aller Stammfunktionen gegeben durch

$$\{G \in C^1(I) : G = F + c, c \in \mathbb{R}\}.$$

Diese Menge wird *unbestimmtes Integral* von f genannt und wir schreiben

$$\int f(x) dx := \{F + c : c \in \mathbb{R}\}$$

oder, wie allgemein gebräuchlich,

$$\int f(x) dx = F(x) + c.$$

Zur Unterscheidung heißt daher $\int_a^b f(x) dx$ auch *bestimmtes Integral* von $f \in \mathcal{R}([a, b])$ (zwischen den Grenzen a und b).

Satz 9.4.6: Fundamentalsatz der Differential- und Integralrechnung

Seien $I \subset \mathbb{R}$ ein Intervall und $F \in C^1(I)$ eine beliebige Stammfunktion von $f \in C^0(I)$. Für beliebige $\alpha, \beta \in I$ gilt dann

$$\int_{\alpha}^{\beta} f(x) dx = F(\beta) - F(\alpha) =: F(x) \Big|_{\alpha}^{\beta}.$$

Beweis. Nach Satz 9.4.2 ist $F_0(x) := \int_{\alpha}^x f(t) dt, x \in I$, Stammfunktion von f , und nach Satz 9.4.4 gilt $F = F_0 + c$ auf I mit einer Konstanten $c \in \mathbb{R}$. Wegen $F_0(\alpha) = 0$, siehe Definition 9.3.3, folgt also

$$\int_{\alpha}^{\beta} f(x) dx = F_0(\beta) = F_0(\beta) - F_0(\alpha) = F(\beta) - F(\alpha),$$

wie behauptet. □

Beispiel 9.4.7. (a) Für $\mu \in \mathbb{R} \setminus \{-1\}$ gilt

$$\int x^\mu dx = \frac{x^{\mu+1}}{\mu+1} + c, \quad x > 0,$$

denn $\frac{d}{dx}(\frac{x^{\mu+1}}{\mu+1}) = x^\mu$ nach (8.2.8). D.h. $F(x) = \frac{x^{\mu+1}}{\mu+1}$ ist Stammfunktion von $f(x) = x^\mu$ auf $(0, +\infty)$. Satz 9.4.6 liefert also

$$\int_\alpha^\beta x^\mu dx = \frac{x^{\mu+1}}{\mu+1} \Big|_\alpha^\beta = \frac{1}{\mu+1} (\beta^{\mu+1} - \alpha^{\mu+1})$$

für beliebige $\alpha, \beta > 0$.

(b) Wir geben einige weitere Beispiele von Stammfunktionen an, die sich sofort aus den Ableitungen elementarer Funktionen ergeben (vergleiche Formeln (8.2.3), (8.2.5), (8.3.1), (8.5.1) und (8.5.3)):

$$\begin{aligned} \int \exp x dx &= \exp x + c, \quad x \in \mathbb{R}, \\ \int \frac{1}{x} dx &= \log x + c, \quad x > 0, \\ \int \cos x dx &= \sin x + c, \quad x \in \mathbb{R}, \\ \int \sin x dx &= -\cos x + c, \quad x \in \mathbb{R}, \\ \int \frac{dx}{\cos^2 x} &= \tan x + c, \quad x \in \left(-\frac{\pi}{2} + k\pi, \frac{\pi}{2} + k\pi\right), \quad k \in \mathbb{Z}, \\ \int \frac{dx}{\sqrt{1-x^2}} &= \arcsin x + c, \quad x \in (-1, 1), \\ \int \frac{dx}{1+x^2} &= \arctan x + c, \quad x \in \mathbb{R}. \end{aligned}$$

Bemerkung 9.4.8. Die Definition der Stammfunktion und des unbestimmten Integrals sowie die Aussagen der Sätze 9.4.2, 9.4.4 und 9.4.6 übertragen sich sofort auf vektorwertige Funktionen $f \in C^0(I, \mathbb{R}^d)$, in dem man komponentenweise argumentiert (\rightarrow Nachprüfen).

9.5. Drei Anwendungen des Fundamentalsatzes

In diesem Abschnitt wollen wir einerseits zwei wichtige Rechenregeln aus dem Fundamentalsatz ableiten, andererseits wenden wir diesen auch auf Funktionenfolgen an um den offenen Beweis von Satz 8.1.1 nachzuliefern.

Satz 9.5.1: Partielle Integration

Ist $I \subset \mathbb{R}$ ein Intervall und sind $f, g \in C^1(I, \mathbb{R}^d)$ gegeben, so gilt

$$\int_a^b \langle f'(x), g(x) \rangle dx = \langle f(x), g(x) \rangle \Big|_a^b - \int_a^b \langle f(x), g'(x) \rangle dx, \quad a, b \in I. \quad (9.5.1)$$

Beweis. Die Produktregel (7.2.1) und Satz 9.4.6 liefern sofort

$$\begin{aligned} \langle f(x), g(x) \rangle \Big|_a^b &= \int_a^b \frac{d}{dx} \langle f(x), g(x) \rangle dx \\ &= \int_a^b \langle f'(x), g(x) \rangle dx + \int_a^b \langle f(x), g'(x) \rangle dx \end{aligned}$$

und nach Umstellen die Behauptung (9.5.1). \square

Bemerkung 9.5.2. Die zuweilen nützliche „unbestimmte Version“ von (9.5.1) ist

$$\int \langle f'(x), g(x) \rangle dx = \langle f(x), g(x) \rangle - \int \langle f(x), g'(x) \rangle dx. \quad (9.5.2)$$

Diese gewinnt man wieder sofort aus der Produktregel (7.2.1) und der Definition des unbestimmten Integrals.

Beispiel 9.5.3. (a) $\int_0^{\frac{\pi}{2}} \sin^2 x dx = \int_0^{\frac{\pi}{2}} \cos^2 x dx = \frac{\pi}{4}$. Denn Satz 9.5.1 liefert

$$\begin{aligned} \int_0^{\frac{\pi}{2}} \sin^2 x dx &= [-\cos x \cdot \sin x] \Big|_0^{\frac{\pi}{2}} + \int_0^{\frac{\pi}{2}} \cos^2 x dx = \int_0^{\frac{\pi}{2}} \cos^2 x dx \\ &= \int_0^{\frac{\pi}{2}} (1 - \sin^2 x) dx = \int_0^{\frac{\pi}{2}} dx - \int_0^{\frac{\pi}{2}} \sin^2 x dx \end{aligned}$$

bzw.

$$\int_0^{\frac{\pi}{2}} \sin^2 x dx = \frac{1}{2} \int_0^{\frac{\pi}{2}} dx = \frac{1}{2} x \Big|_0^{\frac{\pi}{2}} = \frac{\pi}{4}.$$

(b) Für $x > 0$ haben wir gemäß (9.5.2):

$$\int \log x dx = \int 1 \cdot \log x dx = x \log x - \int x \cdot \frac{1}{x} dx = x \log x - x + \text{const},$$

also

$$\int_1^a \log x dx = a(\log a - 1) + 1 \quad \text{für } a > 0.$$

Satz 9.5.4: Substitutionsregel oder Transformationsformel

Es seien $I, I^* \subset \mathbb{R}$ zwei Intervalle, und $f \in C^0(I, \mathbb{R}^d)$ sowie $\varphi \in C^1(I^*)$ mit $\varphi(I^*) \subset I$ seien gegeben. Dann gilt für beliebige $\alpha, \beta \in I^*$:

$$\int_{\varphi(\alpha)}^{\varphi(\beta)} f(x) dx = \int_{\alpha}^{\beta} f(\varphi(t)) \varphi'(t) dt. \quad (9.5.3)$$

Beweis. Es sei $F \in C^1(I, \mathbb{R}^d)$ eine Stammfunktion von f . Für $g := F \circ \varphi$ gilt dann $g \in C^1(I^*, \mathbb{R}^d)$ und der Kettenregel, Satz 7.2.4, entnehmen wir

$$g'(t) = F'(\varphi(t)) \varphi'(t) = f(\varphi(t)) \varphi'(t), \quad t \in I^*.$$

Der Fundamentalsatz, Satz 9.4.6, liefert also für beliebige $\alpha, \beta \in I^*$:

$$\begin{aligned} \int_{\alpha}^{\beta} f(\varphi(t)) \varphi'(t) dt &= \int_{\alpha}^{\beta} g'(t) dt = g(\beta) - g(\alpha) \\ &= F(\varphi(\beta)) - F(\varphi(\alpha)) = \int_{\varphi(\alpha)}^{\varphi(\beta)} f(x) dx, \end{aligned}$$

wie behauptet. □

Beispiel 9.5.5. (a) Zu berechnen sei $\int_0^r \sqrt{r^2 - x^2} dx$, $r > 0$.

Wir betrachten die Transformation $x = \varphi(t) = r \sin t$, $t \in [0, \frac{\pi}{2}]$. Dann ist $\varphi(0) = 0$, $\varphi(\frac{\pi}{2}) = r$ und $\varphi'(t) = r \cos t$. Formel (9.5.3) und Beispiel 9.5.3 (a) liefern also

$$\int_0^r \sqrt{r^2 - x^2} dx = \int_0^{\frac{\pi}{2}} \sqrt{r^2 - r^2 \sin^2 t} \cdot r \cos t dt = r^2 \int_0^{\frac{\pi}{2}} \cos^2 t dt = \frac{\pi r^2}{4}.$$

(\rightarrow Flächeninhalt der Kreisscheibe vom Radius $r > 0$ ist $4 \cdot \frac{\pi r^2}{4} = \pi r^2$).

(b) Zu berechnen sei $\int_0^1 (1+t^2)^\alpha t dt$, $\alpha \in \mathbb{R} \setminus \{-1\}$.

Wir beachten

$$h(t) := (1+t^2)^\alpha t = \frac{1}{2} (1+t^2)^\alpha \frac{d}{dt} (1+t^2), \quad t \in \mathbb{R}.$$

Mit $f(x) := x^\alpha$, $x > 0$, und $\varphi(t) := 1+t^2$, $t \in \mathbb{R}$, haben wir also

$$h(t) = \frac{1}{2} f(\varphi(t)) \varphi'(t), \quad t \in \mathbb{R}.$$

Beachten wir noch $\varphi(0) = 1$, $\varphi(1) = 2$, so folgt

$$\int_0^1 (1+t^2)^\alpha t \, dt \stackrel{(9.5.3)}{=} \frac{1}{2} \int_1^2 x^\alpha \, dx = \frac{1}{2} \frac{x^{\alpha+1}}{\alpha+1} \Big|_1^2 = \frac{2^{\alpha+1} - 1}{2(\alpha+1)}, \quad \alpha \neq -1.$$

Bemerkung 9.5.6. (a) Die unbestimmte Form der Substitutionsregel (9.5.3), nämlich

$$\int f(x) \, dx = \int f(\varphi(t)) \varphi'(t) \, dt \quad (9.5.4)$$

ist häufig ebenfalls hilfreich. Mit Stammfunktionen F von f und Ψ von $(f \circ \varphi) \varphi'$ bedeutet (9.5.4) gerade $F \circ \varphi = \Psi + c$. Genau wie die bestimmte Form (9.5.3) (siehe Beispiel 9.5.5) hat auch (9.5.4) zwei Lesarten, nämlich:

- ▶ „von rechts nach links“: Eine Stammfunktion von $(f \circ \varphi) \varphi'$ lässt sich sofort als $F \circ \varphi$ ablesen.
- ▶ „von links nach rechts“: Möchte man eine Stammfunktion F von f mittels (9.5.4) bestimmen, so muss φ bijektiv sein: Kennt man Ψ , so finden wir dann $F = \Psi \circ \varphi^{-1} + c$.

Beispiel.

$$\begin{aligned} \int \frac{dx}{r^2 + x^2} &\stackrel{x=\varphi(t)=rt}{=} \int \frac{r \, dt}{r^2 + r^2 t^2} = \frac{1}{r} \int \frac{dt}{1+t^2} \\ &\stackrel{(8.5.3)}{=} \frac{1}{r} \arctan t + \text{const} \stackrel{\varphi^{-1}(x)=\frac{x}{r}}{=} \frac{1}{r} \arctan \frac{x}{r} + \text{const}. \end{aligned}$$

Sowohl φ als auch $\psi := \varphi^{-1}$ werden *Substitution* oder *Transformation* genannt.

- (b) Es gibt eine Vielzahl von Kunstgriffen zur Bestimmung von Integralen bzw. Stammfunktionen, auf die wir nicht im Einzelnen eingehen können. Wir verweisen auf [Hil1] Abschnitt 3.10 und [Sa] Kap. III, § 9 und Kap. V, § 1 für einige Beispiele, insbesondere die Integration rationaler Funktionen (\rightarrow *Partialbruchzerlegung*).

Abschließend betrachten wir wieder Funktionenfolgen. Zunächst soll untersucht werden unter welchen Voraussetzungen Integration und Grenzwertbildung vertauscht werden können:

Satz 9.5.7

Sei $\{f_n\}_n$ eine Folge von Funktionen $f_n \in \mathcal{R}(I, \mathbb{R}^d)$, $I = [a, b]$, mit der Eigenschaft $f_n \rightrightarrows f$ ($n \rightarrow \infty$) auf I . Dann folgt $f \in \mathcal{R}(I, \mathbb{R}^d)$ und

$$\int_a^b \lim_{n \rightarrow \infty} f_n(x) \, dx = \int_a^b f(x) \, dx = \lim_{n \rightarrow \infty} \int_a^b f_n(x) \, dx. \quad (9.5.5)$$

Beweis. → Ergänzungen. □

Bemerkung 9.5.8. (a) Die Aussage von Satz 9.5.7 wird i.A. falsch bei nur punktweise konvergenten Funktionenfolgen.

(b) Satz 9.5.7 lässt sich natürlich wieder auf gleichmäßig konvergente Funktionenreihen übertragen.

Wir sind nun in der Lage, den ausgelassenen Beweis von Satz 8.1.1 nachzuliefern, der uns erst die Differenzierbarkeit der elementaren Funktionen in Kapitel 8 sicherte. Zur Erinnerung formulieren wir ihn noch einmal als

Satz 9.5.9

Sei $I = [a, b]$ und $\{f_n\}_n$ eine Folge von Funktionen $f_n \in C^1(I, \mathbb{R}^d)$ für alle $n \in \mathbb{N}$. Falls dann gilt

$$f_n \rightarrow f \ (n \rightarrow \infty), \quad f'_n \rightrightarrows g \ (n \rightarrow \infty) \quad \text{auf } I,$$

so folgt für den punktweisen Limes $f \in C^1(I, \mathbb{R}^d)$, und es gilt

$$\frac{d}{dx} \left(\lim_{n \rightarrow \infty} f_n(x) \right) = f' = g = \lim_{n \rightarrow \infty} \left(\frac{d}{dx} f_n(x) \right) \quad \text{auf } I.$$

Beweis. Da $f'_n \in C^0(I, \mathbb{R}^d) \subset \mathcal{R}(I, \mathbb{R}^d)$ gilt - siehe Satz 9.2.10, liefert Satz 9.5.7 zunächst $g \in \mathcal{R}(I, \mathbb{R}^d)$ sowie

$$\lim_{n \rightarrow \infty} \int_a^x f'_n(t) dt = \int_a^x g(t) dt \quad \text{für alle } x \in [a, b].$$

Andererseits entnehmen wir dem Fundamentalsatz, Satz 9.4.6, die Relation

$$f_n(x) = f_n(a) + \int_a^x f'_n(t) dt, \quad x \in [a, b], \quad n \in \mathbb{N}, \quad (9.5.6)$$

und nach Grenzübergang $n \rightarrow \infty$ somit

$$f(x) = f(a) + \int_a^x g(t) dt, \quad x \in [a, b].$$

Wegen $f'_n \rightrightarrows g \ (n \rightarrow \infty)$ auf I gilt für die Grenzfunktion $g \in C^0(I, \mathbb{R}^d)$ nach dem Weierstraßschen Konvergenzssatz, Satz 6.6.5. Gemäß Satz 9.4.2 ist also $f \in C^1(I, \mathbb{R}^d)$ eine Stammfunktion von g , d.h. es gilt $f' = g$ auf I , wie behauptet. □

Bemerkung 9.5.10. Der Beweis zeigt, dass es genügt statt $f_n \rightarrow f$ ($n \rightarrow \infty$) auf I die Konvergenz der Punkte $\{f_n(a)\}_n$ zu fordern. Die Beziehung $f_n \rightarrow f$ ($n \rightarrow \infty$) auf I folgt dann aus (9.5.6); genauer definiert (9.5.6) in diesem Fall für $n \rightarrow \infty$ die Grenzfunktion. Dabei kann a noch durch einen beliebigen Punkt $p \in I$ ersetzt werden.

9.6. Uneigentliche Integrale

Bisher haben wir nur beschränkte Funktionen über kompakte Intervalle integriert, beide Einschränkungen sollen jetzt aufgeweicht werden. Wir beschränken uns dabei auf die Betrachtung reellwertiger Funktionen; die Definitionen und Aussagen können - wo sinnvoll - direkt auf komplex- und vektorwertige Funktionen übertragen werden.

Definition 9.6.1: Uneigentliches Integral - halboffene Intervalle

Seien $-\infty \leq a < b \leq +\infty$ gegeben.

- Ist $-\infty < a$, so heißt ein $f : [a, b) \rightarrow \mathbb{R}$ über $[a, b)$ *uneigentlich integrierbar*, wenn $f \in \mathcal{R}([a, p])$ für alle $p \in [a, b)$ gilt und der Grenzwert

$$\int_a^b f(x) dx := \lim_{p \rightarrow b-} \int_a^p f(x) dx$$

existiert; dieser Grenzwert wird dann *uneigentliches Integral* von f über $[a, b)$ genannt.

- Ist $b < +\infty$, so heißt $f : (a, b] \rightarrow \mathbb{R}$ über $(a, b]$ *uneigentlich integrierbar*, wenn $f \in \mathcal{R}([p, b])$ für alle $a < p \leq b$ gilt und der Grenzwert

$$\int_a^b f(x) dx := \lim_{p \rightarrow a+} \int_p^b f(x) dx$$

existiert; dieser Grenzwert wird dann *uneigentliches Integral* von f über $(a, b]$ genannt.

Wir sagen auch, das uneigentliche Integral $\int_a^b f(x) dx$ *konvergiert* (bzw. *divergiert*), wenn der angegebene Grenzwert existiert (bzw. nicht existiert). Falls $\int_a^p f(x) dx \rightarrow \pm\infty$ ($p \rightarrow b-$) bzw. $\int_p^b f(x) dx \rightarrow \pm\infty$ ($p \rightarrow a+$) gilt, sagen wir $\int_a^b f(x) dx$ ist *bestimmt divergent* und schreiben

$$\int_a^b f(x) dx = \pm\infty.$$

Schließlich heit $\int_a^b f(x) dx$ *absolut konvergent*, wenn $\int_a^b |f(x)| dx$ konvergiert.

Bemerkung 9.6.2. Fr $-\infty < a < b < +\infty$ nutzen wir dieselbe Notation fr das uneigentliche Integral ber $[a, b)$ und $(a, b]$ wie fr das bestimmte Integral ber das kompakte Intervall $[a, b]$ aus Definition 9.1.9. Das ist legitim, denn fr ein $f \in \mathcal{R}([a, b])$ liefern die verschiedenen Definitionen denselben Integralwert (\rightarrow bungsaufgabe).

Im Folgenden beziehen wir uns immer auf Integrale ber $[a, b)$; analoge Aussagen gelten natrlich auch fr Integrale ber $(a, b]$.

Satz 9.6.3

Sind $-\infty < a < b \leq +\infty$ und $f : [a, b) \rightarrow \mathbb{R}$ gegeben, so ist f genau dann uneigentlich integrierbar ber $[a, b)$, wenn zu jedem $\varepsilon > 0$ ein $\xi = \xi(\varepsilon) \in [a, b)$ so existiert, dass gilt

$$\left| \int_p^{p'} f dx \right| = \left| \int_a^{p'} f dx - \int_a^p f dx \right| < \varepsilon \quad \text{fr alle } p, p' \in (\xi, b). \quad (9.6.1)$$

Beweis. Wir betrachten die Funktion $F(p) := \int_a^p f dx$, $p \in [a, b)$, und zeigen beide Richtungen getrennt.

- „ \Rightarrow “: Nach Voraussetzung existiert $F_0 := \lim_{p \rightarrow b-} F(p)$, d.h. fr beliebiges $\varepsilon > 0$ existiert ein $\xi = \xi(\varepsilon) \in [a, b)$ mit $|F(p) - F_0| < \frac{\varepsilon}{2}$ fr alle $p \in (\xi, b)$. Daraus folgt sofort

$$|F(p') - F(p)| \leq |F(p') - F_0| + |F(p) - F_0| < \varepsilon \quad \text{fr alle } p, p' \in (\xi, b),$$

also die Behauptung (9.6.1).

- „ \Leftarrow “: Sei nun (9.6.1) fr beliebiges $\varepsilon > 0$ und geeignetes $\xi = \xi(\varepsilon) \in [a, b)$ erfllt. Ist $\{p_n\}_n \subset [a, b)$ eine beliebige Folge mit $p_n \rightarrow b - (n \rightarrow \infty)$, so finden wir ein $N = N(\varepsilon) \in \mathbb{N}$ mit $p_n \in (\xi, b)$ fr alle $n \geq N$, und (9.6.1) entnehmen wir somit

$$|F(p_n) - F(p_m)| < \varepsilon \quad \text{fr alle } n, m \geq N.$$

$\{F(p_n)\}_n \subset \mathbb{R}$ ist also eine CAUCHYfolge und folglich existiert auch $\lim_{n \rightarrow \infty} F(p_n)$. Die Charakterisierung der Funktionsgrenzwerte durch Folgengrenzwerte, Satz 6.1.2 - siehe auch Bemerkung 6.3.2 (b) -, liefert die Existenz von $\lim_{p \rightarrow b-} F(p)$, wie behauptet. \square

Bemerkung 9.6.4. ► Das in Satz 9.6.3 angegebene CAUCHYkriterium fr den speziellen (einseitigen) Grenzwert des uneigentlichen Integrals gilt allgemein fr die Funktionsgrenzwerte aus Abschnitt 6.1; der hier angegebene Beweis lsst sich wrtlich bertragen.

- Die Funktion $F : [a, b) \rightarrow \mathbb{R}$ aus dem Beweis von Satz 9.6.3 ist i.A. keine Stammfunktion von f ; siehe Bemerkung 9.4.3, aber auch Satz 9.4.2.

Folgerung 9.6.5

Seien $-\infty < a < b \leq +\infty$ und $f : [a, b) \rightarrow \mathbb{R}$ gegeben. Falls dann das uneigentliche Integral $\int_a^{+\infty} f(x) dx$ absolut konvergiert, so konvergiert es auch im gewöhnlichen Sinne.

Beweis. Sofort aus Satz 9.6.3 und Satz 9.2.6 (c). \square

Beispiel 9.6.6. $\int_0^{+\infty} \frac{\sin x}{x} dx$ ist konvergent, aber nicht absolut konvergent; die Umkehrung von Folgerung 9.6.5 gilt also nicht (\rightarrow Ergänzungen). Übrigens ist $\int_0^{+\infty} \frac{\sin x}{x} dx = \frac{\pi}{2}$ richtig; siehe [Hil1] Abschnitt 3.11, Beispiel \square .

Satz 9.6.7

Seien $-\infty < a < b \leq +\infty$ und eine *nichtnegative* Funktion $\varphi : [a, b) \rightarrow [0, +\infty)$ gegeben. Dann ist $\int_a^b \varphi dx$ genau dann konvergent, wenn die Funktion $\Phi(p) := \int_a^p \varphi dx$, $p \in [a, b)$, beschränkt ist. Wir schreiben in diesem Fall auch

$$\int_a^b \varphi(x) dx < +\infty.$$

Beweis. Sind $a \leq p < p' < b$ beliebig, so liefern die Additivität, Satz 9.3.2, und die Monotonie, Satz 9.2.6 (b), des Integrals:

$$\Phi(p') = \int_a^{p'} \varphi dx = \int_a^p \varphi dx + \int_p^{p'} \varphi dx \geq \int_a^p \varphi dx + \int_p^{p'} 0 dx \geq \int_a^p \varphi dx = \Phi(p).$$

Folglich ist $\Phi : [a, b) \rightarrow \mathbb{R}$ monoton wachsend, die Behauptung ergibt sich nun sofort aus einem Monotoniekriterium für Funktionsgrenzwerte; siehe Bemerkung 9.6.8. \square

Bemerkung 9.6.8. Das besagte *Monotoniekriterium* für Funktionsgrenzwerte lautet: Sind $-\infty < a < b \leq +\infty$ beliebig und ist $f : [a, b) \rightarrow \mathbb{R}$ monoton wachsend, so gilt:

$$\lim_{x \rightarrow b-} f(x) \text{ existiert} \quad \Leftrightarrow \quad f \text{ ist beschränkt.}$$

Wir belassen den einfachen Beweis als kleine *Übungsaufgabe*.

Die Sätze 9.6.3 und 9.6.7 zeigen eine starke Analogie zwischen dem uneigentlichen Integral und unendlichen Reihen; das Riemannsche Integral über die Intervalle $[a, p]$, $a \leq p < b$, übernimmt dabei die Rolle der Partialsummen einer Reihe. Diese Analogie zeigt sich auch im folgenden

Satz 9.6.9: Majorantenkriterium

Seien $-\infty < a < b \leq +\infty$ und $f : [a, b) \rightarrow \mathbb{R}$ mit $f \in \mathcal{R}([a, p])$ für alle $p \in [a, b)$ gegeben. Existiert eine nichtnegative Funktion $\varphi : [a, b) \rightarrow [0, +\infty)$ mit $\varphi \in \mathcal{R}([a, p])$ für alle $p \in [a, b)$, für die $\int_a^b \varphi \, dx < +\infty$ und

$$|f(x)| \leq \varphi(x) \quad \text{für alle } x \in [a, b)$$

gilt, so konvergiert das uneigentliche Integral $\int_a^{+\infty} f(x) \, dx$ absolut; φ heißt dann *integrable Majorante* für f .

Beweis. Wir betrachten die Funktion $F^*(p) := \int_a^p |f| \, dx$, $p \in [a, b)$. Die Monotonie des Integrals liefert

$$F^*(p) = \int_a^p |f| \, dx \leq \int_a^p \varphi \, dx =: \Phi(p), \quad p \in [a, b).$$

Nach Voraussetzung ist Φ gemäß Satz 9.6.7 beschränkt. Folglich ist auch F^* beschränkt und wiederum nach Satz 9.6.7 konvergiert das uneigentliche Integral $\int_a^b |f| \, dx$, wie behauptet. \square

Bemerkung 9.6.10. Die Formulierung und den Beweis eines entsprechenden Minorantenkriteriums überlassen wir dem Leser.

Beispiel 9.6.11. (a) Es gilt $\int_0^1 \frac{dx}{\sqrt{1-x^2}} = \frac{\pi}{2}$, denn wir haben

$$\int_0^p \frac{dx}{\sqrt{1-x^2}} = \int_0^p \arcsin' x \, dx = \arcsin x \Big|_0^p = \arcsin p \rightarrow \frac{\pi}{2} \quad (p \rightarrow 1-).$$

Analog folgt $\int_{-1}^0 \frac{dx}{\sqrt{1-x^2}} = \frac{\pi}{2}$.

(b) Für $\alpha \in (0, 1)$ gilt $\int_0^1 \frac{dx}{x^\alpha} = \frac{1}{1-\alpha}$, denn

$$\int_p^1 \frac{dx}{x^\alpha} = \frac{x^{1-\alpha}}{1-\alpha} \Big|_p^1 = \frac{1}{1-\alpha} (1 - p^{1-\alpha}) \rightarrow \frac{1}{1-\alpha} \quad (p \rightarrow 0+).$$

Für $\alpha \in [1, +\infty)$ divergiert dieses Integral.

(c) Es gilt $\int_0^{+\infty} \frac{dx}{1+x^2} = \frac{\pi}{2}$, denn

$$\int_0^p \frac{dx}{1+x^2} = \int_0^p \arctan' x \, dx = \arctan x \Big|_0^p = \arctan p \rightarrow \frac{\pi}{2} \quad (p \rightarrow +\infty).$$

Analog folgt $\int_{-\infty}^0 \frac{dx}{1+x^2} = \frac{\pi}{2}$.

(d) Die uneigentlichen Integrale $\int_{-\infty}^0 e^{-x^2} dx$, $\int_0^{+\infty} e^{-x^2} dx$ konvergieren, denn:

Aus der Reihendarstellung der e -Funktion folgt $e^{x^2} \geq 1 + x^2 > 0$ bzw. $0 < e^{-x^2} \leq \frac{1}{1+x^2}$ für alle $x \in \mathbb{R}$. Beispiel (c) und das Majorantenkriterium, Satz 9.6.11, liefern die Behauptung.

Definition 9.6.12: Uneigentliches Integral - offene Intervalle

Seien $-\infty \leq a < b \leq +\infty$ und $f : (a, b) \rightarrow \mathbb{R}$ so gegeben, dass f über jedes kompakte Teilintervall $I \subset (a, b)$ integrierbar ist. Gibt es dann ein $c \in (a, b)$, so dass die uneigentlichen Integrale $\int_a^c f \, dx$ und $\int_c^b f \, dx$ existieren, so heißt f *uneigentlich integrierbar* über (a, b) und wir setzen

$$\int_a^b f(x) \, dx := \int_a^c f(x) \, dx + \int_c^b f(x) \, dx$$

für das *uneigentliche Integral* von f über (a, b) . Die weiteren Schreib- und Sprechweisen aus Definition 9.6.1 übertragen sich entsprechend.

Beispiel 9.6.13. (a) Aus Beispiel 9.6.11 (a) und (c) lesen wir ab:

$$\begin{aligned} \int_{-1}^1 \frac{dx}{\sqrt{1-x^2}} &= \int_{-1}^0 \frac{dx}{\sqrt{1-x^2}} + \int_0^1 \frac{dx}{\sqrt{1-x^2}} = \frac{\pi}{2} + \frac{\pi}{2} = \pi, \\ \int_{-\infty}^{+\infty} \frac{dx}{1+x^2} &= \int_{-\infty}^0 \frac{dx}{1+x^2} + \int_0^{+\infty} \frac{dx}{1+x^2} = \pi. \end{aligned}$$

(b) Gemäß Beispiel 9.6.11 (d) existiert das *GAUSSsche Fehlerintegral*

$$\int_{-\infty}^{+\infty} e^{-x^2} dx = \int_{-\infty}^0 e^{-x^2} dx + \int_0^{+\infty} e^{-x^2} dx.$$

In Kapitel 12 werden wir sehen, dass $\int_{-\infty}^{+\infty} e^{-x^2} dx = \sqrt{\pi}$ gilt.

Bemerkung 9.6.14. Naheliegender erscheint möglicherweise eine Definition z.B. des uneigentlichen Integrals $\int_{-\infty}^{+\infty} f(x) \, dx$ als Grenzwert

$$\lim_{p \rightarrow +\infty} \int_{-p}^p f(x) \, dx.$$

Vorsicht: Dieser sogenannte *CAUCHYSche Hauptwert* kann existieren, ohne dass das in Definition 9.6.12 erklärte uneigentliche Integral $\int_{-\infty}^{+\infty} f(x) dx$ existiert. Z.B. für $f = \text{Id}$ gilt

$$\int_{-p}^p x dx = \frac{x^2}{2} \Big|_{-p}^p = 0 \rightarrow 0 \quad (p \rightarrow +\infty),$$

aber mit beliebigem $c \in \mathbb{R}$:

$$\begin{aligned} \int_{-p}^c x dx &= \frac{c^2}{2} - \frac{p^2}{2} \rightarrow -\infty \quad (p \rightarrow +\infty), \\ \int_c^p x dx &= \frac{p^2}{2} - \frac{c^2}{2} \rightarrow +\infty \quad (p \rightarrow +\infty). \end{aligned}$$

CAUCHYSche Hauptwerte lassen sich für beliebige offene Intervalle (a, b) angeben.

Wir schließen unsere Betrachtungen mit einem Resultat, das noch einmal die Verwandtschaft uneigentlicher Integrale mit unendlichen Reihen verdeutlicht:

Satz 9.6.15: RIEMANNSches Integralkriterium

Sei $f : [1, +\infty) \rightarrow \mathbb{R}$ eine monoton fallende, nichtnegative Funktion mit $f \in \mathcal{R}([1, p])$ für alle $p \in [1, +\infty)$. Dann konvergiert die Reihe $\sum_{k=1}^{\infty} a_k$ mit den Gliedern $a_k := f(k)$, $k \in \mathbb{N}$, genau dann, wenn das uneigentliche Integral $\int_1^{+\infty} f(x) dx$ konvergiert. In diesem Fall gilt

$$\sum_{k=2}^{\infty} a_k \leq \int_1^{+\infty} f(x) dx \leq \sum_{k=1}^{\infty} a_k.$$

Beweis. → Ergänzungen. □

Beispiel 9.6.16. Für $\alpha > 1$ berechnen wir

$$\int_1^{+\infty} \frac{dx}{x^\alpha} = \lim_{p \rightarrow +\infty} \frac{x^{1-\alpha}}{1-\alpha} \Big|_1^p = \frac{1}{\alpha-1}.$$

Nach Satz 9.6.15 konvergiert somit die Reihe $\sum_{k=1}^{\infty} \frac{1}{k^\alpha}$ für alle $\alpha > 1$ und für ihren Wert gilt:

$$\sum_{k=1}^{\infty} \frac{1}{k^\alpha} \in \left[\frac{1}{\alpha-1}, \frac{\alpha}{\alpha-1} \right].$$

9.7. Die TAYLORSche Formel

Sei $I \subset \mathbb{R}$ ein Intervall, $f : I \rightarrow \mathbb{R}$ eine Funktion und $x_0 \in I$ sei fixiert. Wir wissen bereits (Satz 6.2.3 und Bemerkung 7.1.4):

- Ist f stetig in x_0 , so gilt $f(x) = f(x_0) + \varphi(x)$, $x \in I$,
- Ist f differenzierbar in x_0 , so gilt $f(x) = f(x_0) + (x - x_0)f'(x_0) + (x - x_0)\varphi(x)$, $x \in I$.

Dabei ist $\varphi : I \rightarrow \mathbb{R}$ jeweils eine (von f und x_0 abhängige) Funktion, die in x_0 stetig ist und $\varphi(x_0) = 0$ erfüllt. Wir können also f nahe x_0 durch eine konstante (bei Stetigkeit) bzw. affin-lineare (bei Differenzierbarkeit) Funktion approximieren.

Mit der TAYLORformel - die Idee geht auf B. TAYLOR zurück - soll allgemeiner eine Approximation einer vorgegebenen, von höherer Ordnung differenzierbaren Funktion durch ein geeignetes Polynom angegeben werden; wesentlich ist dabei die Kontrolle (also die Abschätzung) des Fehlerterms.

Motivation. Es sei $f(x) = \sum_{k=0}^n a_k (x - x_0)^k$, $x \in \mathbb{R}$, ein Polynom n -ten Grades mit $a_0, \dots, a_n \in \mathbb{R}$. Für die Ableitungen im Punkt x_0 folgt dann

$$f^{(k)}(x_0) = k! a_k, \quad k = 0, \dots, n,$$

Umstellen und Einsetzen liefert also

$$f(x) = \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k, \quad x \in \mathbb{R}.$$

Für $n = 0, 1$ stimmt die rechte Seite mit den oben angegebenen Approximationen überein. Außerdem ist für ein Polynom n -ten Grades der Approximationsfehler durch die rechte Seite identisch 0, die Näherung ist also die bestmögliche. Wir halten daher fest:

Definition 9.7.1

Seien $n \in \mathbb{N}_0$ und $f : I \rightarrow \mathbb{R}$ n -mal differenzierbar auf dem Intervall $I \subset \mathbb{R}$. Wir erklären das n -te TAYLORpolynom $T_{n,x_0}f$ von f zum Entwicklungspunkt $x_0 \in I$ als

$$T_{n,x_0}f(x) := \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k, \quad x \in I. \quad (9.7.1)$$

Der Fehler

$$R_{n,x_0}f(x) := f(x) - T_{n,x_0}f(x), \quad x \in I, \quad (9.7.2)$$

heißt n -tes Restglied von f zu x_0 .

Bemerkung 9.7.2. Offenbar gilt die TAYLORSche Formel

$$f(x) = T_{n,x_0}f(x) + R_{n,x_0}f(x), \quad x \in I, \quad (9.7.3)$$

für ein n -mal differenzierbares $f : I \rightarrow \mathbb{R}$ und $x_0 \in I$. Erst durch die Untersuchung des Restglieds (analog zu φ in der Einleitung) erhält diese Formel Aussagekraft. Wir beginnen mit dem

Satz 9.7.3

Seien $n \in \mathbb{N}_0$, $f \in C^{n+1}(I)$ und $x_0 \in I$ gegeben. Für das n -te Restglied haben wir dann die Darstellung

$$R_{n,x_0}f(x) = \frac{1}{n!} \int_{x_0}^x (x-t)^n f^{(n+1)}(t) dt, \quad x \in I. \quad (9.7.4)$$

Beweis. Wir zeigen (9.7.3) mit der Darstellung (9.7.4) für $R_{n,x_0}f$ durch vollständige Induktion über n :

- $n = 0$: Der Fundamentalsatz der Differential- und Integralrechnung, Satz 9.4.6, liefert

$$f(x) = f(x_0) + \int_{x_0}^x f'(t) dt = T_{0,x_0}f(x) + \int_{x_0}^x f'(t) dt, \quad x \in I,$$

also die Behauptung für $n = 0$.

- $n \rightarrow n + 1$: Dann haben wir für $f \in C^{n+2}(I)$

$$\begin{aligned} R_{n,x_0}f(x) &\stackrel{(IV)}{=} \frac{1}{n!} \int_{x_0}^x (x-t)^n f^{(n+1)}(t) dt = - \int_{x_0}^x f^{(n+1)}(t) \frac{d}{dt} \left[\frac{(x-t)^{n+1}}{(n+1)!} \right] dt \\ &\stackrel{\text{Satz 9.5.1}}{=} - \left[f^{(n+1)}(t) \frac{(x-t)^{n+1}}{(n+1)!} \right] \Big|_{x_0}^x + \int_{x_0}^x f^{(n+2)}(t) \frac{(x-t)^{n+1}}{(n+1)!} dt \\ &= \frac{f^{(n+1)}(x_0)}{(n+1)!} (x-x_0)^{n+1} + \frac{1}{(n+1)!} \int_{x_0}^x (x-t)^{n+1} f^{(n+2)}(t) dt, \end{aligned}$$

also

$$\begin{aligned} f(x) &\stackrel{(IV)}{=} T_{n,x_0}f(x) + R_{n,x_0}f(x) \\ &= T_{n+1,x_0}f(x) + \frac{1}{(n+1)!} \int_{x_0}^x (x-t)^{n+1} f^{(n+2)}(t) dt, \quad x \in I, \end{aligned}$$

d.h. die Behauptung für $n + 1$. □

Satz 9.7.4: LAGRANGESche Restgliedformel

Unter den Voraussetzungen von Satz 9.7.3 gibt es zu jedem $x \in I$ ein $\vartheta = \vartheta(x) \in (0, 1)$, so

dass das Restglied geschrieben werden kann als

$$R_{n,x_0}(x) = \frac{f^{(n+1)}(x_0 + \vartheta(x - x_0))}{(n+1)!} (x - x_0)^{n+1}, \quad x \in I. \quad (9.7.5)$$

Beweis. Für $x = x_0$ ist nichts zu zeigen, sei $x > x_0$. Nach dem Mittelwertsatz der Integralrechnung, Satz 9.2.12, gibt es ein $\xi \in (x_0, x)$, so dass gilt:

$$\begin{aligned} R_{n,x_0}(x) &\stackrel{(9.7.4)}{=} \int_{x_0}^x \frac{(x-t)^n}{n!} f^{(n+1)}(t) dt = f^{(n+1)}(\xi) \int_{x_0}^x \frac{(x-t)^n}{n!} dt \\ &= -f^{(n+1)}(\xi) \frac{(x-t)^{n+1}}{(n+1)!} \Big|_{x_0}^x = \frac{f^{(n+1)}(\xi)}{(n+1)!} (x - x_0)^{n+1}. \end{aligned}$$

Die Behauptung folgt mit $\vartheta := \frac{\xi - x_0}{x - x_0} \in (0, 1)$. Im Falle $x < x_0$ wendet man zunächst Definition 9.3.3 an und verfährt anschließend wie oben. \square

Mit Satz 9.7.4 können wir nun die eingangs erwähnten Approximationsaussagen verallgemeinern:

Folgerung 9.7.5

Seien $n \in \mathbb{N}_0$, $f \in C^n(I)$ und $x_0 \in I$ gegeben. Dann existiert eine Funktion $\varphi : I \rightarrow \mathbb{R}$, die in x_0 stetig ist mit $\varphi(x_0) = 0$, so dass gilt

$$f(x) = T_{n,x_0}f(x) + (x - x_0)^n \varphi(x), \quad x \in I.$$

Beweis. Wir wenden Satz 9.7.4 mit $n - 1$ an und erhalten

$$\begin{aligned} f(x) &= T_{n-1,x_0}f(x) + \frac{f^{(n)}(x_0 + \vartheta(x - x_0))}{n!} (x - x_0)^n \\ &= T_{n,x_0}(x) + \frac{f^{(n)}(x_0 + \vartheta(x - x_0)) - f^{(n)}(x_0)}{n!} (x - x_0)^n \quad \text{für } x \in I \end{aligned}$$

mit einem $\vartheta = \vartheta(x) \in (0, 1)$. Wegen $f^{(n)}(x_0 + \vartheta(x - x_0)) \rightarrow f^{(n)}(x_0)$ für $x \rightarrow x_0$ ergibt sich die Behauptung, wenn wir

$$\varphi(x) := \frac{f^{(n)}(x_0 + \vartheta(x - x_0)) - f^{(n)}(x_0)}{n!}, \quad x \in I,$$

setzen. \square

Definition 9.7.6: LANDAUSymbole

Sei $M \subset \mathbb{R}^n$ eine beliebige Menge, $x_0 \in M$ und $\psi : M \setminus \{x_0\} \rightarrow \mathbb{R}$ eine Funktion mit $\psi(x) \neq 0$ für $x \in (M \cap B_r(x_0)) \setminus \{x_0\}$ mit einem $r > 0$. Eine Funktion $g : M \setminus \{x_0\} \rightarrow \mathbb{R}^d$ hat dann:

(a) die *Ordnung „groß O von $\psi(x)$ für $x \rightarrow x_0$ “*, wenn gilt

$$\left\| \frac{g(x)}{\psi(x)} \right\| \leq \text{const} \quad \text{für alle } x \in (M \cap B_\varepsilon(x_0)) \setminus \{x_0\}$$

mit einem $\varepsilon \in (0, r)$; wir schreiben

$$g(x) = O(\psi(x)) \quad \text{für } x \rightarrow x_0.$$

(b) die *Ordnung „klein o von $\psi(x)$ für $x \rightarrow x_0$ “*, wenn sogar

$$\lim_{x \rightarrow x_0} \frac{g(x)}{\psi(x)} = 0$$

erfüllt ist; wir schreiben

$$g(x) = o(\psi(x)) \quad \text{für } x \rightarrow x_0.$$

Bemerkung 9.7.7. (a) Offenbar impliziert $g(x) = o(\psi(x))$ die Relation $g(x) = O(\psi(x))$ für $x \rightarrow x_0$, die Umkehrung stimmt nicht. Ein Standardbeispiel für die Vergleichsfunktion ψ ist $\psi(x) = \|x - x_0\|^\alpha$ für $\alpha \neq 0$.

(b) Die Aussage von Folgerung 9.7.5 lässt sich nun äquivalent schreiben als

$$f(x) = T_{n,x_0}f(x) + o(|x - x_0|^n) \quad \text{für } x \rightarrow x_0$$

(\rightarrow Nachrechnen). Man sagt: $T_{n,x_0}f$ approximiert f von n -ter Ordnung.

9.8. TAYLORreihen

Der naheliegende Versuch, für Funktionen $f \in C^\infty(I)$ in der Taylorformel (9.7.3) zur Grenze $n \rightarrow \infty$ überzugehen, um eine Reihendarstellung für f zu finden, führt auf folgenden Begriff:

Definition 9.8.1

Seien $I \subset \mathbb{R}$ ein Intervall und $f \in C^\infty(I)$ eine gegebene Funktion. Ist $x_0 \in I$ gewählt, so

heißt

$$T_{x_0}f(x) := \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k, \quad x \in I,$$

TAYLORreihe von f zum Entwicklungspunkt x_0 .

Natürlich muss die TAYLORreihe einer C^∞ -Funktion nicht konvergieren (außerhalb vom Entwicklungspunkt). Es gilt der

Satz 9.8.2: Reihenentwicklung

Seien $f \in C^\infty(I)$ und $x_0 \in I$ gewählt. Falls dann in einem Punkt $x \in I$ für die Restglieder von f zu x_0 gilt

$$\lim_{n \rightarrow \infty} R_{n,x_0}f(x) = 0, \quad (9.8.1)$$

so konvergiert die TAYLORreihe von f zu x_0 in x und es gilt

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k. \quad (9.8.2)$$

Konvergiert umgekehrt die TAYLORreihe und gilt (9.8.2), so folgt auch Relation (9.8.1).

Beweis. Sofort aus der TAYLORformel (9.7.3). □

Beispiel 9.8.3. Die Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ mit

$$f(x) = \begin{cases} \exp(-\frac{1}{x^2}), & x \neq 0, \\ 0, & x = 0 \end{cases}$$

gehört zur Klasse $C^\infty(\mathbb{R})$ und erfüllt $f^{(k)}(0) = 0$ für alle $k \in \mathbb{N}_0$ (\rightarrow Übungsaufgabe). Für die TAYLORreihe bei $x_0 = 0$ finden wir also

$$T_0f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} x^k = 0 \neq f(x) \quad \text{für } x \neq 0.$$

Relation (9.8.1) gilt also *nicht* für jedes $f \in C^\infty(\mathbb{R})$, selbst wenn die TAYLORreihe konvergiert.

Bemerkung 9.8.4. Ist $f \in C^\infty(I)$ auf $I = (-r, r)$ definiert oder dargestellt durch eine Potenzreihe (um $x_0 = 0$), d.h.

$$f(x) = \sum_{k=0}^{\infty} a_k x^k,$$

so folgt $a_k = \frac{1}{k!} f^{(k)}(0)$ aus Folgerung 8.1.5. In diesem Fall stimmt f also notwendig mit ihrer TAYLORreihe (um $x_0 = 0$) überein; z.B. sind durch

$$\sum_{k=0}^{\infty} \frac{x^k}{k!}, \quad \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!}, \quad \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!}, \quad x \in \mathbb{R},$$

die TAYLORreihen um $x_0 = 0$ von \exp , \sin bzw. \cos gegeben.

Beispiel 9.8.5. (a) *Die Binomialreihe.* Sei $\alpha \in \mathbb{R} \setminus \{0\}$. Wir wollen die Identität

$$(1+x)^\alpha = \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k, \quad x \in (-1, 1), \quad (9.8.3)$$

beweisen - die rechte Seite heißt *Binomialreihe*. Dabei bezeichnen

$$\binom{\alpha}{0} := 1, \quad \binom{\alpha}{k} := \prod_{l=1}^k \frac{\alpha - l + 1}{l} = \frac{\alpha(\alpha-1)\dots(\alpha-k+1)}{k!}, \quad k \in \mathbb{N},$$

die *allgemeinen Binomialkoeffizienten*. Für $\alpha = n \in \mathbb{N}$ erhalten wir einen Spezialfall des Binomischen Satzes, für $\alpha = -1$ und $x \rightarrow -x$ finden wir die geometrische Reihe.

Zum Beweis von (9.8.3) schreiben wir $f(x) := (1+x)^\alpha$, $x > -1$. Nach Satz 8.2.11 gilt $f \in C^\infty((-1, +\infty))$ und aus (8.2.8) folgt induktiv

$$f^{(k)}(x) = \alpha(\alpha-1)\dots(\alpha-k+1)(1+x)^{\alpha-k} = k! \binom{\alpha}{k} (1+x)^{\alpha-k}, \quad x > -1, \quad k \in \mathbb{N}.$$

Einsetzen liefert für das n -te Taylorpolynom von f zum Entwicklungspunkt $x_0 = 0$:

$$T_{n,0}f(x) = \sum_{k=0}^n \binom{\alpha}{k} x^k, \quad x > -1, \quad n \in \mathbb{N}_0.$$

Formel (9.8.3) folgt somit gemäß Satz 9.8.2 aus $R_{n,0}f(x) \rightarrow 0$ ($n \rightarrow \infty$) für $x \in (-1, 1)$ (\rightarrow Ergänzungen).

(b) *Die Logarithmus-Reihe.* Wir behaupten

$$\log(1+x) = - \sum_{k=1}^{\infty} \frac{(-1)^k}{k} x^k, \quad x \in (-1, 1). \quad (9.8.4)$$

Dazu betrachten wir die Funktion $g(x) := \log(1+x)$, $x > -1$, und beachten

$$g'(x) = \frac{1}{1+x} = \sum_{k=0}^{\infty} (-1)^k x^k \quad \text{für } x \in (-1, 1).$$

Der Fundamentalsatz, Satz 9.4.6, liefert also

$$g(x) = \int_0^x \frac{dt}{1+t} = \int_0^x \left[\sum_{k=0}^{\infty} (-1)^k t^k \right] dt \quad \text{für } x \in (-1, 1).$$

Da $\sum_k (-1)^k t^k$ auf $[0, x]$ bzw. $[x, 0]$ mit beliebigem $x \in (-1, 1)$ gleichmäßig konvergiert (Satz 6.6.9), liefert Satz 9.5.9:

$$\begin{aligned} g(x) &= \int_0^x \lim_{n \rightarrow \infty} \left[\sum_{k=0}^n (-1)^k t^k \right] dt = \lim_{n \rightarrow \infty} \int_0^x \left[\sum_{k=0}^n (-1)^k t^k \right] dt \\ &= \lim_{n \rightarrow \infty} \left[\sum_{k=0}^n (-1)^k \int_0^x t^k dt \right] \\ &= \sum_{k=0}^{\infty} (-1)^k \frac{x^{k+1}}{k+1}, \end{aligned}$$

also (9.8.4). Die Reihe auf der rechten Seite ist die TAYLORreihe von $\log(1+x)$, siehe Bemerkung 9.8.4.

(c) *Die Arcus-Tangens-Reihe.* Analog zu (b) zeigt man (\rightarrow Nachrechnen!)

$$\arctan x = \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} x^{2k+1} \quad \text{für } x \in (-1, 1). \quad (9.8.5)$$

Bemerkung 9.8.6. Mit Hilfe des ABELSchen Stetigkeitssatzes kann man zeigen, dass die Darstellungen (9.8.4), (9.8.5) für $x = 1$ richtig bleiben (\rightarrow Ergänzungen). Damit finden wir die in Beispiel 4.3.4 angekündigten Reihenwerte

$$\sum_{k=1}^{\infty} \frac{(-1)^k}{k} = -\log 2, \quad \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} = \arctan 1 = \frac{\pi}{4}.$$

Auch auf die Binomialreihe ist der Abelsche Stetigkeitssatz anwendbar, falls $\alpha > -1$ gilt. Aus (9.8.3) folgt dann die Darstellung

$$2^\alpha = \sum_{k=0}^{\infty} \binom{\alpha}{k}, \quad \alpha > -1.$$

Teil IV.

Differential- und Integralrechnung in mehreren Veränderlichen

In Teil IV werden unsere Ableitungs- und Integralbegriffe auf Funktionen mehrerer reeller Veränderlicher übertragen. In Kapitel 10 führen wir die verschiedenen Ableitungsbegriffe ein und verallgemeinern die TAYLORSche Formel auf Funktionen mehrerer Veränderlicher. In Kapitel 11 wenden wir die Differentialrechnung einerseits auf Extremwertaufgaben, andererseits auf nichtlineare Gleichungen an. Kapitel 12 ist dann dem mehrdimensionalen Riemannschen Integral gewidmet; den Höhepunkt der Vorlesung stellt die Transformationsformel dar.

Kapitel 10.

Differenzierbare Funktionen mehrerer Veränderlicher

In diesem Kapitel führen wir Ableitungsbegriffe für Funktionen mehrerer Veränderlicher ein. Abschnitt 10.1 ist der partiellen Ableitung gewidmet. In Abschnitt 10.2 geben wir den Mittelwertsatz in mehreren Veränderlichen an und lassen uns dadurch zur Definition der Richtungsableitungen inspirieren. Die totale Ableitung (bzw. das Differential) führen wir in Abschnitt 10.3 ein. Partielle Ableitungen höherer Ordnung erklären wir in Abschnitt 10.4 und beweisen den wichtigen Satz von SCHWARZ. Schließlich wird in Abschnitt 10.5 die TAYLORformel bewiesen.

10.1. Partielle Ableitungen

In Kapitel 2 haben wir bereits Funktionen

$$f = f(x) = (f_1(x_1, \dots, x_n), \dots, f_d(x_1, \dots, x_n)) : \Omega \rightarrow \mathbb{R}^d$$

auf einer Menge $\Omega \subset \mathbb{R}^n$ für $n, d \in \mathbb{N}$ kennengelernt und auf Stetigkeitseigenschaften untersucht. Um auch Wachstums- und Krümmungsverhalten solcher Funktionen mehrerer Veränderlicher beschreiben zu können, benötigen wir einen Ableitungsbegriff. Dazu verallgemeinern wir zunächst Definition 7.1.1 direkt auf Funktionen $f : \Omega \rightarrow \mathbb{R}^d$; einen alternativen Ansatz lernen wir in Abschnitt 10.3 kennen.

Im Folgenden sei $\Omega \subset \mathbb{R}^n$ stets offen. Mit $e_j \in \mathbb{R}^n$ bezeichnen wir immer den j -ten Standardeinheitsvektor, d.h.

$$e_j := (0, \dots, 0, \underset{\hookrightarrow j\text{-te Stelle}}{1}, 0, \dots, 0), \quad j = 1, \dots, n.$$

Dann existiert zu jedem $x^0 \in \Omega$ ein $\varepsilon = \varepsilon(x^0) > 0$ mit $B_\varepsilon(x^0) \subset \Omega$ und somit

$$x^0 + he_j = (x_1^0, \dots, x_{j-1}^0, x_j^0 + h, x_{j+1}^0, \dots, x_n^0) \in \Omega \quad \text{für alle } |h| < \varepsilon, \quad j = 1, \dots, n.$$

Definition 10.1.1: Partielle Ableitung

Sei $f : \Omega \rightarrow \mathbb{R}^d$ gegeben und $x^0 \in \Omega$ sei gewählt. Zu $j \in \{1, \dots, n\}$ nennen wir

$$D_j f(x^0) := \left. \frac{d}{dt} f(x^0 + t e_j) \right|_{t=0} = \lim_{h \rightarrow 0} \frac{f(x^0 + h e_j) - f(x^0)}{h}$$

die *partielle Ableitung* von f nach x_j in x^0 , falls der Grenzwert existiert. f heißt dann in x^0 *partiell differenzierbar* nach x_j . Alternative Schreibweisen sind

$$D_j f(x^0) = \frac{\partial f}{\partial x_j}(x^0) = f_{x_j}(x^0).$$

Bemerkung 10.1.2. (a) Schreiben wir unter den Voraussetzungen von Definition 10.1.1

$$\varphi_j(t) := f(x_1^0, \dots, x_{j-1}^0, t, x_{j+1}^0, \dots, x_n^0), \quad t \in (x_j^0 - \varepsilon, x_j^0 + \varepsilon),$$

mit $B_\varepsilon(x^0) \subset \Omega$, so folgt $D_j f(x^0) = \varphi_j'(x_j^0)$. Die partielle Ableitung von f nach x_j ist also einfach die gewöhnliche Ableitung der Einschränkung φ_j von f auf die j -te Variable x_j bei Festhalten der übrigen $n - 1$ Variablen. Insbesondere stimmt für $n = 1$ die partielle mit der gewöhnlichen Ableitung überein.

(b) Eine Funktion $f = (f_1, \dots, f_d) : \Omega \rightarrow \mathbb{R}^d$ ist genau dann in $x^0 \in \Omega$ nach x_j partiell differenzierbar, wenn f_1, \dots, f_d dort nach x_j partiell differenzierbar sind. Es gilt dann

$$D_j f(x^0) = (D_j f_1(x^0), \dots, D_j f_d(x^0)).$$

Definition 10.1.3

Falls für $f : \Omega \rightarrow \mathbb{R}^d$ die partiellen Ableitungen $D_1 f(x), \dots, D_n f(x)$ für alle $x \in \Omega \subset \mathbb{R}^n$ existieren und stetig sind, so heißt f (*einmal*) *stetig differenzierbar* und wir schreiben $f \in C^1(\Omega, \mathbb{R}^d)$. Für $d = 1$ schreiben wir kurz $C^1(\Omega) := C^1(\Omega, \mathbb{R})$ und für $d = 2$ identifizieren wir wieder $C^1(\Omega, \mathbb{C}) := C^1(\Omega, \mathbb{R}^2)$.

Satz 10.1.4

Jede Funktion $f \in C^1(\Omega, \mathbb{R}^d)$ ist stetig in Ω , d.h. wir haben die Inklusion $C^1(\Omega, \mathbb{R}^d) \subset C^0(\Omega, \mathbb{R}^d)$.

Beweis. Indem wir jede Komponente einzeln betrachten, genügt es den Fall $d = 1$ zu untersuchen.

Sei $x^0 \in \Omega$ beliebig und $\varepsilon > 0$ mit $\overline{B_\varepsilon(x^0)} \subset \Omega$ sei gewählt. Zu beliebigem $x \in B_\varepsilon(x^0)$ schreiben wir I_j für die abgeschlossenen Intervalle zwischen x_j^0 und x_j mit $j = 1, \dots, n$ (also $I_j = [x_j^0, x_j]$ für $x_j^0 \leq x_j$, sonst $I_j = [x_j, x_j^0]$). Die Funktionen

$$\begin{aligned}\varphi_1(t) &:= f(t, x_2^0, \dots, x_n^0), & t \in I_1, \\ \varphi_2(t) &:= f(x_1, t, x_3^0, \dots, x_n^0), & t \in I_2, \\ &\vdots \\ \varphi_n(t) &:= f(x_1, \dots, x_{n-1}, t), & t \in I_n,\end{aligned}$$

sind differenzierbar in I_j und damit dort stetig gemäß Folgerung 7.1.5. Weiter gilt

$$\begin{aligned}f(x^0) - f(x) &= \varphi_1(x_1^0) - \varphi_n(x_n) \\ &= (\varphi_1(x_1^0) - \varphi_1(x_1)) + (\varphi_2(x_2^0) - \varphi_2(x_2)) + \dots + (\varphi_n(x_n^0) - \varphi_n(x_n)) \\ &= \sum_{j=1}^n (\varphi_j(x_j^0) - \varphi_j(x_j)).\end{aligned}\tag{10.1.1}$$

Mit $x \rightarrow x^0$ gilt $x_j \rightarrow x_j^0$ und daher $\varphi_j(x_j) \rightarrow \varphi_j(x_j^0)$, so dass (10.1.1) $f(x) \rightarrow f(x^0)$ liefert, d.h. f ist stetig im beliebig gewählten Punkt $x^0 \in \Omega$. \square

Bemerkung 10.1.5. ▶ Die Existenz von $D_1 f(x^0), \dots, D_n f(x^0)$ in einem Punkt $x^0 \in \Omega$ impliziert i.A. nicht die Stetigkeit von f in x^0 , siehe Beispiel 10.3.6.

- ▶ Produkt-, Quotienten- und Linearitätsregel für differenzierbare Funktionen einer Veränderlichen übertragen sich sofort auf die partiellen Ableitungen; insbesondere ist $C^1(\Omega, \mathbb{R}^d)$ ein linearer Raum. Die Kettenregel erhält folgende Form:

Satz 10.1.6: Kettenregel

Seien $m, n, d \in \mathbb{N}$, offene Mengen $\Omega \subset \mathbb{R}^n$, $\Theta \subset \mathbb{R}^m$ und Funktionen $f = f(x) \in C^1(\Omega, \mathbb{R}^m)$ und $g = g(y) \in C^1(\Theta, \mathbb{R}^d)$ mit $f(\Omega) \subset \Theta$ gegeben. Dann gehört auch die Funktion $h := g \circ f : \Omega \rightarrow \mathbb{R}^d$ zur Klasse $C^1(\Omega, \mathbb{R}^d)$ und es gilt

$$\frac{\partial h}{\partial x_j}(x) = \sum_{l=1}^m \frac{\partial g}{\partial y_l}(f(x)) \frac{\partial f_l}{\partial x_j}(x) \quad \text{für alle } x \in \Omega, j = 1, \dots, n.\tag{10.1.2}$$

Beweis. Analog zum Beweis von Satz 10.1.4 → Ergänzungen. \square

Für folgende Bezeichnungen verwenden wir oft den formal als Vektor erklärten *Nabla-Operator* oder *Nabla-Vektor*

$$\nabla := (D_1, \dots, D_n).$$

Definition 10.1.7

- (a) Zu $f \in C^1(\Omega)$, $\Omega \subset \mathbb{R}^n$, erklären wir den **Gradienten** $\text{grad } f : \Omega \rightarrow \mathbb{R}^n \in C^0(\Omega, \mathbb{R}^n)$ von f als das Vektorfeld

$$\text{grad } f(x) := (D_1 f(x), \dots, D_n f(x)) = \nabla f(x), \quad x \in \Omega.$$

- (b) Für eine Funktion $f = (f_1, \dots, f_d) \in C^1(\Omega, \mathbb{R}^d)$, $\Omega \subset \mathbb{R}^n$, erklären wir die **JACOBI-Matrix** oder **Funktionalmatrix**

$$Df(x) := \begin{pmatrix} D_1 f_1(x) & \dots & D_n f_1(x) \\ \vdots & \ddots & \vdots \\ D_1 f_d(x) & \dots & D_n f_d(x) \end{pmatrix} = \begin{pmatrix} \text{grad } f_1(x) \\ \vdots \\ \text{grad } f_d(x) \end{pmatrix}.$$

Im Fall $d = n$ bezeichnen wir die der quadratischen Matrix Df zugeordnete Determinante

$$J_f := \det(Df) = \begin{vmatrix} D_1 f_1(x) & \dots & D_n f_1(x) \\ \vdots & \ddots & \vdots \\ D_1 f_n(x) & \dots & D_n f_n(x) \end{vmatrix},$$

als **JACOBIdeterminante** oder **Funktionaldeterminante**.

Bemerkung 10.1.8. ► Wenn in Satz 10.1.6 $d = 1$ gilt, d.h. g und h skalare Funktionen sind, liest sich die Kettenregel (10.1.2) als

$$\frac{\partial h}{\partial x_j}(x) = \left\langle \nabla g(f(x)), \frac{\partial f}{\partial x_j}(x) \right\rangle, \quad x \in \Omega, \quad j = 1, \dots, n.$$

wobei $\nabla g = \text{grad } g$ als Gradient bez. y aufzufassen ist.

- Im allgemeinen Fall $d \in \mathbb{N}$ lässt sich (10.1.2) zusammenfassen zu

$$Dh(x) = Dg(f(x)) \circ Df(x), \quad x \in \Omega. \quad (10.1.3)$$

Dabei beziehen sich die Ableitungen bei h und f auf x und bei g auf $y \in \Theta \supset f(\Omega)$, weshalb man auch manchmal zur Verdeutlichung

$$D_x h(x) = D_y g(f(x)) \circ D_x f(x)$$

schreibt.

- Falls in Satz 10.1.6 gerade $m = n = d$ gilt, also Df , Dg und Dh quadratische Matrizen sind, liefern (10.1.3) und der Produktsatz für Determinanten noch

$$J_h(x) = J_g(f(x)) J_f(x), \quad x \in \Omega. \quad (10.1.4)$$

Beispiel 10.1.9. (a) *Abstandsfunktion.* $r : \mathbb{R}^n \rightarrow \mathbb{R}, x \mapsto r(x) := \|x\| = \sqrt{x_1^2 + \dots + x_n^2}$, gehört zur Klasse $C^1(\mathbb{R}^n \setminus \{0\}) \cap C^0(\mathbb{R}^n)$. Für die Ableitung nach x_j erhalten wir

$$\frac{\partial}{\partial x_j} r(x) = \frac{1}{2\sqrt{x_1^2 + \dots + x_n^2}} \cdot 2x_j = \frac{x_j}{\|x\|}$$

bzw.

$$\nabla r(x) = \text{grad } r(x) = \frac{x}{\|x\|} \quad \text{für } x \neq 0.$$

(b) *Polarkoordinaten.* Für beliebiges $z \in \mathbb{C}$ haben wir die Darstellung $z = r e^{i\varphi}$ mit $r = |z| \geq 0$, $\varphi \in \mathbb{R}$; siehe Satz 8.4.1. Für Real- und Imaginärteil heißt das

$$x = r \cos \varphi =: f_1(r, \varphi),$$

$$y = r \sin \varphi =: f_2(r, \varphi).$$

f_1 und f_2 sind aus $C^1((0, +\infty) \times \mathbb{R})$. Für die Abbildung $f := (f_1, f_2) : (0, +\infty) \times \mathbb{R} \rightarrow \mathbb{R}^2$ folgt

$$Df(r, \varphi) = \begin{pmatrix} f_{1r}(r, \varphi) & f_{1\varphi}(r, \varphi) \\ f_{2r}(r, \varphi) & f_{2\varphi}(r, \varphi) \end{pmatrix} = \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix},$$

und für die Jacobideterminante erhalten wir

$$J_f(r, \varphi) = r(\cos^2 \varphi + \sin^2 \varphi) = r > 0.$$

Wir schließen mit zwei speziellen Begriffen für Vektorfelder.

Definition 10.1.10

Ist $\Omega \subset \mathbb{R}^n$ und $f = (f_1, \dots, f_n) \in C^1(\Omega, \mathbb{R}^n)$ ein Vektorfeld, so erklären wir die *Divergenz* $\text{div } f : \Omega \rightarrow \mathbb{R} \in C^0(\Omega)$ gemäß

$$\text{div } f(x) := D_1 f_1(x) + \dots + D_n f_n(x) = \sum_{j=1}^n D_j f_j(x) = \langle \nabla, f \rangle(x), \quad x \in \Omega.$$

Für $n = 3$ erklären wir auch die *Rotation* $\text{rot } f : \Omega \rightarrow \mathbb{R}^3 \in C^0(\Omega, \mathbb{R}^3)$ gemäß

$$\text{rot } f := (D_2 f_3 - D_3 f_2, D_3 f_1 - D_1 f_3, D_1 f_2 - D_2 f_1) = \nabla \times f \quad \text{in } \Omega.$$

10.2. Mittelwertsatz und Richtungsableitung

Wir erinnern an den Mittelwertsatz (der Differentialrechnung) für reellwertige Funktionen einer reellen Veränderlichen, Satz 7.3.8, den wir jetzt verallgemeinern:

Satz 10.2.1: Mittelwertsatz

Es sei $\Omega \subset \mathbb{R}^n$ offen, $f \in C^1(\Omega)$ eine reellwertige Funktion und zwei verschiedene Punkte $x, y \in \Omega$ seien so gewählt, dass für die Verbindungsstrecke

$$[x, y] := \{\lambda x + (1 - \lambda)y : \lambda \in [0, 1]\}$$

zwischen x und y gilt $[x, y] \subset \Omega$. Dann gibt es ein $z \in (x, y) := [x, y] \setminus \{x, y\}$ mit der Eigenschaft

$$f(y) - f(x) = \langle \nabla f(z), y - x \rangle. \quad (10.2.1)$$

Beweis. Für die Funktion

$$\varphi(t) := f(x + t(y - x)), \quad t \in [0, 1].$$

gilt nach Satz 10.1.6 und Satz 6.2.6 $\varphi \in C^1((0, 1)) \cap C^0([0, 1])$ und - siehe auch Bemerkung 10.1.8 -

$$\varphi'(t) = \langle \nabla f(x + t(y - x)), y - x \rangle, \quad t \in (0, 1).$$

Andererseits liefert der Mittelwertsatz für Funktionen einer Veränderlichen, Satz 7.3.8,

$$\varphi(1) - \varphi(0) = \varphi'(\xi)$$

mit einem $\xi \in (0, 1)$. Setzen wir $z := x + \xi(y - x) \in (x, y)$, so folgt

$$f(y) - f(x) = \varphi(1) - \varphi(0) = \varphi'(\xi) = \langle \nabla f(z), y - x \rangle,$$

wie behauptet. □

Für Funktionen $f : I \rightarrow \mathbb{R}^d \in C^1(I, \mathbb{R}^d)$ auf einem Intervall $I \subset \mathbb{R}$ gilt bekanntlich (Folgerung 7.3.11 (c)):

$$f' = 0 \text{ auf } I \Leftrightarrow f = \text{const auf } I.$$

Satz 10.2.1 soll angewendet werden, um diese Aussage auf Funktionen mehrerer Veränderlicher zu verallgemeinern. Dazu benötigen wir noch die grundlegende

Definition 10.2.2

- Eine Menge $M \subset \mathbb{R}^n$ heißt *(bogenweise) zusammenhängend*, wenn für je zwei Punkte $x, y \in M$ ein stetiger Weg $\varphi \in C^0([0, 1], \mathbb{R}^n)$ existiert mit

$$\varphi([0, 1]) \subset M, \quad \varphi(0) = x, \quad \varphi(1) = y.$$

- Eine offene, zusammenhängende Menge im \mathbb{R}^n heißt *Gebiet* und wird meist mit G bezeichnet.

Bemerkung 10.2.3. Intervalle sind offenbar zusammenhängend. Die Richtung „ \Rightarrow “ obiger Aussage wird z.B. auf der Vereinigung $I_1 \cup I_2$ zweier disjunkter Intervalle I_1, I_2 falsch.

Satz 10.2.4

Für eine Funktion $f \in C^1(G, \mathbb{R}^d)$ auf einem Gebiet $G \subset \mathbb{R}^n$ gilt

$$Df = 0 \text{ in } G \Leftrightarrow f = \text{const in } G.$$

Beweis. Falls f konstant ist in G , verschwinden alle partiellen Ableitungen, d.h. $Df = 0$ in G . Für den Beweis der Umkehrung können wir uns offenbar auf den Fall $d = 1$ beschränken, d.h. wir setzen $f \in C^1(G)$ und $\nabla f = 0$ in G voraus. Der Beweis der Konstanz von f erfolgt in zwei Schritten:

1. *Lokale Konstanz.* Sei $x \in G$ fixiert und sei ein $\varepsilon > 0$ mit $B_\varepsilon(x) \subset G$ gewählt. Zu beliebigem $y \in B_\varepsilon(x)$ gilt dann $[x, y] \subset B_\varepsilon(x) \subset G$, und nach dem Mittelwertsatz folgt mit einem $z \in (x, y)$:

$$f(y) - f(x) = \langle \nabla f(z), y - x \rangle = 0,$$

also $f = \text{const in } B_\varepsilon(x)$.

2. *Globale Konstanz.* Mit einer Kontinuitätsmethode \rightarrow *Ergänzungen*. □

Wir wollen die rechte Seite von (10.2.1) noch ein wenig interpretieren. Dazu verallgemeinern wir den Begriff der partiellen Ableitung:

Definition 10.2.5: Richtungsableitung

Seien $f : \Omega \rightarrow \mathbb{R}$ gegeben und $x \in \Omega \subset \mathbb{R}^n$ gewählt. Zu beliebigem $a \in \mathbb{R}^n$ mit $\|a\| = 1$ erklären wir dann die *Richtungsableitung* von f an der Stelle x in Richtung a gemäß

$$\frac{\partial f}{\partial a}(x) := \lim_{h \rightarrow 0} \frac{f(x + ha) - f(x)}{h},$$

falls dieser Grenzwert existiert.

Bemerkung 10.2.6. (a) Mit dem j -ten Standardeinheitsvektor $a = e_j$ haben wir offenbar $\frac{\partial f}{\partial e_j}(x) = D_j f(x)$; siehe Definition 10.1.1. Partielle Ableitungen sind also spezielle Richtungsableitungen.

- (b) Existiert für ein $a \in S^{n-1} = \{\xi \in \mathbb{R}^n : \|\xi\| = 1\}$ die Richtungsableitung $\frac{\partial f}{\partial a}(x)$, so existiert auch $\frac{\partial f}{\partial(-a)}(x)$ und es gilt

$$\frac{\partial f}{\partial(-a)}(x) = -\frac{\partial f}{\partial a}(x).$$

Satz 10.2.7

Ist $f \in C^1(\Omega)$, so existiert $\frac{\partial f}{\partial a}(x)$ für alle $x \in \Omega$ und alle Richtungsvektoren $a \in S^{n-1}$, und es gilt

$$\frac{\partial f}{\partial a}(x) = \langle \nabla f(x), a \rangle.$$

Beweis. Zu festem $x \in \Omega$ existiert ein $\varepsilon > 0$ mit $B_\varepsilon(x) \subset \Omega$. Wir betrachten dann $\varphi(t) := f(x+ta)$, $t \in (-\varepsilon, \varepsilon)$ und $a \in S^{n-1}$, und folgern aus Satz 10.1.6: $\varphi \in C^1((-\varepsilon, \varepsilon))$ sowie

$$\varphi'(t) = \langle \nabla f(x+ta), a \rangle, \quad t \in (-\varepsilon, \varepsilon).$$

Insbesondere für $t = 0$ erhalten wir

$$\langle \nabla f(x), a \rangle = \varphi'(0) = \lim_{h \rightarrow 0} \frac{\varphi(h) - \varphi(0)}{h} \stackrel{\text{Def. 10.2.5}}{=} \frac{\partial f}{\partial a}(x),$$

wie behauptet. □

Bemerkung 10.2.8. ▶ Die rechte Seite in (10.2.1) lässt sich also als $\frac{\partial f}{\partial a}(z) \cdot \|y - x\|$ mit dem Richtungsvektor $a = \frac{y-x}{\|y-x\|}$ schreiben.

- ▶ Die Richtungsableitung $\frac{\partial f}{\partial a}(x)$ beschreibt den Anstieg der Funktion f in x , eingeschränkt auf das Segment $\{x + ta : t \in (-\varepsilon, \varepsilon)\}$ mit kleinem $\varepsilon > 0$ und einem $a \in S^{n-1}$. Dieser ist i.a. für jede Richtung a unterschiedlich groß. Die Richtung von $\text{grad } f$ ist dabei ausgezeichnet:

Folgerung 10.2.9

Sei $f \in C^1(\Omega)$ und $x \in \Omega$ mit $\nabla f(x) \neq 0$ gewählt. Mit $v := \|\nabla f(x)\|^{-1} \nabla f(x) \in S^{n-1}$ gilt dann

$$\frac{\partial f}{\partial(-v)}(x) < \frac{\partial f}{\partial a}(x) < \frac{\partial f}{\partial v}(x) \quad \text{für alle } a \in S^{n-1} \setminus \{\pm v\}.$$

D.h. $\nabla f(x)$ zeigt in die Richtung des größten, $-\nabla f(x)$ in die Richtung des kleinsten Anstiegs von f in x .

Beweis. → Ergänzungen □

Bemerkung 10.2.10. ▶ Falls $\nabla f(x) = 0$ in einem Punkt $x \in \Omega$ gilt, verschwinden nach Satz 10.2.7 dort alle Richtungsableitungen.

- ▶ Der Begriff der Richtungsableitung lässt sich offenbar direkt auf Abbildungen $f : \Omega \rightarrow \mathbb{R}^d$ erweitern. Satz 10.2.7 bleibt gültig, allerdings mit der Darstellung

$$\frac{\partial f}{\partial a}(x) = Df(x)a, \quad (10.2.2)$$

wobei man $\frac{\partial f}{\partial a}$ und a als Spaltenvektoren auffasst. Die geometrische Interpretation aus Folgerung 10.2.9 verliert aber ihren Sinn.

Abschließend halten wir noch eine Art „integrale Version“ von Satz 10.2.1 für vektorwertige Funktionen fest:

Satz 10.2.11: HADAMARDS Lemma

Sind $f \in C^1(\Omega, \mathbb{R}^d)$ und $x, y \in \Omega$ mit $x \neq y$ und $[x, y] \subset \Omega$ gewählt, so folgt

$$f(y) - f(x) = A \circ (y - x).$$

Dabei sind $f(y) - f(x)$, $y - x$ als Spaltenvektoren aufzufassen und die Matrix

$$A := \int_0^1 Df(x + t(y - x)) dt$$

wurde erklärt (komponentenweise Integration).

Beweis. → Ergänzungen □

10.3. Die totale Ableitung

Wir wollen noch einen weiteren Ableitungsbegriff einführen, der geometrisch als lineare Approximation motiviert ist:

Definition 10.3.1: Totale Differenzierbarkeit

Seien $\Omega \subset \mathbb{R}^n$ offen und $x \in \Omega$ gewählt. Eine Funktion $f : \Omega \rightarrow \mathbb{R}^d$ heißt in $x \in \Omega$ *total differenzierbar*, wenn eine lineare Abbildung $L : \mathbb{R}^n \rightarrow \mathbb{R}^d \in \text{Lin}(\mathbb{R}^n, \mathbb{R}^d)$ und ein $r > 0$

mit $B_r(x) \subset \Omega$ so existieren, dass gilt

$$f(x+h) = f(x) + Lh + R(h) \quad \text{für alle } h \in B_r(0). \quad (10.3.1)$$

Hierbei gelte für das Restglied $R = R(h) : B_r(0) \rightarrow \mathbb{R}^d$ die Relation

$$R(h) = o(\|h\|) \quad \text{für } h \rightarrow 0. \quad (10.3.2)$$

Die Abbildung L heißt dann *totale Ableitung* oder *Differential* von f an der Stelle x und wird mit $df(x)$ bezeichnet.

Bemerkung 10.3.2. (a) Die lineare Abbildung $L : \mathbb{R}^n \rightarrow \mathbb{R}^d$ ist durch (10.3.1), (10.3.2) eindeutig bestimmt (\rightarrow Ergänzungen).

(b) Wir können Formeln (10.3.1), (10.3.2) äquivalent schreiben als

$$\begin{aligned} f(x+h) &= f(x) + df(x)h + \|h\|\varepsilon(h) \quad \text{für alle } h \in B_r(0) \\ \text{mit } \varepsilon &= \varepsilon(h) : B_r(0) \rightarrow \mathbb{R}^d, \quad \varepsilon(h) = o(1) \text{ für } h \rightarrow 0, \end{aligned} \quad (10.3.3)$$

indem wir setzen

$$\varepsilon(h) := \begin{cases} \|h\|^{-1}R(h), & h \neq 0 \\ 0, & h = 0 \end{cases}.$$

Eine (10.3.3) entsprechende Darstellung haben wir in Satz 7.1.3 und Bemerkung 7.1.4 für Funktionen einer Veränderlichen als äquivalent zur Differenzierbarkeit erkannt. Wir werden unten sehen, dass die totale Differenzierbarkeit einer Funktion mehrerer Veränderlicher *nicht* äquivalent zur partiellen Differenzierbarkeit ist.

Satz 10.3.3

Ist $f : \Omega \rightarrow \mathbb{R}^d$ in $x \in \Omega$ total differenzierbar, so ist f auch stetig in x .

Beweis. Grenzübergang $h \rightarrow 0$ in (10.3.3) liefert sofort $\lim_{h \rightarrow 0} f(x+h) = f(x)$, wie behauptet. \square

Ein Teil des Zusammenhangs zwischen totaler und partieller Differenzierbarkeit ist enthalten im folgenden

Satz 10.3.4

Ist $f : \Omega \rightarrow \mathbb{R}^d$ in $x \in \Omega$ total differenzierbar, so existieren für alle $a \in S^{n-1}$ die Richtungs-

ableitungen $\frac{\partial f}{\partial a}(x)$, und es gilt

$$df(x)a = \frac{\partial f}{\partial a}(x). \quad (10.3.4)$$

Insbesondere existieren alle partiellen Ableitungen $D_1 f(x), \dots, D_n f(x)$ mit

$$df(x)h = \sum_{j=1}^n D_j f(x)h_j = Df(x)h \quad \text{für alle } h = (h_1, \dots, h_n) \in \mathbb{R}^n. \quad (10.3.5)$$

Speziell für $d = 1$ halten wir noch fest

$$df(x)h = \langle \nabla f(x), h \rangle \quad \text{für alle } h \in \mathbb{R}^n.$$

Beweis. Wir setzen $h = ta \in B_r(0)$, $t \in (-r, r)$, in (10.3.1) ein und erhalten aus (10.3.2):

$$\left\| \frac{f(x+ta) - f(x)}{t} - df(x)a \right\| = \left\| \frac{R(ta)}{t} \right\| \rightarrow 0 \quad (t \rightarrow 0);$$

also existiert $\frac{\partial f}{\partial a}(x)$ gemäß Definition 10.2.5 und es gilt (10.3.4); siehe auch Bemerkung 10.2.10. Für $a = e_j$ folgt speziell die Existenz von $D_j f(x) = \frac{\partial f}{\partial e_j}(x) = df(x)e_j$, $j = 1, \dots, n$, und wir finden für beliebiges $h \in \mathbb{R}^n$:

$$df(x)h = df(x)\left(\sum_{j=1}^n h_j e_j\right) = \sum_{j=1}^n h_j df(x)e_j = \sum_{j=1}^n h_j D_j f(x) = Df(x)h,$$

wie behauptet. □

Geometrische Interpretation. Ist $f : \Omega \rightarrow \mathbb{R}$ in $x^0 \in \Omega$ total differenzierbar, so erklären wir die affin-lineare Funktion

$$\varphi(x) := f(x^0) + df(x^0)(x - x^0) = f(x^0) + \langle \nabla f(x^0), x - x^0 \rangle, \quad x \in \mathbb{R}^n.$$

Der Graph $T := \text{graph } \varphi$ beschreibt eine *Hyperebene* im \mathbb{R}^{n+1} – d.h. einen n -dimensionalen affinen Unterraum –, die durch $(x^0, f(x^0))$ und senkrecht zum Vektor

$$v := (-\nabla f(x^0), 1)$$

verläuft. Dabei ist φ und damit T durch die Forderung

$$f(x) - \varphi(x) = o(\|x - x^0\|) \quad \text{für } x \rightarrow x^0$$

eindeutig festgelegt; siehe Bemerkung 10.3.2 (b). T approximiert also graph f nahe x^0 von erster Ordnung, entsprechend der Tangente bei differenzierbaren Funktionen einer Veränderlichen. T heißt die *Tangentialebene* an f (genauer an graph f) im Punkt x^0 . Der zu T senkrechte Vektor v wird als *Normalenvektor* von f in x^0 bezeichnet.

In Abschnitt 10.5 werden wir Approximationen höherer Ordnung durch Taylorpolynome gewinnen.

Bemerkung 10.3.5. Die Umkehrung von Satz 10.3.4 gilt nicht: Eine Funktion, die in einem Punkt alle Richtungsableitungen besitzt, muss dort nicht stetig und somit nach Satz 10.3.3 auch nicht total differenzierbar sein; dazu das folgende

Beispiel 10.3.6. Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ erklärt als

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

Wir zeigen, dass

$$\frac{\partial f}{\partial a}(0, 0) = \lim_{t \rightarrow 0} \frac{f(ta_1, ta_2) - f(0, 0)}{t} = \lim_{t \rightarrow 0} \frac{f(ta_1, ta_2)}{t}$$

für alle $a = (a_1, a_2) \in S^1$ existiert:

- Für $a = (0, \pm 1)$ haben wir $f(0, \pm t) = 0$ für alle $t \in \mathbb{R} \setminus \{0\}$ und folglich $\frac{\partial f}{\partial a}(0, 0) = 0$.
- Für $a \neq (0, \pm 1)$ und somit $a_1 \neq 0$ erhalten wir

$$\frac{\partial f}{\partial a}(0, 0) = \lim_{t \rightarrow 0} \frac{t^3 a_1 a_2^2}{t^3 (a_1^2 + t^2 a_2^4)} = \frac{a_2^2}{a_1}.$$

Andererseits ist f wegen $f(y^2, y) = \frac{y^4}{y^4+y^4} = \frac{1}{2}$ für alle $y \in \mathbb{R} \setminus \{0\}$ und $f(0, 0) = 0$ im Punkt $(0, 0)$ unstetig und folglich nicht total differenzierbar.

Unter einer zusätzlichen Voraussetzung können wir jedoch aus der partiellen die totale Differenzierbarkeit folgern:

Satz 10.3.7

Gehört $f : \Omega \rightarrow \mathbb{R}^d$ zur Klasse $C^1(\Omega, \mathbb{R}^d)$, so ist f total differenzierbar in Ω .

Beweis. → Ergänzungen. □

10.4. Partielle Ableitungen höherer Ordnung, der Satz von Schwarz

Wir erklären partielle Ableitungen höherer Ordnung wieder induktiv; vgl. Abschnitt 7.4. Seien $\Omega \subset \mathbb{R}^n$ offen und $f : \Omega \rightarrow \mathbb{R}^d$ gegeben.

- $s = 1$: Sei $j \in \{1, \dots, n\}$. Ist f in $x^0 \in \Omega$ partiell nach x_j differenzierbar, so ist $D_j f(x^0)$ gemäß Definition 10.1.1 erklärt.

- $s = 2$: Seien $j_1, j_2 \in \{1, \dots, n\}$. Ist f in allen Punkten $x \in \Omega$ nach x_{j_1} partiell differenzierbar und ist $D_{j_1} : \Omega \rightarrow \mathbb{R}^d$ in $x^0 \in \Omega$ nach x_{j_2} partiell differenzierbar, so setzen wir

$$D_{j_2} D_{j_1} f(x^0) := D_{j_2}(D_{j_1} f)(x^0) = \lim_{h \rightarrow 0} \frac{D_{j_1} f(x^0 + h e_{j_2}) - D_{j_1} f(x^0)}{h}.$$

- $s \rightarrow s + 1$: Seien $j_1, \dots, j_s, j_{s+1} \in \{1, \dots, n\}$. Existiert $D_{j_s} \dots D_{j_1} f : \Omega \rightarrow \mathbb{R}^d$ und ist die Funktion in $x^0 \in \Omega$ nach $x_{j_{s+1}}$ partiell differenzierbar, so setzen wir

$$D_{j_{s+1}} D_{j_s} \dots D_{j_1} f(x^0) := D_{j_{s+1}}(D_{j_s} \dots D_{j_1} f)(x^0).$$

Definition 10.4.1

Seien $\Omega \subset \mathbb{R}^n$ und $f : \Omega \rightarrow \mathbb{R}^d$ gegeben. Existiert für ein $s \in \mathbb{N}$ und für $j_1, \dots, j_s \in \{1, \dots, n\}$ die Größe $D_{j_s} \dots D_{j_1} f(x^0)$, so heißt sie *partielle Ableitung s-ter Ordnung* oder *s-te partielle Ableitung* von f in $x^0 \in \Omega$. Alternative Schreibweisen sind

$$D_{j_s} \dots D_{j_1} f(x^0) = \frac{\partial^s f}{\partial x_{j_s} \dots \partial x_{j_1}}(x^0) = f_{x_{j_1} \dots x_{j_s}}(x^0)$$

Existieren alle s -ten partiellen Ableitung von f in x^0 , so schreiben wir

$$D^s f(x^0) = (D_{j_s} \dots D_{j_1} f(x^0))_{j_1, \dots, j_s=1, \dots, n}$$

für die *s-te Ableitung* von f in x^0 .

Bemerkung 10.4.2. (a) Wir können - wenn existent - $D^s f(x^0) \in \mathbb{R}^{d \cdot n^s}$ als Punkt im Euklidischen Raum der Dimension $d \cdot n^s$ auffassen, wenn wir eine Anordnung der Komponenten festlegen.

- (b) Für $f : \Omega \rightarrow \mathbb{R}$ wird $D^2 f(x^0)$ als quadratische Matrix aufgefasst:

$$H_f(x^0) = D^2 f(x^0) = \begin{pmatrix} f_{x_1 x_1}(x^0) & \dots & f_{x_1 x_n}(x^0) \\ \vdots & \ddots & \vdots \\ f_{x_n x_1}(x^0) & \dots & f_{x_n x_n}(x^0) \end{pmatrix}$$

wird als *HESSE-Matrix* von f bezeichnet.

- (c) Für $f : \Omega \rightarrow \mathbb{R}^d$ wird die erste Ableitung als $D^1 f(x^0) = Df(x^0)$ aufgefasst, sie stimmt also mit der Funktionalmatrix überein. Wir setzen noch $D^0 f(x^0) := f(x^0)$ für die *nullte Ableitung* von f in x^0 .

Definition 10.4.3

Für $s \in \mathbb{N}_0$ erklären wir den Raum der *s-mal stetig differenzierbaren Funktionen* $C^s(\Omega, \mathbb{R}^d)$ als den Vektorraum der Funktionen $f : \Omega \rightarrow \mathbb{R}^d$, deren Ableitungen $Df, \dots, D^s f$ auf Ω existieren und für die $D^s f : \Omega \rightarrow \mathbb{R}^{d \cdot n^s}$ stetig ist. Der Raum der *unendlich oft differenzierbaren Funktionen* ist dann gegeben durch

$$C^\infty(\Omega, \mathbb{R}^d) := \bigcap_{s \in \mathbb{N}_0} C^s(\Omega, \mathbb{R}^d).$$

Für $d = 1$ bzw. $d = 2$ schreiben wir auch $C^s(\Omega) := C^s(\Omega, \mathbb{R})$ bzw. $C^s(\Omega, \mathbb{C}) := C^s(\Omega, \mathbb{R}^2)$.

Bemerkung 10.4.4. ▶ Gemäß Satz 10.1.4 sind *alle* Ableitungen $D^0 f, \dots, D^s f$ einer Funktion $f \in C^s(\Omega, \mathbb{R}^d)$ stetig.

- ▶ Linearkombinationen, Produkte und Kompositionen von s -mal stetig differenzierbaren Funktionen sind wieder s -mal stetig differenzierbar (\rightarrow Nachrechnen).

Beispiel 10.4.5. Für die Funktion

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

gilt $f_{xy}(0, 0) = -1 \neq 1 = f_{yx}(0, 0)$.

Denn: Für die ersten partiellen Ableitungen berechnen wir

$$f_x(x, y) = \frac{x^4 y + 4x^2 y^3 - y^5}{(x^2 + y^2)^2}, \quad f_y(x, y) = \frac{x^5 - 4x^3 y^2 - x y^4}{(x^2 + y^2)^2}, \quad (x, y) \in \mathbb{R}^2 \setminus \{(0, 0)\},$$

sowie

$$f_x(0, 0) = \lim_{h \rightarrow 0} \frac{f(h, 0)}{h} = 0 = \lim_{h \rightarrow 0} \frac{f(0, h)}{h} = f_y(0, 0).$$

Also ist f in \mathbb{R}^2 einmal partiell differenzierbar. Im Nullpunkt erhalten wir weiter

$$\begin{aligned} f_{xy}(0, 0) &= \lim_{h \rightarrow 0} \frac{f_x(0, h)}{h} = \lim_{h \rightarrow 0} \frac{-h}{h} = -1, \\ f_{yx}(0, 0) &= \lim_{h \rightarrow 0} \frac{f_y(h, 0)}{h} = \lim_{h \rightarrow 0} \frac{h}{h} = 1, \end{aligned}$$

wie behauptet □

Ist also $f : \Omega \rightarrow \mathbb{R}^d$ gegeben und existieren $D_j D_k f$ und $D_k D_j f$ für gewisse $j \neq k$, so ist i.A. *nicht* $D_j D_k f = D_k D_j f$. Unter gewissen Voraussetzungen ist die Reihenfolge der zweiten Ableitungen aber doch vertauschbar, wir beginnen mit dem

Hilfssatz 10.4.6

Es sei $B_\delta := B_\delta(0, 0) \subset \mathbb{R}^2$ die Kreisscheibe vom Radius $\delta > 0$ um den Nullpunkt und $\varphi = \varphi(y, z) : B_\delta \rightarrow \mathbb{R}$ sei gegeben. Für φ sollen die partiellen Ableitungen φ_y , φ_z und φ_{yz} in B_δ existieren. Dann gibt es zu beliebigen $h, k \neq 0$ mit $(h, k) \in B_\delta$ einen Punkt $(\xi, \eta) \in B_\delta$, so dass gilt

$$\varphi_{yz}(\xi, \eta) = \frac{\varphi(h, k) - \varphi(h, 0) - \varphi(0, k) + \varphi(0, 0)}{hk}. \quad (10.4.1)$$

Beweis. Mit dem Mittelwertsatz (\rightarrow Ergänzungen). \square

Bemerkung 10.4.7. Die rechte Seite in (10.4.1) kann als *Differenzenquotient zweiter Ordnung* im Punkt $(0, 0)$ aufgefasst werden.

Satz 10.4.8: H. A. SCHWARZ

Seien $\Omega \subset \mathbb{R}^n$ offen und $f : \Omega \rightarrow \mathbb{R}^d$ gegeben. Für ein $x^0 \in \Omega$ und $i, j \in \{1, \dots, n\}$ sollen die partiellen Ableitungen $D_i f$, $D_j f$ und $D_j D_i f$ in einer Kugel $B_r(x^0) \subset \Omega$ existieren, und $D_j D_i f$ sei stetig in x^0 . Dann existiert auch $D_i D_j f(x^0)$ und es gilt

$$D_i D_j f(x^0) = D_j D_i f(x^0).$$

Beweis. O.B.d.A. seien $d = 1$ und $i \neq j$. Wir betrachten die Funktion

$$\varphi(y, z) := f(x^0 + ye_i + ze_j) \quad \text{für } (y, z) \in B_r := B_r(0, 0).$$

Dann existieren φ_y , φ_z und φ_{yz} in B_r , und φ_{yz} ist stetig in $(0, 0)$. Zu gegebenem $\varepsilon > 0$ können wir daher $\delta = \delta(\varepsilon) \in (0, r]$ so klein wählen, dass

$$|\varphi_{yz}(y, z) - \varphi_{yz}(0, 0)| < \varepsilon \quad \text{für alle } (y, z) \in B_\delta \quad (10.4.2)$$

erfüllt ist. Wir wenden nun Hilfssatz 10.4.6 an: Zu beliebigem $(h, k) \in B_\delta$ mit $h, k \neq 0$ ergeben (10.4.1) und (10.4.2)

$$\begin{aligned} & \left| \frac{1}{h} \left(\frac{\varphi(h, k) - \varphi(h, 0)}{k} - \frac{\varphi(0, k) - \varphi(0, 0)}{k} \right) - \varphi_{yz}(0, 0) \right| \\ &= \left| \frac{\varphi(h, k) - \varphi(h, 0) - \varphi(0, k) + \varphi(0, 0)}{hk} - \varphi_{yz}(0, 0) \right| \\ &= |\varphi_{yz}(\xi, \eta) - \varphi_{yz}(0, 0)| < \varepsilon \end{aligned}$$

mit einem $(\xi, \eta) \in B_\delta$. Grenzübergang $k \rightarrow 0$ liefert

$$\left| \frac{1}{h} [\varphi_z(h, 0) - \varphi_z(0, 0)] - \varphi_{yz}(0, 0) \right| \leq \varepsilon \quad \text{für } h \in (-\delta, \delta) \setminus \{0\}.$$

Also existiert

$$\begin{aligned} D_i D_j f(x^0) &= \varphi_{zy}(0, 0) = \lim_{h \rightarrow 0} \frac{1}{h} [\varphi_z(h, 0) - \varphi_z(0, 0)] \\ &= \varphi_{yz}(0, 0) = D_j D_i f(x^0), \end{aligned}$$

wie behauptet. □

Bemerkung 10.4.9. Insbesondere bei einer Funktion $f \in C^2(\Omega, \mathbb{R}^d)$ können daher immer die Ableitungen – genauer die Reihenfolge der Ableitungen – vertauscht werden, die HESSE-Matrix ist dann also symmetrisch; siehe Bemerkung 10.4.2 (b).

Entsprechendes gilt auch für die höheren Ableitungen einer Funktion $f \in C^s(\Omega, \mathbb{R}^d)$ mit $s > 2$, wie man durch vollständige Induktion leicht sieht (\rightarrow Nachrechnen). Dies rechtfertigt auch die folgende

Notation (Multiindices). Zu einem sogenannten *Multiindex* $\alpha := (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n$, d.h. $\alpha_j \in \mathbb{N}_0$ für alle $j = 1, \dots, n$, erklären wir dessen *Länge* $|\alpha|$ gemäß

$$|\alpha| := \alpha_1 + \dots + \alpha_n.$$

Dann schreiben wir für ein $f \in C^{|\alpha|}(\Omega, \mathbb{R}^d)$ abkürzend

$$D_j^{\alpha_j} f := (D_j)^{\alpha_j} f := \underbrace{D_j D_j \dots D_j}_{\alpha_j\text{-mal}} f$$

und

$$D^\alpha f := (D_1)^{\alpha_1} (D_2)^{\alpha_2} \dots (D_n)^{\alpha_n} f = D_1^{\alpha_1} D_2^{\alpha_2} \dots D_n^{\alpha_n} f.$$

Das bedeutet, f wird α_j -mal nach x_j abgeleitet, wobei die Reihenfolge der Differentiation nach Satz 10.4.8 keine Rolle spielt.

Beispiel 10.4.10. Für eine Funktion $f : \Omega \rightarrow \mathbb{R} \in C^2(\Omega)$, $\Omega \subset \mathbb{R}^n$, erklärt man den *LAPLACE-Operator*

$$\Delta f(x) := D_1^2 f(x) + \dots + D_n^2 f(x) = \sum_{k=1}^n f_{x_k x_k}(x), \quad x \in \Omega.$$

Dieser ordnet jeder Funktion $f \in C^2(\Omega)$ eine Funktion $\Delta f \in C^0(\Omega)$ zu, weshalb man auch $\Delta : C^2(\Omega) \rightarrow C^0(\Omega)$ schreibt. Der Laplace-Operator ist (wie der Nabla-Operator $\nabla : C^1(\Omega) \rightarrow$

$C^0(\Omega, \mathbb{R}^n)$) ein Beispiel eines – und zwar eines wichtigen! – *Differentialoperators*. Die zugehörige Gleichung

$$\Delta f(x) = 0 \quad \text{für alle } x \in \Omega$$

heißt *Laplacegleichung* und eine Lösung f nennt man *harmonische Funktion*. Die Laplacegleichung ist eine der wichtigsten *partiellen Differentialgleichungen*; dies sind Gleichungen zwischen Funktionen mehrerer Veränderlicher und ihren partiellen Ableitungen. Im Gegensatz hierzu werden Gleichungen zwischen Funktionen einer Variablen und ihren gewöhnlichen Ableitungen als *gewöhnliche Differentialgleichungen* bezeichnet. Die meisten Prozesse in der Natur lassen sich (zumindest näherungsweise) nur mit Hilfe von Differentialgleichungen beschreiben.

10.5. Die TAYLORSche Formel im \mathbb{R}^n

Zum Abschluss dieses Kapitels geben wir ein Analogon zur Taylorschen Formel für Funktionen einer Veränderlichen an, siehe Abschnitt 9.7. Dazu erklären wir noch Differentiale höherer Ordnung, wobei wir von der Darstellung (10.3.5) ausgehen: Für $f \in C^1(\Omega, \mathbb{R}^d)$, $\Omega \subset \mathbb{R}^n$ offen, gilt

$$df(x)h = \sum_{j=1}^n D_j f(x) h_j \quad \text{für } x \in \Omega, h = (h_1, \dots, h_n) \in \mathbb{R}^n.$$

Definition 10.5.1

Für $f \in C^s(\Omega, \mathbb{R}^d)$ für ein $s \in \mathbb{N}$ erklären wir das *s-te Differential* oder das *Differential der Ordnung s* gemäß

$$d^s f(x)(h^1, \dots, h^s) := \sum_{j_1, \dots, j_s=1}^n D_{j_s} \dots D_{j_1} f(x) h_{j_1}^1 \cdot \dots \cdot h_{j_s}^s$$

für $x \in \Omega$ und $h^l = (h_1^l, \dots, h_n^l) \in \mathbb{R}^n$ mit $l = 1, \dots, s$.

Ist speziell $h^1 = h^2 = \dots = h^s =: h$, so schreiben wir abkürzend

$$d^s f(x)h := d^s f(x)(h, \dots, h) = \sum_{j_1, \dots, j_s=1}^n D_{j_s} \dots D_{j_1} f(x) h_{j_1} \cdot \dots \cdot h_{j_s} \quad (10.5.1)$$

für $x \in \Omega$ und $h = (h_1, \dots, h_n) \in \mathbb{R}^n$.

Bemerkung 10.5.2. ► Das *s-te Differential* ist (für festes x) eine multilineare Abbildung, also linear in jedem Eintrag h^l , $l = 1, \dots, s$. Wegen des Satzes von Schwarz, Satz 10.4.8, ist sie symmetrisch, d.h. invariant unter Permutationen der Einträge h^1, \dots, h^s .

- Offenbar gilt $d^1 f(x)h = df(x)h$ für $f \in C^1(\Omega, \mathbb{R}^d)$, $x \in \Omega$ und $h \in \mathbb{R}^n$. Aus technischen Gründen setzen wir noch

$$d^0 f(x)h := f(x), \quad x \in \Omega, \quad h \in \mathbb{R}^n.$$

- Differentiale höherer Ordnung lassen sich auch direkt durch Verallgemeinerung von Definition 10.3.1 erklären; unter den Differenzierbarkeitsvoraussetzungen aus Definition 10.5.1 stimmen diese dann mit den multilinearen Abbildungen unserer Definition überein.

Beispiel 10.5.3. Mit $s = 2$, $h := h^1 \in \mathbb{R}^n$, $g := h^2 \in \mathbb{R}^n$ haben wir für $f \in C^2(\Omega, \mathbb{R}^d)$:

$$d^2 f(x)(h, g) = \sum_{i,j=1}^n D_j D_i f(x) h_i g_j. \quad (10.5.2)$$

Insbesondere für $h = e_i$, $g = e_j$, also i -ter bzw. j -ter Standardeinheitsvektor, folgt

$$d^2 f(x)(e_i, e_j) = D_j D_i f(x).$$

Allgemeiner gilt: Ist $h^l = e_{i_l}$ der i_l -te Standardeinheitsvektor für $l = 1, \dots, s$, so folgt für $f \in C^s(\Omega, \mathbb{R}^d)$:

$$d^s f(x)(e_{i_1}, \dots, e_{i_s}) = f_{x_{i_1} \dots x_{i_s}}(x).$$

Für die Verbindungsstrecke zwischen zwei Punkten $x^1, x^2 \in \mathbb{R}^n$ schreiben wir wieder

$$[x^1, x^2] = \{x^1 + t(x^2 - x^1) : t \in [0, 1]\}, \quad (x^1, x^2) := [x^1, x^2] \setminus \{x^1, x^2\}.$$

Satz 10.5.4: TAYLORformel im \mathbb{R}^n

Seien $s \in \mathbb{N}_0$, $f \in C^{s+1}(\Omega)$ und $x^0 \in \Omega$ gegeben. Für jedes $x \in \Omega$ mit $[x^0, x] \subset \Omega$ gilt dann die *TAYLORSche Formel*

$$f(x) = T_{s,x^0} f(x) + R_{s,x^0} f(x)$$

mit dem s -ten *Taylorpolynom* $T_{s,x^0} f$ von f zum Entwicklungspunkt x^0 ,

$$T_{s,x^0} f(x) := \sum_{\sigma=0}^s \frac{1}{\sigma!} d^\sigma f(x^0)(x - x^0),$$

und dem s -ten *Restglied* $R_{s,x^0} f$ von f zu x^0 ,

$$R_{s,x^0} f(x) = \frac{1}{(s+1)!} d^{s+1} f(y)(x - x^0).$$

Dabei ist $y = x^0 + \vartheta(x - x^0) \in (x^0, x)$ für ein $\vartheta = \vartheta(x) \in (0, 1)$ geeignet gewählt.

Beweis. Mit $h := x - x^0$ gilt $x^0 + th \in [x^0, x] \subset \Omega$ für $t \in [0, 1]$. Die Funktion $\phi(t) := f(x^0 + th)$, $t \in [0, 1]$, ist also wohldefiniert und induktiv sieht man mit der Kettenregel, Satz 10.1.5, $\phi \in C^{s+1}([0, 1])$ sowie $\sigma \in \{0, \dots, s+1\}$:

$$\phi^{(\sigma)}(t) = \sum_{j_1, \dots, j_\sigma=1}^n D_{j_\sigma} \dots D_{j_1} f(x^0 + th) h_{j_1} \cdot \dots \cdot h_{j_\sigma} \stackrel{(10.5.1)}{=} d^\sigma f(x^0 + th) h, \quad t \in [0, 1]. \quad (10.5.3)$$

Wir entwickeln ϕ um $t = 0$; Formeln (9.7.3), (9.7.1) und Satz 9.7.4 entnehmen wir dann:

$$\phi(t) = T_{s,0}\phi(t) + R_{s,0}\phi(t) = \sum_{\sigma=0}^s \frac{\phi^{(\sigma)}(0)}{\sigma!} t^\sigma + \frac{\phi^{(s+1)}(\vartheta t)}{(s+1)!} t^{s+1}$$

für $t \in [0, 1]$ und geeignetes $\vartheta = \vartheta(t) \in (0, 1)$. Für $t = 1$ und mit $y := x^0 + \vartheta(1)h \in (x^0, x)$ liefert (10.5.3) also

$$f(x) = \phi(1) = \sum_{\sigma=0}^s \frac{1}{\sigma!} d^\sigma f(x^0)(x - x^0) + \frac{1}{(s+1)!} d^{s+1}(y)(x - x^0),$$

wie behauptet. □

Beispiel 10.5.5. Für $f \in C^2(\Omega)$ erhalten wir speziell: Sind $x^0, x \in \Omega$ mit $[x^0, x] \subset \Omega$ gewählt, so gilt mit $h = x - x^0$ und einem $\vartheta \in (0, 1)$

$$\begin{aligned} f(x) &= d^0 f(x^0)h + d^1 f(x^0)h + \frac{1}{2} d^2 f(x^0 + \vartheta h)h \\ &= f(x^0) + \sum_{j=1}^n f_{x_j}(x^0)h_j + \frac{1}{2} \sum_{i,j=1}^n f_{x_i x_j}(x^0 + \vartheta h)h_i h_j \\ &= f(x^0) + \langle \nabla f(x^0), h \rangle + \frac{1}{2} \langle h, H_f(x^0 + \vartheta h)h \rangle \end{aligned} \quad (10.5.4)$$

mit der (symmetrischen) HESSE-Matrix $H_f = (f_{x_i x_j})_{i,j=1, \dots, n}$.

Folgerung 10.5.6

Sind $s \in \mathbb{N}_0$, $f \in C^s(\Omega)$ und $x^0 \in \Omega$ beliebig gewählt, so gilt

$$f(x) = T_{s,x^0} f(x) + o(\|x - x^0\|^s) \quad \text{für } x \rightarrow x^0.$$

Beweis. Analog zum Beweis von Folgerung 9.7.5 → Ergänzung. □

Kapitel 11.

Anwendungen der mehrdimensionalen Differentialrechnung

In diesem Kapitel geben wir Anwendungen der Ableitungsbegriffe des letzten Kapitels an. In den Abschnitten 11.1 und 11.2 untersuchen wir Extremwertaufgaben. In Abschnitt 11.3 werden wir den Begriff des Diffeomorphismus einführen und erste wichtige Eigenschaften diskutieren. Abschnitt 11.4 ist dann dem zentralen Umkehrsatz gewidmet, in Abschnitt 11.5 wenden wir diesen an, um den Satz über implizite Funktionen zu beweisen. Abschließend betrachten wir in Abschnitt 11.6 Extremwertaufgaben unter Nebenbedingungen, die durch Mannigfaltigkeiten definiert werden.

11.1. Lokale Extrema: Bedingung erster Ordnung

Wir betrachten nun Extremwertaufgaben zur Bestimmung von Minima und Maxima einer Funktion mit $n \in \mathbb{N}$ Veränderlichen. In Definition 7.3.1 haben wir bereits die grundlegenden Begriffe wie (*strikte*) *lokale Minima* und *Maxima* eingeführt. Analog zum Satz von FERMAT, Satz 7.3.3, haben wir

Satz 11.1.1: Notwendige Extremalbedingung 1. Ordnung

Eine Funktion $f \in C^1(\Omega)$ besitze in $x^0 \in \Omega$ ein lokales Extremum. Dann gilt

$$\nabla f(x^0) = 0.$$

Beweis. Sei $r > 0$ mit $B_r(x^0) \subset \Omega$ gewählt. Dann besitzt für jedes $j \in \{1, \dots, n\}$ auch die Funktion $f(x^0 + te_j)$, $t \in (-r, r)$, in $t = 0$ ein lokales Extremum. Der Satz von FERMAT liefert also

$$0 = \frac{d}{dt} f(x^0 + te_j) \Big|_{t=0} \stackrel{(10.1.1)}{=} f_{x_j}(x^0), \quad j = 1, \dots, n,$$

wie behauptet. □

Definition 11.1.2

Ein Punkt $x^0 \in \Omega$ heißt *kritischer* oder *stationärer Punkt* von $f \in C^1(\Omega)$, falls $\nabla f(x^0) = 0$ gilt.

Wir wenden Satz 11.1.1 auf reelle, symmetrische $n \times n$ -Matrizen $A = (a_{jk})_{j,k=1,\dots,n} \in \mathbb{R}^{n \times n}$ an. Ein Vektor $\xi \in \mathbb{R}^n \setminus \{0\}$ heißt bekanntlich *Eigenvektor* von A zum *Eigenwert* $\lambda \in \mathbb{R}$, wenn gilt

$$A\xi = \lambda\xi.$$

Durch Normierung kann man o.B.d.A. $\xi \in S^{n-1}$, d.h. $\|\xi\| = 1$, annehmen.

Satz 11.1.3

Für jede symmetrische Matrix $A \in \mathbb{R}^{n \times n}$ ist

$$\lambda := \sup_{\eta \in \mathbb{R}^n \setminus \{0\}} \frac{\langle \eta, A\eta \rangle}{\|\eta\|^2}$$

ein Eigenwert. Diese Größe wird *RAYLEIGH-Quotient*

Bemerkung 11.1.4. Für den Beweis ist es hilfreich, einer reellen Matrix $A = (a_{ij})_{i,j}$ ihre *Euklidische Länge* oder *Frobeniusnorm* zuzuordnen:

$$\|A\| := \left(\sum_{i,j} a_{ij}^2 \right)^{\frac{1}{2}}.$$

Dann gilt für beliebige $A \in \mathbb{R}^{n \times m}$ und $y \in \mathbb{R}^m$ (\rightarrow Nachrechnen):

$$\|Ay\| \leq \|A\| \|y\|.$$

Beweis von Satz 11.1.3. Wir betrachten die Funktion $f \in C^1(\mathbb{R}^n \setminus \{0\})$ mit

$$f(\eta) := \frac{\langle \eta, A\eta \rangle}{\|\eta\|^2}, \quad \eta \in \mathbb{R}^n \setminus \{0\},$$

und erhalten zunächst mit Bemerkung 11.1.4:

$$\|f(\eta)\| \leq \frac{\|\eta\| \|A\eta\|}{\|\eta\|^2} \leq \|A\|, \quad \eta \in \mathbb{R}^n \setminus \{0\}.$$

Es gilt also $\lambda \leq \|A\| < +\infty$. Wir bemerken weiter, dass f entlang der Strahlen $\{\lambda\eta : \lambda > 0\}$ für jedes $\eta \in \mathbb{R}^n \setminus \{0\}$ konstant ist:

$$f(\lambda\eta) = \frac{\langle \lambda\eta, A(\lambda\eta) \rangle}{\|\lambda\eta\|^2} = \frac{\langle \eta, A\eta \rangle}{\|\eta\|^2} = f(\eta), \quad \lambda > 0, \quad (11.1.1)$$

d.h. f ist *positiv homogen vom Grad 0* auf $\mathbb{R}^n \setminus \{0\}$. Auf der kompakten Menge $S^{n-1} \subset \mathbb{R}^n \setminus \{0\}$ nimmt nun f in einem Punkt $\xi \in S^{n-1}$ ihr Maximum an nach dem Weierstraßschen Hauptsatz, Satz 6.4.4. Sei $\varepsilon > 0$ beliebig. Wegen $\lambda = \sup_{\eta \in \mathbb{R}^n \setminus \{0\}} f(\eta)$ existiert ein $\eta \in \mathbb{R}^n \setminus \{0\}$ mit $f(\eta) \geq \lambda - \varepsilon$. Die Homogenität von f liefert dann

$$\lambda - \varepsilon \leq f(\eta) \stackrel{(11.1.1)}{=} f\left(\frac{\eta}{\|\eta\|}\right) \leq f(\xi) \leq \lambda,$$

also $f(\xi) = \lambda$ für $\varepsilon \rightarrow 0$. Somit ist $\xi \in S^{n-1}$ (globaler) Maximalpunkt von f auf $\mathbb{R}^n \setminus \{0\}$. Satz 11.1.1 liefert daher

$$\nabla f(\xi) = \frac{2}{\|\xi\|^2} (A\xi - f(\xi)\xi) = 0,$$

bzw. $A\xi = f(\xi)\xi = \lambda\xi$, wie behauptet. \square

Durch wiederholte Anwendung dieses Resultats erhalten wir den

Satz 11.1.5

Zu jeder symmetrischen Matrix $A \in \mathbb{R}^{n \times n}$ existieren n Eigenvektoren $\xi_1, \dots, \xi_n \in \mathbb{R}^n$ und n reelle Eigenwerte $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ mit $A\xi_j = \lambda_j\xi_j$ für $j = 1, \dots, n$. Die Eigenvektoren $\{\xi_1, \dots, \xi_n\}$ bilden eine Orthonormalbasis des \mathbb{R}^n , d.h.

$$\langle \xi_j, \xi_k \rangle = \delta_{jk} := \begin{cases} 1, & \text{falls } j = k \\ 0, & \text{falls } j \neq k \end{cases}, \quad j, k = 1, \dots, n.$$

Beweis. Wir schreiben $\lambda_1 := \lambda$ für den in Satz 11.1.3 konstruierten Eigenwert und $\xi_1 \in S^{n-1}$ für einen zugehörigen normierten Eigenvektor. Den zu ξ_1 senkrechten Unterraum $U := \{\eta \in \mathbb{R}^n : \langle \eta, \xi_1 \rangle = 0\}$ können wir durch Wahl einer Orthonormalbasis mit dem \mathbb{R}^{n-1} identifizieren. Setzen wir

$$\lambda_2 := \sup_{\eta \in U \setminus \{0\}} \frac{\langle \eta, A\eta \rangle}{\|\eta\|^2},$$

so gilt offenbar $\lambda_2 \leq \lambda_1$. Mit Satz 11.1.3 erhalten wir wieder einen Eigenvektor $\xi_2 \in S^{n-1} \cap U$ zum Eigenwert λ_2 , für den gilt

$$A\xi_2 = \lambda_2\xi_2, \quad \langle \xi_2, \xi_1 \rangle = 0.$$

Durch Fortsetzung des Verfahrens folgt nach n Schritten die Behauptung. \square

Bemerkung 11.1.6. Offenbar ist λ genau dann Eigenwert von A , wenn das homogene lineare Gleichungssystem

$$(A - \lambda E)\xi = 0$$

mit der Einheitsmatrix \mathbb{E} eine nichttriviale Lösung $\xi \in \mathbb{R}^n \setminus \{0\}$ besitzt, was bekanntlich äquivalent zur Forderung

$$\det(A - \lambda \mathbb{E}) = 0$$

ist. Die linke Seite ist ein Polynom n -ten Grades in λ , das *charakteristische Polynom*.

11.2. Lokale Extrema: Bedingungen zweiter Ordnung

Wir wollen die hinreichende und die notwendige Extremalbedingung zweiter Ordnung aus Satz 7.4.8 bzw. Folgerung 7.4.9 auf Funktionen $f : \Omega \rightarrow \mathbb{R}$ mehrerer Veränderlicher verallgemeinern; die zentrale Rolle kommt dabei der HESSE-Matrix H_f zu. Zur kompakten Formulierung verwenden wir die

Definition 11.2.1

Eine symmetrische Matrix $A \in \mathbb{R}^{n \times n}$ heißt *positiv definit* (i.Z. $A > 0$), wenn gilt

$$\langle \eta, A\eta \rangle > 0 \quad \text{für alle } \eta \in \mathbb{R}^n \setminus \{0\}.$$

A heißt *positiv semidefinit* (i.Z. $A \geq 0$), wenn gilt

$$\langle \eta, A\eta \rangle \geq 0 \quad \text{für alle } \eta \in \mathbb{R}^n.$$

Weiter nennen wir A *negativ definit* bzw. *negativ semidefinit* (i.Z. $A < 0$ bzw. $A \leq 0$), falls $-A > 0$ bzw. $-A \geq 0$ erfüllt ist. Schließlich heißt A *indefinit*, wenn $\langle \eta, A\eta \rangle$ sowohl positive als auch negative Werte annimmt.

Satz 11.2.2: Notwendige Extremalbedingung 2. Ordnung

Eine Funktion $f \in C^2(\Omega)$ nehme in $x^0 \in \Omega \subset \mathbb{R}^n$ ein lokales Minimum (bzw. Maximum) an. Dann gilt für die HESSE-Matrix:

$$H_f(x^0) = (f_{x_j x_k}(x^0))_{j,k=1,\dots,n} \geq 0 \quad (\text{bzw. } H_f(x^0) \leq 0).$$

Beweis. Gemäß Satz 11.1.1 gilt $\nabla f(x^0) = 0$. Ist x^0 lokale Minimalstelle von f , so existiert ein $r > 0$ mit $B_r(x^0) \subset \Omega$ und $f(x) \geq f(x^0)$ für alle $x \in B_r(x^0)$. Die Taylorentwicklung in Beispiel 10.5.5 liefert daher

$$0 \leq f(x) - f(x^0) = \frac{1}{2} \langle h, H_f(x^0 + \vartheta h) h \rangle \quad \text{für alle } h \in \mathbb{R}^n : \|h\| < r$$

mit $x = x^0 + h \in B_r(x^0)$ und einem $\vartheta = \vartheta(h) \in (0, 1)$. Ist nun $\eta \in \mathbb{R}^n$ beliebig, so setzen wir $h := t\eta$ mit hinreichend kleinem $t > 0$ ein und erhalten

$$0 \leq \lim_{t \rightarrow 0^+} \langle \eta, H_f(x^0 + \vartheta t\eta)\eta \rangle = \langle \eta, H_f(x^0)\eta \rangle,$$

also $H_f(x^0) \geq 0$; man beachte, dass H_f nach dem Satz von Schwarz symmetrisch ist.

Entsprechend zeigt man $H_f(x^0) \leq 0$ für eine lokale Maximalstelle x^0 von f . \square

Satz 11.2.3

Es sei $f \in C^2(\Omega)$, $x^0 \in \Omega$ kritischer Punkt und in $B_r(x^0) \subset \Omega$ gelte $H_f \geq 0$ (bzw. $H_f > 0$, $H_f \leq 0$, $H_f < 0$). Dann ist x^0 lokaler Minimierer (bzw. strikter lokaler Minimierer, lokaler Maximierer, strikter lokaler Maximierer) von f .

Beweis. Wie im Beweis von Satz 11.2.2 entnehmen wir Beispiel 10.5.5

$$f(x) - f(x^0) = \frac{1}{2} \langle h, H_f(x^0 + \vartheta h)h \rangle \quad \text{für alle } h \in \mathbb{R}^n : \|h\| < r$$

mit $x = x^0 + h \in B_r(x^0)$ und einem $\vartheta = \vartheta(h) \in (0, 1)$. Aus $H_f \geq 0$ in $B_r(x^0)$ folgt also $f(x) \geq f(x^0)$ für alle $x \in B_r(x^0)$, d.h. x^0 ist lokaler Minimierer von f . Analog ergeben sich die übrigen Aussagen. \square

Wir benötigen noch Aussagen zum Zusammenhang zwischen Eigenwerten und Definitheit einer symmetrischen Matrix, die sich aus Satz 11.1.5 ergeben.

Hilfssatz 11.2.4

Sei $A \in \mathbb{R}^{n \times n}$ eine symmetrische Matrix mit den Eigenwerten $\lambda_1 \geq \dots \geq \lambda_n$ und normierten Eigenvektoren $\xi_1, \dots, \xi_n \in S^{n-1}$ gemäß Satz 11.1.5.

(a) Für $\eta \in \mathbb{R}^n$ setzen wir $c_j := \langle \eta, \xi_j \rangle$, $j = 1, \dots, n$. Dann gilt

$$\|\eta\|^2 = \sum_{j=1}^n c_j^2, \quad \langle \eta, A\eta \rangle = \sum_{j=1}^n \lambda_j c_j^2. \quad (11.2.1)$$

(b) Es gelten die Äquivalenzen

$$A > 0 \ (\geq 0) \Leftrightarrow \lambda_n > 0 \ (\geq 0), \quad (11.2.2)$$

$$A < 0 \ (\leq 0) \Leftrightarrow \lambda_1 < 0 \ (\leq 0), \quad (11.2.3)$$

$$A \text{ indefinit} \Leftrightarrow \lambda_n < 0 \text{ und } \lambda_1 > 0. \quad (11.2.4)$$

Beweis. (a) \rightarrow Ergänzungen.

(b) Am Beispiel von (11.2.2): Ist $A > 0$, so folgt $0 < \langle \xi_n, A\xi_n \rangle = \lambda_n \|\xi_n\|^2 = \lambda_n$. Gilt umgekehrt $\lambda_n > 0$, so entnehmen wir (11.2.1) für beliebiges $\eta \in \mathbb{R}^n \setminus \{0\}$:

$$\langle \eta, A\eta \rangle \geq \lambda_n \sum_{j=1}^n c_j^2 = \lambda_n \|\eta\|^2 > 0,$$

also $A > 0$. Entsprechend sieht man $A \geq 0 \Leftrightarrow \lambda_n \geq 0$.

(11.2.3), (11.2.4): \rightarrow Ergänzungen. □

Beispiel 11.2.5. $\lambda \in \mathbb{R}$ ist genau dann Eigenwert von $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix} \in \mathbb{R}^{2 \times 2}$, wenn gilt

$$0 = \det(A - \lambda \mathbb{E}) = \lambda^2 - (a + c)\lambda + ac - b^2.$$

Für die Eigenwerte $\lambda_1 \geq \lambda_2$ von A folgt somit

$$\lambda_1 + \lambda_2 = a + c, \quad \lambda_1 \lambda_2 = ac - b^2 = \det A,$$

z.B. aus der „ p - q -Formel“. Wir entnehmen also Hilfssatz 11.2.4 die

Folgerung 11.2.6

Für $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix} \in \mathbb{R}^{2 \times 2}$ gilt:

$$A > 0 \Leftrightarrow \det A > 0 \text{ und } a > 0,$$

$$A < 0 \Leftrightarrow \det A > 0 \text{ und } a < 0,$$

$$A \geq 0 \text{ oder } A \leq 0 \Leftrightarrow \det A \geq 0,$$

$$A \text{ indefinit} \Leftrightarrow \det A < 0.$$

Wir kommen nun zur angekündigten Verallgemeinerung von Satz 7.4.8:

Satz 11.2.7: Hinreichende Extremalbedingung

Sei $x^0 \in \Omega$ kritischer Punkt der Funktion $f \in C^2(\Omega)$ und es gelte $H_f(x^0) > 0$ (bzw. $H_f(x^0) < 0$). Dann besitzt f in x^0 ein striktes lokales Minimum (bzw. Maximum).

Beweis. Wir zeigen, dass ein $r > 0$ existiert mit $B_r(x^0) \subset \Omega$ und $H_f > 0$ (bzw. $H_f < 0$) in $B_r(x^0)$; die Behauptung folgt dann aus Satz 11.2.3.

Sei also $H_f(x^0) > 0$. Nach Hilfssatz 11.2.5 gilt dann $\lambda_n > 0$ für den kleinsten Eigenwert von $H_f(x^0)$. Und aus Formel (11.2.1) erhalten wir mit den dortigen Bezeichnungen

$$\langle \eta, H_f(x^0)\eta \rangle \geq \lambda_n \sum_{j=1}^n c_j^2 = \lambda_n \|\eta\|^2 \quad \text{für alle } \eta \in \mathbb{R}^n$$

und folglich

$$\begin{aligned} \langle \eta, H_f(x)\eta \rangle &= \langle \eta, H_f(x^0)\eta \rangle + \langle \eta, (H_f(x) - H_f(x^0))\eta \rangle \\ &\geq \lambda_n \|\eta\|^2 - \|\eta\| \|(H_f(x) - H_f(x^0))\eta\| \\ &\stackrel{\text{Bem. 11.1.4}}{\geq} (\lambda_n - \|H_f(x) - H_f(x^0)\|) \|\eta\|^2 \quad \text{für alle } \eta \in \mathbb{R}^n, x \in \Omega. \end{aligned} \quad (11.2.5)$$

Wegen $f \in C^2(\Omega)$ existiert ein $r > 0$ mit $B_r(x^0) \subset \Omega$ und

$$\|H_f(x) - H_f(x^0)\| \leq \frac{\lambda_n}{2} \quad \text{für alle } x \in B_r(x^0).$$

Einsetzen in (11.2.5) ergibt also

$$\langle \eta, H_f(x)\eta \rangle \geq \frac{\lambda_n}{2} \|\eta\|^2 > 0 \quad \text{für alle } \eta \in \mathbb{R}^n \setminus \{0\}, x \in B_r(x^0),$$

d.h. $H_f > 0$ in $B_r(x^0)$. Der Fall $H_f(x^0) < 0$ wird entsprechend oder durch Übergang zu $-f$ behandelt. \square

11.3. Diffeomorphismen

Eine 1-1-Zuordnung der Elemente zweier Mengen gelingt mittels Bijektion. In der Differentialrechnung benötigt man zusätzlich Differenzierbarkeitseigenschaften, das führt auf

Definition 11.3.1

Seien $\Omega, \Theta \subset \mathbb{R}^n$ offene Mengen und $s \in \mathbb{N} \cup \{\infty\}$. Eine Abbildung $f : \Omega \rightarrow \mathbb{R}^n$ heißt *Diffeomorphismus (der Klasse C^s)* von Ω auf $\Theta = f(\Omega)$ oder *C^s -Diffeomorphismus*, wenn:

- ▶ $f : \Omega \rightarrow \Theta$ ist bijektiv,
- ▶ $f \in C^s(\Omega, \mathbb{R}^n)$ und $f^{-1} \in C^s(\Theta, \mathbb{R}^n)$.

Bemerkung 11.3.2. Man kann die Definition allgemeiner für Abbildungen $f : \Omega \rightarrow \Theta$ mit $\Omega \subset \mathbb{R}^n$ und $\Theta \subset \mathbb{R}^m$ für beliebige Dimensionen $n, m \in \mathbb{N}$ formulieren. Allerdings folgt aus der Existenz eines solchen Diffeomorphismus dann sofort $m = n$ (\rightarrow Ergänzungen).

Hilfssatz 11.3.3

Sei $f : \Omega \rightarrow \Theta$ ein Diffeomorphismus, $\Omega, \Theta \subset \mathbb{R}^n$. Dann ist $Df : \Omega \rightarrow \mathbb{R}^{n \times n}$ in jedem Punkt aus Ω invertierbar, und es gelten

$$(Df(x))^{-1} = Df^{-1}(f(x)), \quad J_f(x) = \frac{1}{J_{f^{-1}}(f(x))}, \quad x \in \Omega.$$

Beweis. Die Kettenregel (10.1.2) und die Relation $f^{-1} \circ f = \text{Id}$ auf Ω liefern

$$\mathbb{E} = Df^{-1}(f(x)) \circ Df(x), \quad x \in \Omega,$$

mit der Einheitsmatrix $\mathbb{E} \in \mathbb{R}^{n \times n}$. Folglich ist Df in Ω invertierbar mit der Inversen $(Df^{-1}) \circ f$. Die zweite Relation ergibt sich sofort aus (10.1.4) und $\det \mathbb{E} = 1$. \square

Bemerkung 11.3.4. Die Umkehrung von Hilfssatz 11.3.3 gilt nicht: Die Polarkoordinatenabbildung $f(r, \varphi) = (r \cos \varphi, r \sin \varphi)$ erfüllt $J_f(r, \varphi) = r > 0$ für $(r, \varphi) \in (0, +\infty) \times \mathbb{R}$, siehe Beispiel 10.1.9 (b). Wegen der 2π -Periodizität von \cos und \sin ist f aber nicht einmal bijektiv. Jedoch gilt der

Satz 11.3.5: Satz von der Diffeomorphie

Die Funktion $f \in C^1(\Omega, \mathbb{R}^n)$ bilde $\Omega \subset \mathbb{R}^n$ bijektiv auf die offene Menge $\Theta = f(\Omega)$ ab. Weiter seien $J_f \neq 0$ in Ω und $f^{-1} : \Theta \rightarrow \mathbb{R}^n \in C^0(\Theta, \mathbb{R}^n)$. Dann ist f ein C^1 -Diffeomorphismus.

Beweis. 1. Wir haben $f^{-1} \in C^1(\Theta, \mathbb{R}^n)$ zu zeigen. Dazu beweisen wir, dass f^{-1} in Θ total differenzierbar ist; genauer behaupten wir: Zu jedem $y \in \Theta$ gibt es ein $\varepsilon = \varepsilon(y) > 0$ mit $B_\varepsilon(y) \subset \Theta$, so dass für alle $k \in \mathbb{R}^n$ mit $\|k\| < \varepsilon$ gilt

$$f^{-1}(y + k) = f^{-1}(y) + A \cdot k + R^*(k) \quad \text{mit } R^*(k) = o(\|k\|) \text{ für } k \rightarrow 0, \quad (11.3.1)$$

wobei $A := Df(f^{-1}(y))^{-1}$ gesetzt wurde. Gemäß Satz 10.3.4 ist dann f^{-1} partiell differenzierbar auf Θ mit Funktionalmatrix

$$Df^{-1} = ((Df) \circ f^{-1})^{-1} \quad \text{in } \Theta.$$

Da die rechte Seite nach Voraussetzung stetig ist, folgt $f^{-1} \in C^1(\Theta, \mathbb{R}^n)$.

2. Zu zeigen bleibt (11.3.1): Wir setzen $x := f^{-1}(y) \in \Omega$. Gemäß Satz 10.3.7 ist $f \in C^1(\Omega, \mathbb{R}^n)$ total differenzierbar in x , d.h. es gibt ein $r > 0$ mit $B_r(x) \subset \Omega$, so dass für alle $h \in B_r(0)$ gilt:

$$f(x+h) = f(x) + Df(x)h + R(h) \quad \text{mit } R(h) = o(\|h\|) \text{ für } h \rightarrow 0. \quad (11.3.2)$$

Wir setzen ferner $m := \|A\| > 0$ und wählen $r > 0$ zusätzlich so klein, dass

$$\|R(h)\| \leq \frac{1}{2m} \|h\| \quad \text{für alle } h \in B_r(0). \quad (11.3.3)$$

Da f^{-1} stetig ist, existiert $\varepsilon > 0$ so, dass für alle $k \in B_\varepsilon(0)$ gilt $\|f^{-1}(y+k) - f^{-1}(y)\| < r$. Wir setzen dann

$$h := f^{-1}(y+k) - f^{-1}(y) \in B_r(0) \quad \text{bzw.} \quad k = f(x+h) - f(x), \quad k \in B_\varepsilon(0).$$

Da $Df(x)$ invertierbar ist, können wir (11.3.2) für beliebiges $k \in B_\varepsilon(0)$ und dieses h umschreiben zu

$$f^{-1}(y+k) - f^{-1}(y) = h = Df(x)^{-1} \circ (f(x+h) - f(x) - R(h)) = Ak + R^*(k), \quad (11.3.4)$$

wobei wir $R^*(k) := -AR(f^{-1}(y+k) - f^{-1}(y)) = -AR(h)$ gesetzt haben.

3. Für den Beweis von (11.3.1) ist wegen (11.3.4) nur noch $R^*(k) = o(\|k\|)$ für $k \rightarrow 0$ zu zeigen. Für beliebiges $k \in B_\varepsilon(0)$ und $h = f^{-1}(y+k) - f^{-1}(y)$ beachten wir dazu

$$\|h\| \stackrel{(11.3.4)}{\leq} \|A\| \|k\| + \|R^*(k)\| \leq m \|k\| + m \|R(h)\| \stackrel{(11.3.3)}{\leq} m \|k\| + \frac{1}{2} \|h\|$$

bzw. $\|h\| \leq 2m \|k\|$. Folglich gilt $h \rightarrow 0$ ($k \rightarrow 0$) und schließlich

$$\frac{\|R^*(k)\|}{\|k\|} \leq m \frac{\|R(h)\|}{\|k\|} \leq 2m^2 \frac{\|R(h)\|}{\|h\|} \stackrel{(11.3.2)}{\rightarrow} 0 \quad (k \rightarrow 0)$$

bzw. $R^*(k) = o(\|h\|)$ für $k \rightarrow 0$. Damit ist alles gezeigt. \square

Bemerkung 11.3.6. Wir werden im nächsten Abschnitt sehen, dass die Aussage von Satz 11.3.5 richtig bleibt, wenn wir die Voraussetzungen der Offenheit von Θ und Stetigkeit von f^{-1} fallenlassen (Folgerung 11.4.8).

Folgerung 11.3.7

Erfüllt $f : \Omega \rightarrow \mathbb{R}^n$ die Voraussetzungen von Satz 11.3.5 und gilt zusätzlich $f \in C^s(\Omega, \mathbb{R}^n)$ für ein $s \in \mathbb{N} \cup \{\infty\}$, so ist f ein C^s -Diffeomorphismus.

Beweis. Nach Satz 11.3.5 ist f zunächst ein C^1 -Diffeomorphismus. Hilfssatz 11.3.3 entnehmen wir daher

$$Df^{-1}(y) = Df(f^{-1}(y))^{-1} \quad \text{für alle } y \in \Theta.$$

Mit vollständiger Induktion folgt $f^{-1} \in C^s(\Theta, \mathbb{R}^n)$; dabei verwendet man noch, dass die Komposition zweier C^s -Funktionen wieder eine C^s -Funktion ist (\rightarrow Übungsaufgabe). \square

11.4. Der Umkehrsatz

In diesem Abschnitt wollen wir untersuchen, unter welchen Voraussetzungen eine Abbildung $f \in C^1(\Omega, \mathbb{R}^n)$ eine Umkehrfunktion besitzt. Dazu zunächst ein

Beispiel 11.4.1. Zu $A \in \mathbb{R}^{n \times n}$, $b \in \mathbb{R}^n$ betrachten wir die affin-lineare Abbildung $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$, $x \mapsto Ax + b$. Die lineare Algebra sagt uns, dass f genau dann injektiv ist mit $f(\mathbb{R}^n) = \mathbb{R}^n$, wenn A invertierbar ist, d.h. wenn $\det A \neq 0$ gilt. Wegen $Df = A$ auf \mathbb{R}^n heißt das: $f^{-1} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ existiert genau dann, wenn $J_f \neq 0$ auf \mathbb{R}^n gilt. Ferner ist $f^{-1}(y) = A^{-1}(y - b)$, $y \in \mathbb{R}^n$, richtig, d.h. f ist ein Diffeomorphismus.

In Bemerkung 11.3.4 haben wir gesehen, dass wir die Aussage von Beispiel 11.4.1 nicht auf beliebige C^1 -Funktionen übertragen können. Im „Kleinen“ geht das jedoch, wir erhalten den zentralen

Satz 11.4.2: Umkehrsatz

Sei $\Omega \subset \mathbb{R}^n$ offen, $f : \Omega \rightarrow \mathbb{R}^n \in C^s(\Omega, \mathbb{R}^n)$ mit $s \in \mathbb{N} \cup \{\infty\}$ sei gegeben, und in einem Punkt $x^0 \in \Omega$ gelte

$$J_f(x^0) = \det Df(x^0) \neq 0.$$

Dann gibt es eine offene Umgebung $U = U(x^0) \subset \Omega$ von x^0 , so dass $V := f(U)$ offen ist und $f|_U$ einen C^s -Diffeomorphismus von U auf V darstellt.

Bemerkung 11.4.3. Eine Menge $U \subset \mathbb{R}^n$ heißt *Umgebung* von $x^0 \in \mathbb{R}^n$, wenn es eine r -Umgebung $B_r(x^0)$ von x^0 gibt mit $B_r(x^0) \subset U$; insbesondere ist jede r -Umgebung also selbst auch Umgebung.

Zum Beweis des Umkehrsatzes zeigen wir in zwei Hilfssätzen, dass wir U so wählen können, dass die Voraussetzungen von Folgerung 11.3.7 erfüllt sind. Für deren elegante Formulierung geben wir noch die

Definition 11.4.4

Seien $\Omega \subset \mathbb{R}^n$ offen und $f : \Omega \rightarrow \mathbb{R}^d$ gegeben.

- (a) f heißt *offen*, wenn das Bild $f(\Omega')$ jeder offenen Teilmenge Ω' von Ω wieder offen ist.
- (b) Ist $f \in C^1(\Omega, \mathbb{R}^d)$, so nennen wir einen Punkt $x^0 \in \Omega$ *regulär*, wenn $df(x^0)$ surjektiv ist; anderenfalls heißt x^0 *singulär*. Sind alle Punkte von Ω regulär, so heißt f *regulär (in Ω)*.

Bemerkung 11.4.5. (a) Für beliebiges $f : \Omega \rightarrow \mathbb{R}^d$ sind äquivalent (\rightarrow Übungsaufgabe):

- (i) f ist offen.
- (ii) Sind $x^0 \in \Omega$, $r > 0$ mit $B_r(x^0) \subset \Omega$ beliebig, so gibt es zu jedem $\varepsilon \in (0, r)$ ein $\delta > 0$ mit $B_\delta(f(x^0)) \subset f(B_\varepsilon(x^0))$.
- (b) Wegen Satz 10.3.4 ist $x^0 \in \Omega$ genau dann regulärer Punkt von $f \in C^1(\Omega, \mathbb{R}^d)$, wenn $n \geq d$ und $\text{rang } Df(x^0) = d$ gilt. Speziell folgt:
 - ▶ $d = n$: x^0 ist regulär $\Leftrightarrow J_f(x^0) \neq 0$.
 - ▶ $d = 1$: x^0 ist regulär $\Leftrightarrow \nabla f(x^0) \neq 0$.

Hilfssatz 11.4.6

Sei $x^0 \in \Omega$ regulärer Punkt von $f \in C^1(\Omega, \mathbb{R}^n)$. Dann gibt es eine Umgebung $U = U(x^0) \subset \Omega$ von x^0 so, dass $f|_U$ injektiv ist und die Umkehrabbildung $f^{-1} : V \rightarrow \mathbb{R}^n$ stetig ist auf $V := f(U)$.

Beweis. Wir betrachten die Funktion

$$\psi(x) := f(x) - f(x^0) - Df(x^0) \cdot (x - x^0), \quad x \in \Omega.$$

Dann gilt $\psi \in C^1(\Omega, \mathbb{R}^n)$ und $D\psi(x^0) = 0$. In einer Kugel $B_R(x^0) \subset \Omega$ liefert Hadamards Lemma

$$\psi(x) - \psi(x') = A \cdot (x - x'), \quad x, x' \in B_R(x^0), \quad (11.4.1)$$

mit

$$A := \int_0^1 D\psi(x' + t(x - x')) dt.$$

Da x^0 regulär ist, ist $Df(x^0) \in \mathbb{R}^{n \times n}$ invertierbar. Wir setzen $\mu := \frac{1}{2\|Df(x^0)^{-1}\|} > 0$. Wegen $D\psi(x^0) = 0$ und der Stetigkeit von $D\psi$ existiert ein $r \in (0, R)$ mit der Eigenschaft $\|D\psi\| \leq \mu$ in $B_r(x^0)$. Wir erhalten

$$\|A\| \leq \int_0^1 \|D\psi(x' + t(x - x'))\| dt \leq \mu \quad \text{für alle } x, x' \in B_r(x^0). \quad (11.4.2)$$

Insgesamt finden wir für beliebige $x, x' \in B_r(x^0)$:

$$\begin{aligned} \|f(x) - f(x')\| &= \|\psi(x) - \psi(x') + Df(x^0)(x - x')\| \\ &\geq \|Df(x^0)(x - x')\| - \|\psi(x) - \psi(x')\| \\ &\stackrel{(11.4.1)}{\geq} \|Df(x^0)(x - x')\| - \|A\|\|x - x'\| \\ &\stackrel{(11.4.2)}{\geq} \|Df(x^0)(x - x')\| - \mu\|x - x'\| \\ &\geq 2\mu\|Df(x^0)^{-1}Df(x^0)(x - x')\| - \mu\|x - x'\| \\ &= \mu\|x - x'\|. \end{aligned} \quad (11.4.3)$$

Folglich ist $f|_U$ injektiv auf $U = B_r(x^0)$. Für die Umkehrabbildung liefert (11.4.3) noch die behauptete Stetigkeit auf $V = f(U)$. \square

Hilfssatz 11.4.7

Jede reguläre Abbildung $f \in C^1(\Omega, \mathbb{R}^n)$ ist offen.

Beweis. Man verwende die äquivalente Eigenschaft aus Bemerkung 11.4.5 (a) und Satz 11.1.1 \rightarrow Ergänzungen. \square

Wir kommen nun zum

Beweis von Satz 11.4.2. Wegen $J_f(x^0) \neq 0$ und der Stetigkeit von J_f existiert eine Umgebung $U = B_r(x^0) \subset \Omega$, so dass $f|_U$ regulär ist. Nach den Hilfssätzen 11.4.6 und 11.4.7 können wir U noch so wählen, dass $V = f(U)$ offen und $f : U \rightarrow V$ bijektiv mit Umkehrfunktion $f^{-1} : V \rightarrow U \in C^0(V, \mathbb{R}^n)$ ist. Folgerung 11.3.7 liefert somit die Behauptung. \square

Wir halten noch eine wichtige Beobachtung fest, die eine abgeschwächte, globale Version von Satz 11.4.2 und gleichzeitig eine Verschärfung von Folgerung 11.3.7 darstellt:

Folgerung 11.4.8

Jede reguläre, bijektive Abbildung $f : \Omega \rightarrow \Theta \in C^s(\Omega, \mathbb{R}^n)$, $s \in \mathbb{N} \cup \{\infty\}$, ist offen und ein C^s -Diffeomorphismus.

Beweis. Nach Hilfssatz 11.4.7 ist f offen. Nach Hilfssatz 11.4.6 ist $f^{-1} : \Theta \rightarrow \Omega$ in einer Umgebung $V = V(y^0) \subset \Theta$ jedes Punktes $y^0 \in \Theta$ stetig, es folgt also $f^{-1} \in C^0(\Theta, \mathbb{R}^n)$. Folgerung 11.3.7 liefert die Behauptung. \square

11.5. Implizite Funktionen

Sei wie immer $\Omega \subset \mathbb{R}^n$ offen. Den Umkehrsatz, Satz 11.4.2, können wir dahingehend interpretieren, dass das Gleichungssystem $f(x) = y$ für $f : \Omega \rightarrow \mathbb{R}^n$, $y \in \mathbb{R}^n$ unter den angegebenen Voraussetzungen lokal eindeutig gelöst werden kann: Die n Unbekannten (x_1, \dots, x_n) werden aus den n Gleichungen $f_j(x_1, \dots, x_n) = y_j$, $j = 1, \dots, n$, eindeutig bestimmt.

Analog betrachten wir nun dieses System $f(x) = y$ für $f : \Omega \rightarrow \mathbb{R}^d$, also d Gleichungen zur Bestimmung von n Unbekannten, wobei $n \neq d$ gelte. Auch hierzu wieder das Beispiel affin-linearer Abbildungen:

Beispiel 11.5.1. Sei $f(x) := Ax + b$, $x \in \mathbb{R}^n$, mit $A \in \mathbb{R}^{d \times n}$, $b \in \mathbb{R}^d$.

- Für $d > n$ haben wir mehr Gleichungen als Unbekannte, das System ist *überbestimmt* und bekanntlich nicht für alle Werte y lösbar. Diesen Fall wollen wir daher auch i.F. nicht betrachten.
- Für $d < n$, also weniger Gleichungen als Unbekannte, spricht man vom *unterbestimmten* System. Falls $\text{rg } A = d$ ist, gibt es eine Spaltenumordnung - wieder mit A bezeichnet -, so dass wir $A = (B, C)$ mit Matrizen $B \in \mathbb{R}^{d \times (n-d)}$, $C \in \mathbb{R}^{d \times d}$ und $\det C \neq 0$ schreiben können. Ordnen wir die Koordinaten von x entsprechend um und schreiben $x = (u, v)$ mit $u \in \mathbb{R}^{n-d}$, $v \in \mathbb{R}^d$, so liest sich $y = Ax + b$ als

$$y = (B, C) \begin{pmatrix} u \\ v \end{pmatrix} + b = Bu + Cv + b$$

und nach Multiplikation mit C^{-1} von links folgt

$$v = -C^{-1}Bu + C^{-1}(y - b) =: \varphi(u, y).$$

Wir haben also die Gleichung $y = f(u, v)$ für beliebiges $y \in \mathbb{R}^d$ nach v aufgelöst:

$$\{(u, v) \in \mathbb{R}^{n-d} \times \mathbb{R}^d : y = f(u, v)\} = \{(u, \varphi(u, y)) : u \in \mathbb{R}^{n-d}\}.$$

Der Lösungsraum der Gleichung ist also (für festes y) $n - d$ -dimensional.

Wie beim Umkehrsatz lässt sich dieses Beispiel wieder lokal auf nichtlineare Systeme übertragen. Wir setzen zu $n, d \in \mathbb{N}$ mit $n > d$ zur besseren Übersichtlichkeit $m := n - d \in \mathbb{N}$ bzw. $n = m + d$. Im Folgenden ist also $\Omega \subset \mathbb{R}^{m+d}$ offen; für ein $f = f(u, v) \in C^1(\Omega, \mathbb{R}^d)$ schreiben wir

$$Df = (D_u f, D_v f) \quad \text{mit} \quad D_u f := \begin{pmatrix} f_{1u_1} & \cdots & f_{1u_m} \\ \vdots & \ddots & \vdots \\ f_{du_1} & \cdots & f_{du_m} \end{pmatrix}, \quad D_v f := \begin{pmatrix} f_{1v_1} & \cdots & f_{1v_d} \\ \vdots & \ddots & \vdots \\ f_{dv_1} & \cdots & f_{dv_d} \end{pmatrix} \quad \text{in } \Omega.$$

Wir erhalten den wichtigen

Satz 11.5.2: Satz über implizite Funktionen

Seien $s \in \mathbb{N} \cup \{\infty\}$, $\Omega \subset \mathbb{R}^{m+d}$ offen und $f = f(u, v) : \Omega \rightarrow \mathbb{R}^d \in C^s(\Omega, \mathbb{R}^d)$. In einem Punkt $(u^0, v^0) \in \Omega$ sei $f(u^0, v^0) = y^0$, und es gelte

$$\det D_v f(u^0, v^0) \neq 0. \quad (11.5.1)$$

Dann gibt es offene Umgebungen $U = U(u^0), V = V(v^0)$ mit $U \times V \subset \Omega$ und $Y = Y(y^0) \subset \mathbb{R}^d$ sowie eine Funktion $\varphi = \varphi(u, y) \in C^s(U \times Y, \mathbb{R}^d)$, so dass

$$\{(u, v) \in U \times V : f(u, v) = y\} = \{(u, \varphi(u, y)) : u \in U\} \quad \text{für alle } y \in Y.$$

Beweis. Wir betrachten die Funktion

$$F(u, v) := (u, f(u, v)) \in C^s(\Omega, \mathbb{R}^{m+d}).$$

Für die Funktionalmatrix haben wir

$$DF(u, v) = \begin{pmatrix} \mathbb{E} & 0 \\ D_u f(u, v) & D_v f(u, v) \end{pmatrix}, \quad (u, v) \in \Omega,$$

die Regularitätsbedingung (11.5.1) liefert also nach dem Determinantenentwicklungssatz

$$J_F(u^0, v^0) = \det D_v f(u^0, v^0) \neq 0.$$

Nach dem Umkehrsatz, Satz 11.4.2, existiert somit eine offene Umgebung $W = W(u^0, v^0) \subset \Omega$, so dass $F|_W$ ein C^s -Diffeomorphismus von W auf die offene Umgebung $Z = Z(u^0, y^0) := F(W)$ ist.

Wir wählen zunächst $r > 0$ so klein, dass $B_r(u^0) \times B_r(v^0) \subset B_{2r}(u^0, v^0) \subset W$ richtig ist. Dann ist $F(B_r(u^0) \times B_r(v^0)) \subset Z$ offene Umgebung von (u^0, y^0) . Nun wählen wir $\varepsilon \in (0, r]$ so klein, dass $B_\varepsilon(u^0) \times B_\varepsilon(y^0) \subset B_{2\varepsilon}(u^0, y^0) \subset F(B_r(u^0) \times B_r(v^0))$ gilt. Schließlich setzen wir

$$U := B_\varepsilon(u^0), \quad V := B_r(v^0), \quad Y := B_\varepsilon(y^0).$$

Dann ist $U \times V \subset W$ und $U \times Y \subset Z' := F(U \times V)$. Wir bezeichnen die Einschränkung $F|_{U \times V}$ wieder mit F und schreiben $\Phi := F^{-1} : Z' \rightarrow U \times V$ für die Umkehrabbildung. Wegen $F = (\text{Id}, f)$ ist auch $\Phi = (\text{Id}, \varphi)$ mit einer Funktion $\varphi \in C^s(Z', \mathbb{R}^d)$. Für beliebiges $y \in Y$ setzen wir $M := \{(u, v) \in U \times V : f(u, v) = y\}$ und erhalten schließlich

$$\begin{aligned} M &= \Phi(F(M)) = \Phi(\{(u, \eta) \in Z' : \eta = y\}) \\ &= \Phi(U \times \{y\}) = \{(u, \varphi(u, y)) : u \in U\}, \end{aligned}$$

wie behauptet. \square

Bemerkung 11.5.3. Aus Satz 11.5.2 lesen wir für jedes $y \in Y$ ab:

$$y = f(u, \varphi(u, y)), \quad u \in U.$$

Leiten wir nach u ab, so folgt also

$$0 = D_u f(u, \varphi(u, y)) + D_v f(u, \varphi(u, y)) D_u \varphi(u, y)$$

bzw.

$$D_u \varphi(u, y) = -D_v f(u, \varphi(u, y))^{-1} \cdot D_u f(u, \varphi(u, y)), \quad y \in U, \quad (11.5.2)$$

da nach Konstruktion im Beweis von Satz 11.5.2 und wegen Hilfssatz 11.3.3 $\det D_v f \neq 0$ in $U \times V$ gilt.

Folgerung 11.5.4

Ist $f \in C^s(\Omega, \mathbb{R}^d)$ gegeben, $s \in \mathbb{N} \cup \{\infty\}$, und ist $x^0 \in \Omega \subset \mathbb{R}^{m+d}$ regulärer Punkt von f mit $f(x^0) = y^0$, so gibt es Umgebungen $X = X(x^0) \subset \Omega$ und $Y = Y(y^0) \subset \mathbb{R}^d$, so dass sich die (lokalen) *Niveaumengen*

$$(f|_X)^{-1}(y) := \{x \in X : f(x) = y\}, \quad y \in Y,$$

als Graph von C^s -Funktionen über (ein und derselben) m -dimensionalen Koordinatenebene schreiben lassen.

Beweis. Da x^0 regulär ist, gilt $\text{rang } Df(x^0) = d$. Nach Umsortierung der Koordinaten können wir also mit $x = (u, v) \in \mathbb{R}^m \times \mathbb{R}^d$ annehmen:

$$\det D_v f(u^0, v^0) \neq 0.$$

Wir können somit Satz 11.5.2 auf f anwenden: es existieren offene Umgebungen $U = U(u^0)$, $V = V(v^0)$, $Y = Y(y^0)$ und eine Funktion $\varphi \in C^s(U \times Y, \mathbb{R}^d)$, so dass

$$(f|_{U \times V})^{-1}(y) = \{(u, v) \in U \times V : f(u, v) = y\} = \{(u, \varphi(u, y)) : u \in U\}, \quad y \in Y.$$

Mit $X := U \times V$ ist das die Behauptung. \square

Beispiel 11.5.5. Wir betrachten $f(x) = f(u, v) := u^2 - v^2$, $(u, v) \in \mathbb{R}^2$. Dann gilt $\nabla f(u, v) = 2(u, -v)$; alle Punkte $x^0 = (u^0, v^0) \in \mathbb{R}^2 \setminus \{(0, 0)\}$ sind also regulär und die Niveaulächen von f lassen sich lokal um diese Punkte als Graphen von C^∞ -Funktionen darstellen.

Wegen $f(0, 0) = 0$ und $f^{-1}(0) = \{(u, v) \in \mathbb{R}^2 : u = \pm v\}$ sehen wir, dass Folgerung 11.5.4 ohne Voraussetzung der Regularität von x^0 falsch wird: Für keine Umgebung $X = X(0, 0)$ lässt sich $(f|_X)^{-1}(0)$ als Graph über einer der Koordinatenachsen schreiben.

11.6. Mannigfaltigkeiten und lokale Extrema unter Nebenbedingungen

In Folgerung 11.5.4 haben wir gesehen, dass sich die Niveaumengen von C^s -Funktionen lokal um reguläre Punkte als Graphen schreiben lassen. Geometrisch sind solche Funktionsgraphen *Flächen* (bzw. im Fall $m = 1$ *Kurven*, die wir hier subsummieren wollen). Objekte, die sich aus Flächen zusammensetzen, werden *Mannigfaltigkeiten* genannt. Das erklärt unsere folgende Begriffsbildung:

Definition 11.6.1: Gleichungsdefinierte Mannigfaltigkeit

Seien immer $m, d \in \mathbb{N}$ und $s \in \mathbb{N} \cup \{\infty\}$. Eine nichtleere Menge $M \subset \mathbb{R}^{m+d}$ heißt *m-dimensionale gleichungsdefinierte Mannigfaltigkeit* der Klasse C^s oder *C^s -Mannigfaltigkeit*, wenn eine offene Menge $\Omega \subset \mathbb{R}^{m+d}$ und eine Funktion $f \in C^s(\Omega, \mathbb{R}^d)$ so existieren, dass

$$M = f^{-1}(0) = \{x \in \Omega : f(x) = 0\} \subset \Omega$$

gilt und alle $x^0 \in M$ reguläre Punkte von f sind. m ist dann die *Dimension*, d die *Kodimension* von M .

Bemerkung 11.6.2. (a) Aus Stetigkeitsgründen können wir (bei Bedarf) annehmen, dass f regulär (in Ω) ist, indem wir ggf. Ω durch eine offene Menge Ω' mit $M \subset \Omega' \subset \Omega$ und $\text{rang } Df = d$ auf Ω' ersetzen; siehe Definition 11.4.4 und Bemerkung 11.4.5 (b).

(b) Der in der Geometrie gebräuchliche Begriff der Mannigfaltigkeit ist deutlich allgemeiner.

Beispiel 11.6.3. $\Theta \subset \mathbb{R}^m$ sei offen und $\varphi : \Theta \rightarrow \mathbb{R}^d \in C^s(\Theta, \mathbb{R}^d)$ sei gegeben. Wir erklären $\Omega := \Theta \times \mathbb{R}^d$ und $f \in C^s(\Omega, \mathbb{R}^d)$ mittels

$$f(x) := v - \varphi(u), \quad x := (u, v) \in \Omega.$$

Wegen $Df = (-D_u \varphi, \mathbb{E})$ in Ω ist f regulär, $M = f^{-1}(0) = \text{graph } \varphi$ ist also eine m -dimensionale Mannigfaltigkeit der Klasse C^s . Die Umkehrung gilt immerhin lokal:

Folgerung 11.6.4

Jede m -dimensionale Mannigfaltigkeit $M \subset \mathbb{R}^{m+d}$ der Klasse C^s lässt sich lokal als Graph einer Funktion $\varphi : U \rightarrow \mathbb{R}^d \in C^s(U, \mathbb{R}^d)$ mit offener Menge $U \subset \mathbb{R}^m$ schreiben.

Beweis. Da jedes $x^0 \in M \subset \Omega$ regulärer Punkt von $f \in C^s(\Omega, \mathbb{R}^d)$ mit $f(x^0) = 0$ ist, existiert gemäß Folgerung 11.5.4 eine offene Umgebung $X = X(x^0) \subset \Omega$, so dass

$$(f|_X)^{-1}(0) = \{x \in X : f(x) = 0\} = M \cap X$$

als Graph einer C^s -Funktion über einer m -dimensionalen Koordinatenebene geschrieben werden kann. \square

Bemerkung 11.6.5. Die Regularitätsforderung ist hier wesentlich, wir erinnern an Beispiel 11.5.5.

Wir wollen Mannigfaltigkeiten noch ein bisschen besser verstehen, wie bei Funktionen dient dazu zunächst die erste Näherung:

Definition 11.6.6: Tangential- und Normalraum

Es sei $M \subset \mathbb{R}^{m+d}$ eine m -dimensionale Mannigfaltigkeit der Klasse C^1 mit Kodimension $d \in \mathbb{N}$.

- (a) Ein Vektor $v \in \mathbb{R}^{m+d}$ heißt *Tangentialvektor* von M im Punkt $x \in M$, wenn eine Kurve $c \in C^1((-\delta, \delta), \mathbb{R}^{m+d})$ mit

$$c(0) = x, \quad c'(0) = v \quad \text{und} \quad c((-\delta, \delta)) \subset M \quad (11.6.1)$$

existiert. Die Menge aller solcher Vektoren heißt *Tangentialraum* $T_x M$ von M im Punkt x .

- (b) Das orthogonale Komplement

$$T_x^\perp M := \{\xi \in \mathbb{R}^{m+d} : \langle \xi, v \rangle = 0 \text{ für alle } v \in T_x M\}$$

heißt *Normalraum* von M in x ; seine Elemente werden *Normalenvektoren* von M in x genannt.

Satz 11.6.7

Sei $M = f^{-1}(0) \subset \Omega \subset \mathbb{R}^{m+d}$ eine m -dimensionale Mannigfaltigkeit mit Kodimension $d \in \mathbb{N}$ und $f = (f_1, \dots, f_d) \in C^1(\Omega, \mathbb{R}^d)$. Dann gelten für alle $x \in M$:

- (a) $T_x M, T_x^\perp M$ sind lineare Unterräume des \mathbb{R}^{m+d} mit $\dim T_x M = m, \dim T_x^\perp M = d$.
- (b) $T_x^\perp M = \text{span}\{\nabla f_1(x), \dots, \nabla f_d(x)\}$.
- (c) $T_x M = \{v \in \mathbb{R}^{m+d} : Df(x)v = 0\}$, d.h. $T_x M = \text{Kern}(df(x))$.

Beweis. (a) Mittels lokaler Graphendarstellung \rightarrow Ergänzungen.

- (b) Sind $x \in M, v \in T_x M$ und $c \in C^1((-\delta, \delta), \mathbb{R}^{m+d})$ mit (11.6.1) beliebig, so finden wir für jede Komponentenfunktion f_k von f :

$$0 = \frac{d}{dt} f_k(c(t)) \Big|_{t=0} = \langle \nabla f_k(x), v \rangle, \quad k = 1, \dots, d.$$

Folglich gilt $\nabla f_1(x), \dots, \nabla f_d(x) \in T_x^\perp M$. Wegen $\text{rang } Df(x) = d$ sind die Vektoren $\nabla f_1(x), \dots, \nabla f_d(x)$ linear unabhängig, aus $\dim T_x^\perp M = d$ folgt somit

$$\text{span}\{\nabla f_1(x), \dots, \nabla f_d(x)\} = T_x^\perp M.$$

- (c) Für beliebiges $x \in M$ erhalten wir nun sofort die Äquivalenzen

$$\begin{aligned} v \in T_x M = (T_x^\perp M)^\perp &\stackrel{(b)}{\Leftrightarrow} \langle \nabla f_k(x), v \rangle = 0 \text{ für alle } k = 1, \dots, d \\ &\Leftrightarrow Df(x)v = 0, \end{aligned}$$

also die Behauptung. □

Beispiel 11.6.8. Die Einheitssphäre $S^{n-1} = \{x \in \mathbb{R}^n : \|x\| = 1\} \subset \mathbb{R}^n$ ist eine $(n-1)$ -dimensionale Mannigfaltigkeit der Klasse C^∞ , denn: Mit $f(x) := \|x\|^2 - 1 \in C^\infty(\mathbb{R}^{(n-1)+1}, \mathbb{R}^1)$ können wir $S^{n-1} = f^{-1}(0)$ schreiben, und wir haben $\nabla f(x) = 2x \neq 0$ für alle $x \in S^{n-1}$. Darüber hinaus entnehmen wir Satz 11.6.7:

$$\begin{aligned} T_x S^{n-1} &= \{v \in \mathbb{R}^n : \langle x, v \rangle = 0\}, \\ T_x^\perp S^{n-1} &= \text{span}\{x\} \quad \text{für } x \in S^{n-1}. \end{aligned}$$

Abschließend kehren wir noch einmal zum Thema „lokale Extrema“ zurück: Betrachtet wird jetzt eine reellwertige Funktion $\phi \in C^1(\Omega)$ auf einer offenen Menge Ω . Wir suchen nach Extrema von $\phi|_M$ für eine Teilmenge $M \subset \Omega$. Die Aufgabe unterliegt also der *Nebenbedingung*, dass nur Punkte aus M für den Vergleich der Funktionswerte zulässig sind. Da M nicht offen sein muss, sind die Methoden aus den Abschnitten 11.1, 11.2 nicht anwendbar. Für den Fall einer Mannigfaltigkeit M erhalten wir den

Satz 11.6.9: Extrema unter Nebenbedingungen, LAGRANGESche Multiplikatoren

Es sei $\phi : \Omega \rightarrow \mathbb{R} \in C^1(\Omega)$ auf der offenen Menge $\Omega \subset \mathbb{R}^n$, $n = m + d$, gegeben. Weiter sei $M = f^{-1}(0)$ mit einer Funktion $f = (f_1, \dots, f_d) \in C^1(\Omega, \mathbb{R}^d)$ eine m -dimensionale Mannigfaltigkeit. Schließlich besitze die Einschränkung $\phi|_M : M \rightarrow \mathbb{R}$ in $x^0 \in M$ ein lokales Extremum. Dann gibt es Zahlen $\lambda_1, \dots, \lambda_d \in \mathbb{R}$, so dass gilt

$$\nabla \phi(x^0) + \lambda_1 \nabla f_1(x^0) + \dots + \lambda_d \nabla f_d(x^0) = 0, \quad (11.6.2)$$

d.h. $x^0 \in M \subset \Omega$ ist kritischer Punkt der Funktion $\psi := \phi + \lambda_1 f_1 + \dots + \lambda_d f_d \in C^1(\Omega)$. Die Zahlen $\lambda_1, \dots, \lambda_d$ heißen *LAGRANGESche Multiplikatoren*.

Beweis. Sei $v \in T_{x^0}M$ beliebig und $c \in C^1((-\delta, \delta), \mathbb{R}^n)$ mit (11.6.1) für $x = x^0$ gewählt. Dann wird $\phi \circ c : (-\delta, \delta) \rightarrow \mathbb{R}$ in $t = 0$ extremal, der Satz von FERMAT, Satz 7.3.3, liefert somit:

$$0 = \left. \frac{d}{dt} \phi(c(t)) \right|_{t=0} = \langle \nabla \phi(x^0), v \rangle.$$

Also ist $\nabla \phi(x^0) \in T_{x^0}^\perp M$ richtig. Nach Satz 11.6.7 (b) gibt es daher Zahlen $\mu_1, \dots, \mu_d \in \mathbb{R}$ mit

$$\nabla \phi(x^0) = \sum_{k=1}^d \mu_k \nabla f_k(x^0),$$

und (11.6.2) folgt mit $\lambda_k := -\mu_k$ für $k = 1, \dots, d$. □

Bemerkung 11.6.10. ► Neben der notwendigen Bedingung 1. Ordnung aus Satz 11.6.9 gibt es auch Extremalbedingungen 2. Ordnung für Extrema unter Nebenbedingungen (→ *Ergänzungen*).

- Geometrisch besagt Satz 11.6.9: *Notwendig für eine Extremalstelle $x^0 \in M$ von $\phi|_M$ ist, dass der Vektor $\nabla \phi(x^0)$ Normalenvektor von M im Punkt x^0 ist; vgl. Satz 11.6.7 (b).*
- In Satz 11.6.9 treten $n + d = m + 2d$ Unbekannte $x_1^0, \dots, x_n^0, \lambda_1, \dots, \lambda_d \in \mathbb{R}$ auf, die aus den $n + d$ Gleichungen

$$\begin{aligned} \phi_{x_j}(x^0) + \sum_{k=1}^d \lambda_k f_{k,x_j}(x^0) &= 0, \quad j = 1, \dots, n, \\ f_k(x^0) &= 0, \quad k = 1, \dots, d, \end{aligned}$$

zu bestimmen sind. Oft ist es sinnvoll, zunächst alle kritischen Punkte $x^0 \in \Omega$ der Funktion $\psi = \phi + \lambda_1 f_1 + \dots + \lambda_d f_d$ mit beliebigen $\lambda_1, \dots, \lambda_d \in \mathbb{R}$ zu bestimmen und anschließend jene auszuwählen, die zusätzlich die *Bindungsgleichungen* $f_k(x^0) = 0$, $k = 1, \dots, d$, erfüllen.

- Die Nebenbedingung $x^0 \in M$ wird durch Gleichungen bestimmt (nämlich die Bindungsgleichungen, die die Mannigfaltigkeit definieren). Daneben können mit dieser Methode oft auch Nebenbedingungen der Form

$$f_k \leq 0, \quad k = 1, \dots, d,$$

behandelt werden. Dazu untersucht man die Fälle $f_k < 0$ und $f_k = 0$ getrennt voneinander. Wir illustrieren das Vorgehen am abschließenden

Beispiel 11.6.11. Seien $a \in S^{n-1}$, $R > 0$ und $\phi \in C^\infty(\mathbb{R}^n)$ mit $\phi(x) = \langle x, a \rangle^2 + 1$, $x \in \mathbb{R}^n$, gegeben. Gesucht ist das Maximum von ϕ auf $\overline{B_R(0)} = \{x \in \mathbb{R}^n : |x| \leq R\}$.

Lösung: Da ϕ stetig ist und $\overline{B_R(0)}$ kompakt ist, wird das Maximum $\sigma := \sup \{\phi(x) : x \in \overline{B_R(0)}\}$ in einem Punkt $x^0 \in \overline{B_R(0)}$ angenommen, d.h. $\phi(x^0) = \sigma$. Wir unterscheiden zwei mögliche Fälle:

- $x^0 \in B_R(0)$. Dann wäre x^0 auch lokaler Maximierer von $\phi|_{B_R(0)}$, wir können also Satz 11.1.1 auf der offenen Menge $B_R(0)$ anwenden. Wegen

$$\nabla \phi(x) = 2\langle x, a \rangle a, \quad x \in \mathbb{R}^n,$$

ist $K := \{x \in B_R(0) : \langle x, a \rangle = 0\}$ die Menge der kritischen Punkte von $\phi|_{B_R(0)}$.

- $x^0 \in \partial B_R(0) = f^{-1}(0)$ mit $f(x) := \|x\|^2 - R^2$, $x \in \mathbb{R}^n$. Wegen $\nabla f(x) = 2x \neq 0$ für $x \neq 0$ ist $\partial B_R(0)$ eine $n - 1$ -dimensionale Mannigfaltigkeit (der Klasse C^∞). Nach Voraussetzung wäre auch $\phi|_{\partial B_R(0)}$ in x^0 maximal, gemäß Satz 11.6.9 existierte also ein $\lambda \in \mathbb{R}$ mit

$$0 = \phi(x^0) + \lambda \nabla f(x^0) = 2\langle x^0, a \rangle a + 2\lambda x^0. \quad (11.6.3)$$

Für $\langle x^0, a \rangle = 0$ ist das System mit $\lambda = 0$ erfüllt; wegen $x^0 \in \partial B_R(0)$ erweitert sich die Menge K der Kandidaten zu

$$\overline{K} := \{x \in \overline{B_R(0)} : \langle x, a \rangle = 0\}.$$

Für $\langle x^0, a \rangle \neq 0$ liefert Multiplikation von (11.6.3) mit a sofort $\lambda = -\|a\|^2 = -1$ und somit $x^0 = \langle x^0, a \rangle a$, d.h. x^0 ist ein Vielfaches von a . Wegen $x^0 \in \partial B_R(0)$, also $\|x^0\|^2 = R^2$, folgt schließlich $\langle x^0, a \rangle^2 = R^2$ und somit $x^0 = \pm Ra$.

Abschließend vergleichen wir die Funktionswerte der Kandidaten $\overline{K} \cup \{\pm Ra\}$ für globale Maximalstellen. Wegen

$$\phi(\pm Ra) = R^2 + 1 > 1 = \phi(x), \quad x \in \overline{K},$$

folgt $\sigma = R^2 + 1$ und die globalen (und einzigen lokalen) Maximierer von ϕ auf $\overline{B_R(0)}$ sind $\pm Ra \in \partial B_R(0)$.

Kapitel 12.

Das RIEMANNSche Integral im \mathbb{R}^n

In diesem Kapitel verallgemeinern wir das RIEMANNSche Integral auf mehrdimensionale Mengen. In Abschnitt 12.1 übertragen wir den RIEMANNSchen Integralbegriff von kompakten Intervallen auf kompakte Quader inklusive der Integrabilitätskriterien. In Abschnitt 12.2 verallgemeinern wir zunächst die bekannten Rechenregeln für Integrale auf Quader; anschließend beweisen wir den Satz über die iterierte Integration, mit dem mehrdimensionale auf eindimensionale Integrale zurückgeführt werden können. Abschnitt 12.3 enthält wichtige technische Grundlagen, insbesondere den Satz von HEINE-BOREL, Satz 12.3.3. In Abschnitt 12.4 wenden wir diese Hilfsmittel an, um die Menge der Unstetigkeitsstellen integrierbare Funktionen auf Quadern zu charakterisieren. Dies hilft uns in Abschnitt 12.5 den Begriff der *quadrierbaren* Mengen besser zu verstehen - hier wird das RIEMANN-Integral auf eben diesen quadrierbaren Mengen definiert. In Abschnitt 12.6 führen wir spezielle quadrierbare Mengen ein, die *Normalbereiche*, und erweitern die bekannte Additivität des eindimensionalen Integrals auf Integrale über quadrierbaren Mengen. Die abschließenden beiden Abschnitte sind der *Transformationsformel* gewidmet. In Abschnitt 12.7 beschreiben wir zunächst kurz die Beweisidee der Formel für sogenannte Testfunktionen, bevor wir in Abschnitt 12.8 den Weg zur allgemeinen Transformationsformel beschreiben und dabei den Begriff des uneigentlichen mehrdimensionalen Integrals erklären.

12.1. Das Integral auf Quadern

Das Konzept des RIEMANNSchen Integrals aus Kapitel 9 lässt sich von kompakten Intervallen direkt auf *Quader* übertragen. Damit meinen wir:

Definition 12.1.1

Sind $I_j = [a_j, b_j] \subset \mathbb{R}$ nichtleere, abgeschlossene Intervalle für $j = 1, \dots, n$, so heißt

$$Q := I_1 \times \dots \times I_n = \{(x_1, \dots, x_n) \in \mathbb{R}^n : x_j \in [a_j, b_j], j = 1, \dots, n\}$$

Quader im \mathbb{R}^n . Mit

$$\text{vol}(Q) = |Q| := \prod_{j=1}^n |I_j| = \prod_{j=1}^n (b_j - a_j)$$

bezeichnen wir das **Volumen** oder den **Inhalt** von Q .

Zur Definition des Integrals beginnen wir wieder mit der

Definition 12.1.2: Zerlegungen

Es seien $\mathcal{Z}^{(j)}$ Zerlegungen von $I_j = [a_j, b_j]$ für $j = 1, \dots, n$ mit

$$a_j = x_{j,0} < x_{j,1} < \dots < x_{j,N_j} = b_j.$$

Dann heißt $\mathcal{Z} := \mathcal{Z}^{(1)} \times \dots \times \mathcal{Z}^{(n)}$ **Zerlegung von Q** . Wir schreiben $I_{j,\alpha_j} = [x_{j,\alpha_j-1}, x_{j,\alpha_j}]$ mit $\alpha_j \in \{1, \dots, N_j\}$, $j = 1, \dots, n$, für das α_j -te Teilintervall der Zerlegung $\mathcal{Z}^{(j)}$ und erklären die Teilquader

$$Q_\alpha := I_{1,\alpha_1} \times \dots \times I_{n,\alpha_n}, \quad \alpha = (\alpha_1, \dots, \alpha_n) \in A,$$

wobei $A = A(\mathcal{Z})$ die Menge der auftretenden Multiindizes angibt:

$$A := \{\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n : 1 \leq \alpha_j \leq N_j, j = 1, \dots, n\}.$$

Wir schreiben dann auch $\mathcal{Z} = \{Q_\alpha\}_{\alpha \in A}$. Schließlich setzen wir

$$\Delta(\mathcal{Z}) := \max \{\Delta \mathcal{Z}^{(1)}, \dots, \Delta \mathcal{Z}^{(n)}\}$$

für die **Feinheit** der Zerlegung \mathcal{Z} .

Definition 12.1.3: Unter-, Ober- und Zwischensumme

Sei $Q \subset \mathbb{R}^n$ ein Quader und $f : Q \rightarrow \mathbb{R}$ beschränkt. Weiter sei $\mathcal{Z} = \{Q_\alpha\}_{\alpha \in A}$ eine Zerlegung von Q .

(a) Mit den Abkürzungen

$$\underline{m}_\alpha := \inf_{Q_\alpha} f, \quad \overline{m}_\alpha := \sup_{Q_\alpha} f, \quad \alpha \in A,$$

bilden wir die *Untersumme*

$$\underline{S}_Z(f) := \sum_{\alpha \in A} \underline{m}_\alpha |Q_\alpha|$$

und die *Obersumme*

$$\overline{S}_Z(f) := \sum_{\alpha \in A} \overline{m}_\alpha |Q_\alpha|$$

von f zur Zerlegung Z .

- (b) Aus jedem Teilquader Q_α wählen wir ein $\xi_\alpha \in Q_\alpha$, $\alpha \in A$, und schreiben $\xi = (\xi_\alpha)_{\alpha \in A}$. Dann nennen wir

$$S_Z(f) = S_Z(f, \xi) := \sum_{\alpha \in A} f(\xi_\alpha) |Q_\alpha|$$

eine *RIEMANNSCHE Zwischensumme* von f zu Z und ξ .

Bemerkung 12.1.4. (a) Die in Definition 10.1.2 beschriebene Zerlegung $Z = Z^{(1)} \times \dots \times Z^{(n)} = \{Q_\alpha\}_{\alpha \in A}$ hat genau $N_1 \cdot N_2 \cdot \dots \cdot N_n = \# A$ Teilquader Q_α , die sich nach Konstruktion offenbar nicht überlappen, also höchstens Randpunkte gemein haben, und für die gilt

$$|Q| = \sum_{\alpha \in A} |Q_\alpha|. \quad (12.1.1)$$

- (b) Offenbar gilt

$$\underline{S}_Z(f) \leq S_Z(f) \leq \overline{S}_Z(f)$$

für jede Riemannsche Zwischensumme. Ferner können wir zu jedem $\varepsilon > 0$ und jeder Zerlegung Z Zwischenwerte $\underline{\xi} = (\underline{\xi}_\alpha)_{\alpha \in A}$, $\overline{\xi} = (\overline{\xi}_\alpha)_{\alpha \in A}$ so angeben, dass gilt

$$S_Z(f, \underline{\xi}) < \underline{S}_Z(f) + \varepsilon, \quad S_Z(f, \overline{\xi}) > \overline{S}_Z(f) - \varepsilon.$$

Definition 12.1.5

- Eine Zerlegung $Z_* = Z_*^{(1)} \times \dots \times Z_*^{(n)}$ von Q heißt *Verfeinerung* einer Zerlegung $Z = Z^{(1)} \times \dots \times Z^{(n)}$ von Q , wenn $Z_*^{(j)}$ Verfeinerungen von $Z^{(j)}$ sind für alle $j = 1, \dots, n$.
- Eine *gemeinsame Verfeinerung* $Z \vee Z_*$ von Z und Z_* ist erklärt als

$$Z \vee Z_* := (Z^{(1)} \vee Z_*^{(1)}) \times \dots \times (Z^{(n)} \vee Z_*^{(n)}).$$

Völlig analog zu Hilfssatz 9.1.6 beweist man den

Hilfssatz 12.1.6

(a) Ist \mathcal{Z}_* Verfeinerung der Zerlegung \mathcal{Z} von Q , so gilt

$$\underline{S}_{\mathcal{Z}}(f) \leq \underline{S}_{\mathcal{Z}_*}(f) \leq \bar{S}_{\mathcal{Z}_*}(f) \leq \bar{S}_{\mathcal{Z}}(f).$$

(b) Sind $\mathcal{Z}_1, \mathcal{Z}_2$ zwei beliebige Zerlegungen von Q , so gilt

$$\underline{S}_{\mathcal{Z}_1}(f) \leq \bar{S}_{\mathcal{Z}_2}(f).$$

Definition 12.1.7: Unter- und Oberintegral

Ist $f : Q \rightarrow \mathbb{R}$ beschränkt, so erklären wir das *Unterintegral* $\underline{\mathcal{I}}(f)$ bzw. *Oberintegral* $\bar{\mathcal{I}}(f)$ von f als

$$\begin{aligned} \underline{\mathcal{I}}(f) &= \underline{\mathcal{I}}_Q(f) := \sup \{ \underline{S}_{\mathcal{Z}}(f) : \mathcal{Z} \text{ ist Zerlegung von } Q \}, \\ \bar{\mathcal{I}}(f) &= \bar{\mathcal{I}}_Q(f) := \inf \{ \bar{S}_{\mathcal{Z}}(f) : \mathcal{Z} \text{ ist Zerlegung von } Q \}. \end{aligned}$$

Bemerkung 12.1.8. Für jede beschränkte Funktion $f : Q \rightarrow \mathbb{R}$ und jede Zerlegung \mathcal{Z} von Q gilt nach Hilfssatz 12.1.6 (a):

$$-\infty < |Q| \inf_Q f \leq \underline{S}_{\mathcal{Z}}(f) \leq \bar{S}_{\mathcal{Z}}(f) \leq |Q| \sup_Q f < +\infty.$$

Also sind $\underline{\mathcal{I}}(f), \bar{\mathcal{I}}(f) \in \mathbb{R}$ wohldefiniert. Hilfssatz 12.1.6 (b) entnehmen wir noch

$$\underline{S}_{\mathcal{Z}}(f) \leq \underline{\mathcal{I}}(f) \leq \bar{\mathcal{I}}(f) \leq \bar{S}_{\mathcal{Z}}(f) \quad (12.1.2)$$

für alle Zerlegungen \mathcal{Z} von Q .

Definition 12.1.9: Integrierbarkeit und Integral

Eine beschränkte Funktion $f : Q \rightarrow \mathbb{R}$ auf dem Quader $Q \subset \mathbb{R}^n$ heißt *(RIEMANN-)integrierbar*, wenn $\underline{\mathcal{I}}(f) = \bar{\mathcal{I}}(f)$ erfüllt ist. Wir setzen dann

$$\mathcal{I}(f) := \underline{\mathcal{I}}(f) = \bar{\mathcal{I}}(f)$$

für das **(RIEMANN-)Integral** von f auf Q und schreiben auch

$$\mathcal{I}(f) = \int_Q f(x) dx = \int_Q f d(x_1, \dots, x_n) = \int_Q f dV,$$

wobei $dV = dx = d(x_1, \dots, x_n)$ *Volumenelement* genannt wird. Die Klasse aller RIEMANN-integrierbaren Funktionen auf Q bezeichnen wir mit $\mathcal{R}(Q)$.

Beispiel 12.1.10. Für eine konstante Funktion $f : Q \rightarrow \mathbb{R}, x \mapsto c$, auf einem Quader $Q \subset \mathbb{R}^n$ gilt

$$\underline{S}_Z(f) = \sum_{\alpha \in A} c|Q_\alpha| \stackrel{(12.1.1)}{=} c|Q|, \quad \bar{S}_Z(f) = c|Q|,$$

wobei $Z = \{Q_\alpha\}_{\alpha \in A}$ eine beliebige Zerlegung von Q ist. Folglich haben wir $\underline{I}(f) = \bar{I}(f) = c|Q|$ und somit $f \in \mathcal{R}(Q)$ sowie

$$\int_Q c dx = c|Q|.$$

Auch die drei Integrabilitätskriterien I-III für beschränkte Funktionen $f : [a, b] \rightarrow \mathbb{R}$ lassen sich direkt übertragen. Zur Formulierung nennen wir eine Folge $\{Z_p\}_p$ von Zerlegungen von Q wieder *ausgezeichnet*, wenn $\{\Delta(Z_p)\}_p$ eine Nullfolge ist; siehe Definition 9.2.4.

Satz 12.1.11: Integrabilitätskriterien

Für eine Funktion $f : Q \rightarrow \mathbb{R}$ auf einem Quader $Q \subset \mathbb{R}^n$ sind äquivalent:

- (i) $f \in \mathcal{R}(Q)$.
- (ii) Für alle $\varepsilon > 0$ existiert eine Zerlegung Z von Q mit $\bar{S}_Z(f) - \underline{S}_Z(f) < \varepsilon$.
- (iii) Für alle $\varepsilon > 0$ existiert ein $\delta = \delta(\varepsilon) > 0$, so dass für jede Zerlegung Z von Q mit $\Delta(Z) < \delta$ gilt: $\bar{S}_Z(f) - \underline{S}_Z(f) < \varepsilon$.
- (iv) Für jede ausgezeichnete Zerlegungsfolge $\{Z_p\}_p$ und jede Wahl von Zwischenwerten $\{\xi_p\}_p$ konvergiert die Folge $\{S_{Z_p}(f, \xi_p)\}_p$.

Ist f integrierbar über Q , so gilt für jede ausgezeichnete Zerlegungsfolge $\{Z_p\}_p$ und jede Wahl von Zwischenwerten $\{\xi_p\}_p$:

$$\int_Q f(x) dx = \lim_{p \rightarrow \infty} S_{Z_p}(f, \xi_p) = \lim_{p \rightarrow \infty} \underline{S}_{Z_p}(f) = \lim_{p \rightarrow \infty} \bar{S}_{Z_p}(f). \quad (12.1.3)$$

Beweis. Wörtliches Übertragen bzw. leichte Anpassung (bei Kriterium (iii)) der Beweise der Sätze 9.2.1, 9.2.3 und 9.2.5. □

12.2. Rechenregeln und iterierte Integration auf Quadern

Mit Hilfe der Integrabilitätskriterien aus Satz 12.1.11 lassen sich viele Ergebnisse der Integration über Intervalle direkt auf Integrale über Quader übertragen. Wir verzichten daher auch auf die Beweise der folgenden beiden Sätze; vgl. Sätze 9.2.6, 9.2.9, 9.2.10:

Satz 12.2.1: Rechenregeln

Sei $Q \subset \mathbb{R}^n$ ein Quader im Sinne von Definition 12.1.1.

(a) Gilt $f, g \in \mathcal{R}(Q)$, so ist $\alpha f + \beta g \in \mathcal{R}(Q)$ für beliebige $\alpha, \beta \in \mathbb{R}$ erfüllt. Es gilt dann

$$\mathcal{I}(\alpha f + \beta g) = \alpha \mathcal{I}(f) + \beta \mathcal{I}(g);$$

$\mathcal{R}(Q)$ ist also ein reeller Vektorraum.

(b) Sind $f, g \in \mathcal{R}(Q)$ mit $f \leq g$ auf Q gegeben, so folgt

$$\mathcal{I}(f) \leq \mathcal{I}(g).$$

(c) Mit $f \in \mathcal{R}(Q)$ ist auch $|f| \in \mathcal{R}(Q)$ richtig mit

$$|\mathcal{I}(f)| \leq \mathcal{I}(|f|).$$

(d) Sind $f, g \in \mathcal{R}(Q)$, so ist auch $f \cdot g \in \mathcal{R}(Q)$ erfüllt. Es gelten dann

$$\mathcal{I}(|fg|) \leq \left(\sup_Q |g| \right) \cdot \mathcal{I}(|f|)$$

sowie die *SCHWARZsche Ungleichung*

$$\mathcal{I}(|fg|) \leq \sqrt{\mathcal{I}(f^2)} \sqrt{\mathcal{I}(g^2)}.$$

(e) Gilt $f, g \in \mathcal{R}(Q)$ sowie $|g| \geq c > 0$ auf Q mit einer Konstanten $c > 0$, so folgt auch $\frac{f}{g} \in \mathcal{R}(Q)$ mit

$$\mathcal{I}\left(\left|\frac{f}{g}\right|\right) \leq \frac{1}{c} \mathcal{I}(|f|).$$

Satz 12.2.2

Es gilt $C^0(Q) \subset \mathcal{R}(Q)$.

Wir wollen nun Integrale über Quader im \mathbb{R}^n auf Integrale über niederdimensionale Quader und letztlich Intervalle zurückführen. Sei dazu $f = f(x, y) : Q \times R \rightarrow \mathbb{R} \in \mathcal{R}(Q \times R)$ für Quader $Q \subset \mathbb{R}^q$, $R \subset \mathbb{R}^r$ mit $q + r = n$ vorgelegt. Wir schreiben dann

$$\underline{\varphi}(x) := \underline{\mathcal{I}}_R(f(x, \cdot)), \quad \overline{\varphi}(x) := \overline{\mathcal{I}}_R(f(x, \cdot)), \quad x \in Q, \quad (12.2.1)$$

für das Unter- bzw. Oberintegral von $f(x, \cdot) : R \rightarrow \mathbb{R}$ zu $x \in Q$, welche nicht übereinstimmen müssen. Die so erklärten Funktionen $\underline{\varphi}, \overline{\varphi} : Q \rightarrow \mathbb{R}$ sind beschränkt und erfüllen $\underline{\varphi} \leq \overline{\varphi}$ auf Q gemäß (12.1.2). Es gilt nun der

Satz 12.2.3: Iterierte Integration

Seien $Q \subset \mathbb{R}^q$, $R \subset \mathbb{R}^r$ Quader. Für beliebiges $f = f(x, y) : Q \times R \rightarrow \mathbb{R} \in \mathcal{R}(Q \times R)$ sind die in (12.2.1) erklärten Funktionen $\underline{\varphi}, \overline{\varphi} : Q \rightarrow \mathbb{R}$ integrierbar über Q , und es gilt

$$\int_{Q \times R} f(x, y) d(x, y) = \int_Q \underline{\varphi}(x) dx = \int_Q \overline{\varphi}(x) dx.$$

Beweis. Wir betrachten Zerlegungen \mathcal{Z}_Q von Q in Teilquader Q_α , $\alpha \in A$, und \mathcal{Z}_R von R in Teilquader R_β , $\beta \in B$. Dann ist $\mathcal{Z}_T := \mathcal{Z}_Q \times \mathcal{Z}_R$ eine Zerlegung von $T := Q \times R$ mit Teilquadranten $T_{\alpha\beta} := Q_\alpha \times R_\beta$, $(\alpha, \beta) \in A \times B$. Umgekehrt lässt sich so jede Zerlegung von T durch Zerlegungen von Q und R darstellen. Wir bemerken noch

$$\Delta(\mathcal{Z}_T) = \max \{ \Delta(\mathcal{Z}_Q), \Delta(\mathcal{Z}_R) \}. \quad (12.2.2)$$

Nun erklären wir die Größen

$$\underline{m}_{\alpha\beta} := \inf_{T_{\alpha\beta}} f, \quad \overline{m}_{\alpha\beta} := \sup_{T_{\alpha\beta}} f, \quad (\alpha, \beta) \in A \times B.$$

Für beliebiges $x \in Q_\alpha$ mit einem $\alpha \in A$ folgt dann

$$\sum_{\beta \in B} \underline{m}_{\alpha\beta} |R_\beta| \leq \sum_{\beta \in B} \left(\inf_{y \in R_\beta} f(x, y) \right) |R_\beta| \leq \underline{\mathcal{I}}_R(f(x, \cdot)) = \underline{\varphi}(x)$$

und entsprechend

$$\sum_{\beta \in B} \overline{m}_{\alpha\beta} |R_\beta| \geq \overline{\varphi}(x) \geq \underline{\varphi}(x) \geq \sum_{\beta \in B} \underline{m}_{\alpha\beta} |R_\beta|, \quad x \in Q_\alpha. \quad (12.2.3)$$

Seien $\xi = \{\xi_\alpha\}_{\alpha \in A}$ beliebige Zwischenwerte. Setzen wir $x = \xi_\alpha \in Q_\alpha$ in (12.2.3) ein, multiplizieren mit $|Q_\alpha|$ und summieren über $\alpha \in A$, so folgt (man beachte $|T_{\alpha\beta}| = |Q_\alpha| |R_\beta|$):

$$\underline{S}_{\mathcal{Z}_T}(f) \leq S_{\mathcal{Z}_Q}(\underline{\varphi}, \xi) \leq S_{\mathcal{Z}_Q}(\overline{\varphi}, \xi) \leq \overline{S}_{\mathcal{Z}_T}(f). \quad (12.2.4)$$

Sei nun $\{Z_{Q,p}\}_p$ eine ausgezeichnete Zerlegungsfolge von Q und $\{\xi_p\}_p$ eine zugehörige Folge von Zwischenwerten. Ist weiter $\{Z_{R,p}\}_p$ eine ausgezeichnete Zerlegungsfolge von R und setzen wir $Z_{T,p} = Z_{Q,p} \times Z_{R,p}$, $p \in \mathbb{N}$, so ist $\{Z_{T,p}\}_p$ gemäß (12.2.2) ausgezeichnete Zerlegungsfolge von T . Wegen $f \in \mathcal{R}(T)$ liefern (12.2.4) und Satz 12.1.11 $\underline{\varphi}, \bar{\varphi} \in \mathcal{R}(Q)$ sowie

$$\int_T f(x, y) d(x, y) = \lim_{p \rightarrow \infty} S_{Z_{T,p}}(f) = \lim_{p \rightarrow \infty} S_{Z_{Q,p}}(\underline{\varphi}, \xi_p) = \int_Q \underline{\varphi}(x) dx = \int_Q \bar{\varphi}(x) dx.$$

Damit ist alles gezeigt. \square

Bemerkung 12.2.4. $\underline{\varphi}$ und $\bar{\varphi}$ stimmen genau dann in Q überein, wenn $f(x, \cdot) : R \rightarrow \mathbb{R}$ für jedes $x \in Q$ integrierbar ist; dann ist $\underline{\varphi}(x) = \mathcal{I}_R(f(x, \cdot)) = \bar{\varphi}(x)$ auf Q richtig. Da nach Satz 12.2.2 jede stetige Funktion integrierbar ist, erhalten wir sofort den folgenden

Satz 12.2.5: Iterierte Integration stetiger Funktionen

Ist $f = f(x, y) : Q \times R \rightarrow \mathbb{R} \in C^0(Q \times R)$ gegeben, so ist $f(x, \cdot) : R \rightarrow \mathbb{R}$ für jedes $x \in Q$ über R integrierbar und es gilt

$$\int_{Q \times R} f(x, y) d(x, y) = \int_Q \left(\int_R f(x, y) dy \right) dx. \quad (12.2.5)$$

Bemerkung 12.2.6. (a) Durch Umbezeichnung $x \leftrightarrow y$, also Vertauschen der Koordinaten, entnimmt man (12.2.5):

$$\int_Q \left(\int_R f(x, y) dy \right) dx = \int_{Q \times R} f(x, y) d(x, y) = \int_R \left(\int_Q f(x, y) dx \right) dy. \quad (12.2.6)$$

Auf die Reihenfolge der Integration kommt es also nicht an!

(b) Ist $Q = I_1 \times \dots \times I_n \subset \mathbb{R}^n$ ein Quader und $f : Q \rightarrow \mathbb{R} \in C^0(Q)$, so folgt aus Satz 12.2.5

$$\int_Q f d(x_1, \dots, x_n) = \int_{I_1} \left(\int_{I_2} \left(\dots \left(\int_{I_n} f dx_n \right) \dots \right) dx_2 \right) dx_1. \quad (12.2.7)$$

Wir können also jedes Integral einer stetigen Funktion auf einem Quader im \mathbb{R}^n durch sukzessives eindimensionales Integrieren auswerten. Wie in (12.2.6) spielt die Reihenfolge der Integration dabei keine Rolle.

Beispiel 12.2.7. (a) Sei $f(x, y) = xy$, $Q = [0, 2] \times [0, 1]$. Dann gilt

$$\begin{aligned} \int_Q f \, d(x, y) &\stackrel{(12.2.7)}{=} \int_0^2 \left(\int_0^1 xy \, dy \right) dx \stackrel{(12.2.6)}{=} \int_0^1 \left(\int_0^2 xy \, dx \right) dy \\ &= \int_0^1 \left(y \int_0^2 x \, dx \right) dy = \int_0^1 2y \, dy = 1. \end{aligned}$$

(b) Sei $f(x, y) = xye^{x^2y}$, $Q = [0, 1] \times [0, 1]$. Dann gilt

$$\begin{aligned} \int_Q f \, d(x, y) &= \int_0^1 \left(\int_0^1 xye^{x^2y} \, dx \right) dy = \int_0^1 \left(\int_0^1 \frac{\partial}{\partial x} \left[\frac{1}{2} e^{x^2y} \right] dx \right) dy \\ &= \int_0^1 \frac{1}{2} (e^y - 1) \, dy = \frac{1}{2} (e - 2). \end{aligned}$$

Wir schließen mit der üblichen koordinatenweisen Übertragung des Integralbegriffs auf vektor- bzw. komplexwertige Funktionen:

Definition 12.2.8

Für einen Quader $Q \subset \mathbb{R}^n$ erklären wir den *Raum der integrierbaren Funktionen* $f : Q \rightarrow \mathbb{R}^d$,

$$\mathcal{R}(Q, \mathbb{R}^d) := \{f = (f_1, \dots, f_d) : Q \rightarrow \mathbb{R}^d : f_l \in \mathcal{R}(Q), l = 1, \dots, d\},$$

und setzen

$$\int_Q f \, dx := \left(\int_Q f_1 \, dx, \dots, \int_Q f_d \, dx \right).$$

Speziell für $n = 2$ identifizieren wir $\mathcal{R}(Q, \mathbb{C}) = \mathcal{R}(Q, \mathbb{R}^2)$ und haben

$$\int_Q f \, dx = \int_Q \operatorname{Re} f \, dx + i \int_Q \operatorname{Im} f \, dx.$$

Bemerkung 12.2.9. Wo sinnvoll, übertragen sich die Aussagen der Sätze dieses Abschnitts auf vektor- bzw. komplexwertige Funktionen.

12.3. Der Satz von HEINE-BOREL, Nullmengen und Mengen vom Inhalt Null

Wir stellen hier technisches Rüstzeug bereit, um den Integralbegriff in den kommenden Abschnitten weiter entwickeln zu können. Zentral ist dabei der wichtige *Satz von HEINE-BOREL*, für dessen Formulierung wir zunächst noch einen Begriff benötigen:

Definition 12.3.1

Ist $M \subset \mathbb{R}^n$ eine Menge und ist J eine beliebige Indexmenge, so heißt eine Familie $\mathcal{F} = \{M_j\}_{j \in J}$ von Mengen $M_j \subset \mathbb{R}^n$ *Überdeckung* von M , wenn gilt

$$M \subset \bigcup_{j \in J} M_j.$$

Die Überdeckung heißt *offen*, wenn alle Mengen M_j , $j \in J$, offen sind. Die Überdeckung heißt *endlich* (bzw. *abzählbar*), wenn die Indexmenge J endlich (bzw. abzählbar) ist.

Beispiel 12.3.2. Für jedes nichtleere $M \subset \mathbb{R}^n$ ist die Familie $\mathcal{F} = \{B_r(x) : x \in M\}$ mit Kugeln $B_r(x)$ beliebiger Radien $r = r(x) > 0$ eine offene Überdeckung. Eine (triviale) endliche Überdeckung von M ist immer $\mathcal{F} = \{\mathbb{R}^n\}$.

Satz 12.3.3: HEINE-BOREL

Eine Menge $K \subset \mathbb{R}^n$ ist genau dann kompakt, wenn sich aus jeder offenen Überdeckung von K eine endliche Überdeckung von K auswählen lässt.

Bemerkung 12.3.4. Die angegebene äquivalente Eigenschaft wird als *HEINE-BOREL-Eigenschaft* bezeichnet. Sie wird insbesondere in unendlich-dimensionalen Räumen als Definition für Kompaktheit verwendet.

Beweis von Satz 12.3.3. ▶ „ \Leftarrow “: *Ergänzungen.*

- ▶ „ \Rightarrow “: Sei $K \subset \mathbb{R}^n$ kompakt. Angenommen, es gäbe eine offene Überdeckung \mathcal{F} von K , aus der sich *keine* endliche Überdeckung auswählen lässt. Mittels einer *Würfelschachtelung* führen wir dies zum Widerspruch.

1. Da K beschränkt ist, existiert ein (abgeschlossener) Würfel $W \subset \mathbb{R}^n$ mit $K \subset W$. Durch Halbierung aller Seiten zerlegen wir W in $N := 2^n$ Teilwürfel W_1^*, \dots, W_N^* mit $|W_j^*| = 2^{-n}|W|$. \mathcal{F} ist auch Überdeckung der Mengen $W_j^* \cap K$. Und nach Annahme

existiert mindestens ein $j_1 \in \{1, \dots, N\}$, so dass keine endliche Überdeckung von $W_{j_1}^* \cap K$ aus \mathcal{F} ausgewählt werden kann. Wir schreiben $W_1 := W_{j_1}^*$.

2. Durch Halbierung der Seiten zerlegen wir W_1 in $N = 2^n$ Teilwürfel $W_1^{**}, \dots, W_N^{**}$ mit $|W_j^{**}| = 2^{-n}|W_1| = 2^{-2n}|W|$. \mathcal{F} überdeckt wieder alle $W_j^{**} \cap K$ und nach 1. existiert mindestens ein $j_2 \in \{1, \dots, N\}$, so dass $W_{j_2}^{**} \cap K$ nicht durch eine endliche Unterfamilie von \mathcal{F} überdeckt werden kann. Wir setzen dann $W_2 := W_{j_2}^{**}$.
3. Fortsetzung des Verfahrens liefert eine Folge $W_1 \supset W_2 \supset W_3 \supset \dots$ mit $|W_l| = 2^{-ln}|W|$ und folgender Eigenschaft:

Für alle $l \in \mathbb{N}$ ist \mathcal{F} offene Überdeckung von $W_l \cap K$, aus der keine endliche Überdeckung von $W_l \cap K$ ausgewählt werden kann. (12.3.1)

Insbesondere gilt also $\lim_{l \rightarrow \infty} |W_l| = 0$ und damit auch

$$\text{diam } W_l = \sqrt[n]{n} \sqrt[n]{|W_l|} \rightarrow 0 \quad (l \rightarrow \infty). \quad (12.3.2)$$

Wegen (12.3.1) gilt $W_l \cap K \neq \emptyset$ für alle $l \in \mathbb{N}$. Also existieren $x_l \in W_l \cap K$ und gemäß (12.3.2) bildet $\{x_l\}_l \subset K$ eine Cauchyfolge: Zu beliebigem $\varepsilon > 0$ existiert nämlich ein $N = N(\varepsilon) \in \mathbb{N}$ mit $\text{diam } W_l < \varepsilon$ für alle $l \geq N$. Sind dann $k, l \geq N$ gewählt, so folgt $x_k, x_l \in W_N$ und damit

$$\|x_k - x_l\| \leq \text{diam } W_N < \varepsilon \quad \text{für alle } k, l \geq N.$$

Da \mathbb{R}^n vollständig ist, existiert ein x_0 mit $\lim_{l \rightarrow \infty} x_l = x_0$. Und da K und alle W_l abgeschlossen sind, gilt $x_0 \in W_l \cap K$ für alle $l \in \mathbb{N}$.

Da \mathcal{F} Überdeckung von K ist, gibt es ein $\Omega \in \mathcal{F}$ mit $x_0 \in \Omega$. Da Ω offen ist, finden wir ein $\varrho > 0$ mit $B_\varrho(x_0) \subset \Omega$. Und wegen (12.3.2) existiert ein $l_0 \in \mathbb{N}$ mit $\text{diam } W_{l_0} < \varrho$ und somit

$$W_{l_0} \cap K \subset W_{l_0} \subset B_\varrho(x_0) \subset \Omega \in \mathcal{F}.$$

Also haben wir $W_{l_0} \cap K$ durch die endliche Teilüberdeckung $\{\Omega\}$ von \mathcal{F} überdeckt, im Widerspruch zu (12.3.1). Somit war die Annahme falsch, und der Satz ist bewiesen. \square

Wir betrachten nun Teilmengen des \mathbb{R}^n , die in gewissem Sinne „klein“ sind.

Definition 12.3.5

Sei $M \subset \mathbb{R}^n$ gegeben. Wir sagen:

- $M \subset \mathbb{R}^n$ hat den **Inhalt Null**, wenn es zu jedem $\varepsilon > 0$ Quader Q_1, \dots, Q_N mit einem $N = N(\varepsilon) \in \mathbb{N}$ so gibt, dass $\sum_{j=1}^N |Q_j| < \varepsilon$ gilt und $\{\mathring{Q}_1, \dots, \mathring{Q}_N\}$ eine endliche offene Überdeckung von M ist. Wir schreiben dann $|M| = 0$.

- M hat das **Maß Null**, falls zu jedem $\varepsilon > 0$ abzählbar viele Quader Q_1, Q_2, \dots so existieren, dass $\sum_j |Q_j| < \varepsilon$ gilt und $\{\mathring{Q}_j\}_j$ eine abzählbare offene Überdeckung von M ist. Wir schreiben dann $\text{meas } M = 0$ und nennen M **Nullmenge**.

Offenbar ist jede Menge vom Inhalt Null auch Nullmenge. Die Umkehrung ist i.A. falsch, jedoch gilt die

Folgerung 12.3.6

Eine kompakte Menge $K \subset \mathbb{R}^n$ ist genau dann Nullmenge, wenn K den Inhalt Null hat.

Beweis. Wir haben nur die Richtung „ \Rightarrow “ zu zeigen. Ist also K Nullmenge, so gibt es zu beliebigem $\varepsilon > 0$ eine Überdeckung von K durch abzählbar viele offene Quader $\{\mathring{Q}_j\}_j$ mit Gesamtvolumen $\sum_j |Q_j| < \varepsilon$. Nach dem Satz von HEINE-BOREL können wir endlich viele Quader Q_{j_1}, \dots, Q_{j_N} auswählen mit

$$K \subset \bigcup_{k=1}^N \mathring{Q}_{j_k}, \quad \sum_{k=1}^N |Q_{j_k}| \leq \sum_j |Q_j| < \varepsilon,$$

d.h. K hat den Inhalt Null, wie behauptet. \square

Beispiel 12.3.7. Ist $K \subset \mathbb{R}^m$ kompakt und $\varphi \in C^0(K, \mathbb{R}^d)$, so ist $\text{graph } \varphi \subset \mathbb{R}^{m+d}$ eine Menge vom Inhalt Null (\rightarrow Ergänzungen).

Aus bekannten Mengen vom Inhalt bzw. Maß Null kann man leicht weitere solcher Mengen konstruieren:

Hilfssatz 12.3.8

- (a) Jede Teilmenge einer Menge mit Inhalt Null (bzw. Maß Null) hat ebenfalls Inhalt Null (bzw. Maß Null).
- (b) Die Vereinigung endlich vieler Mengen vom Inhalt Null hat ebenfalls Inhalt Null.
- (c) Die Vereinigung abzählbar vieler Nullmengen ist wieder Nullmenge.

Beweis. (a), (b) Klar.

- (c) Seien $M_1, M_2, \dots \subset \mathbb{R}^n$ mit $\text{meas } M_k = 0$ für alle k gegeben und $M = \bigcup_k M_k$. Zu beliebigem $\varepsilon > 0$ und jedem k existieren dann abzählbar viele Quader Q_{k1}, Q_{k2}, \dots mit

$$M_k \subset \bigcup_j \mathring{Q}_{kj} \quad \text{und} \quad \sum_j |Q_{kj}| < 2^{-k} \varepsilon.$$

Die Menge der auftretenden Indizes (j, k) ist Teilmenge von $\mathbb{N} \times \mathbb{N}$ und somit abzählbar. Es gilt nun

$$M = \bigcup_k M_k \subset \bigcup_{(j,k)} \mathring{Q}_{kj}$$

sowie

$$\sum_{(j,k)} |\mathcal{Q}_{kj}| < \varepsilon \sum_k 2^{-k} \leq \varepsilon \sum_{k=1}^{\infty} \left(\frac{1}{2}\right)^k = \varepsilon,$$

wie behauptet. \square

Definition 12.3.9

Eine kompakte Menge $K \subset \mathbb{R}^n$, $n = m + d$, heißt *dünn*, wenn zu jedem $x^0 \in K$ ein Radius $r = r(x^0) > 0$ und eine Umordnung $x = (u, v)$ mit $u \in \mathbb{R}^m$ und $v \in \mathbb{R}^d$ so existieren, dass gilt: Es gibt eine kompakte Menge $K_0 \subset \mathbb{R}^m$ und eine stetige Funktion $\varphi : K_0 \rightarrow \mathbb{R}^d$ mit

$$K \cap \overline{B_r(x^0)} = \{(u, v) \in \mathbb{R}^{m+d} : v = \varphi(u), u \in K_0\}.$$

Beispiel 12.3.10. Gemäß Folgerung 11.6.4 ist insbesondere jede m -dimensionale *kompatte* gleichungsdefinierte Mannigfaltigkeit $M \subset \mathbb{R}^{m+d}$ eine dünne Menge; dazu wählt man im dortigen Beweis zur offenen Umgebung $X = X(x^0)$ einen Radius $r = r(x^0) > 0$ mit $\overline{B_r(x^0)} \subset X$.

Satz 12.3.11

Jede dünne Menge $K \subset \mathbb{R}^n$ hat den Inhalt Null.

Beweis. Mit dem Satz von HEINE-BOREL wählen wir aus der offenen Überdeckung $\{B_r(x) : x \in K\}$ mit den in Definition 12.3.9 angegebenen Radien $r = r(x) > 0$ endlich viele Kugeln $B_{r_1}(x^1), \dots, B_{r_p}(x^p)$, die K überdecken. Nach Beispiel 12.3.7 haben $K \cap \overline{B_{r_l}(x^l)}$ den Inhalt Null für alle $l = 1, \dots, p$. Und nach Hilfssatz 12.3.8 (a), (b) gilt dies auch für

$$K \subset \bigcup_{l=1}^p \{K \cap \overline{B_{r_l}(x^l)}\},$$

wie behauptet. \square

Folgerung 12.3.12

Jede kompakte Mannigfaltigkeit ist eine dünne Menge und hat somit den Inhalt Null.

Beweis. Sofort aus Beispiel 12.3.10 und Satz 12.3.11. □

12.4. Unstetigkeitsstellen integrierbarer Funktionen auf Quadern

Nach Satz 12.2.2 (und Bemerkung 12.2.8) ist jede stetige Funktion $f : Q \rightarrow \mathbb{R}^d$ auf einem Quader $Q \subset \mathbb{R}^n$ integrierbar. Aber wie „groß“ darf die Menge der Unstetigkeitsstellen einer nichtstetigen Funktion f werden, damit sie integrierbar bleibt? Eine Antwort auf diese Frage wird in Abschnitt 12.5 auch wichtig für die Integration über allgemeinere Teilmengen des \mathbb{R}^n .

Wir beginnen mit der

Definition 12.4.1

Seien $M \subset \mathbb{R}^n$ beliebig und $f : M \rightarrow \mathbb{R}$ beschränkt.

- Die *Oszillation* von f auf $N \subset M$ ist erklärt als

$$\operatorname{osc}_N f := \sup_N f - \inf_N f \stackrel{(9.2.2)}{=} \sup_{x, x' \in N} |f(x) - f(x')|.$$

- Ein Punkt $x \in M$ heißt *Unstetigkeitsstelle* von f , wenn gilt

$$\sigma_f(x) := \lim_{r \rightarrow 0+} \left(\operatorname{osc}_{M \cap B_r(x)} f \right) > 0. \quad (12.4.1)$$

Die Menge aller Unstetigkeitsstellen von f in $N \subset M$ bezeichnen wir mit

$$\mathcal{S}_N(f) := \{x \in N : \sigma_f(x) > 0\}, \quad \mathcal{S}(f) := \mathcal{S}_M(f).$$

Bemerkung 12.4.2. (a) $\sigma_f : M \rightarrow \mathbb{R}$ ist wohl definiert, da $\operatorname{osc}_{M \cap B_r(x)} f$ nichtnegativ und monoton wachsend in r ist.

(b) f ist in $x \in M$ genau dann stetig, wenn $\sigma_f(x) = 0$ gilt (\rightarrow Übungsaufgabe).

Ziel dieses Abschnitts ist der Beweis von

Satz 12.4.3: Unstetigkeitsstellen integrierbarer Funktionen

Eine beschränkte Funktion $f : Q \rightarrow \mathbb{R}$ ist genau dann integrierbar, wenn die Menge $\mathcal{S}(f)$ ihrer Unstetigkeitsstellen eine Nullmenge ist.

Der Beweis von Satz 12.4.3 beruht auf dem folgenden Hilfssatz:

Hilfssatz 12.4.4

Zu beschränktem $f : Q \rightarrow \mathbb{R}$ erklären wir $\sigma_f : Q \rightarrow \mathbb{R}$ wie in (12.4.1) und setzen

$$Q(\varepsilon) := \{x \in Q : \sigma_f(x) \geq \varepsilon\}, \quad \varepsilon > 0.$$

Dann ist $\mathcal{S}(f)$ genau dann Nullmenge, wenn $Q(\varepsilon)$ für alle $\varepsilon > 0$ den Inhalt Null hat.

Beweis. ▶ „ \Rightarrow “: Ist $\mathcal{S}(f)$ Nullmenge, so ist für jedes $\varepsilon > 0$ auch $Q(\varepsilon) \subset \mathcal{S}(f)$ Nullmenge.

Da $Q(\varepsilon)$ kompakt ist (\rightarrow Übungsaufgabe), hat $Q(\varepsilon)$ nach Folgerung 12.3.6 für alle $\varepsilon > 0$ den Inhalt Null.

▶ „ \Leftarrow “: Gilt andererseits $|Q(\varepsilon)| = 0$ für alle $\varepsilon > 0$, so insbesondere auch $|Q(\frac{1}{k})| = 0$ für $k \in \mathbb{N}$. Wegen

$$\mathcal{S}(f) = \bigcup_{k=1}^{\infty} Q\left(\frac{1}{k}\right) \quad (12.4.2)$$

ist $\mathcal{S}(f)$ gemäß Hilfssatz 12.3.8 (c) Nullmenge. \square

Beweis von Satz 12.4.3. ▶ „ \Rightarrow “: Ergänzungen.

▶ „ \Leftarrow “: Sei $\mathcal{S}(f)$ Nullmenge und $\varepsilon > 0$ sei beliebig fixiert. Nach Hilfssatz 12.4.4 ist dann $|Q(\varepsilon)| = 0$ für beliebiges $\varepsilon > 0$. Also existiert eine endliche Überdeckung $\bigcup_{j=1}^p \mathring{Q}_j$ von $Q(\varepsilon)$ durch $p \in \mathbb{N}$ Quader mit Inhaltssumme $\sum_{j=1}^p |Q_j| < \varepsilon$.

Wir betrachten die kompakte Menge $\hat{Q} := Q \setminus (\mathring{Q}_1 \cup \dots \cup \mathring{Q}_p)$. Nach Konstruktion gilt $\sigma_f(x) < \varepsilon$ für jedes $x \in \hat{Q}$. Also gibt es zu jedem $x \in \hat{Q}$ einen (abgeschlossenen) Würfel W_x mit Mittelpunkt x , so dass $\text{osc}_{Q \cap W_x} f < \varepsilon$ erfüllt ist. Nun liefert $\bigcup_{x \in \hat{Q}} \mathring{W}_x$ eine offene Überdeckung von \hat{Q} , aus der wir nach dem Satz von HEINE-BOREL endlich viele Würfel W_{x_1}, \dots, W_{x_r} mit $\mathring{W}_{x_1} \cup \dots \cup \mathring{W}_{x_r} \supset \hat{Q}$ auswählen können.

Insgesamt ist also $\{Q_1, \dots, Q_p, W_{x_1}, \dots, W_{x_r}\}$ eine Überdeckung von Q durch Quader. Wir ordnen nun eine Zerlegung \mathcal{Z}^* von Q in Teilquader Q_α^* , $\alpha \in A$, so zu, dass die Indexmenge gemäß $A = A' \cup A''$ in zwei disjunkte Teilmengen zerfällt, für die gilt:

$$\triangleright \bigcup_{\alpha \in A'} Q_\alpha^* \subset \bigcup_{j=1}^p Q_j \text{ und folglich } \sum_{\alpha \in A'} |Q_\alpha^*| \leq \sum_{j=1}^p |Q_j| < \varepsilon.$$

▷ Für jedes $\alpha \in A''$ gilt $Q_\alpha^* \subset Q \cap W_{x_k}$ mit einem $k = k(\alpha) \in \{1, \dots, r\}$ und folglich $\operatorname{osc}_{Q_\alpha^*} f < \varepsilon$.

Für die zu Z^* gehörigen Ober- und Untersummen erhalten wir dann:

$$\begin{aligned} \bar{S}_{Z^*}(f) - \underline{S}_{Z^*}(f) &= \sum_{\alpha \in A'} (\operatorname{osc}_{Q_\alpha^*} f) |Q_\alpha^*| + \sum_{\alpha \in A''} (\operatorname{osc}_{Q_\alpha^*} f) |Q_\alpha^*| \\ &< 2 \left(\sup_Q |f| \right) \sum_{\alpha \in A'} |Q_\alpha^*| + \varepsilon \sum_{\alpha \in A''} |Q_\alpha^*| \\ &\leq \varepsilon (2 \sup_Q |f| + |Q|). \end{aligned}$$

Da $\varepsilon > 0$ beliebig gewählt war, ist also f nach Satz 12.1.11 integrierbar. \square

Bemerkung 12.4.5. Wir können das Ergebnis von Satz 12.4.3 direkt auf vektor- bzw. komplexwertige Funktionen verallgemeinern. Ist etwa $f = (f_1, \dots, f_d) : Q \rightarrow \mathbb{R}^d$ beschränkt, so ist f genau dann unstetig, wenn mindestens eine der Komponenten f_k in x unstetig ist, d.h. $\mathcal{S}(f) = \bigcup_{k=1}^d \mathcal{S}(f_k)$. Nun gilt

$$\begin{aligned} f \in \mathcal{R}(Q, \mathbb{R}^d) &\Leftrightarrow f_1, \dots, f_d \in \mathcal{R}(Q) \\ &\stackrel{\text{Satz 12.4.3}}{\Leftrightarrow} \operatorname{meas} \mathcal{S}(f_1) = \dots = \operatorname{meas} \mathcal{S}(f_d) = 0 \\ &\stackrel{\text{Hilfss. 12.3.8(c)}}{\Leftrightarrow} \operatorname{meas} \mathcal{S}(f) = 0. \end{aligned}$$

12.5. Integration über quadrierbare Mengen

Wir wollen nun allgemeinere Mengen $M \subset \mathbb{R}^n$ als Integrationsbereiche wählen. Diese erklären wir mit dem Riemann-Integral charakteristischer Funktionen (siehe Beispiel 1.5.3 (d)):

Definition 12.5.1

Eine beschränkte Menge $M \subset \mathbb{R}^n$ heißt *quadrierbar* (oder *JORDAN-messbar*, *JORDANmenge*), wenn ihre charakteristische Funktion $\chi_M : \mathbb{R}^n \rightarrow \mathbb{R}$,

$$\chi_M(x) := \begin{cases} 1, & \text{für } x \in M, \\ 0, & \text{für } x \in \mathbb{R}^n \setminus M \end{cases}, \quad x \in \mathbb{R}^n,$$

auf einem Quader $Q \subset \mathbb{R}^n$ mit $\overline{M} \subset \overset{\circ}{Q}$ integrierbar ist. Den Wert

$$|M| := v(M) := \int_Q \chi_M(x) dx \tag{12.5.1}$$

nennen wir den (*n-dimensionalen*) *Inhalt* (oder *Volumen* oder *JORDAN-Maß*) von M .

Bemerkung 12.5.2. (a) Obige Definition ist von der Wahl des Quaders $Q \subset \mathbb{R}^n$ mit $\overline{M} \subset \overset{\circ}{Q}$ unabhängig.

(b) Jeder Quader $Q = [a_1, b_1] \times \dots \times [a_n, b_n]$ ist quadrierbar mit

$$v(Q) = \prod_{j=1}^n (b_j - a_j) = |Q|.$$

Die ursprüngliche Definition des Inhalts eines Quaders stimmt also mit der in (12.5.1) überein.

(c) Jede Menge M vom Inhalt Null ist quadrierbar mit $v(M) = 0$ (\rightarrow *Ergänzungen*); auch hier stimmt also $v(M)$ mit der ursprünglichen Definition von $|M|$ überein.

Satz 12.5.3: Quadrierbarkeitskriterium

Eine beschränkte Menge $M \subset \mathbb{R}^n$ ist genau dann quadrierbar, wenn ihr Rand ∂M Nullmenge ist ($\Leftrightarrow |\partial M| = 0$).

Beweis. Da ∂M abgeschlossen ist, ist $\mathbb{R}^n \setminus \partial M$ offen. Folglich ist χ_M in einer Umgebung jedes Punktes $x \in \mathbb{R}^n \setminus \partial M$ konstant, es folgt $\mathcal{S}(\chi_M) \subset \partial M$. Andererseits ist auch $\partial M \subset \mathcal{S}(\chi_M)$ richtig, denn offenbar gilt $\sigma_{\chi_M}(x) = 1$ für alle $x \in \partial M$; siehe Definition 12.4.1.

Wählen wir $Q \subset \mathbb{R}^n$ mit $\overline{M} \subset \overset{\circ}{Q}$, so ist χ_M nach Satz 12.4.3 genau dann auf Q integrierbar, wenn $\partial M = \mathcal{S}(\chi_M)$ eine Nullmenge ist. Da ∂M kompakt ist, ist dies nach Folgerung 12.3.6 äquivalent zu $|\partial M| = 0$. \square

Folgerung 12.5.4

Ist $M \subset \mathbb{R}^n$ beschränkt und ∂M eine dünne Menge im Sinne von Definition 12.3.9, so ist M quadrierbar.

Beweis. Satz 12.3.11 liefert $|\partial M| = 0$, nach Satz 12.5.3 ist M also quadrierbar. \square

Folgerung 12.5.5

Sind $M, N \subset \mathbb{R}^n$ quadrierbar, so sind auch $M \cup N$, $M \cap N$ und $M \setminus N$ quadrierbar.

Beweis. Offenbar sind $M \cup N$, $M \cap N$ und $M \setminus N$ beschränkt, da M, N beschränkt sind. Wegen

$$\partial(M \cup N), \partial(M \cap N), \partial(M \setminus N) \subset \partial M \cup \partial N$$

folgt die Behauptung sofort aus Satz 12.5.3 und Hilfssatz 12.3.8. \square

Beispiel 12.5.6. Wegen der Folgerungen 12.3.12, 12.5.4 und 12.5.5 sind Durchschnitte, Vereinigungen und Differenzen von endlich vielen beschränkten Mengen quadrierbar, deren Ränder Mannigfaltigkeiten sind.

Zum Beispiel ist also der Kreisring $R(a, b) := \{x \in \mathbb{R}^n : a < |x| < b\} = B_b(0) \setminus \overline{B_a(0)}$ mit $0 < a < b < +\infty$ eine quadrierbare Menge, da $\partial B_a(0)$, $\partial B_b(0)$ Mannigfaltigkeiten sind.

Definition 12.5.7: Integrierbarkeit auf quadrierbaren Mengen

Seien $M \subset \mathbb{R}^n$ quadrierbar und $f : M \rightarrow \mathbb{R}^d$ beschränkt.

- Die *kanonische Fortsetzung* $\bar{f}_M : \mathbb{R}^n \rightarrow \mathbb{R}^d$ von f ist erklärt als

$$\bar{f}_M(x) := \begin{cases} f(x), & \text{für } x \in M \\ 0, & \text{für } x \in \mathbb{R}^n \setminus M \end{cases}.$$

- f heißt *RIEMANN-integrierbar* oder *integrierbar* auf M , wenn die kanonische Fortsetzung \bar{f}_M auf einem Quader $Q \subset \mathbb{R}^n$ mit $\bar{M} \subset \overset{\circ}{Q}$ integrierbar ist. Wir schreiben dann $f \in \mathcal{R}(M, \mathbb{R}^d)$ und erklären das (*RIEMANNsche*) *Integral* von f auf M gemäß

$$\int_M f \, dV = \int_M f \, d(x_1, \dots, x_n) = \int_M f \, dx := \int_Q \bar{f}_M \, dx.$$

- Für $d = 1$ und $d = 2$ schreiben wir noch $\mathcal{R}(M) := \mathcal{R}(M, \mathbb{R})$ bzw. $\mathcal{R}(M, \mathbb{C}) := \mathcal{R}(M, \mathbb{R}^2)$ und bemerken

$$\int_M f \, dx = \int_M \operatorname{Re} f \, dx + i \int_M \operatorname{Im} f \, dx, \quad f \in \mathcal{R}(M, \mathbb{C}).$$

Bemerkung 12.5.8. (a) Die Definition ist unabhängig von der Wahl des Quaders $Q \subset \mathbb{R}^n$ mit $\bar{M} \subset \overset{\circ}{Q}$.

- (b) Für die konstante Funktion $f = 1 : M \rightarrow \mathbb{R}$, gilt $\bar{f}_M = \chi_M$. Nach Definition ist also $1 \in \mathcal{R}(M)$ und es gilt

$$\int_M 1 \, dx = \int_Q \chi_M \, dx = |M|.$$

Satz 12.5.9: LEBESGUESCHES Integrabilitätskriterium

Sei $M \subset \mathbb{R}^n$ quadrierbar und $f : M \rightarrow \mathbb{R}^d$ beschränkt. Dann gilt

$$f \in \mathcal{R}(M, \mathbb{R}^d) \Leftrightarrow \text{meas } \mathcal{S}_M^\circ(f) = 0.$$

Beweis. Wir betrachten die kanonische Fortsetzung $\bar{f}_M : \mathbb{R}^n \rightarrow \mathbb{R}^d$, für deren Unstetigkeitsstellen offenbar gilt

$$\mathcal{S}_M^\circ(f) \subset \mathcal{S}(\bar{f}_M) \subset \mathcal{S}_M^\circ(f) \cup \partial M.$$

Ist $Q \subset \mathbb{R}^n$ ein Quader mit $\bar{M} \subset \overset{\circ}{Q}$, so folgt

$$\begin{aligned} f \in \mathcal{R}(M, \mathbb{R}^d) &\stackrel{\text{Def. 12.5.7}}{\Leftrightarrow} \bar{f}_M \in \mathcal{R}(Q, \mathbb{R}^d) \\ &\stackrel{\text{Satz 12.4.3}}{\Leftrightarrow} \mathcal{S}(\bar{f}_M) \text{ ist Nullmenge} \\ &\stackrel{\text{HS 12.3.8, Satz 12.5.3}}{\Leftrightarrow} \mathcal{S}_M^\circ(f) \text{ ist Nullmenge,} \end{aligned}$$

wie behauptet. □

Folgerung 12.5.10

Jede beschränkte Funktion $f : M \rightarrow \mathbb{R}^d \in C^0(M, \mathbb{R}^d)$ auf der quadrierbaren Menge $M \subset \mathbb{R}^n$ ist integrierbar.

Beweis. Sofort klar aus Satz 12.5.9 wegen $\mathcal{S}_M^\circ(f) = \emptyset$. □

Durch Definition 12.5.7 lassen sich die Rechenregeln aus Satz 12.2.1 übertragen; siehe auch Bemerkung 12.2.8:

Satz 12.5.11: Rechenregeln

Sei $M \subset \mathbb{R}^n$ quadrierbar.

(a) Für $f, g \in \mathcal{R}(M, \mathbb{R}^d)$ und $\alpha, \beta \in \mathbb{R}$ gilt $\alpha f + \beta g \in \mathcal{R}(M, \mathbb{R}^d)$ und

$$\int_M (\alpha f + \beta g) dx = \alpha \int_M f dx + \beta \int_M g dx.$$

(b) Für $d = 1$: Sind $f, g \in \mathcal{R}(M, \mathbb{R}) =: \mathcal{R}(M)$ mit $f \leq g$ auf M gegeben, so folgt

$$\int_M f \, dx \leq \int_M g \, dx.$$

(c) Für jedes $f \in \mathcal{R}(M, \mathbb{R}^d)$ gilt auch $\|f\| \in \mathcal{R}(M)$ und

$$\left\| \int_M f \, dx \right\| \leq \int_M \|f\| \, dx.$$

(d) Gilt $f, g \in \mathcal{R}(M, \mathbb{R}^d)$, so folgt $\langle f, g \rangle \in \mathcal{R}(M)$ und

$$\int_M |\langle f, g \rangle| \, dx \leq \left(\sup_M \|g\| \right) \int_M \|f\| \, dx$$

sowie die *SCHWARZSCHE Ungleichung*

$$\int_M |\langle f, g \rangle| \, dx \leq \left(\int_M \|f\|^2 \, dx \right)^{\frac{1}{2}} \left(\int_M \|g\|^2 \, dx \right)^{\frac{1}{2}}.$$

(e) Für $d = 2$: Gilt $f, g \in \mathcal{R}(M, \mathbb{C})$ und $|g| \geq c > 0$ auf M mit einer Konstante $c > 0$, so folgt $\frac{f}{g} \in \mathcal{R}(M, \mathbb{C})$ und

$$\int_M \left| \frac{f}{g} \right| \, dx \leq \frac{1}{c} \int_M |f| \, dx.$$

Beweis. → Ergänzungen. □

Zum Abschluss geben wir noch die Verallgemeinerung von Satz 9.5.7 auf Funktionen auf quadrierbaren Mengen an:

Satz 12.5.12

Sei $M \subset \mathbb{R}^n$ quadrierbar und die Folge $\{f_k\}_k \subset \mathcal{R}(M, \mathbb{R}^d)$ konvergiere gleichmäßig gegen eine Funktion $f : M \rightarrow \mathbb{R}^d$. Dann folgen $f \in \mathcal{R}(M, \mathbb{R}^d)$ und

$$\int_M \lim_{k \rightarrow \infty} f_k(x) \, dx = \int_M f(x) \, dx = \lim_{k \rightarrow \infty} \int_M f_k(x) \, dx.$$

Beweis. → Übungsaufgabe. □

12.6. Normalbereiche und Additivität des Integrals

Eine interessante Klasse von quadrierbaren Mengen enthält die

Definition 12.6.1

Eine Menge $M \subset \mathbb{R}^n$ heißt *Normalbereich (bez. der x_j -Achse)*, wenn es eine quadrierbare, kompakte Menge $K \subset \mathbb{R}^{n-1}$ und Funktionen $\psi, \chi : K \rightarrow \mathbb{R} \in C^0(K)$ mit $\psi \leq \chi$ auf K so gibt, dass M die folgende Form hat:

$$M = \{x \in \mathbb{R}^n : y = (x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n) \in K, \psi(y) \leq x_j \leq \chi(y)\}. \quad (12.6.1)$$

Bemerkung 12.6.2. Jeder Normalbereich $M \subset \mathbb{R}^n$ ist quadrierbar.

Denn: $M \subset \mathbb{R}^n$ ist offenbar kompakt und es gilt

$$\partial M = \text{graph } \psi \cup \text{graph } \chi \cup \{x \in \mathbb{R}^n : y \in \partial K, \psi(y) \leq x_j \leq \chi(y)\}.$$

Nach Beispiel 12.3.7 haben $\text{graph } \psi$ und $\text{graph } \chi$ den Inhalt Null. Und nach Satz 12.5.3 hat $\partial K \subset \mathbb{R}^{n-1}$ und folglich auch $\{x \in \mathbb{R}^n : y \in \partial K, \psi(y) \leq x_j \leq \chi(y)\}$ den Inhalt Null (\rightarrow Nachrechnen). Hilfssatz 12.3.8 (b) liefert also die Behauptung. \square

Integrale über Normalbereiche lassen sich auf Integrale über niederdimensionale Mengen zurückführen (vgl. auch Satz 12.2.5):

Satz 12.6.3: Fubini

Es sei $M \subset \mathbb{R}^n$ Normalbereich der Form (12.6.1) und $f \in C^0(M, \mathbb{R}^d)$ sei gegeben. Dann gilt

$$\int_M f \, dx = \int_K \left(\int_{\psi(y)}^{\chi(y)} f \, dx_j \right) dy.$$

Beweis. Anwendung von Satz 12.2.3 auf die kanonische Fortsetzung \bar{f}_M (\rightarrow Ergänzungen). \square

Beispiel 12.6.4. Gesucht ist $|B_R(0)|$ für die Kugel $B_R(0) \subset \mathbb{R}^3$ mit Radius $R > 0$. Wir können schreiben

$$\overline{B_R(0)} = \{(x, y, z) \in \mathbb{R}^3 : (x, y) \in K_R(0), -\sqrt{R^2 - x^2 - y^2} \leq z \leq \sqrt{R^2 - x^2 - y^2}\}$$

mit der abgeschlossenen Kreisscheibe

$$\begin{aligned} K_R(0) &:= \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq R^2\} \\ &= \{(x, y) \in \mathbb{R}^2 : -R \leq x \leq R, -\sqrt{R^2 - x^2} \leq y \leq \sqrt{R^2 - x^2}\}. \end{aligned}$$

Satz 12.6.3 liefert also

$$\begin{aligned} |\overline{B_R(0)}| &= \int_{\overline{B_R(0)}} 1 \, d(x, y, z) = \int_{K_R(0)} \left(\int_{-\sqrt{R^2-x^2-y^2}}^{\sqrt{R^2-x^2-y^2}} 1 \, dz \right) d(x, y) \\ &= 2 \int_{K_R(0)} \sqrt{R^2-x^2-y^2} \, d(x, y) = 2 \int_R^R \left(\int_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} \sqrt{R^2-x^2-y^2} \, dy \right) dx. \end{aligned}$$

Für beliebiges $\varrho > 0$ gilt nun (substituiere $y = \varrho \cos t$, $t \in [0, \pi]$):

$$\int_{-\varrho}^{\varrho} \sqrt{\varrho^2 - y^2} \, dy = \left(-\frac{\varrho^2}{2} \arccos \frac{y}{\varrho} + \frac{y}{2} \sqrt{\varrho^2 - y^2} \right) \Big|_{-\varrho}^{\varrho} = \frac{\pi \varrho^2}{2}.$$

Einsetzen mit $\varrho = \sqrt{R^2 - x^2}$ liefert somit

$$|\overline{B_R(0)}| = \pi \int_{-R}^R (R^2 - x^2) \, dx = \pi \left[R^2 x - \frac{x^3}{3} \right]_{-R}^R = \frac{4\pi R^3}{3}.$$

Bemerkung 12.5.2 (c) und Hilfssatz 12.6.6 unten entnehmen wir schließlich

$$|B_R(0)| = |\overline{B_R(0)}| - |\partial B_R(0)| = |\overline{B_R(0)}| = \frac{4\pi R^3}{3}.$$

Wir wollen noch Satz 9.3.2 auf die mehrdimensionale Integration übertragen. Dazu benötigen wir einige Hilfsaussagen, wir beginnen mit der

Definition 12.6.5

Sei $f : M \rightarrow \mathbb{R}^d$ auf einer Menge $M \subset \mathbb{R}^n$ erklärt und sei $M' \subset M$ quadrierbar. Dann heißt f *auf M' integrierbar*, i.Z. $f \in \mathcal{R}(M', \mathbb{R}^d)$, wenn $f|_{M'} \in \mathcal{R}(M', \mathbb{R}^d)$ gilt, und wir setzen

$$\int_{M'} f \, dx := \int_{M'} f|_{M'} \, dx.$$

Hilfssatz 12.6.6

Seien $M_1, M_2 \subset \mathbb{R}^n$ quadrierbar mit $M_1 \cap M_2 = \emptyset$, und eine Funktion $f : M_1 \cup M_2 \rightarrow \mathbb{R}^d$

sei gegeben. Gilt dann $f \in \mathcal{R}(M_j, \mathbb{R}^d)$ für $j = 1, 2$, so folgt $f \in \mathcal{R}(M_1 \cup M_2, \mathbb{R}^d)$ und

$$\int_{M_1 \cup M_2} f \, dx = \int_{M_1} f \, dx + \int_{M_2} f \, dx.$$

Beweis. Nach Folgerung 12.5.5 ist $M_1 \cup M_2$ quadrierbar. Für die kanonische Fortsetzung $\bar{f}_{M_1 \cup M_2}$ von f haben wir

$$\bar{f}_{M_1 \cup M_2} = \chi_{M_1 \cup M_2} \bar{f}_{M_1 \cup M_2} = \chi_{M_1} \bar{f}_{M_1 \cup M_2} + \chi_{M_2} \bar{f}_{M_1 \cup M_2} =: f_1 + f_2. \quad (12.6.2)$$

Sei $Q \subset \mathbb{R}^n$ ein beliebiger Quader mit $\overline{M_1 \cup M_2} \subset \overset{\circ}{Q}$. Wegen

$$f_j := \chi_{M_j} \bar{f}_{M_1 \cup M_2} = \overline{(f|_{M_j})}_{M_j}, \quad j = 1, 2, \quad (12.6.3)$$

gilt nach Voraussetzung $f_j \in \mathcal{R}(Q, \mathbb{R}^d)$, $j = 1, 2$, Satz 12.2.1 (a) liefert also $f \in \mathcal{R}(Q, \mathbb{R}^d)$ bzw. $f \in \mathcal{R}(M_1 \cup M_2, \mathbb{R}^d)$ sowie

$$\begin{aligned} \int_{M_1 \cup M_2} f \, dx &= \int_Q \bar{f}_{M_1 \cup M_2} \, dx \stackrel{(12.6.2)}{=} \int_Q f_1 \, dx + \int_Q f_2 \, dx \\ &\stackrel{(12.6.3)}{=} \int_{M_1} f|_{M_1} \, dx + \int_{M_2} f|_{M_2} \, dx = \int_{M_1} f \, dx + \int_{M_2} f \, dx, \end{aligned}$$

wie behauptet. \square

Bemerkung 12.6.7. Wendet man Hilfssatz 12.6.6 auf charakteristische Funktionen an, so erhält man verschiedene Inhalts-Beziehungen zwischen quadrierbaren Mengen. Sind z.B. $M, N \subset \mathbb{R}^n$ quadrierbar mit $M \subset N$, so folgt $|M| \leq |N|$ (\rightarrow Übungsaufgabe).

Hilfssatz 12.6.8

Ist $N \subset \mathbb{R}^n$ quadrierbare Nullmenge, so ist jedes beschränkte $f : N \rightarrow \mathbb{R}^d$ integrierbar auf N und es gilt

$$\int_N f \, dx = 0.$$

Beweis. \rightarrow Ergänzungen. \square

Bemerkung 12.6.9. Insbesondere gilt dies für jede Menge vom Inhalt Null gemäß Bemerkung 12.5.2 (c). Nach Satz 12.5.3 und Hilfssatz 12.3.8 (a) gilt die Aussage von Hilfssatz 12.6.8 somit für jede Teilmenge $N \subset \partial M$ beliebiger quadrierbarer Mengen $M \subset \mathbb{R}^n$.

Hilfssatz 12.6.10

Sei $M \subset \mathbb{R}^n$ quadrierbar und sei $f : M \rightarrow \mathbb{R}^d$ beschränkt.

- (a) Falls $f \in \mathcal{R}(M, \mathbb{R}^d)$ gilt, so folgt $f \in \mathcal{R}(M', \mathbb{R}^d)$ für jede quadrierbare Teilmenge $M' \subset M$.
- (b) Ist $N \subset M$ quadrierbare Nullmenge und gilt $f \in \mathcal{R}(M \setminus N, \mathbb{R}^d)$, so folgt $f \in \mathcal{R}(M, \mathbb{R}^d)$ und

$$\int_M f \, dx = \int_{M \setminus N} f \, dx. \quad (12.6.4)$$

Beweis. (a) Sei $Q \subset \mathbb{R}^n$ mit $\overline{M} \subset \overset{\circ}{Q}$ gewählt. Nach Satz 12.2.1 (d) gilt dann

$$(\overline{f|_{M'}})_{M'} = \chi_{M'} \overline{f}_M \in \mathcal{R}(Q, \mathbb{R}^d)$$

und somit $f \in \mathcal{R}(M', \mathbb{R}^d)$.

- (b) Nach Hilfssatz 12.6.8 gilt $f \in \mathcal{R}(N, \mathbb{R}^d)$ mit $\int_N f \, dx = 0$. Wegen $M = (M \setminus N) \cup N$ liefert Hilfssatz 12.6.6 somit $f \in \mathcal{R}(M, \mathbb{R}^d)$ und

$$\int_M f \, dx = \int_{M \setminus N} f \, dx + \int_N f \, dx = \int_{M \setminus N} f \, dx,$$

wie behauptet. □

Bemerkung 12.6.11. Eine beliebige Menge $M \subset \mathbb{R}^n$ ist genau dann quadrierbar, wenn $\overset{\circ}{M}$ quadrierbar ist, denn es gilt $M = \overset{\circ}{M} \cup N$ mit $N = M \setminus \overset{\circ}{M} \subset \partial M$. Hilfssatz 11.6.10 liefert daher: Ist $M \subset \mathbb{R}^n$ quadrierbar und $f : M \rightarrow \mathbb{R}^d$ beschränkt, so gilt $f \in \mathcal{R}(M, \mathbb{R}^d) \Leftrightarrow f \in \mathcal{R}(\overset{\circ}{M}, \mathbb{R}^d)$, und wir haben

$$\int_M f \, dx = \int_{\overset{\circ}{M}} f \, dx, \quad f \in \mathcal{R}(M, \mathbb{R}^d).$$

Wir kommen nun zum wichtigen

Satz 12.6.12: Additivität des Integrals

Sei $M = \bigcup_{l=1}^p M_l$ mit quadrierbaren Mengen $M_1, \dots, M_p \subset \mathbb{R}^n$, die $\overset{\circ}{M}_l \cap \overset{\circ}{M}_k = \emptyset$ für $l \neq k$ erfüllen, und sei $f : M \rightarrow \mathbb{R}^d$ beschränkt. Dann gilt

$$f \in \mathcal{R}(M, \mathbb{R}^d) \Leftrightarrow f \in \mathcal{R}(M_l, \mathbb{R}^d) \quad \text{für alle } l = 1, \dots, p, \quad (12.6.5)$$

und wir haben

$$\int_M f \, dx = \sum_{l=1}^p \int_{M_l} f \, dx = \sum_{l=1}^p \int_{\dot{M}_l} f \, dx.$$

Beweis. Sei $f : M \rightarrow \mathbb{R}^d$ beschränkt. Gemäß Folgerung 12.5.5 ist M quadrierbar. Nach Bemerkung 12.6.11 sind auch $\dot{M}_1, \dots, \dot{M}_p$ quadrierbar, und es gilt $f \in \mathcal{R}(M_l, \mathbb{R}^d) \Leftrightarrow f \in \mathcal{R}(\dot{M}_l, \mathbb{R}^d)$ sowie

$$\int_{M_l} f \, dx = \int_{\dot{M}_l} f \, dx, \quad l = 1, \dots, p.$$

Weiter ist dann $\dot{M}_1 \cup \dots \cup \dot{M}_p$ quadrierbar, und wir können schreiben

$$M = \dot{M}_1 \cup \dots \cup \dot{M}_p \cup N, \quad N \subset \bigcup_{l=1}^p \partial M_l.$$

Nach Satz 12.5.3 und Hilfssatz 12.3.8 ist N Menge vom Inhalt Null und damit quadrierbare Nullmenge. Den Hilfssätzen 12.6.6 und 12.6.10 entnehmen wir daher (12.6.5) sowie

$$\int_M f \, dx = \int_{M \setminus N} f \, dx = \int_{\dot{M}_1 \cup \dots \cup \dot{M}_p} f \, dx = \sum_{l=1}^p \int_{\dot{M}_l} f \, dx = \sum_{l=1}^p \int_{M_l} f \, dx, \quad f \in \mathcal{R}(M, \mathbb{R}^d),$$

wie behauptet. \square

Bemerkung 12.6.13. Ist $M \subset \mathbb{R}^n$ darstellbar als Vereinigung $M = M_1 \cup \dots \cup M_p$ endlich vieler Normalbereiche mit $\dot{M}_l \cap \dot{M}_k = \emptyset$ für $l \neq k$, d.h. kann man M in $p \in \mathbb{N}$ Normalbereiche „zerschneiden“, so kann man z.B. für jedes $f \in C^0(M)$ zunächst $\int_{M_l} f \, dx = \int_{\dot{M}_l} f \, dx, l = 1, \dots, p$, mit Satz 12.6.3 und anschließend $\int_M f \, dx$ mit Satz 12.6.12 bestimmen.

12.7. Die Transformationsformel für Testfunktionen - Kurzfassung

In den abschließenden beiden Abschnitten beschreiben wir den Beweisweg zur Transformationsformel, Satz 12.8.1 unten. Der vollständige Beweis wird als Anhang angefügt.

Zunächst skizzieren wir den Beweis der folgenden speziellen Version der Transformationsformel:

Satz 12.7.1: Transformationsformel für Testfunktionen

Seien $\Omega, \Theta \subset \mathbb{R}^n$ offene, quadrierbare Mengen und $\phi : \Omega \rightarrow \Theta$ ein Diffeomorphismus. Dann

gilt für beliebige Testfunktionen $f \in C_c^0(\Theta)$ die Identität

$$\int_{\Theta} f(y) dy = \int_{\Omega} f(\phi(x)) |J_{\phi}(x)| dx. \quad (12.7.1)$$

Zu erklären ist hier zunächst der Begriff der *Testfunktion*:

Definition 12.7.2

- Ist $M \subset \mathbb{R}^n$ beliebig und $f : M \rightarrow \mathbb{R}^d$ gegeben. Dann heißt die Menge

$$\text{supp } f := \overline{\{x \in M : f(x) \neq 0\}}$$

der *Träger* oder *Support* von f .

- Ist $M \subset \mathbb{R}^n$ beliebig, so heißt $M' \subset M$ *kompakt enthalten* in M , i.Z. $M' \subset\subset M$, wenn $\overline{M'}$ kompakt ist und $\overline{M'} \subset M$ erfüllt.
- Sind $\Omega \subset \mathbb{R}^n$ offen und $s \in \mathbb{N}_0 \cup \{\infty\}$, so bezeichnet

$$C_c^s(\Omega, \mathbb{R}^d) := \{f \in C^s(\Omega, \mathbb{R}^d) : \text{supp } f \subset\subset \Omega\}$$

die Menge der *s-mal stetig differenzierbaren, \mathbb{R}^d -wertigen Funktionen mit kompaktem Träger in Ω* . Ein solches $f \in C_c^s(\Omega, \mathbb{R}^d)$ nennen wir auch kurz *Testfunktion*. Wir setzen noch $C_c^s(\Omega) := C_c^s(\Omega, \mathbb{R})$.

Bemerkung 12.7.3. ► Jede Testfunktion $g \in C_c^0(\Omega)$ ist offenbar beschränkt, die Integrale in Satz 12.7.1 sind also gemäß Folgerung 12.5.10 wohl definiert.

- Wir können uns eine Funktion $f \in C_c^s(\Omega, \mathbb{R}^d)$ immer auf ganz \mathbb{R}^n erklärt denken, indem wir f zu 0 auf $\mathbb{R}^n \setminus \Omega$ fortsetzen. Dann ist $f \in C^s(\mathbb{R}^n, \mathbb{R}^d)$ und $\text{supp } f \subset\subset \Omega$ richtig.

Beweis von Satz 12.7.1. Dieser ist technisch aufwendig und besteht i.W. aus zwei Schritten:

1. Schritt: „Lokalisierung“. Es genügt, die Transformationsformel für Testfunktionen f zu beweisen, die außerhalb einer beliebig kleinen Kugel verschwinden (also = 0 sind). Dazu bedienen wir uns der folgenden

Definition 12.7.4: Zerlegung der Eins

Sei $M \subset \mathbb{R}^n$ nichtleer. Eine *Zerlegung der Eins* auf M ist eine Familie $\{\eta_{\alpha}\}_{\alpha \in I}$ von Funktionen $\eta_{\alpha} \in C_c^{\infty}(\mathbb{R}^n)$ mit beliebiger Indexmenge I und den folgenden Eigenschaften:

- (i) Für alle $\alpha \in I$ gilt $0 \leq \eta_\alpha \leq 1$ auf \mathbb{R}^n .
- (ii) Für jedes $x \in \mathbb{R}^n$ existieren höchstens endlich viele $\alpha \in I$ mit $\eta_\alpha(x) \neq 0$.
- (iii) Es gilt $\sum_{\alpha \in I} \eta_\alpha(x) = 1$ für alle $x \in M$ sowie $0 \leq \sum_{\alpha \in I} \eta_\alpha \leq 1$ auf \mathbb{R}^n .

Die Zerlegung der Eins heißt *endlich*, wenn I endlich ist.

Hilfssatz 12.7.5

Sei $K \subset \mathbb{R}^n$ kompakt und $\mathcal{F} = \{O_x : x \in K\}$ eine beliebige offene Überdeckung, wobei O_x eine offene Umgebung von $x \in K$ bezeichne. Dann gibt es eine endliche Zerlegung der Eins $\{\eta_\alpha\}_{\alpha=1,\dots,p}$ auf K mit der zusätzlichen Eigenschaft $\eta_\alpha \in C_c^\infty(O_{x_\alpha})$ für geeignete $x_\alpha \in K$, $\alpha = 1, \dots, p$.

Beweis. Dieser gelingt mit dem Satz von HEINE-BOREL und expliziter Konstruktion der η_α mit Hilfe der e -Funktion. ■

Hilfssatz 12.7.6: Lokalisierung

Es seien $\Omega, \Theta \subset \mathbb{R}^n$ und $\phi : \Omega \rightarrow \Theta$ wie in Satz 12.7.1. Dann gilt Formel (12.7.1) genau dann für alle $f \in C_c^0(\Theta)$, wenn zu jedem $y^0 \in \Theta$ ein $\varrho = \varrho(y^0) > 0$ mit $B_\varrho(y^0) \subset \Theta$ so existiert, dass Formel (12.7.1) für alle $f \in C_c^0(B_\varrho(y^0))$ erfüllt ist.

Beweis. ▶ „ \Rightarrow “: Klar, wegen $C_c^0(B_\varrho(y^0)) \subset C_c^0(\Theta)$ für $B_\varrho(y^0) \subset \Theta$.

- ▶ „ \Leftarrow “: Sei also $f \in C_c^0(\Theta)$ gewählt, d.h. $K := \text{supp } f \subset \Theta$ ist kompakt. Dann existiert nach Hilfssatz 12.7.5 eine endliche Zerlegung der Eins $\{\eta_\alpha\}_{\alpha=1,\dots,p}$ auf K mit $\eta_\alpha \in C_c^\infty(B_{\varrho(y_\alpha)}(y_\alpha))$ für $\alpha = 1, \dots, p$, wobei $\varrho(y_\alpha) > 0$ wie in der Voraussetzung gewählt seien. Es folgt nun mit $g_\alpha := f\eta_\alpha \in C_c^0(B_{\varrho(y_\alpha)}(y_\alpha))$ nach Voraussetzung:

$$\begin{aligned}
 0 &= \sum_{\alpha=1}^p \left(\int_{\Theta} g_\alpha(y) dy - \int_{\Omega} g_\alpha(\phi(x)) |J_\phi(x)| dx \right) \\
 &= \int_{\Theta} \left[\sum_{\alpha=1}^p \eta_\alpha(y) \right] f(y) dy - \int_{\Omega} \left[\sum_{\alpha=1}^p \eta_\alpha(\phi(x)) \right] f(\phi(x)) |J_\phi(x)| dx \\
 &= \int_{\Theta} f(y) dy - \int_{\Omega} f(\phi(x)) |J_\phi(x)| dx,
 \end{aligned}$$

also die Behauptung. ■

2. Schritt: „Induktion über die Raumdimension“. Der Beweis der Formel (12.7.1) für $f \in C_c^0(B_\varrho(y^0))$ wird zunächst für den Fall $n = 1$ geführt; hier reduziert sie sich auf die Substitutionsformel für Einfachintegrale:

Zu $y^0 \in \Theta \subset \mathbb{R}$ wählen wir $\varrho > 0$ mit $[y^0 - \varrho, y^0 + \varrho] \subset \Theta$. Setzen wir $a := \phi^{-1}(y^0 - \varrho)$, $b := \phi^{-1}(y^0 + \varrho)$, so können wir o.B.d.A. $\phi' > 0$ und somit $a < b$ annehmen. Dann folgt $[a, b] \subset \Omega$ und für beliebiges $f \in C_c^0((y^0 - \varrho, y^0 + \varrho))$ berechnen wir mit der Substitutionsformel in einer Veränderlichen, Satz 9.5.4:

$$\begin{aligned} \int_{\Theta} f(y) dy &= \int_{y^0 - \varrho}^{y^0 + \varrho} f(y) dy = \int_a^b f(\phi(x)) \phi'(x) dx \\ &\stackrel{\phi' > 0}{=} \int_{\Omega} f(\phi(x)) |J_{\phi}(x)| dx. \end{aligned}$$

Im Induktionsschritt wird angenommen, dass der Satz für Teilmengen des \mathbb{R}^{n-1} gilt und gezeigt, dass er dann auch für Teilmengen des \mathbb{R}^n richtig bleibt; dies gelingt durch iterierte Integration, nachdem der Satz über implizite Funktionen geeignet eingesetzt wurde. \square

12.8. Die allgemeine Transformationsformel - Kurzfassung

Nun wollen wir skizzieren, wie man mit Satz 12.7.1 die folgende allgemeine Version der Transformationsformel beweisen kann:

Satz 12.8.1: Transformationsformel

Es seien $\Omega, \Theta \subset \mathbb{R}^n$ offene Mengen und $\phi = \phi(x) : \Omega \rightarrow \Theta$ ein C^1 -Diffeomorphismus. Dann gilt für jedes $f \in C^0(\Theta)$ mit $\int_{\Theta} |f| dy < +\infty$ die Identität

$$\int_{\Theta} f(y) dy = \int_{\Omega} f(\phi(x)) |J_{\phi}(x)| dx. \quad (12.8.1)$$

Dabei handelt es sich bei den angegebenen Integralen um *uneigentliche Integrale*, die im \mathbb{R}^n wie folgt erklärt sind:

Definition 12.8.2: Ausschöpfungen

Es seien $\Omega \subset \mathbb{R}^n$ offen und $M_j \subset \subset \Omega, j \in \mathbb{N}$, gewählt. Dann heißt $\{M_j\}_{j \in \mathbb{N}}$ eine *Ausschöpfung*

von Ω , wenn für jedes Kompaktum $K \subset \Omega$ ein $j_0 = j_0(K) \in \mathbb{N}$ existiert mit

$$K \subset M_j \quad \text{für alle } j \geq j_0.$$

Wir schreiben dann $M_j \rightarrow \Omega$ ($j \rightarrow \infty$). Sind die Mengen M_j quadrierbar, so heißt die Ausschöpfung $\{M_j\}_j$ *quadrierbar*.

Bemerkung 12.8.3. Man zeigt recht leicht, dass zu jeder offenen Menge $\Omega \subset \mathbb{R}^n$ eine quadrierbare Ausschöpfung $\{M_j\}_j$ von Ω existiert (jedes M_j ist dabei eine endliche Vereinigung von Quadern).

Definition 12.8.4: Uneigentliches Integral

Seien $\Omega \subset \mathbb{R}^n$ offen und $f \in C^0(\Omega, \mathbb{R}^d)$ gegeben. Wenn für jede quadrierbare Ausschöpfung $\{M_j\}_j$ von Ω die Folge $\left\{ \int_{M_j} f \, dx \right\}_j$ konvergiert, so setzen wir

$$\int_{\Omega} f(x) \, dx := \lim_{j \rightarrow \infty} \int_{M_j} f(x) \, dx \quad (12.8.2)$$

für das *uneigentliche Integral* von f über Ω . Wir sagen dann auch, f ist *uneigentlich integrierbar* über Ω oder das uneigentliche Integral $\int_{\Omega} f \, dx$ *existiert* oder *konvergiert*. Schließlich heißt $\int_{\Omega} f \, dx$ *absolut konvergent*, wenn das Integral $\int_{\Omega} \|f(x)\| \, dx$ konvergiert. Wir schreiben dann auch $\int_{\Omega} \|f(x)\| \, dx < +\infty$.

Bemerkung 12.8.5. Man kann zeigen:

- (a) Die Definition ist unabhängig von der Wahl der quadrierbaren Ausschöpfung $\{M_j\}_j$ von Ω .
- (b) Ist Ω quadrierbar und $f \in C^0(\Omega, \mathbb{R}^d)$ beschränkt, so stimmt das uneigentliche mit dem Riemannschen Integral von f auf Ω überein.
- (c) Absolute Konvergenz von $\int_{\Omega} f \, dx$ impliziert Konvergenz im gewöhnlichen Sinne.
- (d) $\int_{\Omega} f \, dx$ konvergiert genau dann absolut, wenn $\int_M \|f\| \, dx \leq c$ für ein $c < +\infty$ und alle quadrierbaren $M \subset \subset \Omega$ gilt.
- (e) Ist $f \in C_c^0(\Omega)$, so existiert ein offenes, *quadrierbares* $\Omega' \subset \Omega$ mit $f \in C_c^0(\Omega')$ und das uneigentliche Integral (12.8.2) existiert mit $\int_{\Omega} f \, dx = \int_{\Omega'} f \, dx$. Satz 12.7.1 gilt daher auch für offene, nicht notwendig quadrierbare Mengen $\Omega, \Theta \subset \mathbb{R}^n$.

Wir werden Satz 12.8.1 aus Satz 12.7.1 durch Approximation gewinnen. Der dafür angemessene Konvergenzbegriff ist der folgende:

Definition 12.8.6

Eine Folge $\{f_j\}_j$ mit $f_j : \Omega \rightarrow \mathbb{R}^d \in C^0(\Omega, \mathbb{R}^d)$, $j = 1, 2, \dots$, heißt *kompakt gleichmäßig konvergent*, wenn für jede kompakte Teilmenge $K \subset \Omega$ die Einschränkungen $\{f_j|_K\}_j$ gleichmäßig konvergieren.

Bemerkung 12.8.7. Insbesondere existiert dann eine Grenzfunktion $f(x) := \lim_{j \rightarrow \infty} f_j(x)$, $x \in \Omega$, und nach dem WEIERSTRASSschen Konvergenzsatz, Satz 6.6.5, gilt $f \in C^0(\Omega, \mathbb{R}^d)$.

Die Vertauschbarkeit von Grenzwertbildung und Integration bei kompakt gleichmäßiger Konvergenz liefert der

Hilfssatz 12.8.8

Seien $\Omega \subset \mathbb{R}^n$ offen und $f_j \in C^0(\Omega, \mathbb{R}^d)$, $j = 1, 2, \dots$, kompakt gleichmäßig konvergent gegen $f \in C^0(\Omega, \mathbb{R}^d)$. Weiter sei $F \in C^0(\Omega)$ mit $\int_{\Omega} F \, dx < +\infty$ *integrable Majorante*, d.h. es gelte

$$\|f_j\| \leq F \quad \text{auf } \Omega, \quad j \in \mathbb{N}.$$

Dann konvergieren die uneigentlichen Integrale $\int_{\Omega} f_j \, dx$ für $j = 1, 2, \dots$ und $\int_{\Omega} f \, dx$ (sogar absolut), und es gilt

$$\int_{\Omega} \left(\lim_{j \rightarrow \infty} f_j \right) dx = \int_{\Omega} f \, dx = \lim_{j \rightarrow \infty} \left(\int_{\Omega} f_j \, dx \right).$$

Beweis von Satz 12.8.1. Wir wählen zu Ω eine quadrierbare Ausschöpfung $\{M_j\}_j$. Dann ist auch $\{N_j\}_j$ mit $N_j := \phi(M_j)$ eine quadrierbare Ausschöpfung von $\Theta = \phi(\Omega)$. Mit der Zerlegung der Eins findet man zu jedem $N_j \subset \subset \Theta$ ein $\eta_j \in C_c^\infty(\Theta)$ mit $\eta_j = 1$ auf N_j und $\eta_j(\Theta) \subset [0, 1]$. Die Funktionen $f_j := f\eta_j \in C_c^0(\Theta)$ sind dann über Θ integrierbar und es folgt aus Satz 12.7.1 und Bemerkung 12.8.5 (e):

$$\int_{\Theta} f_j \, dy = \int_{\Omega} (f_j \circ \phi) |J_{\phi}| \, dx, \quad j \in \mathbb{N}. \quad (12.8.3)$$

Die gleiche Relation gilt offenbar für $|f_j|$. Ferner konvergiert $\{f_j\}_j$ kompakt gleichmäßig gegen f auf Θ , und es gilt $|f_j| \leq |f|$ auf Θ sowie $\int_{\Theta} |f| \, dy < +\infty$. Hilfssatz 12.8.8 liefert also

$$\lim_{j \rightarrow \infty} \int_{\Theta} f_j \, dy = \int_{\Theta} f \, dy. \quad (12.8.4)$$

Analog konvergiert $\{(f_j \circ \phi)|J_\phi|\}_j$ kompakt gleichmäßig auf Ω gegen $(f \circ \phi)|J_\phi|$. Ist schließlich $M \subset \subset \Omega$ quadrierbar, so existiert ein $j \in \mathbb{N}$ mit $M \subset M_j$ und es folgt:

$$\begin{aligned} 0 &\leq \int_M |f \circ \phi| |J_\phi| dx \stackrel{f_j = f \eta_j}{=} \int_M |f_j \circ \phi| |J_\phi| dx \leq \int_\Omega |f_j \circ \phi| |J_\phi| dx \\ &\stackrel{(12.8.3) \text{ für } |f_j|}{=} \int_\Theta |f_j| dy \leq \int_\Theta |f| dy < +\infty. \end{aligned}$$

Nach Bemerkung 12.8.5 (d) existiert also das uneigentliche Integral $\int_\Omega |f \circ \phi| |J_\phi| dx$ und Hilfssatz 12.8.8 liefert

$$\int_\Omega (f \circ \phi) |J_\phi| dx = \lim_{j \rightarrow \infty} \left(\int_\Omega (f_j \circ \phi) |J_\phi| dx \right). \quad (12.8.5)$$

Die Formeln (12.8.3)-(12.8.5) ergeben schließlich die behauptete Identität (12.8.1). \square

Beispiel 12.8.9. (a) *Polarkoordinaten:* Es sei $f \in C^0(B_R)$ mit $\int_{B_R} |f| dy < +\infty$ auf der Kreisscheibe $B_R := B_R(0,0) \subset \mathbb{R}^2$ gegeben. Dann gilt

$$\int_{B_R} f(y) dy = \int_0^R \left(\int_0^{2\pi} f(r \cos \theta, r \sin \theta) r d\theta \right) dr. \quad (12.8.6)$$

Denn: Gemäß Aufgabe 1 (a) von Übungsblatt 11 ist die Abbildung $\phi : (0, R) \times (0, 2\pi) \rightarrow B_R \setminus ([0, R] \times \{0\})$ mit

$$\phi = \phi(r, \theta) := (r \cos \theta, r \sin \theta), \quad (r, \theta) \in (0, R) \times (0, 2\pi),$$

ein Diffeomorphismus mit $J_\phi = r$ auf $(0, R) \times (0, 2\pi)$. Da B_R quadrierbar ist und $[0, R] \times \{0\}$ den Inhalt Null hat, entnehmen wir Bemerkung 12.5.2 (c), Hilfssatz 12.6.10 (b) und der Transformationsformel

$$\int_{B_R} f dy = \int_{B_R \setminus ([0, R] \times \{0\})} f dy = \int_{(0, R) \times (0, 2\pi)} (f \circ \phi) |J_\phi| d(r, \theta).$$

Iterierte Integration liefert also die Behauptung (12.8.6). \square

Ist speziell f radialsymmetrisch, d.h. $f(y) = g(|y|)$ mit einem $g = g(r) \in C^0([0, R])$, so folgt

$$\int_{B_R} f(y) dy = \int_0^R \left(\int_0^{2\pi} g(r) r d\theta \right) dr = 2\pi \int_0^R g(r) r dr. \quad (12.8.7)$$

Noch spezieller, für $f \equiv 1$ auf B_R , erhalten wir somit

$$|B_R| = \int_{B_R} 1 dx = 2\pi \int_0^R r dr = \pi R^2.$$

(b) *GAUSSsches Fehlerintegral*: Wir behaupten

$$\int_{-\infty}^{+\infty} e^{-x^2} dx = \sqrt{\pi}.$$

Denn: In Beispiel 9.6.13 (b) haben wir bereits gesehen, dass das GAUSSsche Fehlerintegral existiert. Zur Berechnung setzen wir $W_R := [-R, R] \times [-R, R]$ für beliebiges $R > 0$ und beachten

$$\int_{W_R} e^{-(y_1^2+y_2^2)} dy \stackrel{\text{Satz 12.2.5}}{=} \left[\int_{-R}^R e^{-x^2} dx \right]^2 \leq c < +\infty \quad (12.8.8)$$

mit einer von $R > 0$ unabhängigen Konstanten $c \in [0, +\infty)$. Ist $M \subset \subset \mathbb{R}^2$ quadrierbar, so existiert ein $R > 0$ mit $M \subset W_R$, so dass nach (12.8.8) und Bemerkung 12.8.5 (d) das uneigentliche Integral von $e^{-(y_1^2+y_2^2)}$ über \mathbb{R}^2 existiert. Da $\{W_j\}_{j=1,2,\dots}$ quadrierbare Ausschöpfung des \mathbb{R}^2 ist, erhalten wir

$$\int_{\mathbb{R}^2} e^{-(y_1^2+y_2^2)} dy = \lim_{j \rightarrow \infty} \int_{W_j} e^{-(y_1^2+y_2^2)} dy = \left[\int_{-\infty}^{+\infty} e^{-x^2} dx \right]^2. \quad (12.8.9)$$

Andererseits ist auch $\{B_j\}_{j=1,2,\dots}$ mit den Kreisscheiben $B_j = B_j(0,0)$ eine quadrierbare Ausschöpfung von \mathbb{R}^2 . Formeln (12.8.7) und (12.8.9) liefern also

$$\begin{aligned} \int_{-\infty}^{+\infty} e^{-x^2} dx &= \left[\int_{\mathbb{R}^2} e^{-(y_1^2+y_2^2)} dy \right]^{\frac{1}{2}} = \left[\lim_{j \rightarrow \infty} \int_{B_j} e^{-(y_1^2+y_2^2)} dy \right]^{\frac{1}{2}} \\ &= \left[\pi \lim_{j \rightarrow \infty} \int_0^j e^{-r^2} (2r) dr \right]^{\frac{1}{2}} = \sqrt{\pi} \left[- \lim_{j \rightarrow \infty} \int_0^j \frac{d}{dr} (e^{-r^2}) dr \right]^{\frac{1}{2}} \\ &= \sqrt{\pi} \left[\lim_{j \rightarrow \infty} (1 - e^{-j^2}) \right]^{\frac{1}{2}} = \sqrt{\pi}, \end{aligned}$$

wie behauptet. □

Anhang A.

Die Transformationsformel

In diesem soll der in den Abschnitten 12.7 und 12.8 skizzierte Beweis der Transformationsformel für Mehrfachintegrale vollständig ausgeführt werden; dieser Teil ist informativ zu verstehen, enthält also weder Klausur- noch Prüfungsstoff.

In Abschnitt A.1.1 diskutieren wir dazu zunächst den Begriff der Testfunktion und beweisen den wichtigen Satz A.1.5 über die Existenz einer (geeigneten) Zerlegung der Eins. In Abschnitt A.2 untersuchen wir das Verhalten quadrierbarer Mengen unter Diffeomorphismen und konstruieren quadrierbare Ausschöpfungen offener Mengen. Abschnitt A.3 enthält dann den Beweis der Transformationsformel für Testfunktionen, Satz A.3.1. Mit Hilfe der quadrierbaren Ausschöpfungen aus Abschnitt A.2 erklären wir in Abschnitt A.4 uneigentliche Integrale im \mathbb{R}^n und geben eine Auswahl ihrer Eigenschaften an. Schließlich wird in Abschnitt A.5 die allgemeine Transformationsformel, Satz A.5.1, bewiesen.

A.1. Testfunktionen

Vorweg sei erwähnt, dass wir in diesem Anhang die Ergebnisse aus Abschnitt 12.6 zur Additivität des Integrals und zur Integration über quadrierbare Nullmengen immer wieder *kommentarlos* anwenden werden.

Wir werden die Transformationsformel in Abschnitt A.3 zunächst für sogenannte *Testfunktionen* beweisen, siehe Satz A.3.1; diese sollen hier erklärt und untersucht werden.

Definition A.1.1

- Ist $M \subset \mathbb{R}^n$ beliebig und $f : M \rightarrow \mathbb{R}^d$ gegeben. Dann heißt die Menge

$$\operatorname{supp} f := \overline{\{x \in M : f(x) \neq 0\}}$$

der *Träger* oder *Support* von f .

- Ist $M \subset \mathbb{R}^n$ beliebig, so heißt $M' \subset M$ *kompakt enthalten* in M , i.Z. $M' \subset\subset M$, wenn

$\overline{M'}$ kompakt ist und $\overline{M'} \subset M$ erfüllt.

- Sind $\Omega \subset \mathbb{R}^n$ offen und $s \in \mathbb{N}_0 \cup \{\infty\}$, so bezeichnet

$$C_c^s(\Omega, \mathbb{R}^d) := \{f \in C^s(\Omega, \mathbb{R}^d) : \text{supp } f \subset\subset \Omega\}$$

die Menge der *s-mal stetig differenzierbaren, \mathbb{R}^d -wertigen Funktionen mit kompaktem Träger in Ω* . Ein solches $f \in C_c^s(\Omega, \mathbb{R}^d)$ nennen wir auch kurz *Testfunktion*. Wir setzen noch $C_c^s(\Omega) := C_c^s(\Omega, \mathbb{R})$.

Bemerkung A.1.2. ► Jede Testfunktion $g \in C_c^0(\Omega)$ ist offenbar beschränkt und somit gemäß Folgerung 12.5.10 über quadrierbares Ω integrierbar.

- Wir können uns eine Funktion $f \in C_c^s(\Omega, \mathbb{R}^d)$ immer auf ganz \mathbb{R}^n erklärt denken, indem wir f zu 0 auf $\mathbb{R}^n \setminus \Omega$ fortsetzen. Dann ist $f \in C^s(\mathbb{R}^n, \mathbb{R}^d)$ und $\text{supp } f \subset\subset \Omega$ richtig.

Wir wollen hier einige wichtige Eigenschaften von Testfunktionen angeben. Dazu beginnen wir mit der

Definition A.1.3: Zerlegung der Eins

Sei $M \subset \mathbb{R}^n$ nichtleer. Eine *Zerlegung der Eins* auf M ist eine Familie $\{\eta_\alpha\}_{\alpha \in I}$ von Funktionen $\eta_\alpha \in C_c^\infty(\mathbb{R}^n)$ mit beliebiger Indexmenge I und den folgenden Eigenschaften:

- (a) Für alle $\alpha \in I$ gilt $0 \leq \eta_\alpha \leq 1$ auf \mathbb{R}^n .
- (b) Für jedes $x \in \mathbb{R}^n$ existieren höchstens endlich viele $\alpha \in I$ mit $\eta_\alpha(x) \neq 0$.
- (c) Es gilt $\sum_{\alpha \in I} \eta_\alpha(x) = 1$ für alle $x \in M$ sowie $0 \leq \sum_{\alpha \in I} \eta_\alpha \leq 1$ auf \mathbb{R}^n .

Die Zerlegung der Eins heißt *endlich*, wenn I endlich ist.

Bemerkung A.1.4. Ist $\{\eta_\alpha\}_\alpha$ Zerlegung der Eins auf M und $M' \subset M$ nichtleer, so ist $\{\eta_\alpha\}_\alpha$ auch Zerlegung der Eins auf M' .

Satz A.1.5

Sei $K \subset \mathbb{R}^n$ kompakt und $\mathcal{F} = \{U(x) : x \in K\}$ eine beliebige offene Überdeckung, wobei $U(x)$ eine offene Umgebung von $x \in K$ bezeichne. Dann gibt es $p \in \mathbb{N}$ Punkte $x_1, \dots, x_p \in K$ und eine endliche Zerlegung der Eins $\{\eta_\alpha\}_{\alpha=1, \dots, p}$ auf K mit der zusätzlichen Eigenschaft $\eta_\alpha \in C_c^\infty(U(x_\alpha))$ für $\alpha = 1, \dots, p$.

Beweis. 1. Da die $U(x)$ offen sind, existiert zu jedem $x \in K$ ein $r(x) \in (0, 1)$ mit $B(x) := B_{r(x)}(x) \subset\subset U(x)$. Ferner gibt es eine Kugel $B := B_R(0)$ mit

$$\bigcup_{x \in K} B(x) \subset\subset B.$$

Und zu jedem $x \in \bar{B} \setminus K$ finden wir eine Kugel $B(x) = B_{r(x)}(x) \subset \mathbb{R}^n \setminus K$. Aus der so gewonnen offenen Überdeckung $\hat{\mathcal{F}} := \{B(x) : x \in \bar{B}\}$ der kompakten Kugel \bar{B} können wir nach Satz 12.3.3 von HEINE-BOREL endlich viele Kugeln $B_\alpha := B(x_\alpha)$, $\alpha = 1, \dots, N$, auswählen, die \bar{B} überdecken. Diese sortieren wir so, dass

$$x_\alpha \in \begin{cases} K, & \text{für } \alpha = 1, \dots, p \\ \bar{B} \setminus K, & \text{für } \alpha = p+1, \dots, N \end{cases}$$

mit einem $p < N$ richtig ist. Nach Konstruktion gilt dann $B_\alpha \cap K = \emptyset$ für $\alpha = p+1, \dots, N$.

2. Mit dem Radius $r_\alpha := r(x_\alpha) > 0$ der Kugel B_α setzen wir

$$\xi_\alpha(x) := \psi(r_\alpha^2 - |x - x_\alpha|^2), \quad x \in \mathbb{R}^n,$$

mit der Funktion

$$\psi(t) := \begin{cases} e^{-\frac{1}{t}}, & \text{für } t > 0 \\ 0, & \text{für } t \leq 0 \end{cases} \in C^\infty(\mathbb{R}).$$

Dann gilt $\xi_\alpha \in C_c^\infty(\mathbb{R}^n)$ für alle $\alpha = 1, \dots, N$ (vgl. Beispiel 9.8.3) und wir bemerken

$$\xi_\alpha(x) > 0 \Leftrightarrow x \in B_\alpha \quad \text{für alle } \alpha = 1, \dots, N.$$

Insbesondere haben wir also $\text{supp } \xi_\alpha = \bar{B}_\alpha \subset B \cap U(x_\alpha)$ für $\alpha = 1, \dots, p$ und $\xi_\alpha = 0$ auf K für $\alpha = p+1, \dots, N$. Schließlich setzen wir noch

$$\xi(x) := \sum_{\alpha=1}^N \xi_\alpha(x), \quad x \in \mathbb{R}^n,$$

und beachten $\xi > 0$ auf \bar{B} . Für

$$\eta_\alpha(x) := \begin{cases} \frac{\xi_\alpha(x)}{\xi(x)}, & \text{für } x \in B \\ 0, & \text{für } x \in \mathbb{R}^n \setminus B \end{cases} \in C_c^\infty(U(x_\alpha)), \quad \alpha = 1, \dots, p,$$

sind dann die Eigenschaften (a)-(c) aus Definition A.1.3 erfüllt. \square

Wir schließen mit der praktischen

Folgerung A.1.6

Seien $\Omega \subset \mathbb{R}^n$ offen und $K \subset \Omega$ kompakt. Dann existiert ein $\eta \in C_c^\infty(\Omega)$ mit $\eta = 1$ auf K und $0 \leq \eta \leq 1$ auf Ω .

Beweis. Zu jedem $x \in K$ gibt es ein $r(x) > 0$ mit $B(x) := B_{r(x)}(x) \subset \Omega$. Nach Hilfssatz A.1.5 können wir zur Überdeckung $\mathcal{F} = \{B(x) : x \in K\}$ Punkte $x_1, \dots, x_p \in K$ und eine endliche Zerlegung der Eins $\{\eta_\alpha\}_{\alpha=1, \dots, p}$ auf K mit $\eta_\alpha \in C_c^\infty(B(x_\alpha))$, $\alpha = 1, \dots, p$, angeben. Die Funktion $\eta := \sum_{\alpha=1}^p \eta_\alpha$ leistet dann das Gewünschte. \square

A.2. Quadrierbarkeit und Diffeomorphismen

Für die folgenden Untersuchungen müssen wir wissen, wie sich quadrierbare (Teil-)Mengen unter diffeomorphen Abbildungen verhalten. Dazu beginnen wir mit folgendem Begriff:

Definition A.2.1: Ausschöpfungen

Es seien $\Omega \subset \mathbb{R}^n$ offen und $M_j \subset \Omega$, $j \in \mathbb{N}$, gewählt. Dann heißt $\{M_j\}_{j \in \mathbb{N}}$ eine *Ausschöpfung* von Ω , wenn für jedes Kompaktum $K \subset \Omega$ ein $j_0 = j_0(K) \in \mathbb{N}$ existiert mit

$$K \subset M_j \quad \text{für alle } j \geq j_0.$$

Wir schreiben dann $M_j \rightarrow \Omega$ ($j \rightarrow \infty$). Sind die Mengen M_j quadrierbar, so heißt die Ausschöpfung $\{M_j\}_j$ *quadrierbar*.

Natürlich existiert immer die triviale Ausschöpfung $\{M_j\}_j = \{\Omega\}_j$. Wir wollen zeigen, dass es zu jeder offenen Menge auch eine *quadrierbare* Ausschöpfung gibt. Dazu benötigen wir noch den

Hilfssatz A.2.2

Seien $A, K \subset \mathbb{R}^n$ nichtleer, A abgeschlossen, K kompakt und es gelte $A \cap K = \emptyset$. Dann folgt für die *Distanz zwischen A und K* :

$$\text{dist}(A, K) := \inf \{\|x - y\| : x \in A, y \in K\} > 0.$$

Beweis. Wäre $\text{dist}(A, K) = 0$, so existierten Folgen $\{x_l\}_l \subset A$ und $\{y_l\}_l \subset K$ mit

$$\|x_l - y_l\| \rightarrow 0 \quad (l \rightarrow \infty). \quad (\text{A.2.1})$$

Da K kompakt ist, können wir nach Satz 5.3.8 eine konvergente Teilfolge $\{y_{l_k}\}_k$ auswählen mit $\xi := \lim_{k \rightarrow \infty} y_{l_k} \in K$. Aus (A.2.1) folgt dann auch $x_{l_k} \rightarrow \xi$ ($k \rightarrow \infty$), d.h. $x \in A$, denn A ist abgeschlossen. Wir haben also den Widerspruch $\xi \in A \cap K$ und somit $\text{dist}(A, K) > 0$, wie behauptet. \square

Satz A.2.3: Ausschöpfungslemma

Zu jeder offenen Menge $\Omega \subset \mathbb{R}^n$ existiert eine quadrierbare Ausschöpfung $\{M_j\}_j$ von Ω . Dabei können die M_j als endliche Vereinigungen von Würfeln - d.h. insbesondere kompakt - gewählt werden.

Beweis. Wir setzen $W_j := [-j, j] \times \dots \times [-j, j] \subset \mathbb{R}^n$, $j \in \mathbb{N}$. Für $\Omega = \mathbb{R}^n$ ist $\{W_j\}_j$ bereits die gesuchte quadrierbare Ausschöpfung. Für $\Omega \neq \mathbb{R}^n$ betrachten wir eine äquidistante Zerlegung \mathcal{Z}_j des Würfels W_j in Teilwürfel $W_{j\alpha}$, $\alpha \in A_j$, mit $\text{diam } W_{j\alpha} \leq \frac{1}{j}$ für alle $\alpha \in A_j$. Die Mengen

$$M_j := \bigcup_{\alpha \in A_j: W_{j\alpha} \subset \Omega} W_{j\alpha}, \quad j \in \mathbb{N},$$

sind dann kompakte quadrierbare Mengen mit $M_j = \overline{M_j} \subset \Omega$.

Ist $K \subset \Omega$ eine beliebige kompakte Teilmenge, so existiert ein $j_1 = j_1(K) \in \mathbb{N}$ mit $K \subset W_j$ für alle $j \geq j_1$. Nach Hilfssatz A.2.2 gilt weiter

$$d := \text{dist}(\mathbb{R}^n \setminus \Omega, K) > 0.$$

Wählen wir $j_0 \geq j_1$ mit $\frac{1}{j_0} < d$, so folgt

$$\text{diam } W_{j\alpha} \leq \frac{1}{j} < d \quad \text{für alle } j \geq j_0, \alpha \in A_j.$$

Sind nun $x \in K$ und $j \geq j_0$ beliebig gewählt, so existiert ein $\alpha \in A_j$ mit $x \in W_{j\alpha}$. Für alle $y \in W_{j\alpha}$ folgt dann $\|y - x\| < d$, also $y \in \Omega$, d.h. $W_{j\alpha} \subset \Omega$. Folglich ist $x \in M_j$, d.h. $K \subset M_j$ für alle $j \geq j_0$, wie behauptet. \square

Bemerkung A.2.4. (a) Zu jeder offenen Menge existiert sogar eine *monotone* quadrierbare Ausschöpfung, d.h. eine quadrierbare Ausschöpfung $\{M_j\}_j$ von Ω , für die gilt $M_j \subset M_{j+1}$ für alle $j \in \mathbb{N}$.

Denn: Ist $\{M'_l\}_l$ eine gemäß Satz A.2.3 existente, beliebige quadrierbare Ausschöpfung, so setzen wir einfach $M_j := \bigcup_{l=1}^j M'_l$, $j \in \mathbb{N}$. \square

- (b) Ist $\{M_j\}_j$ eine Ausschöpfung von Ω , so existiert zu jedem Kompaktum $K \subset \Omega$ ein $j'_0 = j'_0(K) \in \mathbb{N}$ so, dass sogar gilt

$$K \subset \mathring{M}_j \quad \text{für alle } j \geq j'_0.$$

Denn: Mit dem Satz von HEINE-BOREL finden wir für ein beliebiges positives $r = r(K) < \text{dist}(\mathbb{R}^n \setminus \Omega, K)$ Punkte $x_1, \dots, x_p \in K$, so dass $K \subset \bigcup_{l=1}^p B_r(x_l) \subset \subset \Omega$ gilt. Zu $K' := \overline{\bigcup_{l=1}^p B_r(x_l)} \subset \Omega$ finden wir dann gemäß Definition A.2.1 ein $j_0 = j_0(K')$ mit $K' \subset M_{j_0}$ und folglich auch

$$K \subset K' \subset \mathring{M}_j \quad \text{für alle } j \geq j'_0(K) := j_0(K'),$$

wie behauptet. \square

Satz A.2.5

Seien $\Omega, \Theta \subset \mathbb{R}^n$ offen und $\phi : \Omega \rightarrow \Theta$ ein Diffeomorphismus. Ist dann $M \subset \subset \Omega$ eine quadrierbare Teilmenge, so gilt $\phi(\partial M) = \partial \phi(M)$ und auch $\phi(M) \subset \subset \Theta$ ist quadrierbar.

Beweis. 1. Zunächst bemerken wir $\partial \phi(M) \subset \phi(\partial M)$. Sind nämlich $y_0 \in \partial \phi(M)$ und $\varepsilon > 0$ beliebig, so existieren Folgen $\{y_j^I\}_j \subset \phi(M)$ und $\{y_j^A\}_j \subset \Theta \setminus \phi(M)$ mit $y_j^I \rightarrow y_0$, $y_j^A \rightarrow y_0$ ($j \rightarrow \infty$). Sind $x_j^I := \phi^{-1}(y_j^I) \in M$, $x_j^A := \phi^{-1}(y_j^A) \in \Omega \setminus M$ für $j \in \mathbb{N}$ und $x^0 := \phi^{-1}(y_0)$ gesetzt, so liefert die Stetigkeit von ϕ^{-1} auch $x_j^I \rightarrow x^0$, $x_j^A \rightarrow x^0$ ($j \rightarrow \infty$). Folglich ist $x^0 \in \partial M$ bzw. $y_0 = \phi(x^0) \in \phi(\partial M)$.

Völlig analog erhalten wir $\partial \phi^{-1}(\phi(M)) \subset \phi^{-1}(\partial \phi(M))$ und somit $\phi(\partial M) \subset \partial \phi(M)$, insgesamt gilt also

$$\phi(\partial M) = \partial \phi(M).$$

2. Mit Satz A.2.3 wählen wir eine quadrierbare Ausschöpfung $\{M_j\}_j$ von Ω . Dann haben wir also $\overline{M} \subset \Omega' \subset \subset \Omega$ mit $\Omega' := \mathring{M}_j$ für ein hinreichend großes $j \in \mathbb{N}$, siehe Bemerkung A.2.4. Nach Voraussetzung gilt $|\partial M| = 0$ gemäß Satz 12.5.3, zu beliebigem $\varepsilon > 0$ existieren also Quader $Q_1, \dots, Q_p \subset \mathbb{R}^n$ mit $\partial M \subset \bigcup_{l=1}^p \mathring{Q}_l$ und $\sum_{l=1}^p |Q_l| < \varepsilon$. Wir können durch eventuelle Verkleinerung $\bigcup_l \mathring{Q}_l \subset \Omega'$ annehmen und zusätzlich, dass alle Kantenlängen der Q_l rational sind. Dann können wir die Q_l in (gleichgroße) Würfel W_1, \dots, W_r zerlegen, es gilt also

$$\partial M \subset \bigcup_{k=1}^r W_k, \quad \sum_{k=1}^r |W_k| < \varepsilon.$$

Setzen wir $L := \sup_{\Omega'} \|D\phi\| \in [0, +\infty)$, so liefern Satz 10.2.11 und die Konvexität der $W_k \subset \Omega'$:

$$\|\phi(x_1) - \phi(x_2)\| \leq L \|x_1 - x_2\| \quad \text{für alle } x_1, x_2 \in W_k, \quad k = 1, \dots, r.$$

Folglich gilt $\phi(W_k) \subset W'_k$, $k = 1, \dots, r$, für Würfel W'_k mit $\sqrt{n}L$ -facher Kantenlänge, d.h. $|W'_k| = (\sqrt{2}L)^n |W_k|$, $k = 1, \dots, r$. Insgesamt haben wir also

$$\phi(\partial M) = \phi\left(\bigcup_{k=1}^r (\partial M \cap W_k)\right) \subset \bigcup_{k=1}^r \phi(\partial M \cap W_k) \subset \bigcup_{k=1}^r W'_k, \quad \sum_{k=1}^r |W'_k| < (\sqrt{2}L)^n \varepsilon.$$

Wählen wir konzentrische Würfel W''_k zu W'_k mit doppelter Kantenlänge, so liefert Teil 1 schließlich

$$\partial \phi(M) = \phi(\partial M) \subset \bigcup_{k=1}^r W''_k, \quad \sum_{k=1}^r |W''_k| < (2\sqrt{2}L)^n \varepsilon.$$

Da $\varepsilon > 0$ beliebig war, ist also $|\partial \phi(M)| = 0$, d.h. $\phi(M)$ ist quadrierbar gemäß Satz 12.5.3. \square

Folgerung A.2.6

Seien $\Omega, \Theta \subset \mathbb{R}^n$ offen, $\phi : \Omega \rightarrow \Theta$ diffeomorph und $\{M_j\}_j$ eine quadrierbare Ausschöpfung von Ω mit $M_j \subset \subset \Omega$ für alle $j \in \mathbb{N}$. Dann ist auch $\{\phi(M_j)\}_j$ eine quadrierbare Ausschöpfung von Θ .

Beweis. Zunächst ist $\phi(M_j)$ für alle $j \in \mathbb{N}$ quadrierbar gemäß Satz A.2.5. Weiter ist $\{\phi(M_j)\}_j$ Ausschöpfung von Θ . Ist nämlich $K \subset \Theta$ ein Kompaktum, so ist auch $\phi^{-1}(K) \subset \Omega$ kompakt. Nach Definition existiert somit ein $j_0 = j_0(K) \in \mathbb{N}$ mit $\phi^{-1}(K) \subset M_{j_0}$ bzw. $K = \phi(\phi^{-1}(K)) \subset \phi(M_{j_0})$ für alle $j \geq j_0$. Damit ist schon alles gezeigt. \square

A.3. Die Transformationsformel für Testfunktionen

In diesem Abschnitt beweisen wir den

Satz A.3.1: Transformationsformel für Testfunktionen

Seien $\Omega, \Theta \subset \mathbb{R}^n$ offene, quadrierbare Mengen und $\phi : \Omega \rightarrow \Theta$ ein Diffeomorphismus. Dann gilt für beliebige Testfunktionen $f \in C_c^0(\Theta)$ die Identität

$$\int_{\Theta} f(y) dy = \int_{\Omega} f(\phi(x)) |J_{\phi}(x)| dx. \quad (\text{A.3.1})$$

Für den Beweis dieses Spezialfalls des allgemeinen Satzes A.5.1 benötigen wir einige Hilfssätze.

Hilfssatz A.3.2: Lokalisierung

Seien $\Omega, \Theta \subset \mathbb{R}^n$ offene, quadrierbare Mengen und $\phi : \Omega \rightarrow \Theta$ ein Diffeomorphismus. Dann gilt (A.2.1) genau dann für alle $f \in C_c^0(\Theta)$, wenn zu jedem $y^0 \in \Theta$ ein $\varrho = \varrho(y^0) > 0$ mit $B_{\varrho}(y^0) \subset \Theta$ so existiert, dass (A.2.1) für alle $f \in C_c^0(B_{\varrho}(y^0))$ erfüllt ist.

Beweis. ▶ „ \Rightarrow “: Klar, wegen $C_c^0(B_{\varrho}(y^0)) \subset C_c^0(\Theta)$ für $B_{\varrho}(y^0) \subset \Theta$.

▶ „ \Leftarrow “: Sei also $f \in C_c^0(\Theta)$ gewählt, d.h. $K := \text{supp } f \subset \Theta$ ist kompakt. Dann existiert nach Satz A.1.5 eine endliche Zerlegung der Eins $\{\eta_{\alpha}\}_{\alpha=1,\dots,p}$ auf K mit $\eta_{\alpha} \in C_c^{\infty}(B_{\varrho(y_{\alpha})}(y_{\alpha}))$ für $\alpha = 1, \dots, p$, wobei $\varrho(y_{\alpha}) > 0$ wie in der Voraussetzung gewählt seien. Es folgt nun mit $g_{\alpha} := f\eta_{\alpha} \in C_c^0(B_{\varrho(y_{\alpha})}(y_{\alpha}))$ nach Voraussetzung:

$$\begin{aligned} 0 &= \sum_{\alpha=1}^p \left(\int_{\Theta} g_{\alpha}(y) dy - \int_{\Omega} g_{\alpha}(\phi(x)) |J_{\phi}(x)| dx \right) \\ &= \int_{\Theta} \left[\sum_{\alpha=1}^p \eta_{\alpha}(y) \right] f(y) dy - \int_{\Omega} \left[\sum_{\alpha=1}^p \eta_{\alpha}(\phi(x)) \right] f(\phi(x)) |J_{\phi}(x)| dx \\ &= \int_{\Theta} f(y) dy - \int_{\Omega} f(\phi(x)) |J_{\phi}(x)| dx, \end{aligned}$$

also die Behauptung. \square

Hilfssatz A.3.3

Seien $\Omega, \Theta \subset \mathbb{R}^{1+n}$ offene quadrierbare Mengen und $\phi = (\sigma, \text{Id}) : \Omega \rightarrow \Theta$ ein Diffeomorphismus mit einem $\sigma \in C^1(\Omega)$. Weiter seien $I \subset \mathbb{R}$ und $S \subset \mathbb{R}^n$ Quader mit $Q = I \times S \subset \Omega$. Setzen wir noch $K := \phi(Q) \subset \Theta$, so gilt mit diesem ϕ die Transformationsformel (A.2.1) für jedes $f \in C_c^0(\overset{\circ}{K})$.

Beweis. Wir beachten zunächst

$$D\phi = \begin{pmatrix} \sigma_t & D_z \sigma \\ 0 & \mathbb{E} \end{pmatrix} \quad \text{in } \Omega$$

und folglich

$$J_{\phi} = \sigma_t \quad \text{in } \Omega. \quad (\text{A.3.2})$$

Wegen $J_{\phi} \neq 0$ in Ω kann also σ_t in der zusammenhängenden Menge Q ihr Vorzeichen nicht wechseln. O.B.d.A. sei $\sigma_t > 0$ in Q . Schreiben wir $I = [a, b]$, so gilt dann insbesondere

$\sigma(a, z) < \sigma(b, z)$ für alle $z \in R$, und

$$K = \phi(Q) = \{(\sigma(t, z), z) : t \in I, z \in Q\} = \{(s, z) : z \in S, \sigma(a, z) \leq s \leq \sigma(b, z)\}$$

ist ein Normalbereich; siehe Definition 12.6.1. Für beliebiges $f \in C_c^0(\Theta)$ mit $\text{supp } f \subset K$ erhalten wir daher

$$\begin{aligned} \int_{\Theta} f d(s, z) &= \int_K f d(s, z) \stackrel{\text{Satz 12.6.3}}{=} \int_S \left(\int_{\sigma(a, z)}^{\sigma(b, z)} f(s, z) ds \right) dz \\ &\stackrel{\text{Satz 9.5.4}}{=} \int_S \left(\int_a^b f(\sigma(t, z), z) \sigma_t(t, z) dt \right) dz \\ &\stackrel{\text{(A.3.2)}}{=} \int_S \left(\int_I (f \circ \phi) |J_{\phi}| dt \right) dz \stackrel{\text{Satz 12.2.5}}{=} \int_{\Omega} (f \circ \phi) |J_{\phi}| d(t, z), \end{aligned}$$

wie behauptet. \square

Hilfssatz A.3.4

Seien $\Omega, \Theta \subset \mathbb{R}^{1+n}$ offene quadrierbare Mengen und $\phi = (\text{Id}, \gamma) : \Omega \rightarrow \Theta$ eine Diffeomorphismus mit einem $\gamma \in C^1(\Omega, \mathbb{R}^n)$. Weiter seien $I \subset \mathbb{R}$ und $S \subset \mathbb{R}^n$ Quader mit $Q = I \times S \subset \Omega$.

- (a) Dann existiert eine offene quadrierbare Menge $\underline{\Omega} \supset S$ mit $I \times \underline{\Omega} \subset \subset \Omega$ so, dass $\gamma(t, \cdot) : \underline{\Omega} \rightarrow \Theta(t)$ für jedes $t \in I$ diffeomorph ist; dabei haben wir noch $\Theta(t) := \gamma(t, \underline{\Omega}) \subset \mathbb{R}^n$, $t \in I$, gesetzt.
- (b) Wir setzen $K := \phi(I \times S)$. Ist dann für jedes $f \in C_c^0(\Theta)$ mit $\text{supp } f \subset K$ und jedes $t \in I$ die Gleichung

$$\int_{\Theta(t)} f(t, \zeta) d\zeta = \int_{\underline{\Omega}} f(t, \gamma(t, z)) |\det D\gamma_z(t, z)| dz \quad (\text{A.3.3})$$

erfüllt, so gilt mit diesen f und ϕ die Transformationsformel (A.3.1).

Beweis. (a) Wir können $\underline{\Omega} = \overset{\circ}{S}_0$ mit einem zu S konzentrischen Quader $S_0 \subset \mathbb{R}^n$ wählen, der S im Innern enthält und dessen Durchmesser $\text{diam } S_0 \leq \text{diam } S + \varepsilon$ mit einem $\varepsilon < \text{dist}(Q, \Omega^c)$ erfüllt. Wegen

$$D\phi = \begin{pmatrix} 1 & 0 \\ \gamma_t & D_z \gamma \end{pmatrix}, \quad J_{\phi} = \det D_z \gamma \neq 0 \quad \text{in } \Omega \quad (\text{A.3.4})$$

ist dann $\gamma(t, \cdot) : \underline{\Omega} \rightarrow \underline{\Theta}(t)$ für jedes $t \in I$ gemäß Folgerung 11.4.8 ein Diffeomorphismus. Wir bemerken noch, dass $\underline{\Theta}(t)$ für alle $t \in I$ offen und nach Satz A.2.5 auch quadrierbar ist.

(b) Wir setzen $\underline{K}(t) := \gamma(t, S) \subset \underline{\Theta}(t)$, $t \in I$. Wegen der speziellen Form von ϕ gilt

$$K = \phi(I \times S) = \{(t, \gamma(t, z)) : t \in I, z \in S\} \subset I \times \mathbb{R}^n.$$

Da K kompakt ist, existiert also ein Quader T mit $K \subset I \times T$; insbesondere gilt dann $\underline{K}(t) \subset T$ für alle $t \in I$. Gemäß Satz A.2.5 sind ferner K , $\underline{K}(t)$ quadrierbar für alle $t \in I$. Denken wir uns f noch zu 0 auf den \mathbb{R}^{1+n} fortgesetzt so folgt

$$\begin{aligned} \int_{\underline{\Theta}} f d(t, \zeta) &= \int_K f d(t, \zeta) = \int_{I \times T} f d(t, \zeta) \\ &\stackrel{\text{Satz 12.2.5}}{=} \int_I \left(\int_{\underline{K}(t)} f d\zeta \right) dt = \int_I \left(\int_{\underline{\Theta}(t)} f d\zeta \right) dt \\ &\stackrel{(\text{A.3.3})}{=} \int_I \left(\int_{\underline{\Omega}} f(t, \gamma(t, z)) |\det D\gamma_z(t, z)| dz \right) dt \\ &= \int_I \left(\int_S f(t, \gamma(t, z)) |\det D\gamma_z(t, z)| dz \right) dt \\ &\stackrel{(\text{A.3.4})}{=} \int_I \left(\int_S (f \circ \phi) |J_\phi| dz \right) dt \stackrel{\text{Satz 12.2.5}}{=} \int_{\Omega} (f \circ \phi) |J_\phi| d(t, z), \end{aligned}$$

wie behauptet. □

Hilfssatz A.3.5

Seien $\Omega, \Theta, \Delta \subset \mathbb{R}^n$ offene quadrierbare Mengen und $\phi : \Omega \rightarrow \Theta$, $\psi : \Delta \rightarrow \Omega$, $\chi : \Delta \rightarrow \Theta$ Diffeomorphismen mit $\phi \circ \psi = \chi$ auf Δ . Weiter sei $f \in C_c^0(\Theta)$ gewählt und $g := (f \circ \phi) |J_\phi| \in C_c^0(\Omega)$ gesetzt. Wenn dann die Transformationsformeln

$$\begin{aligned} \int_{\Theta} f dy &= \int_{\Delta} (f \circ \chi) |J_\chi| dw, \\ \int_{\Omega} g dx &= \int_{\Delta} (g \circ \psi) |J_\psi| dw \end{aligned}$$

gelten, so gilt auch (A.3.1) mit diesen f und ϕ .

Beweis. Wegen $\chi = \phi \circ \psi$ gilt $J_\chi = (J_\phi \circ \psi) J_\psi$ in Δ . Wir berechnen daher mit Hilfe der

Voraussetzungen:

$$\begin{aligned} \int_{\Theta} f \, dy &= \int_{\Delta} (f \circ \chi) |J_{\chi}| \, dw = \int_{\Delta} (f \circ \phi \circ \psi) |J_{\phi \circ \psi}| |J_{\psi}| \, dw \\ &= \int_{\Delta} (g \circ \psi) |J_{\psi}| \, dw = \int_{\Omega} g \, dx = \int_{\Omega} (f \circ \phi) |J_{\phi}| \, dx, \end{aligned}$$

wie behauptet. \square

Beweis von Satz A.3.1. Wegen Hilfssatz A.3.2 genügt es, die Aussage für $f \in C_c^0(B_{\varrho}(y^0))$ mit beliebigem $y^0 \in \Theta$ und beliebig wählbarem Radius $\varrho = \varrho(y^0) > 0$ mit $B_{\varrho}(y^0) \subset \Theta$ zu zeigen. Für diesen Beweis nutzen wir eine vollständige Induktion über die Raumdimension $n \in \mathbb{N}$:

- **$n = 1$:** Zu $y^0 \in \Theta \subset \mathbb{R}$ wählen wir $\varrho > 0$ mit $[y^0 - \varrho, y^0 + \varrho] \subset \Theta$. Setzen wir $a := \phi^{-1}(y^0 - \varrho)$, $b := \phi^{-1}(y^0 + \varrho)$, so können wir o.B.d.A. $a < b$ annehmen. Dann gilt $[a, b] \subset \Omega$ und für beliebiges $f \in C_c^0((y^0 - \varrho, y^0 + \varrho))$ berechnen wir mit der Transformationsformel in einer Veränderlichen, Satz 9.5.4:

$$\int_{\Theta} f \, dy = \int_{y^0 - \varrho}^{y^0 + \varrho} f(y) \, dy = \int_a^b f(\phi(x)) \phi'(x) \, dx \stackrel{\phi' > 0}{=} \int_{\Omega} (\phi) |J_{\phi}| \, dx.$$

- **$n \rightarrow n + 1$:** Wir fixieren $y^0 \in \Theta \subset \mathbb{R}^{n+1}$ und schreiben $y = (s, z) \in \Theta$ mit $s \in \mathbb{R}$, $z \in \mathbb{R}^n$, speziell also $y^0 = (s^0, z^0)$. Ferner schreiben wir $\phi = (\omega, \zeta) : \Omega \rightarrow \Theta$, $\Omega \subset \mathbb{R}^{n+1}$, wobei $\omega \in C^1(\Omega)$ die erste und $\zeta \in C^1(\Omega, \mathbb{R}^n)$ die letzten n Komponenten von ϕ bezeichnen. Sei $x^0 = \phi^{-1}(y^0)$. Wegen $\text{rg } D\zeta(x^0) = n$ können wir o.B.d.A. annehmen, dass $D_2\zeta(x^0), \dots, D_{n+1}\zeta(x^0)$ linear unabhängig sind. Schreiben wir $x = (t, v) \in \Omega$ mit $t \in \mathbb{R}$, $v \in \mathbb{R}^n$, so erfüllt ζ die Voraussetzungen des Satzes über implizite Funktionen, Satz 11.5.2. Wir finden also Umgebungen $U = U(t^0)$, $V = V(v^0)$ mit $U \times V \subset \Omega$ und $Z = Z(z^0)$ sowie eine Funktion $\gamma = \gamma(t, z) \in C^1(U \times Z, \mathbb{R}^n)$, so dass $\det D_v \zeta \neq 0$ auf $U \times V$ gilt und

$$\{(t, v) \in U \times V : \zeta(t, v) = z\} = \{(t, \gamma(t, z)) : t \in U\} \quad \text{für alle } z \in Z$$

erfüllt ist. Wir setzen nun

$$\psi := (\text{Id}, \gamma) : U \times Z \rightarrow \mathbb{R}^{1+n}, \quad \chi := (\omega \circ \psi, \text{Id}) : U \times Z \rightarrow \mathbb{R}^{1+n}.$$

Wegen $\psi(U \times Z) \subset U \times V \subset \Omega$ ist χ wohldefiniert und wir bemerken

$$\phi \circ \psi = (\omega \circ \psi, \zeta \circ \psi) = (\omega \circ \psi, \text{Id}) = \chi \quad \text{in } U \times Z.$$

Insbesondere ist also $\chi(U \times Z) \subset \Theta$. Wegen $\zeta \circ \psi = \text{Id}$ ist weiter $(D_v \zeta \circ \psi) D_z \gamma = \mathbb{E}$ auf $U \times Z$. Wir folgern daher

$$J_{\psi} = \det D_z \gamma \neq 0, \quad J_{\chi} = J_{\phi} J_{\psi} \neq 0 \quad \text{in } U \times Z,$$

d.h. ψ und χ bilden $U \times Z$ diffeomorph auf ihr Bild ab; siehe Folgerung 11.4.8.

Wir setzen noch $\Delta := U \times Z$, $\Omega' := \psi(U \times Z) \subset \Omega$ und $\Theta' := \chi(U \times Z) \subset \Theta$. Sei $Q = I \times S \subset \Delta$ ein um (t^0, z^0) zentrierter Würfel und $K := \chi(Q)$. Dann gibt es ein $\varrho > 0$ mit $B_{\varrho}(y^0) \subset K$; man beachte $\chi(t^0, z^0) = (\omega \circ \psi(t^0, z^0), z^0) = (s^0, z^0) = y^0$. Für beliebiges $f \in C_c^0(B_{\varrho}(y^0))$ gilt also gemäß Hilfssatz A.3.3

$$\int_{\Theta'} f dy = \int_{\Delta} (f \circ \chi) |J_{\chi}| d(t, z).$$

Sei weiter $K' := \psi(Q) = \phi^{-1}(K)$. Für $g := (f \circ \phi) |J_{\phi}| \in C_c^0(\Omega')$ gilt dann $\text{supp } g \subset \phi^{-1}(B_{\varrho}(y^0)) \subset \phi^{-1}(K) = K'$. Weiter ist $\psi = (\text{Id}, \gamma) : \Delta \rightarrow \Omega'$ von der in Hilfssatz A.3.4 betrachteten Form. Wählen wir $\underline{\Delta} \subset S$ mit $I \times \underline{\Delta} \subset \Delta$ wie im dortigen Teil (a), so wird $\underline{\Delta} \subset \mathbb{R}^n$ von $\gamma(t, \cdot)$ diffeomorph auf $\underline{\Omega}(t) = \gamma(t, \underline{\Delta}) \subset \mathbb{R}^n$ abgebildet. Nach Induktionsvoraussetzung gilt also insbesondere für $g(t, \cdot) \in C_c^0(\underline{\Omega}(t))$ die Transformationsformel

$$\int_{\underline{\Omega}(t)} g(t, v) dv = \int_{\underline{\Delta}} g(t, \gamma(t, z)) |\det D\gamma_z(t, z)| dz, \quad t \in I.$$

Hilfssatz A.3.4 (b) liefert somit

$$\int_{\Omega'} g dx = \int_{\Delta} (g \circ \psi) |J_{\psi}| d(t, z).$$

Folglich sind alle Voraussetzungen von Hilfssatz A.3.5 erfüllt, wir erhalten für beliebiges $f \in C_c^0(B_{\varrho}(y^0))$:

$$\int_{\Theta} f dy = \int_{\Theta'} f dy \stackrel{\text{Hilfssatz A.3.5}}{=} \int_{\Omega'} (f \circ \phi) |J_{\phi}| dx = \int_{\Omega} (f \circ \phi) |J_{\phi}| dx.$$

Damit ist der Induktionsschritt bewiesen. □

A.4. Uneigentliche Integrale im \mathbb{R}^n

Durch einen Ausschöpfungsprozess wollen wir nun (ähnlich wie im Eindimensionalen in Abschnitt 9.6) uneigentliche Integrale erklären, auf die wir die Transformationsformel im nächsten Abschnitt ausdehnen wollen. Wir erinnern an den Begriff der *quadrierbaren Ausschöpfung* aus Definition A.2.1 und an Satz A.2.3.

Definition A.4.1: Uneigentliches Integral

Seien $\Omega \subset \mathbb{R}^n$ offen und $f \in C^0(\Omega, \mathbb{R}^d)$ gegeben. Wenn für jede quadrierbare Ausschöpfung

$\{M_j\}_j$ von Ω die Folge $\left\{\int_{M_j} f \, dx\right\}_j$ konvergiert, so setzen wir

$$\int_{\Omega} f(x) \, dx := \lim_{j \rightarrow \infty} \int_{M_j} f(x) \, dx \quad (\text{A.4.1})$$

für das *uneigentliche Integral* von f über Ω . Wir sagen dann auch, f ist (*uneigentlich*) *integrierbar* über Ω oder das uneigentliche Integral $\int_{\Omega} f \, dx$ *existiert* oder *konvergiert*. Schließlich heißt $\int_{\Omega} f \, dx$ *absolut konvergent*, wenn das Integral $\int_{\Omega} \|f\| \, dx$ konvergiert. Wir schreiben dann auch $\int_{\Omega} \|f\| \, dx < +\infty$.

Bemerkung A.4.2. (a) Die Definition ist unabhängig von der Wahl der quadrierbaren Ausschöpfung $\{M_j\}_j$ von Ω .

Denn: Ist $\{M'_j\}_j$ eine weitere quadrierbare Ausschöpfung, so ist offenbar auch die gemischte Folge $\{\tilde{M}_j\}_j := \{M_1, M'_1, M_2, M'_2, \dots\}$ quadrierbare Ausschöpfung von Ω . Ist f über Ω uneigentlich integrierbar, so existiert also auch der Grenzwert $\lim_{j \rightarrow \infty} \int_{\tilde{M}_j} f \, dx$ und es gilt insbesondere

$$\lim_{j \rightarrow \infty} \int_{M_j} f(x) \, dx = \lim_{j \rightarrow \infty} \int_{\tilde{M}_j} f(x) \, dx = \lim_{j \rightarrow \infty} \int_{M'_j} f(x) \, dx,$$

wie behauptet. \square

(b) Die Linearität und Monotonie (für $d = 1$) aus Satz 12.5.11 (a), (b) übertragen sich sofort auf das uneigentliche Integral aufgrund der entsprechenden Grenzwertregeln. Die übrigen Abschätzungen aus Satz 12.5.12 (c)-(e) bleiben hingegen nur dann richtig, wenn man die Existenz aller auftretenden Integral voraussetzt.

Hilfssatz A.4.3

Seien $\Omega \subset \mathbb{R}^n$ offen und quadrierbar und $f \in C^0(\Omega, \mathbb{R}^d)$ beschränkt. Dann stimmt das in (A.4.1) erklärte uneigentliche Integral von f über Ω mit dem Riemannschen Integral $\int_{\Omega} f \, dx$ aus Definition 12.5.7 überein.

Beweis. Sei $\varepsilon > 0$ beliebig fixiert. Wegen $|\partial\Omega| = 0$ existieren Quader Q_1, \dots, Q_p mit $\partial\Omega \subset \bigcup_{l=1}^p \mathring{Q}_l =: \Theta$ und $|\Theta| \leq \sum_{l=1}^p |Q_l| < \varepsilon$. Wir setzen $K := \Omega \setminus \Theta = \overline{\Omega} \setminus \Theta$. Dann ist $K \subset \Omega$ kompakt, quadrierbar und es gilt $\Omega \setminus K = \Omega \cap \Theta \subset \Theta$ und daher

$$|\Omega \setminus K| \leq |\Theta| \leq \sum_{p=1}^l |Q_l| < \varepsilon.$$

Ist nun $\{M_j\}_j$ eine beliebige quadrierbare Ausschöpfung von Ω , so gilt $K \subset M_j$ und folglich $|\Omega \setminus M_j| \leq |\Omega \setminus K| < \varepsilon$ für $j \geq j_0(\varepsilon)$. Wir erhalten

$$\left\| \int_{\Omega} f \, dx - \int_{M_j} f \, dx \right\| = \left\| \int_{\Omega \setminus M_j} f \, dx \right\| \leq \left(\sup_{\Omega} \|f\| \right) |\Omega \setminus M_j| \leq \left(\sup_{\Omega} \|f\| \right) \varepsilon, \quad j \geq j_0,$$

also $\int_{\Omega} f \, dx = \lim_{j \rightarrow \infty} \int_{M_j} f \, dx$, wie behauptet. \square

Es gibt eine Reihe von Aussagen zur Additivität des uneigentlichen Integrals, wir wollen hier nur ein für uns relevantes Ergebnis festhalten:

Hilfssatz A.4.4

Ist $K \subset \Omega$ kompakt und quadrierbar und $f \in C^0(\Omega, \mathbb{R}^d)$ sowohl über Ω als auch über $\Omega \setminus K$ uneigentlich integrierbar, so gilt

$$\int_{\Omega} f \, dx = \int_{\Omega \setminus K} f \, dx + \int_K f \, dx.$$

Beweis. Ist $\{M_j\}_j$ quadrierbare Ausschöpfung von Ω , so ist $\{M_j \setminus K\}_j$ quadrierbare Ausschöpfung von $\Omega \setminus K$. Wegen $K \subset M_j$ für $j \geq j_0$ mit einem $j_0 = j_0(K) \in \mathbb{N}$ gilt weiter

$$\int_{M_j} f \, dx = \int_{M_j \setminus K} f \, dx + \int_K f \, dx \quad \text{für } j \geq j_0.$$

Grenzübergang $j \rightarrow \infty$ liefert also die Behauptung. \square

Hilfssatz A.4.5

Falls das uneigentliche Integral $\int_{\Omega} f \, dx$ für ein $f \in C^0(\Omega, \mathbb{R}^d)$ absolut konvergiert, so konvergiert es auch im gewöhnlichen Sinn.

Beweis. Sei zunächst $\{M'_l\}_l$ eine *monotone* quadrierbare Ausschöpfung von Ω , siehe Bemerkung A.2.4 (a). Nach Voraussetzung ist $\{\int_{M'_l} \|f\| \, dx\}_l$ konvergent und daher auch Cauchyfolge. Zu beliebigem $\varepsilon > 0$ existiert somit ein $L = L(\varepsilon) \in \mathbb{N}$, so dass gilt

$$\int_{M'_l \setminus M'_L} \|f\| \, dx = \int_{M'_l} \|f\| \, dx - \int_{M'_L} \|f\| \, dx < \frac{\varepsilon}{2} \quad \text{für alle } l \geq L. \quad (\text{A.4.2})$$

Sei nun $\{M_j\}_j$ eine beliebige quadrierbare Ausschöpfung von Ω . Dann existiert ein $J = J(\varepsilon) \in \mathbb{N}$, so dass

$$M'_L \subset \subset M_j \quad \text{für alle } j \geq J$$

erfüllt ist. Weiter existiert zu beliebigen $j, k \geq J$ ein $K = K(\varepsilon) \geq L$, so dass gilt

$$M_j, M_k \subset\subset M'_K.$$

Wir erhalten somit

$$\begin{aligned} \left\| \int_{M_j} f \, dx - \int_{M_k} f \, dx \right\| &= \left\| \int_{M_j \setminus M'_L} f \, dx - \int_{M_k \setminus M'_L} f \, dx \right\| \\ &\leq \int_{M_j \setminus M'_L} \|f\| \, dx + \int_{M_k \setminus M'_L} \|f\| \, dx \\ &\leq 2 \int_{M'_K \setminus M'_L} \|f\| \, dx \stackrel{(A.4.2)}{<} \varepsilon \quad \text{für alle } j, k \geq J. \end{aligned}$$

Folglich ist $\{\int_{M_j} f \, dx\}_j$ Cauchyfolge und somit konvergent, wie behauptet. \square

Hilfssatz A.4.6: Charakterisierung absoluter Konvergenz

Für beliebiges $f \in C^0(\Omega, \mathbb{R}^d)$ ist $\int_{\Omega} f \, dx$ genau dann absolut konvergent, wenn eine Konstante $c \in [0, +\infty)$ so existiert, dass gilt

$$\int_M \|f\| \, dx \leq c \quad \text{für alle quadrierbaren } M \subset\subset \Omega. \quad (A.4.3)$$

Beweis. ▶ „ \Rightarrow “: Sei $\int_{\Omega} \|f\| \, dx$ konvergent, so setzen wir $c := \int_{\Omega} \|f\| \, dx$. Ist dann $M \subset\subset \Omega$ quadrierbar und $\{M_j\}_j$ quadrierbare Ausschöpfung von Ω , so existiert ein $j_0 = j_0(M) \in \mathbb{N}$ mit $M \subset\subset M_j$ für alle $j \geq j_0$. Es folgt

$$\int_M \|f\| \, dx \leq \int_{M_j} \|f\| \, dx \quad \text{für alle } j \geq j_0,$$

Grenzübergang $j \rightarrow \infty$ liefert also (A.4.3).

▶ „ \Leftarrow “: Ist andererseits (A.4.3) erfüllt, so gilt dies insbesondere für die Elemente $M_j \subset\subset \Omega$ einer monotonen quadrierbaren Ausschöpfung $\{M_j\}_j$ von Ω . Die Folge $\{\int_{M_j} \|f\| \, dx\}_j$ ist dann monoton wachsend und nach oben durch c beschränkt, also auch konvergent. Wie im Beweis von Hilfssatz A.4.5 sieht man, dass daraus die Konvergenz von $\{\int_{M_j} \|f\| \, dx\}_j$ für jede quadrierbare Ausschöpfung $\{M_j\}_j$ von Ω folgt. \square

Folgerung A.4.7

Für $f \in C^0(\Omega, \mathbb{R}^d)$ konvergiere $\int_{\Omega} f \, dx$ absolut. Dann konvergiert auch $\int_{\Omega'} f \, dx$ für jedes $\Omega' \subset \Omega$ absolut (und somit auch im gewöhnlichen Sinne).

Beweis. Sofort aus Hilfssatz A.4.6, denn (A.4.3) gilt insbesondere auch für jedes quadrierbare $M \subset \subset \Omega'$. \square

Haben wir eine Folge $f_j : \Omega \rightarrow \mathbb{R}^d$ stetiger Funktionen, für die die uneigentlichen Integrale über Ω existieren, so fragen wir wieder nach der Vertauschbarkeit von Grenzwertbildung und Integration. Der dafür angemessene Konvergenzbegriff ist der folgende:

Definition A.4.8

Eine Folge $f_j : \Omega \rightarrow \mathbb{R}^d \in C^0(\Omega, \mathbb{R}^d)$, $j = 1, 2, \dots$, heißt *kompakt gleichmäßig konvergent*, wenn für jede kompakte Teilmenge $K \subset \Omega$ die Einschränkungen $\{f_j|_K\}_j$ gleichmäßig konvergieren.

Bemerkung A.4.9. Insbesondere existiert dann eine Grenzfunktion $f(x) := \lim_{j \rightarrow \infty} f_j(x)$, $x \in \Omega$, und nach dem Weierstraßschen Konvergenzsatz gilt $f \in C^0(\Omega, \mathbb{R}^d)$.

Satz A.4.10

Sei $f_j \in C^0(\Omega, \mathbb{R}^d)$, $j = 1, 2, \dots$, kompakt gleichmäßig konvergent gegen $f \in C^0(\Omega, \mathbb{R}^d)$. Weiter existiere eine *integrable Majorante* für $\{f_j\}_j$, d.h. es gibt ein nichtnegatives $F \in C^0(\Omega)$ mit $\int_{\Omega} F \, dx < +\infty$, so dass gilt

$$\|f_j(x)\| \leq F(x) \quad \text{für alle } x \in \Omega \text{ und } j \in \mathbb{N}. \quad (\text{A.4.4})$$

Dann konvergieren auch die uneigentlichen Integrale $\int_{\Omega} f_j \, dx$ für $j = 1, 2, \dots$ und $\int_{\Omega} f \, dx$ (sogar absolut), und es gilt

$$\int_{\Omega} \left(\lim_{j \rightarrow \infty} f_j \right) dx = \int_{\Omega} f \, dx = \lim_{j \rightarrow \infty} \left(\int_{\Omega} f_j \, dx \right).$$

Beweis. Zunächst existiert nach Hilfssatz A.4.6 ein $c \in [0, +\infty)$, so dass gilt

$$\int_M \|f_j\| \, dx \stackrel{(\text{A.4.4})}{\leq} \int_M F \, dx \leq c \quad \text{für alle } M \subset \subset \Omega \text{ und } j \in \mathbb{N}.$$

Wieder nach Hilfssatz A.4.6 ist also $\int_{\Omega} f_j dx$ für alle $j \in \mathbb{N}$ absolut konvergent. Ferner liefert Grenzübergang $j \rightarrow \infty$ in (A.4.4) $\|f\| \leq F$ auf Ω , so dass analog die absolute Konvergenz von $\int_{\Omega} f dx$ folgt. Zu beliebigem Kompaktum $K \subset \Omega$ konvergieren dann auch $\int_{\Omega \setminus K} F dx$, $\int_{\Omega \setminus K} f dx$ und $\int_{\Omega \setminus K} f_j dx$ für $j \in \mathbb{N}$ absolut.

Nun wählen wir zu vorgegebenem $\varepsilon > 0$ eine kompakte, quadrierbare Menge $K \subset \Omega$ so, dass gilt

$$\int_{\Omega \setminus K} F(x) dx = \left| \int_{\Omega \setminus K} F(x) dx \right| \stackrel{\text{Hilfssatz A.4.4}}{=} \left| \int_{\Omega} F(x) dx - \int_K F(x) dx \right| < \varepsilon;$$

man kann $K = M_j$ z.B. als Element der quadrierbaren Ausschöpfung $\{M_j\}_j$ aus Satz A.2.3 wählen. Dann folgt aus der Monotonie des Integrals auch

$$\int_{\Omega \setminus K} \|f_j(x)\| dx < \varepsilon \quad \text{für } j = 1, 2, \dots, \quad \int_{\Omega \setminus K} \|f(x)\| dx < \varepsilon. \quad (\text{A.4.5})$$

Da andererseits $f_j \rightrightarrows f$ auf $K \subset \Omega$ gilt, existiert nach Satz 12.5.12 ein $j_0 = j_0(\varepsilon) \in \mathbb{N}$ mit

$$\left\| \int_K f_j dx - \int_K f dx \right\| < \varepsilon \quad \text{für alle } j \geq j_0. \quad (\text{A.4.6})$$

Kombination von (A.4.5) und (A.4.6) liefert schließlich

$$\begin{aligned} \left\| \int_{\Omega} f_j dx - \int_{\Omega} f dx \right\| &\stackrel{\text{Hilfssatz A.4.6}}{\leq} \left\| \int_{\Omega \setminus K} (f_j - f) dx \right\| + \left\| \int_K (f_j - f) dx \right\| \\ &\leq \int_{\Omega \setminus K} \|f_j\| dx + \int_{\Omega \setminus K} \|f\| dx + \left\| \int_K (f_j - f) dx \right\| \\ &< 3\varepsilon, \quad j \geq j_0, \end{aligned}$$

also die Behauptung. □

A.5. Die allgemeine Transformationsformel

Unser Hauptziel dieses Anhangs ist nun der Beweis des folgenden Satzes:

Satz A.5.1: Transformationsformel

Es seien $\Omega, \Theta \subset \mathbb{R}^n$ offene Mengen und $\phi = \phi(x) : \Omega \rightarrow \Theta$ ein C^1 -Diffeomorphismus. Dann

gilt für jedes $f \in C^0(\Theta)$ mit $\int_{\Theta} |f| dy < +\infty$ die Identität

$$\int_{\Theta} f(y) dy = \int_{\Omega} f(\phi(x)) |J_{\phi}(x)| dx. \quad (\text{A.5.1})$$

Bemerkung A.5.2. Durch komponentenweise Betrachtung überträgt sich (A.5.1) offenbar sofort auf komplex- bzw. vektorwertige Funktionen f .

Zur Vorbereitung benötigen wir noch den einfachen

Hilfssatz A.5.3

Zu jeder offenen Menge $\Omega \subset \mathbb{R}^n$ und $f \in C_c^0(\Omega)$ existiert eine offene, quadrierbare Menge $\Omega' \subset \Omega$ mit $\text{supp } f \subset \Omega'$. Das uneigentliche Integral $\int_{\Omega} f dx$ existiert, und es gilt

$$\int_{\Omega} f dx = \int_{\Omega'} f dx.$$

Beweis. Wir setzen $K := \text{supp } f \subset \Omega$. Ist $\{M_j\}_j$ eine quadrierbare Ausschöpfung, so existiert gemäß Bemerkung A.2.4 (b) ein $j_0 \in \mathbb{N}$ mit $K \subset \overset{\circ}{M}_j$ für alle $j \geq j_0$. Setzen wir z.B. $\Omega' := \overset{\circ}{M}_{j_0} \subset \Omega$, so haben wir also $\text{supp } f \subset \Omega'$ und beachten

$$\int_{M_j} f dx = \int_{M_j \setminus K} f dx + \int_K f dx = \int_K f dx = \int_{M_{j_0}} f dx = \int_{\Omega'} f dx \quad \text{für alle } j \geq j_0.$$

Grenzübergang $j \rightarrow \infty$ liefert die behauptete Integralidentität. \square

Folgerung A.5.4

Die Transformationsformel für Testfunktionen, Satz A.3.1 bleibt für offene, nicht notwendig quadrierbare Mengen $\Omega, \Theta \subset \mathbb{R}^n$ richtig.

Beweis. Mit Hilfssatz A.5.3 wählen wir zu $f \in C_c^0(\Theta)$ eine offene, quadrierbare Menge $\Theta' \subset \Theta$ mit $\text{supp } f \subset \Theta'$ und setzen $\Omega' := \phi^{-1}(\Theta')$. Dann ist $\phi|_{\Omega'} : \Omega' \rightarrow \Theta'$ diffeomorph und $f \in C_c^0(\Theta')$, $(f \circ \phi)|J_{\phi}| \in C_c^0(\Omega')$ erfüllt. Wenden wir Satz A.3.1 auf den Mengen Ω', Θ' an, so liefert schließlich Hilfssatz A.5.3:

$$\int_{\Theta} f(y) dy = \int_{\Theta'} f(y) dy = \int_{\Omega'} f(\phi(x)) |J_{\phi}(x)| dx = \int_{\Omega} f(\phi(x)) |J_{\phi}(x)| dx,$$

wie behauptet. \square

Beweis von Satz A.5.1. Wir wählen zu Ω eine quadrierbare Ausschöpfung $\{M_j\}_j$ mit $M_j \subset\subset \Omega$ für alle $j \in \mathbb{N}$; siehe Satz A.2.3. Dann ist gemäß Folgerung A.2.6 auch $\{N_j\}_j$ mit $N_j := \phi(M_j)$ eine quadrierbare Ausschöpfung von $\Theta = \phi(\Omega)$ mit $N_j \subset\subset \Theta$ für alle $j \in \mathbb{N}$. Mit Folgerung A.1.6 finden wir zu jedem $N_j \subset\subset \Theta$ ein $\eta_j \in C_c^\infty(\Theta)$ mit $\eta_j = 1$ auf N_j und $\eta_j(\Theta) \subset [0, 1]$. Die Funktionen $f_j := f\eta_j \in C_c^0(\Theta)$ sind dann über Θ uneigentlich integrierbar gemäß Hilfssatz A.5.3, und Folgerung A.5.4 liefert

$$\int_{\Theta} f_j dy = \int_{\Omega} (f_j \circ \phi) |J_\phi| dx, \quad j \in \mathbb{N}. \quad (\text{A.5.2})$$

Die gleiche Relation gilt offenbar für $|f_j|$, $j \in \mathbb{N}$. Ferner konvergiert $\{f_j\}_j$ kompakt gleichmäßig gegen f auf Θ , und es gilt $|f_j| \leq |f|$ auf Θ sowie $\int_{\Theta} |f| dy < +\infty$. Satz A.4.10 liefert also

$$\lim_{j \rightarrow \infty} \int_{\Theta} f_j dy = \int_{\Theta} f dy. \quad (\text{A.5.3})$$

Analog konvergiert $\{(f_j \circ \phi) |J_\phi|\}_j$ kompakt gleichmäßig auf Ω gegen $(f \circ \phi) |J_\phi|$. Ist schließlich $M \subset\subset \Omega$ quadrierbar, so existiert ein $j \in \mathbb{N}$ mit $M \subset M_j$ und es folgt:

$$\begin{aligned} 0 &\leq \int_M |f \circ \phi| |J_\phi| dx \stackrel{f_j = f\eta_j}{=} \int_M |f_j \circ \phi| |J_\phi| dx \\ &\stackrel{\text{Hilfssatz A.4.4}}{\leq} \int_{\Omega} |f_j \circ \phi| |J_\phi| dx \stackrel{(\text{A.5.2}) \text{ für } |f_j|}{=} \int_{\Theta} |f_j| dy \leq \int_{\Theta} |f| dy < +\infty. \end{aligned}$$

Nach Hilfssatz A.4.6 konvergiert also das uneigentliche Integral $\int_{\Omega} (f \circ \phi) |J_\phi| dx$ absolut und Satz A.4.10 liefert

$$\int_{\Omega} (f \circ \phi) |J_\phi| dx = \lim_{j \rightarrow \infty} \left(\int_{\Omega} (f_j \circ \phi) |J_\phi| dx \right). \quad (\text{A.5.4})$$

Aus (A.5.2)-(A.5.4) ergibt sich schließlich die behauptete Transformationsformel (A.5.1). \square

Beispiel A.5.5. (a) *Polarkoordinaten:* Es sei $f \in C^0(B_R)$ mit $\int_{B_R} |f| dy < +\infty$ auf der Kreisscheibe $B_R := B_R(0, 0) \subset \mathbb{R}^2$ gegeben. Dann gilt

$$\int_{B_R} f(y) dy = \int_0^R \left(\int_0^{2\pi} f(r \cos \theta, r \sin \theta) r d\theta \right) dr. \quad (\text{A.5.5})$$

Denn: Gemäß Aufgabe 1 (a) von Übungsblatt 11 ist die Abbildung $\phi : (0, R) \times (0, 2\pi) \rightarrow B_R \setminus ([0, R) \times \{0\})$ mit

$$\phi = \phi(r, \theta) := (r \cos \theta, r \sin \theta), \quad (r, \theta) \in (0, R) \times (0, 2\pi),$$

ein Diffeomorphismus mit $J_\phi = r$ auf $(0, R) \times (0, 2\pi)$. Da B_R quadrierbar ist und $[0, R) \times \{0\}$ den Inhalt Null hat, entnehmen wir der Transformationsformel

$$\int_{B_R} f \, dy = \int_{B_R \setminus ([0, R) \times \{0\})} f \, dy = \int_{(0, R) \times (0, 2\pi)} (f \circ \phi) |J_\phi| \, d(r, \theta).$$

Iterierte Integration liefert also [A.5.5](#). \square

Ist speziell f radialsymmetrisch, d.h. $f(y) = g(|y|)$ mit einem $g = g(r) \in C^0([0, R))$, so folgt

$$\int_{B_R} f(y) \, dy = \int_0^R \left(\int_0^{2\pi} g(r)r \, d\theta \right) dr = 2\pi \int_0^R g(r)r \, dr. \quad (\text{A.5.6})$$

Noch spezieller, für $f \equiv 1$ auf K_R , erhalten wir somit

$$|B_R| = \int_{B_R} 1 \, dx = 2\pi \int_0^R r \, dr = \pi R^2.$$

(b) *GAUSSSches Fehlerintegral*: Wir behaupten

$$\int_{-\infty}^{+\infty} e^{-x^2} \, dx = \sqrt{\pi}.$$

Denn: In Beispiel [9.6.13](#) (b) haben wir bereits gesehen, dass das GAUSSSche Fehlerintegral existiert. Zur Berechnung setzen wir $W_R := [-R, R] \times [-R, R]$ für beliebiges $R > 0$ und beachten

$$\int_{W_R} e^{-(y_1^2 + y_2^2)} \, dy \stackrel{\text{Satz 12.2.5}}{=} \left[\int_{-R}^R e^{-x^2} \, dx \right]^2 \leq c < +\infty \quad (\text{A.5.7})$$

mit einer von $R > 0$ unabhängigen Konstanten $c \in [0, +\infty)$. Ist $M \subset\subset \mathbb{R}^2$ quadrierbar, so existiert ein $R > 0$ mit $M \subset W_R$, so dass nach [\(A.5.7\)](#) und Hilfssatz [A.4.6](#) das uneigentliche Integral $\int_{\mathbb{R}^2} e^{-(y_1^2 + y_2^2)} \, dy$ existiert. Da $\{W_j\}_{j=1,2,\dots}$ quadrierbare Ausschöpfung des \mathbb{R}^2 ist, erhalten wir

$$\int_{\mathbb{R}^2} e^{-(y_1^2 + y_2^2)} \, dy = \lim_{j \rightarrow \infty} \int_{W_j} e^{-(y_1^2 + y_2^2)} \, dy = \left[\int_{-\infty}^{+\infty} e^{-x^2} \, dx \right]^2. \quad (\text{A.5.8})$$

Andererseits ist auch $\{B_j\}_{j=1,2,\dots}$ mit den Kreisscheiben $B_j = B_j(0, 0)$ eine quadrierbare

Ausschöpfung von \mathbb{R}^2 . Formeln (A.5.6) und (A.5.8) liefern also

$$\begin{aligned}
\int_{-\infty}^{+\infty} e^{-x^2} dx &= \left[\int_{\mathbb{R}^2} e^{-(y_1^2+y_2^2)} dy \right]^{\frac{1}{2}} = \left[\lim_{j \rightarrow \infty} \int_{B_j} e^{-(y_1^2+y_2^2)} dy \right]^{\frac{1}{2}} \\
&= \left[\pi \lim_{j \rightarrow \infty} \int_0^j e^{-r^2} (2r) dr \right]^{\frac{1}{2}} = \sqrt{\pi} \left[- \lim_{j \rightarrow \infty} \int_0^j \frac{d}{dr} (e^{-r^2}) dr \right]^{\frac{1}{2}} \\
&= \sqrt{\pi} \left[\lim_{j \rightarrow \infty} (1 - e^{-j^2}) \right]^{\frac{1}{2}} = \sqrt{\pi},
\end{aligned}$$

wie behauptet. □

Literaturverzeichnis

- [Dei] O. DEISER: *Reelle Zahlen*. Springer-Verlag, 2008.
- [DeRa] R. DENK, R. RACKE: *Kompendium der Analysis*. Vieweg+Teubner Verlag, 2011.
- [Fo1] O. FORSTER: *Analysis 1*. Springer Spektrum, 2016.
- [Fo2] O. FORSTER: *Analysis 2*. Springer Spektrum, 2017.
- [Hie1] M. HIEBER: *Analysis I*. Springer Spektrum, 2018.
- [Hie2] M. HIEBER: *Analysis II*. Springer Spektrum, 2019.
- [Hil1] S. HILDEBRANDT: *Analysis 1*. Springer-Verlag, 2006.
- [Hil2] S. HILDEBRANDT: *Analysis 2*. Springer-Verlag, 2003.
- [Kö1] K. KÖNIGSBERGER: *Analysis 1*. Springer-Verlag, 2004.
- [Kö2] K. KÖNIGSBERGER: *Analysis 2*. Springer-Verlag, 2004.
- [Pö1] J. PÖSCHEL: *Etwas Analysis*. Springer Spektrum, 2014.
- [Pö2] J. PÖSCHEL: *Etwas mehr Analysis*. Springer Spektrum, 2014.
- [Sa] F. SAUVIGNY: *Analysis*. Springer Spektrum, 2014.
- [Wa1] W. WALTER: *Analysis 1*. Springer-Verlag, 1997.
- [Wa2] W. WALTER: *Analysis 2*. Springer-Verlag, 2002.

Index

Überdeckung

- abzählbare, 256
- endliche, 256
- offene, 256

Lagrangsche Multiplikatoren, 245

Laplace-Operator, 222

Rayleigh-Quotient, 228

Abbildung, 20

- bijektive, 23
- identische, 20
- injektive, 23
- offene, 237
- reguläre, 237
- surjektive, 23
- umkehrbare, 24

Ableitung, 133

- gewöhnliche, 133
- höherer Ordnung, 143, 219
- Linearität der, 136
- partielle, 208
 - höherer Ordnung, 219
- Richtungs-, 213
- totale, 216

Absolutbetrag, 36

Additionstheoreme

- für \sin , \cos , 157
- für \tan , \cot , 164

Allquantor, 10

Anordnungsaxiome, 33

äquivalent, 19

Äquivalenz

- logische, 6

Äquivalenzklasse, 19

Äquivalenzrelation, 19

Archimedisches Axiom, 35

Arcus Cosinus, 165

Arcus Cotangens, 165

Arcus Sinus, 165

Arcus Tangens, 165

- Reihendarstellung, 203

Assoziativität, 21

Aussage, 5

Aussageform, 10

Ausschöpfung

- quadrierbare, 275, 282

BERNOULLISCHE Ungleichung, 37

Betrag

- in \mathbb{R} , 36
- in angeordnetem Körper, 36
- komplexer Zahlen, 45

Betragsfunktion, 46

Bild, Bildmenge, 22

Bildpunkt, 21

Binomialkoeffizient, 40

- allgemeiner, 202
- Binomialreihe, 202
- BINOMIScher Lehrsatz, 41
- CAUCHY-SCHWARZsche Ungleichung, 97
- CAUCHYfolge, 56
 - im \mathbb{R}^d , 99
- CAUCHYsche Produktformel, 89
 - für Potenzreihen, 93
- CAUCHYsches Konvergenzkriterium, 57
 - für Funktionenfolgen, 128
 - im \mathbb{R}^d , 100
- Cosinus Hyperbolicus, 166
- Cosinusfunktion, 156
- Cotangens, 163
- d -Tupel, 18
- Definitionsbereich, 21
- Dezimalbruch
 - endlicher, 52
 - unendlicher, 52
- Diffeomorphismus, 233
- Differential, 216
 - s-ter Ordnung, 223
- Differentialquotient, 133
- Differenz von Mengen, 14
- Direkter Beweis, 9
- DIRICHLETSche Sprungfunktion, 114
- disjunkt, 14
- Distributivgesetze
 - für Mengen, 15
- Divergenz eines Vektorfeldes, 211
- Doppelkomplementgesetz, 15
- Durchschnitt
 - von Mengenfamilien, 17
 - von zwei Mengen, 14
- e -Funktion, 153
- Einschränkung, 20
- Element
 - einer Menge, 11
 - negatives, 33
 - positives, 33
- scshape Eulersche Formel, 157
- EULERSche Zahl e , 152
- Existenzquantor, 10
- Exponentialfunktion
 - allgemeine, 156
 - Funktionalgleichung der, 152
 - komplexe, 130
 - reelle, 153
- Exponentialreihe, 86, 92, 130
- Extremalbedingung
 - hinreichende, 146
 - im \mathbb{R}^n , 233
 - notwendige 1. Ordnung, 139
 - im \mathbb{R}^n , 227
 - notwendige 2. Ordnung, 146
 - im \mathbb{R}^n , 230
- Folge, 51
 - beschränkte, 53
 - im \mathbb{R}^d , 99
 - bestimmt divergente, 55
 - CAUCHY-, 56
 - im \mathbb{R}^d , 99
 - divergente, 52
 - im \mathbb{R}^d , 99
 - Glied einer, 51
 - konstante, 52
 - konvergente, 52
 - im \mathbb{R}^d , 99
 - monotone, 64
 - Null-, 53
 - im \mathbb{R}^d , 99
 - Produkt-, 52

-
- rekursiv definierte, 65
 - streng monotone, 64
 - unbeschränkte, 54
 - von oben beschränkt, 54
 - von unten beschränkt, 54
 - Folgerung, 6
 - Fundamentalsatz der Algebra, 123
 - Funktion, 20
 - beschränkte, 120
 - bestimmt divergente, 112
 - bijektive, 23
 - charakteristische, 20
 - differenzierbare, 133
 - divergente, 110
 - gleichmäßig stetige, 121
 - injektive, 23
 - integrierbare, 173, 250, 264, 268
 - komplexwertig, 173
 - vektorwertig, 173
 - k -mal differenzierbare, 143
 - k -mal stetig differenzierbare, 144
 - konkave, 147
 - konvergente, 110
 - konvexe, 147
 - monotone, 119
 - Oszillation einer, 260
 - partiell differenzierbare, 208
 - Potenz-, 119
 - RIEMANN-integrierbare, 172, 250, 264
 - stückweise stetige, 182
 - stetig differenzierbar, 208
 - stetige, 115
 - stetige (in einem Punkt), 112
 - streng monotone, 119
 - strikt konkave, 147
 - strikt konvexe, 147
 - surjektive, 23
 - total differenzierbare, 215
 - umkehrbare, 24
 - uneigentlich integrierbar, 191
 - uneigentlich integrierbare, 195, 275, 291
 - unendlich oft differenzierbare, 144
 - unstetige, 113
 - Funktionaldeterminante, 210
 - Funktionalmatrix, 210
 - Funktionenfolge
 - gleichmäßig konvergente, 127
 - kompakt gleichmäßig konvergente, 276, 294
 - punktweise konvergente, 127
 - Funktionenreihe
 - gleichmäßig konvergente, 129
 - punktweise konvergente, 129
 - Funktionsgrenzwert, 110
 - links- bzw. rechtsseitiger, 116
 - Funktionswert, 21
 - Funkzion
 - mit kompaktem Träger, 272, 280
 - GAUSSsche Summenformel, 39
 - GAUSSsches Fehlerintegral, 278, 298
 - GAUSSsches Fehlerintegral, 195
 - Gebiet, 213
 - geordnetes Paar, 17
 - Gradient, 210
 - Graph, 22
 - Grenzwert
 - einer Folge, 52
 - im \mathbb{R}^d , 99
 - einer Funktion, 110
 - im Unendlichen, 117
 - links- bzw. rechtsseitiger, 116
 - punktweiser einer Funktionenfolge, 127
 - Häufungspunkt, 105
-

- Häufungswert, 70
 - Folgen im \mathbb{R}^d , 99
- HADAMARDS Lemma, 215
- HEINE-BOREL-Eigenschaft, 256
- HESSE-Matrix, 219
- imaginäre Achse, 43
- imaginäre Einheit, 44
- Imaginärteil, 44
- Implikation, 6
- Indirekter Beweis, 9
- Infimum, 67
- innerer Punkt, 105
- Integral, 173, 251
 - bestimmtes, 185
 - unbestimmtes, 185
 - uneigentliches, 191, 195, 275, 291
 - absolut konvergentes, 192, 275, 291
 - bestimmt divergentes, 191
 - divergentes, 191
 - konvergentes, 191, 291
- Intervall
 - abgeschlossenes, 62
 - halboffenes, 62
 - Länge, Durchmesser, 62
 - offenes, 62
- Intervallschachtelungsprinzip, 62
- isolierter Punkt, 105
- JACOBI-determinante, 210
- JACOBI-matrix, 210
- JORDAN-Maß, 263
- JORDAN-menge, 262
- Körper, 30
 - angeordneter, 33
 - archimedisch angeordneter, 35
 - bewerteter, 45
 - vollständig (archimedisch) angeordnet, 57
 - vollständig bewerteter, 57
- Körperaxiome, 30
- kanonische Fortsetzung, 264
- kartesisches Produkt, 17
- Kettenregel, 137
- Kettenschluss, 9
- Kommutativität, 21
- kompakt enthalten, 272, 279
- Kompaktum, 104
- Komplement, 14
 - relatives, 14
- komplexe Addition, 42
- komplexe Multiplikation, 42
- Komposition, 22
- Kontradiktion, 7
- Kontrapositionsregel, 9
- Konvergenzradius, 92
- Kreiszahl π , 158
- kritischer Punkt, 139
 - im \mathbb{R}^n , 228
- L'HOSPITALSche Regel, 142
- LAGRANGESche Restgliedformel, 198
- LEBESGUESches Integrabilitätskriterium, 265
- LEIBNIZ-Kriterium, 87
- LEIBNIZ-Regel, 145
- Limes, 52
 - einer Folge, 52
 - im \mathbb{R}^d , 99
 - einer Funktion, 110
 - im Unendlichen, 117
 - links- bzw. rechtsseitiger, 116
 - punktweiser einer Funktionenfolge, 127
- Limes inferior, 70
- Limes superior, 70
 - Charakterisierung, 71

-
- Logarithmus, 154
 - Logarithmusfunktion, 154
 - Funktionalgleichung der, 154
 - Reihendarstellung, 202
 - Mächtigkeit einer endlichen Menge, 13
 - Majorante, 83
 - für Funktionenreihen, 129
 - integrable, 194, 294
 - Majorantenkriterium, 83
 - für Funktionenreihen, 129
 - für uneigentliche Integrale, 194
 - Mannigfaltigkeit, 242
 - Codimension einer, 242
 - Dimension einer, 242
 - Normalenvektor einer, 243
 - Normalraum einer, 243
 - Tangentenvektor einer, 243
 - Tangentialraum einer, 243
 - Matrix
 - indefinite, 230
 - negativ semidefinite, 230
 - negative definite, 230
 - positiv definite, 230
 - positiv semidefinit, 230
 - Maximum
 - einer Menge, 69
 - globales, 138
 - lokales, 138
 - striktes globales, 138
 - striktes lokales, 138
 - Menge, 11
 - abgeschlossene, 102
 - Abschluss einer, 105
 - abzählbar unendliche, 75
 - abzählbare, 75
 - beschränkte, 66
 - im \mathbb{R}^d , 104
 - dünne, 259
 - Durchmesser einer, 104
 - endliche, 75
 - folgenkompakte, 104
 - Häufungspunkt einer, 105
 - Inhalt einer, 263
 - innerer Punkt einer, 105
 - Inneres einer, 105
 - isolierter Punkt einer, 105
 - JORDAN-messbare, 262
 - kompakte, 104
 - mit Inhalt Null, 257
 - mit Maß Null, 258
 - nach oben beschränkte, 66
 - nach unten beschränkte, 66
 - offene, 101
 - quadrierbare, 262
 - Rand einer, 105
 - Randpunkt einer, 105
 - Volumen einer, 263
 - zusammenhängende, 212
 - Minimum
 - einer Menge, 69
 - globales, 138
 - lokales, 138
 - striktes globales, 138
 - striktes lokales, 138
 - Mittelwertsatz, 140
 - der Integralrechnung, 179
 - in mehreren Veränderlichen, 212
 - Multiindex, 222
 - Multindex
 - Länge eines, 222
 - Nabla-Operator, 209
 - Negation, 6
 - Niveaumenge, 241
 - Norm, 98
-

- Normalenvektor, 243
- Normalraum, 243
- Nullfolge, 53
 - im \mathbb{R}^d , 99
- Nullmenge, 258
- obere Schranke, 66
 - kleinste, 67
- Oberintegral, 172, 250
- Obersumme, 170, 249
- ODER-Verknüpfung, 6
- Operation, 21
- Oszillation, 260
- Partialsumme, 79
- Partielle Ableitung, 208
 - höherer Ordnung, 219
- Partielle Integration, 187
- Permutation, 89
- Polarkoordinaten, 277, 297
- Potenz
 - allgemeine, 155
- Potenz, q -te, 66
- Potenzfunktion, 119, 155
- Potenzmenge, 13
- Potenzreihe, 91
 - Koeffizienten einer, 91
 - Konvergenzgebiet einer, 92
 - Konvergenzradius einer, 92
- Produkt
 - inneres, 96
- Produkt \prod , 39
- Produktregel, 136
- Quader im \mathbb{R}^n , 248
 - Volumen, Inhalt eines, 248
- Quadrierbarkeitskriterium, 263
- Quotientenkriterium, 84
- Quotientenregel, 136
- Randpunkt, 105
- Realteil, 44
- reelle Achse, 43
- Regeln von de Morgan
 - für Mengenfamilien, 17
 - für zwei Mengen, 15
- regulärer Punkt, 237
- Reihe, 79
 - LEIBNIZ-, 87
 - absolut konvergente, 83
 - alternierende harmonische, 87
 - bedingt konvergente, 90
 - beschränkte, 87
 - bestimmt divergente, 80
 - divergente, 80
 - Exponential-, 86, 92
 - geometrische, 80, 91
 - Summenformel der, 40
 - harmonische, 82
 - konvergente, 79
 - Minorante einer, 84
 - Potenz-, 91
 - Umordnung einer, 89
 - unbedingt konvergente, 90
 - Wert der, 79
- Reihenglieder, 79
- Relation, 18
- Restglied, 197
 - im \mathbb{R}^n , 224
- Restklassenmenge, 19
- Richtungsableitung, 213
- RIEMANN-Integral, 172, 251, 264
 - bestimmtes, 172
- RIEMANNsche Zwischensumme, 170, 249
- RIEMANNsches Integralkriterium, 196
- Rotation eines Vektorfeldes, 211

- Sandwich-Lemma, 61
- Satz
- von der iterierten Integration, 253, 254
 - von der monotonen Konvergenz, 65
 - VON FERMAT, 139
 - von Fubini, 267
 - VON HEINE-BOREL, 256
 - VON HEINE, 122
 - VON MOIVRE, 162
 - VON ROLLE, 140
 - VON SCHWARZ, 221
 - VON BOLZANO-WEIERSTRASS, 63
 - VON BOLZANO-WEIERSTRASS
 - im \mathbb{R}^d , 101
 - VON CAUCHY-HADAMARD, 92
- SCHWARZsche Ungleichung, 178, 252, 266
- Signumfunktion, 114
- singulärer Punkt, 237
- Sinus Hyperbolicus, 166
- Sinusfunktion, 156
- Skalarprodukt, 96
- Stammfunktion, 183
- stationärer Punkt, 139
 - im \mathbb{R}^n , 228
- Substitutionsregel, 188
- Summe Σ , 39
- Support einer Funktion, 272, 279
- Supremum, 67
- Tangens, 163
- Tangente, 134
- Tangentenvektor, 243
- Tangentialraum, 243
- Tautologie, 7
- TAYLORpolynom, 197
 - im \mathbb{R}^n , 224
- TAYLORreihe, 201
- TAYLORSche Formel, 197
 - im \mathbb{R}^n , 224
- Teilfolge, 63
- Teilmenge, 13
 - beschränkte, 66
 - im \mathbb{R}^d , 104
 - dichte, 73
 - Durchmesser einer, 104
 - echte, 13
 - folgenkompakte, 104
 - kompakte, 104
 - nach oben beschränkte, 66
 - nach unten beschränkte, 66
- Testfunktion, 272, 280
- Totale Ableitung, 216
- Träger einer Funktion, 272, 279
- Transformationsformel, 188
 - allgemeine, 274, 295
 - für Testfunktionen, 271, 285
- Umgebung, 236
 - r -, 53
- Umkehrabbildung, 24
- Umkehrfunktion, 24
- Umordnungssatz
 - DIRICHLETScher, 90
 - RIEMANNScher, 90
- UND-Verknüpfung, 6
- Unendliche Reihe, 79
- untere Schranke, 66
 - größte, 67
- Unterintegral, 172, 250
- Untersumme, 170, 249
- Urbild, 22
- VENN-Diagramm, 14
- Vereinigung
 - von Mengenfamilien, 17
 - von zwei Mengen, 14

- Verschmelzungsgesetze, 15
- Vollständige Induktion, 37
- Vollständigkeit
 - von \mathbb{C} , 59
- Vollständigkeitsaxiom, 57

- Wahrheitstafel, 6
- Wahrheitswert, 5
- WEIERSTRASSscher Konvergenzsat, 128
- WEIERSTRASSscher Hauptsatz, 121
- Wertebereich, 21
- Widerspruchsbeweis, 9
- Wurzel
 - n -te komplexe, 163
- Wurzel, s -te reelle, 66
- Wurzelkriterium, 86

- Zahl
 - konjugiert komplexe, 45
- Zahlen, 30
 - ganze \mathbb{Z} , 27
 - irrationale, 29
 - komplexe \mathbb{C} , 42
 - natürliche \mathbb{N} , 27
 - rationale \mathbb{Q} , 28
 - reelle \mathbb{R} , 29
- Zahlenfolge, 51
 - komplexe, 51
 - rationale, 51
 - reelle, 51
- Zerlegung, 170
 - eines Quaders, 248
 - Feinheit einer, 170, 248
 - gemeinsame Verfeinerung, 171, 249
 - Teilpunkte einer, 170
 - Verfeinerung einer, 171, 249
- Zerlegung der Eins, 272, 280
 - endliche, 273, 280
- Zerlegungsfolge, 175
 - ausgezeichnete, 175, 251
- Zwischenwertsatz, 118