Skript Analysis I - III

Wintersemester 2001/2002 - Wintersemester 2002/2003

TECHNISCHE UNIVERSITÄT
CHEMNITZ

LaTeX
Jens Rckert 2004

Vorlesungsskript zu den Grundvorlesungen Analysis I - III

Vorlesender: Prof. Dr. Schneider

TU-Chemnitz
Fakultät für Mathematik
Wintersemester 2002/2003

GeTEXt von Jens Rückert ..... jens.rueckert@mathematik.tu-chemnitz.de

Überarbeitet und gesetzt von Dr. Helmut Harbrecht .....
..... helmut.harbrecht@mathematik.tu-chemnitz.de

Inhaltsverzeichnis

1 Grundlagen **9**

1.1 Elementare Aussagenlogik 9

1.1.1 Logische Operation 9

1.1.2 Quantoren 12

1.2 Mengenlehre 12

1.3 Funktionen 16

2 Reelle Zahlen **23**

2.1 Algebraische Grundbegriffe und Körper 23

2.2 Geordnete Mengen 28

2.3 Geordnete Körper 29

2.4 Natürliche Zahlen 30

2.5 Komplexe Zahlen 35

2.6 Euklidische Räume 39

2.7 Zusatz: Dedekindsche Schnitte 41

3 Einführung in die Topologie **45**

3.1 Metrische Räume 45

3.2 Kompakte Mengen 50

3.3 Perfekte Mengen 56

3.4 Zusatz: Relative Offenheit und Kompaktheit 56

4 Folgen und Reihen **59**

4.1 Allgemeine Zahlenfolgen 59

4.2 Reelle Zahlenfolgen 61

4.3 Cauchy–Folgen 66

4.4 Reihen 69

4.5 Das Wurzel– und Quotientenkriterium 73

4.6 Alternierende Reihen 74

4.7 Summen und Produkte von Reihen 75

4.8 Umordnen von Reihen 77

4.9 Potenzreihen 78

5 Stetige Funktionen **81**

5.1 Funktionsgrenzwerte 81

5.2 Stetigkeit von Funktionen 82

5.3 Einseitiger Grenzwert und Stetigkeit 84

5.4 Zusammenhängende Mengen und Stetigkeit 85

5.5 Stetigkeit und Kompaktheit 86

6 Differentiation von Funktionen einer Veränderlichen **89**

6.1 Differenzenquotient und Differentialquotient 89

6.2 Mittelwertsätze 93

6.3 Die l'Hospitalsche Regel 95

7 Integration von Funktionen **99**

7.1 Das unbestimmte Integral 99

7.2 Integration gebrochen rationaler Funktionen 102

7.3 Das Riemann-Stieltjes-Integral 104

7.4 Eigenschaften des Riemann-Stieltjes-Integral 112

7.5 Integration und Differentiation 117

7.6 Rektifizierbare Kurven 121
7.7 Uneigentliche Integrale 123

8 Folgen und Reihen von Funktionen 125

8.1 Gleichmäßige Konvergenz 125
8.2 Potenzreihen 134
8.3 Fourier-Reihen 145
8.4 Gleichgradige Stetigkeit 155
8.5 Der Approximationssatz von Weierstraß 158

9 Mehrdimensionale Differentiation 161

9.1 (Partielle) Ableitungen und Richtungsableitungen 161
9.2 Höhere (partielle) Ableitungen 171
9.3 Nichtlineare Gleichungen 176
9.4 Extremalwertprobleme 184

10 Integration im $\mathbb{R}^n$ 195

10.1 Grundlagen 195
10.2 Die Substitutionsregel im $\mathbb{R}^n$ 202

11 Differentialformen 211

11.1 Grundlegende Definitionen 211
11.2 Volumen- und Oberflächenintegrale 212
11.3 Differentialformenkalkül 220
11.3.1 Elementare Eigenschaften 220
11.3.2 Normaldarstellungen 221
11.3.3 Äußere Multiplikation 222
11.3.4 Äußere Differentiation 224
11.3.5 Substitution 225
11.4 Simplizes und Ketten 230

11.5 Der Satz von Stokes 235
11.6 Anwendungen des Satzes von Stokes 239

12 Maßtheorie **245**
12.1 Ringe und Algebren von Mengen 245
12.2 Mengenfunktionen 248
12.3 Inhalt, Maß und Konstuktion von Maßen 253
12.4 Messbare Funktionen 265
12.4.1 Definitionen und Eigenschaften 265
12.4.2 Konvergenzsätze 270
12.5 Das Lebesgue-Integral 272
12.6 Grenzwertsätze für Integrale 277
12.7 Vergleich des Lebesgue-Integrals mit dem Riemann-Integral 280
12.8 L^p-Räume 283
12.9 Integration in Produkträumen 287
12.9.1 Produktmaße 287
12.10 Zusatz: Zerlegung von Mengenfunktionen 293

Einleitung

Gegenstände der Analysis sind im wesentlichen Funktionen. Eine Funktion ist ein mathematisches Konstrukt, das zu einem gegebenen *Input* einen eindeutig definierten *Output* liefert:

$$\text{Input} \xrightarrow{\text{Funktion}} \text{Output}$$

(wir werden das später noch genauer definieren. Normalerweise sind die Inputdaten Zahlen x und zu jeder dieser Zahlen x liefert eine Funktion f eine eindeutige Zahl $f(x)$. Daher verlangt die *Analysis* zuerst

- als Fundament die Einführung von Zahlen
- sowie eine klare Definition des Konzeptes einer Funktion.

Die natürlichen Zahlen gehören, abgesehen von zwei- und dreidimensionalen geometrischen Gebilden, im Grunde zu den ganz wenigen Objekten der Mathematik, die unserer allgemeinen und natürlichen Anschauung zugänglich sind.

Um aber in Bereiche vorzustoßen, die unserer Anschauung schwer zugänglich sind, bedürfen wir einer besonders strengen Methodik. Ausgehend von ganz wenigen Gesetzmässigkeiten, sogenannte *Axiomen*, die wir postulieren müssen, versuchen wir durch ganz strenge *Logik* in Dimensionen vorzudringen, die sich unserer Vorstellung vollkommen entziehen.

$$\boxed{\textbf{Axiome + Definitionen + Voraussetzungen}}$$
$$\Big\Downarrow \text{(Beweis)}$$
$$\boxed{\textbf{Lemma} \text{ (Hilfssatz) + } \textbf{Proposition} \text{ (kleiner Satz, Bemerkung)}}$$
$$\Big\Downarrow \text{(Beweis)}$$
$$\boxed{\textbf{Theorem} \text{ (Satz)}}$$
$$\Big\Downarrow \text{(Beweis)}$$
$$\boxed{\textbf{Korollar} \text{ (Folgerung)}}$$

Diese ideale Vorstellung, dass wir von wenigen Axiome ausgehen und daraus mittels logischen Schlußfolgerungen alle weiteren Aussagen beweisen entspricht aber auch nicht ganz der Realität. In Wirklichkeit gleicht die Mathematik mehr einem Baum (Russel): *oberhalb der Erde entfaltet sich die Baumkrone als ein verästeltes und kompliziertes Gebilde, (das wir studieren wollen). Wenn man allerdings nach den Grundlagen forscht, so stößt man unter der Erde ebenfalls auf ein filigranes Wurzelwerk.*

Wir können in einer einführenden Veranstaltung die Grundlegung der Mathematik lediglich bis zu einer gewissen Stufe betreiben. Wir werden beginnen mit der *Aussagenlogik*, um uns einen strengen und formalen Rahmen zu schaffen. Die *Mengenlehre* erlaubt uns die Einbeziehung vielfältigster Objekte und ermöglicht eine große Allgemeinheit, die uns von dem engen Rahmen der Zahlen befreit. Den Begriff einer *Funktion* werden wir mengentheoretisch definieren. Wesentlich ist, dass sie uns dabei im Umgang mit dem so strapazierten Begriff des *Unendlichen* enorm hilfreich ist. Die Begriffe wie *Approximation* und *Konvergenz* sind von geradezu fundamentaler Bedeutung für die Analysis. Hierzu behandeln wir ausführlich das Kapitel über *Metriken*, bevor wir uns konkret *Folgen* und *stetigen Funktionen einer rellen Variablen* zuwenden.

Zitat (Jospeh Fourier: Die analytische Theorie der Wärme, 1807). *Die Analysis ist genauso allumfassend wie die Natur selbst. Sie betrifft unsere sämtlichen Sinneserscheinungen, die Zeit ebenso wie die Längen, Kräfte oder Temperaturen. Dies ist eine schwierige Wissenschaftsdisziplin, die nur langsam Fortschritte macht, bei der aber andererseits jedes einmal gewonnene Grundprinzip Bestand hat.*

Das Hauptmerkmal der Analysis ist ihre Klarheit; für nebulöse Fromulierungen läßt diese Sprache keinen Raum. Sie verknüpft höchst unterschiedliche Erscheinungen und zeigt dadurch gemeinsame Analogien auf. Die Analysis kann sogar Erscheinungen beschreiben, die wie Luft oder Raum flüchtig sind, oder von Körpern, die fern im Weltraum schweben; sie kann helfen, Himmelserscheinungen zu verschiedenen, oft viele Jahrhunderte getrennten Epochen zu erkennen, oder auch die Schwerkraft und oder Wärme tief im Erdinneren, die uns niemals zugänglich sind. All diese Erscheinungen macht die Analysis erfassbar und messbar. Sie ist wohl eine Fähigkeit des menschlichen Geistes, die diesem mitgegeben wurde, um sein kurzes Leben und seine unvollkommene Sinneswahrnehmung zu ergänzen. Noch bemerkenswerter ist, dass bei der Untersuchung sämtlicher Erscheinungen prinzipiell das gleiche Vorgehen möglich ist. Sie alle werden in der gleichen Sprache beschrieben, als wollte die Analysis die Einheit und Einfachheit des Weltplanes aufzeigen ....

Kapitel 1

Grundlagen

1.1 Elementare Aussagenlogik

Eine Aussage im Sinne der Aussagenlogik ist eine sprachliche Aussage, bei der klar entschieden werden kann, ob sie wahr oder falsch ist. Für die Aussagenlogik spielt der Inhalt keine Rolle sondern lediglich der Wahrheitswert "wahr" oder "falsch".

Beispiel 1.1.1.

1. *1 plus 1 ist 2.*
2. *1 plus 1 ist 1.*
3. *1 plus 1 ist nicht gleich 0.*

Keine Aussagen sind dagegen Sätze wie "morgen regnet es", "Gehen Sie bitte zur Seite!", oder auch widersprüchliche Sätze, sogenannte Antinomien, wie z.B. "Minos, der König von Kreta, sagt, dass alle Kreter lügen".

1.1.1 Logische Operation

Negation: Sei A eine Aussage, so heißt $\neg A$ ("nicht" A) die Negation von A. Der Sachverhalt wird durch eine Wahrheitstafel dargestellt:

A	$\neg A$
w ("wahr")	f
f ("falsch")	w

Verknüpfung "und": Sind A und B Aussagen, so ist "A und B" (mathematisch $A \wedge B$) ebenfalls eine Aussage. Wir definieren: $A \wedge B$, gelesen A und B, ist genau dann wahr, falls die Aussagen A und B beide wahr sind. Wahrheitstafel:

A	B	$A \wedge B$
w	w	w
w	f	f
f	w	f
f	f	f

Verknüpfung "oder": Sind A und B Aussagen, so ist $A \vee B$, gelesen "A oder B", ebenfalls eine Aussage. Diese ist genau dann wahr, wenn mindestens eine der beiden Aussagen A oder B wahr ist. Wahrheitstafel:

A	B	$A \vee B$
w	w	w
w	f	w
f	w	w
f	f	f

Äquivalenz: Zwei Aussagen A und B heißen äquivalent, wenn der Wahrheitswert übereinstimmt. Wir schreiben in diesem Fall $A \Leftrightarrow B$. Wahrheitstafel:

A	B	$A \Leftrightarrow B$
w	w	w
w	f	f
f	w	f
f	f	w

Satz 1.1.2. *Seien A, B und C Aussagen. Dann gelten die folgenden Äquivalenzen:*

1. *Doppelte Negation* $\quad \neg(\neg A) \Leftrightarrow A$
2. *Kommutativgesetze* $\quad A \wedge B \Leftrightarrow B \wedge A$
   $\quad A \vee B \Leftrightarrow B \vee A$
3. *Assoziativgesetze* $\quad A \wedge (B \wedge C) \Leftrightarrow (A \wedge B) \wedge C$
   $\quad A \vee (B \vee C) \Leftrightarrow (A \vee B) \vee C$
4. *Distributivgesetze* $\quad A \wedge (B \vee C) \Leftrightarrow (A \wedge B) \vee (A \wedge C)$
   $\quad A \vee (B \wedge C) \Leftrightarrow (A \vee B) \wedge (A \vee C)$
5. *De Morgansche Regeln* $\quad \neg(A \wedge B) \Leftrightarrow (\neg A) \vee (\neg B)$
   $\quad \neg(A \vee B) \Leftrightarrow (\neg A) \wedge (\neg B)$

Beweis. Man beweist alle Aussagen mittels Wahrheitstafeln, z.B. gilt bzgl. der Negation

A	$\neg A$	$\neg(\neg A)$
w	f	w
f	w	f

Rest als Übung. □

Bemerkung 1.1.3. *Man vergleiche die Regeln (2) bis (4) mit analogen Regeln der Arithmetik* $(+, \cdot)$.

Implikation: Sind A und B Aussagen, so ist $A \Rightarrow B$, gelesen "A impliziert B", äquivalent zur Aussage $B \vee (\neg A)$. Wahrheitstafel:

A	B	$A \Rightarrow B$
w	w	w
w	f	f
f	w	w
f	f	w

Satz 1.1.4. *Seien A, B und C Aussagen. Dann gilt:*

$$(A \Rightarrow B) \wedge (B \Rightarrow A) \quad \Leftrightarrow \quad (A \Leftrightarrow B)$$

Beweis. Übung. □

Bemerkung 1.1.5.

1. *Eine Aussage $A \Leftrightarrow B$ beweist man mit Hilfe von Satz 1.1.4. Man zeigt zuerst A impliziert B und dann B impliziert A.*
2. *Falls A falsch ist, so ist $A \Rightarrow B$ immer wahr ("e falso quod libet").*

Satz 1.1.6. *Für beliebige Aussagen A und B gilt:*

1.	*bejahende Abtrennungsregel*	$(A \wedge (A \Rightarrow B))$	$\Rightarrow$	B
2.	*verneinende Abtrennungsregel*	$(\neg B \wedge (A \Rightarrow B))$	$\Rightarrow$	$\neg A$
3.	*Regel vom indirekten Beweis*	$(A \Rightarrow B)$	$\Leftrightarrow$	$(\neg B \Rightarrow \neg A)$
4.	*Regel vom Kettenschluss*	$(A \Rightarrow B) \wedge (B \Rightarrow C)$	$\Rightarrow$	$(A \Rightarrow C)$

Beweis. Die erste Formel folgt aus

$$\begin{aligned}(A \wedge (A \Rightarrow B)) &\Leftrightarrow A \wedge (\neg A \vee B)\\ &\Leftrightarrow ((A \wedge \neg A) \vee (A \wedge B))\\ &\Leftrightarrow (f \vee (A \wedge B))\\ &\Leftrightarrow A \wedge B \Rightarrow B.\end{aligned}$$

Rest als Übung. □

1.1.2 Quantoren

Sei A_j, $j \in J$, eine Menge von Aussagen, z.B. A_1, $A_2, \ldots$. Dann definieren wir $\forall_{j\in J}\ A_j$, gelesen "für alle A_j mit j aus J", als genau dann wahr, falls jede Aussage A_j, $j \in J$, wahr ist. Falls $J = \{1, \ldots, N\}$ endlich ist, dann können wir diese Aussage durch endlich viele "und"-Verknüpfungen ausdrücken

$$\forall_{j\in J}\ A_j \Leftrightarrow \forall_{j=1}^{N}\ A_j \Leftrightarrow (A_1 \wedge A_2 \wedge \ldots \wedge A_N).$$

"$\forall$" heißt der *Allquantor*.

Demgegenüber definieren wir $\exists_{j\in J}\ A_j$, gelesen "es existiert ein A_j mit j aus J", als genau dann wahr, falls eine wahre Aussage A_j für mindestens ein $j \in J$ existiert. Für endlich viele A_j, d.h. $J = \{1, \ldots, N\}$, gilt

$$\exists_{j\in J}\ A_j \Leftrightarrow \exists_{j=1}^{N}\ A_j \Leftrightarrow A_1 \vee A_2 \vee \ldots \vee A_N.$$

"$\exists$" heißt der *Existenzquantor*.

Beachte: Die Reihenfolge der Quantoren darf im allgemeinen nicht vertauscht werden, d.h.

$$(\forall_{j\in J}\ \exists_{i\in I}\ A_{i,j}) \not\Leftrightarrow (\exists_{i\in I}\ \forall_{j\in J}\ A_{i,j}).$$

Die erste Aussage heißt "für alle $j \in J$ gibt es ein $i \in I$ mit $A_{i,j}$" während die zweite Aussage lautet "es existiert ein $i \in I$ für das für alle $j \in J$ gilt $A_{i,j}$".

Beispiel 1.1.7. *Man vergleiche die Aussagen "für alle Studenten gibt es einen Studentenausweis" und "es gibt ein Einwohnermeldeamt für alle Bewohner der Stadt".*

Satz 1.1.8. *Seien A_j, $j \in J$, Aussagen. Dann gelten die de Morganschen Regeln*

$$\begin{aligned} \neg(\forall_{j\in J}\ A_j) &\iff (\exists_{j\in J}\ \neg A_j), \\ \neg(\exists_{j\in J}\ A_j) &\iff (\forall_{j\in J}\ \neg A_j). \end{aligned}$$

1.2 Mengenlehre

Neben der Methodik der Logik ist die Sprache der Mengenlehre für die moderne Mathematik unerläßlich. Den Begriff einer Menge werden wir uns ganz naiv gegenüberstellen.

Eine Menge ist eine Zusammenfassung von Objekten unserer Anschauung.

Diese Objekte nennen wir die *Elemente* der Menge. Dies bedeutet, dass wir eine Menge genau dann kennen, wenn uns gesagt wird, welche Elemente sie enthält. Zum Beispiel sind $M = \{a, b, 10\}$ oder auch $A = \{1, 2, -1\}$ Mengen. Dabei ist zu beachten, dass die Reihenfolge, in welcher die Elemente aufgezählt werden, unerheblich ist.

Zumeist sind Mengen über Aussageformen $A(x)$ definiert, wie z.B. $A(x)$: "x ist grün". Wir schreiben kurz

$$M = \{x : A(x)\}$$

für die Menge aller x für die die Aussage $A(x)$ wahr ist.

Gleichheit von Mengen: Wir schreiben $x \in M$ falls x ein Element von M ist. Für $\neg(x \in M)$ schreiben wir $x \notin M$. Zwei Mengen M und N sind genau dann gleich, wenn sie die gleichen Elemente haben.

Teilmengen: Einge Menge N heißt eine *Teilmenge* von einer Menge M, kurz $N \subseteq M$, falls aus $x \in N$ stets folgt $x \in M$, kurz

$$N \subseteq M \Longleftrightarrow \forall_{x \in N}\, x \in M.$$

Wir schreiben für $\neg(A \subseteq B)$ abkürzend $A \not\subseteq B$.

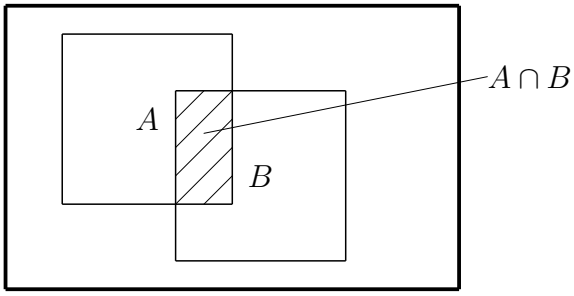
Satz 1.2.1 (Transsitivität von "$\subseteq$"). *Seien A, B und C Mengen mit $A \subseteq B$ und $B \subseteq C$, so gilt auch $A \subseteq C$.*

Beweis. Sei $x \in A$ beliebig, dann ist $x \in B$ wegen $A \subseteq B$. Wegen $x \in B$ und $B \subseteq C$ gilt auch $x \in C$. □

Durchschnitt von Mengen: Die Menge

$$A \cap B = \{x : (x \in A) \wedge (x \in B)\}$$

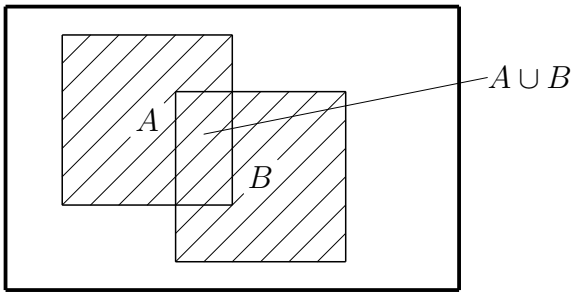
heißt die *Schnittmenge* von A und B oder einfach "A geschnitten B". Eine anschauliche Darstellung dieses Begriffes liefert das sogenannte "Venn–Diagramm".



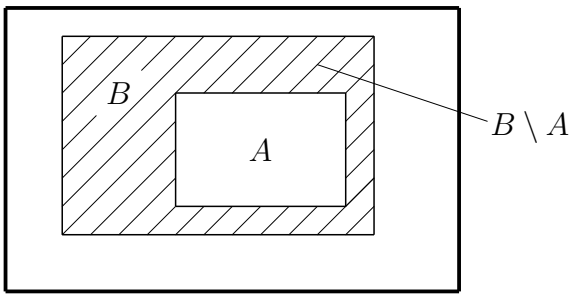
Vereinigung von Mengen: Seien A und B Mengen. Dann ist

$$A \cup B = \{x : (x \in A) \vee (x \in B)\}$$

die *Vereinigungsmenge* von A und B, oder kurz "A vereinigt B".



Komplementärmengen: Sei A und B Mengen, so ist die Menge $\{x \in B : x \notin A\}$ ebenfalls eine Menge, die wir mit $B \setminus A$, gelesen "B ohne A" oder auch "B minus A", bezeichen. Falls $A \subseteq B$, so heißt $B \setminus A$ das *Komplement* von A in der Menge B.



Leere Menge: Sei $F(x)$ eine Aussageform, die für jedes $x \in A$ falsch ist. Dann enthält die Menge $M := \{x \in A : F(x)\}$ kein einziges Element. Eine solche Menge, die keine Elemente enthält, heißt *leer*.

Bemerkung 1.2.2. *Sei $\emptyset$ eine leere Menge.*

1. *Die Aussage $x \in \emptyset$ ist immer falsch.*

2. *Es gilt $\emptyset \subseteq A$ für jede Menge A.*

Proposition 1.2.3. *Sind A und B beides leere Mengen, so ist $A = B$. Folglich ist die leere Menge eindeutig. Sie wird mit $\emptyset$ bezeichnet.*

Beweis. Wir zeigen $A \subseteq B$ und $B \subseteq A$. $A \subseteq B$ ist äquivalent zur Aussage $\forall x : x \in A \Rightarrow x \in B$. Da die Aussage $x \in A$ immer falsch ist, ist die obige Implikation richtig. Analog zeigt man $B \subseteq A$. $\square$

Disjunkte Menge: Zwei Mengen A und B heißen *disjunkt*, falls der Durchschnitt $A \cap B = \emptyset$.

Satz 1.2.4. *Seien A, B und C Teilmengen einer Menge X. Dann gelten:*

1.	*Doppeltes Komplement*	$X \setminus (X \setminus A)$	$=$	A
2.	*Kommutativgesetze*	$A \cup B$	$=$	$B \cup A$
		$A \cap B$	$=$	$B \cap A$
3.	*Assoziativgesetze*	$(A \cup B) \cup C$	$=$	$A \cup (B \cup C)$
		$(A \cap B) \cap C$	$=$	$A \cap (B \cap C)$
4.	*Distributivgesetze*	$A \cup (B \cap C)$	$=$	$(A \cup B) \cap (A \cup C)$
		$A \cap (B \cup C)$	$=$	$(A \cap B) \cup (A \cap C)$
5.	*Dominanzgesetze*	$A \cup \emptyset$	$=$	A
		$A \cap \emptyset$	$=$	$\emptyset$
6.	*De Morgansche Regeln*	$X \setminus (A \cap B)$	$=$	$(X \setminus A) \cup (X \setminus B)$
		$X \setminus (A \cup B)$	$=$	$(X \setminus A) \cap (X \setminus B)$

Beweis. Die erste Behauptung beweist man wie folgt. Per Definition ist $X \setminus (X \setminus A) = \{x \in X : x \notin (X \setminus A)\}$. Für $x \in X$ gilt nun $x \in X \setminus A$ genau dann, wenn $x \notin A$. Durch Negation folgt hieraus $x \in A$. Da $A \subseteq X$ ist $X \setminus (X \setminus A) = A$.

Der Beweis der restlichen Behauptungen sei als Übung überlassen. □

Beliebige Vereinigungen und Durchschnitte: Sei $\{A_j : j \in J\}$ eine Familie von Mengen, dann ist

$$\bigcup_{j \in J} A_j := \{x : \exists_{j \in J}\, x \in A_j\}$$

die Vereinigung der Mengen A_j, $j \in J$, bzw.

$$\bigcap_{j \in J} A_j := \{x : \forall_{j \in J}\, x \in A_j\}$$

die *Schnittmenge* aller Mengen A_j, $j \in J$.

Produktmenge: Seien $a \in A$ und $b \in B$, so heißt (a, b) ein *geordnetes Paar*. Dabei beachte man, dass $(a, b) \neq (b, a)$. Die Menge

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

heißt die *Paarmenge* von A und B bzw. das *kartesische Produkt* von A und B. Wir sagen hierfür auch "A kreuz B".

Potenzmenge: Sei A eine Menge, so heißt die Menge aller Teilmengen von A

$$\mathcal{P}(A) = \{B : B \subseteq A\}$$

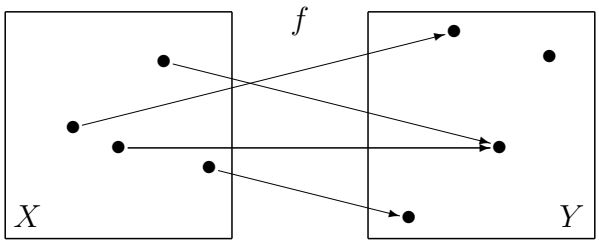
die *Potenzmenge* von A.

Bemerkung 1.2.5. *Es gelten die in der nachstehenden Tabelle aufgeführten Analogien zwischen der Aussagenlogik und der Mengenlehre.*

Aussagenalgebra	Mengenalgebra
Aussagen A, B	Mengen A, B
Verknüpfung $\wedge$	Durchschnitt $\cap$
Verknüpfung $\vee$	Vereinigung $\cup$
Negation $\neg$	Komplement $X \setminus A$
Kontradiktion	Leere Menge $\emptyset$
Tautologie	Grundmenge X
Implikation $A \Rightarrow B$	Teilmenge $A \subseteq B$
Äquivalenz $A \Leftrightarrow B$	Gleichheit $A = B$

1.3 Funktionen

Eine Funktion $f : X \to Y$ ist eine Vorschrift, die jedem Element $x \in X$ einer Menge X ein und auch nur ein Element $f(x) \in Y$ zuordnet.



Definition 1.3.1. *Seien $X, Y \neq \emptyset$ beliebige Mengen und $G \subseteq X \times Y$.*

1. *Das Tripel (G,X,Y) heißt eine* **Relation** *zwischen den Mengen X und Y.*

2. *Eine Relation (G, X, Y) zwischen den Mengen X und Y heißt* **Funktion** *$f : X \to Y$, falls*

   (a) für alle $x \in X$ ein $y = f(x)$ existiert mit $(x, y) \in G$, kurz

   $$\forall_{x \in X}\, \exists_{y \in Y} : (x, y) \in G;$$

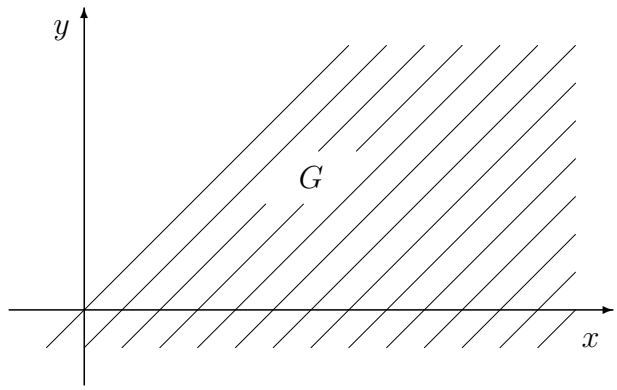
   (b) für alle $x \in X$ und $y, y' \in Y$ aus $(x, y), (x, y') \in G$ folgt, dass $y = y'$ eindeutig ist, kurz

   $$\forall_{x \in X, y, y' \in Y}((x, y) \in G \wedge (x, y') \in G) \Longrightarrow y = y').$$

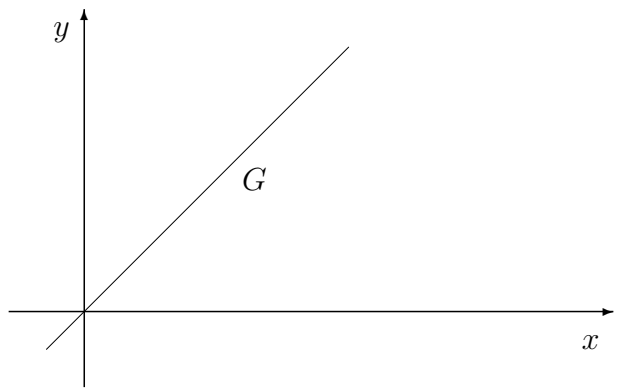
Funktionen nennt man auch oftmals **Abbildungen**.

Beispiel 1.3.2.

1. *Wählen wir $X = Y = \mathbb{N}$, so ist durch $G = \{(x, y) \in X \times Y : x \geq y\}$ eine Relation gegeben.*



2. *Für $X = Y \neq \emptyset$ ist durch $G = \{(x, y) \in X \times Y : x = y\}$ eine Funktion definiert.*



Ist $f = (G, X, Y)$ so schreiben wir $f : X \to Y$ und nennen X den *Definitionsbreich* und Y den *Bildbereich* von f. Die Menge $G = \{(x, y) : y = f(x), x \in X\}$ heißt der *Graph* von f und es gilt $y = f(x) \Leftrightarrow (x, y) \in G$. Zur präzisen Angabe einer Funktion gehört immer auch die Angabe des Definitionsbereiches und des Bildbereiches. Zwei Funktionen $f : X \to Y$ und $g : A \to B$ sind genau dann gleich, falls $A = X$ und $B = Y$ ist und $f(x) = g(x)$ für alle $x \in X$ gilt.

Definition 1.3.3. *Sei A eine Teilmenge des Definitionsbereiches $A \subseteq X$. Dann heißt die Menge $f(A) = \{y \in Y : y = f(x), x \in A\} \subseteq Y$ das Bild von A unter der Funktion f. Sei $B \subseteq Y$, so heißt die Menge*

$$f^{-1}(B) = \{x \in X : f(x) = y, y \in B\} \subseteq X$$

das Urbild der Menge B unter f.

Definition 1.3.4. *Eine Funktion $f : X \to Y$ heißt*

1. **injektiv**, *falls für alle $y \in f(X)$ und $x, x' \in X$ mit $f(x) = f(x') = y$ folgt $x = x'$;*

2. **surjektiv**, *falls $f(X) = Y$ ist;*

3. **bijektiv**, *falls f injektiv und surjektiv ist.*

Sei $f : X \to Y$ eine bijektive Funktion. Dann existiert zu jedem $y \in Y$ genau ein $x \in X$ mit $y = f(x)$. Die Funktion $g : Y \to X$ mit $g(y) = x \Leftrightarrow f(x) = y$ heißt *Umkehrfunktion*. Wir schreiben hierfür $f^{-1} := g : Y \to X$.

Zwei Mengen X und Y haben die gleiche *Kardinalität*, wir schreiben $\mathrm{card}(X) = \mathrm{card}(Y)$, falls eine **bijektive** Abbbildung $f : X \to Y$ zwischen beiden Mengen existiert.

Satz 1.3.5. *Sei $\mathbb{N} = \{1, 2, 3, \ldots\}$ die Menge der natürlichen Zahlen. Die Produktmenge $\mathbb{N}^2 := \mathbb{N} \times \mathbb{N}$ hat die gleiche Kardinalität wie $\mathbb{N}$.*

Beweis. Seien $p, q \in \mathbb{N}$ und somit $(p, q) \in \mathbb{N}^2$ beliebig. Wir definieren das "Cantorsches Diagonalverfahren" mittels der Funktion

$$f(p, q) := \frac{(p + q - 1)(p + q - 2)}{2} + p.$$

Die Namensgebung dieser Funktion ist motiviert durch folgendes Rechteckschema:

	1	2	3	4
1	(1,1)	(1,2)	(1,3)	(1,4)
2	(2,1)	(2,2)	(2,3)	
3	(3,1)	(3,2)		
4	(4,1)			$\ddots$

	1	2	3	4
1	1	2	4	7
2	3	5	8	
3	6	9		
4	10			$\ddots$

Für $n \in \mathbb{N}$ seien $M_n = \{(1, n), (2, n - 1), \ldots, (n - 1, 2), (n, 1)\}$, Teilmengen in $\mathbb{N}^2$. Wir sehen aus dem obigen Schema, dass die Mengen M_n paarweise disjunkt sind und dass $\bigcup_{n\in\mathbb{N}} M_n = \mathbb{N}^2$. Seien ferner $N_n := \{\frac{(n-1)n}{2} + 1, \frac{(n-1)n}{2} + 2, \ldots, \frac{(n-1)n}{2} + n\}$ Teilmengen in $\mathbb{N}$. Aufgrund der Beziehung $\frac{(n-1)n}{2} + n = \frac{n(n+1)}{2}$ ersehen wir, dass die Mengen N_n ebenfalls paarweise disjunkt sind und dass gilt $\bigcup_{n\in\mathbb{N}} N_n = \mathbb{N}$. Wie man unschwer sieht bildet $f : M_n \to N_n$ die beiden Mengen M_n und N_n bijektiv aufeinander ab. □

Definition 1.3.6. *Eine Menge A heißt (höchstens)* **abzählbar**, *falls A endlich ist, d.h. aus endlich viele Elementen besteht, oder aber falls $\mathrm{card}(A) = \mathrm{card}(\mathbb{N})$ gilt.*

Korollar 1.3.7.

1. *Die Menge der ganzen Zahlen*

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

   ist abzählbar.

2. *Die Menge aller positiven Brüche hat die gleiche Mächtigkeit wie $\mathbb{N}$.*

3. *Die Menge der rationalen Zahlen*

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

ist abzählbar.

Bemerkung 1.3.8. *Mengen A und B mit* **endlich** *vielen Elementen haben gleiche Kardinalität genau dann, wenn sie gleichviele Elemente enthalten, d.h.* $\operatorname{card}(A) = \operatorname{card}(B)$.

Die folgende Aussage ist für Mengen mit endlich vielen Elementen trivial, im allgemeinen Fall aber ein tiefliegendes Resultat.

Satz 1.3.9 (Cantor und Russell). *Sei X eine beliebige Menge. Dann existiert*

1. **keine** *surjektive Funktion*
$$f : X \to \mathcal{P}(X);$$

2. **keine** *injektive Funktion*
$$g : \mathcal{P}(X) \to X.$$

Beweis. 1. Sei $f : X \to \mathcal{P}(X)$ eine beliebige Funktion. Wir zeigen, dass es immer ein Element $Y \in \mathcal{P}(X)$ gibt mit $Y \notin f(X)$, d.h. das nicht im Bild von f liegt. Wir bemerken, dass für jedes $x \in X$ das Bild $f(x) \in \mathcal{P}(X)$ eine Teilmenge von X ist, $f(x) \subseteq X$, und definieren die Menge
$$A := \{x \in X : x \notin f(x)\} \in \mathcal{P}(X).$$

Wir führen den Beweis indirekt und dazu nehmen folgendes wir an,

$$\text{es existiert ein } x \in X \text{ mit } f(x) = A.$$

Dann gibt es gibt nur zwei Möglichkeiten für beliebiges $x \in X$:

(a) $x \in f(x)$, dann ist aber $x \notin A$. Somit ist $f(x) \not\subseteq A$, d.h. keine Teilmenge von A, und damit $f(x) \neq A$.

(b) Oder aber $x \notin f(x)$ und somit ist $x \in A$. Aufgrund von $x \notin f(x)$ ist $A \not\subseteq f(x)$ keine Teilmenge von $f(x)$. In diesem Fall ist ebenfalls $f(x) \neq A$.

Somit haben wir die obige Annahme zu einem Widerspruch geführt.

2. Wir beweisen den zweiten Teil ähnlich wie Teil eins. Angenommen, es gebe eine injektive Abbildung $g : \mathcal{P}(X) \to X$. Wir definieren die Menge

$$B := \{g(Y) : Y \in \mathcal{P}(X) \text{und } g(Y) \notin Y\} \subseteq X.$$

Dann gibt es gibt nur zwei Möglichkeiten:

(a) $g(B) \in B$, dann ist existiert $Y \subseteq X$ mit $g(B) = g(Y)$. Andererseits ist aber $g(B) \in B$ und aufgrund der Definition von B auch $B \neq Y$. In diesem Fall liegt ein Widerspruch zur Annahme der Injektivität vor.

(b) $g(B) \notin B$, ist ein Widerspruch in sich, da aus $g(B) \notin B \in \mathcal{P}(X)$ folgt $g(B) \in B$.

□

Korollar 1.3.10. *Die Potenzmenge $\mathcal{P}(\mathbb{N})$ der natürlichen Zahlen ist überabzählbar.*

Satz 1.3.11 (Bernsteinscher Äquivalenzsatz). *Falls die Funktionen $f : X \to Y$ und $g : Y \to X$ beide* **injektiv** *sind, so gilt $\operatorname{card}(X) = \operatorname{card}(Y)$, d.h. es existiert eine bijektive Funktion*

$$F : X \to Y.$$

Beweis. Wir nennen ein $u \in f^{-1}(Y) \subseteq X$ (bzw. $v \in g^{-1}(X) \subseteq X$) einen *Vorgänger* von $y \in Y$ (bzw. von $x \in X$), falls $f(u) = y$ (bzw. $g(v) = x$). Ein Element $a \in X \cup Y$ heißt ein *Vorfahre* von $x \in X \cup Y$, falls es endlich viele Vorgänger a_i, $i = 0, 1, \ldots, n$ gibt mit $a_0 = a$, $a_n = b$ und a_i Vorgänger von a_{i+1} ist.

$X_0 \subseteq X$ sei die Menge der Elemente in X mit einer geraden oder ungeraden Anzahl von Vorfahren. Weiter sei $X_1 \subseteq X$ die Menge aller $x \in X$ mit einer ungeraden Anzahl von Vorfahren und $Y_2 \subseteq Y$ bestehe aus allen $y \in Y$ mit einer geraden Anzahl von Vorfahren. Dann gilt erstens $X_1 \cap X_2 = \emptyset$ und zweitens ist $g : Y_2 \to X_1$ eine bijektive Funktion, d.h. es existiert $\gamma : X_1 \to Y_2$ mit $\gamma\big(g(y)\big) = y$, $y \in Y_2$, und $g\big(\gamma(x)\big) = x$, $x \in X_1$. Wir setzen

$$F(x) = \begin{cases} f(x), & x \in X_0, \\ \gamma(x), & x \in X_1. \end{cases}$$

Die Funktion F ist injektiv und $F(X_0) \subseteq Y$ ist die Menge der Elemente $y \in Y$ mit einer ungeraden Anzahl oder unendlichen Anzahl von Vorfahren. Somit ist Y die disjunkte Vereinigung $Y = Y_2 \cup F(X_0)$. Hierbei beachte man, dass jedes Element $a \in X \cup Y$ mindestens einen Vorfahren hat. □

Sei $f : X \to Y$ eine injektive Funktion, dann existiert $g : Y \to X$ mit $g\big(f(x)\big) = x$ für jedes $x \in X$ (Linksinverse). Somit ist die Funktion $f : X \to f(X)$ bijektiv, d.h. es existiert $g : f(X) \to X$ mit $g\big(f(x)\big) = x$. Falls $Y \setminus f(X) = \emptyset$ haben wir g konstruiert. Falls $Y \setminus f(x) \neq \emptyset$, wählen wir ein beliebiges $a \in X$ und setzen

$$g(y) = \begin{cases} g(y), & y \in f(X), \\ a, & y \in Y \setminus f(X). \end{cases}$$

Es stellt sich die Frage, ob sich nun für jede surjektive Funktion $f : X \to Y$ ebenfalls eine Funktion $g : Y \to X$ mit $f\big(g(y)\big) = y$, $y \in Y$, konstruieren lässt. Da $Y = f(X)$, muss man dazu zu jedem $y \in Y$ in der Urbildmenge der Funktion $f^{-1}(\{y\}) \subseteq X$ ein Element $g(y)$ auswählen.

Die Frage, ob dies auch für unendliche Mengen $f^{-1}(\{y\})$ immer möglich ist, kann leider *nicht* direkt beantwortet werden. Wir werden diese Möglichkeit in dem sogenannten *Auswahlaxiom* postulieren.

Axiom 1.3.12 (Auswahlaxiom). *Zu jeder surjektiven Funktion $f : X \to Y$ existiert eine Funktion $g : Y \to X$ mit $f\big(g(y)\big) = y$ für alle $y \in Y$ (Rechtsinverse).*

Das Auswahlaxiom ist ein wichtiges, aber auch umstrittenes, Hilfsmittel, z.B. in der *Funktionalanalysis* etc. Im Rahmen der Analysis-Grundvorlesungen werden wir es kaum verwenden müssen.

Bemerkung 1.3.13. *Die "naive Mengenlehre" hat ihre Grenzen. Wir betrachten die sogenannte "universelle Menge"*

$$\mathcal{U} = \{X : X \text{ ist eine Menge}\}.$$

Dann ist $\mathcal{P}(\mathcal{U})$ ebenfalls eine Menge. Aber $\mathcal{P}(\mathcal{U}) \subseteq \mathcal{U}$, da $\mathcal{U}$ ja alle Mengen enthält. Es gibt aber immer eine injektive Abbildung $p : \mathcal{U} \to \mathcal{P}(\mathcal{U})$, nämlich $p(U) := \{U\}$ für alle $U \in \mathcal{U}$, und eine ebenfalls injektive Funktion $g : \mathcal{P}(\mathcal{U}) \to \mathcal{U}$, nämlich $g(\mathcal{U}) = \mathcal{U}$ für alle $U \in \mathcal{P}(\mathcal{U})$. Somit ist $\operatorname{card}\mathcal{P}(\mathcal{U}) = \operatorname{card}\mathcal{U}$. Dies steht im Widerspruch zum Satz von Cantor und Russell 1.3.9. Damit sehen wir, dass man den naiv eingeführten Begriff einer Menge, insbesondere Mengen von Mengen, i.a. etwas vorsichtig behandeln muß.

Ein ähnliches Beispiel ist die Russellsche Antinomie: Sei $\mathcal{B}$ die Menge aller Mengen, die sich nicht selbst enthalten. Eine solche Formulierung ist natürlich in sich widersprüchlich.

Die Entdeckung dieser Widersprüche führte zur Entwicklung der "Axiomatischen Mengenlehre", deren eingehende Behandlung in einer einführenden Veranstaltung allerdings zu weit gehen würde.

Kapitel 2

Reelle Zahlen

2.1 Algebraische Grundbegriffe und Körper

Definition 2.1.1. *Sei $M \neq \emptyset$ eine Menge. Jede Funktion $f : M \times M \to M$ heißt eine (binäre, innere)* **Verknüpfung** *oder eine Operation auf M. Wir schreiben für $(a, b) \in M \times M, a * b = f(a, b) \in M$. Dabei kann das Zeichen $*$ jeweils durch $\circ, +, \cdot$ oder ähnlichem ersetzt werden.*

Beispiel 2.1.2.

1. *Es sei $M = \mathbb{N}$, $M = \mathbb{Q}$ oder $M = \mathbb{R}$. Die* **Addition** *auf M*

$$a, b \in M : (a, b) \mapsto a + b \in M$$

   ist eine Verknüpfung.

2. *Auf $\mathbb{Q}$ sind durch*

$$a * b = \max\{a, b\}$$

   und

$$a \circ b = \frac{1}{2}(a + b)$$

   ebenfalls Verknüpfungen bzw. Operationen definiert.

Versehen wir eine Menge M mit einer Operation $*$, so reden wir von einer *algebraischen Struktur* $(M, *)$.

Bemerkung 2.1.3. *Da vorausgesetzt wurde, dass $\cdot * \cdot : M \times M \to M$ eine Funktion ist, folgt für eine Operation $*$ auf M stets*

$$a, b \in M \Rightarrow a * b \in M.$$

Definition 2.1.4. *Sei $(M, *)$ eine algebraische Struktur.*

(G1) **Kommutativität:** *Für alle $a, b \in M$ gilt*

$$a * b = b * a.$$

(G2) **Assoziativität:** *Für alle $a, b, c \in M$ gilt*

$$(a * b) * c = a * (b * c).$$

(G3) **Neutrales Element** *oder* **Einselement***: Es existiert ein Element $e \in M$, so dass für alle $a \in M$ gilt*

$$a * e = e * a = a.$$

(G4) **Inverses Element:** *Zu jedem $a \in M$ existiert ein $a^{-1} \in M$ mit*

$$a * a^{-1} = a^{-1} * a = e.$$

Ist eine algebraische Struktur $(M, *)$ assoziativ, d.h. ist (G2) erfüllt, so nennen wir $(M, *)$ eine **Halbgruppe**. *Eine Halbgruppe mit Einselement $(M, *, e)$ heißt auch* **Monoid**. *Sind die Eigenschaften (G2),(G3) und (G4) erfüllt, so heißt $(M, *)$ eine* **Gruppe**. *Falls zusätzlich auch (G1) gilt, so reden wir von einer* **kommutativen** *oder* **abelschen Gruppe**.

Beispiel 2.1.5.

1. *Die Menge der natürlichen Zahlen $\mathbb{N}$ versehen mit der Multiplikation $\cdot$ ist eine Halbgruppe mit dem neutralen Element $e = 1$.*

2. *Die Mengen $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$ versehen mit der Addition $+$ ist eine Halbgruppe mit dem neutralen Element $e = 0$.*

3. *$(\mathbb{Z} \setminus \{0\}, \cdot)$ ist eine Halbgruppe.*

4. *Sowohl $(\mathbb{Z}, +)$ als auch $(\mathbb{Q} \setminus \{0\}, \cdot)$ sind abelsche Gruppen.*

5. *Sei X eine Menge und $G = \{f : X \to X : f \text{ bijektiv}\}$ die Menge der bijektiven Funktion von X nach X. G versehen mit der Komposition $\circ : x \mapsto (f \circ g)(x) := f\big(g(x)\big)$, $x \in X$, ist ebenfalls eine Gruppe. Das neutrale Element ist hierbei die identische Abbildung $\mathrm{id} : X \to X$ gegeben durch $\mathrm{id}(x) := x$ für alle $x \in X$. Das inverse Element zu f ist die Umkehrfunktion f^{-1} zu f. Man beachte allerdings, dass diese Gruppe i.a. nicht kommutativ ist.*

Satz 2.1.6.

1. *Sei $(M, *)$ eine Halbgruppe mit Einselement, dann gibt es höchstens ein neutrales Element.*

2. *Sei $(M, *)$ eine Halbgruppe mit Einselement, dann gibt es zu jedem $x \in M$ höchstens ein inverses Element x^{-1}.*

3. *Sei $(M, *)$ eine Gruppe, dann gilt für alle $x, y, z \in M$*

$$x * y = x * z \quad \Leftrightarrow \quad y = z.$$

4. *Sei $(M, *)$ eine Gruppe, dann folgt für alle $x, y \in M$ mit $x * y = x$ dass $y = e$.*

5. *In einer Gruppe $(M, *)$ gilt für alle $x \in G$*

$$(x^{-1})^{-1} = x.$$

Beweis. 1. Seien $e, e' \in M$ neutrale Elemente. Dann gelten für alle $x \in M$ die Gleichungen

$$x * e = e * x = x$$

und

$$x * e' = e' * x = x.$$

Wir setzen in die erste Gleichung $x := e'$ und in die zweite $x := e$ ein und erhalten durch Gleichsetzen

$$e' = e' * e = e.$$

2. Sei $x \in M$ beliebig und seien $y, y' \in M$, sofern sie existieren, beides inverse Elemente zu x, d.h.

$$y * x = x * y = x * y' = y' * x = e.$$

Wir verwenden das Assoziativgesetz (G2)

$$y = y * e = y * (x * y') = (y * x) * y' = e * y' = y'.$$

3. Wir zeigen zunächst die Implikation "$\Leftarrow$". Aus $y = z$ folgt durch Einsetzen $x * y = x * z$. Um "$\Rightarrow$" zu zeigen, bemerken wir zuerst, dass zu jedem $x \in M$ das inverse Element $x^{-1} \in M$ exisitiert. Hiermit folgt dann

$$\begin{aligned} y = e * y = (x^{-1} * x) * y = x^{-1} * (x * y) = \\ = x^{-1} * (x * z) = (x^{-1} * x) * z = e * z = z. \end{aligned}$$

4. Die Behauptung folgt aus Punkt 3 mit $z = e$.

5. Per Definition gilt

$$e = x^{-1} * x = (x^{-1}) * x = x * (x^{-1}).$$

□

Definition 2.1.7 (Körper). *Eine Menge M versehen mit zwei Operationen $+$ und $\cdot$ heißt ein* **Körper** *$(M, +, \cdot)$, wenn*

1. *$(M, +)$ eine kommutative Gruppe ist, d.h. für alle $a, b, c \in M$ gelten*

(A1) $a + b = b + a$,

(A2) $a + (b + c) = (a + b) + c$,

(A3) $a + 0 = a$ *(0 ist das neutrale Element bezüglich der Addition),*

(A4) $a + (-a) = 0$ *($-a$ ist das inverse Element zu a bezüglich der Addition),*

2. $(M \setminus \{0\}, \cdot)$ *ebenfalls eine kommutative Gruppe ist, d.h. für alle $a, b, c \in M \setminus \{0\}$ gelten*

(M1) $a \cdot b = b \cdot a$

(M2) $a \cdot (b \cdot c) = (a \cdot b) \cdot c$

(M3) $a \cdot 1 = a$ *(1 ist das neutrale Element bezüglich der Multiplikation),*

(M4) $a \cdot \frac{1}{a} = 1$ *($\frac{1}{a}$ ist das inverse Element bezüglich der Multiplikation),*

3. *es gilt das Distributivgesetz, d.h. für alle $a, b, c \in M$ gilt*

(D) $(a + b) \cdot c = (a \cdot c) + (b \cdot c) =: a \cdot c + b \cdot c$.

Beispiel 2.1.8.

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

2. $M = \{0, 1\}$ *versehen mit der folgenden Addition und Multiplikation*

a	b	$a + b$
0	0	0
0	1	1
1	0	1
1	1	0

a	b	$a \cdot b$
0	0	0
0	1	0
1	0	0
1	1	1

ist ebenfalls ein Körper.

Satz 2.1.9. *Sei $(K, +, \cdot)$ ein Körper, dann gelten für beliebige $x, y \in K$ die folgenden Aussagen.*

1. *Es gilt $-(-x) = x$ und $-0 = 0$.*

2. *Für alle $x \in K$ gilt $0 \cdot x = x \cdot 0 = 0$.*

3. *Für alle $x, y \in K$ mit $x \neq 0$ und $y \neq 0$ folgt $x \cdot y \neq 0$.*

4. *Aus der Gleichung $x \cdot y = 0$ folgt $x = 0$ oder $y = 0$.*

5. *Falls gilt $xy = xz = 1$ ist $y = z$.*

6. *Es ist $(-x)y = -(xy)$.*

Beweis.

1. Diese Eigenschaft folgt aus der Eindeutigkeit des neutralen und inversen Elementes einer Gruppe, vgl. Satz 2.1.6.

2. Es gilt
$$0 \cdot x + 0 \cdot x = (0 + 0) \cdot x = 0 \cdot x = 0 \cdot x + 0$$
und wegen Satz 2.1.6, Punkt 4, folgt hieraus $0 \cdot x = 0$.

3. Die Aussage beweisen wir indirekt. Wir nehmen an, dass $x \cdot y = 0$. Dann existieren die inversen Elemente $x^{-1}, y^{-1} \in K$ und es ist
$$1 = (x \cdot x^{-1}) \cdot (y \cdot y^{-1}) = (x^{-1} \cdot y^{-1}) \cdot (x \cdot y) = (x^{-1} \cdot y^{-1}) \cdot 0 = 0.$$
Aus diesem Widerspruch folgt die Behauptung.

4. Folgt aus 3. durch Negation.

5. Aus $x \cdot y = x \cdot z = 1$ folgt $x \neq 0$. Somit existiert x^{-1} und es folgt wie gehabt
$$y = 1 \cdot y = (x^{-1} \cdot x) \cdot y = x^{-1} \cdot (x \cdot y) = x^{-1} \cdot (x \cdot z) = (x^{-1} \cdot x) \cdot z = 1 \cdot z = z.$$

6. Es gilt
$$(-x) \cdot y + x \cdot y = (x - x) \cdot y = 0 \cdot y = 0.$$

□

Falls $(K, +, \cdot)$ ein Körper ist, so schreiben wir $x^2 := x \cdot x$.

Proposition 2.1.10. *In einem Körper gelten*

1. $\forall x \in K : (-1) \cdot x = -x$,

2. $(-1)^2 = 1$ *und*

3. $\forall x \in K : (-x)^2 = x^2$.

Beweis. Übung. □

Wir werden im folgenden einfachheitshalber kurz xy statt $x \cdot y$ schreiben.

2.2 Geordnete Mengen

Definition 2.2.1. *Sei M eine Menge. Eine* **Ordnung** *auf M ist eine Relation, üblicherweise mit "$<$" bezeichnet, mit den beiden folgenden Eigenschaften:*

(O1) Für beliebige $x, y \in M$ gilt genau eine (!) der folgenden Eigenschaften

$$x < y, \quad x = y \quad \textit{oder} \quad x > y.$$

(O2) Sind $x, y, z \in M$, so folgt aus $x < y$ und $y < z$ dass auch $x < z$ gilt (Transitivität).

Die Aussage "$x < y$" wird als "x kleiner als y" bzw. "$x > y$" als "x größer als y" gelesen. "$x \leq y$" bedeutet "$x < y$ oder $x = y$". Eine geordnete M, kurz $(M, <)$, ist eine Menge M auf der eine Ordnung "$<$" definiert ist.

Beispiel 2.2.2. *Sowohl die natürlichen Zahlen $\mathbb{N}$, die ganzen Zahlen $\mathbb{Z}$ als auch die rationalen Zahlen $\mathbb{Q}$ bilden geordnete Mengen. Die Aussage "$r < s$" ist gleichbedeutend mit $s - r$ ist positiv.*

Definition 2.2.3. *Sei $(M, <)$ eine geordnete Menge und $N \subseteq M$. Falls ein $y \in M$ existiert mit $x \leq y$ für alle $x \in N$, so heißt N* **nach oben beschränkt** *und y eine* **obere Schranke** *von N.* **Untere Schranken** *werden analog definiert.*

Definition 2.2.4. *Sei M eine geordnete Menge und $N \subseteq M$ eine nach oben beschränkte Teilmenge von M. Falls ein Element $y \in M$ derart existiert, dass*

1. *y ist eine obere Schranke von N,*
2. *für alle $x \in M$ mit $x < y$ ist x keine obere Schranke von N,*

das heißt, y ist die kleinste obere Schranke von N, dann heißt y das **Supremum** *von N, kurz $y = \sup N$.*

In der gleichen Art und Weise definiert man das **Infimum** *$y = \inf N$ einer nach unten beschränkten Menge N, als die größte untere Schranke von N.*

Bemerkung 2.2.5. *Das Infimum bzw. das Supremum von N, sofern es existiert, muss selbst kein Element von N sein.*

Beispiel 2.2.6. *Sei $M = \mathbb{Q}$:*

1. $A = \{1, 2, \frac{3}{2}, \frac{1}{2}\}, \quad \sup A = 2 \in A, \quad \inf A = \frac{1}{2} \in A.$
2. $B = \{\frac{1}{n} : n \in \mathbb{N}\}, \quad \sup B = 1 \in B$ *aber* $\inf B = 0 \notin B.$

3. *$C = \{x \in \mathbb{Q} : x^2 < 2\}$: C ist beschränkt, aber es existiert in $M = \mathbb{Q}$ weder ein Infimum noch ein Supremum.*

Definition 2.2.7. *Eine geordnete Menge M besitzt die* **Supremumseigenschaft***, falls zu jeder nichtleeren und nach oben beschränkten Teilmenge $N \subseteq M$ ein Supremum $y = \sup N \in M$ existiert.*

Satz 2.2.8. *Sei M eine geordnete Menge mit der Supremumseigenschaft und sei $N \subseteq M$ nach unten beschränkt. Sei $U \subseteq M$ die Menge aller unteren Schranken von N, d.h.*

$$U = \{y \in M : y \leq x \text{ für alle } x \in N\}.$$

Dann existiert $y = \inf N$ und es gilt $y = \sup U \in U$.

Beweis. Da N nach unten beschränkt ist, ist $U \neq \emptyset$. Andererseits ist $U \subseteq M$ nach oben beschränkt, denn jedes $x \in N$ ist eine obere Schranke von U. Infolge der Supremumseigenschaft existiert

$$y = \sup U.$$

Sei $z < y$ beliebig, dann ist z keine obere Schranke von U, d.h. $z \notin N$ und folglich gilt $z < x$ für alle $x \in N$. Daraus folgt $y \leq x$ für alle $x \in N$ und somit $y \in U$ nach Definition von U. Aber es existiert wegen $y = \sup U$ kein $z > y$ mit $z \leq x$ für alle $x \in N$. Das heißt, y ist die größte untere Schranke, also das Infimum von N. □

2.3 Geordnete Körper

Definition 2.3.1. *Sei K ein Körper und zugleich eine geordnete Menge $(K, <)$. K heißt ein* **geordneter Körper***, falls die Ordnungsrelation den folgenden Bedingungen genügt:*

(O3) Für $x, y, z \in K$ mit $x < y$ folgt $x + z < y + z$.

(O4) Für alle $x, y \in K$ mit $x > 0$ und $y > 0$ folgt $x \cdot y > 0$.

Ist $x > 0$, so heißt x *positiv*, falls $x < 0$ heißt x *negativ*. Es folgen einige wichtige Regeln zur Behandlung von Ungleichungen.

Satz 2.3.2. *In einem geordneten Körper gelten folgende Aussagen für $x, y, z \in K$:*

1. *Ist $x > 0$, so ist $-x < 0$ und umgekehrt.*
2. *Ist $x > 0$ und $y < z$, so gilt $x \cdot y < x \cdot z$.*
3. *Ist $x < 0$ und $y < z$, so folgt $x \cdot y > x \cdot z$.*
4. *Ist $x \neq 0$, so ist $x^2 > 0$. Insbesondere ist $1 > 0$.*

5. *Ist* $0 < x < y$*, so folgt* $0 < \frac{1}{y} < \frac{1}{x}$.

Beweis.

1. Aus $x > 0$ folgt $0 = -x + x > -x + 0 = -x$, d.h. $-x < 0$. Die Umkehrung beweist man analog.

2. Aus $z > y$ folgt $z - y > y - y = 0$ und somit $x \cdot (z - y) > 0$, also $x \cdot z - x \cdot y > 0$ bzw. $x \cdot z > x \cdot y$.

3. Aus $z > y$ und $-x > 0$ folgt $(-x) \cdot (z - y) > 0$, d.h. $x \cdot y - x \cdot z > 0$ bzw. $x \cdot y > x \cdot z$.

4. Ist $x > 0$, so folgt $x^2 = x \cdot x > 0$. Ist $x < 0$, dann ist $-x > 0$ und $x^2 = (-x) \cdot (-x) > 0$.

5. Angenommen es ist $\frac{1}{x} < 0$. Daraus folgt der Widerspruch $1 = x \cdot \frac{1}{x} < 0$. Damit gilt $\frac{1}{x}, \frac{1}{y} > 0$ und $\frac{1}{x} \cdot \frac{1}{y} > 0$. Hieraus schließt man aus $x < y$ dass

$$\frac{x}{x \cdot y} = \frac{1}{y} < \frac{y}{x \cdot y} = \frac{1}{x}.$$

□

Definition 2.3.3. *Wir definieren die* **reellen Zahlen** *als einen geordneten Körper mit der Supremunmseigenschaft. Dies bedeutet, eine Menge* $\mathbb{R}$*, die*

1. *die Körperaxiome (A1) – (A4), (M1) – (M4), sowie (D) erfüllt, d.h.* $(\mathbb{R}, +, .)$ *ist ein Körper,*

2. *den Ordnungsaxiomen (O1) – (O4) genügt, d.h.* $(\mathbb{R}, +, \cdot, <)$ *ist ein geordneter Körper,*

3. *die Supremumseigenschaft aus Definition 2.2.7 besitzt,*

bezeichnen wir als reelle Zahlen $\mathbb{R}$.

2.4 Natürliche Zahlen

Haben wir die reellen Zahlen axiomatisch eingeführt müssen wir darauf aufbauend die natürlichen Zahlen definieren. Die ganzen und die rationalen Zahlen folgen dann unmittelbar aus dieser Definition.

Definition 2.4.1. *Eine Teilmenge* $N \subseteq \mathbb{R}$ *heißt eine* **induktive Menge**, *falls*

1. $1 \in N$

2. *aus* $a \in N$ *folgt* $a + 1 \in N$.

Der Durchschnitt aller induktiven Mengen heißt die **Menge der natürlichen Zahlen***, die mit* $\mathbb{N}$ *bezeichnet wird.*

Satz 2.4.2 (Archimedische Eigenschaft). *Seien* $x, y \in \mathbb{R}$ *mit* $x > 0$*, so existiert eine natürliche Zahl* $n \in \mathbb{N}$*, so daß*

$$nx > y.$$

Beweis. Wir führen einen indirekten Beweis. Sei $A = \{n \cdot x : n \in \mathbb{N}\}$. Angenommen die Behauptung sei falsch, dann ist A beschränkt und y eine obere Schranke von A. Wegen der Supremumseigenschaft existiert $\alpha = \sup A$. Da $x > 0$ gilt $\alpha - x < \alpha$, also ist $\alpha - x$ keine obere Schranke von A, d.h. es existiert $m \in \mathbb{N}$ mit $\alpha - x < mx \in A$. Dann ist aber $\alpha < (m+1)x$. Folglich α ist keine obere Schranke. Dies steht im Widerspruch zur Annahme und daher ist die Menge A unbeschränkt. □

Satz 2.4.3. *Sei* $M \subseteq \mathbb{N}$ *eine Teilmenge der natürlichen Zahlen. Dann besitzt* M *ein Minimum, d.h. es existiert* $m \in M$*, so dass für alle* $n \in M$ *gilt* $m \leq n$*. Dieses Minimum wird mit* $m = \min M$ *bezeichnet und es gilt* $\min M = \inf M$*.*

Beweis. Wir bemerken zunächst, dass $M \subseteq \mathbb{N}$ nach unten durch 1 beschränkt ist, folglich ein Infimum besitzt. Sei $n = \inf M \in \mathbb{R}$. Angenommen es gelte $n \notin M$, dann existert ein $m \in M$ mit $n < m < n + 1$. Nun ist entweder $m = \min M$ und somit $m = \inf M$ im Widerspruch zu $n < m$ oder es existiert ein $l \in M$ mit $l < m$. Dann ist aber auch $m - 1 \in M$, was im Widerspruch steht zu $m < n + 1 \Leftrightarrow m - 1 < n$. □

Satz 2.4.4 (Dichtheit von $\mathbb{Q}$). *Seien* $x, y \in \mathbb{R}$ *mit* $x < y$*, so existiert* $q \in \mathbb{Q}$ *mit* $x < q < y$*.*

Beweis. Aus $x < y$ folgt $y - x > 0$. Wegen der Archimedischen Eigenschaft gibt es eine natürliche Zahl n mit $n(y - x) > 1$. Desweiteren folgt aus dem gleichen Grund, dass $m \in \mathbb{N}$ existiert mit

$$nx < m.$$

Aufgrund von Satz 2.4.3 existiert eine kleinste derartige natürliche Zahl m. Diese erfüllt dann sogar

$$m - 1 \leq nx < m.$$

Somit gilt

$$nx < m \leq 1 + nx < ny,$$

und hieraus erhalten wir

$$x < \frac{m}{n} < y.$$

□

Die Menge $\mathbb{N}$ ist natürlich wieder eine induktive Menge. Eine wichtige Anwendung induktiver Mengen ist das Prinzip der vollständigen Induktion, das wir vielfach als Beweismethode verwenden werden. Der *Beweis durch vollständige Induktion* beruht auf dem Nachweis, dass diejenigen Zahlen n, für die $A(n)$ richtig ist eine induktive Menge bilden.

Satz 2.4.5. *Seien $A(n)$, $n \in \mathbb{N}$, Aussagen, dann sind die beiden folgenden Aussagen äquivalent*

(i) $\forall_{n\in\mathbb{N}} A(n)$,

(ii) $A(1)$ *und* $\forall_{n\in\mathbb{N}} A(n) \Rightarrow A(n+1)$.

Beweis. Die Implikation "(i) $\Rightarrow$ (ii)" ist klar, d.h. wir müssen nur noch "(i) $\Leftarrow$ (ii)" zeigen. Dazu sei $X := \{n \in \mathbb{N} : A(n)\}$ und $Y := \mathbb{N} \setminus X$. Aus (ii) folgt $1 \notin Y$ und $n+1 \in Y \Rightarrow n \in Y$. Wir nehmen an, dass $Y \neq \emptyset$. Wegen Satz 2.4.3 besitzt die Menge $Y \subseteq \mathbb{N}$ ein Minimum m mit $m > 1$. Damit ist aber auch $m - 1 \in Y$, was im Widerspruch zu $m = \min Y$ steht. □

Um das Induktionsprinzip zu demonstrieren beweisen wir die *Bernoullische Ungleichung*.

Lemma 2.4.6 (Bernoullische Ungleichung). *Für beliebige $n \in \mathbb{N}$ und $x > -1$ gilt*

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

Dabei gilt die strenge Ungleichung nur für $n > 1$ und $x \neq 0$.

Beweis.

(i) Induktionsverankerung: Für $n = 1$ lautet die Behauptung

$$1 + x = 1 + x,$$

was für alle $x \in \mathbb{R}$ richtig ist.

(ii) Induktionsschritt $n \mapsto n+1$: Wir nehmen an, dass für ein festes $n \in \mathbb{N}$ gilt

$$(1+x)^n \geq 1 + nx \quad \forall x > -1.$$

Wir müssen nun die Behauptung für $n+1$ zeigen. Diese folgt aus der für alle $x > -1$ gültigen Ungleichungskette

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

Den zweiten Teil der Aussage überprüft man analog mit der Induktionsverankerung $n = 2$. □

Satz 2.4.7. *Zu jeder reellen Zahl $x > 0$, $x \in \mathbb{R}$ und jeder natürlichen Zahl $n \in \mathbb{N}$ existiert genau eine positive reelle Zahl y mit*

$$y^n = x.$$

Wir schreiben hierfür $y = \sqrt[n]{x}$ oder $y = x^{\frac{1}{n}}$.

Beweis. Für $0 < y_1 < y_2$ folgt $0 < y_1^n < y_2^n$. Hieraus ersehen wir sofort, dass es höchstens ein $y > 0$ geben kann mit $y^n = x$. Damit ist die Eindeutigkeit von y bewiesen.

Wir zeigen nun die Existenz. Sei $E = \{t : t^n < x\}$. Für $t = \frac{x}{1+x}$ gilt $0 < t < \min\{1, x\}$. Hieraus folgt $0 < t^n \leq t < x$, somit $t \in E$ und $E \neq \emptyset$. Die Supremumseigenschaft sichert die Existenz von $y := \sup E \in \mathbb{R}$. Es bleibt zu zeigen, dass hierfür $y^n = x$ gilt. Dies geschieht indirekt indem wir die Annahmen $y^n < x$, $y^n > x$ beide ad absurdum führen. Wir verwenden hierzu die folgende Identität

$$b^n - a^n = (b-a)(b^{n-1} + b^{n-2}a + \ldots + ba^{n-2} + a^{n-1}), \tag{2.1}$$

woraus für $0 < a < b$ folgt

$$b^n - a^n < (b-a)nb^{n-1}. \tag{2.2}$$

1. Sei $y^n < x$ angenommen. Wir wählen ein h mit $0 < h < 1$ und

   $$h = \frac{x - y^n}{n(y+1)^{n-1}} > 0.$$

   Wir setzen in (2.1) $a = y$ und $b = y + h$. Dann gilt wegen (2.2)

   $$(y+h)^n - y^n < hn(y+h)^{n-1} < hn(y+1)^{n-1} < x - y^n,$$

   also $(y+h)^n < x$, d.h. $y + h \in E$. Wegen $y + h > y$ liegt ein Widerspruch zu $y = \sup E$ vor.

2. Im Falle $y^n > x$ setzen wir

   $$k = \frac{y^n - x}{ny^{n-1}} > 0.$$

   Dann ist $0 < k < y$ und für beliebiges $t \leq y - k > 0$, also $t < y$, gilt

   $$y^n - t^n < y^n - (y-k)^n < kny^{n-1} = y^n - x.$$

   Somit gilt $t^n > x$ und $t \notin E$ und folglich ist $y - k$ eine obere Schranke zu E. Wegen $y - k < y = \sup E$ liegt hier ebenfalls ein Widerspruch vor.

□

Korollar 2.4.8. *Für $x, y \in \mathbb{R}$, $x, y > 0$, und $n \in \mathbb{N}$ gilt*

$$(xy)^{\frac{1}{n}} = x^{\frac{1}{n}} \cdot y^{\frac{1}{n}}.$$

Beweis. Setze $a := x^{\frac{1}{n}}$ und $b := y^{\frac{1}{n}}$. Dann gilt $x \cdot y = a^n \cdot b^n = (a \cdot b)^n$. Wegen der Eindeutigkeitsaussage im letzten Satz folgt $(x \cdot y)^{\frac{1}{n}} = ab = x^{\frac{1}{n}} y^{\frac{1}{n}}$. □

Sei $A = \{a_1, a_2, \ldots, a_N\}$. Die Summe aller Zahlen aus A bezeichnen wir abkürzend mit

$$\sum_{i=1}^{N} a_i := a_1 + a_2 + \ldots + a_N$$

und das Produkt aller Zahlen aus A mit

$$\prod_{i=1}^{N} a_i := a_1 \cdot a_2 \cdot \ldots \cdot a_N.$$

Insbesondere schreiben wir im Falle $a_1 = 1$, $a_2 = 2$, $\ldots a_N = N$ auch

$$N! := 1 \cdot 2 \cdot \ldots \cdot N = \prod_{i=1}^{N} i.$$

Dabei wird $N!$ als *Fakultät* von N bezeichnet.

Definition 2.4.9. *Seien $n, k \in \mathbb{N}$ natürliche Zahlen. Wir definieren für $n \geq k$ die* **Binomialkoeffizienten** *durch*

$$\binom{n}{0} := \binom{n}{n} := 1,$$

und

$$\binom{n}{k} := \frac{n \cdot (n-1) \cdot \ldots \cdot (n-k+1)}{1 \cdot 2 \cdot \ldots \cdot k} = \prod_{i=1}^{k} \frac{n+1-i}{i} = \frac{n!}{k!(n-k)!}.$$

Proposition 2.4.10. *Es gelten die Rechenregeln*

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

und

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

Beweis. Der Beweis beruht auf vollständiger Induktion, wir verzichten allerdings auf ihn. Man beachte, dass sich beide Formeln durch das Pascalsche Dreieck motivieren lassen.

$$\begin{array}{lccccccccccccc}
n=0: & & & & & & & 1 & & & & & & \\
 & & & & & & / & & \backslash & & & & & \\
n=1: & & & & & 1 & & & & 1 & & & & \\
 & & & & / & & \backslash & & / & & \backslash & & & \\
n=2: & & & 1 & & & & 2 & & & & 1 & & \\
 & & / & & \backslash & & / & & \backslash & & / & & \backslash & \\
n=3: & 1 & & & & 3 & & & & 3 & & & & 1 \quad \longrightarrow k \\
 & k=0 & & & & k=1 & & & & k=2 & & & & k=3 \\
\downarrow & & & & & & & & & & & & & \\
n & & & & & & & & & & & & &
\end{array}$$

□

Satz 2.4.11 (Binomischer Satz). *Seien $a, b \in \mathbb{R}$ und $n \in \mathbb{N}$, dann gilt*

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

Beweis. Wir beweisen diesen Satz mittels vollständiger Induktion.

(i) Induktionsverankerung: Für $n = 1$ ist die Behauptung

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

offensichtlich richtig.

(ii) Induktionsschritt $n \mapsto n+1$: Wir nehmen an, dass für n die Behauptung richtig ist. Dann folgt mit Proposition 2.4.10

$$\begin{aligned}
(a+b)^{n+1} &= (a+b)^n (a+b) \\
&= \left(\sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k} \right) (a+b) \\
&= \sum_{k=0}^{n} \binom{n}{k} a^{k+1} b^{n-k} + \sum_{k=0}^{n} \binom{n}{k} a^k b^{n+1-k} \\
&= a^{n+1} + b^{n+1} + \sum_{k=0}^{n-1} \binom{n}{k} a^{k+1} b^{n-k} + \sum_{k=0}^{n-1} \binom{n}{k+1} a^{k+1} b^{n-k} \\
&= a^{n+1} + b^{n+1} + \sum_{k=0}^{n-1} \binom{n+1}{k+1} a^{k+1} b^{n-k} \\
&= \sum_{k=0}^{n+1} \binom{n+1}{k} a^k b^{n+1-k}.
\end{aligned}$$

□

2.5 Komplexe Zahlen

Wir betrachten die Menge

$$\mathbb{C} := \mathbb{R} \times \mathbb{R}$$

versehen mit der folgenden Addition und Multiplikation: Für $z = (a, b), w = (c, d) \in \mathbb{C}$ soll

$$z + w := (a + c, b + d) \in \mathbb{C}$$

und

$$z \cdot w := (ac - bd, ad + bc) \in \mathbb{C}$$

gelten.

Satz 2.5.1. *Die Menge $\mathbb{C}$ versehen mit der obigen Addition und Multiplikation bildet den Körper $(\mathbb{C}, +, \cdot)$ der komplexen Zahlen mit dem Nullelement $(0, 0)$ und Einselement $(1, 0)$.*

Beweis. Wir zeigen die Körperaxiome. Zunächst bemerken wir, dass sowohl die Addition als auch die Multiplikation für alle komplexen Zahlen definiert ist und immer eine komplexe Zahl liefert. Im folgenden bezeichnen $z = (a, b)$, $w = (b, d)$ und $v = (e, f)$ stets komplexe Zahlen.

(A1) Die Addition ist kommutativ, da $z + w = (a + c, b + d) = (c + a, d + b) = w + z$.

(A2) Die Addition ist assoziativ, da

$$\begin{aligned}(z + w) + v &= (a + c, b + d) + (e, f) = (a + (c + e), b + (d + f)) \\ &= (a, b) + ((c + e), (d + f)) = z + (w + v).\end{aligned}$$

(A3) $(0, 0)$ ist das Nullelement, da $z + (0, 0) = (a + 0, b + 0) = (a, b) = z$.

(A4) Setze $-z = (-a, -b)$, dann gilt $z + (-z) = (a - a, b - b) = (0, 0)$.

(M1) Die Multipliaktion ist kommutativ, da $z \cdot w = (ac - bd, ad + bc) = (ca - db, da + cb) = w \cdot z$

(M2) Die Multipliaktion ist assoziativ, da

$$\begin{aligned}(z \cdot w) \cdot v &= (ac - bd, ad + bc) \cdot (e, f) \\ &= \big(ac - bd)e - (ad + bc)f, (ac - bd)f + (ad + bc)e\big) \\ &= \big(a(ce - df) - b(cf + de), a(cf + de) + b(ce - df)\big) \\ &= (a, b) \cdot (ce - df, cf + de) \\ &= z \cdot (w \cdot v)\end{aligned}$$

(M3) Wegen $(1, 0) \cdot (a, b) = (a, b) = z$ ist $(1, 0)$ das Einselement.

(M4) Aus $z \neq (0, 0)$ folgt $a \neq 0$ oder $b \neq 0$ und damit ist $a^2 + b^2 > 0$. Wir definieren

$$z^{-1} = \frac{1}{z} = \Big(\frac{a}{a^2 + b^2}, \frac{-b}{a^2 + b^2}\Big).$$

Nachrechnen ergibt

$$z \cdot \frac{1}{z} = (a, b) \cdot \Big(\frac{a}{a^2 + b^2} \frac{-b}{a^2 + b^2}\Big) = (1, 0)$$

und analog $\frac{1}{z} \cdot z = (1, 0)$.

(D) Das Distributivgesetz folgt mittels

$$\begin{aligned} z \cdot (w+v) &= (a,b) \cdot (c+e, d+f) \\ &= (ac+ae-bd-bf, ad+af+bc+be) \\ &= (ac-bd, ad+bc) + (ae-bf, af+be) \\ &= z \cdot w + z \cdot v. \end{aligned}$$

□

Definition 2.5.2. *Wir setzen $i = (0,1)$. i bezeichnet die* **imginäre Einheit**. *Ferner identifizieren eine reelle Zahl $a \in \mathbb{R}$ mit der komplexen Zahl $(a,0) \in \mathbb{C}$.*

Satz 2.5.3. *Es gilt für alle $a, c \in \mathbb{R}$*

1. $i^2 = i \cdot i = (-1,0) = -1$,
2. $w = (a,b) = a + ib$.

Beweis. Die erste Aussage ergibt sich aus $i^2 = (0,1) \cdot (0,1) = (-1,0) = -1$ während die zweite Aussage aus $a + ib = (a,0) + (0,1) \cdot (b,0) = (a,b)$ folgt. □

Anstelle von $z = (a,b)$, $a, b \in \mathbb{R}$, schreiben wir meist $z = a + ib \in \mathbb{C}$. Wenn man die Konvention $i^2 = -1$ und die obige Schreibweise benutzt, kann man mit den üblichen Arithmetiken $+$ bzw. $\cdot$ die komplexe Addition bzw. Multiplikation auch folgendermaßen schreiben. Seien $z = a + ib, w = c + id \in \mathbb{C}$ mit $a, b, c, d \in \mathbb{R}$, dann ist

$$z + w = (a+ib) + (c+id) = (a+c) + i(b+d)$$

und

$$z \cdot w = (a+id) \cdot (c+id) = ac + ibc + iad + i^2bd = (ac-bd) + i \cdot (bc+ad).$$

Definition 2.5.4. *Sei $z = a + ib \in \mathbb{C}$, $a, b \in \mathbb{R}$, eine komplexe Zahl. Dann heißt a der* **Realteil** *und b der* **Imaginärteil** *von z, kurz*

$$a := \operatorname{Re} z, \qquad b = \operatorname{Im} z,$$

und $\overline{z} = a - ib$ die zu z **konjugiert komplexe Zahl**.

Satz 2.5.5. *Seien $z, w \in \mathbb{C}$, dann gelten die Rechenregeln*

1. $\overline{z+w} = \overline{z} + \overline{w}$ *und* $\overline{z \cdot w} = \overline{z} \cdot \overline{w}$,
2. $\operatorname{Re} z = \frac{1}{2}(z + \overline{z})$ *und* $\operatorname{Im} z = \frac{1}{2i}(z - \overline{z})$,
3. *das Produkt $z \cdot \overline{z}$ ist reell und positiv falls $z \neq 0$.*

Beweis. Sei $z = a + ib$ und $w = c + id$ dann überprüft man die ersten beiden Aussagen leicht durch Nachrechnen. Die dritte Aussage folgt aus

$$z\overline{z} = (a + ib) \cdot (a + ib) = a^2 - (i^2) \cdot b^2 = a^2 + b^2 > 0,$$

falls $z \neq 0$. □

Definition 2.5.6. *Sei $z \in \mathbb{C}$, so heißt*

$$|z| := \sqrt{z \cdot \overline{z}} = (z \cdot \overline{z})^{\frac{1}{2}} = \sqrt{a^2 + b^2}$$

der **Absolutbetrag** *von z.*

Bemerkung 2.5.7. *Diese Definition ist konsistent mit dem Absolutbetrag einer reellen Zahl a, da $|a| = \sqrt{a^2} \geq 0$.*

Satz 2.5.8. *Seien $z, w \in \mathbb{C}$ komplexe Zahlen. Dann gilt*

1. $|z| > 0$ *für* $z \neq 0$ *und* $|0| = 0$,
2. $|\overline{z}| = |z|$,
3. $|z \cdot w| = |z||w|$,
4. $\operatorname{Re} z, \operatorname{Im} z \leq |z|$,
5. $|z + w| \leq |z| + |w|$ *(Dreiecks-Ungleichung),*
6. $||z| - |w|| \leq |z - w|$.
7. $|z||w| \leq \frac{1}{2}(|z|^2 + |w|^2)$.

Beweis. Die Aussagen 1,2 und 4 sind trivial. Die dritte Aussage folgt aus $|z \cdot w|^2 = z \cdot w \cdot \overline{z \cdot w} = z \cdot \overline{z} \cdot w \cdot \overline{w} = |z|^2|w|^2$. Die fünfte Aussage ergibt sich aus

$$\begin{aligned} |z + w|^2 &= (z + w)\overline{(z + w)} \\ &= z\overline{z} + z\overline{w} + w\overline{z} + w\overline{w} \\ &= |z|^2 + 2\operatorname{Re}(w\overline{z}) + |w|^2 \\ &\leq |z|^2 + 2|w||z| + |w|^2 \\ &= (|z| + |w|)^2. \end{aligned}$$

Die sechste Aussage verbleibt als Übung. Aussage 7 ergibt sich mit $0 \leq (|z| - |w|)^2 = |z|^2 - 2|z||w| + |w|^2$. □

Als Abschluss dieses Abschnittes geben wir noch zwei bedeutende Sätze an, deren Beweis allerdings unsere bisherigen Möglichkeiten übersteigt. Die Beweise werden wir an späterer Stelle nachreichen.

Satz 2.5.9 (Eulersche Formel). *Sei $\varphi \in \mathbb{R}$, dann gilt die sogenannte Eulersche Formel*

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

Satz 2.5.10 (Funsamentalsatz der Algebra). *Seien $a_0, a_1, \ldots, a_{n-1} \in \mathbb{C}$. Dann hat die Gleichung*

$$z^n + a_{n-1}z^{n-1} + \ldots + a_1 z + a_0 = 0$$

genau n nicht notwendigerweise verschiedene komplexe Lösungen.

2.6 Euklidische Räume

Definition 2.6.1. *Sei $n \in \mathbb{N}$, dann definieren wir das n-Tupel*

$$\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$$

mit $x_1, x_2, \ldots, x_n \in \mathbb{R}$. Die Elemente in $\mathbb{R}^n$ nennen wir **Vektoren** *oder* **Punkte** *im $\mathbb{R}^n$. Für diese definieren wir die folgende Operation der Addition $+ : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$,*

$$\mathbf{x} + \mathbf{y} = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n), \quad \mathbf{x}, \mathbf{y} \in \mathbb{R}^n,$$

und die Multiplikation mit einem Skalar $\cdot : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$

$$\alpha\mathbf{x} = (\alpha x_1, \alpha x_2, \ldots, \alpha x_n), \quad \mathbf{x} \in \mathbb{R}^n, \alpha \in \mathbb{R}.$$

Desweiteren führen wir ein **inneres Produkt**

$$\mathbf{x}, \mathbf{y} \in \mathbb{R}^n \to \langle \mathbf{x}, \mathbf{y} \rangle := \sum_{i=1}^{n} x_i y_i,$$

sowie die Norm von x,

$$\|\mathbf{x}\| = \langle \mathbf{x}, \mathbf{x} \rangle^{\frac{1}{2}} = \left(\sum_{i=1}^{n} x_i^2 \right)^{\frac{1}{2}}.$$

Die so definierte Struktur eines Vektorraumes mit einem inneren Produkt heißt ein n-dimensionaler **Euklidischer Raum**.

Satz 2.6.2. *Es seien $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$ und $\alpha \in \mathbb{R}$.*

1. *Kommutativität: $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$.*
2. *Linearität: $\langle \mathbf{x} + \mathbf{y}, \mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{z} \rangle + \langle \mathbf{y}, \mathbf{z} \rangle$ und $\langle \alpha\mathbf{x}, \mathbf{y} \rangle = \alpha\langle \mathbf{x}, \mathbf{y} \rangle$.*
3. *Es ist $\|\mathbf{x}\| > 0$ genau dann wenn $\mathbf{x} \neq \mathbf{0} = (0, 0, \ldots, 0)$. Im Falle $\mathbf{x} = \mathbf{0}$ gilt $\|\mathbf{x}\| = 0$.*
4. *Homogenität: Es gilt $\|\alpha\mathbf{x}\| = |\alpha|\|\mathbf{x}\|$.*

5. *Cauchy-Schwarzsche-Ungleichung: Es gilt* $|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \|\mathbf{x}\| \cdot \|\mathbf{y}\|$.

6. *Dreiecksungleichung: Es gilt* $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$.

Beweis.

1. Beweis erfolgt durch Nachrechnen.

2. Beweis erfolgt durch Nachrechnen.

3. trivial

4. $\|\alpha \mathbf{x}\|^2 = \sum_{i=1}^n (\alpha x_i)^2 = \alpha^2 \sum_{i=1}^n x_i^2 = \alpha^2 \|\mathbf{x}\|^2$

5. Offensichtlich ist die Behauptung richtig im Falle $\mathbf{y} = \mathbf{0}$. Ohne Beschränkung der Allgemeinheit sei nun $\mathbf{y} \neq \mathbf{0}$. Für beliebiges $\alpha \in \mathbb{R}$ gilt

$$\begin{aligned} 0 &\leq \|\mathbf{x} + \alpha \mathbf{y}\|^2 \\ &= \langle \mathbf{x} + \alpha \mathbf{y}, \mathbf{x} + \alpha \mathbf{y} \rangle \\ &= \|\mathbf{x}\|^2 + 2\alpha \langle \mathbf{x}, \mathbf{y} \rangle + \alpha^2 \|\mathbf{y}\|^2. \end{aligned}$$

   Wir wählen $\alpha = \frac{-\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{y}\|^2}$ und erhalten

$$0 \leq \|\mathbf{x}\|^2 - 2\frac{\langle \mathbf{x}, \mathbf{y} \rangle^2}{\|\mathbf{y}\|^2} + \frac{\langle \mathbf{x}, \mathbf{y} \rangle^2}{\|\mathbf{y}\|^4} \|\mathbf{y}\|^2.$$

   Wegen $\|\mathbf{y}\| > 0$ folgt daraus

$$0 \leq \|\mathbf{x}\|^2 \|\mathbf{y}\|^2 - 2\langle \mathbf{x}, \mathbf{y} \rangle^2 + \langle \mathbf{x}, \mathbf{y} \rangle^2 = (\|\mathbf{x}\| \cdot \|\mathbf{y}\|)^2 - |\langle \mathbf{x}, \mathbf{y} \rangle|^2.$$

6. Unter Zuhilfenahme der dritten Aussage finden wir

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

□

2.7 Zusatz: Dedekindsche Schnitte

Wir haben die reellen Zahlen rein axiomatisch als einen geordneten Körper mit der Supremumseigenschaft eingeführt, d.h. wir definieren sie allein aus den Körperaxiomen (A1) – (A4), (M1) – (M4) und (D), den Ordnungsaxiomen (O1) – (O4), sowie der Supremumseigenschaft aus Definition 2.2.7, die wir als letztes Axiom dazugenommen haben. Diese axiomatische Einführung ist zwar mathematisch streng und exakt, aber im Grunde doch recht abstrakt.

Anschaulich würden nur die natürlichen Zahlen axiomatisch eingeführt, und mittels des Umweges über die rationalen Zahlen würden dann die reellen Zahlen hergeleitet werden. Dadurch wird der Stoff allerdings nicht einfacher, sondern nur länger. Für diejenigen, die sich damit zufrieden geben, dass man die rationalen Zahlen schon definiert hat, sei ein Weg aufgezeigt, wie man streng mathematisch die reellen Zahlen $\mathbb{R}$ aus den rationalen Zahlen entwickeln kann. Es sei darauf hingewiesen, dass der anscheinend einfachste Weg über die Dezimaldarstellungen der reellen Zahlen nicht ganz ohne Tücken ist. Dies kann man beispielsweise daran erkennen, dass $2 = 2.0 = 1.\overline{9}$, d.h. die Dezimaldarstellung ist nicht eindeutig. Historisch gesehen gibt es im wesentlichen die Konstruktion von Kantor mittels Äquivalenzklassen von Cauchyfolgen (Cauchyfolgen werden wir erst etwas später kennenlernen) und die Konstruktion von Dedekind, die an dieser Stelle kurz vorgestellt werden soll.

Satz 2.7.1 (Dedekindsche Schnitte). *Es gibt einen geordneten Körper $\mathbb{R}$, der den Körper der rationalen Zahlen enthält und die Supremumseigenschaft besitzt.*

Der Beweis dieses Satz erfordert einige Vorbereitungen.

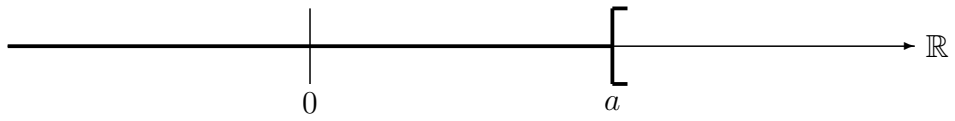
Definition 2.7.2 (Dedekindsche Schnitte). *Eine Teilmenge $\alpha \subseteq \mathbb{Q}$ heißt ein* **(Dedekindscher) Schnitt**, *falls gilt:*

1. *$\alpha \neq \emptyset$ ist nicht leer und $\alpha \neq \mathbb{Q}$.*
2. *Ist $p \in \alpha$, so folgt für alle $q \in \mathbb{Q}$ mit $q < p$ dass $q \in \alpha$.*
3. *Zu jedem $p \in \alpha$ existiert ein $r \in \alpha$ mit $p < r$.*

Bemerkung 2.7.3. *Punkt 3 ist gleichbedeutend damit, dass α kein größtes Element enthält. Punkt 2 impliziert die beiden folgenden Eigenschaften: Erstens, aus $p \in \alpha$ und $q \notin \alpha$ folgt $p < q$, und zweitens, aus $r \notin \alpha$ und $r < s$ folgt $s \notin \alpha$.*

Wir werden später die Schnitte mit Elementen von $\mathbb{R}$ identifizieren. Man überzeugt sich leicht, dass für $q \in \mathbb{Q}$ die Menge $q^\star = \{p \in \mathbb{Q} : p < q\}$ einen Schnitt definiert. Wir wollen nun die Menge der Schnitte $\mathbb{R}$ näher betrachten.

Bemerkung 2.7.4. *Wir können uns diesen Sachverhalt wie folgt veranschaulichen. Wir identifizieren eine rationale Zahl a mit der Menge $\{x \in \mathbb{R} : x < a\}$.*



Beachte, dass dies bloß eine Vorstellung ist, da wir weder die reelle Zahlengerade noch $\mathbb{R}$ definiert haben. Wenn wir die reellen Zahlen streng axiomatisch einführen, müssten wir die natürlichen Zahlen $\mathbb{N}$ (bzw. die ganzen Zahlen $\mathbb{Z}$ und die rationalen Zahlen $\mathbb{Q}$) definieren. Dennoch erleichtern Vorstellungen vielfach das Denken, Verständnis und die Beweise.

Wir definieren eine Ordnung auf der Menge $\mathbb{R}$ aller Schnitte durch

$$\alpha < \beta : \quad \alpha \text{ ist echte Teilmenge von } \beta.$$

Wir zeigen zuerst, dass "$<$" eine Ordnung auf $\mathbb{R}$ definiert.

Lemma 2.7.5. *Die Relation "$<$" definiert eine Ordnung auf der Menge der Schnitte $\mathbb{R}$.*

Beweis. Seien $\alpha, \beta \in \mathbb{R}$. Angenommen es gilt nicht $\alpha \leq \beta$, d.h. α ist keine Teilmenge von β (und echt verschieden von β). Dann existiert ein $p \in \mathbb{Q}$ mit $p \in \alpha$ aber $p \notin \beta$. Für jedes $q \in \beta$ gilt wegen Punkt 2 aus Definition 2.7.2 $q < p$ und aus dem gleichen Grund folgt $q \in \alpha$, d.h. $\beta \subseteq \alpha$ bzw. $\beta < \alpha$. Da die Rollen von α und β austauschbar sind, ist das Lemma hiermit bewiesen. $\square$

Das nächste Lemma sichert die Supremumseigenschaft.

Lemma 2.7.6. *Die geordnete Menge $(\mathbb{R}, <)$ besitzt die Supremumseigenschaft.*

Beweis. Sei $A \subseteq \mathbb{R}$ eine nichtleere Teilmenge, die eine obere Schranke $\beta \in \mathbb{R}$ besitzt. Wir bilden $\gamma = \bigcup_{\alpha \in A} \alpha$ als die Vereinigung aller Schnitte $\alpha \in A$. Wir zeigen zunächst, dass γ ein Schnitt ist. Da A nicht leer und nach oben durch β beschränkt ist, ist γ nicht leer und ebenfalls nach oben beschränkt. In der Tat ist β auch eine obere Schranke für γ. Insbesondere gilt $\gamma \neq \mathbb{Q}$.

Für $p \in \gamma$ gibt es ein $\alpha \in A$ mit $p \in \alpha$. Da α ein Schnitt ist, gilt für jedes $q \in \mathbb{Q}$ mit $q < p$ dass $q \in \alpha$ und somit auch $q \in \gamma$. Analog existiert zu beliebigen $p \in \gamma$ ein α mit $p \in \alpha$ und damit auch ein $r \in \alpha$ bzw. $r \in \gamma$ mit $p < r$. Die Tatsache, dass γ eine obere Schranke ist, ist klar, da jedes $\beta \subseteq \gamma = \bigcup_{\alpha \in A} \alpha$.

Wir zeigen, dass γ die kleinste obere Schranke ist. Angenommen, es gebe eine kleinere obere Schranke δ mit $\delta < \gamma$. Dies bedeutet, dass δ eine echte Teilmenge von γ ist. Folglich existiert $p \in \gamma$ mit $p \notin \delta$. Da $p \in \gamma$ existiert $\alpha \in A$ mit $p \in \alpha$. Dann gilt, $\delta < \alpha$, d.h. δ ist keine obere Schranke für A im Widerspruch zur Annahme. $\square$

Wir definieren nun die Operation "$+$" auf $\mathbb{R}$ durch $\alpha \pm \beta = \{p \pm q : p \in \alpha, q \in \beta\}$ und $0^\star = \{p \in \mathbb{Q} : p < 0\}$ sowie $-\alpha = 0^\star - \alpha$. Zunächst ist zu zeigen, dass $\alpha + \beta$ ebenfalls ein Schnitt ist.

Lemma 2.7.7. *Aus* $\alpha, \beta \in \mathbb{R}$ *folgt* $\alpha + \beta \in \mathbb{R}$.

Beweis. $\alpha + \beta$ ist nicht leer und $\alpha + \beta \neq \mathbb{Q}$.

Zu jedem $p \in \alpha + \beta$ existiert ein $a \in \alpha$ und ein $b \in \beta$ mit $p = a + b$. Für jedes $q \in \mathbb{Q}$ mit $q < p$ gilt $q - b < p - b = a$. Also ist $q - b \in \alpha$ und $(q - b) + b = q \in \alpha + \beta$.

Zu $p = a + b$ existieren $q_1 \in \alpha$ mit $q_1 > a$ und $q_2 \in \beta$ mit $q_2 > b$. Damit ist $q_1 + q_2 \in \alpha + \beta$ und $q_1 + q_2 > a + q_2 > a + b$. □

Die Definition der Multiplikation ist etwas subtiler. Wir beschränken uns zunächst auf die positiven reellen "Zahlen". Seien $\alpha, \beta \in \mathbb{R}^+ := \{\alpha \in \mathbb{R} : 0^\star < \alpha\}$, dann ist

$$\alpha \cdot \beta := \{p \cdot q \in \mathbb{Q} : p \in \alpha, q \in \beta\}$$

und $1^\star = \{q \in \mathbb{Q} : q < 1\}$. Ferner definieren wir

$$\alpha \cdot 0^\star := 0^\star.$$

Wir vervollständigen die Definition durch die Festlegung

$$\alpha \cdot \beta = \begin{cases} -[(-\alpha) \cdot \beta] & \text{falls } \alpha < 0^\star, \beta > 0^\star \\ (-\alpha) \cdot (-\beta) & \text{falls } \alpha, \beta < 0^\star. \end{cases}$$

Lemma 2.7.8. *Aus* $\alpha, \beta \in \mathbb{R}$ *folgt* $\alpha \cdot \beta \in \mathbb{R}$.

Beweis. Wir betrachten zunächst nur $\alpha, \beta \in \mathbb{R}^+ := \{\alpha \in \mathbb{R} : \alpha > 0^\star\}$. Hierfür erfolgt der Beweis analog zur Addition. Man erhält hieraus sogar, dass aus $\alpha, \beta > 0^\star$ folgt $\alpha \cdot \beta > 0^\star$. Die verbleibenden Fälle $\alpha, \beta \in \mathbb{R}$ kann man mittels der Defintion auf den obigen Fall $\alpha, \beta \in \mathbb{R}^+$ zurückspielen. □

Nun sind die "Körperaxiome" zu prüfen. Nun gilt es die Körpereigenschaften alle zu prüfen.

Lemma 2.7.9. *Die Menge* $\mathbb{R}^\star$ *mit den Operationen "+" und "·" ist ein geordneter Körper.*

Beweis. Da die Beweise der einzelnen Eigenschaften alle ähnlich verlaufen, wollen wir hier nur einige dieser Eigenschaften beispielshaft und ganz knapp behandeln. Der vollständige Beweis verbleibt als Übungsaufgabe.

Multiplikation: Man zeigt beispielsweise die Gruppen-Axiome (M1) – (M4) erst für $(\mathbb{R}^+, \cdot)$. Den allgemeinen Fall erhält man hieraus vermittels der Definition der Multiplikation auf ganz $\mathbb{R}$ unter Verwendung der Beziehung $-(-\alpha) = \alpha$.

Den Nachweis des Distributivgesetzes führen wir lediglich für den Fall, dass $\alpha, \beta < 0^\star$ und $\beta + \gamma > 0^\star$. Dann ist

$$\gamma = (\gamma + \beta) + (-\beta) > 0^\star$$

und

$$\alpha \cdot \gamma = \alpha \cdot (\beta + \gamma) + \alpha \cdot (-\beta).$$

Wegen $\alpha \cdot (-\beta) = -(\alpha \cdot \beta)$ folgt

$$\alpha \cdot (\beta + \gamma) = \alpha \cdot \beta + \alpha \cdot \gamma.$$

In analoger Art und Weise behandelt man die anderen Fälle, was zugegebenermassen sehr aufwendig ist. □

Damit ist in etwas abgekürzter Form gezeigt, dass die oben definierte Menge $\mathbb{R}$ versehen mit einer Addition und einer Multiplikation ein geordneter Körper ist $(\mathbb{R}, +, \cdot)$ bzw. $(\mathbb{R}, +, \cdot, >)$.

Jeder rationalen Zahl $r \in \mathbb{Q}$ ist durch die Definition des *Schnittes* die Menge aller rationalen Zahlen $p \in \mathbb{Q}$ mit $p < r$, d.h. $r^\star = \{p \in \mathbb{Q} : p < r\}$, ein-eindeutig zugeordnet.

Lemma 2.7.10. *Für alle rationalen Zahlen $r, s \in \mathbb{Q}$ gilt*

1. $r^\star + s^\star = (r + s)^\star \in \mathbb{R}$,
2. $r^\star \cdot s^\star = (r \cdot s)^\star \in \mathbb{R}$,
3. $r^\star < s^\star$ *gilt genau dann wenn* $r < s$.

Beweis.

1. Sei $p \in r^\star + s^\star$ dann existieren, $u, v \in \mathbb{Q}$ mit $u < r$, $v < s$ und $p = u + v$, somit ist $p \in (r + s)^\star$. Sei nun umgekehrt $p \in (r + s)^\star$, dann gilt $p < r + s$. Wir wählen $t \in \mathbb{Q}$ so, dass
$$2t = r + s - p$$
und setzen
$$r' := r - t, \quad s' := s - t.$$
Dann sind $r' \in r^\star$, $s' \in s^\star$ und $p = r' + s'$, d.h. es ist $p \in r^\star + s^\star$.
2. Die Aussage beweist man analog zum vorherigen Punkt.
3. Ist $r < s$, so gilt $r \in s^\star$ aber $r \notin r^\star$, also haben wir $r^\star < s^\star$. Umgekehrt folgt aus $r^\star < s^\star$ die Exisitenz von $p \in s^\star$ mit $p \in r^\star$ und hieraus $r \leq p < s$, also auch $r < s$.

□

Wir haben somit gezeigt, dass die Menge $\mathbb{R}$ alle Axiome bzw. alle erforderlichen Eigenschaften der reellen Zahlen besitzt, und dass die rationalen Zahlen $\mathbb{Q}$ mit einer Teilmenge der so definierten Menge $\mathbb{R}$ identifiziert werden können. Streng genommen existiert eine injektive Abbildung $\phi : \mathbb{Q} \to \mathbb{R}$ mit $\phi(q) = q^\star$, die verträglich mit den Operationen "+" und "·" ist, d.h. $\phi(p + q) = (p + q)^\star = p^\star + q^\star$ und analog für die Multiplikation, und $p < q$ gilt genau dann wenn $p^\star < q^\star$. Die Abbildung ϕ ist damit eine Bijektion von $\phi : \mathbb{Q} \to \phi(\mathbb{Q}) \subseteq \mathbb{R}$.

Die schon vorher nachgewiesene Dichtheit von $\mathbb{Q}$ in $\mathbb{R}$ zeigt, dass $\mathbb{R}$ quasi die Vervollständigung von $\mathbb{Q}$ unter der Hinzunahme der Supremumseigenschaft darstellt.

Kapitel 3

Einführung in die Topologie

3.1 Metrische Räume

Definition 3.1.1. *Sei X eine Menge, dann heißt eine Funktion $d : X \times X \to \mathbb{R}$ eine* **Metrik** *auf X falls für alle $x, y, z \in X$ gilt*

1. $d(x, y) \geq 0$ *und* $d(x, y) = 0 \Leftrightarrow x = y$ *(Definitheit),*
2. $d(x, y) = d(y, x)$ *(Symmetrie),*
3. $d(x, z) \leq d(x, y) + d(y, z)$ *(Dreiecksungleichung).*

Eine Metrik d wird mitunter auch als **Distanzfunktion** *bezeichnet. Eine Menge X versehen mit einer Metrik $d(\cdot, \cdot)$ nennen wir einen* **metrischen Raum** *(X, d). Die Elemente $x \in X$ werden dann manchmal Punkte genannt.*

Beispiel 3.1.2.

1. *Der Vektorraum $\mathbb{R}^n$, $n \in \mathbb{N}$, versehen mit der Euklidischen Norm als Metrik*

$$d(\mathbf{x}, \mathbf{y}) := \|\mathbf{x} - \mathbf{y}\|, \quad \mathbf{x}, \mathbf{y} \in \mathbb{R}^n,$$

   bildet einen metrischen Raum.

2. *Im $\mathbb{R}^n$ bildet*

$$d(\mathbf{x}, \mathbf{y}) := \frac{\|\mathbf{x} - \mathbf{y}\|}{1 + \|\mathbf{x} - \mathbf{y}\|}$$

   ebenfalls eine Metrik.

3. *Sei X eine beliebige Menge, dann wird durch*

$$d(x,y) := \begin{cases} 1, & \textit{falls } x \neq y, \\ 0, & \textit{falls } x = 0, \end{cases}$$

eine Metrik auf X erklärt.

Das erste Beispiel legt folgende Verallgemeinerung von $\|\cdot\|$ nahe.

Definition 3.1.3. *Eine Funktion $\|\cdot\| : \mathbb{R}^n \to \mathbb{R}$ heißt eine* **Norm** *auf $\mathbb{R}^n$, falls für beliebige $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$ und $\alpha \in \mathbb{R}$ gilt*

1. $\|\mathbf{x}\| \geq 0$ *und* $\|\mathbf{x}\| = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$ *(Definitheit),*
2. $\|\alpha\mathbf{x}\| = |\alpha|\|\mathbf{x}\|$ *(Homogenität),*
3. $\|\mathbf{x}+\mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ *(Dreiecksungleichung).*

Beispiel 3.1.4.

1. *Der Vektorraum $\mathbb{R}^n$, $n \in \mathbb{N}$, versehen mit der Euklidischen Norm*

$$\|\mathbf{x}-\mathbf{y}\|_2 := \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2} \quad \mathbf{x}, \mathbf{y} \in \mathbb{R}^n,$$

bildet einen metrischen Raum.

2. *In $\mathbb{R}^n$ definieren*

$$\|\mathbf{x}-\mathbf{y}\|_1 := \sum_{i=1}^{n} |x_i - y_i|, \quad \mathbf{x}, \mathbf{y} \in \mathbb{R}^n,$$

und

$$\|\mathbf{x}-\mathbf{y}\|_\infty := \max_{1\leq i\leq n} |x_i - y_i|, \quad \mathbf{x}, \mathbf{y} \in \mathbb{R}^n,$$

ebenfalls Normen.

Definition 3.1.5 (Intervalle). *Unter einem offenen bzw. abgeschlossenem Intervall $(a,b) \subseteq \mathbb{R}$, bzw. $[a,b] \subseteq \mathbb{R}$, $a < b$, verstehen wir die Menge*

$$(a,b) := \{x \in \mathbb{R} : a < x < b\}$$

bzw.

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

Gelegentlich verwenden wir auch halb offene Intervalle

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

sowie unbeschränkte Intervalle

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

Definition 3.1.6.

1. *Sei (X, d) ein metrischer Raum und $r > 0$, $x \in X$. Dann heißt eine Menge*

$$U_r(x) := \{x \in X : d(x, y) < r\} \subseteq X$$

*eine r-***Umgebung des Punktes** *x, auch* **offene Kugel mit Radius** *r genannt.*

2. *In $\mathbb{R}^n$ heißt für $r > 0$ und $\mathbf{x} \in \mathbb{R}^n$ die Menge $K_r(\mathbf{x}) = \{\mathbf{y} \in \mathbb{R}^n : \|\mathbf{x} - \mathbf{y}\| < r\}$ eine* **offene Kugel** *mit Mittelpunkt $\mathbf{x}$ und Radius r.*

3. *Eine Menge $E \subseteq \mathbb{R}^n$ heißt* **konvex***, falls für alle $\mathbf{x}, \mathbf{y} \in E$ und $0 < \lambda < 1$ stets folgt*

$$\lambda\mathbf{x} + (1 - \lambda)\mathbf{y} \in E.$$

Beispiel 3.1.7. *Kugeln $K_r(\mathbf{x})$ sind konvexe Mengen. Denn für $\mathbf{y}, \mathbf{z} \in K_r(\mathbf{x})$ gilt $\|\mathbf{x} - \mathbf{y}\| < r$ und $\|\mathbf{x} - \mathbf{z}\| < r$. Daraus folgt nun für $\lambda \in (0, 1)$*

$$\begin{aligned}
\|\mathbf{x} - (\lambda\mathbf{y} + (1 - \lambda)\mathbf{z})\| &= \|\lambda\mathbf{x} + (1 - \lambda)x - \lambda\mathbf{y} - (1 - \lambda)\mathbf{z}\| \\
&\leq \|\lambda\mathbf{x} - \lambda\mathbf{y}\| + \|(1 - \lambda)(\mathbf{x} - \mathbf{z})\| \\
&= \lambda\|\mathbf{x} - \mathbf{y}\| + (1 - \lambda)\|\mathbf{x} - \mathbf{z}\| \\
&< \lambda r + (1 - \lambda)r = r.
\end{aligned}$$

Definition 3.1.8. *Sei (X, d) ein metrischer Raum und $E \subseteq X$.*

1. *Ein Punkt $p \in X$ heißt* **Häufungspunkt** *der Menge E, falls in jeder r-Umgebung $U_r(p)$ stets ein $q \in E$ existiert mit $p \neq q$, das heißt*

$$U_r(p) \cap (E \setminus \{x\}) \neq \emptyset \quad \forall r > 0.$$

2. *Ist $p \in E$ kein Häufungspunkt von E, dann heißt p ein* **isolierter Punkt** *von E.*

3. *Die Menge E heißt* **abgeschlossen** *in X, wenn jeder Häufungspunkt von E selbst in E liegt.*

4. *Die Menge*

$$\overline{E} := E \cup \{p \in X : p \text{ ist Häufungspunkt von } E\}$$

heißt der **Abschluss** *oder die* **Abschließung** *von E.*

5. *Ein Punkt $p \in E$ heißt ein* **innerer Punkt** *von E, falls es eine Umgebung $U_r(p)$, $r > 0$, gibt mit $U_r(p) \subseteq E$.*

6. *Die Menge E heißt* **offen** *in X, falls jeder Punkt $p \in E$ ein innerer Punkt von E ist.*

7. *Die Menge $E^\circ := \{p \in E : p$ ist innerer Punkt von $E\}$ heißt das* **Innere** *von E.*

8. *Die Menge $\partial E := \overline{E} \cap \overline{X \setminus E}$ heißt der* **Rand** *von E. Dies bedeutet, jede r-Umgebung $U_r(p) \subseteq X$ enthält sowohl innere Punkte $q \in E^\circ$ als auch "äußere" Punkte $q \in (X \setminus \overline{E})^\circ$.*

9. *Die Menge E heißt* **beschränkt**, *falls $r > 0$ und $p \in E$ existieren, so dass $E \subseteq U_r(p)$.*

10. *Die Menge E heißt eine in X* **dichte Menge**, *falls jeder Punkt $p \in X$ ein Häufungspunkt von E ist, das heißt, wenn gilt $\overline{E} = X$.*

Satz 3.1.9. *Jede r-Umgebung ist eine offene Menge.*

Beweis. Sei $U_r(p)$, $r > 0$, eine Umgebung um $p \in X$ und $q \in U_r(p)$ ein beliebiger Punkt darin. Dann existiert ein $h > 0$ mit $d(p,q) = r - h < r$. Somit gilt für beliebige Punkte x mit $d(q,x) < h$, das heißt $x \in U_h(q)$, dass

$$d(p,x) \leq d(p,q) + d(q,x) < r - h + h = r.$$

Folglich ist q ist ein innerer Punkt von $U_r(q)$. □

Satz 3.1.10. *Eine Menge $E \subseteq X$ ist genau dann offen, wenn ihr Komplement $X \setminus E$ abgeschlossen in X ist.*

Beweis. "$\Longleftarrow$" Sei $X \setminus E$ abgeschlossen und $x \in E$ beliebig. Dann ist $x \notin X \setminus E$ und somit kein Häufungspunkt dieser Menge. Das heißt, es existiert eine Umgebung $U_r(x)$ mit $U_r(x) \bigcap (X \setminus E) = \emptyset$. Damit ist $U_r(x) \subseteq E$ und somit $x \in E$ ein innerer Punkt von E.

"$\Longrightarrow$" Sei E offen und x ein Häufungspunkt von $X \setminus E$. Dann liegt in jeder Umgebung $U_r(x)$ ein $y \in X \setminus E$ mit $x \neq y$. In diesem Fall ist x kein innerer Punkt von E. Da E offen ist, d.h. nur aus inneren Punkten besteht, gilt $x \in X \setminus E$. □

Korollar 3.1.11. *Das Komplement einer Menge E ist genau dann offen, wenn E abgeschlossen ist.*

Proposition 3.1.12. *Sei (X,d) ein metrischer Raum. Falls $x \in X$ ein Häufungspunkt der Menge $E \subseteq X$ ist, so enthält jede Umgebung $U_r(x)$ unendlich viele Elemente von E.*

Beweis. Angenommen es existiert eine Umgebung $U_r(x)$, die endlich viele Elemente $p_1, p_2, \ldots, p_n \neq x$ von E enthält. Dann ist $r^\star := \min_{1 \leq i \leq n} d(x, p_i) > 0$ und $U_{r^\star}(x) \cap E = \{x\}$ im Widerspruch zur Annahme x ist Häufungspunkt von E. □

Korollar 3.1.13. *Jede endliche Menge ist abgeschlossen.*

Satz 3.1.14. *Sei (X,d) ein metrischer Raum.*

1. *Sei $\{G_i : i \in I\}$ eine Familie in X offener Mengen, dann ist die Vereinigungsmenge $\bigcup_{i\in I} G_i$ ebenfalls offen in X.*

2. *Für jede Familie $\{F_i : i \in I\}$ in X abgeschlossener Mengen ist die Schnittmenge $\bigcap_{i\in I} F_i$ wieder abgeschlossen in X.*

3. *Für jede* **endliche** *Familie $\{G_i : i = 1, 2, \ldots, N\}$ in X offener Mengen ist $\bigcap_{i=1}^{N} G_i = G_1 \cap G_2 \cap \ldots \cap G_N$ offen in X.*

4. *Für jede* **endliche** *Familie $\{F_i : i = 1, 2, \ldots, N\}$ in X abgeschlossener Mengen ist die Vereinigung $\bigcup_{i=1}^{N} F_i = F_1 \cup F_2 \cup \ldots \cup F_N$ ebenfalls abgeschlossen in X.*

Beweis.

1. Sei $G = \bigcup_i G_i$. Zu jedem $x \in G$ existiert ein Index $i \in I$ mit $x \in G_i$. Da $x \in G_i$ ein innerer Punkt in $G_i \subseteq G$ ist, ist x ebenfalls innerer Punkt in G. Somit besteht G nur aus inneren Punkten und ist daher offen.

2. Wir wenden die De Morgenschen Regeln an. Dazu setzen wir $X \setminus F_i := G_i$. Da für alle $i \in I$ die Menge G_i offen ist, folgt dass $X \setminus \bigcap_{i\in I} F_i = \bigcup_{i\in I}(X \setminus F_i)$ ebenfalls offen ist, d.h. $F = \bigcap_{i=I} F_i$ ist nach Satz 3.1.10 bzw. Korollar 3.1.11 abgeschlossen.

3. Sei $G = \bigcap_{i=1}^{N} G_i$ und $x \in G$ beliebig. Da für alle $i = 1, \ldots, N$ gilt dass $x \in G_i$, existieren Radien r_i mit $U_{r_i}(x) \subseteq G_i$, $i = 1, 2, \ldots, N$. Wir setzen $r = \min\{r_1, r_2, \ldots, r_N\} > 0$, dann ist $U_r(x) \subseteq U_{r_i}(x)$ für alle i. Somit folgt schließlich $U_r(x) \subseteq \bigcap_{i=1}^{N} U_{r_i}(x) \subseteq \bigcap_{i=1}^{N} G_i = G$.

4. Folgt durch Komplementbildung aus Punkt 3.

□

Bemerkung 3.1.15. *Die Voraussetzung "die Indexmenge ist endlich" ist für die dritte und vierte Aussage* **wesentlich**. *In den ersten beiden Aussagen kann dagegen die Indexmenge I beliebig gewählt werden, insbesondere darf I überabzählbar unendlich sein.*

Beispiel 3.1.16. *Für $k \in \mathbb{N}$ sei G_k das offene Intervall $\left(-\frac{1}{k}, \frac{1}{k}\right)$. Offensichtlich ist jedes G_k eine offene Menge in $\mathbb{R}$ und $G = \bigcup_{k\in\mathbb{N}} G_k = (-1, 1)$ ist offen in $\mathbb{R}$. Aber die Menge $F = \bigcap_{k\in\mathbb{N}} F_k = \{0\}$ ist nicht offen sondern abgeschlossen in $\mathbb{R}$.*

Satz 3.1.17. *Ist (X, d) ein metrischer und $E, F, G \subseteq X$, dann gelten die folgenden Aussagen.*

1. *Aus $E \subseteq F$ folgt $\overline{E} \subseteq \overline{F}$.*

2. *Ist F eine in X abgeschlossene Menge und $E \subseteq F$, so gilt auch $\overline{E} \subseteq F$.*

3. *Die Menge $\overline{E}$ ist abgeschlossen in X.*

4. *Es ist $E = \overline{E}$ genau dann, wenn E abgeschlossen ist.*

5. *$\overline{E}$ ist die kleinste abgeschlossene Menge aus X, die E enthält.*

Beweis.

1. Es sei $x \in \overline{E}$ beliebig, aber fest. Dann ist entweder $x \in E$ oder x ist ein Häufungspunkt von E. Im ersten Fall gilt $x \in F$, also auch $x \in \overline{E}$. Sei nun also x ein Häufungspunkt von E. Dann ist für alle r-Umgebungen von x der Schnitt $U_r(x) \cap (E \setminus \{x\})$ nicht leer. Wegen $E \subseteq G$ ist aber auch der Schnitt $U_r(x) \cap (F \setminus \{x\})$ nicht leer. Folglich ist x ein Häufungspunkt von F und damit $x \in \overline{F}$.

2. Die Behauptung folgt aus der ersten Aussage mit $F = \overline{F}$.

3. Eine Menge ist genau dann abgeschlossen, wenn ihr Komplement offen ist. Also genügt es zu zeigen, dass $X \setminus \overline{E}$ eine in X offene Menge ist. Sei hierzu $x \in X \setminus \overline{E}$ beliebig gewählt. Zuerst bemerken wir, dass $x \notin \overline{E}$ und somit kein Häufungspunkt von E ist. Es existiert also eine r-Umegbung von x mit $U_r(x) \cap E = \emptyset$, d.h. $E \subseteq X \setminus U_r(x)$. Da $U_r(x)$ in X offen ist, ist $X \setminus U_r(x)$ in X abgeschlossen, woraus mit der zweiten Aussage folgt, dass $\overline{E} \subseteq X \setminus U_r(x)$, also auch $U_r(x) \subseteq X \setminus \overline{E}$. Demnach ist x ein innerer Punkt von $X \setminus \overline{E}$ und folglich ist $X \setminus \overline{E}$ offen.

4. Für $x \notin \overline{E}$ ist $x \notin E$ und somit ist x auch kein Häufungspunkt von E. Damit existiert eine Umgebung $U_r(x)$ mit $U_r(x) \cap \overline{E} = \emptyset$. Daher ist x für alle $x \in X \setminus \overline{E}$ ein innerer Punkt von $X \setminus \overline{E}$. Also ist die Menge $X \setminus \overline{E}$ offen und daher $\overline{E}$ abgeschlossen.

5. Die Behauptung folgt unvermittels aus der zweiten und dritten Aussage.

□

3.2 Kompakte Mengen

Definition 3.2.1. *Unter einer* **offenen Überdeckung** *einer Menge E in einem metrischen Raum X verstehen wir eine Familie offener Mengen $\{G_i : i \in I, G_i \subseteq X \text{ offen}\}$, so dass*

$$E \subseteq \bigcup_{i \in I} G_i.$$

Definition 3.2.2. *Eine Teilmenge K eines metrischen Raumes heißt* **kompakt**, *falls jede offene Überdeckung von K eine endliche Teilmenge besitzt, die selbst wieder eine offene Überdeckung von K ist. In anderen Worten, zu jeder offenen Überdeckung $\{G_\alpha : \alpha \in I\}$ existieren Indizes $\alpha_1, \alpha_2, \ldots, \alpha_N \in I$ mit*

$$K \subseteq \bigcup_{i=1}^{N} G_{\alpha_i}.$$

Beispiel 3.2.3. *Jede endliche Teilmenge ist kompakt.*

Satz 3.2.4. *Sei $K \subseteq X$ eine kompakte Teilmenge eines metrischen Raumes (X, d). Dann ist K beschränkt und abgeschlossen.*

Beweis. Wir zeigen zuerst, dass K beschränkt ist. Zu jedem $x \in K$ definieren wir eine Umgebung $U_x := U_1(x) = \{y \in Y : d(x,y) < 1\}$. Dann wird K von den Mengen U_x, $x \in K$, überdeckt,

$$K \subseteq \bigcup_{x \in K} U_x.$$

Aufgrund der Kompaktheit von K existiert eine endliche Teilüberdeckung $\{U_{x_i} : i = 1, \ldots, n\}$ mit $K \subseteq \bigcup_{i=1}^n U_{x_i}$. Wir wählen $y \in K$ und definieren

$$\rho := \max\{d(x_i, y) : i = 1, \ldots, n\} < \infty.$$

Dann ist $U_{x_i} \subseteq U_{\rho+1}(y)$ für jedes x_i, $i = 1, \ldots, n$. Hieraus folgt wegen

$$K \subseteq \bigcup_{i=1}^n U_{x_i} \subseteq U_{\rho+1}(y)$$

die Behauptung.

Als nächstes zeigen wir die Abgeschlossenheit von K, indem wir beweisen, dass $X \setminus K$ eine in X offene Menge ist. Sei $y \in X \setminus K$ ein beliebiges Element im Komplement von K. Zu jedem $x \in X$ definieren wir zwei Umgebungen, mit $V_x = U_r(y)$ und $W_x = U_r(x)$ mit $r = r_x := \frac{1}{2}d(x,y)$. Dann gilt $y \in V_x$ und $x \in W_x$. Somit wird K von den Mengen W_x, $x \in K$, überdeckt

$$K \subseteq \bigcup_{x \in K} W_x.$$

Die Kompaktheit sichert die Existenz endlich vieler x_i, $i = 1, \ldots, n$ mit

$$K \subseteq \bigcup_{i=1}^n W_{x_i}.$$

Wir definieren nun $\varepsilon := \min\{r_i = \frac{1}{2}d(x_i, y) : 1 \leq i \leq n\}$ und $V := U_\varepsilon(y)$. Dann gilt $V \cap W_{x_i} = \emptyset$ für alle $i = 1, \ldots, n$. Wegen $V \cap K \subseteq V \cap \bigcup_{i=1}^n W = \emptyset$ ist y ein innerer Punkt von $X \setminus K$. $\square$

Satz 3.2.5. *Jede abgeschlossene Teilmenge $A \subseteq K$ einer kompakten Menge ist selbst kompakt.*

Beweis. Sei $\{V_\alpha, \alpha \in I\}$ eine beliebige offene Überdeckung von A, d.h. $A \subseteq \bigcup_{\alpha \in I} V_\alpha$. Wegen der Abgeschlossenheit von A ist mit Satz 3.1.10 die Komplementärmenge $X \setminus A$ ist offen in X. Dann liefern $(X \setminus A) \cup \bigcup_{\alpha \in I} V_\alpha = \bigcup_{\alpha \in I}(V_\alpha \cup X \setminus A)$ offene Überdeckungen der kompakten Menge K. Die Menge K kann wegen ihrer Kompaktheit mit endlich vielen $G_{\alpha_i} := \big(V_{\alpha_i} \cup (X \setminus A)\big)$, $i = 1, \ldots, n$ überdeckt werden,

$$K \subseteq \bigcup_{i=1}^n G_{\alpha_i} = \bigcup_{i=1}^n \big(V_{\alpha_i} \cup (X \setminus A)\big).$$

Wegen $A \subseteq K$ gilt

$$\begin{aligned} A = A \cap K = A \cap \bigcup_{i=1}^{n} G_{\alpha_i} &= A \cap \bigcup_{i=1}^{n} \big((V_{\alpha_i} \cup (X \setminus A)\big) \\ &= \bigcup_{i=1}^{n} (A \cap V_{\alpha_i}) = A \cap \bigcup_{i=1}^{n} V_{\alpha_i} \end{aligned}$$

und damit ist $A \subseteq \bigcup_{i=1}^{n} V_{\alpha_i}$. Folglich haben wir eine endliche Teilüberdeckung von A gefunden. □

Korollar 3.2.6. *Sei $K \subseteq X$ eine kompakte und $A \subseteq X$ eine in X abgeschlossene Menge. Dann ist $A \cap K$ kompakt.*

Beweis. Wegen Satz 3.2.4 ist K und damit nach Satz 3.1.14 auch $A \cap K$ abgeschlossen. Da $A \cap K \subseteq K$ ist diese Menge nach Satz 3.2.5 ebenfalls kompakt. □

Satz 3.2.7 (Zentriertes System kompakter Mengen). *Sei $\{K_\alpha : \alpha \in I\}$ eine Familie kompakter Mengen derart, dass zu jeder endlichen Teilfamilie $\{K_{\alpha_i} : i = 1, \ldots, n\}$ der Durchschnitt $\bigcap_{i=1}^{n} K_{\alpha_i} \neq \emptyset$. Dann ist ebenfalls $\bigcap_{\alpha \in I} K_\alpha \neq \emptyset$.*

Beweis. Sei $K_1 \in \{K_\alpha : \alpha \in I\}$ beliebig gewählt. Dann sind $G_\alpha := X \setminus K_\alpha$ für alle $\alpha \in I' := I \setminus \{1\}$ nach Satz 3.1.10 offene Mengen. Wir nehmen nun an, dass

$$K_1 \cap \bigcap_{\alpha \in I'} K_\alpha = \emptyset.$$

Dann gilt

$$K_1 \subseteq X \setminus \bigcap_{\alpha \in I'} K_\alpha = \bigcup_{\alpha \in I'} (X \setminus K_\alpha) = \bigcup_{\alpha \in I'} G_\alpha$$

d.h. $\{G_\alpha : \alpha \in I'\}$ ist eine offene Überdeckung von K_1. Da K_1 kompakt ist, existiert eine endliche Teilüberdeckung $\{G_{\alpha_i} : i = 1, \ldots, n\}$ mit $K \subseteq \bigcup_{i=1}^{n} G_{\alpha_i}$. Wir wenden jetzt die de Morganschen Regeln an und erhalten

$$\begin{aligned} \emptyset &= K_1 \cap \Big(X \setminus \bigcup_{i=1}^{n} G_{\alpha_i}\Big) \\ &= K_1 \cap \bigcap_{i=1}^{n} (X \cap G_{\alpha_i}) \\ &= K_1 \cap \bigcap_{i=1}^{n} K_{\alpha_\alpha}. \end{aligned}$$

Dies steht im Widerspruch zur Voraussetzung, dass jeder endliche Durchschnitt nichtleer ist. □

Korollar 3.2.8. *Seien $\emptyset \neq K_n \subseteq X$ kompakte Mengen mit $K_{n+1} \subseteq K_n$, $n \in \mathbb{N}$, dann ist $\bigcap_{n\in\mathbb{N}} K_n \neq \emptyset$.*

Satz 3.2.9 (Cantor). *Sei $E \subseteq K$ eine unendliche Teilmenge einer kompakten Menge K, dann besitzt E mindestens einen Häufungspunkt in K.*

Beweis. Angenommen es existiert kein Häufungspunkt $x \in K$ von E. Dann existiert zu jedem Punkt $x \in E$ eine Umgebung $U_x = U_r(x)$, $r = r_x > 0$, derart, dass $U_x \cap E = \{x\}$. Ferner existiert zu jedem Punkt $x \in K \setminus E$ eine Umgebung $U_x = U_r(x)$, $r = r_x > 0$, derart, dass $U_x \cap E = \emptyset$. Dann bildet wegen $K \subseteq \bigcup_{x\in K} U_x$ die Familie $\mathcal{U} := \{U_x, x \in K\}$ eine Überdeckung von K. Jeder Punkt $x \in E$ liegt in genau einer Umgebung $U_x \in \mathcal{U}$. Daher kann $\mathcal{U}$ keine endliche Teilüberdeckung besizen, da E unendlich ist. Dies widerspricht natürlich der Kompaktheit von K. □

Definition 3.2.10. *Seien $-\infty < a_i < b_i < \infty$, $i = 1, \ldots, k$, dann heißen die Produktmengen*

$$\begin{aligned} I &= \prod_{i=1}^{k} [a_i, b_i] := [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_k, b_k] \\ &= \{\mathbf{x} = (x_1, \ldots, x_k) : a_i \leq x_i \leq b_i, i = 1, \ldots, k\} \subseteq \mathbb{R}^k \end{aligned}$$

k–Zellen.

Lemma 3.2.11.

1. *Seien $\emptyset \neq I_l := [a_l, b_l] \subseteq \mathbb{R}$, $l \in \mathbb{N}$, abgeschlossene Intervalle mit $I_{l+1} \subseteq I_l$. Dann ist $\bigcap_{i=1}^{\infty} I_l \neq \emptyset$ nichtleer.*

2. *Seien $I_l = \prod_{i=1}^{k} [a_{i,l}, b_{i,l}] \subseteq \mathbb{R}^k$, $l \in \mathbb{N}$, abgeschlossene k–Zellen mit $I_{l+1} \subseteq I_l$. Dann ist $\bigcap_{l=1}^{\infty} I_l \neq \emptyset$.*

Beweis.

1. Sei $M := \{a_l, l \in \mathbb{N}\}$, dann ist M nach oben beschränkt durch alle b_l, $l \in \mathbb{N}$. Somit ist $a_l \leq a := \sup M \leq b_l$ für alle $l \in \mathbb{N}$. Dies impliziert $a \in \bigcap_{l\in\mathbb{N}} [a_k, b_k]$.

2. Analog definieren wir zu jedem $i = 1, \ldots, k$ $a_i := \sup\{a_{i,l}, l \in \mathbb{N}\}$. Dann ist $\mathbf{a} = (a_1, \ldots, a_k) \in I_l$ für alle $l \in \mathbb{N}$ auch im Schnitt $\mathbf{a} \in \bigcap_{l=1}^{\infty} I^l$.

□

Lemma 3.2.12.

1. *Jedes abgeschlossene und beschränkte Intervall $[a, b]$ mit $-\infty < a < b < \infty$ ist kompakt.*

2. *Jede nichtleere, beschränkte k–Zelle ist kompakt.*

Beweis. 1. Sei $\delta := b - a$ die Länge des Intervalls. Dann gilt für alle $x, y \in I = [a, b]$ dass $|x - y| \leq \delta$. Wir nehmen nun an, dass $I = I_0 = [a, b]$ nicht kompakt sei. Dann gibt es eine Überdeckung $\{G_\alpha, \alpha \in J\}$ von I_0, die keine endliche Teilüberdeckung enthält. Wir zerlegen I_0 in zwei gleichgroße Teilintervalle

$$I_0 = [a, b] = \left[a, \frac{a+b}{2}\right] \cup \left[\frac{a+b}{2}, b\right] =: I_{1,l} \cup I_{1,r}.$$

Nun existiert, sagen wir ohne Beschränkung der Allgemeinheit, zu $I_1 := I_{1,l}$ keine endliche Teilüberdeckung. Wir teilen nun I_1 wiederum in zwei gleichgroße Teilintervalle $I_1 = I_{2,l} \cup I_{2,r}$ und finden unter diesen ein $I_2 \subseteq I_1$, zu dem ebenfalls keine endliche Teilüberdeckung existiert. Diese Vorgehensweise setzen wir fort und erhalten damit eine Familie abgeschlossener Teilintervalle I_l, $l \in \mathbb{N}_0$, mit den folgenden Eigenschaften:

- $I_{l+1} \subseteq I_l \subseteq I_0$ für alle $l \in \mathbb{N}_0$.
- Es gilt $|x - y| \leq 2^{-l}\delta$ für alle $x, y \in I_l$.
- Die Intervalle I_l besitzen alle keine endliche Teilüberdeckung.

Aufgrund von Lemma 3.2.11 existiert ein $x^\star \in I$ mit $x^\star \in \bigcap_{l=0}^{\infty} I_l$. In der Überdeckung $\{G_\alpha : \alpha \in J\}$ existiert eine offene Menge G_α mit $x^\star \in G_\alpha$. Nun ist G_α offen und somit ist $x^\star$ ein innerer Punkt von G_α, das heißt, es existiert ein $r > 0$ mit $U_r(x) \subseteq G_\alpha$.

Wir wählen $n \in \mathbb{N}$, so dass $2^{-n}\delta < r$ ist. Da aufgrund unserer Konstruktion die Intervalle I_l die Länge $2^{-l} \cdot \delta$ haben, gilt für alle $x, y \in I_n$

$$|x - y| \leq 2^{-n}\delta < r.$$

Wegen $x^\star \in I_n$ ist $I_n \subseteq G_\alpha$. Dies bedeutet I_n wird von einer einzigen offenen Menge überdeckt. Dies steht aber im Widerspruch zur Konstruktion der Intervalle I_n, da diese alle keine endliche Teilüberdeckung besitzen.

2. Im k–dimensionalen Fall führen wir den Beweis analog. Dazu benötigen wir nur die folgende Modifikation. Wir setzen

$$\delta := \left(\sum_{i=1}^{k} |b_i - a_i|^2\right)^{\frac{1}{2}},$$

so dass $\|\mathbf{x} - \mathbf{y}\|_2 \leq \delta$ für beliebige Vektoren $\mathbf{x}, \mathbf{y} \in I$ gilt. Wir zerteilen I in 2^k Teilzellen, wie zum Beispiel $I_{1,l} = \prod_{i=1}^{k} \left[a_i, \frac{a_i+b_i}{2}\right]$. Dann folgt die Behauptung mittels der gleichen Argumentation wie in oben.

□

Satz 3.2.13 (Heine–Borel). *Sei $\emptyset \neq E \subseteq \mathbb{R}^k$ und $d(\cdot, \cdot)$ die kanonische Metrik in $\mathbb{R}^k$,*

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

Dann sind die drei folgenden Eigenschaften alle äquivalent:

1. *E ist abgeschlossen und beschränkt.*
2. *E ist kompakt.*
3. *Jede unendliche Teilmenge $A \subseteq E$ von E hat einen Häufungspunkt in E.*

Beweis. 1. $\Longrightarrow$ 2.: Zu E existiert aufgrund der Beschränktheit eine Kugel $U_R(\mathbf{x})$ mit $E \subseteq U_R(\mathbf{x})$ und somit wiederum eine Zelle I mit $E \subseteq U_R(\mathbf{x}) \subseteq I$. Lemma 3.2.12 besagt, dass I kompakt ist. Da $E \subseteq I$ abgeschlossen in $\mathbb{R}^k$ ist, ist wegen Satz 3.2.5 E kompakt.

2. $\Longrightarrow$ 3.: Dies ist die Aussage von Satz 3.2.9.

3. $\Longrightarrow$ 1.:

- Wir nehmen an, dass E nicht beschränkt sei. In diesem Fall existieren $\mathbf{x}_n \in E$ mit $\|\mathbf{x}_n\| > n$, $n \in \mathbb{N}$. Die Mengen $M = \{\mathbf{x}_n : n \in \mathbb{N}\}$ ist unendlich und besitzt keinen Häufungspunkt. Denn in keiner endlichen Umgebung liegen unendlich viele Punkte von M.
- Angenommen, es sei E nicht abgeschlossen. Dann existiert ein Häufungspunkt $\mathbf{x} \in \mathbb{R}^k \setminus E$. Damit gibt es für jedes $n \in \mathbb{N}$ ein $\mathbf{x}_n \in E$ mit $\|\mathbf{x} - \mathbf{x}_n\| < \frac{1}{n}$. Die Menge $M := \{\mathbf{x}_n : n \in \mathbb{N}\} \subseteq E$ ist eine unendliche Teilmenge von E, die wegen der Voraussetzung einen Häufungspunkt $\mathbf{x}^\star \in E$ besitzt. Sei nun $n \in \mathbb{N}$ so gewählt, dass $\frac{1}{n} < \frac{1}{2}\|\mathbf{x} - \mathbf{x}^\star\|$ gilt. Hieraus folgt für alls $l \geq n$
$$\begin{aligned}\|\mathbf{x}_l - \mathbf{x}^\star\| &\geq \big|\|\mathbf{x} - \mathbf{x}^\star\| - \|\mathbf{x} - \mathbf{x}_l\|\big| \\ &\geq \left|\|\mathbf{x} - \mathbf{x}^\star\| - \frac{1}{l}\right| \\ &\geq \frac{1}{2}\|\mathbf{x} - \mathbf{x}^\star\| > 0\end{aligned}$$
im Widerspruch zur Definition $\mathbf{x}^\star$ ist Häufungspunkt von M.

□

Bemerkung 3.2.14. *Die Voraussetzung, dass die Dimension des Raumes $X = \mathbb{R}^k$ endlich ist, ist wesentlich für die Äquivalenz "K abgeschlossen und beschränkt $\Longleftrightarrow$ K ist kompakt". In allgemeinen metrischen Räumen (X, d) folgt nicht, dass jede abgeschlossene und beschränkte Menge auch kompakt ist. Der Nachweis der Kompaktheit ist dann i.a. schwieriger.*

Als Korollar können wir jetzt den Satz von Bolzano–Weierstraß formulieren.

Satz 3.2.15 (Bolzano–Weierstraß). *Jede beschränkte und unendliche Teilmenge E des $(\mathbb{R}^k, d)$ besitzt mindestens einen Häufungspunkt in $\mathbb{R}^k$.*

Beweis. Zu E existiert eine hinreichend große k–Zelle I mit $E \subseteq I$. Da I kompakt ist, folgt die Aussage unmittelbar aus Satz 3.2.13. □

3.3 Perfekte Mengen

Definition 3.3.1. *Sei (X, d) ein metrischer Raum. Die Menge $P \subseteq X$ heißt* **perfekt**, *falls P in X abgeschlossen ist und jeder Punkt von P ein Häufungspunkt von P ist.*

Satz 3.3.2. *Sei $\emptyset \neq P \subseteq \mathbb{R}^k$ eine perfekte Menge. Dann ist P überabzählbar.*

Beweis. Wir bemerken, da P Häufungspunkte besitzt, muss P unendlich sein. Angenommen, P sei abzählbar, d.h. $P = \{\mathbf{x}_n : n \in \mathbb{N}\}$ mit $\mathbf{x}_n \neq \mathbf{x}_m$, $n \neq m$. Weil alle Elemente $\mathbf{x}$ einer perfekten Menge Häufungspunkte sind, existieren Umgebungen $V_n = U_{r_n}(\mathbf{x}_n)$ mit

1. $\overline{V_{n+1}} \subseteq V_n$,
2. $\mathbf{x}_n \notin \overline{V_{n+1}}$,
3. $V_{n+1} \cap P$ ist nichtleer.

Wir setzen $K_n := \overline{V_n} \cap P$. Da $\overline{V_n}$ abgeschlossen und beschränkt ist, sind $\overline{V_n}$ und K_n kompakt. Da $\mathbf{x}_n \notin K_{n+1}$, liegt kein Punkt von P in $\bigcap_{n=1}^{\infty} K_n$. Wegen $K_n \subseteq P$ ist daher $\bigcap_{n=1}^{\infty} K_n = \emptyset$ im Widerspruch zu Korollar 3.2.8. □

Korollar 3.3.3. *Jedes Intervall $[a, b]$ mit $a < b$ ist überabzählbar.*

Beweis. Die Menge $I := [a, b]$ ist perfekt. □

3.4 Zusatz: Relative Offenheit und Kompaktheit

Definition 3.4.1. *Es sei $E \subseteq Y \subseteq X$ und (X, d) ein metrischer Raum. Die Menge E heißt* **relativ offen** *in Y, falls zu jedem Punkt $p \in X$ ein $r > 0$ existiert mit*

$$\{y \in Y : d(p, y) < r\} \subseteq E.$$

Beispiel 3.4.2. *Für $a, b \in \mathbb{R}$ ist das Intervall $I = (a, b)$ eine offene Menge in $\mathbb{R}$, aber $I = \{z \in \mathbb{C} : z = x + i \cdot 0, x \in (a, b)\}$ ist nicht offen in $\mathbb{C}$.*

Satz 3.4.3. *Sei $Y \subseteq X$. Eine Teilmenge $E \subseteq Y$ ist genau dann relativ offen in Y, wenn eine offene Menge $G \subseteq Y$ existiert mit*

$$E = Y \cap G.$$

Beweis. Sei E relativ offen in Y. Dann existiert zu jedem $x \in E$ ein $r_x > 0$ mit

$$\{y \in Y : d(x, y) < r_x\} \subseteq E.$$

Wir definieren die offenen Kugeln

$$K_x = \{y \in X : d(x,y) < r\}.$$

Es gilt wegen Satz 3.1.14, dass die Menge $G = \bigcup_{x \in E} K_x$ offen ist. Aufgrund der Voraussetzung gilt $K_x \cap Y \subseteq E$, woraus sich

$$Y \cap G = Y \cap \bigcup_{x \in E} K_x = \bigcup_{x \in E} (Y \cap K_x) \subseteq E$$

ergibt. Da für jedes $x \in E \subseteq Y$ gilt $x \in K_x$, ist

$$E \subseteq G \cap Y.$$

Sei nun umgekehrt G eine offene Menge in X mit $E = Y \cap G$, so besitzt jedes $x \in E$ eine offene Umgebung $U_x \subseteq G$ und es gilt weiter

$$U_x \cap Y \subseteq Y \cap G = E$$

d.h. E ist relativ offen. □

Definition 3.4.4. *Es sei $E \subseteq Y \subseteq X$ und (X,d) ein metrischer Raum. Die Menge E heißt* **relativ kompakt** *in Y, falls jede in Y relativ offene Überdeckung eine endliche Teilüberdeckung besitzt.*

Satz 3.4.5. *Sei $K \subseteq Y \subseteq X$. Dann ist die Menge K genau dann kompakt in X, wenn sie relativ kompakt in Y ist.*

Beweis. Sei $K \subseteq X$ eine kompakte Menge. Zu jeder beliebigen offenen Überdeckung $\{U_\alpha : \alpha \in I\}$ existiert eine endliche Teilüberdeckung

$$K \subseteq U_{\alpha_1} \cup \ldots \cup U_{\alpha_n}.$$

Wir setzen $V_{\alpha_i} = Y \cap U_{\alpha_i}$ bzw. $V_\alpha = Y \cap U_\alpha$. Da diese Mengen in Y relativ offen sind, erhalten wir wegen $K \subseteq Y$ die Aussage, dass zu jeder Überdeckung $\{V_\alpha : \alpha \in I\}$ ebenfalls $\{V_{\alpha_i}\}$ eine Überdeckung ist

$$K \subseteq V_{\alpha_1} \cup \ldots \cup V_{\alpha_n}.$$

Sei $K \subseteq Y$ relativ kompakt in Y und $\{G_{\alpha_i} : \alpha \in I\}$ eine in Y relativ offene Überdeckung mit endlicher Teilüberdeckung $G_{\alpha_1}, \ldots, G_{\alpha_n}$. Dann existieren aufgrund von Satz 3.4.3 zu jedem G_α eine in X offene Menge mit $V_\alpha = Y \cap G_\alpha$. Und wegen $V_\alpha \subseteq G_\alpha$ gilt $K = \bigcup_{i=1}^N V_{\alpha_i}$. □

Kapitel 4

Folgen und Reihen

4.1 Allgemeine Zahlenfolgen

Definition 4.1.1. *Eine Folge* $(a_k)_{k\in\mathbb{N}}$ *in* X *ist eine Abbildung* $f : \mathbb{N} \to X$ *mit* $a_k := f(k)$.

Definition 4.1.2. *Sei* (X, d) *ein metrischer Raum, dann heißt die Folge* $(a_k)_{k\in\mathbb{N}}$ **konvergent**, *falls ein* **Grenzelement** *oder* **Grenzwert** $a \in X$ *existiert, so dass für alle* $\varepsilon > 0$ *ein* $N_0 \in \mathbb{N}$ *gibt mit*

$$d(a, a_k) < \varepsilon \quad \forall\, k > N_0.$$

Mit anderen Worten, es gilt $a_k \in U_\varepsilon(a)$ *für alle* $k > N_0$. *Wir schreiben dann* $a = \lim_{k\to\infty} a_k$. *Eine Folge, die nicht konvergent ist, nennen wir* **divergent**. *Eine Folge ist* **beschränkt**, *falls* $\{a_k : k \in \mathbb{N}\}$ *eine beschränkte Menge in* (X, d) *ist.*

Beispiel 4.1.3. *Sei* $X = \mathbb{R}$.

1. *Die stationäre Folge* $(a_k)_{k\in\mathbb{N}}$ *mit* $a_k = a \in \mathbb{R}$ *ist konvergent mit Grenzwert* $\lim_{k\to\infty} a_k = a$. *Man beachte:* a *ist kein Häufungspunkt der Menge* $\{a_k : k \in \mathbb{N}\} = \{a\}$.

2. *Die Folge* $(a_k)_{k\in\mathbb{N}}$ *mit* $a_k = k$ *ist unbeschränkt und folglich divergent.*

3. *Sei* $a_k = \frac{1}{k}$, *dann ist* $(a_k)_{k\in\mathbb{N}}$ *konvergent mit* $\lim_{k\to\infty} a_k = 0$. *Denn sei* $\varepsilon > 0$ *beliebig aber fest gewählt. Nach der Archimedischen Eigenschaft (Satz 2.4.2) existiert ein* $N_0 \in \mathbb{N}$ *mit* $N_0 > 1/\varepsilon$. *Somit gilt für alle* $k > N_0$

$$0 < |a_k - 0| = \frac{1}{k} < \frac{1}{N_0} < \varepsilon.$$

   In diesem Fall ist 0 *der einzige Häufungspunkt der Menge* $\{a_k : k \in \mathbb{N}\}$.

4. *Für* $a_k = (-1)^k(1 - \frac{1}{k})$ *ist die Folge* $(a_k)_{k\in\mathbb{N}}$ *divergent. Die Menge* $\{a_k : k \in \mathbb{N}\}$ *besitzt allerdings zwei Häufungspunkte, nämlich* $+1$ *und* -1.

Satz 4.1.4. *Jede konvergente Folge $(a_k)_{k\in\mathbb{N}}$, $a_k \in (X,d)$, ist auch beschränkt.*

Beweis. Sei a das Grenzelement und $\varepsilon = 1$, dann existiert ein N_0 so dass $d(a_k, a) < \varepsilon = 1$ für alle $k > \mathbb{N}_0$ gilt. Setzen wir

$$R = \max\left(\{d(a_k, a) : k = 1, \ldots, N_0\} \cup \{1\}\right)$$

so folgt $d(a_k, a) < R + 1$ für alle $k \in \mathbb{N}$, d.h. $\{a_k : k \in \mathbb{N}\} \subseteq U_{R+1}(a)$. □

Satz 4.1.5. *Sei $(a_k)_{n\in\mathbb{N}}$ eine Folge in (X,d).*

1. *Die Folge $(a_k)_{k\in\mathbb{N}}$ konvergiert genau dann gegen $a \in X$, wenn außerhalb jeder Umgebung $U_\varepsilon(a)$ nur endlich viele Folgenglieder a_k liegen.*

2. *Der Grenzwert einer konvergenten Folge ist eindeutig.*

3. *Sei $E \subseteq X$ und $a \in X$ ein Häufungspunkt von E, dann existiert eine konvergente Folge $(a_k)_{k\in\mathbb{N}}$ mit $a \neq a_k \in E$ für alle $k \in \mathbb{N}$ und $\lim_{k\to\infty} a_k = a$.*

Beweis.

1. "$\Rightarrow$" ist klar.
   "$\Leftarrow$": Sei $U_\varepsilon(a)$ eine beliebige Umgebung von a. Nun sei $N_0 = \max\{k \in \mathbb{N} : a_k \notin U_\varepsilon(a)\}$. Dann gilt für alle $k > N_0$, dass $d(a, a_k) < \varepsilon$.

2. Seien a und a' zwei Grenzwerte einer konvergenten Folge $(a_k)_{k\in\mathbb{N}}$. Angenommen es gelte $a \neq a'$, dann ist $\varepsilon = d(a, a') > 0$ und es existiert ein $N_0 \in \mathbb{N}$, so dass sowohl $d(a_k, a) < \frac{\varepsilon}{2}$ als auch $d(a', a_k) < \frac{\varepsilon}{2}$ gilt für alle $k > N_0$. Damit folgt aber der Widerspruch

$$\varepsilon = d(a, a') \leq d(a, a_k) + d(a_k, a') \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

3. Sei a ein Häufungspunkt von E, dann existiert zu jedem $k \in \mathbb{N}$ ein $a_k \in E \setminus \{a\}$ mit $a_k \in U_{1/k}(a)$, d.h. $d(a, a_k) < \frac{1}{k}$. Die hierdurch definierte Folge $(a_k)_{k\in\mathbb{N}}$ konvergiert gegen a. Denn zu jedem $\varepsilon > 0$ existiert ein $N_0 \in \mathbb{N}$ mit $N_0 > 1/\varepsilon$ und damit gilt für alle $k > N_0$ dass $|d(a, a_k)| < \frac{1}{k} < \frac{1}{N_0} < \varepsilon$.

□

Definition 4.1.6. *Seien $p_k \in \mathbb{N}$, $k \in \mathbb{N}$, beliebig gewählt mit $p_k < p_{k+1}$. Dann heißt $(a_{p_k})_{k\in\mathbb{N}}$ eine* **Teilfolge** *von $(a_k)_{k\in\mathbb{N}}$.*

Es ist klar, dass für konvergente Folgen mit Grenzelement a auch alle Teilfolgen gegen das gleiche Grenzelement konvergieren. Umgekehrt kann eine divergente Folge jedoch konvergente Teilfolgen haben, man betrachte beispielsweise die Folge $(a_k)_{k\in\mathbb{N}}$ mit $a_k = (-1)^k(1 - \frac{1}{k})$. Falls alle Teilfolgen einer Folge gegen das gleiche Grenzelement a konvergieren, ist die Folge selbst konvergent gegen a.

4.2 Reelle Zahlenfolgen

Satz 4.2.1 (Dreifolgensatz). *Seien a_k, x_k, b_k, $k \in \mathbb{N}$, reelle Zahlen mit*

$$a_k \leq x_k \leq b_k \quad \textit{für alle } k \in \mathbb{N},$$

und seien die Folgen $(a_k)_{k\in\mathbb{N}}$ und $(b_k)_{k\in\mathbb{N}}$ beide konvergent mit

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

Dann konvergiert die Folge $(x_k)_{k\in\mathbb{N}}$ ebenfalls gegen x.

Beweis. Sei $\varepsilon > 0$ beliebig aber fest gewählt. Dann gibt es ein $N_0 \in \mathbb{N}$ mit $d(x, a_k) = |x - a_k| < \varepsilon$ und $|x - b_k| < \varepsilon$ für alle $k > N_0$. Es gilt also

$$-\varepsilon < a_k - x < \varepsilon, \quad -\varepsilon < b_k - x < \varepsilon.$$

für alle $k > N_0$. Wegen $a_k \leq x_k \leq b_k$ folgt hieraus

$$-\varepsilon < a_k - x < x_k - x < b_k - x < \varepsilon,$$

d.h. $|x_k - x| < \varepsilon$ für alle $k > N_0$. □

Definition 4.2.2. *Eine reelle Folge $(a_k)_{k\in\mathbb{N}}$ heißt* **monoton wachsend**, *falls $a_k \leq a_{k+1}$ für alle $k \in \mathbb{N}$ gilt. Sie heißt* **monoton fallend**, *falls $a_{k+1} \leq a_k$ für alle $k \in \mathbb{N}$ gilt. Man spricht von* **strenger Monotonie**, *falls jeweils die strenge Ungleichung gilt.*

Satz 4.2.3 (Monotoniekriterium). *Eine monotone und beschränkte reelle Folge $(a_k)_{k\in\mathbb{N}}$ ist konvergent.*

Beweis. Ohne Beschränkung der Allgemeinheit sei $(a_k)_{k\in\mathbb{N}}$ monoton wachsend. Da die Menge $M = \{a_k : k \in \mathbb{N}\}$ nach Voraussetzung eine beschränkte Teilmenge in $\mathbb{R}$ ist, existiert $a := \sup M$. Wir zeigen, dass $a = \lim_{k\to\infty} a_k$.

Aufgrund der Definition des Supremums existiert für beliebiges $\varepsilon > 0$ ein $a_{N_0} \in M$ mit

$$a - \varepsilon < a_{N_0} \leq a.$$

Wegen der Monotonie folgt hieraus für alle $k > N_0$, dass $a_k \leq a$ und

$$a - \varepsilon < a_{N_0} \leq a_k \leq a$$

für alle $k > N_0$ bzw. $|a_k - a| < \varepsilon$. □

Satz 4.2.4 (Limesregeln). *Seien $(a_k)_{k\in\mathbb{N}}$ und $(b_k)_{k\in\mathbb{N}}$* **konvergente** *reelle Folgen mit $\lim_{k\to\infty} a_k = a$ und $\lim_{k\to\infty} b_k = b$. Dann gelten die folgenden Rechenregeln:*

1. *Die Folge* $(a_k + b_k)_{k\in\mathbb{N}}$ *ist ebenfalls konvergent mit*
$$\lim_{k\to\infty}(a_k + b_k) = \lim_{k\to\infty} a_k + \lim_{k\to\infty} b_k = a + b.$$

2. *Für beliebige* $\alpha, \beta \in \mathbb{R}$ *konvergieren* $(a_k + \alpha)_{k\in\mathbb{N}}$ *und* $(\beta a_k)_{k\in\mathbb{N}}$ *mit*
$$\lim_{k\to\infty}(a_k + \alpha) = a + \alpha$$
und
$$\lim_{k\to\infty}(\beta a_k) = \beta a.$$

3. *Die Folge* $(a_k \cdot b_k)$ *konvergiert gegen*
$$\lim_{k\to\infty} a_k \cdot b_k = \lim_{k\to\infty} a_k \cdot \lim_{k\to\infty} b_k = a \cdot b.$$

4. *Existiert ein* $c > 0$*, so dass* $|a_k| > c$ *für alle* $k \in \mathbb{N}$*, so konvergiert die Folge* $(\frac{1}{a_k})_{k\in\mathbb{N}}$ *gegen*
$$\lim_{k\to\infty} \frac{1}{a_k} = \frac{1}{\lim_{k\to\infty} a_k} = \frac{1}{a}.$$

Beweis.

1. Zu jedem $\varepsilon > 0$ gibt es ein $N_0 \in \mathbb{N}$ mit $|a_k - a| < \frac{\varepsilon}{2}$ und $|b_k - b| < \frac{\varepsilon}{2}$ für alle $k > N_0$. Aus der Dreiecksungleichung folgt
$$|a_k + b_k - (a + b)| \leq |a_k - a| + |b_k - b| \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$
für alle $k > N_0$.

2. Diese Aussage ist trivial.

3. Wir zerlegen
$$\begin{aligned} a_k b_k - ab &= a_k b_k - a_k b + a_k b - ab \\ &= a_k(b_k - b) + (a_k - a)b. \end{aligned}$$
Die Folge $\big((a_k - a)b\big)_{k\in\mathbb{N}}$ konvergiert infolge von Punkt 1 und 2 gegen 0. Wir zeigen nun $\lim_{k\to\infty} a_k(b_k - b) = 0$. Sei hierzu $\varepsilon > 0$ beliebig aber fest gewählt. Dann existiert ein $N \in \mathbb{N}$, mit
$$|a_k - a| < \varepsilon,$$
d.h. $|a_k| < |a| + \varepsilon =: C$, für alle $k > N$. Ferner, gibt es ein $M \in \mathbb{N}$ mit
$$|b - b_k| < \frac{\varepsilon}{C}$$
für alle $k > M$. Setzen wir $N_0 = \max\{M, N\}$, so gilt demnach
$$|a_k(b_k - b) - 0| \leq C \cdot \frac{\varepsilon}{C} < \varepsilon.$$

4. Wir bemerken zuerst, dass $a \neq 0$, denn sonst haben wir für alle $0 < \varepsilon < c$ den Widerspruch

$$|a_k - a| = |a_k| > c > \varepsilon$$

für alle $k \in \mathbb{N}$.

Wir wählen nun ein $M \in \mathbb{N}$ derart, dass für alle $k > M$ gilt

$$|a - a_k| \leq \frac{1}{2}|a|.$$

Dann folgt $|a_k| \geq |a|/2$ für alle $k > M$. Ferner existiert zu beliebigen $\varepsilon > 0$ ein $N \in \mathbb{N}$ mit

$$|a_k - a| \leq \frac{1}{2}|a|^2\varepsilon$$

für alle $k \geq N$. Wählen wir nun $N_0 = \max\{M, N\}$, so erhalten wir für $k > N_0$ die Abschätzung

$$\left|\frac{1}{a} - \frac{1}{a_k}\right| = \left|\frac{a_k - a}{aa_k}\right| < \frac{\varepsilon|a|^2}{2|a_k||a|} \leq \varepsilon.$$

□

Satz 4.2.5. *1. Für alle $p > 0$ gilt*

$$\lim_{k\to\infty} \frac{1}{k^p} = 0.$$

2. Für alle $p > 0$ gilt

$$\lim_{k\to\infty} \sqrt[k]{p} = \lim_{k\to\infty} p^{\frac{1}{k}} = 1.$$

3. Es gilt

$$\lim_{k\to\infty} \sqrt[k]{k} = 1.$$

4. Für alle $p > 0$ und $\alpha \in \mathbb{R}$ ist

$$\lim_{k\to\infty} \frac{k^\alpha}{(1+p)^k} = 0.$$

5. Für alle $x \in \mathbb{R}$ mit $|x| < 1$ ist

$$\lim_{k\to\infty} x^k = 0.$$

Beweis.

1. Sei $\varepsilon > 0$ beliebig aber fest. Dann existiert ein $N_0 \in \mathbb{N}$ mit $N_0 > (\frac{1}{\varepsilon})^{\frac{1}{p}}$ und deshalb folgt für alle $k > N_0$ dass

$$\left|\frac{1}{k^p} - 0\right| \leq \frac{1}{N^p} < \varepsilon.$$

2. Im Falle $p = 1$ ist die Folge stationär und es ist nichts zu zeigen. Sei nun $p > 1$, dann betrachten wir $x_k := \sqrt[k]{p} - 1$. Die Bernoullische Ungleichung (Lemma 2.4.6) liefert die Abschätzung
$$1 + kx_k \leq (1 + x_k)^k = p.$$
Somit gilt
$$0 < x_k \leq \frac{p-1}{k}.$$
Da $\lim_{k\to\infty} \frac{p-1}{k} = (p-1) \lim_{k\to\infty} \frac{1}{k} = 0$ folgt mit dem Dreifolgensatz 4.2.1 $\lim_{k\to\infty} x_k = 0$ und hieraus die Behauptung.

   Falls $p < 1$, so setzen wir $q := \frac{1}{p} > 1$ und mit dem eben bewiesenen Resultat erhalten wir
$$1 = \lim_{k\to\infty} \sqrt[k]{q} = \lim_{k\to\infty} \sqrt[k]{\frac{1}{p}} = (\lim_{k\to\infty} \sqrt[k]{p})^{-1}.$$

3. Wir setzen $x_k := \sqrt[k]{k} - 1 \geq 0$ und wenden den binomischen Satz 2.4.11 an
$$k = (1 + x_k)^k \geq \binom{k}{2} x_k^2 = \frac{(k-1) \cdot k}{2} x_k^2.$$
Hieraus erhalten wir die Ungleichung $0 \leq x_k \leq \sqrt{\frac{2}{k-1}}$ für alle $k > 1$. Da $\lim_{k\to\infty} \sqrt{\frac{2}{k-1}} = 0$, folgt $\lim_{k\to\infty} x_k = 0$ mit Hilfe des Dreifolgensatzes 4.2.1.

4. Wir wählen ein $m \in \mathbb{N}$ derart, dass $\alpha < m$. Aus dem binomischen Satz 2.4.11 folgt für alle $k > 2m$ die Abschätzung
$$(1 + p)^k \geq \binom{k}{m} p^m = \frac{k(k-1)\dots(k-m+1)}{1 \cdot 2 \cdot \ldots \cdot m} p^m > \frac{k^m}{2^m m!} p^m.$$
Hieraus erhalten wir
$$0 < \frac{k^\alpha}{(1+p)^k} < \frac{2^m m!}{p^m} k^{(\alpha - m)}.$$
Wegen $k^{(\alpha-m)} \to 0$ für $k \to \infty$ gilt
$$\lim_{k\to\infty} \frac{2^m m!}{p^m} k^{(\alpha-m)}$$
und damit die Behauptung.

5. Diese Aussage folgt aus der vierten Aussage mit $\alpha = 0$.

□

Beispiel 4.2.6.

1. *Sei*
$$a_k = \frac{3k^2+1}{2k^2+2k+4} = \frac{k^2(3+\frac{1}{k^2})}{k^2(2+\frac{2}{k}+\frac{4}{k^2})}.$$
Dann ist
$$\lim_{k\to\infty} a_k = \frac{\lim_{k\to\infty}(3+\frac{1}{k^2})}{\lim_{k\to\infty}(2+\frac{2}{k}+\frac{4}{k^2})} = \frac{3}{2}$$

2. *Sei*
$$\begin{aligned} a_k = \sqrt{k+1}-\sqrt{k} &= \frac{(\sqrt{k+1}-\sqrt{k})(\sqrt{k+1}+\sqrt{k})}{\sqrt{k+1}+\sqrt{k}} \\ &= \frac{1}{\sqrt{k+1}+\sqrt{k}}, \end{aligned}$$
dann folgt
$$\lim_{k\to\infty} a_k = \lim_{k\to\infty} \frac{1}{\sqrt{k+1}+\sqrt{k}} = 0.$$

Seien $a_k \in \mathbb{R}$ für alle $k \in \mathbb{N}$. Wir schreiben $a_k \to \infty$ für $k \to \infty$, falls zu jedem $R > 0$ ein $N_0 \in \mathbb{N}$ existiert mit $a_k > R$ fü alle $k > N_0$. Analog definieren wir $a_k \to -\infty$.

Definition 4.2.7. *Sei $(a_k)_{k\in\mathbb{N}}$ eine reelle Folge und sei E die Menge aller Teilfolgengrenzwerte. Dann heißt*
$$\overline{s} = \sup E := \limsup_{k\to\infty} a_k$$
der **Limes superior** *und*
$$\underline{s} = \inf E =: \liminf_{k\to\infty} a_k$$
der **Limes inferior** *von $(a_k)_{k\in\mathbb{N}}$.*

Beispiel 4.2.8. *Für $k \in \mathbb{N}$ sei $a_k = (-1)^k(1-\frac{1}{k})$, dann ist $E = \{-1, 1\}$ und besteht aus den Häufungspunkten von $\{a_k : k \in \mathbb{N}\}$. Es gilt*
$$\overline{s} = \limsup_{k\to\infty} a_k = 1$$
und
$$\underline{s} = \liminf_{k\to\infty} a_k = -1.$$

Satz 4.2.9. *Sei $\mathbb{R}^n$ versehen mit der Euklidischen Norm $\|\cdot\| = \|\cdot\|_2$. Ferner seien $(\mathbf{a}_k)_{k\in\mathbb{N}}$ und $(\mathbf{b}_k)_{k\in\mathbb{N}}$ zwei Folgen in $\mathbb{R}^n$, d.h. $\mathbf{a}_k = (a_{k,1}, a_{k,2}, \ldots, a_{k,n})$ bzw. $\mathbf{b}_k = (b_{k,1}, b_{k,2}, \ldots, b_{k,n})$ mit $a_{i,k}, b_{i,k} \in \mathbb{R}$ für alle $i, k \in \mathbb{N}$. Dann gilt:*

1. *Die Folge $(\mathbf{a}_k)_{k\in\mathbb{N}}$ konvergiert genau dann gegen $\mathbf{a} = (a_1, a_2, \ldots, a_n) \in \mathbb{R}^n$, falls für alle $i = 1, 2, \ldots, n$ die Komponentenfolge $(a_{i,k})_{k\in\mathbb{N}}$ gegen a_i konvergiert.*

2. *Ist die Folge* $(\mathbf{a}_k)_{k\in\mathbb{N}}$ *kovergent mit* $\lim_{k\to\infty}\mathbf{a}_k = \mathbf{a}$ *und ist ferner durch* $(\alpha_k)_{k\in\mathbb{N}}$ *eine konvergente reelle Zahlenfolge mit Grenzwert* α *gegeben, dann folgt*

$$\lim_{k\to\infty}\alpha_k\mathbf{a}_k = \alpha\mathbf{a}.$$

3. *Sind die Folgen* $(\mathbf{a}_k)_{k\in\mathbb{N}}$ *und* $(\mathbf{b}_k)_{k\in\mathbb{N}}$ *beide konvergent mit*

$$\lim_{k\to\infty}\mathbf{a}_k = \mathbf{a}, \quad \lim_{k\to\infty}\mathbf{b}_k = \mathbf{b},$$

so gilt

$$\lim_{k\to\infty}(\mathbf{a}_k + \mathbf{b}_k) = \mathbf{a} + \mathbf{b},$$
$$\lim_{k\to\infty}\langle\mathbf{a}_k, \mathbf{b}_k\rangle = \langle\mathbf{a}, \mathbf{b}\rangle.$$

Beweis.

1. "$\Rightarrow$". Die Konvergenz der einzelnen Komponentenfolgen $(a_{i,k})_{k\in\mathbb{N}}$ folgt aus der Abschätzung

$$|a_i - a_{i,k}|^2 \leq \sum_{i=1}^{n}|a_i - a_{i,k}|^2 = \|\mathbf{a} - \mathbf{a}_k\|^2.$$

"$\Leftarrow$". Sei $\varepsilon > 0$ beliebig aber fest gewählt und $N_0 \in \mathbb{N}$ so dass

$$|a_{i,k} - a_i| < \frac{\varepsilon}{\sqrt{n}}, \quad k > N_0$$

für alle $i = 1, 2, \ldots, n$. Dann gilt

$$\|\mathbf{a}_k - a\| = \sqrt{\left(\sum_{i=1}^{n}|a_{i,k} - a_i|^2\right)} \leq \frac{\varepsilon}{\sqrt{n}}\cdot\sqrt{n} = \varepsilon.$$

2. Die Behauptung folgt aus der ersten Aussage und den Limesregeln (Satz 4.2.4).

3. Die Behauptung folgt ebenfalls aus der ersten Aussage und den Limesregeln.

□

4.3 Cauchy–Folgen

Definition 4.3.1. *Sei* $(a_k)_{k\in\mathbb{N}}$ *eine Folge in einem metrischen Raum* (X, d)*, d.h.* $a_k \in X$ *für alle* $k \in \mathbb{N}$*. Dann heißt die Folge eine* **Cauchy–Folge** *in* X*, falls zu jedem* $\varepsilon > 0$ *ein* $N_0 \in \mathbb{N}$ *existiert derart, dass für beliebige* $m, n > N_0$ *gilt*

$$d(a_m, a_n) < \varepsilon.$$

Satz 4.3.2. *Jede konvergente Folge $(a_k)_{k\in\mathbb{N}}$ ist eine Cauchy–Folge.*

Beweis. Sei $a \in X$ der Grenzwert von $(a_k)_{k\in\mathbb{N}}$, d.h. $\lim_{k\to\infty} a_k = a$. Zu beliebigem $\varepsilon > 0$ existiert ein $N_0 \in \mathbb{N}$ mit

$$d(a, a_n) < \frac{\varepsilon}{2}$$

für alle $n > N$. Folglich gilt für alle $m, n > N_0$

$$d(a_m, a_n) \leq d(a_m, a) + d(a, a_n) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

□

Bemerkung 4.3.3. *Man beachte, dass eine Cauchy–Folge $(a_k)_{k\in\mathbb{N}}$ nicht konvergent sein muss. Sei beispielsweise $(a_k)_{k\in\mathbb{N}}$ mit $a_k \in \mathbb{Q}$ eine Folge mit $\lim_{a_k\to k} = \sqrt{2}$. Dann ist $(a_k)_{k\in\mathbb{N}}$ eine Cauchy–Folge in $(\mathbb{Q}, |\cdot|)$, die jedoch in $\mathbb{Q}$ nicht konvergiert.*

Definition 4.3.4. *Ein metrischer Raum (X, d) heißt* **vollständig***, wenn jede Cauchy–Folge $(a_k)_{k\in\mathbb{N}}$ aus X einen Grenzwert in X besitzt.*

Satz 4.3.5. *Sei (X, d) ein vollständiger, metrischer Raum und sei $K \subseteq X$ eine abgeschlossene Teilmenge in X. Dann ist K (bzw. (K, d)) ebenfalls vollständig.*

Beweis. Sei $a_k \in K$ für alle $k \in \mathbb{N}$ und $(a_k)_{k\in\mathbb{N}}$ eine Cauchy–Folge in $K \subseteq X$. Dann existiert das Grenzelement $a \in X$, da X vollständig ist. Nun ist aber a ein Häufungspunkt der Menge $M = \{a_k : k \in \mathbb{N}\} \subseteq K$. Vorausgesetzt M ist unendlich, was wir ohne Beschränkung der Allgemeinheit annehmen dürfen, ist a auch ein Häufungspunkt von K. Da K abgeschlossen ist, liegt a in K. □

Definition 4.3.6. *Seien (X, d) ein metrischer Raum und $E \subseteq X$ eine beschränkte Menge, dann heißt das Supremum*

$$\operatorname{diam} E := \sup\{d(x, y) : x, y \in E\}$$

der Durchmesser von E.

Lemma 4.3.7. *Seien (X, d) ein metrischer Raum und $\overline{E}$ der Abschluß einer beschränkten Menge $E \subseteq X$. Dann gilt*

$$\operatorname{diam} \overline{E} = \operatorname{diam} E.$$

Beweis. Aus $E \subseteq \overline{E}$ folgt direkt

$$\operatorname{diam} E \leq \operatorname{diam} \overline{E}.$$

Sei nun $\varepsilon > 0$ beliebig aber fest. Wir wählen zwei Punkte $x, y \in \overline{E}$. Zu diesen gibt es ein $x', y' \in E$ mit $d(x, x') < \varepsilon$ und $d(y, y') < \varepsilon$. Wir verwenden die Dreiecksungleichung um $d(x, y)$ abzuschäutzen

$$d(x, y) \leq d(x, x') + d(x', y') + d(y', y) < 2\varepsilon + \operatorname{diam} E.$$

Hieraus folgt

$$\operatorname{diam} \overline{E} \leq 2\varepsilon + \operatorname{diam} \overline{E}$$

für alle $\varepsilon > 0$, und durch Grenzübergang $\varepsilon \to 0$ erhalten wir die Behauptung. □

Lemma 4.3.8. *Seien (X, d) ein metrischer Raum und $K_n \subseteq X$, $n \in \mathbb{N}$, kompakte Mengen mit $K_{n+1} \subseteq K_n$ und $\lim_{k\to\infty} \operatorname{diam} K_n = 0$. Dann besteht $\bigcap_{n=1}^{\infty} K_n$ aus genau einem Punkt.*

Beweis. Nach dem Satz über das zentrierte System kompakter Mengen 3.2.7 folgt $K := \bigcap_{n=1}^{\infty} K_n \neq \emptyset$. Enthält K mehr als einen Punkt, so ist $\operatorname{diam} K > 0$. Wegen $K \supseteq K_n \supseteq K_{n+1}$ gilt aber

$$0 \leq \operatorname{diam} K \leq \operatorname{diam} K_n \leq \operatorname{diam} K_{n+1} \to 0$$

für $n \to \infty$. □

Satz 4.3.9. *Sei (X, d) ein metrischer Raum und $K \subseteq X$ eine kompakte Teilmenge. Die Folge $(p_k)_{k\in\mathbb{N}}$ sei eine Cauchy–Folge in K. Dann konvergiert $(p_k)_{k\in\mathbb{N}}$ gegen ein Element in K, d.h. kompakte metrische Räume (K, d) sind vollständig.*

Beweis. Sei $(p_k)_{k\in\mathbb{N}}$ eine Cauchy–Folge und $P_n = \{p_k : k \geq n\}$, dann gilt $P_{n+1} \subseteq P_n$ und wegen $\operatorname{diam} P_n = \operatorname{diam} \overline{P}_n$ gilt $\operatorname{diam} \overline{P}_n \to 0$ für $n \to \infty$. Die Mengen $\overline{P}_n \subseteq K$ sind beschränkt und abgeschlossen, also nach Satz 3.2.5 auch kompakt. Wegen $\operatorname{diam} \overline{P}_n \to 0$ für $n \to \infty$ besteht nach Lemma 4.3.8 $P := \bigcap_{n\in\mathbb{N}} \overline{P}_n$ aus genau einem Element $p \in K$. Wir müssen noch zeigen, dass $(p_k)_{k\in\mathbb{N}}$ gegen p konvergiert. Sei $\varepsilon > 0$, dann existiert ein $N_0 \in \mathbb{N}$ mit $\operatorname{diam} \overline{P}_{N_0} < \varepsilon$. Da $p \in P_n$ für alle $n \in \mathbb{N}$, erhalten wir $d(p, p_k) < \varepsilon$ für alle $k > N_0$. □

Satz 4.3.10. *Sei (X, d) ein metrischer Raum und $(p_k)_{k\in\mathbb{N}}$ eine Cauchy–Folge in X. Dann ist die Folge $(p_k)_{k\in\mathbb{N}}$ in (X, d) beschränkt.*

Beweis. Wie im vorigen Satz definieren wir die Menge $P_n = \{p_k : k \geq n\}$. Wie wir schon bemerkt haben, folgt für eine Cauchy–Folge $(p_k)_{k\in\mathbb{N}}$ dann $\operatorname{diam} \overline{P_n} \to 0$ für $n \to \infty$, d.h. es existiert ein $N_0 \in \mathbb{N}$ derart, dass $\operatorname{diam} \overline{P_n} < 1$ für alle $n > N_0$. Setzen wir nun

$$R = 1 + \max\{d(p_n, p_{N_0}) : n = 0, 1, \ldots, N_0\},$$

dann gilt $d(p_n, p_{N_0}) < R$ für alle p_n, $n \in \mathbb{N}$. □

Satz 4.3.11. *Die Euklidischen Räume $\mathbb{R}^n$ versehen mit der kanonischen Metrik sind vollständig.*

Beweis. Sei $(\mathbf{p}_k)_{k\in\mathbb{N}}$ eine beliebige Cauchy–Folge in $\mathbb{R}^n$. Dann ist sie aufgrund des lezten Satzes auch beschränkt. Folglich existiert eine abgeschlossene und beschränkte Menge $K \subseteq \mathbb{R}^n$ mit $\mathbf{p}_k \in K$ für alle $k \in \mathbb{N}$. Aufgrund des Satzes von Heine-Borel 3.2.13 ist die Menge K auch kompakt. Nach Satz 4.3.9 konvergiert daher die Cauchy–Folge gegen ein Grenzelement $\mathbf{p} \in K \subseteq \mathbb{R}^n$. □

Satz 4.3.12. *Eine Menge $K \subseteq \mathbb{R}^n$ ist genau dann kompakt, falls jede Folge $(\mathbf{a}_k)_{k\in\mathbb{N}}$ mit $\mathbf{a}_k \in K$ eine in K konvergente Teilfolge $(\mathbf{a}_{p_k})_{k\in\mathbb{N}}$ enthält, d.h. $\lim_{k\to\infty} \mathbf{a}_{p_k} = \mathbf{a} \in K$.*

Beweis. "$\Rightarrow$": Sei K kompakt, so besitzt jede unendliche Teilmenge E einen Häufungspunkt in K (Satz von Cantor 3.2.9). Sei $E = \{\mathbf{a}_k : k \in \mathbb{N}\}$ und $\mathbf{a} \in K$ ein Häufungspunkt von E. Dann existiert zu jedem $k \in \mathbb{N}$ ein $\mathbf{a}_{p_k} \in E$ mit $\|\mathbf{a}_{p_k} - \mathbf{a}\| < \frac{1}{k}$. Folglich konvergiert $(\mathbf{a}_{p_k})_{k\in\mathbb{N}}$ gegen $\mathbf{a}$.

"$\Leftarrow$": Falls jede Folge eine konvergente Teilfolge besitzt, so ist K beschränkt. Denn anderenfalls existieren $\mathbf{a}_k \in K$ mit $\|\mathbf{a}_k\|_2 > k$. Und diese Folge kann keine konvergente Teilfolge besitzen. K ist aber auch abgeschlossen. Nun sei $\mathbf{a}$ ein Häufungspunkt von K, dann existieren $\mathbf{a}_k \in K$ mit $\|\mathbf{a} - \mathbf{a}_k\| < \frac{1}{k}$. Hieraus folgt, dass $\lim_{k\to\infty} \mathbf{a}_k = \mathbf{a}$ in K liegt. Mit dem Satz von Heine-Borel 3.2.13 folgt die Kompaktheit der Menge K. □

Bemerkung 4.3.13. *Die Aussage gilt sogar in allgemeinen metrischen Räumen. Allerdings ist der Beweis aufwendiger.*

Korollar 4.3.14 (Satz von Weierstraß). *Jede beschränkte Folge $(\mathbf{a}_k)_{k\in\mathbb{N}}$ mit $\mathbf{a}_k \in \mathbb{R}^n$ besitzt eine konvergente Teilfolge.*

4.4 Reihen

Definition 4.4.1. *Sei $(a_k)_{k\in\mathbb{N}}$ eine Folge und $n \in \mathbb{N}$. Dann heißt*

$$s_n = \sum_{k=1}^{n} a_k$$

die **Partialsummenfolge** *von $(a_k)_{k\in\mathbb{N}}$. Falls die Folge $(s_n)_{n\in\mathbb{N}}$ konvergiert mit $s = \lim_{n\to\infty} s_n$, so definieren wir*

$$\sum_{k=1}^{\infty} a_k := s = \lim_{n\to\infty} \sum_{k=1}^{n} a_k$$

und sprechen von der **konvergenten Reihe** $\sum_{k=1}^{\infty} a_k$. *Falls die Folge der Partialsummen $(s_n)_{n\in\mathbb{N}}$ divergiert, so heißt die Reihe $\sum_{k=1}^{\infty} a_k$ diveregent. Wir schreiben oft kurz $\sum a_k$ für die obige Reihe.*

Das Cauchy-Kriterium (Satz 4.3.2) kann hierfür folgendermaßen formuliert werden.

Satz 4.4.2. *Die Reihe $\sum a_k$ konvergiert genau dann, wenn für jedes $\varepsilon > 0$ eine natürliche Zahl N_0 existiert, so dass für alle $m \geq n > N_0$ gilt*

$$\left| \sum_{k=n}^{m} a_k \right| < \varepsilon.$$

Beweis. Für die Partialsummenfolge $(s_n)_{n\in\mathbb{N}}$ gilt

$$s_m - s_{n-1} = \sum_{k=n}^{m} a_k.$$

Somit folgt die Aussage unmittelbar aus Satz 4.3.2. □

Satz 4.4.3. *Falls die Reihe $\sum a_k$ konvergiert, so gilt $\lim_{k\to\infty} a_k = 0$.*

Beweis. Zu jedem $\varepsilon > 0$ exisitert nach Satz 4.4.2 ein $N_0 \in \mathbb{N}$, so dass

$$|s_n - s_{n-1}| = |a_n| < \varepsilon$$

für alle $n > N_0$. Das heißt, die Folge $(a_k)_{k\in\mathbb{N}}$ konvergiert gegen 0. □

Die Bedingung $a_k \to 0$ ist also notwendig für die Konvergenz der Reihe $\sum a_k$. Allerdings ist sie *nicht* hinreichend. Wie wir sehen werden, *divergiert* die Reihe $\sum_{k=1}^{\infty} \frac{1}{k}$, genauer $\sum_{k=1}^{n} \frac{1}{k} \to \infty$ für $n \to \infty$.

Satz 4.4.4. *Seien $a_k \geq 0$ für alle $k \in \mathbb{N}$, so konvergiert die Reihe $\sum a_k$ genau dann, wenn die Partialsummenfolge $(s_n) = \left(\sum_{k=1}^{n} a_k\right)_{n\in\mathbb{N}}$ beschränkt ist.*

Beweis. Die Partialsummenfolge ist hier monoton wachsend, da $s_{n+1} - s_n = a_{n+1} \geq 0$ für alle $n \in \mathbb{N}$ ist. Damit folgt die Behauptung aus dem Monotoniekriterium (Satz 4.2.3). □

Satz 4.4.5 (Majoranten–/Minorantenkriterium).

1. *Falls $|a_k| \leq c_k$ für alle $k \in \mathbb{N}$ gilt, und die Reihe $\sum c_k$ (Majorante) konvergiert, so konvergiert auch $\sum a_k$.*

2. *Falls $0 \leq d_k \leq a_k$, $k \in \mathbb{N}$ ist und die Reihe $\sum d_k$ (Minorante) divergiert, so divergiert auch $\sum a_k$.*

Beweis. 1. Sei $\varepsilon > 0$, so gibt es nach Satz 4.4.2 ein $N_0 \in \mathbb{N}$, so dass für alle $m \geq n > N$ gilt

$$\sum_{k=n}^{m} c_k < \varepsilon.$$

Hieraus folgt mit der Dreiecksungleichung

$$\left|\sum_{k=n}^{m} a_k\right| \leq \sum_{k=n}^{m} |a_k| \leq \sum_{k=n}^{m} c_k < \varepsilon,$$

woraus die Konvergenz folgt.

2. Angenommen $\sum a_k$ ist konvergent, so folgt wegen der Aussage auch die Konvergenz von $\sum d_k$ im Widerspruch zur Voraussetzung.

□

Satz 4.4.6 (Geometrische Reihe). *Die* **geometrische Reihe** $\sum_{k=0}^{\infty} x^k$ *ist für* $|x| < 1$ *konvergent und für* $|x| \geq 1$ *divergent. Falls* $|x| < 1$ *ist* $\sum_{k=0}^{\infty} x^k = \frac{1}{1-x}$.

Beweis. Für $x = \pm 1$ ist die Reihe divergent. Für $x \neq 1$ gilt

$$s_n - xs_n = \sum_{k=0}^{n} x^k - \sum_{k=1}^{n+1} x^k = 1 - x^{n+1},$$

d.h. $s_n = \frac{1-x^{n+1}}{1-x}$. Demnach divergiert (s_n) für $|x| > 1$. Im Falle $|x| < 1$ erhalten wir wegen $\lim_{n\to\infty} x^{n+1} = 0$ das Resultat

$$\sum_{k=0}^{\infty} x^k = \lim_{n\to\infty} s_n = \frac{1}{1-x}.$$

□

Satz 4.4.7 (Verdichtungskriterium). *Sei* $0 \leq a_{k+1} \leq a_k$ *für alle* $k \in \mathbb{N}$. *Dann konvergiert die Reihe* $\sum a_k$ *genau dann, wenn* $\sum_{k=1}^{\infty} 2^k a_{2^k}$ *konvergiert.*

Beweis. Es genügt wegen Satz 4.4.4 die Beschränktheit der Partialsummenfolgen

$$s_n = a_1 + a_2 + \ldots + a_n = \sum_{k=0}^{n} a_k$$

und

$$t_n = a_1 + 2a_2 + 4a_4 + \ldots + 2^n a_{2n} = \sum_{k=0}^{n} 2^k a_{2^k}$$

zu untersuchen. Für $n < 2^{k+1}$ folgt wegen $a_l \geq a_{l+1}$, $l \in \mathbb{N}$, dass

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

Ist die Folge $(t_k)_{k\in\mathbb{N}}$ beschränkt, so ist demnach auch $(s_n)_{n\in\mathbb{N}}$ beschränkt. Andererseits gilt für $n > 2^k$ dass

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

d.h. $2s_n \geq t_k$ für $n > 2^k$. Ist also $(t_k)_{k\in\mathbb{N}}$ unbeschränkt, so ist dies auch $(s_n)_{n\in\mathbb{N}}$ und umgekehrt, womit der Satz bewiesen ist. □

Proposition 4.4.8. *Die Reihe* $\sum_{k=1}^{\infty} \frac{1}{k^\alpha}$, $\alpha \in \mathbb{R}$, *ist divergent im Falle* $\alpha \leq 1$ *und konvergent im Falle* $\alpha < 1$.

Beweis. Für $\alpha \leq 0$ ist $a_k \geq 1$ und es ist nichts zu zeigen. Sei also $\alpha > 0$, dann gilt $0 \leq a_{k+1} \leq a_k$ und wir können das Verdichtungskriterium anwenden. Es ergibt sich

$$\sum_{k=1}^{\infty} 2^k a_{2^k} = \sum_{k=1}^{\infty} \frac{2^k}{(2^k)^\alpha} = \sum_{k=1}^{\infty} (2^{(1-\alpha)})^k,$$

also eine geometrische Reihe. Diese ist nach Satz 4.4.6 genau dann konvergent, wenn $2^{(1-\alpha)} < 1$, d.h. wenn $\alpha > 1$. □

Bemerkung 4.4.9. *Insbesondere erhalten wir im Falle* $\alpha = 1$ *die Aussage, dass die sogenannte* **harmonische Reihe** $\sum_{k=1}^{\infty} \frac{1}{k} = 1 + \frac{1}{2} + \frac{1}{3} + \ldots$ *divergiert.*

Proposition 4.4.10. *Die Reihe* $\sum_{k=0}^{\infty} \frac{x^k}{k!}$ *konvergiert für alle* $x \in \mathbb{R}$.

Beweis. Zu jedem $x \in \mathbb{R}$ existiert ein $N_0 \in \mathbb{N}$ mit $N_0! \geq (|x|+1)^{N_0}$. Ferner ist für alle $k > N_0$

$$|a_k| = \frac{|x|^k}{k!} \leq \frac{|x|^k}{(|x|+1)^k} = q^k$$

mit $0 \leq q := \frac{|x|}{(|x|+1)} < 1$. Da die Reihe $\sum q_k$ nach Staz 4.4.6 konvergent ist, folgt die Konvergenz von $\sum_{k=0}^{\infty} \frac{x^k}{k!}$ aus dem Majorantenkriterium 4.4.5. □

Definition 4.4.11. *Die Zahl*

$$e := \sum_{k=0}^{\infty} \frac{1}{k!}$$

heißt die **Eulersche Zahl**.

Satz 4.4.12. *Die Zahl* e *ist irrational.*

Beweis. Angenommen e ist rational, dann exisiteren teilerfremde natürliche Zahlen $p, q \in \mathbb{N}$ mit $e = \frac{p}{q}$. Sei $s_n = \sum_{k=0}^{n} \frac{1}{k!}$, so gilt

$$\begin{aligned} 0 < e - s_n &= \frac{1}{(n+1)!} + \frac{1}{(n+2)!} + \frac{1}{(n+3)!} + \ldots \\ &\leq \frac{1}{(n+1)!} \left(1 + \frac{1}{(n+1)} + \frac{1}{(n+1)^2} + \ldots \right). \end{aligned}$$

Wählen wir nun $n = q$, so folgt hieraus die Ungleichung

$$0 < e - s_n \leq \frac{1}{(q+1)!} \cdot \frac{1}{1 - \frac{1}{q+1}} = \frac{1}{q \cdot q!}.$$

Da $q! s_n = q!(1 + \frac{1}{1!} + \frac{1}{2!} + \ldots + \frac{1}{q!})$ und $q!e$ natürliche Zahl sind mit $q!e < q! s_n$, ist auch $q!(e - s_n) \in \mathbb{N}$ im Widerspruch zu $q!(e - s_n) \leq \frac{1}{q} < 1$. □

4.5 Das Wurzel– und Quotientenkriterium

Satz 4.5.1 (Wurzelkriterium). *Seien $a_k \in \mathbb{R}$, $k \in \mathbb{N}$, und $\sum_{k=1}^{\infty} a_k$ die zugehörige Reihe. Sei ferner*

$$q = \lim_{k\to\infty} \sup \sqrt[k]{|a_k|}.$$

Dann gilt:

1. *Falls $q < 1$, konvergieren die Reihen $\sum a_k$ und $\sum |a_k|$.*
2. *Falls $q > 1$, divergiert die Reihe $\sum a_k$.*

Beweis. 1. Sei $q < 1$, dann existiert ein $r \in \mathbb{R}$ mit $q < r < 1$ und es gilt $\sqrt[k]{|a_k|} < r$, d.h. $|a_k| < r^k$, für alle $k > N_0$. Die Reihe $\sum_{k=0}^{\infty} r^k = \frac{1}{1-r}$ ist damit eine konvergierende Majorante von $\sum_{k=0}^{\infty} a_k$. Es folgt somit die Konvergenz.

2. Falls $q > 1$ existiert ein $r \in \mathbb{R}$ mit $1 < r < q$ und eine Teilfolge (a_{k_n}) von (a_k), so dass $r^{k_n} < |a_{k_n}|$ für alle $n \geq N_0 \in \mathbb{N}$ gilt. Somit ist a_{k_n} kein Nullfolge und somit die Reihe $\sum_{k=1}^{\infty} a_k$ divergent.

□

Bemerkung 4.5.2. *Im Fall $q = \lim_{k\to\infty} \sup \sqrt[k]{|a_k|} = 1$ läßt sich* **keine** *Aussage über Konvergenz machen, wie man am Beispiel der Reihe $\sum \frac{1}{k^\alpha}$, $\alpha > 0$, sieht.*

Satz 4.5.3 (Quotientenkriterium).

1. *Falls $\lim_{k\to\infty} \sup \left|\frac{a_{k+1}}{a_k}\right| < 1$, so konvergieren die Reihen $\sum a_k$ und $\sum |a_k|$.*
2. *Falls es ein $N_0 \in \mathbb{N}$ gibt, so dass*
$$\left|\frac{a_{k+1}}{a_k}\right| \geq 1$$
für alle $k > N_0$ ist, so divergiert die Reihe $\sum a_k$.

Beweis. 1. Sei $q = \lim_{k\to\infty} \sup \left|\frac{a_{k+1}}{a_k}\right| < r < 1$ mit einem geeignetem $r \in \mathbb{R}$, dann existiert ein $N_0 \in \mathbb{N}$ mit $\left|\frac{a_{k+1}}{a_k}\right| < r$, d.h. $|a_{k+1}| \leq r|a_k|$, für alle $k > N_0$. Somit ist $|a_k| \leq r^k |a_{N_0}|$ für alle $k > M_0$ und $\sum_{k=N_0}^{\infty} r^k$ eine konvergente Majorante.

2. Falls $\left|\frac{a_{k+1}}{a_k}\right| \geq 1$ für alle $k \geq N_0$, so folgt $|a_{k+1}| \geq |a_k| \geq |a_{N_0}| \geq 1$ und $\lim_{k\to\infty} a_k \neq 0$, falls dieser Grenwert überhaupt exisitert. Somit ist in diesem Fall die Reihe divergent.

□

Das Quotienten– und das Wurzelkriterium sind äußerst hilfreiche Instrumente zur Prüfung der Konvergenz einer Reihe. Das Quotientenkriterium läßt sich zumeist leichter anwenden. Demgegenüber besitzt das Wurzelkriterium einen größeren Wirkungsbereich.

4.6 Alternierende Reihen

Definition 4.6.1. *Eine Reihe* $\sum_{k=1}^{\infty} c_k$ *mit* $c_k \in \mathbb{R}$ *nennen wir* **alternierend**, *falls jeweils die Vorzeichen aufeinander folgender* c_k *verschieden sind, d.h.,* $c_k \cdot c_{k+1} < 0$ *für alle* $k \in \mathbb{N}$.

Für alternierende Reihen ist der Nachweis von Konvergenz sehr einfach mittels des *Leibniz–Kriteriums* durchzuführen. Um dieses allerdings beweisen zu können, benötigen wir zuerst zwei Lemmeta.

Lemma 4.6.2. *Seien* $a_k, b_k \in \mathbb{R}$ *für alle* $k \in \mathbb{N}_0$. *Setze*

$$A_n := \sum_{k=0}^{n} a_k$$

für $n \geq 0$ *und* $A_{-1} := 0$. *Dann gilt für* $M \geq m \in \mathbb{N}_0$

$$\sum_{n=m}^{M} a_n b_n = A_M b_M - A_{m-1} b_m + \sum_{n=m}^{M-1} A_n (b_n - b_{n+1}).$$

Beweis. Es ist

$$\begin{aligned}
\sum_{n=m}^{M} a_n b_n &= \sum_{n=m}^{M} (A_n - A_{n-1}) b_n \\
&= \sum_{n=m}^{M} A_n b_n - \sum_{n=m-1}^{M-1} A_n b_{n+1} \\
&= A_M b_M - A_{m-1} b_m + \sum_{n=m}^{M-1} A_n (b_n - b_{n+1}).
\end{aligned}$$

□

Lemma 4.6.3 (Dirichlet–Kriterium). *Seien* $a_k, b_k \in \mathbb{R}$ *für alle* $k \in \mathbb{N}_0$ *und es gelte:*

1. *Es gibt ein* $R > 0$ *mit* $A_n := |\sum_{k=0}^{n} a_k| < R$ *für alle* $n \in \mathbb{N}_0$.
2. *Die Folge* $(b_k)_{k \in \mathbb{N}}$ *ist monoton fallend mit* $\lim_{k \to \infty} b_k = 0$.

Dann konvergiert die Reihe $\sum a_k b_k$.

Beweis. Wegen $b_k \to 0$ für $k \to \infty$, existiert zu beliebigem $\varepsilon > 0$ ein $N_0 \in \mathbb{N}$, so dass

$b_k \leq b_{N_0} < \frac{\varepsilon}{2R}$ für alle $k > N_0$. Seien $q \geq p > N_0$, so gilt mit Hilfe des obigen Lemmas

$$\begin{aligned}\left|\sum_{n=p}^{q} a_n b_n\right| &\leq |A_q b_q - A_{p-1} b_p| + \left|\sum_{n=p}^{q-1} A_n (b_n - b_{n+1})\right| \\ &\leq R\left(b_p + b_q + \sum_{n=p}^{q-1}(b_n - b_{n+1})\right) \\ &= R(b_p + b_q + b_p - b_q) \\ &< 2R\frac{\varepsilon}{pR} < \varepsilon.\end{aligned}$$

Hieraus folgt die Konvergenz nach dem Cauchy-Kriterium (Satz 4.4.2). □

Satz 4.6.4 (Leibniz–Kriterium). *Es sei $\sum_{k=1}^{\infty} c_k$ mit $c_k \in \mathbb{R}$ eine alternierende Reihe. Es gelte:*

1. *Die Folge $(|c_k|)_{k\in\mathbb{N}}$ ist monoton fallend.*
2. $\lim_{k\to\infty} c_k = 0$.

Dann konvergiert die Reihe $\sum c_k$.

Beweis. Es gelte ohne Beschränkung der Allgemeinheit $c_1 \geq 0$. Dann ist

$$\sum_{k=1}^{\infty} c_k = \sum_{k=1}^{\infty} (-1)^{k+1} |c_k|.$$

Setzen wir $a_k = (-1)^{k+1}$ und $b_k = |c_k|$ in Lemma 4.6.3, so folgt damit die Behauptung. □

Beispiel 4.6.5. *Die Reihe $\sum\limits_{k=1}^{\infty} \frac{1}{k}(-1)^k$ ist konvergent, obwohl $\sum\limits_{k=1}^{\infty} \frac{1}{k}$ divergiert.*

Definition 4.6.6. *Eine Reihe $\sum a_k$ heißt* **absolut konvergent**, *falls die Reihe $\sum |a_k|$ (also die Reihe der Absolutbeträge) konvergiert.*

Proposition 4.6.7. *Ist die Reihe $\sum a_k$ absolut konvergent, so ist sie auch konvergent.*

Der Beweis dieser Proposition ist trivial. Aber man beachte, dass die Umkehrung im allgemeinen nicht gilt, siehe Beispiel 4.6.5.

4.7 Summen und Produkte von Reihen

Satz 4.7.1. *Konvergieren die beiden Reihen $A = \sum_{k=0}^{\infty} a_k$ und $A = \sum_{k=0}^{\infty} b_k$, so konvergiert auch deren Summe $A + B = \sum_{k=0}^{\infty}(a_k + b_k)$.*

Beweis. Der Beweis ist eine einfache Anwendung der Limesregeln für Folgen (Satz 4.2.4). □

Definition 4.7.2. *Zu den beiden Reihen* $\sum_{k=0}^{\infty} a_k$ *und* $\sum_{k=0}^{\infty} b_k$ *definieren wir das* **Cauchy–Produkt** $\sum_{n=0}^{\infty} c_n$ *durch*

$$c_n = \sum_{k=0}^{n} a_k b_{n-k}.$$

Satz 4.7.3. *Konvergieren die beiden Reihen* $A = \sum_{k=0}^{\infty} a_k$ *und* $B = \sum_{k=0}^{\infty} b_k$ *beide* **absolut**, *dann konvergiert auch das Cauchy–Produkt*

$$\sum_{n=0}^{\infty} c_n = A \cdot B.$$

Beweis. Wir setzen $A_n = \sum_{k=0}^{n} a_k$, $B_n = \sum_{k=0}^{n} b_k$ und $C_n = \sum\limits_{k=0}^{n} c_k$ und $\beta_n = B_n - B$. Dann gilt

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

Ferner setzen wir $\gamma_n = a_0\beta_n + a_1\beta_{n-1} + \cdots + a_n\beta_0$ und $\alpha = \sum_{k=0}^{\infty} |a_k|$. Aufgrund der Konvergenz der Reihe $\sum_{k=0}^{\infty} b_k$ gilt $\lim_{n\to\infty} \beta_n = 0$. Sei nun $\varepsilon > 0$ beliebig aber fest. Dann können wir ein $N_0 \in \mathbb{N}$ so wählen, dass $|\beta_n| < \varepsilon$ für alle $n > N_0$ gilt. Hieraus folgt

$$\begin{aligned} |\gamma_n| &\leq |\beta_0 a_n + \cdots + \beta_{N_0} a_{n-N_0}| + |\beta_{N+1} a_{n-N-1} + \cdots + \beta_n a_0| \\ &\leq |\beta_0 a_n + \cdots + \beta_n a_{n-N}| + \varepsilon\alpha. \end{aligned}$$

Für festes N_0 gilt nun aber $\limsup_{n\to\infty} |\gamma|_n \leq \varepsilon\alpha$, da $a_k \to 0$ für $k \to \infty$. Dies impliziert $\lim_{n\to\infty} \gamma_n = 0$ und damit

$$C = \lim_{n\to\infty} C_n = \lim_{n\to\infty} A_n B + \lim_{n\to\infty} \gamma_n = AB.$$

□

Die Voraussetzung der absoluten Konvergenz ist wichtig, wie folgendes Beispiel zeigt.

Beispiel 4.7.4. *Sei* $a_k = b_k = \frac{(-1)^k}{\sqrt{k+1}}$. *Die beiden Reihen* $\sum a_k$ *und* $\sum b_k$ *konvergieren nach dem Leibniz–Kriterium, aber sie sind nicht absolut konvergent. Nun gilt für*

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

und

$$|c_n| \geq \sum_{k=0}^{n} \frac{1}{n+1} = \frac{n+1}{n+1} = 1.$$

Das heißt $\sum c_k$ *ist divergent.*

4.8 Umordnen von Reihen

Definition 4.8.1. *Sei $s : \mathbb{N} \to \mathbb{N}$ eine bijektive Abbildung. Dann heißt $(a_{s(n)})_{n\in\mathbb{N}}$ eine* **Umordnung der Folge** $(a_n)_{n\in\mathbb{N}}$ *und $\sum_{n=0}^{\infty} a_{s(n)}$ eine* **Umordnung der Reihe** $\sum a_n$.

Satz 4.8.2 (Umordnungssatz). *Jede Umordnung $\sum_{n=1}^{\infty} a_{s(n)}$ einer absoluten konvergenten Reihe $\sum_{n=1}^{\infty} a_n$ konvergiert ebenfalls absolut und es gilt $\sum_{n=1}^{\infty} a_n = \sum_{n=1}^{\infty} a_{s(n)}$.*

Beweis. Wir bezeichnen die Teilsummen von $\sum a_n$ und $\sum a_{s(n)}$ mit $s_m = \sum_{n=1}^{m} a_n$ und $t_m = \sum_{n=1}^{m} a_{s(n)}$. Zu vorgegebenen $\varepsilon > 0$ gibt es ein $N_0 \in \mathbb{N}$ derart, dass $|a_{N_0+1}| + |a_{N_0+2}| + \cdots + |a_{N_0+m}| < \varepsilon$ ist für beliebiges $m \in \mathbb{N}$. Demnach ist $\sum_{n=N_0+1}^{\infty} |a_n| < \varepsilon$. Wir wählen nun $M > N_0$ so groß, dass unter den Zahlen $s(1), s(2), \ldots, s(M)$ alle Zahlen $1, 2, \ldots, N_0$ vorkommen.

Für alle $n > M$ heben sich dann in der Differenz $s_n - t_n$ alle Glieder a_k mit einem Index $k \leq M$ auf, da diese sowohl in s_n als auch in t_n erscheinen, d.h. es gilt also

$$|s_n - t_n| < \varepsilon.$$

Damit ist gezeigt dass $(s_n - t_n)_{n\in\mathbb{N}}$ eine Nullfolge ist mit $\lim_{n\to\infty} s_n = \lim_{n\to\infty} t_n$ bzw. $\sum_{n=1}^{\infty} a_n = \sum_{n=1}^{\infty} a_{s(n)}$. Diese Schlussweise läßt sich auch auf die Reihe $\sum_{n=1}^{\infty} |a_n|$ anwenden, womit der Satz bewiesen ist. □

Problematisch wird die Situation allerdings, falls $\sum a_n$ konvergiert, aber nicht absolut konvergiert. Beispielsweise ist nach dem Leibniz–Kriterium die Reihe $\sum_{n=1}^{\infty} (-1)^n \frac{1}{n}$ konvergent, aber $\sum_{l=1}^{\infty} \frac{1}{2l} - \sum_{l=0}^{\infty} \frac{1}{2l+1}$ divergent. In $\mathbb{R}$ gilt sogar folgender Sachverhalt.

Satz 4.8.3 (Riemann). *Sei $\sum_{n=1}^{\infty} a_n$ konvergent, aber nicht absolut konvergent. Zu beliebigen $-\infty \leq \alpha \leq \beta \leq +\infty$ existiert eine Umordnung $s(n)$ derart, dass*

$$\liminf_{N\to\infty} \sum_{n=1}^{N} a_{s(n)} = \alpha$$

und

$$\limsup_{N\to\infty} \sum_{n=1}^{N} a_{s(n)} = \beta.$$

Beweis. Wir setzen

$$p_n = \frac{|a_n| + a_n}{2}, \quad q_n = \frac{|a_n| - a_n}{2}$$

für alle $n \in \mathbb{N}$. Dann ist $p_n - q_n = a_n$ und $p_n + q_n = |a_n|$, sowie $p_n, q_n \geq 0$.

Da $\sum p_n + q_n = \sum |a_n|$ divergiert, muss mindestens eine der beiden Reihen $\sum p_n$ bzw. $\sum q_n$ divergieren. Angenommen $\sum q_n$ wäre konvergent, so müsste auch $\sum p_n = \sum a_n + \sum q_n$ konvergieren. Hieraus folgt, dass beide Reihen $\sum p_n$ und $\sum q_n$ divergent sind.

Seien P_i, $i \in \mathbb{N}$, nun die nichtnegativen Glieder von $(a_k)_{k\in\mathbb{N}}$ und Q_i, $i \in \mathbb{N}$, die Absolutbeträge der negativen Glieder. Die Reihen $\sum p_n$ und $\sum q_n$ unterscheiden sich von $\sum P_n$ und $\sum Q_n$ nur durch Nullglieder und sind daher ebenfalls beide divergent.

Wir wählen konvergente reelle Zahlenfolgen $(\alpha_n)_{n\in\mathbb{N}}$ und $(\beta_n)_{n\in\mathbb{N}}$ mit $\lim_{n\to\infty} \alpha_n = \alpha$ und $\lim_{n\to\infty} \beta_n = \beta$. Seien m_1 und k_1 die kleinsten ganzen Zahlen derart, dass

$$B_1 := \sum_{i=1}^{m_1} P_i > \beta_1$$

und

$$A_1 := \sum_{i=1}^{m_1} P_i - \sum_{i=1}^{k_1} Q_i < \alpha_1.$$

Nun definieren wir m_2 und k_2 also die kleinsten ganzen Zahlen mit

$$B_2 := \sum_{i=1}^{m_1} P_i - \sum_{i=1}^{k_1} Q_i + \sum_{i=m_1+1}^{m_2} P_i > \beta_2$$

und

$$A_2 := \sum_{i=1}^{m_1} P_i - \sum_{i=1}^{k_1} Q_i + \sum_{i=m_1+1}^{m_2} P_i - \sum_{i=k_1}^{k_2} Q_i < \alpha_2.$$

Aufgrund der Divergenz der Reihen $\sum P_i$ und $\sum Q_i$ läßt sich diese Konstruktion beliebig fortsetzen, und es gilt die Abschätzung

$$|B_n - \beta_n| \leq P_{m_n}$$

und

$$|A_n - \alpha_n| \leq Q_{k_n}$$

für alle $n \in \mathbb{N}$. Wegen der Konvergenz von $\sum a_n$ folgt $P_n \to 0$ und $Q_n \to 0$ für $n \to \infty$. Daher konvergieren die Partialsummen B_n bzw. A_n nach β bzw. α, also $\sum_{k=1}^{\infty} B_n = \beta$ und $\sum_{k=1}^{\infty} A_n = \alpha$. Wie man leicht sieht handelt es sich bei den beiden konstruierten Reihen um Umordnungen der Reihe $\sum a_n$. □

4.9 Potenzreihen

Definition 4.9.1. *Sei $(a_k)_{k\in\mathbb{N}}$ eine komplexe Zahlenfolge. Dann nennen wir die Reihe*

$$\sum_{k=0}^{\infty} a_k z^k, \quad z \in \mathbb{C}$$

eine **Potenzreihe** *und* a_k *die* **Koeffizienten** *der Potenzreihe. Die Zahl*

$$R := \left(\limsup_{k\to\infty} \sqrt[k]{|a_k|} \right)^{-1}$$

heißt der **Konvergenzradius** *der Potenzreihe.*

Satz 4.9.2. *Sei* $\sum_{k=0}^{\infty} a_k z^k$ *eine Potenzreihe mit Konvergenzradius* R*. Dann konvergiert diese Reihe für alle* $|z| < R$ *absolut und divergiert für* $|z| > R$*.*

Beweis. Wir setzen $c_k = a_k z^k$ und wenden das Wurzelkriterium an

$$\limsup_{k\to\infty} \sqrt[k]{|c_k|} = \limsup_{k\to\infty} \sqrt[k]{|a_k|} \cdot |z| = \frac{|z|}{R}.$$

□

Beispiel 4.9.3.

1. *Die Reihe* $\sum_{k=1}^{\infty} k^k z^k$ *hat den Konvergenzradius* $R = 0$*.*

2. *Die Reihen*
$$e^z = \sum_{k=0}^{\infty} \frac{z^k}{k!}$$
und
$$\sin z = \sum_{k=0}^{\infty} (-1)^k \frac{z^{2k+1}}{(2k+1)!}, \quad \cos z = \sum_{k=0}^{\infty} (-1)^k \frac{z^{2k}}{(2k)!},$$
haben den Konvergenzradius $R = \infty$*.*

3. *Die Reihe* $\sum_{k=0}^{\infty} z^k = \frac{1}{1-z}$ *hat den Konvergenzradius* $R = 1$ *(geometrischer Reihe).*

4. *Die Reihen* $\sum_{k=1}^{\infty} \frac{z^k}{k}$ *und* $\sum_{k=1}^{\infty} \frac{z^k}{k^2}$ *haben beide den Konvergenzradius* $R = 1$*. Allerdings divergiert die erste Reihe für* $|z| = 1$ *während die zweite Reihe für alle* $|z| \leq 1$ *konvergiert.*

Bemerkung 4.9.4. *Für den Fall* $|z| = R$ *ist im allgemeinen keine Aussage möglich.*

Kapitel 5

Stetige Funktionen

5.1 Funktionsgrenzwerte

Definition 5.1.1. *Seien (X, d_X) und (Y, d_Y) metrische Räume. Desweiteren seien E eine Teilmenge von X, $f : E \to Y$ eine Funktion und p ein Häufungspunkt der Menge E. Wir sagen f hat den* **Funktionsgrenzwert** *y an der Stelle p, kurz $f(x) \to y$ für $x \to p$ bzw. $\lim_{x\to p} f(x) = y$, falls ein $y \in Y$ existiert, so dass es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt derart, dass*

$$d_Y\big(f(x), y\big) < \varepsilon$$

für alle $x \in \{x \in E : 0 < d_X(x, p) < \delta\} = \big(U_\delta(p) \setminus \{p\}\big) \cap E$.

Satz 5.1.2. *Es gilt $f(x) \to y$ für $x \to p$ genau dann, wenn für jede Folge $(x_k)_{k\in\mathbb{N}}$ aus $E \setminus \{p\}$ mit $\lim_{k\to\infty} x_k = p$ gilt $\lim_{k\to\infty} f(x_k) = y$.*

Beweis. "$\Rightarrow$": Falls $\lim_{x\to p} f(x) = y$, so folgt nach Definition für beliebiges $\varepsilon > 0$, dass ein $\delta > 0$ existiert mit $d_Y\big(f(x), y\big) < \varepsilon$ für alle $x \in \big(U_\delta(p) \setminus \{p\}\big) \cap E$. Falls die Folge $(x_k)_{k\in\mathbb{N}}$ für $k \to \infty$ gegen p konvergiert, gibt es ein $N_\delta \in \mathbb{N}$ mit $0 < d_X(x_k, p) < \delta$ für alle $k > N_\delta$. Die bedeutet $x_k \in \big(U_\delta(p) \setminus \{p\}\big) \cap E$ und folglich gilt $d_Y\big(f(x_k), y\big) < \varepsilon$ für alle $k > N_\delta$.

"$\Leftarrow$": Angenommen, es gilt $f(x) \not\to y$ für $x \to p$. Dann existiert ein $\varepsilon > 0$, so dass für jedes $\delta > 0$ ein $x \in E \setminus \{p\}$ existiert mit $d_Y\big(f(x), y\big) \geq \varepsilon$, aber $0 < d_X(x, p) < \delta$. Wir setzen $\delta = \delta_n = \frac{1}{n}$ und nehmen einen zugehörigen Punkt $x_n = x$. Dann gilt $x_n \to x$ für $n \to \infty$, aber $d_Y\big(f(x_n), y\big) \geq \varepsilon > 0$ für alle $n \in \mathbb{N}$. Dies steht im Widerspruch zur Voraussetzung $f(x_n) \to y$. $\square$

Korollar 5.1.3. *Hat eine Funktion f an der Stelle p einen Funktionsgrenzwert, so ist dieser eindeutig.*

Die Limesregeln (Satz 4.2.4) übertragen sich analog auf die Funktionsgrenzwerte.

Satz 5.1.4. *Sei $\mathbb{R}^m$ und $\mathbb{R}^n$, $m, n \in \mathbb{N}$, versehen mit der Euklidischen Norm. Desweiteren seien $E \subseteq \mathbb{R}^m$ und $\mathbf{p}$ ein Häufungspunkt der Menge E. Dann gilt für die Funktionen $f : E \to \mathbb{R}^n$, $g : E \to \mathbb{R}^n$ und beliebigen $\mathbf{a} \in \mathbb{R}^n$ und $\alpha \in \mathbb{R}$, dass*

$$\begin{aligned}
\lim_{\mathbf{x}\to\mathbf{p}} \big(f(\mathbf{x}) + g(\mathbf{x})\big) &= \lim_{\mathbf{x}\to\mathbf{p}} f(\mathbf{x}) + \lim_{\mathbf{x}\to\mathbf{p}} g(\mathbf{x}),\\
\lim_{\mathbf{x}\to\mathbf{p}} \big(f(\mathbf{x}) + \mathbf{a}\big) &= \lim_{\mathbf{x}\to\mathbf{p}} f(\mathbf{x}) + \mathbf{a},\\
\lim_{\mathbf{x}\to\mathbf{p}} \big(\alpha f(\mathbf{x})\big) &= \alpha \lim_{\mathbf{x}\to\mathbf{p}} f(\mathbf{x}),\\
\lim_{\mathbf{x}\to\mathbf{p}} \langle f(\mathbf{x}), g(\mathbf{x})\rangle &= \langle \lim_{\mathbf{x}\to\mathbf{p}} f(\mathbf{x}), \lim_{\mathbf{x}\to\mathbf{p}} g(\mathbf{x})\rangle.
\end{aligned}$$

Beweis. Der Beweis bleibt dem Leser als Übung überlassen. □

5.2 Stetigkeit von Funktionen

Definition 5.2.1. *Sei $E \subseteq X$ und $f : E \to Y$ eine Funktion sowie $x_0 \in E$. Die Funktion f heißt* **stetig an der Stelle** *$x_0 \in E$, falls es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt derart, dass $d_y\big(f(x), f(x_0)\big) < \varepsilon$ für alle $d_x(x, x_0) < \delta$ gilt. Die Funktion f heißt* **stetig in** *E, falls f in jedem Punkt $x \in E$ stetig ist.*

Bemerkung 5.2.2. *In einem isolierten Punkt $x_0 \in E$ ist jede Funktion stetig, wie man leicht sieht.*

Satz 5.2.3. *Unter der Voraussetzung des vorigen Satzes gilt in einem Häufungspunkt x_0 von E, dass f dort genau dann stetig ist, falls*

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

ist.

Beweis. Vergleiche beide Definitionen! □

Satz 5.2.4. *Es seien (X, d_X), (Y, d_Y) und (Z, d_Z) metrische Räume. Für $E \subseteq X$ seien durch $f : E \to f(E) \subseteq Y$ und $g : f(E) \to Z$ Abbildungen gegeben und $h = g \circ f : E \to Z$ bezeichne ihre Komposition. Falls f in $x_0 \in E$ und g in $f(x_0)$ beide stetig sind, so ist h an der Stelle x_0 ebenfalls stetig.*

Beweis. Da g an der Stelle $f(x_0) \in X$ stetig ist, gibt es zu jedem $\varepsilon > 0$ ein $\gamma > 0$ mit $d_Z\big(g(y), g(f(x_0))\big) < \varepsilon$ für alle $y \in Y$ mit $d_Y\big(y, f(x_0)\big) < \gamma$. Aufgrund der Stetigkeit von f in $x_0 \in E$ existiert zu diesem $\gamma > 0$ ein $\delta > 0$ mit $d_Y\big(f(x), f(x_0)\big) < \gamma$ für alle $x \in E$ mit $d_X(x, x_0) < \delta$. Dies bedeutet, es gilt

$$d_Z\big(g(f(x)), g(f(x_0))\big) = d_Z\big(h(x), h(x_0)\big) < \varepsilon$$

für alle $x \in E$ mit $d_X(x, x_0) < \delta$. □

Satz 5.2.5. *Die Abbildung $f : X \to Y$ ist genau stetig auf X, falls für jede offene Menge $W \subseteq Y$ das Urbild $f^{-1}(W)$ eine offene Teilmenge in X ist.*

Beweis. "$\Rightarrow$": Sei f stetig auf X und $W \subseteq Y$ eine offene Teilmenge von Y. Wir zeigen, dass jeder Punkt $p \in f^{-1}(W)$ ein innerer Punkt der Menge $f^{-1}(W)$ ist. Sei $p \in X$ mit $f(p) \in W$. Dann ist $f(p)$ ein innerer Punkt in W, d.h. es existiert $\varepsilon > 0$ mit $U_\varepsilon\big(f(p)\big) \subseteq W$. Aufgrund der Stetigkeit existiert ein $\delta > 0$ mit $f(x) \in U_\varepsilon\big(f(p)\big)$ für alle x mit $d_X(x,p) < \delta$. Damit gilt für alle $x \in U_\delta(p)$ dass $f(x) \in U_\varepsilon\big(f(p)\big) \subseteq f(W)$, d.h. $x \in f^{-1}(W)$ bzw. $U_\delta(p) \subseteq f^{-1}(W)$. Da $U_\delta(p)$ eine offene Umgebung von p definiert, ist dieser Teil des Satzes bewiesen.

"$\Leftarrow$": Sei $f^{-1}(W)$ eine offene Menge in X für jede offene Menge $W \subseteq Y$. Zu $x_0 \in X$ und beliebigem $\varepsilon > 0$ sei $W := \big\{y \in Y : d_Y\big(y, f(x_0)\big) < \varepsilon\big\}$. Da W eine offene Teilmenge in Y, ist $U = f^{-1}(W) \subseteq X$ ebenfalls offen und $x_0 \in W$ ist ein innerer Punkt von U. Das heißt, es existiert eine offene Umgebung

$$U_\delta(x_0) = \{x \in X : d_X(x, x_0) < \delta\} \subseteq W.$$

Somit gilt für alle $x \in U_\delta(x_0)$, das heißt $d_X(x, x_0) < \delta$, dass $d_X\big(f(x), f(x_0)\big) < \varepsilon$ ist. □

Satz 5.2.6. *Seien f und $g : X \to \mathbb{C}$ (bzw. $\mathbb{R}$) stetige Funktionen, dann sind ebenfalls $f + g$ und $f \cdot g$ stetig. Falls $g(x) \neq 0$ für alle $x \in X$, dann ist sogar f/g eine stetig Funktion.*

Beweis. Der Beweis ist eine einfache Anwendung der Limesregeln für Folgen. □

Satz 5.2.7. *Seien $f_1, f_2, \ldots, f_n : X \to \mathbb{R}$ stetige Funktionen, dann ist $\mathbf{f} = (f_1, f_2, \ldots, f_n) : X \to \mathbb{R}^n$ stetig. Sei $\mathbf{g} : X \to \mathbb{R}^n$ ebenfalls stetig, so gilt dies auch für die Funktionen $\mathbf{f} + \mathbf{g} : X \to \mathbb{R}^n$ und $\langle \mathbf{f}, \mathbf{g} \rangle : X \to \mathbb{R}$.*

Beweis. Der erste Teil des Beweises folgt aus der Ungleichung

$$|f_i(x) - f_i(x_0)| \leq \|\mathbf{f}(x) - \mathbf{f}(x_0)\|.$$

Der Rest verbleibt dem Leser als Übung. □

Beispiel 5.2.8.

1. *Sei $f : \mathbb{R} \to \mathbb{R}$ ein Polynom, so ist f stetig.*
2. *Sei $f : \mathbb{R} \to \mathbb{R}$ gegeben durch*

$$f(x) = \operatorname{sgn}(x) = \begin{cases} +1, & x > 0, \\ 0, & x = 0, \\ -1, & x < 0, \end{cases}$$

so ist f in $x = 0$ nicht stetig, aber in $\mathbb{R} \setminus \{0\}$ stetig.

3. *Sei*
$$f(x) = \begin{cases} \frac{1}{x}, & x \neq 0, \\ 0, & x = 0, \end{cases},$$
so ist f in $x = 0$ nicht stetig, aber in $\mathbb{R} \setminus \{0\}$.

4. *Die Funktion*
$$f(x) = \begin{cases} \sin \frac{1}{x}, & x \neq 0, \\ 0, & x = 0, \end{cases},$$
ist in $x = 0$ nicht stetig, aber in $\mathbb{R} \setminus \{0\}$.

5. *Die Funktion*
$$f(x) = \begin{cases} x \sin \frac{1}{x}, & x \neq 0, \\ 0, & x = 0, \end{cases},$$
ist in $\mathbb{R}$ stetig.

5.3 Einseitiger Grenzwert und Stetigkeit

Sei $E \subseteq \mathbb{R}$ und $f : E \to Y$ eine Funktion. Häufig treten Grenzwerte auf, bei denen x von rechts oder von links gegen x_0 strebt, bei denen also E ein Intervall $(x_0, x_0 + a)$ bzw. $(x_0, x_0 - a)$ mit $a > 0$ ist. Man spricht dann von einem *rechtsseitigen* bzw. *linksseitigen Grenzwert* und verwendet zu seiner Kennzeichnung das Symbol $x \to x_0+$ bzw. $x \to x_0-$ (gelegentlich auch $x \to x_0 + 0$ bzw. $x \to x_0 - 0$). Es existiert also zum Beispiel $y = \lim_{x \to x_0+} f(x)$, wenn es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt, so dass für alle $x_0 < x < x_0 + \delta$ stets die Ungleichung $|f(x) - y| < \varepsilon$ besteht. Diese Begriffe werden auch dann benutzt, wenn f in einer vollen Umgebung von x_0 definiert ist, aber der Grenzwert $\lim_{x \to x_0} f(x)$ nicht existiert. Man überzeugt sich leicht, dass genau dann $\lim_{x \to x_0} f(x) = y$ existiert, wenn $\lim_{x \to x_0-} f(x) = \lim_{x \to x_0+} f(x) = y$ ist.

Analog ist die *einseitige Stetigkeit* definiert. Eine Funktion ist *linksseitig stetig*, wenn
$$\lim_{x \to x_0-} f(x) = f(x_0),$$
bzw. *rechtsseitig stetig*, wenn
$$\lim_{x \to x_0+} f(x) = f(x_0).$$
Wenn f in einer vollen Umgebung von x_0 definiert ist, so besteht ein ähnlicher Zusammenhang wie beim Grenzwert, kurz "rechtsseitig stetig und linksseitig stetig $\Leftrightarrow$ stetig".

Beispiel 5.3.1. *Für die Signumfunktion aus Beispiel 5.2.8 gilt*
$$\lim_{x \to 0-} \operatorname{sgn}(x) = -1, \qquad \lim_{x \to 0+} \operatorname{sgn}(x) = +1.$$
Insbesondere ist sie rechtseitig stetig im Nullpunkt (aber nicht linkseitig stetig, denn dies würde ja Stetigkeit bedeuten).

5.4 Zusammenhängende Mengen und Stetigkeit

Definition 5.4.1. *Sei (X, d) ein metrischer Raum und $A, B \subseteq X$. Die beiden Mengen A und B heißen* **getrennt***, falls*

$$\overline{A} \cap B = \emptyset \quad \text{und} \quad A \cap \overline{B} = \emptyset.$$

Insbesondere ist somit die leere Menge von jeder anderen Menge getrennt.

Eine Menge $M \subseteq X$ heißt **zusammenhängend***, wenn sie nicht als Vereinigung zweier nichtleerer getrennter Mengen geschrieben werden kann.*

Beispiel 5.4.2.

1. *Die beiden Mengen $A = (0, 1)$ und $B = (1, 2)$ sind getrennt.*
2. *Die beiden Mengen $A = [0, 1]$ und $B = (1, 2)$ sind zwar disjunkt, aber nicht getrennt.*

Satz 5.4.3. *Eine Teilmenge $M \subseteq \mathbb{R}$ ist genau dann zusammenhängend, falls für alle $x, y \in M$ mit $x < y$ jedes $z \in (x, y)$ auch $z \in M$ ist.*

Beweis. "$\Rightarrow$": Seien $x, y \in M$ mit $x < y$ und $z \in (x, y)$ mit $z \notin M$. Wir setzen

$$A := (-\infty, z) \cap M, \quad B := (z, \infty) \cap M.$$

Dann gilt $M = A \cup B$ und $\overline{A} \cap B = \emptyset$ sowie $A \cap \overline{B} = \emptyset$, das heißt M ist nicht zusammenhängend. Dies steht aber im Widerspruch zur Voraussetzung.

"$\Leftarrow$": Sei M nicht zusammenhängend. Dann existieren zwei getrennte Mengen A und B mit $M = A \cup B$. Sei $x \in A$ und $y \in B$. Ohne Beschränkung der Allgemeinheit gelte $x < y$. Wir setzen

$$z := \sup(A \cap (x, y)).$$

Dann ist $z \in \overline{A}$, aber wegen der Getrenntheit $z \notin B$. Falls $z \notin A$, so gilt $x < z < y$ und $z \notin M$. Falls $z \in A$, so existiert wegen $z \notin \overline{B}$ ein weiteres $z' \notin B$ mit $z < z' < y$. Das heißt $z' \notin A$und somit $z' \notin M$. □

Satz 5.4.4. *Sei $f : X \to Y$ eine auf X stetige Funktion. Zu jeder zusammenhängenden Menge $M \subseteq X$ ist auch $f(M) \subseteq Y$ eine zusammenhängende Teilmenge von Y.*

Beweis. Angenommen, $f(M)$ sei nicht zusammenhängend, d.h. $f(M) = A \cup B$ mit geeigneten, getrennten Mengen $A, B \subseteq Y$. Wir definieren

$$G := f^{-1}(A) \cap M \neq \emptyset \quad \text{und} \quad H := f^{-1}(B) \cap M \neq \emptyset.$$

Dann gilt $M = G \cup H$, da $M \subseteq f^{-1}(A) \cup f^{-1}(B)$. Aufgrund der Stetigkeit von f ist $f^{-1}(\overline{A})$ in X abgeschlossen und es gilt $\overline{G} \subseteq f^{-1}(\overline{A})$, somit $f(\overline{G}) \subseteq \overline{A}$. Da $f(H) = B$, folgt aus $\overline{A} \cap B = \emptyset$ auch $\overline{G} \cap H = \emptyset$. Analog zeigt man $G \cap \overline{H} = \emptyset$, d.h. die Menge M ist nicht zusammenhängend. □

Beispiel 5.4.5. *Sei $f : \mathbb{R} \to \mathbb{R}$ definiert durch*

$$f(x) := \begin{cases} \sin \frac{1}{x}, & x \neq 0, \\ 0, & x = 0. \end{cases}$$

Dann ist f in $x_0 = 0$ nicht stetig. Dennoch wird jede zusammenhängende Menge M wieder auf eine zusammenhängende Menge abgebildet.

Satz 5.4.6 (Zwischenwertsatz). *Sei $f : [a, b] \to \mathbb{R}$ eine stetige Funktion mit $f(a) < f(b)$. Dann existiert zu jedem $c \in \mathbb{R}$ mit $f(a) < c < f(b)$ ein $x \in (a, b)$ mit $f(x) = c$.*

Beweis. Da das Intervall $[a, b]$ nach Satz 5.4.3 zusammenhängend ist, ist auch $f([a, b])$ eine zusammenhängende Menge, d.h. $f([a, b]) = [f(a), f(b)]$. □

Definition 5.4.7. *Eine Funktion $f : [a, b] \to \mathbb{R}$ heißt* **monoton wachsend** *bzw.* **monoton fallend** *falls $f(x) \leq f(y)$ bzw. $f(x) \geq f(y)$. Wir sprechen von* **strenger Monotonie** *falls die strenge Ungleichung gilt.*

Korollar 5.4.8. *Sei $f : [a, b] \to \mathbb{R}$ streng monoton wachsend und stetig, dann ist*

$$f : [a, b] \to [f(a), f(b)]$$

bijektiv.

5.5 Stetigkeit und Kompaktheit

Satz 5.5.1. *Seien (X, d_X) und (Y, d_Y) metrische Räume und die Menge $K \subseteq X$ kompakt. Ist $f : K \to Y$ eine stetige Abbildung, dann ist $f(K) \subseteq Y$ kompakt.*

Beweis. Wir betrachten eine beliebige offene Überdeckung $\{V_\alpha\}_{\alpha \in I}$ von $f(K)$. Aufgrund von Satz 5.2.5 sind die Mengen $f^{-1}(V_\alpha)$ für alle $\alpha \in I$ offen in X. Wegen $f(K) \subseteq \bigcup_{\alpha \in I} V_\alpha$ gilt $K \subseteq \bigcup_{\alpha \in I} f^{-1}(V_\alpha)$, d.h. $\{f^{-1}(V_\alpha) : \alpha \in I\}$ bildet eine offene Überdeckung von X. Infolge der Kompaktheit von K existiert eine endliche Teilüberdeckung $X \subseteq \bigcup_{i=1}^{N} f^{-1}(V_{\alpha_i})$. Damit gilt

$$f(K) \subseteq \bigcup_{i=1}^{N} f\big(f^{-1}(V_{\alpha_i})\big) = \bigcup_{i=1}^{N} V_{\alpha_i}.$$

Die bedeutet, dass $\{V_{\alpha_i}, i = 1, \ldots, N\}$ eine endliche Teilüberdeckung von $f(K)$ bildet. □

Definition 5.5.2. *Seien (X, d_X) und (Y, d_Y) metrische Räume und $E \subseteq X$. Die Abbildung $f : E \to Y$ heißt* **beschränkt**, *falls $R > 0$ und $y_0 \in Y$ existieren mit $d_Y\big(f(x), y_0\big) < R$ für alle $x \in E$.*

Satz 5.5.3. *Seien (X, d_X) und (Y, d_Y) metrische Räume und die Menge $K \subseteq X$ kompakt. Sei ferner $f : K \to Y$ eine stetige Funktion, dann ist $f(K) \subseteq Y$ eine abgeschlossene und beschränkte Menge, d.h. f ist eine beschränkte Funktion.*

Beweis. Wegen Satz 5.5.1 ist $f(K) \subseteq Y$ kompakt. Folglich ist $f(K)$ auch beschränkt und abgeschlossen, vergleiche Satz 3.2.4. □

Satz 5.5.4. *Sei (X, d) ein metrischer Raum, $K \subseteq X$ eine kompakte Menge und $f : K \to \mathbb{R}$ eine stetige Funktion. Ferner seien $m := \inf_{x \in X} f(x)$ und $M := \sup_{x \in X} f(x)$. Dann existieren $p, q \in X$ mit $f(p) = m$ und $f(q) = M$.*

Beweis. Nach Satz 5.5.3 ist $f(x) \subseteq \mathbb{R}$ eine abgeschlossene und beschränkte Menge. Somit gilt $m, M \in f(K)$, d.h. es existieren $p, q \in X$ mit $f(p) = m$ und $f(q) = M$. □

Satz 5.5.5. *Seien (X, d_X) und (Y, d_Y) metrische Räume und die Menge $K \subseteq X$ kompakt. Die Funktion $f : K \to Y$ sei stetig und bijektiv. Dann ist die inverse Abbildung $f^{-1} : Y \to X$ ebenfalls stetig auf Y.*

Beweis. Wegen Satz 5.2.5 genügt es zu zeigen, dass die Bildmenge $f(V)$ einer offenen Menge $V \subseteq K$ in Y offen ist. Nun ist das Komplement von V in K, d.h. $K \setminus V$, abgeschlossen in X. Eine abgeschlossene Teilmenge einer kompakten Menge ist wieder kompakt (Satz 3.2.5) und infolge Satz 5.5.1 ist daher $f(K \setminus V) \subseteq Y$ kompakt und somit abgeschlossen in Y. Wegen der Bijektivität gilt

$$Y \setminus f(K \setminus V) = f(V)$$

und daher ist die Menge $f(V)$ offen. □

Definition 5.5.6. *Seien (X, d_X) und (Y, d_Y) metrische Räume. Eine Funktion $f : X \to Y$ heißt* **gleichmäßig stetig**, *falls zu jedem $\varepsilon > 0$ ein $\delta > 0$ existiert derart, dass für alle $x, y \in X$ mit $d_X(x, y) < \delta$ gilt*

$$d_Y\big(f(x), f(y)\big) < \varepsilon.$$

Bemerkung 5.5.7. *Die gleichmäßige Stetigkeit ist eine globale Eigenschaft. Die Funktion $f : (0, \infty) \to \mathbb{R}$ gegeben durch $f(x) = \frac{1}{x}$ ist in allen Punkten $x \in (0, \infty)$ stetig, aber nicht gleichmäßig stetig.*

Satz 5.5.8. *Seien (X, d_X) und (Y, d_Y) metrische Räume und die Menge $K \subseteq X$ kompakt. Ist die Funktion $f : K \to Y$ stetig, so ist f auch gleichmäßig stetig auf K.*

Beweis. Sei $\varepsilon > 0$ beliebig aber fest gewählt. Wegen der Stetigkeit gibt es zu jedem $x_0 \in X$ eine Zahl $\delta_{x_0} > 0$ mit $d_Y\big(f(x), f(x_0)\big) < \frac{\varepsilon}{2}$ für alle $x \in K$ mit $d_X(x, x_0) < \delta_{x_0}$. Wir betrachten die Mengen V_{x_0}, $x_0 \in K$, definiert durch

$$V_{x_0} = \left\{ x \in X : d_X(x, x_0) < \frac{\delta_{x_0}}{2} \right\}.$$

Dann ist $\{V_{x_0} : x_0 \in K\}$ eine offene Überdeckung von K. Wegen der Kompaktheit von K existiert eine endliche Teilüberdeckung $\{V_{x_i} : i = 1, \ldots, N\}$. Sei nun $\delta := \min\{\delta_{x_i} : i = 1, \ldots, N\}$, dann gibt es zu jedem $x, y \in K$ mit $d_X(x, y) < \delta$ einen Index $m \in \{1, \ldots, N\}$ mit

$$d_X(x, x_m) < \delta_{x_m}$$

und

$$d_X(y, x_m) < \delta_{x_m}.$$

Somit gilt

$$d_Y\big(f(x), f(y)\big) \leq d_Y\big(f(x), f(x_m)\big) + d_Y\big(f(x_m), f(y)\big) < \varepsilon.$$

$\square$

Kapitel 6

Differentiation von Funktionen einer Veränderlichen

6.1 Differenzenquotient und Differentialquotient

Motivation: Sei ein Graph G der Funktion $f : (a, b) \to \mathbb{R}$ gegeben. Gesucht ist die Gleichung der Tangente an G in dem Punkt $P\big(x_0, f(x_0)\big)$ mit $x_0 \in (a, b)$.

Der Anstieg der Sekante durch die Punkte $\big(x_0, f(x_0)\big)$ und $\big(x, f(x)\big)$ ist durch den *Differenzenquotienten* von f an der Stelle x_0 gegeben

$$\frac{\Delta x}{\Delta y} = \frac{f(x) - f(x_0)}{x - x_0}.$$

Definition 6.1.1. 1. *Die Funktion $f : (a, b) \to \mathbb{R}$ heißt an der Stelle $x_0 \in (a, b)$* **differenzierbar**, *wenn der Grenzwert $\lim_{x \to x_0} \frac{f(x)-f(x_0)}{x-x_0} =: f'(x_0) \in \mathbb{R}$ existiert (Differentialquotient $\frac{dy}{dx}$). Die Zahl $f'(x_0)$ heißt die* **erste Ableitung** *von f an der Stelle x_0.*

2. *Die Funktion heißt* **differenzierbar auf** (a, b), *wenn f an jeder Stelle $x_0 \in (a, b)$ differenzierbar ist. In diesem Fall ordnen wir x_0 eindeutig $f'(x_0)$ zu. Wir erhalten damit eine Funktion $f' : (a, b) \to \mathbb{R}$, genannt die erste Ableitung von f auf (a, b).*

Satz 6.1.2. *Die Funktion $f : (a, b) \to \mathbb{R}$ ist genau dann an der Stelle $x_0 \in (a, b)$ differenzierbar, wenn eine Zahl $A \in \mathbb{R}$ sowie eine Funktion $\varphi : (a, b) \to \mathbb{R}$ mit $\lim_{x \to x_0} \varphi(x) = 0$ existieren, so dass*

$$f(x) = f(x_0) + \big(A + \varphi(x)\big)(x - x_0)$$

für alle $x \in (a, b)$.

Beweis. "$\Rightarrow$": Sei f differenzierbar in x_0, d.h. es existiert $\lim_{x \to x_0} \frac{f(x)-f(x_0)}{x-x_0} =: A$. Wir setzen

nun $\varphi : (a, b) \to \mathbb{R}$ wie folgt

$$\varphi(x) = \begin{cases} \frac{f(x)-f(x_0)}{x-x_0} - A, & x \neq x_0, \\ 0, & x = x_0. \end{cases}$$

Für alle $x \neq x_0$ ist dann in der Tat $f(x) - f(x_0) = (A + \varphi(x))(x - x_0)$. Natürlich gilt auch $\lim_{x\to x_0} \varphi(x) = 0$ und somit die Behauptung auch für $x = x_0$.

"$\Leftarrow$": Es gilt

$$\lim_{x\to x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{x\to x_0} \big(A + \varphi(x)\big) = A + 0 = A,$$

also $f'(x_0) = A$. □

Bemerkung 6.1.3. *Die Funktion $t : \mathbb{R} \to \mathbb{R}$ mit $t(x) = f(x_0) + f'(x_0)(x - x_0)$ ist die Gleichung der Tangente an die Kurve $y = f(x)$ im Punkt $\big(x_0, f(x_0)\big)$.*

Satz 6.1.4. *Wenn die Funktion $f : (a, b) \to \mathbb{R}$ an der Stelle $x_0 \in (a, b)$ differenzierbar ist, dann ist sie stetig in x_0.*

Beweis. Da x_0 Häufungspunkt von (a, b) ist, so gilt

$$\lim_{x\to x_0} f(x) = \lim_{x\to x_0} \big(f(x_0) + (A + \varphi(x))(x - x_0)\big) = f(x_0).$$

Somit folgt gemäß Satz 5.2.3, dass f stetig in x_0 ist. □

Bemerkung 6.1.5. *Die Umkehrung dieses Satzes gilt i.a. nicht, vergleiche $f(x) = |x|$ im Falle $x_0 = 0$.*

Satz 6.1.6. *Wenn die Funktionen $f : (a, b) \to \mathbb{R}$ und $g : (a, b) \to \mathbb{R}$ an einer Stelle $x_0 \in (a, b)$ differenzierbar sind, so sind auch die Funktionen $f + g$, $f \cdot g$, $\alpha \cdot f$ mit $\alpha \in \mathbb{R}$ differenzierbar in x_0. Ist $g(x_0) \neq 0$, so ist ferner auch $\frac{f}{g}$ differenzierbar in x_0. Dabei gelten die folgenden Rechenregeln:*

1. $$(f + g)'(x_0) = f'(x_0) + g'(x_0),$$
2. $$\big(\alpha \cdot f(x_0)\big)' = \alpha \cdot f'(x_0),$$
3. $$(f \cdot g)'(x_0) = f'(x_0)g(x_0) + g'(x_0)f(x_0) \qquad \textit{(Produktregel)},$$
4. $$\Big(\frac{f}{g}\Big)'(x_0) = \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g^2(x_0)} \qquad \textit{(Quotientenregel)}.$$

Beweis. Die ersten beiden Aussagen sind trivial. Der Beweis der Produktregel ergibt sich aus

$$\begin{aligned}\frac{(g\cdot f)(x)-(g\cdot f)(x_0)}{x-x_0} &= \frac{1}{x-x_0}\Big[f(x)g(x)-f(x_0)g(x_0)\Big]\\ &= \frac{1}{x-x_0}\Big[f(x)g(x)-f(x)g(x_0)+f(x)g(x_0)-f(x_0)g(x_0)\Big]\\ &= f(x)\frac{g(x)-g(x_0)}{x-x_0}+g(x_0)\frac{f(x)-f(x_0)}{x-x_0}\\ &\overset{x\to x_0}{\longrightarrow} f(x_0)g'(x_0)+g(x_0)f'(x_0).\end{aligned}$$

Schließlich zeigen wir die Quotientenregel. Da $g(x_0)\neq 0$, ist g in x_0 stetig, d.h. es existiert ein $\epsilon>0$, so dass $g(x)\neq 0$ für alle $x\in(x_0-\epsilon,x_0+\epsilon)\subset(a,b)$. Damit existiert die Funktion $q:=\frac{1}{g}:(x_0-\epsilon,x_0+\epsilon)\to\mathbb{R}$ mit $q(x)=\frac{1}{g(x)}$. Wir betrachten

$$\begin{aligned}\frac{q(x)-q(x_0)}{x-x_0} &= \frac{1}{x-x_0}\Big[\frac{1}{g(x)}-\frac{1}{g(x_0)}\Big]\\ &= \frac{1}{x-x_0}\Big[\frac{g(x_0)-g(x)}{g(x)g(x_0)}\Big]\\ &= -\frac{1}{g(x)g(x_0)}\Big[\frac{g(x)-g(x_0)}{x-x_0}\Big]\\ &\overset{x\to x_0}{\longrightarrow} -\frac{g'(x_0)}{g^2(x)},\end{aligned}$$

d.h. $q'(x)=-g'(x_0)/g^2(x)$. Sei nun $h:(x_0-\epsilon,x_0+\epsilon)\to\mathbb{R}$ gegeben durch $h(x)=\frac{f(x)}{g(x)}=f(x)\cdot q(x)$, dann folgt die Behauptung mit Hilfe der Produktregel. □

Satz 6.1.7 (Kettenregel). *Seien $f:(a,b)\to(c,d)$ und $g:(c,d)\to\mathbb{R}$ Funktionen und $h:(a,b)\to\mathbb{R}$ gegeben durch $h(x)=g(f(x))$, d.h. $h=g\circ f$. Ist die Funktion f an der Stelle $x_0\in(a,b)$ und die Funktion g an der Stelle $y_0=f(x_0)$ differenzierbar, so ist die Funktion h ebenfalls an der Stelle x_0 differenzierbar und es gilt*

$$h'(x_0)=g'(y_0)\cdot f'(x_0)=g'\big(f(x_0)\big)\cdot f'(x_0).$$

$[z=h(x)=g(f(x)),y=f(x)\Rightarrow\frac{dz}{dx}=\frac{dz}{dy}\cdot\frac{dy}{dx}]$

Beweis. Aus der Differenzierbarkeit folgt die Existenz von Funktionen $\varphi:(a,b)\to\mathbb{R}$ und $\psi:(c,d)\to\mathbb{R}$ mit $\lim_{x\to x_0}\varphi(x)=\lim_{y\to y_0}\psi(x)=0$ und

$$\begin{aligned}f(x) &= f(x_0)+[f'(x_0)+\varphi(x)](x-x_0), &&\text{für alle } x\in(a,b),\\ g(y) &= g(y_0)+[g'(y_0)+\psi(y)](y-y_0), &&\text{für alle } y\in(c,d).\end{aligned}$$

Somit ist

$$\begin{aligned} h(x) - h(x_0) &= g\big(f(x)\big) - g\big(f(x_0)\big) \\ &= [g'(y_0) + \psi(y)](y - y_0) \\ &= \big[g'\big(f(x_0)\big) + \psi\big(f(x_0)\big)\big][f'(x_0) + \varphi(x)](x - x_0). \end{aligned}$$

Beim Grenzübergang $x \to x_0$ gilt nun wegen der Stetigkeit von f dass $y = f(x) \to f(x_0) = y_0$ und somit $\psi\big(f(x)\big) \to 0$ und $\varphi(x) \to 0$ für $x \to x_0$. Demnach folgt $\frac{h(x)-h(x_0)}{x-x_0} \to g'(f(x_0))f'(x_0)$ für $x \to x_0$. $\square$

Ableitungen wichtiger Funktionen	
f	f'
x^α	$\alpha x^{\alpha-1}$
$\sin x$	$\cos x$
$\cos x$	$-\sin x$
$\tan x$	$\dfrac{1}{\cos^2 x}$
$\cot x$	$-\dfrac{1}{\sin^2 x}$
$\arcsin x$	$\dfrac{1}{\sqrt{1-x^2}}$
$\arccos x$	$-\dfrac{1}{\sqrt{1-x^2}}$
$\arctan x$	$\dfrac{1}{1+x^2}$
$\operatorname{arccot} x$	$-\dfrac{1}{1+x^2}$
e^x	e^x
a^x $(a > 0)$	$a^x \ln a$
$\ln x$	$\dfrac{1}{x}$
$\log_a x$ $(a > 0, a \neq 1)$	$\dfrac{1}{x \ln a}$

Satz 6.1.8 (Ableitung der Umkehrfunktion). *Sei $f : [a, b] \to [c, d]$ eine bijektive Funktion, die in $x_0 \in [a, b]$ differenzierbar ist. Dann ist die inverse Funktion $f^{(-1)} : [c, d] \to [a, b]$ an der Stelle*

$y_0 := f(x_0)$ *differenzierbar und es gilt*

$$(f^{(-1)})'(y_0) = \frac{1}{f'(f^{(-1)}(y_0))}.$$

Beweis. Die Stetigkeit von f impliziert gemäß Satz 5.5.5 die Stetigkeit der Umkehrfunktion $f^{(-1)}$. Damit gilt für eine beliebige Folge (y_n) mit $y_n \to y_0$ auch $x_n := f^{(-1)}(y_n) \to f^{(-1)}(y_0) =: x_0$ für $n \to \infty$. Daher gilt

$$\frac{f^{(-1)}(y_n) - f^{(-1)}(y_0)}{y_n - y_0} = \frac{x_n - x_0}{f(x_n) - f(x_0)} \overset{n\to\infty}{\longrightarrow} \frac{1}{f'(x_0)},$$

womit die Behauptung bewiesen ist. □

Beispiel 6.1.9. *Die Umkehrkunktion der Exponentialfunktion $y = f(x) = e^x$ ist der Logarithmus $f^{(-1)}(y) = \log y$. Wir finden $f'(x) = e^x$ und somit für die Ableitung des Logarithmus $(f^{(-1)})'(y) = (\log y)' = 1/e^x = 1/y$.*

6.2 Mittelwertsätze

Definition 6.2.1. *Sei $f : X \to \mathbb{R}$ stetig. Diese Funktion hat an der Stelle x_0 ein* **lokales Minimum**, *falls ein $\delta > 0$ existiert, so dass $f(x) \geq f(x_0)$ für alle $x \in X$ mit $d_X(x, x_0) < \delta$. Analog definiert man ein* **lokales Maximum**.

Satz 6.2.2 (Fermat). *Sei $f : [a, b] \to \mathbb{R}$ stetig auf $[a, b]$. Falls f an der Stelle $x_0 \in (a, b)$ ein lokales Minimum bzw. Maximum besitzt und f in x_0 differenzierbar ist, so gilt $f'(x_0) = 0$.*

Beweis. Aufgrund der Voraussetzung existiert ein $\delta < 0$ mit $a < x_0 - \delta < x < x_0 + \delta < b$ und $f(x) \geq f(x_0)$ für alle $x \in (x_0 - \delta, x_0 + \delta)$. Für beliebiges $x_0 < x < x_0 + \delta$ ist somit $\frac{f(x)-f(x_0)}{x-x_0} \geq 0$. Andererseits ist für beliebiges $x_0 - \delta < x < x_0$ aber $\frac{f(x)-f(x_0)}{x-x_0} \leq 0$. Dies bedeutet, dass für $x \to x_0$ wegen Differenzierbarkeit folgt $\frac{f(x)-f(x_0)}{x-x_0} = 0$. □

Satz 6.2.3 (Verallgemeinerter Mittelwertsatz). *Seien $f, g : [a, b] \to \mathbb{R}$ stetig auf $[a, b]$ und in allen Punkten $x \in (a, b)$ differenzierbar, dann existiert ein $x_0 \in (a, b)$ mit $[f(a) - f(b)]g'(x_0) = [g(a) - g(b)]f'(x_0)$.*

Beweis. Wir betrachten die Funktion $h(x) := [f(a) - f(b)]g(x) - [g(a) - g(b)]f(x)$. Diese Funktion ist stetig auf $[a, b]$ und differenzierbar in allen Punkten $x \in (a, b)$. Ferner folgt aus $h(a) = g(b)f(a) - f(b)g(a)$ und $h(b) = f(a)g(b) - f(b)g(a)$, dass $h(a) = h(b)$.

Ist h konstant, dann ist nichts zu zeigen. Sei also h nicht konstant. Dann gibt es ein $y \in (a, b)$ mit $h(y) \neq h(a)$. Ohne Beschränkung der Allgemeinheit können wir annehmen, dass $h(y) > h(a)$. Nach Satz 5.5.4 nimmt h ihr Maximum $M := \max_{x\in[a,b]} h(x)$ für ein $x_0 \in [a, b]$ an. Da aber $h(a) = h(b) < h(y) \leq M$ gilt sogar $x_0 \in (a, b)$. Insbesondere gilt $h'(x_0) = 0$, woraus die Behauptung folgt. □

Korollar 6.2.4 (Mittelwertsatz). *Sei $f : [a.b] \to \mathbb{R}$ stetig und in allen Punkten $x \in (a, b)$ differenzierbar, so existiert eine Zwischenstelle $\xi \in (a, b)$ mit $f(a) - f(b) = (a - b)f'(\xi)$.*

Beweis. Setze $g(x) := x$, so folgt die Aussage direkt aus Satz 6.2.3. □

Satz 6.2.5. *Sei $f : [a, b] \to \mathbb{R}$ differenzierbar und sei $\lambda \in \mathbb{R}$ beliebig mit $f'(a) < \lambda < f'(b)$, dann gibt es ein $x_0 \in (a, b)$ mit $f'(x_0) = \lambda$.*

Beweis. Sei $g(x) := f(x) - \lambda x$, dann ist g ebenfalls differenzierbar auf $[a, b]$ und es gilt $g'(a) = f(a) - \lambda < 0$. Daher existiert ein $a < \underline{x} < b$ mit $g(a) > g(\underline{x})$. Andererseits gilt $g'(b) = f'(b) - \lambda > 0$ und somit gibt es ein $a < \overline{x} < b$ mit $g(\overline{x}) < g(b)$. Da $g : [a, b] \to \mathbb{R}$ auf $[a, b]$ nach Satz 6.1.4 stetig ist, existiert ein $x_0 \in (a, b)$ mit $g(x_0) \leq g(x)$ für alle $x \in (a, b)$, d.h. g nimmt an der Stelle x_0 sein Minimum an. Damit liegt an dieser Stelle ein lokales Minimum vor. Wegen Satz 6.2.2 gilt $g'(x_0) = 0$, d.h. $0 = g'(x_0) = f'(x_0) - \lambda$ und somit $f'(x_0) = \lambda$. □

Definition 6.2.6. *Sei $f : [a, b] \to \mathbb{R}$ und es existiere in allen $x \in [a, b]$ die erste Ableitung $f'(x)$. Falls die Funktion $x \mapsto f'(x)$ ebenfalls in $x_0 \in [a, b]$ differenzierbar ist, dann heißt der Grenzwert*

$$\lim_{x \to x_0} \frac{f'(x) - f'(x_0)}{x - x_0} =: f''(x_0)$$

die **zweite Ableitung** *der Funktion an der Stelle x_0. Analog definieren wir die höheren Ableitungen $f'''(x)$, $f^{(4)}(x), \ldots, f^{(k)}(x)$.*

Falls die Funktion $f^{(k)}(x)$ für alle $x \in [a, b]$ existiert und ebenfalls stetig ist, dann heißt die Funktion f **k-mal stetig differenzierbar** *auf $[a, b]$. Wir schreiben hierfür kurz $f \in C^k([a, b])$ bzw. $f \in C^{(k)}([a, b], \mathbb{R})$. Dabei bezeichnet $f \in C^k([a, b])$ bzw. $C^{(k)}([a, b], \mathbb{R})$* **die Menge (der Raum) der k-mal stetig differenzierbaren Funktion**.

Satz 6.2.7 (Satz von Taylor). *Sei $f \in C^{k-1}([a, b])$ und $f^{(k)}$ existiere in allen $x \in [a, b]$. Zu $x_0 \in (a, b)$ definieren wir das Taylor-Polynom $T_{k-1}(x) = T_{k-1}f(x, x_0)$ via*

$$T_{k-1}(x) := f(x_0) + f'(x_0)(x - x_0) + f''(x_0) \cdot \frac{1}{2!} \cdot (x - x_0)^2 + \ldots + \frac{f^{(k-1)}}{(k-1)!}(x - x_0)^{k-1}$$

(Taylorentwicklung um den Entwicklungspunkt x_0 der Ordnung $k - 1$). Dann gibt es zu jedem $x \in [a, b]$ eine Zwischenstelle $\xi \in (x_0, x)$ bzw. $\xi \in (x, x_0)$ mit $x \neq x_0$ derart, dass

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

Beweis. Seien x und x_0 beliebig, aber fest. Ohne Beschränkung der der Allgemeinheit sei $x > x_0$. Dann sei $\lambda \in \mathbb{R}$ so gewählt, dass $f(x) - T_{k-1}(x) = \lambda(x - x_0)^k$, also $\lambda := \frac{f(x) - T_{k-1}(x)}{(x - x_0)^k}$. Wir

definieren $g : [a, b] \to \mathbb{R}$ durch $g(t) = f(t) - T_{k-1}(t) - \lambda(t - x_0)^k$. Dann gelten

$$\begin{aligned} f(x_0) &= T_{k-1}(x_0), \\ f'(x_0) &= T'_{k-1}(x_0), \\ &\vdots \\ f^{(k)}(x_0) &= T^{(k)}_{k-1}(x_0), \end{aligned}$$

und

$$g^{(k)}(x_0) = f^{(k)}(x_0) - k!\lambda, \qquad g^{(k)}(x) = f^{(k)}(x) - k!\lambda.$$

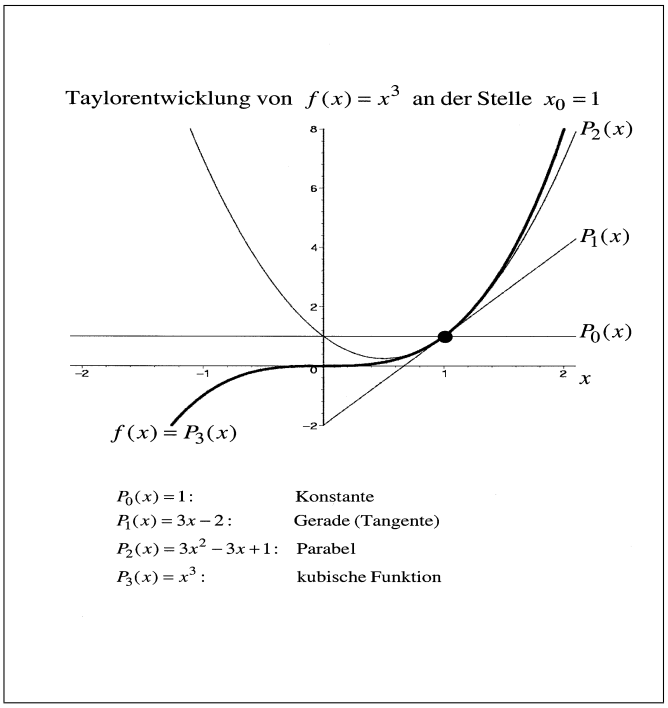
Da weiterhin gilt $g(x_0) = 0$ und $g(x) = f(x) - T_{k-1}(x) - \lambda(x - x_0)^k = 0$, existiert aufgrund des Mittelwertsatzes der Differentialrechnung ein $\xi_1 \in (x, x_0)$ mit $g'(\xi_1) = 0$.

Nun ist aber auch $g'(x_0) = f'(x_0) - T'_{k-1}(x_0) - \lambda k(x - x_0)^{k-1} = 0$. Die Anwendung des Mittelwertsatzes auf die Funktion g' liefert die Existenz einer Zwischenstelle $\xi_2 \in (x_0, \xi_1)$ mit $g''(\xi_2) = 0$.

Da wiederum $g''(x_0) = f''(x_0) - T''_{k-1}(x_0) = 0$ und allgemein $g^{(l)}(x_0) = 0$ für alle $l = 1, 2, \ldots, k-1$ sind, finden wir nach diesem Konstruktionsprinzip Zwischenstellen $\xi_l \in (x_0, \xi_{l-1})$, $l = 1, 2, \ldots, k$ mit $g^{(l)}(\xi_l) = 0$.

Wir setzen schließlich $\xi = \xi_k$ und folgern aus $0 = g^{(k)}(\xi) = f^{(k)}(\xi) - k!\lambda$ die Beziehung $\lambda = \frac{f^{(k)}(\xi)}{k!}$. Die Behauptung folgt dann aus $g(x) = 0 = f(x) - T_{k-1}(x) - \frac{f^{(k)}(\xi)}{k!}(x - x_0)^k$. $\square$

Beispiel 6.2.8.



6.3 Die l'Hospitalsche Regel

Der folgende Satz ist of hilfreich bei der Berechnung von Grenzwerten.

Satz 6.3.1 (l'Hospitalsche Regel). *Seien f und g reell und differenzierbar in (a, b) mit $-\infty \leq a < b \leq +\infty$. Ferner sei $g'(x) \neq 0$ für alle $x \in (a, b)$ und es gelte*

$$\lim_{x \to a} \frac{f'(x)}{g'(x)} = C,$$

wobei $C = \pm\infty$ zugelassen sei. Vorausgesetzt dass

$$\lim_{x \to a} f(x) = 0 \quad \textit{und} \quad \lim_{x \to a} g(x) = 0 \tag{6.1}$$

oder

$$\lim_{x \to a} f(x) = \pm\infty \quad \textit{und} \quad \lim_{x \to a} g(x) = \pm\infty, \tag{6.2}$$

so gilt

$$\lim_{x \to a} \frac{f(x)}{g(x)} = C.$$

Beweis. Wir nehmen an es gelte (6.1). Zunächst betrachten wir den Fall $-\infty \leq C < +\infty$. Wir wählen eine Zahl $\overline{C}$ mit $C < \overline{C}$ und ein $c \in (C, \overline{C})$. Nach Voraussetzung gibt es einen Punkt $\overline{x} \in (a, b)$, so dass für $x \in (a, \overline{x})$ stets

$$\frac{f'(x)}{g'(x)} < c$$

folgt. Für $a < x < y < \overline{x}$ erhält man nach dem verallgemeinerten Mittelwertsatz 6.2.3 eine Zwischenstelle $\xi \in (x, y)$ mit

$$\frac{f(x) - f(y)}{g(x) - g(y)} = \frac{f'(\xi)}{g'(\xi)} < c.$$

Folglich ergibt sich

$$\lim_{x \to a} \frac{f(x) - f(y)}{g(x) - g(y)} = \frac{f(y)}{g(y)} \leq c < \overline{C}$$

für alle $y \in (a, \overline{x})$. Damit ist die Behauptung bewiesen im Falle $C = -\infty$. Im Falle $-\infty < C \leq +\infty$ findet man auf analoge Weise zu einem $\underline{C} < C$ einen Punkt $\underline{x} \in (a, b)$ so, dass

$$\frac{f(y)}{g(y)} > \underline{C}$$

für alle $y \in (a, \underline{x})$. Da $\underline{C}$ und $\overline{C}$ beliebig gewählt werden können, ist somit die Behauptung für alle $-\infty \leq C \leq +\infty$ bewiesen.

Wir müssen den Satz noch im Falle von Voraussetzung (6.2) beweisen. Dabei gelte ohne Einschränkung der Allgemeinheit $\lim_{x\to a} g(x) = +\infty$. Zunächst nehmen wir wieder an, dass $-\infty \leq C < +\infty$. Wie oben finden wir zu $\overline{C} > C$ und $c \in (C, \overline{C})$ ein $\overline{x} \in (a, b)$, so dass

$$\frac{f'(x)}{g'(x)} < c$$

für alle $x \in (a, \overline{x})$ und demnach

$$\frac{f(x) - f(y)}{g(x) - g(y)} < c$$

für alle $a < x < y < \overline{x}$. Wir wollen dabei y so wählen, dass $g(x) > g(y)$ und $g(x) > 0$ für alle $x \in (a, y)$. Dann erhalten wir aus der letzten Beziehung durch Multiplikation mit $\big(g(x) - g(y)\big)/g(x) > 0$ dass

$$\frac{f(x) - f(y)}{g(x)} < c\frac{g(x) - g(y)}{g(x)}$$

bzw.

$$\frac{f(x)}{g(x)} < c - c\frac{g(y)}{g(x)} + \frac{f(y)}{g(x)}.$$

Da $\lim_{x\to a} g(x) = +\infty$ finden wir ein $\overline{x} \in (a, y)$ mit

$$\frac{f(x)}{g(x)} < \overline{C}$$

für alle $x \in (a, \overline{x})$. Damit ist die Behauptung im Falle $C = -\infty$ schon gezeigt. Falls $-\infty < C \leq +\infty$ findet man analog zu einem $\underline{C} < C$ einen Punkt $\underline{x} \in (a, b)$ so, dass

$$\frac{f(x)}{g(x)} > \underline{C}$$

für alle $x \in (a, \underline{x})$. Daraus folgt wieder die Behauptung. □

Bemerkung 6.3.2.

1. *Die analoge Aussage ist natürlich auch wahr für $x \to b$.*

2. *Häufig tritt bei der Anwendung dieser Regel der Fall ein, dass auch $\lim_{x\to a} f'(x)/g'(x)$ vom Typus "$0/0$" bzw. "∞/∞" ist. Man wendet dann die Regel von l'Hospital nochmals an, das heißt man betrachtet den Grenzwert $\lim_{x\to a} f''(x)/g''(x)$. Gegebenenfalls muss das Verfahren weiter fortgesetzt werden. Allerdings kann es durchaus vorkommen, dass alle Ableitungen vom unbestimmten Typus sind, wie das Beispiele $f(x) = \sinh(x)$ und $g(x) = \cosh(x)$ im Falle $x \to \pm\infty$ zeigen.*

Kapitel 7

Integration von Funktionen einer reellen Variablen

7.1 Das unbestimmte Integral

Sei im folgenden I eine zusammenhängende Teilmenge von $\mathbb{R}$, d.h. $I = [a, b]$, (a, b), $(-\infty, b]$, etc.

Definition 7.1.1. *Sei $f : I \to \mathbb{R}$. Wenn eine Funktion $F : I \to \mathbb{R}$ in I differenzierbar ist, dann heißt F* **Stammfunktion** *von f, falls $F'(x) = f(x)$ für alle $x \in I$ gilt.*

Proposition 7.1.2. *Sei $F \to \mathbb{R}$ eine Stammfunktion von f, dann gilt:*

1. *Für jedes beliebiges $C \in \mathbb{R}$ ist die Funktion $x \mapsto F(x) + C$ ebenfalls eine Stammfunktion von f.*
2. *Sei G eine weitere Stammfunktion von f, dann existiert ein $C \in \mathbb{R}$ mit $G(x) = F(x) + C$.*

Beweis.

1. Es gilt $\big(F(x) + C\big)' = F'(x) + 0 = f(x)$.
2. Sei G eine weitere Stammfunkion von f. Wir betrachten $F - G$ und finden
$$\big(F(x) - G(x)\big)' = F'(x) - G'(x) = f(x) - f(x) = 0,$$
d.h. $F(x) - G(x) = C$ für ein $C \in \mathbb{R}$.

□

Definition 7.1.3. *Die Gesamtheit aller Stammfunktionen von f bezeichnen wir als das* **unbestimmte Integral** *von f, kurz $\int f dx = \int f(x)dx = F(x) + C$ mit $C \in \mathbb{R}$ beliebig.*

Bemerkung 7.1.4. *Sei $f \in C^1(I)$, dann ist die Operation $D := \frac{d}{dx}$, $f \mapsto D(f) := f'$ nicht injektiv. Aber die (inverse) Operation $D^{-1} := \int \cdot\, dx$, $f \mapsto D^{-1} := F$ (F Stammfunktion von f) ist die Rechtsinverse zu D, da $D(D^{-1})(f) = D(F) = F' = f$.*

Satz 7.1.5. *Seien $f, g : I \to \mathbb{R}$ zwei Funktionen mit den Stammfunktionen F und G, d.h. $F' = f$ bzw. $G' = g$, und $a \in \mathbb{R}$.*

1. *Es gilt*
$$\int af(x)dx = a \cdot \int f(x)dx.$$

2. *Es gilt*
$$\int (f+g)dx = \int fdx + \int gdx.$$

3. *Seien f, g differenzierbar, dann gilt*
$$\int fg'dx = f \cdot g - \int f'gdx$$
(partielle Integration).

4. *Sei $\varphi : J \to I$ differenzierbar und bijektiv, dann gilt*
$$\int f\big(\varphi(x)\big)\varphi'(x)dx = F(\varphi(x)) + C, \quad C \in \mathbb{R},$$
(Substitutionsregel).

Beweis.

1. Es gilt $\big(a \cdot \int f(x)dx\big)' = aF'(x) = af(x) = \big(\int af(x)dx\big)'$.

2. Die Behauptung folgt aus
$$\begin{aligned}\left(\int fdx + \int gdx\right)' &= F'(x) + G'(x) = f(x) + g(x) = (f+g)(x)\\ &= \left(\int (f+g)dx\right)'.\end{aligned}$$

3. Mit der Produktregel folgt
$$(f \cdot g)' = f'g + fg' = \left(\int f'gdx\right)' + \left(\int fg'dx\right)' = \left(\int f'gdx + \int fg'dx\right)'.$$

4. Mit der Kettenregel folgt

$$\Big(F\big(\varphi(x)\big)+C\Big)' = \Big(F\big(\varphi(x)\big)\Big)' = F'\big(\varphi(x)\big)\cdot\varphi'(x) = f\big(\varphi(x)\big)\cdot\varphi'(x).$$

□

Grundintegrale	
$\int x^\alpha dx = \frac{1}{\alpha+1}x^{\alpha+1}+C \quad (\alpha \neq -1)$	$\int \frac{1}{x}dx = \ln\lvert x\rvert + C \quad (x \neq 0)$
$\int \sin x dx = -\cos x + C$	$\int e^x dx = e^x + C$
$\int \cos x dx = \sin x + C$	$\int a^x dx = \frac{a^x}{\ln a} \quad (a>0,\ a\neq 1)$
$\int \frac{dx}{\cos^2 x} = \tan x + C$	$\int \frac{dx}{1+x^2} = \arctan x + C$
$\int \frac{dx}{\sin^2 x} = -\cot x + C$	$\int \frac{dx}{\sqrt{1-x^2}} = \arcsin x + C \quad (\lvert x\rvert < 1)$
Weitere Integrale $\quad (a>0)$	
$\int \frac{f'(x)}{f(x)}dx = \ln\lvert f(x)\rvert + C$	$\int \frac{f'(x)}{1+f^2(x)}dx = \arctan f(x) + C$
$\int \frac{1}{a^2-x^2}dx = \frac{1}{2a}\ln\left\lvert\frac{a+x}{a-x}\right\rvert + C$	$\int \frac{1}{a^2+x^2}dx = \frac{1}{a}\arctan\frac{x}{a} + C$
$\int \sin^2 x dx = \frac{1}{2}(x-\sin x\cos x)+C$	$\int \cos^2 x dx = \frac{1}{2}(x+\sin x\cos x)+C$
$\int \tan x dx = -\ln\lvert\cos x\rvert + C$	$\int \cot x dx = \ln\lvert\sin x\rvert + C$
$\int \ln x dx = x\ln x - x + C$	$\int \frac{x}{\sqrt{a^2-x^2}}dx = -\sqrt{a^2-x^2}+C$
$\int \frac{1}{\sqrt{a^2-x^2}}dx = \arcsin\frac{x}{a} + C$	$\int \frac{1}{\sqrt{x^2\pm a^2}}dx = \ln\lvert x+\sqrt{x^2\pm a^2}\rvert + C$
$\int \sqrt{a^2-x^2}dx = \frac{x}{2}\sqrt{a^2-x^2} + \frac{a^2}{2}\arcsin\frac{x}{a} + C$	
$\int \sqrt{a^2+x^2}dx = \frac{x}{2}\sqrt{a^2+x^2} + \frac{a^2}{2}\ln\lvert x+\sqrt{a^2+x^2}\rvert + C$	

Bemerkung 7.1.6.

1. *Der natürliche Logarithmus* $x \mapsto \ln x$ *ist die Umkehrfunktion zu* $x \mapsto e^x$.

2. *Die Funktion* $x \mapsto \arcsin x$, $x \in (-1, 1)$, *ist die Umkehrfunktion zu* $x \mapsto \sin x$ *mit* $x \in (-\frac{\pi}{2}, \frac{\pi}{2})$ *(Hauptteil des* $\arcsin$*).*

3. *Die Funktion* $x \mapsto \arctan x$ *mit* $x \in \mathbb{R}$ *ist die Umkehrfunktion zu* $x \mapsto \tan x$, $x \in (-\frac{\pi}{2}, \frac{\pi}{2})$.

7.2 Integration gebrochen rationaler Funktionen

Definition 7.2.1. *Eine Funktion der Form* $x \mapsto p(x)/q(x)$, *wobei* p *und* q *Polynome bezeichnen und der Grad von* p *kleiner dem Grad von* q *ist, heißt* **echt gebrochen rational**.

Wir betrachten zunächst komplexe Funktionen $z \mapsto f(z) = p(z)/q(z)$ mit $z \in \mathbb{C}$.

Satz 7.2.2. *Jede echt gebrochen rationale Funktion* $f = \frac{p}{q}$, p *und* q *Polynome, lässt sich in der folgenden Form eindeutig darstellen*

$$f(z) = \sum_{i=1}^{k} \sum_{j=1}^{l_i} \frac{a_{i,j}}{(z - z_i)^j}, \quad k, l_i \in \mathbb{N}, \quad a_{i,j}, z_i \in \mathbb{C},$$

wobei z_i, $i = 1, \ldots, k$, *die Nullstellen des Nennerpolynoms* q *darstellen.*

Beweis. Wir führen den Existenzbeweis induktiv bezüglich dem Polynomgrad n des Nennerpolynoms q durch. Wir bemerken, der Fundamentalsatz der Algebra lautet

$$q(z) = (z - z_i)^{l_1} \ldots (z - z_i)^{l_k}, \quad l_1 + \ldots + l_k = n,$$

d.h. jede Funktion hat soviele Nullstellen, wie ihr Grad beträgt.

(i) Induktionsverankerung $n = 1$: Es ist $f(z) = \frac{a}{z+b}$, folglich $z_1 = -b$ und $a_{1,1} = a$.

(ii) Induktionsschritt $n \mapsto n + 1$: Wir gehen davon aus, dass die Aussage für $\operatorname{Grad} q = n$ bewiesen ist. Sei z_1 eine beliebige Nullstelle von q mit Vielfachheit l_1, d.h. $q(z) = (z - z_1)^{l_1} r(z)$ mit einem Polynom $r(z)$ vom Grad $r(z) = n - l_1$ und $r(z_1) \neq 0$. Nun ist

$$\frac{p(z)}{r(z)} - \frac{p(z_1)}{r(z_1)} = \frac{p(z)r(z_1) - p(z_1)r(z)}{r(z)r(z_1)} = \frac{(z - z_1)^{l_1} s(z)}{r(z)},$$

mit einem Polynom s mit $\operatorname{Grad} t < n - l_1$. Damit haben wir

$$g(z) := \frac{p(z)}{(z - z_1)^{l^1} r(z)} - \frac{p(z_1)}{(z - z_1)^{l_1} r(z_1)} = \frac{s(z)}{(z - z_1)^{l_1 - 1} r(z)}.$$

Nach der Induktionsvoraussetzung existiert die Partialbruchzerlegung von $g(z)$, d.h.

$$f(z) = \frac{p(z)}{(z - z_1)^p r(z)} = g(z) + \frac{p(z_1)}{(z - z_1)^{l_1} r(z_1)}$$

ist Partialbruchzerlegung von f.

Um die Eindeutigkeit dieser Zerlegung zu zeigen, nehmen wir an, $f(z)$ erlaube eine weitere Partialbruchzerlegung mit Koeffizienten $b_{i,j}$. Wir substrahieren die beiden Zerlegungen und erhalten

$$0 = \sum_{i=1}^{k} \sum_{j=1}^{l_i} \frac{a_{i,j} - b_{i,j}}{(z - z_i)^j}.$$

Für beliebiges $1 \leq m \leq k$ führt Multiplikation dieser Gleichung mit $(z - z_m)^{l_m}$ und Einsetzen von $z := z_m$ auf $a_{m,l_m} = b_{m,l_m}$. Berücksichtigt man dies, liefert die Multiplikation der Gleichung mit $(z - z_m)^{l_m - 1}$ und Einsetzen von $z := z_m$ auf $a_{m,l_m-1} = b_{m,l_m-1}$. So fortfahrend schließt man auf die Identität aller Koeffizienten, d.h. die Eindeutigkeit der Partialbruchzerlegung. □

Korollar 7.2.3 (Reelle Partialbruchzerlegung). *Jede echt gebrochene rationale reelle Funktion $f(x) = \frac{p(x)}{q(x)}$ besitzt eine Partialbruchzerlegung*

$$f(x) = \sum_{i=1}^{k} \sum_{j=1}^{l_i} \frac{a_{i,j}}{(x - x_i)^j} + \sum_{i=1}^{k'} \sum_{j=1}^{l'_i} \frac{b_{i,j}x + c_{i,j}}{(x^2 + \alpha_i x + \beta_i)^j},$$
$$k, k', l_i, l'_i \in \mathbb{N}, \quad a_{i,j}, b_{i,j}, c_{i,j}, \alpha_i, \beta_i \in \mathbb{R}.$$

Mit Hilfe der Partialbruchzerlegung kann man nun Integrale beliebiger rationaler Funktionen $f(x) := p(x)/q(x)$ mit Polynomen p und q bestimmen. Falls $\operatorname{Grad} p \geq \operatorname{Grad} q$ erhält man mittels Polynomdivision die Darstellung

$$f(x) = p(x)/q(x) = r(x) + s(x)/t(x),$$

wobei nun $\operatorname{Grad} s < \operatorname{Grad} t$. Dann berechnet man bezüglich $s(x)/t(x)$ die Partialbruchzerlegung. Unter Beachtung der Stammfunktionen $\int (x-a)^{-1} dx = \ln|x - a| + C$, $\int (x-a)^{-j} dx = 1/(1-j)(x-a)^{1-j}] + C$ $(j \neq 1)$ und

$$\int \frac{x + a}{x^2 + bx + c} dx = \frac{1}{2} \ln|x^2 + bx + c| + \frac{2a - b}{\sqrt{4c - b^2}} \arctan \frac{2x + b}{\sqrt{4c - b^2}}, \quad 4c - b^2 > 0,$$

kann man nun das Integral $\int f(x)dx$ berechnen falls $l'_i = 1$ für alle $1 \leq i \leq k'$, d.h. Terme der Form $(x+a)/(x^2+bx+c)^j$ für $j > 1$ nicht auftreten. Es sei jedoch angemerkt, dass letztere mit Hilfe partieller Integration ebenfalls berechenbar sind.

Beispiel 7.2.4. *Wir betrachten die rationale Funktion*

$$f(x) := \frac{x^5 - 2x^4 + 2x^3 - 2x^2 + x + 1}{(x^2 + 1)(x - 1)^2}.$$

Da der Polynomgrad des Zählers größer ist als der Polynomgrad des Nenners, führen wir zurest eine Polynomdivision mit Rest aus

$$f(x) = x + \frac{1}{(x^2 + 1)(x - 1)^2}.$$

Nun müssen wir vom echt gebrochenen Anteil die Partialbruchzerlegung bestimmen. Der Nenner hat die doppelte Nullstelle 1 *und die (konjugiert) komplexen Nullstellen* $\pm i$*. Gemäß Korollar 7.2.3 machen wir den Ansatz*

$$\frac{1}{(x^2+1)(x-1)^2} = \frac{Ax+B}{x^2+1} + \frac{C}{x-1} + \frac{D}{(x-1)^2}.$$

Bringen wir nun die rechte Seite der Gleichung auf den Hauptnenner ergibt sich

$$\begin{aligned}
&\frac{1}{(x^2+1)(x-1)^2} \\
&\quad = \frac{(Ax+B)(x-1)^2 + C(x^2+1)(x-1) + D(x^2+1)}{(x^2+1)(x-1)^2} \\
&\quad = \frac{x^3(A+C) + x^2(-2A+B-2C+D) + x(A-2B+C) + (B-C+D)}{(x^2+1)(x-1)^2}.
\end{aligned}$$

Die Polynome der Zähler der linken und rechten Seite müssen gleich sein, das heisst, die Koeffizienten müssen übereinstimmen. Dies führt uns auf die vier (linearen) Gleichungen

$$\begin{aligned}
A + C &= 0, \\
-2A + B - C + D &= 0, \\
A - 2B + C &= 0, \\
B - C + D &= 1.
\end{aligned}$$

Lösen wir dieses lineare (4×4)*-Gleichungssystem erhalten wir*

$$A = 1/2, \quad B = 0, \quad C = -1/2, \quad D = 1/2,$$

das heißt

$$f(x) = x + \frac{x}{2(x^2+1)} - \frac{1}{2(x-1)} + \frac{1}{2(x-1)^2}$$

Damit ergibt sich für f *die Stammfunktion*

$$\int f dx = \frac{x^2}{2} + \frac{1}{4}\ln|x^2+1| - \frac{1}{2}\ln|x-1| - \frac{1}{2(x-1)} + C.$$

7.3 Das Riemann-Stieltjes-Integral

Wir behandeln die Integration von (reellen) Funktionen einer reellen Variablen über einem Intervall $I = [a, b]$. Die Behandlung der Integration über Mengen werden wir erst später in einem allgemeineren Zusammenhang kennenlernen.

Definition 7.3.1. *Gegeben sei das Intervall $I = [a, b]$. Unter einer* **Zerlegung** *bzw.* **Partition** *von I verstehen wir eine endliche Menge von Punkten*

$$Z := \{x_0 = a \leq x_1 \leq \cdots \leq x_i \leq \cdots \leq x_n = b\} \subseteq I.$$

Wir schreiben $\Delta x_i := x_i - x_{i-1}$ und setzen für beschränkte Funktionen $f : I \to \mathbb{R}$

$$m_i := \inf_{t \in [x_{i-1}, x_i]} f(t), \qquad M_i := \sup_{t \in [x_{i-1}, x_i]} f(t),$$

Sei $\varphi : I \to \mathbb{R}$ eine monotone wachsende Funktion und $\Delta\varphi_i := \varphi(x_i) - \varphi(x_{i-1})$. Wir definieren für eine Zerlegung Z die **Untersumme** *durch*

$$s(Z, f, \varphi) := \sum_{i=1}^{n} m_i \Delta\varphi_i$$

bzw. die **Obersumme** *durch*

$$S(Z, f, \varphi) := \sum_{i=1}^{n} M_i \Delta\varphi_i.$$

Hiermit definieren wir

$$\overline{\int_a^b} f d\varphi := \inf_Z S(Z, f, \varphi), \qquad \underline{\int_a^b} f d\varphi := \sup_Z s(Z, f, \varphi),$$

wobei das Infimum bzw. das Supremum über alle möglichen (endlichen) Zerlegungen Z von I zu nehmen ist.

Falls

$$\overline{\int_a^b} f d\varphi = \underline{\int_a^b} f d\varphi$$

ist, heißt f bezüglich φ **Riemann-Stieltjes-integrierbar**, *kurz $f \in R(\varphi, I) = R(\varphi)$ und*

$$\int_a^b f d\varphi := \overline{\int_a^b} f d\varphi = \underline{\int_a^b} f d\varphi$$

heißt das **Riemann-Stieltjes-Integral**.

Im Sonderfall $\varphi(x) \equiv x$ ist

$$\int_a^b f d\varphi = \int_a^b f(x) dx$$

und wir sprechen von einem **Riemann-Integral**. *Die Funktion ist dann* **Riemann-integrierbar**, *kurz $f \in R(I)$.*

Bemerkung 7.3.2. *Wir bemerken, dass* $\overline{\int_a^b} f\,d\varphi$ *und* $\underline{\int_a^b} f\,d\varphi$ *definiert sind. Aus der Beschränktheit von* f *folgt für ein hinreichend großes* $R > 0$*, dass* $|f(x)| \leq R$ *für alle* $x \in I$ *ist. Somit gilt für eine beliebige Zerlegung* Z

$$\begin{aligned} S(Z,f,\varphi) &= \sum_{i=1}^n M_i \Delta\varphi_i \leq \sum_{i=1}^n R\Delta\varphi_i = R\sum_{i=1}^n \Delta\varphi_i = R\sum_{i=1}^n \varphi(x_i) - \varphi(x_{i-1}) \\ &= R\big(\varphi(x_n) - \varphi(x_0)\big) = R\big(\varphi(b) - \varphi(a)\big) \end{aligned}$$

und analog

$$s(Z,f,\varphi) \geq -R\sum_{i=1}^n \Delta\varphi_i = -R\big(\varphi(b) - \varphi(a)\big).$$

Damit folgt

$$-R\big(\varphi(b) - \varphi(a)\big) \leq s(Z,f,\varphi) \leq S(Z,f,\varphi) \leq R\big(\varphi(b) - \varphi(a)\big),$$

d.h. die Mengen $\{S(Z,f,\varphi) \,:\, Z \text{ ist Zerlegung}\}$ *bzw.* $\{s(Z,f,\varphi) \,:\, Z \text{ ist Zerlegung}\}$ *sind beschränkt und besitzen somit Infimum und Supremum.*

Definition 7.3.3. *Eine Zerlegung* $Z^\star$ *von* $I = [a,b]$ *heißt eine* **Verfeinerung** *der Zerlegung* Z *von* I*, falls* $Z^\star \supseteq Z$.

Proposition 7.3.4. *Sei* $Z^\star$ *eine Verfeinerung von* Z*, d.h.* $Z \subseteq Z^\star$*, dann gilt*

$$S(Z,f,\varphi) \geq S(Z^\star,f,\varphi)$$

und

$$s(Z,f,\varphi) \leq s(Z^\star,f,\varphi).$$

Beweis. Zunächst betrachten wir $Z^\star = Z \cup \{x^\star\}$ mit $x_{i-1} < x^\star < x_i$ für ein gewisses i. Dann ist

$$\begin{aligned} \sup_{x_{i-1}\leq t\leq x_i} f(t)\Delta\varphi_i &= \sup_{x_{i-1}\leq t\leq x_i} f(t)\big(\varphi(x_i) - \varphi(x^\star) + \varphi(x^\star) - \varphi(x_{i-1})\big) \\ &= M_i\big(\varphi(x_i) - \varphi(x^\star)\big) + M_i\big(\varphi(x^\star) - \varphi(x_{i-1})\big) \\ &\geq \sup_{x_{i-1}\leq t\leq x^\star} f(t)\Delta\varphi_{i-1}^\star + \sup_{x^\star\leq t\leq x_{i-1}} f(t)\Delta\varphi_i^\star, \end{aligned}$$

das heißt

$$S(Z,f,\varphi) = \sum M_i\Delta\varphi_i \geq \sum M_i^\star\Delta\varphi_i^\star = S(Z^\star,f,\varphi).$$

Analog zeigt man

$$s(Z,f,\varphi) = \sum m_i\Delta\varphi_i \leq \sum m_i^\star\Delta\varphi_i^\star = s(Z^\star,f,\varphi).$$

Den allgemeinen Fall erhalten wir hieraus durch vollständige Induktion. □

Proposition 7.3.5. *Es gilt*

$$\underline{\int_a^b} f d\varphi \leq \overline{\int_a^b} f d\varphi.$$

Beweis. Seien Z_1 und Z_2 Zerlegungen, dann betrachten wir eine gemeinsame Verfeinerung $Z^\star = Z_1 \cup Z_2$, d.h. $Z_1, Z_2 \supseteq Z^\star$. Wie wir eben gezeigt haben, gilt dann

$$s(Z_1, f, \varphi) \leq s(Z^\star, f, \varphi) \leq S(Z^\star, f, \varphi) \leq S(Z_2, f, \varphi).$$

Folglich ist

$$s(Z_1, f, \varphi) \leq S(Z_2, f, \varphi).$$

Lässt man nun Z_2 fest und nimmt das Supremum über alle Z_1, so folgt nun

$$\underline{\int_a^b} f d\varphi \leq S(Z_2, f, \varphi).$$

Nimmt man nun das Infimum über alle Z_2, so folgt die Behauptung. □

Satz 7.3.6. *Für eine beschränkte Funktion f gilt $f \in R(\varphi, [a, b])$ genau dann, wenn zu jedem $\varepsilon > 0$ eine Zerlegung Z existiert mit*

$$S(Z, f, \varphi) - s(Z, f, \varphi) < \varepsilon.$$

Beweis. "$\Leftarrow$": Zu beliebigen $\varepsilon > 0$ existiere eine Zerlegung Z mit

$$S(Z, f, \varphi) - s(Z, f, \varphi) < \varepsilon.$$

Nach Proposition 7.3.4 gilt

$$s(Z, f, \varphi) \leq \underline{\int_a^b} f d\varphi \leq \overline{\int_a^b} f d\varphi \leq S(Z, f, \varphi)$$

und folglich

$$0 \leq \overline{\int_a^b} f d\varphi - \underline{\int_a^b} f d\varphi \leq S(Z, f, \varphi) - s(Z, f, \varphi) < \varepsilon.$$

Damit folgt aber sofort

$$0 = \overline{\int_a^b} f d\varphi - \underline{\int_a^b} f d\varphi$$

und somit $f \in R(\varphi, [a, b])$.

"$\Rightarrow$": Es sei $f \in R(\varphi, [a, b])$. Dann existieren zu jedem $\varepsilon > 0$ Zerlegungen Z_1 und Z_2 mit

$$S(Z_1, f, \varphi) - \int_a^b f d\varphi < \frac{\varepsilon}{2}, \qquad \int_a^b f d\varphi - s(Z_2, f, \varphi) < \frac{\varepsilon}{2}.$$

Für eine gemeinsame Verfeinerung $Z^\star$ von Z_1 und Z_2 gilt dann

$$S(Z^\star, f, \varphi) \leq S(Z_1, f, \varphi) < \frac{\varepsilon}{2} + \int_a^b f d\varphi < \varepsilon + s(Z_2, f, \varphi) \leq \varepsilon + s(Z^\star, f, \varphi),$$

also

$$0 < S(Z^\star, f, \varphi) - s(Z^\star, f, \varphi) < \varepsilon.$$

□

Satz 7.3.7. *1. Gilt*

$$S(Z, f, \varphi) - s(Z, f, \varphi) < \varepsilon$$

für ein Zerlegung Z von $[a, b]$, so gilt dies auch für jede Verfeinerung.

2. Sei

$$S(Z, f, \varphi) - s(Z, f, \varphi) < \varepsilon$$

für $Z = \{x_0, \ldots, x_n\}$, dann gilt

$$\sum_{i=1}^n |f(t_i) - f(s_i)| \Delta\varphi_i < \varepsilon$$

für beliebige Zwischenwerte $s_i, t_i \in [x_{i-1}, x_i]$.

3. Sei

$$S(Z, f, \varphi) - s(Z, f, \varphi) < \varepsilon$$

für $Z = \{x_0, \ldots, x_n\}$, dann gilt

$$\left| \sum_{i=1}^n f(t_i) \Delta\varphi_i - \int f d\varphi \right| < \varepsilon.$$

für beliebige Zwischenwerte $t_i \in [x_{i-1}, x_i]$.

Beweis. 1. Die erste Aussage folgt direkt aus Proposition 7.3.4.

2. Für bliebig gewählte $s_i, t_i \in [x_{i-1}, x_i]$ gilt

$$|f(t_i) - f(s_i)| \leq M_i - m_i.$$

Damit folgt die Behauptung gemäß

$$0 \leq \sum_{i=1}^n |f(t_i) - f(s_i)| \Delta\varphi_i \leq \sum_{i=1}^n (M_i - m_i) \Delta\varphi_i = S(Z, f, \varphi) - s(Z, f, \varphi) < \varepsilon.$$

3. Wir erinnern uns zunächst, dass für jede Zerlegung gilt

$$s(Z, f, \varphi) \leq \int f d\varphi \leq S(Z, f, \varphi).$$

Ferner ist offensichtlich, dass

$$s(Z, f, \varphi) \leq \sum_{i=1}^{n} f(t_i)\Delta\varphi_i \leq S(Z, f, \varphi)$$

für beliebige $t_i \in [x_{i-1}, x_i]$. Hieraus folgt

$$\left| \int f d\varphi - \sum_{i=1}^{n} f(t_i)\Delta\varphi_i \right| \leq S(Z, f, \varphi) - s(Z, f, \varphi) < \varepsilon.$$

□

Satz 7.3.8. *Sei $f \in C([a, b])$, dann ist $f \in R(\varphi, [a, b])$.*

Beweis. Sei $\varepsilon > 0$ beliebig und $\eta > 0$ so gewählt, dass

$$\eta < \frac{\varepsilon}{\varphi(b) - \varphi(a)}.$$

Wir bemerken, da $[a, b]$ kompakt ist, ist f gleichmäßig stetig auf $[a, b]$, folglich existiert zu $\eta > 0$ ein $\delta > 0$ mit $|f(x) - f(y)| < \eta$ für alle $x, y \in [a, b]$ mit $|x - y| < \delta$. Wir wählen nun eine Zerlegung derart, dass $\Delta x_i < \delta$ für alle $i = 1, \ldots, n$. Dann gilt

$$M_i - m_i = \sup_{x,y \in [x_{i-1}, x_i]} |f(x) - f(y)| < \eta.$$

Damit erhalten wir

$$S(Z, f, \varphi) - s(Z, f, \varphi) = \sum_{i=1}^{n} (M_i - m_i)\Delta\varphi_i \leq \eta \sum_{i=1}^{n} \Delta\varphi_i = \eta\big(\varphi(b) - \varphi(a)\big) < \varepsilon$$

Nach Satz 7.3.6 gilt daher $f \in R(\varphi, [a, b])$. □

Satz 7.3.9. *Ist φ stetig und f monoton auf $[a, b]$, dann ist $f \in R(\varphi, [a, b])$.*

Beweis. Sei $\varepsilon > 0$ beliebig und Z eine Zerlegung derart, dass

$$\Delta\varphi_i = \varphi(x_i) - \varphi(x_{i-1}) = \frac{\varphi(b) - \varphi(a)}{n}.$$

Eine derartige Zerlegung existiert aufgrund des Zwischenwertsatzes. Ohne Beschränkung der Allgemeinheit sei nun f monoton wachsend angenommen. Es folgt dann für hinreichend großes $n \in \mathbb{N}$ dass

$$\begin{aligned} S(Z,f,\varphi) - s(Z,f,\varphi) &= \sum_{i=1}^{n} (M_i - m_i)\Delta\varphi_i \\ &= \sum_{i=1}^{n} \big(f(x_i) - f(x_{i-1})\big)\frac{\varphi(b)-\varphi(a)}{n} \\ &= \frac{\varphi(b)-\varphi(a)}{n}\big(f(b)-f(a)\big) \\ &< \varepsilon. \end{aligned}$$

□

Satz 7.3.10. *Sei f auf $[a,b]$ eine beschränkte Funktion, die bis auf endlich viele Unstetigkeitsstellen y_k, $k = 1,\dots,m$, stetig ist. Ist φ an den Stellen y_k stetig, dann ist $f \in R(\varphi,[a,b])$.*

Beweis. Aufgrund der Stetigkeit von φ in den Punkten y_k existieren Intervalle $I_k := (a_k, b_k)$, $k = 1,\dots,m$, mit $y_k \in I_k$ und

$$\varphi(b_k) - \varphi(a_k) < \frac{\varepsilon}{m}.$$

Da

$$K := [a,b] \setminus \bigcup_{k=1}^{m} I_k$$

kompakt ist, ist $f : K \to \mathbb{R}$ eine gleichmäßig stetige Funktion auf K. Daher existiert ein $\delta > 0$ mit $|f(x) - f(y)| < \varepsilon$ für alle $x, y \in K$ mit $|x - y| < \delta$.

Wir konstruieren nun eine Zerlegung $Z = \{x_0,\dots,x_n\}$ von $[a,b]$ wie folgt. Es gelte $\{a_k, b_k : k = 1,\dots,m\} \subseteq Z$ und für jedes $k \in \{1,\dots,m\}$ gebe es ein $i \in \{1,\dots,n\}$, so dass $[x_{i-1},x_i] = [a_k,b_k]$, d.h. $x_i \notin \bigcup I_k$ für alle $i = 1,\dots,n$. Ferner gelte $\Delta x_i < \delta$ falls $[x_{i-1},x_i] \subseteq K$. Aufgrund der Beschränktheit von f existiert ein $M > 0$ mit $|f(x)| < M$ für alle $x \in [a,b]$. Daher gilt zunächst

$$M_i - m_i = \sup_{t\in[x_{i-1},x_i]} f(t) - \inf_{t\in[x_{i-1},x_i]} f(t) \le 2M$$

für alle $i \in 1,\dots,n$. Im Falle $[x_{i-1},x_i] \subseteq K$ folgt sogar aufgrund der gleichmäßigen Setigkeit von f dass

$$|f(s) - f(t)| < \varepsilon$$

für alle $s, t \in [x_{i-1},x_i]$ und daher

$$M_i - m_i < \varepsilon.$$

Nach Konstruktion der Zerlegung Z erhalten wir nun die Abschätzung

$$\begin{aligned} S(Z,f,\varphi) - s(Z,f,\varphi) &\leq \sum_{i=1}^{n} (M_i - m_i)\Delta\varphi_i \\ &< \varepsilon \sum_{i=1}^{n} \Delta\varphi_i + \sum_{k=1}^{m} 2M\varepsilon = \varepsilon(\varphi(b) - \varphi(a) + 2mM). \end{aligned}$$

Nach Satz 7.3.6 folgt damit $f \in R(\varphi, [a,b])$. □

Satz 7.3.11. *Sei $f \in R(\varphi, [a,b])$ und $f([a,b]) \subseteq [m,M]$. Desweiteren sei $g : [m,M] \to \mathbb{R}$ stetig auf $[a,b]$. Dann folgt für die Komposition $h : [a,b] \to \mathbb{R}$ definiert durch $h(x) = g\big(f(x)\big)$ ebenfalls $h \in R(\varphi, [a,b])$.*

Beweis. Wir bemerken zunächst, dass $g : [m,M] \to \mathbb{R}$ gleichmäßig stetig ist. Somit existiert zu jedem $\varepsilon > 0$ mit $|g(s) - g(t)| < \varepsilon$ für alle $s,t \in [m,M]$ mit $|s-t| < \delta$. Wir dürfen dabei δ so wählen, dass $\delta < \varepsilon$ gilt.

Da $f \in R(\varphi, [a,b])$, existiert zu diesem $\delta > 0$ eine Zerlegung $Z = \{x_0, \ldots, x_n\}$ von $[a,b]$ mit

$$S(Z,f,\varphi) - s(Z,f,\varphi) < \delta^2.$$

Wir schreiben

$$M_i := \sup_{t \in [x_{i-1}, x_i]} f(t), \qquad m_i := \inf_{t \in [x_{i-1}, x_i]} f(t),$$

und

$$M_i^\star := \sup_{t \in [x_{i-1}, x_i]} h(t), \qquad m_i^\star := \inf_{t \in [x_{i-1}, x_i]} h(t).$$

Wir zerlegen nun $Z = V \cup W$, so dass gilt $x_i \in V$ genau dann, wenn $M_i - m_i < \delta$, und $x_i \in W$ genau dann, wenn $M_i - m_i \geq \delta$. Insbesondere gilt dann $V \cap W = \emptyset$.

Für ein beliebiges $x_i \in Z$ gilt dann entweder

$$M_i^\star - m_i^\star < \varepsilon,$$

nämlich wenn $x_i \in V$, oder aber

$$M_i^\star - m_i^\star \leq 2 \sup_{t \in [a,b]} \big|g\big(f(t)\big)\big| = 2 \sup_{t \in [m,M]} |g(t)| =: 2K,$$

nämlich wenn $x_i \in W$.

Wir können nun abschätzen

$$\begin{aligned} S(Z,f,\varphi) - s(Z,f,\varphi) &= \sum_{i=1}^{n} (M_i^\star - m_i^\star)\Delta\varphi_i \\ &= \sum_{x_i \in V} (M_i^\star - m_i^\star)\Delta\varphi_i + \sum_{x_i \in W} (M_i^\star - m_i^\star)\Delta\varphi_i \\ &< \varepsilon\big(\varphi(b) - \varphi(a)\big) + 2K \sum_{x_i \in W} \Delta\varphi_i. \end{aligned}$$

Nun gilt aber

$$\delta \sum_{x_i \in W} \Delta\varphi_i \leq \sum_{x_i \in W} (M_i - m - i)\Delta\varphi_i \leq \delta^2 < \delta\varepsilon$$

aufgrund der Wahl der Zerlegung Z, und damit haben wir

$$S(Z, f, \varphi) - s(Z, f, \varphi) < \varepsilon\big(\varphi(b) - \varphi(a) + 2K\big).$$

Nach Satz 7.3.6 folgt die Behauptung. □

7.4 Eigenschaften des Riemann-Stieltjes-Integral

Satz 7.4.1.

1. *Seien $f_1, f_2 \in R(\varphi, [a, b])$ und $c \in \mathbb{R}$, so gilt $f_1 + f_2 \in R(\varphi, [a, b])$ und $cf_1 \in R(\varphi, [a, b])$ mit*

$$\int_a^b (f_1 + f_2) d\varphi = \int_a^b f_1 d\varphi + \int f_2 d\varphi,$$
$$\int_a^b c f_1 d\varphi = c \int_a^b f_1 d\varphi.$$

2. *Gilt $f_1(x) \leq f_2(x)$ für alle $x \in [a, b]$, so folgt*

$$\int_a^b f_1 d\varphi \leq \int_a^b f_2 d\varphi.$$

3. *Für $f \in R(\varphi, [a, b])$ und $c \in [a, b]$ gilt*

$$\int_a^c f d\varphi + \int_c^b f d\varphi = \int_a^b f d\varphi.$$

4. *Sei $f \in R(\varphi_1, [a, b])$ und $f \in R(\varphi_2, [a, b])$, dann ist $f \in R(\varphi_1 + \varphi_2, [a, b])$ mit*

$$\int_a^b f d(\varphi_1 + \varphi_2) = \int_a^b f d\varphi_1 + \int_a^b f d\varphi_2.$$

Beweis. Wir wollen uns auf den Beweis des ersten Teils der ersten Aussage beschränken, da die Beweise der verbleibenden Aussagen ähnlich sind.

Wir bemerken zuerst, dass für jede beliege Zerlegung Z gilt

$$\begin{aligned} s(Z, f_1, \varphi) + s(Z, f_2, \varphi) &\leq s(Z, f_1 + f_2, \varphi) \\ &\leq S(Z, f_1 + f_2, \varphi) \leq S(Z, f_1, \varphi) + S(Z, f_2, \varphi) \end{aligned} \tag{7.1}$$

Seien nun $f_1, f_2 \in R(\varphi, [a,b])$ und $\varepsilon > 0$ vorgegeben. Es existieren Zerlegungen Z_1 und Z_2 mit

$$S(Z_1, f, \varphi) - s(Z_1, f, \varphi) < \varepsilon/2, \qquad S(Z_2, f, \varphi) - s(Z_2, f, \varphi) < \varepsilon/2.$$

Dies gilt auch für jede gemeinsame Verfeinerung Z von Z_1 und Z_2

$$S(Z, f_1, \varphi) - s(Z, f_1, \varphi) < \varepsilon/2, \qquad S(Z, f_2, \varphi) - s(Z, f_2, \varphi) < \varepsilon/2.$$

Aufgrund von (7.1) folgt damit

$$\begin{aligned} &S(Z, f_1 + f_2, \varphi) - s(Z, f_1 + f_2, \varphi) \\ &\qquad \leq S(Z, f_1, \varphi) + S(Z, f_2, \varphi) - s(Z, f_1, \varphi) - s(Z, f_2, \varphi) < \varepsilon. \end{aligned}$$

Gemäß Satz 7.3.6 schließen wir $f_1 + f_2 \in R(\varphi, [a,b])$.

Es verbleibt noch die Gleichung

$$\int_a^b (f_1 + f_2) d\varphi = \int_a^b f_1 d\varphi + \int_a^b f_2 d\varphi$$

zu zeigen. Für die Zerlegung Z gelten die Ungleichungen

$$S(Z, f_1, \varphi) < \int_a^b f_1 d\varphi + \varepsilon, \qquad S(Z, f_2, \varphi) < \int_a^b f_2 d\varphi + \varepsilon,$$

d.h. nach (7.1)

$$\int_a^b (f_1 + f_2) d\varphi \leq S(Z, f_1 + f_2) < \int_a^b f_1 d\varphi + \int_a^b f_2 d\varphi + 2\varepsilon.$$

Da ε beliebig ist folgt hieraus

$$\int_a^b (f_1 + f_2) d\varphi \leq \int_a^b f_1 d\varphi + \int_a^b f_2 d\varphi.$$

Ersetzt man in dieser Ungleichung f_1 durch $-f_1$ und f_2 durch $-f_2$, so erhalten wir die Umkehrung dieser Ungleichung. Damit haben wir die gewünschte Gleichheit bewiesen. □

Satz 7.4.2. *Für $f, g \in R(\varphi, [a,b])$ folgt*

1. $f \cdot g \in R(\varphi, [a,b])$.
2. $|f| \in R(\varphi, [a,b])$ *und* $\left| \int_a^b f d\varphi \right| \leq \int_a^b |f| d\varphi$.

Beweis. 1. Da $y \mapsto y^2$ stetig auf $\mathbb{R}$ ist und $f \in R(\varphi, [a,b])$, gilt nach Satz 7.3.11, dass $f^2 \in R(\varphi, [a,b])$. Folglich ist auch $(f+g)^2, (f-g)^2 \in R(\varphi, [a,b])$ und damit auch $(f+g)^2 + (f-g)^2 \in R(\varphi, [a,b])$. Aus der Parallelogrammidentität

$$4fg = (f+g)^2 - (f-g)^2$$

folgt dann $f \cdot g \in R(\varphi, [a,b])$.

2. Da $y \mapsto |y|$ auf $\mathbb{R}$ stetig ist, folgt aus Satz Satz 7.3.11, dass $|f| \in R(\varphi, [a,b])$. Sei $c = \pm 1$ derart, dass $c \int_a^b f d\varphi \geq 0$, dann ist

$$\left| \int_a^b f d\varphi \right| = c \int_a^b f d\varphi \int_a^b c f d\varphi \leq \int_a^b |f| d\varphi$$

wegen $cf(x) \leq |f(x)|$ und Satz 7.4.1.

□

Unter gewissen Bedingungen lässt sich das allgemeine Riemann-Stieltjes-Integral auf ein einfaches Riemann-Integral zurückführen.

Satz 7.4.3. *Sei $\varphi : [a,b] \to \mathbb{R}$ monoton wachsend mit $\varphi' \in R([a,b])$. Sei ferner f eine beschränkte Funktion auf $[a,b]$. Dann ist $f \in R(\varphi, [a,b])$ genau dann, falls $f \cdot \varphi' \in R([a,b])$ und es gilt*

$$\int_a^b f d\varphi = \int_a^b f \varphi' dx.$$

Beweis. Da $\varphi' \in R([a,b])$, gibt es zu jedem $\varepsilon > 0$ eine Zerlegung $Z = \{x_0, \ldots, x_n\}$ von $[a,b]$ mit $S(Z, \varphi') - s(Z, \varphi') < \varepsilon$. Aufgrund des Mittelwertsatzes existieren Zwischenstellen $t_i \in (x_{i-1}, x_i)$ mit

$$\varphi'(t_i) = \frac{\varphi(x_i) - \varphi(x_{i-1})}{x_i - x_{i-1}} = \frac{\Delta\varphi_i}{\Delta x_i}.$$

Für beliebiges $s_i \in [x_{i-1}, x_i]$ folgt nun

$$\sum_{i=1}^n |\varphi'(s_i) - \varphi'(t_i)| \Delta x_i < S(Z, \varphi') - s(Z, \varphi') < \varepsilon.$$

Setzen wir $M := \sup_{x \in [a,b]} |f(x)|$ und beachten

$$\sum_{i=1}^n f(s_i) \Delta\varphi_i = \sum_{i=1}^n f(s_i) \varphi'(t_i) \Delta x_i,$$

ergibt sich hieraus

$$\left| \sum_{i=1}^n f(s_i) \Delta\varphi_i - \sum_{i=1}^n f(s_i) \varphi'(s_i) \Delta x_i \right| \leq \sum_{i=1}^n |f(s_i)| |\varphi'(s_i) - \varphi'(t_i)| \Delta x_i < M\varepsilon.$$

Da die s_i beliebig gewählt werden können, folgt hieraus

$$S(Z, f, \varphi) < \sum_{i=1}^n f(s_i) \varphi'(s_i) \Delta x_i + M\varepsilon \leq S(Z, f \cdot \varphi') + M\varepsilon,$$

$$S(Z, f \cdot \varphi') < \sum_{i=1}^n f(s_i) \Delta\varphi_i + M\varepsilon \leq S(Z, f, \varphi) + M\varepsilon,$$

das heisst

$$|S(Z, f, \varphi') - S(Z, f, \varphi)| < M\varepsilon.$$

Alle Schritte bleiben bei jeder Verfeinerung von Z wahr, so dass folgt

$$\left| \overline{\int_a^b} f d\varphi - \overline{\int_a^b} f\varphi' dx \right| < M\varepsilon$$

und somit

$$\overline{\int_a^b} f d\varphi = \overline{\int_a^b} f\varphi' dx.$$

Analog zeigt man, dass

$$\underline{\int_a^b} f d\varphi = \underline{\int_a^b} f\varphi' dx,$$

woraus nun die Behauptung folgt. □

Definition 7.4.4. *Die sogenannte* **Heavysidefunktion** $H : \mathbb{R} \to \mathbb{R}$ *ist definiert durch*

$$H(x) = \begin{cases} 0, & x \leq 0, \\ 1, & x > 0, \end{cases}.$$

Lemma 7.4.5. *Sei $f : [a, b] \to \mathbb{R}$ beschränkt und stetig in $s \in [a, b]$, dann gilt mit $\varphi(x) : H(x-s)$ dass*

$$\int_a^b f d\varphi = \int_a^b f dH(\cdot - s) = f(s).$$

Beweis. Betrachte die Zerlegung $Z = \{a = x_0, x_1, x_2, x_3 = b\}$ mit $x_1 = s$. Es gilt

$$S(Z, f, \varphi) = \sum_{i=1}^{3} \sup_{t \in [x_{i-1}, x_i]} f(t)\Delta H_i = \sup_{t \in [x_1, x_2]} f(t) \cdot 1 =: M_2$$

und

$$s(Z, f, \varphi) = \sum_{i=1}^{3} \inf_{t \in [x_{i-1}, x_i]} f(t)\Delta H_i = \inf_{t \in [x_1, x_2]} f(t) \cdot 1 =: m_2.$$

Da f an der Stelle $x_1 = s$ stetig ist, gilt $M_2 \to f(s)$ und $m_2 \to f(s)$ für $x_2 \to x_1$, dies bedeutet

$$\underline{\int_a^b} f d\varphi = \overline{\int_a^b} f d\varphi = f(s),$$

also die Behauptung. □

Satz 7.4.6. *Seien $c_k \geq 0$ für alle $k \in \mathbb{N}$ mit*

$$\sum_{k=1}^{\infty} c_k < \infty,$$

$\{s_k\}$ eine Folge verschiedener Punkte auf $[a, b]$ und $f : [a, b] \to \mathbb{R}$ stetig auf $[a, b]$. Dann gilt für die monoton wachsende Stufenfunktion

$$\varphi(x) = \sum_{k=1}^{\infty} c_k H(x - s_k)$$

dass

$$\int_a^b f d\varphi = \sum_{k=1}^{\infty} f(s_k) c_k.$$

Beweis. Zuerst bemerken wir, dass die Reihe $\sum_{k=1}^{\infty} c_k H(x - s_k)$ infolge des Majorantenkriteriums für jedes $x \in [a, b]$ absolut konvergent ist. Die Funktion $\varphi(x)$ ist offensichtlich monoton und es gilt $\varphi(a) = 0$ und $\varphi(b) = \sum_{k=1}^{\infty} c_k$.

Zu $\varepsilon > 0$ beliebig, existiert ein $N \in \mathbb{N}$ mit der Eigenschaft

$$\sum_{k=N+1}^{\infty} c_k < \varepsilon.$$

Wir definieren

$$\varphi_1 := \sum_{k=1}^{N} c_k H(x - s_k), \qquad \varphi_2 := \sum_{k=N+1}^{\infty} c_k H(x - s_k).$$

Dann gilt nach Lemma 7.4.5

$$\int_a^b f d\varphi_1 = \sum_{k=1}^{N} c_k f(s_k).$$

Da aus der Stetigkeit von der Funktion f deren Beschränktheit folgt

$$M := \sup_{x \in [a,b]} |f(x)| < \infty,$$

erhalten wir

$$\left| \int_a^b f d\varphi_2 \right| \leq M\varepsilon.$$

Wegen $\varphi = \varphi_1 + \varphi_2$ ist, folgt aus Satz 7.4.1 dass

$$\left| \int_a^b f d\varphi - \sum_{k=1}^{N} c_k f(s_k) \right| \leq M\varepsilon.$$

Für $N \to \infty$ bzw. $\varepsilon \to 0$ ergibt dies die Aussage. □

Satz 7.4.7 (Substitutionsregel). *Sei $\gamma : [A, B] \to [a, b]$ streng monoton wachsend und stetig. Ferner sei $\varphi : [a, b] \to \mathbb{R}$ monoton wachsend und $f \in R(\varphi, [a, b])$. Setzen wir*

$$\psi(y) := \varphi\big(\gamma(y)\big), \qquad g(y) := f\big(\gamma(y)\big),$$

dann ist $g \in R(\psi, [a, b])$ und es gilt

$$\int_A^B g d\psi = \int_a^b f d\varphi.$$

Beweis. Wir bemerken, dass aufgrund der strengen Monotonie die Funktion $\gamma : [A, B] \to [a, b]$ bijektiv ist. Somit gehört zu jeder Zerlegung $Z = \{x_0, \ldots, x_n\}$ von $[a, b]$ eine Zerlegung $W = \{y_0, \ldots, y_n\}$ von $[A, B]$ gegeben durch $y_i = \gamma^{-1}(x_i)$. Ferner gibt es zu jedem $y \in [y_{i-1}, y_i]$ ein $x \in [x_{i-1}, x_i]$ mit $g(y) = f(x)$. Damit gilt $S(W, g, \psi) = S(Z, f, \varphi)$ und $s(W, f, \psi) = s(Z, f, \varphi)$.

Wegen $f \in R(\varphi, [a, b])$ existiert zu jedem $\varepsilon > 0$ eine Zerlegung Z mit

$$S(Z, f, \varphi) \leq \int_a^b f d\varphi + \varepsilon, \qquad s(Z, f, \varphi) \geq \int_a^b f d\varphi - \varepsilon.$$

Daraus folgt aber

$$S(W, g, \psi) \leq \int_A^B g d\psi + \varepsilon, \qquad s(W, g, \psi) \leq \int_A^B g d\psi - \varepsilon,$$

d.h. die Behauptung. □

Bemerkung 7.4.8. *Für die Wahl $\varphi \equiv x$ ist $\psi \equiv \gamma$. Ist nun $\gamma' \in R([A, B])$, so folgt die "bekannte" Substitutionsregel*

$$\int_a^b f(x)dx = \int_A^B f\big(\gamma(y)\big)\gamma'(y)dy.$$

7.5 Integration und Differentiation

Satz 7.5.1. *Sei $f \in R([a, b])$, dann ist die Funktion $x \mapsto F(x) := \int_a^x f(t)dt$ stetig auf $[a, b]$. Ist f an der Stelle $x_0 \in [a, b]$ stetig, so ist F dort differenzierbar mit $F'(x_0) = f(x_0)$.*

Beweis. Wegen $f \in R([a, b])$ ist $M := \sup_{t\in[a,b]} |f(t)| < \infty$. Somit folgt für beliebige $a \leq x \leq y \leq b$ dass

$$|F(x) - F(y)| = \int_x^y f(t)dt \leq M(y - x).$$

Sei $\varepsilon > 0$ beliebig aber fest. Wir wählen $0 < \delta \leq \varepsilon/M$ und erhalten für alle $|x - y| < \delta$, dass $|F(x) - F(y)| < \varepsilon$ ist, d.h. F ist auf $[a, b]$ (gleichmäßig) stetig.

Falls f an der Stelle x_0 stetig ist, wählen wir zu $\varepsilon > 0$ ein $\delta > 0$ derart, dass $|f(x) - f(x_0)| < \varepsilon$ für alle $|x - y| < \delta$. Sei nun $x_0 - \delta < s \leq x_0 \leq t < x_0 + \delta$, dann gilt

$$\left| \frac{F(t) - F(s)}{t - s} - f(x_0) \right| = \left| \frac{1}{t - s} \int_s^t f(u) - f(x_0) du \right| < \varepsilon,$$

also insbesondere für $s = x_0$

$$\left| \frac{F(t) - F(x_0)}{t - x_0} - f(x_0) \right| < \varepsilon.$$

Daraus folgt aber sofort

$$F'(x_0) = \lim_{t \to x_0} \frac{F(t) - F(x_0)}{t - x_0} = f(x_0).$$

□

Satz 7.5.2 (Hauptsatz der Differential- und Integralrechnung). *Sei $f \in R([a, b])$ und $F(x) = \int f(x)dx + C$ eine Stammfunktion von f, d.h. $F'(x) = f(x)$ für alle $x \in [a, b]$. Dann gilt*

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

Beweis. Sei $\varepsilon > 0$ beliebig und Z eine Zerlegung mit

$$S(Z, f) - s(Z, f) < \varepsilon.$$

Nach dem Mittewertsatz der Differentialrechnung existieren für alle $i = 1, \ldots, n$ Zwischenstellen $t_i \in [x_{i-1}, x_i]$ mit

$$F(x_i) - F(x_{i-1}) = F'(t_i)\Delta x_i = f(t_i)\Delta x_i.$$

Daraus folgt

$$\sum_{i=1}^n f(t_i)\Delta x_i = \sum_{i=1}^n \big(F(x_i) - F(x_{i-1})\big) = F(b) - F(a).$$

Aus Satz 7.3.6 folgt daher

$$\left| F(b) - F(a) - \int_a^b f(x)dx \right| < \varepsilon.$$

□

Aus dem Mittelwertsatzes der Differentialrechnung folgt mit Hilfe des Hauptsatzes der Differential- und Integralrechnung sofort der Mittelwertsatzes der Integralrechnung.

Korollar 7.5.3 (Mittelwertsatz der Integralrechnung). *Sei $f \in R([a,b])$ stetig, so gibt es ein $\xi \in (a,b)$ mit*

$$\int_a^b f(x)dx = (b-a)f(\xi).$$

Wir können nun die folgende Variante des Satzes von Taylor beweisen.

Satz 7.5.4 (Lagrange-Darstellung des Restgliedes). *Sei $f : I = [a,b] \to \mathbb{R}$ und $f \in C^{m+1}([a,b])$, dann gilt beliebiges für $x_0 \in (a,b)$*

$$f(x) = T_m(x,x_0) + R_m(x,x_0)$$

mit dem Taylorpolynom

$$T_m(x,x_0) = \sum_{k=0}^{m} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k$$

und dem Lagrangeschen Restglied

$$R_m(x,x_0) = \int_{x_0}^{x} \frac{f^{(m+1)}(t)}{m!}(x-t)^m dt.$$

Beweis. Wir führen den Beweis mittels vollständiger Induktion.

Für $m = 0$ erhalten wir die Aussage

$$f(x) = f(x_0) + \int_{x_0}^{x} f'(t)dt,$$

die äquivalent ist zum Hauptsatz der Differential- und Integralrechnung.

Wir nehmen nun an, dass die Behauptung für $m \in \mathbb{N}$ gilt, und zeigen, dass sie dann auch für $m+1$ folgt. Es gelte also

$$f(x) = T_m(x,x_0) + R_m(x,x_0).$$

Definieren wir

$$F(t) := \frac{f^{(m+1)}(t)}{(m+1)!}(x-t)^{m+1}$$

dann folgt

$$F'(t) = \frac{dF}{dt}(t) = \frac{f^{(m+2)}(t)}{(m+1)!}(x-t)^{m+1} - \frac{f^{(m+1)}(t)}{m!}(x-t)^m$$

und daher ist

$$\begin{aligned} F(x) - F(x_0) &= \int_{x_0}^{x} F'(t)dt \\ &= \int_{x_0}^{x} \frac{f^{(m+2)}(t)}{(m+1)!}(x-t)^{m+1}dt - \int_{x_0}^{x} \frac{f^{(m+1)}(t)}{m!}(x-t)^m dt \\ &= R_{m+1}(x,x_0) - R_m(x,x_0). \end{aligned}$$

Es ergibt sich demnach die Beziehung

$$R_m(x, x_0) = R_{m+1}(x, x_0) + F(x_0).$$

Diese eingesetzt in die Induktionsvoraussetzung liefert

$$\begin{aligned} f(x) &= T_m(x, x_0) + \frac{f^{(m+1)}(x_0)}{(m+1)!}(x - x_0)^{m+1} + R_{m+1}(x, x_0) \\ &= T_{m+1}(x, x_0) + R_{m+1}(x, x_0). \end{aligned}$$

□

Korollar 7.5.5 (Partielle Integration). *Seien F und G auf $[a, b]$ differenzierbare Funktionen mit $F' = f$ und $G' = g$ und $f, g \in R([a, b])$. Dann gilt*

$$\int_a^b F(x)g(x)dx = F(b)G(b) - F(a)G(a) - \int_a^b f(x)G(x)dx.$$

Beweis. Setzen wir $H(x) = F(x)G(x)$, dann folgt die Behauptung direkt aus dem Hauptsatz der Differential- und Integralrechnung

$$\begin{aligned} F(b)G(b) - F(a)G(a) = H(b) - H(a) &= \int_a^b H'(x)dx \\ &= \int_a^b F(x)g(x)dx + \int_a^b f(x)G(x)dx. \end{aligned}$$

□

Definition 7.5.6. *Seien $f_1, f_2, \ldots, f_n : [a, b] \to \mathbb{R}$ und $\mathbf{f} = (f_1, \ldots, f_n) : [a, b] \to \mathbb{R}^n$ sowie $\varphi : [a, b] \to \mathbb{R}$ monoton wachsend. Wir schreiben $\mathbf{f} \in R(\varphi, [a, b])$, falls alle $f_1, \ldots, f_n \in R(\varphi, [a, b])$. In diesem Fall gilt*

$$\int_a^b \mathbf{f} d\varphi = \left(\int_a^b f_1 d\varphi, \ldots, \int_a^b f_n d\varphi \right).$$

Aus dem Eindimensionalen erhalten wir unmittelbar die folgende Proposition.

Proposition 7.5.7. *Seien die Funktionen $\mathbf{f}, \mathbf{F} : [a, b] \to \mathbb{R}^n$ und $\mathbf{f} \in R([a, b])$ mit $\mathbf{F}' = \mathbf{f}$, dann gilt*

$$\int_a^b \mathbf{f}(x)dx = \mathbf{F}(b) - \mathbf{F}(a).$$

7.6 Rektifizierbare Kurven

Definition 7.6.1. *Eine stetige Funktion $\gamma : [a, b] \to \mathbb{R}^k$ heißt* **Kurve** *auf dem Parameterintervall $[a, b]$. Falls γ injektiv ist, heißt die Kurve ein* **Bogen**. *Falls $\gamma(a) = \gamma(b)$ gilt, dann heißt die Kurve* **geschlossen**.

Bemerkung 7.6.2. *Wir haben hier absichtlich Kurven als Abbildungen und nicht als Punktmengen definiert. Zwei Kurven γ_1 und γ_2 können verschieden sein, jedoch den selben Bildbereich ($\in \mathbb{R}^k$) besitzen.*

Definition 7.6.3. *Sei $Z = \{x_0, x_1, \ldots, x_n\}$ eine Zerlegung von $[a, b]$, dann definieren wir $\Lambda(Z, \gamma) := \sum_{i=1}^n \|\gamma(x_i) - \gamma(x_{i-1})\|$, d.h. $\Lambda(Z, \gamma)$ ist die Länge des Polygonzuges durch die Punkte $\gamma(x_i)$. Wir nennen die Kurve rektifizierbar, falls gilt*

$$\Lambda(\gamma) = \sup_Z \Lambda(Z, \gamma) < \infty$$

Die Größe $\Lambda(\gamma)$ heißt **Bogenlänge** *der Kurve γ.*

Satz 7.6.4. *Der Tangentenvektor an die Kurve γ gegeben durch $\gamma' : [a, b] \to \mathbb{R}^n$ sei stetig auf $[a, b]$, kurz $\gamma' \in C^1([a, b]; \mathbb{R}^n)$. Dann ist γ rektifizierbar und es gilt*

$$\Lambda(\gamma) = \int_a^b \|\gamma'(t)\| dt.$$

Beweis. Wir zeigen zunächst, dass gilt

$$\Lambda(\gamma) \leq \int_a^b \|\gamma'(t)\| dt.$$

Hierzu bezeichne $Z = \{t_0, \ldots, t_n\}$ eine beliebige Zerlegung von $[a, b]$ mit

$$a = t_0 < t_1 < \ldots < t_{i-1} < t_i < \ldots < t_n = b.$$

Es gilt

$$\|\gamma(t_i) - \gamma(t_{i-1})\| = \left\| \int_{t_{i-1}}^{t_i} \gamma'(t) dt \right\| \leq \int_{t_{i-1}}^{t_i} \|\gamma'(t)\| dt,$$

und daher $\Lambda(Z, \gamma) \leq \int_a^b \|\gamma'(t)\| dt$ durch aufaddieren über alle Teilintervalle. Damit folgt aber

$$\sup_Z \Lambda(Z, \gamma) = \Lambda(\gamma) \leq \int_a^b \|\gamma'(t)\| dt.$$

Wir zeigen nun, dass auch die Ungleichung

$$\Lambda(\gamma) \geq \int_a^b \|\gamma'(t)\| dt$$

gilt. Wir bemerken zunächst, dass γ' gleichmäßig stetig ist auf $[a, b]$. Demnach existiert zu jedem $\varepsilon > 0$ ein $\delta > 0$ mit

$$\|\gamma'(t) - \gamma'(s)\| < \varepsilon$$

für alle $s, t \in [a, b]$ mit

$$|s - t| < \delta.$$

Wählen wir nun eine Zerlegung $Z = \{t_0, \ldots, t_n\}$ mit der Eigenschaft $\Delta_i < \delta$, dann gilt

$$\|\gamma'(t)\| \leq \|\gamma'(t_i)\| + \varepsilon$$

für alle $t \in [t_{i-1}, t_i]$. Daraus folgt

$$\begin{aligned}
\int_{t_{i-1}}^{t_i} \|\gamma'(t)\| dt &\leq \int_{t_{i-1}}^{t_i} \big(\|\gamma'(t_i)\| + \varepsilon\big) dt \\
&= \left\| \int_{t_{i-1}}^{t_i} \gamma'(t_i) dt \right\| + \varepsilon \Delta t_i \\
&= \left\| \int_{t_{i-1}}^{t_i} \big(\gamma'(t) + \gamma'(t_i) - \gamma'(t)\big) dt \right\| + \varepsilon \Delta t_i \\
&\leq \left\| \int_{t_{i-1}}^{t_i} \gamma'(t) dt \right\| + \left\| \int_{t_{i-1}}^{t_i} \big(\gamma'(t_i) - \gamma'(t)\big) dt \right\| + \varepsilon \Delta t_i \\
&\leq \left\| \int_{t_{i-1}}^{t_i} \gamma'(t) dt \right\| + \int_{t_{i-1}}^{t_i} \|\gamma'(t_i) - \gamma'(t)\| dt + \varepsilon \Delta t_i \\
&\leq \|\gamma(t_i) - \gamma(t_{i-1})\| + 2\varepsilon \Delta t_i.
\end{aligned}$$

Summieren wir nun wider über $i = 1, \ldots, n$, d.h. über alle Teilintervalle, erhalten wir

$$\begin{aligned}
\int_a^b \|\gamma'(t)\| dt &\leq \sum_{i=1}^n \|\gamma(t_i) - \gamma(t_{i-1})\| + 2\varepsilon(b - a) = \Lambda(Z, \gamma) + 2\varepsilon(b - a) \\
&\leq \Lambda(\gamma) + 2\varepsilon(b - a)
\end{aligned}$$

für beliebiges $\varepsilon > 0$. Daraus folgt

$$\int_a^b \|\gamma'(t)\| dt \leq \Lambda(\gamma).$$

□

Beispiel 7.6.5. *Der Einheitskreis in $\mathbb{R}^2$ kann als Kurve vermittels der Abbildung*

$$\gamma := \begin{cases} [0, 2\pi] \to \mathbb{R}^2, \\ t \mapsto (\cos t, \sin t)^T, \end{cases}$$

beschrieben werden. Aus $\gamma'(t) = (-\sin t, \cos t)^T$ *folgt*

$$\|\gamma'(t)\| = \sqrt{\sin^2 t + \cos^2 t} = 1$$

und daher

$$\Lambda(\gamma) = \int_0^{2\pi} \|\gamma'(t)\| dt = \int_0^{2\pi} dt = 2\pi.$$

7.7 Uneigentliche Integrale

Das Riemann-Stieljes-Integral $\int_a^b f(x)d\varphi(x)$ wurde entwickelt unter der doppelten Voraussetzung, dass sowohl die zu integrierende Funktion f als auch das Integrationsintervall $[a, b]$ beschränkt sind. Lässt man eine dieser beiden Voraussetzung fallen, verliert das Integral zunächst seinen Sinn. Es liegt jedoch auf der Hand, auch für unendliche Intervalle das Integral durch einen entsprechenden Grenzübergang zu definieren, etwa für beschränktes f, aber einem unbeschränkten Intervall

$$\int_a^\infty f(x)d\varphi(x) = \lim_{y\to\infty} \int_a^y f(x)d\varphi(x).$$

Entsprechend wird verfahren, wenn das Intervall $I = [a, b]$ zwar beschränkt ist, die Funktion f jedoch unbeschränkt ist auf $[a, b]$. Nehmen wir an, dass f im Punkt $c \in [a, b]$ unbeschränkt ist, dann erhalten wir

$$\begin{aligned}\int_a^b f(x)d\varphi(x) &= \int_a^c f(x)d\varphi(x) + \int_c^b f(x)d\varphi(x)\\ &= \lim_{\varepsilon\to 0+} \int_a^{c-\varepsilon} f(x)d\varphi(x) + \lim_{\varepsilon\to 0+} \int_{c+\varepsilon}^b f(x)d\varphi(x).\end{aligned}$$

Beispiel 7.7.1. *Wir betrachten das Integral* $\int_0^a x^\alpha dx$ *für beliebiges* $a > 0$ *und* $\alpha \in \mathbb{R}$. *Wir finden*

$$\begin{aligned}\int_0^a x^\alpha dx &= \lim_{\varepsilon\to 0+} \int_\varepsilon^a x^\alpha dx\\ &= \lim_{\varepsilon\to 0+} \begin{cases}(\alpha+1)^{-1}(a^{\alpha+1} - \varepsilon^{\alpha+1}), & \alpha \neq -1\\ \ln a - \ln \varepsilon, & \alpha = -1,\end{cases}\\ &= \begin{cases}\frac{a^{\alpha+1}}{\alpha+1}, & \alpha > -1,\\ \infty, & \alpha \leq -1.\end{cases}\end{aligned}$$

Hingegen kehrt sich die Situation für das Integral $\int_a^\infty x^\alpha dx$ *um*

$$\begin{aligned}
\int_a^\infty x^\alpha dx &= \lim_{y\to\infty} \int_a^y x^\alpha dx \\
&= \lim_{y\to\infty} \begin{cases} (\alpha+1)^{-1}(y^{\alpha+1} - a^{\alpha+1}), & \alpha \neq -1 \\ \ln y - \ln a, & \alpha = -1, \end{cases} \\
&= \begin{cases} -\frac{a^{\alpha+1}}{\alpha+1}, & \alpha < -1, \\ \infty, & , \alpha \geq -1. \end{cases}
\end{aligned}$$

Kapitel 8

Folgen und Reihen von Funktionen

8.1 Gleichmäßige Konvergenz

Definition 8.1.1. *Sei $(f_n)_{n\in\mathbb{N}}$ eine Folge von Funktionen, die auf der Menge D definiert sind. Ferner konvergiere die Zahlenfolge $\big(f_n(x)\big)_{n\in\mathbb{N}}$ für jedes $x \in D$. Dann definiert*

$$f(x) = \lim_{n\to\infty} f_n(x)$$

ebenfalls eine Funktion auf D. Wir sagen, die Funktionsfolge $(f_n)_{n\in\mathbb{N}}$ konvergiert **punktweise** *gegen die Grenzwerte von f.*

Konvergiert die Reihe $\sum_{n=1}^{\infty} f_n(x)$ für jedes $x \in D$, so definieren wir analog

$$F(x) = \sum_{n=1}^{\infty} f_n(x)$$

und nennen die Funktion F **die Summe** *der Reihe $\sum_{n=1}^{\infty} f_n$.*

Die Frage ist nun, wieweit sich Eigenschaften wie Stetigkeit, Differenzierbarkeit und Integrierbarkeit der einzelnen Funktionen auf die Grenzfunktion übertragen lassen. Dass man hierbei vorsichtig sein muß, zeigen folgende Beispiele:

Beispiel 8.1.2. *1. Sei*

$$s_{m,n} := \frac{m}{m+n}, \qquad m, n \in \mathbb{N}.$$

Für alle $m \in \mathbb{N}$ gilt

$$\lim_{n\to\infty} s_{m,n} = 0,$$

dass heißt,

$$\lim_{m\to\infty} \left(\lim_{n\to\infty} s_{m,n} \right) = 0.$$

Aber für alle $n \in \mathbb{N}$ gilt

$$\lim_{m\to\infty} \frac{m}{m+n} = 1,$$

dass heißt,

$$\lim_{n\to\infty} \left(\lim_{m\to\infty} s_{m,n} \right) = 1.$$

Die Reihenfolge von Grenzprozessen dürfen wir im allgemeinen also nicht vertauschen.

2. *Wir betrachten für $n \in \mathbb{N}_0$ die Funktionen*

$$f_n(x) := \frac{x^2}{(1+x^2)^n}, \qquad x \in \mathbb{R}.$$

und betrachten

$$F(x) = \sum_{n=0}^{\infty} \frac{x^2}{(1+x^2)^n}.$$

Für $x \neq 0$ erhalten wir

$$F(x) = \sum_{n=0}^{\infty} \frac{x^2}{(1+x^2)^n} = x^2 \frac{1}{1-\frac{1}{1-x^2}} = 1 + x^2.$$

Da jedoch $f_n(0) = 0$ für alle $n \in \mathbb{N}_0$ gilt

$$F(x) = \begin{cases} 0, & x = 0, \\ 1+x^2, & x \neq 0, \end{cases}$$

dass heißt F ist nicht stetig in $x = 0$.

3. *Für $m \in \mathbb{N}_0$ sei*

$$f_m(x) := \lim_{n\to\infty} (\cos m! x\pi)^{2n}.$$

Ist x irrational, so folgt $m!x \notin \mathbb{Z}$ und somit $|\cos(m!x\pi)| < 1$, also $f_m(x) = 0$. Dagegen gilt für rationales x, sagen wir $x = \frac{p}{q}$ mit $p \in \mathbb{Z}$ und $q \in \mathbb{N}$, dass $f_m(x) = 1$ für alle $m \geq q$ da $m!\frac{p}{q} \in \mathbb{Z}$. Dies bedeutet

$$\lim_{m\to\infty} f_m(x) = \begin{cases} 1, & x \in \mathbb{Q}, \\ 0, & x \notin \mathbb{Q}. \end{cases}$$

Es sei angemerkt, dass diese Funktion nicht Riemann-integrierbar ist.

Definition 8.1.3. *Sei $(f_n)_{n\in\mathbb{N}}$ eine Folge von Funktionen $f_n : D \to \mathbb{R}$ bzw. $f_n : D \to \mathbb{C}$. Auf D heißt diese Folge* **gleichmäßig konvergent** *gegen $f : D \to \mathbb{R}$ bzw. $f : D \to \mathbb{C}$, falls zu jedem $\varepsilon > 0$ ein $N = N(\varepsilon) \in \mathbb{N}$ existiert, so dass gilt*

$$|f_n(x) - f(x)| < \varepsilon$$

für alle $n > N$ und alle Stellen $x \in D$.

Proposition 8.1.4. *Die Funktionenfolge $(f_n)_{n\in\mathbb{N}}$ konvergiert genau dann gleichmäßig gegen f, falls die Zahlenfolge*

$$M_n = \sup_{x\in D} |f_n(x) - f(x)|, \qquad n \in \mathbb{N},$$

eine Nullfolge ist, also gilt

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

Satz 8.1.5 (Cauchy-Kriterium für gleichmäßige Konvergenz). *Sei $(f_n)_{n\in\mathbb{N}}$ eine Folge von Funktionen $f_n : D \to \mathbb{R}$ bzw. $f_n : D \to \mathbb{C}$. Die Folge $(f_n)_{n\in\mathbb{N}}$ konvergiert genau dann gleichmäßig gegen $f : D \to \mathbb{R}$ bzw. $f : D \to \mathbb{C}$, falls zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$ existiert, so dass gilt*

$$|f_n(x) - f_m(x)| < \varepsilon$$

für alle $n, m > N$ und für alle $x \in D$.

Beweis. " $\Rightarrow$": Wegen $f_n \to f$ gleichmäßig, existiert zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$, so dass für alle $x \in D$ und für alle $n, m > N$ gilt

$$|f(x) - f_n(x)| < \frac{\varepsilon}{2}, \qquad |f(x) - f_m(x)| < \frac{\varepsilon}{2}.$$

Somit gilt aber auch

$$|f_m(x) - f_n(x)| \leq |f_m(x) - f(x)| + |f(x) - f_n(x)| < \varepsilon$$

für alle $x \in D$ und $n, m > N$.

" $\Leftarrow$": Für beliebiges $x \in D$ folgt aus der Cauchy-Bedingung, dass die reelle bzw. komplexe Zahlenfolge $\big(f_n(x)\big)_{n\in\mathbb{N}}$ eine Cauchy-Folge ist. Da $\mathbb{R}$ bzw. $\mathbb{C}$ ein vollständiger metrischer Raum ist, existiert ein $f(x) \in \mathbb{R}$ bzw. $f(x) \in \mathbb{C}$ mit $f(x) = \lim_{n\to\infty} f_n(x)$. Damit gilt für ein gegebenes $\varepsilon > 0$ und einem geeignetem $N \in \mathbb{N}$ dass

$$|f_n(x) - f_m(x)| < \varepsilon$$

für alle $x \in D$ und $n, m > N$.

Halten wir $n < m$ fest und betrachten den Grenzübergang $m \to \infty$, dann geht $f_m(x)$ in $f(x)$ über und aus der obigen Ungleichung folgt

$$|f_n(x) - f(x)| \leq \varepsilon.$$

□

In vollständiger Analogie zu Zahlenreihen, wird die Definition der gleichmäßigen Konvergenz von Reihen über die gleichmäßige Konvergenz der Partialsummenfolgen definiert.

Definition 8.1.6. *Wir sagen, dass die Funktionenreihe $\sum_{n=1}^{\infty} f_n(x)$ gleichmäßig auf D konvergiert, wenn ihre Partialsummenfolge*

$$s_m(x) = \sum_{n=1}^{m} f_n(x)$$

gleichmäßig auf D konvergiert.

Korollar 8.1.7 (Majorantensatz von Weierstraß). *Sei $M_n = \sup_{x \in D} |f_n(x)|$. Falls die Reihe $\sum_{n=1}^{\infty} M_n$ gegen ein $M \in \mathbb{R}$ konvergiert, so konvergiert die Funktionenreihe $\sum_{n=1}^{\infty} f_n$ gleichmäßig gegen eine Grenzfunktion.*

Beweis. Sei $\varepsilon > 0$ beliebig, dann existiert aufgrund des Cauchy-Kriteriums für Reihen ein $N \in \mathbb{N}$ mit

$$\sum_{k=n+1}^{m} M_k \leq \varepsilon$$

für alle $N < n \leq m$.

Hieraus folgt nun für alle $x \in D$ dass

$$|s_m(x) - s_n(x)| = \left| \sum_{k=n+1}^{m} f_k(x) \right| \leq \sum_{k=n+1}^{m} |f_k(x)| \leq \sum_{k=n+1}^{m} M_k \leq \varepsilon.$$

Die Partialsummenfolge $s_n(x)$ erfüllt also das Cauchy-Kriterium aus Satz 8.1.5. □

Satz 8.1.8.

1. *Sei $(f_n)_{n \in \mathbb{N}}$ eine gegen f gleichmäßig konvergierende Funktionenfolge von auf D stetigen Funktionen, dann ist f ebenfalls stetig.*

2. *Sei x ein Häufungspunkt von D und für alle $n \in \mathbb{N}$ konvergiere $\lim_{t \to x} f_n(t) =: A_n$. Sei $f_n(x)$ gleichmäßig konvergent gegen f und $A_n \to A$ für $n \to \infty$, dann gilt*

$$\lim_{t \to x} f(t) = \lim_{n \to \infty} A_n = A,$$

oder anders ausgedrückt,

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

Beweis. Wir zeigen nur die zweite Aussage, da diese die erste nach sich zieht.

Sei $\varepsilon > 0$ beliebig, dann existiert ein $N \in \mathbb{N}$ derart, dass

$$|f_n(t) - f(t)| < \frac{\varepsilon}{3}$$

und

$$|A_n - A| < \frac{\varepsilon}{3}$$

für alle $t \in D$ und $n > N$. Wegen $\lim_{n\to\infty} f_n(t) = A_n$ existiert eine Umgebung U_x von x, so dass

$$|f_n(t) - A_n| < \frac{\varepsilon}{3}$$

für alle $t \in U_x \cap D$ und $n > N$. Damit ergibt sich schließlich

$$|f(t) - A| \leq |f(t) - f_n(t)| + |f_n(t) - A_n| + |A_n - A| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.$$

□

Definition 8.1.9. *Die Menge aller stetigen und beschränkten Funktionen $f : D \to \mathbb{R}$ bzw. $f : D \to \mathbb{C}$ bezeichnen wir mit $C(D;\mathbb{R})$ bzw. $C(D;\mathbb{C})$, kurz mit $C(D)$. Wir definieren auf dieser Menge die* **Supremumsnorm** *vermittels*

$$\|f\| := \sup_{x\in D} |f(x)|.$$

Satz 8.1.10. *Die Menge $C(D)$ versehen mit der Supremumsnorm $\|\cdot\|$ bildet einen normierten Vektorraum mit der Metrik $d(f,g) := \|f - g\|$.*

Beweis. Seien $f, g \in C(D)$ und $\alpha, \beta \in \mathbb{R}$ bzw. $\alpha, \beta \in \mathbb{C}$, dann ist $\alpha f + \beta g \in C(D)$, d.h. $C(D)$ ist ein linearer Raum.

Da die Funktionen aus $C(D)$ beschränkt sind, gilt

$$0 \leq \|f\| = \sup_{x\in D} |f(x)| < \infty$$

für alle $f \in C(D)$. Aus

$$0 = \|f\| = \sup_{x\in D} |f(x)|$$

folgt sofort, dass $f \equiv 0$ sein muss, also die Nullfunktion ist. Damit ist die Norm definit. Die Norm ist ebenfalls homogen, da für jedes $\alpha \in \mathbb{R}$ bzw. $\alpha \in \mathbb{C}$ gilt

$$\|\alpha f\| = \sup_{x\in D} |\alpha f(x)| = |\alpha| \sup_{x\in D} |f(x)| = |\alpha| \cdot \|f\|.$$

Ferner gilt die Dreiecksungleichung, denn für beliebiges $f, g \in C(D)$ folgt

$$\begin{aligned}
\|f + g\| &= \sup_{x\in D} |f(x) + g(x)| \\
&\leq \sup_{x\in D} \big(|f(x)| + |g(x)|\big) \\
&\leq \sup_{x\in D} |f(x)| + \sup_{x\in D} |g(x)| \\
&= \|f\| + \|g\|.
\end{aligned}$$

Dies zeigt, dass die Supremumsnorm eine Norm auf $C(D)$ definiert. Somit ist $d(f,g) = \|f - g\|$ eine Metrik. □

Satz 8.1.11. *Der Raum $C(D)$ versehen mit der obigen Metrik ist ein vollständiger Raum.*

Beweis. Sei $(f_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge in $C(D)$, d.h. zu jedem $\varepsilon > 0$ existiert ein $N \in \mathbb{N}$, so dass für alle $n, m > N$ gilt

$$d(f_n, f_m) = \|f_n - f_m\| < \varepsilon.$$

Dies bedeutet, dass

$$|f_n(x) - f_m(x)| < \varepsilon$$

für alle $x \in D$. Aus Korollar 8.1.7 folgt, dass $(f_n)_{n\in\mathbb{N}}$ gegen die Grenzwertfunktion f konvergiert, dass heißt

$$\lim_{n\to\infty} \|f_n - f\| = 0.$$

Satz 8.1.8 garantiert, dass f stetig auf D ist. Aus $\|f_n - f\| < 1$ für hinreichend großes n folgt

$$\|f\| \leq 1 + \|f_n\| \leq \infty,$$

d.h. f ist beschränkt und damit gilt für die Grenzfunktion $f \in C(D)$. □

Satz 8.1.12. *Sei $\varphi : [a, b] \to \mathbb{R}$ eine monoton wachsende Funktion und für alle $n \in \mathbb{N}$ gelte $f_n \in \mathcal{R}(\varphi, [a, b])$. Falls $f_n \to f$ gleichmäßig konvergiert, dann ist $f \in \mathcal{R}(\varphi, [a, b])$ und es gilt*

$$\lim_{n\to\infty} \int_a^b f_n d\varphi = \int_a^b f d\varphi.$$

Beweis. Setzen wir

$$\varepsilon_n := \|f_n - f\| = \sup_{x\in[a,b]} |f_n(x) - f(x)|$$

dann gilt

$$f_n(x) - \varepsilon_n \leq f(x) \leq f_n(x) + \varepsilon_n$$

für alle $x \in [a, b]$. Somit folgen die Ungleichungen

$$\int_a^b (f_n - \varepsilon_n) d\varphi \leq \underline{\int_a^b} f d\varphi \leq \overline{\int_a^b} f d\varphi \leq \int_a^b f_n d\varphi + \varepsilon_n\big(\varphi(b) - \varphi(a)\big)$$

und daher

$$\begin{aligned} 0 \leq \overline{\int_a^b} f d\varphi - \underline{\int_a^b} f d\varphi &\leq \int_a^b f_n d\varphi + \varepsilon_n\big(\varphi(b) - \varphi(a)\big) - \int_a^b (f_n - \varepsilon_n) d\varphi \\ &\leq 2\varepsilon_n\big(\varphi(b) - \varphi(a)\big) \overset{n\to\infty}{\longrightarrow} 0, \end{aligned}$$

dass heißt

$$\underline{\int_a^b} f d\varphi = \overline{\int_a^b} f d\varphi = \int_a^b f d\varphi$$

und somit $f \in \mathcal{R}(\varphi, [a,b])$.

Der zweite Teil der Behauptung folgt nun aus

$$\begin{aligned}\left|\int_a^b f d\varphi - \int_a^b f_n d\varphi\right| &= \left|\int_a^b (f-f_n) d\varphi\right| \leq \int_a^b |f-f_n| d\varphi \leq \varepsilon_n \int_a^b d\varphi \\ &= \varepsilon_n\big(\varphi(b)-\varphi(a)\big) \overset{n\to\infty}{\longrightarrow} 0.\end{aligned}$$

□

Dieser Satz überträgt sich natürlich direkt auf gleichmäßig konvergente Funktionenreihen.

Korollar 8.1.13. *Sei $\varphi : [a,b] \to \mathbb{R}$ eine monoton wachsende Funktion und für alle $n \in \mathbb{N}$ gelte $f_n \in \mathcal{R}(\varphi, [a,b])$. Ist $F = \sum_{n=1}^{\infty} f_n$ gleichmäßig konvergent, dann ist $F \in \mathcal{R}(\varphi, [a,b])$ und es gilt*

$$\int_a^b \left(\sum_{n=1}^{\infty} f_n\right) d\varphi = \sum_{n=1}^{\infty} \int_a^b f_n d\varphi.$$

Satz 8.1.14. *Sei $(f_n)_{n\in\mathbb{N}}$ eine Folge von auf $[a,b]$ differenzierbaren Funktionen, so dass $\big(f_n(x_0)\big)_{n\in\mathbb{N}}$ für einen Punkt $x_0 \in [a,b]$ konvergiert. Konvergiert $(f_n')_{n\in\mathbb{N}}$ gleichmäßig auf $[a,b]$, dann konvergiert die Folge $(f_n)_{n\in\mathbb{N}}$ selbst gleichmäßig auf $[a,b]$ gegen eine Funktion f und es gilt*

$$\lim_{n\to\infty} f_n'(x) = f'(x)$$

für alle $x \in [a,b]$.

Beweis. Sei $\varepsilon > 0$, beliebig, dann nehmen wir ein $N \in \mathbb{N}$ derart, dass für alle $n, m > N$ und $t \in [a,b]$ gilt

$$|f_n(x_0) - f_m(x_0)| < \varepsilon \tag{8.1}$$

und

$$|f_n'(t) - f_m'(t)| < \frac{\varepsilon}{2(b-a)}.$$

Wenden nun den Mittelwertsatz auf die Funktion $f_n - f_m$ an, dann gilt für beliebiges $x \in [a,b]$

$$\big|\big(f_n(x) - f_m(x)\big) - \big(f_n(t) - f_m(t)\big)\big| = |(f_n' - f_m')(\xi)||t-x| \leq \frac{\varepsilon|t-x|}{2(b-a)} \leq \frac{\varepsilon}{2}. \tag{8.2}$$

Aus der Ungleichung

$$|f_n(x) - f_m(x)| \leq \big|\big(f_n(x) - f_m(x)\big) - \big(f_n(t) - f_m(t)\big)\big| + |f_n(t) - f_m(t)|$$

folgt nun durch die spezielle Wahl $t = x_0$ mit Hilfe von (8.1) und (8.2) dass

$$|f_n(x) - f_m(x)| < \varepsilon/2 + \varepsilon/2 = \varepsilon$$

für alle $x \in [a, b]$. Damit ist das Cauchy-Kriterium für die gleichmäßige Konvergenz der Funktionenfolge $(f_n)_{n\in\mathbb{N}}$ nachgewiesen und daher konvergiert f_n gleichmäßig gegen eine Grenzfunktion f.

Für festes $x \in [a, b]$ betrachten wir nun die Funktionen

$$\varphi_n(t) = \frac{f_n(t) - f_n(x)}{t - x}, \qquad \varphi(t) = \frac{f(t) - f(x)}{t - x}.$$

Da $f_n(t)$ gegen $f(t)$ konvergiert, folgern wir dass $\varphi_n(t)$ gegen $\varphi(t)$ für $n \to \infty$. Insbesondere ergibt sich aus (8.2), dass

$$|\varphi_n(t) - \varphi_m(t)| = \left| \frac{\big(f_n(x) - f_m(x)\big) - \big(f_n(t) - f_m(t)\big)}{t - x} \right| \leq \frac{\varepsilon}{2(b - a)},$$

das heißt φ_n konvergiert sogar gleichmäßig auf $[a, b] \setminus \{x\}$ gegen φ.

Nun gilt aber auch

$$\lim_{t\to x} \varphi_n(t) = f_n'(t)$$

für alle $n \in \mathbb{N}$. Nach Satz 8.1.8 folgt damit

$$\lim_{n\to\infty} f_n'(x) = \lim_{n\to\infty} \lim_{t\to x} \varphi_n(t) = \lim_{t\to x} \lim_{n\to\infty} \varphi_n(t) = \lim_{t\to x} \varphi(t) = f'(x),$$

was zu zeigen war. □

Satz 8.1.15. *Es existiert eine Funktion $f : \mathbb{R} \to \mathbb{R}$, die stetig auf $\mathbb{R}$ ist, aber an keiner Stelle $x \in \mathbb{R}$ differenzierbar ist.*

Beweis. Wir definieren die 2-periodische Funktion $\varphi : \mathbb{R} \to \mathbb{R}$ durch

$$\varphi(x) := |x|, \qquad -1 \leq x \leq 1,$$

und der Forderung, dass

$$\varphi(x) = \varphi(x + 2).$$

Dann gilt

$$\|\varphi\| = \sup_{x\in\mathbb{R}} |\varphi(x)| = 1$$

und wegen

$$|\varphi(x) - \varphi(t)| \leq |t - x| \tag{8.3}$$

für alle $t, x \in \mathbb{R}$ ist φ stetig auf $\mathbb{R}$.

Wir setzen nun

$$f(x) := \sum_{n=0}^{\infty} \left(\frac{3}{4}\right)^n \varphi(4^n x)$$

Wegen $\|\varphi(4^n x)\| = 1$ finden wir

$$\left|\sum_{n=0}^{\infty}\left(\frac{3}{4}\right)^n \varphi(4^n x)\right| \leq \sum_{n=0}^{\infty}\left(\frac{3}{4}\right)^n \|\varphi(4^n x)\| = \sum_{n=0}^{\infty}\left(\frac{3}{4}\right)^n < \infty,$$

dass heißt, die obige Reihe konvergiert gleichmäßig gegen f. Da alle Partialsummen stetig auf $\mathbb{R}$ sind, ist f eine auf ganz $\mathbb{R}$ stetige Funktion.

Seien $x \in \mathbb{R}$ und $m \in \mathbb{N}$ beliebig gewählt. Wir zeigen, dass f an der Stelle x nicht differenzierbar ist. Dazu nehmen wir

$$\delta_m := \pm\frac{1}{2} \cdot 4^{-m},$$

wobei das Vorzeichen so gewählt sei, dass zwischen $4^m x$ und $4^m(x+\delta_m)$ keine ganze Zahl liegt. Dies ist möglich, da $4^m|\delta_m| = 1/2$.

Wir definieren

$$\gamma_n := \frac{\varphi\big(4^n(x+\delta_m)\big) - \varphi(4^n x)}{\delta_m}.$$

Für $n > m$ ist $4^n\delta_m = \pm 1/2 \cdot 4^{n-m}$ eine gerade ganzzahlige Zahl. Wir erinnern uns, dass $f(x+2) = f(x)$ ist. Somit gilt wegen der Periodizität von φ dass $\gamma_n = 0$ für alle $n > m$. Für $0 \leq n \leq m$ folgt aus (8.3), dass $|\gamma_n| \leq 4^n$ gilt.

Da insbesondere $|\gamma_m| = 4^m$ ist, folgern wir

$$\begin{aligned}
\left|\frac{f(x+\delta_m) - f(x)}{\delta_m}\right| &= \left|\frac{\sum_{n=0}^{\infty}\left(\frac{3}{4}\right)^n \varphi\big(4^n(x+\delta_m)\big) - \sum_{n=0}^{\infty}\left(\frac{3}{4}\right)^n \varphi(4^n x)}{\delta_m}\right| \\
&= \left|\sum_{n=0}^{m}\left(\frac{3}{4}\right)^n \frac{\varphi\big(4^n(x+\delta_m)\big) - \varphi(4^n x)}{\delta_m}\right| \\
&= \left|\sum_{n=0}^{m}\left(\frac{3}{4}\right)^n \gamma_n\right| \\
&\geq 3^m - \sum_{n=0}^{m-1}\left|\left(\frac{3}{4}\right)^n \gamma_n\right| \\
&\geq 3^m - \sum_{n=0}^{m-1} 3^n \\
&= (3^m - 1)/2.
\end{aligned}$$

Für $m \to \infty$ gilt $\delta_m \to 0$. Es folgt daher, dass f an der Stelle $x \in \mathbb{R}$ nicht differenzierbar ist. $\square$

8.2 Potenzreihen

In diesem Abschnitt werden wir einige Eigenschaften von Funktionen betrachten, die durch Potenzreihen dagestellt werden können, d.h. Funktionen der Form

$$f(x) = \sum_{k=0}^{\infty} c_k x^k$$

oder allgemeiner der Form

$$f(x) = \sum_{k=0}^{\infty} c_k (x-a)^k$$

mit $a, c_k \in \mathbb{R}$ oder $a, c_k \in \mathbb{C}$.

Satz 8.2.1. *Falls die Reihe $f(x) = \sum_{k=0}^{\infty} c_k x^k$ für alle $x \in (-R, R)$ konvergiert, dann konvergiert für jedes $\varepsilon > 0$ diese Reihe gleichmäßig auf $[-R+\varepsilon, R-\varepsilon]$ gegen eine stetige Funktion. Für alle $x \in (-R, R)$ ist f differenzierbar mit $f'(x) = \sum_{k=1}^{\infty} k c_k x^{k-1}$.*

Beweis. Es gilt für alle in $x \in [-R+\varepsilon, R-\varepsilon]$ die Abschätzung

$$|c_k x^k| \leq |c_k|(R-\varepsilon)^k.$$

Aufgrund der Voraussetzung ist

$$\sum_{k=0}^{\infty} |c_k||R-\varepsilon|^k$$

konvergent. Mit Korollar 8.1.7 folgt, dass die Reihe gleichmäßig stetig ist. Wegen

$$\lim_{k\to\infty} \sqrt[k]{k} = 1$$

erhalten wir

$$\lim_{k\to\infty} \sup \sqrt[k]{k|c_k|} = \lim_{k\to\infty} \sup \sqrt[k]{|c_k|},$$

d.h. die Reihen $\sum c_k x^k$ und $\sum k c_k x^{k-1}$ haben den gleichen Konvergenzradius. Da die Reihe $\sum k c_k x^{k-1}$ in $[\varepsilon - R, R - \varepsilon]$ gleichmäßig konvergiert, folgt die letzte Aussage mit Satz 8.1.14. □

Korollar 8.2.2. *Sei $f(x) = \sum_{k=0}^{\infty} c_k x^k$ in $(-R, R)$ konvergent, dann ist f unendlich oft differenzierbar in diesem Intervall, d.h. $f \in C^k\big((-R,R)\big)$ für alle $k \in \mathbb{N}_0$, kurz $f \in C^\infty\big((-R,R)\big)$. Es gilt*

$$f^{(n)}(x) = \sum_{k=n}^{\infty} k(k-1)(k-2)\cdots(k-n+1) c_k x^{k-n}$$

und $f^{(n)}(0) = n! c_n$.

Beweis. Diese Aussage folgt durch wiederholte Anwendung des vorhergehenden Satzes. □

Bemerkung 8.2.3. *1. Die Aussagen bezüglich der gleichmäßigen Konvergenz übertragen sich auch ins Komplexe. Dabei werden die Intervalle $(-R, R)$ bzw. $[-R+\varepsilon, R-\varepsilon]$ ersetzt durch die Kreisscheiben $\{z \in \mathbb{C} : |z| < R\}$ bzw. $\{z \in \mathbb{C} : |z| < R - \varepsilon\}$.*

2. Wir haben gesehen, dass Potenzreihen in $C^\infty(I)$ sind, wobei I ein geeignetes Intervall bezeichne. Umgekehrt gibt es aber Funktionen $f \in C^\infty(I)$, die sich nicht durch Potenzreihen darstellen lassen. Funktionen, die sich in Potenzreihen entwickeln lassen, heißen **analytisch**.

Lemma 8.2.4. *Seien $a_{i,j} \in \mathbb{R}$ bzw. $a_{i,j} \in \mathbb{C}$ mit $\sum_{j=0}^\infty |a_{i,j}| =: b_i$ und $\sum b_i < \infty$ konvergent, dann gilt*

$$\sum_{j=0}^\infty b_i = \sum_{i=0}^\infty \left(\sum_{j=0}^\infty a_{i,j} \right).$$

Beweis. Gegeben sei eine reelle Zahlenfolge $(x_n)_{n\in\mathbb{N}}$ mit $x_n \to x_0$ für $n \to \infty$. Auf der Menge $E := \{x_n : n \in \mathbb{N}_0\}$ definieren wir für $i \in \mathbb{N}$ die Funktionen

$$f_i(x_n) := \sum_{j=0}^n a_{i,j}, \qquad n \in \mathbb{N},$$

und

$$f_i(x_0) := \sum_{j=0}^\infty a_{i,j}.$$

Ferner sei für $x \in E$

$$g(x) = \sum_{i=1}^\infty f_i(x).$$

Der Punkt x_0 ist der einzige Häufungspunkt von E. Weil

$$\lim_{x_i \to x_0} f_i(x_n) = f_i(x_0)$$

gilt, sind die Funktionen f_i für alle $i \in \mathbb{N}$ auf E stetig. Da $|f_i(x)| \leq b_i$ für alle $x \in E$ gilt und $\sum b_i$ konvergiert, konvergiert nach Korollar 8.1.7 auch die Reihe $\sum_{i=1}^\infty f_i(x)$ gleichmäßig gegen g. Nach Satz 8.1.8 ist somit g ebenfalls stetig auf E und die Grenzprozesse sind vertauschbar. □

Satz 8.2.5 (Taylor-Reihe). *Sei*

$$f(x) = \sum_{k=0}^\infty c_k x^k$$

konvergent für alle $|x| < R$ und sei $|x_0| < R$. Dann kann f im Punkt x_0 ebenfalls in eine Taylor-Reihe

$$f(x) = \sum_{k=0}^\infty \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k$$

entwickelt werden. Diese Entwicklung ist konvergent für alle $x \in \{x \in \mathbb{R} : |x - x_0| < R - |x_0|\}$.

Beweis. Wir bemerken zuerst, dass

$$\begin{aligned}\sum_{k=0}^{\infty}\sum_{j=0}^{k}|c_k\binom{k}{j}x_0^{k-j}(x-x_0)^j| &= \sum_{k=0}^{\infty}\sum_{j=0}^{k}|c_k|\binom{k}{j}|x_0|^{k-j}|x-x_0|^j \\ &= \sum_{k=0}^{\infty}|c_k|(|x_0|+|x-x_0|)^k\end{aligned}$$

für alle $x \in \{x : |x-x_0|+|x_0| < R\}$ konvergiert. Daher können wir Lemma 8.2.4 benutzen, um die Summationsreihe zu vertauschen und die gewünschte Umentwicklung zu erhalten

$$\begin{aligned}f(x) &= \sum_{k=0}^{\infty}c_kx^k = \sum_{k=0}^{\infty}c_k[(x+x_0)-x_0]^k = \sum_{k=0}^{\infty}\sum_{j=0}^{k}c_k\binom{k}{j}x_0^{k-j}(x-x_0)^j \\ &= \sum_{j=0}^{\infty}\left(\sum_{k=j}^{\infty}c_k\binom{k}{j}x_0^{k-j}\right)(x-x_0)^j.\end{aligned}$$

Es sei angemerkt, dass wir hier

$$\begin{aligned}&\{(k,j)\in\mathbb{N}_0^2 : 0\le k<\infty, 0\le j\le k\} \\ &\quad = \{(k,j)\in\mathbb{N}_0^2 : 0\le j<\infty, j\le k<\infty\} \\ &\quad = \{(k,j)\in\mathbb{N}_0^2 : 0\le j\le k\}\end{aligned}$$

benutzt haben. Gemäß Korollar 8.2.2 gilt nun

$$\begin{aligned}\frac{f^{(j)}(x_0)}{j!} &= \frac{1}{j!}\sum_{k=j}^{\infty}c_kk(k-1)\cdots(k-j+1)x_0^{k-j} = \sum_{k=j}^{\infty}c_k\frac{k!}{j!(k-j)!}x_0^{k-j} \\ &= \sum_{k=j}^{\infty}c_k\binom{k}{j}x_0^{k-j},\end{aligned}$$

womit dann die Behaptung folgt. □

Satz 8.2.6 (Identitätssatz für Potenzreihen). *Seien die Funktionen*

$$f(x) = \sum_{k=0}^{\infty}f_kx^k, \qquad g(x) = \sum_{k=0}^{\infty}g_kx^k,$$

konvergent in $S = \{x : |x| < R\}$ *und sei* $E := \{x \in S : f(x) = g(x)\}$. *Falls ein Häufungspunkt* z *von* E *in* S *liegt, dann gilt* $S \subseteq \{x : f(x) = g(x)\}$ *und* $f_k = g_k$ *für alle* $k \in \mathbb{N}_0$

Beweis. Wir setzen $h_k := f_k - g_k$ für alle $k \in \mathbb{N}_0$ und betrachten für $x \in S$ die Funktion

$$h(x) = \sum_{k=0}^{\infty}h_kx^k = \sum_{k=0}^{\infty}(f_k-g_k)x^k = f(x)-g(x).$$

Sei H die Menge aller Häufungspunkte von E in S. Dann ist H abgeschlossen und $B := S \setminus H$ offen. Falls wir zeigen können, dass H auch eine offene Menge ist, dann sind H und B getrennte Mengen. Allerdings ist $S = B \cup H = \{x : |x| < R\}$ zusammenhängend. Folglich ist $H \neq \emptyset$, d.h. $B = \emptyset$, und somit $H = S$. Hierdurch haben wir die Behauptung bewiesen, falls wir zeigen, dass H offen ist.

Sei $x_0 \in H \subseteq S$ beliebig, dann gilt nach Satz 8.2.5, dass $h(x) = \sum_{k=0}^{\infty} h_k(x - x_0)^k$ diejenige Reihe ist, in die sich h an der Stelle x_0 entwickeln lässt für $|x - x_0| < \delta$ mit $\delta > 0$ geeignet. Ferner gilt nach Voraussetzung $h(x_0) = 0$.

Angenommen, es gibt ein $h_k \neq 0$, dann sei $j = \min\{k \in \mathbb{N}_0 : h_k \neq 0\}$, d.h. es gilt

$$h(x) = \sum_{k=j}^{\infty} h_k(x - x_0)^k = (x - x_0)^j b(x)$$

mit

$$b(x) = \sum_{k=0}^{\infty} h_{j+k}(x - x_0)^k = \frac{h(x)}{(x - x_0)^j}.$$

Die Funktion b ist stetig in x_0 und wegen $b(x_0) = h_j \neq 0$ existiert ein $\delta > 0$ mit $b(x) \neq 0$ für alle $|x - x_0| < \delta$. Daraus folgt $h(x) = (x - x_0)^j b(x) \neq 0$ für alle $0 < |x - x_0| < \delta$. Dies ist ein Widerspruch zu der Tatsache, dass x_0 ist Häufungspunkt von E.

Somit gilt $h_j = 0$ für alle $j \in \mathbb{N}$ und folglich $h(x) = 0$ für alle $|x - x_0| < \delta$, d.h. für alle $x \in U_\delta(x_0) \cap E$. Folglich ist x_0 ein innerer Punkt von E. Wir haben also tatsächlich nachgewiesen, dass H offen ist. □

Definition 8.2.7. *Die Funktion $z \mapsto e^z$ definiert durch*

$$e^z = \sum_{k=0}^{\infty} \frac{z^k}{k!}$$

heißt **Exponentialfunktion**.

Satz 8.2.8.

1. *Es gilt*

$$e^{z+w} = e^z \cdot e^w$$

für alle $z, w \in \mathbb{C}$.

2. *Es gilt*

$$(e^x)'(x) = \frac{d}{dx} e^x = e^x$$

für alle $x \in \mathbb{R}$ und

$$\int e^x dx = e^x + C.$$

3. *Es gilt $e^x > 0$ für alle $x \in \mathbb{R}$ und $x \mapsto e^x$ ist eine streng monoton wachsende Funktion.*

Beweis. Der Beweis verbleibt dem Leser als Übung. □

Definition 8.2.9. *Da die (reelle) Exponentialfunktion streng monoton wachsend ist, ist die Abbildung*

$$f : \begin{cases} \mathbb{R} \to \mathbb{R}^+, \\ x \mapsto e^x, \end{cases}$$

bijektiv. Die Umkehrabbildung $f^{-1} : \mathbb{R}^+ \to \mathbb{R}$ ist die **Logarithmusfunktion** *oder der* **natürliche Logarithmus**

$$f^{-1}(x) = \log(x) = \ln(x).$$

Satz 8.2.10.

1. *Für alle $x, y > 0$ gilt*

$$\log x + \log y = \log(x \cdot y).$$

2. *Es gilt die Umrechnungsformel*

$$a^x = e^{x \log a}$$

für alle $a > 0$.

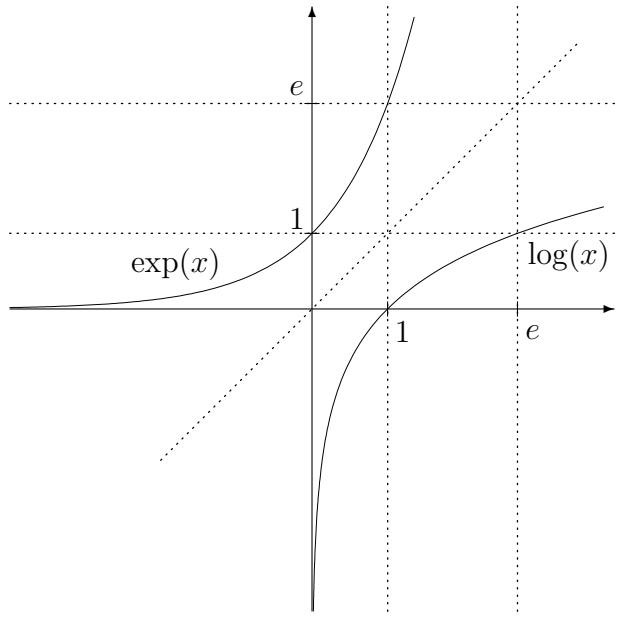
3. *Es gelten*

$$\frac{d}{dx} \log x = \frac{1}{x}$$

bzw.

$$\int \frac{dx}{x} = \log |x| + C.$$

Beweis. Der Beweis verbleibt dem Leser als Übung. □



Definition 8.2.11. *Für $z \in \mathbb{C}$ definieren wir den* **Cosinushyperbolikus** *via*

$$\cosh z = \frac{1}{2}(e^z + e^{-z}),$$

den **Sinushyperbolikus** *via*

$$\sinh z = \frac{1}{2}(e^z - e^{-z}),$$

und die **trigonometrischen Funktionen** *vermittels*

$$\begin{aligned}
\cos z &= \frac{1}{2}(e^{iz} + e^{-iz}), \\
\sin z &= \frac{1}{2i}(e^{iz} - e^{-iz}), \\
\tan z &= \frac{\sin z}{\cos z}, \quad z \neq k\pi + \pi/2, \quad k \in \mathbb{Z}, \\
\cot z &= \frac{\cos z}{\sin z}, \quad z \neq k\pi, \qquad\quad k \in \mathbb{Z}.
\end{aligned}$$

Bemerkung 8.2.12. *Setzen wir die Potenzreihe der Exponentialfunktion in die Definition des Cosinus bzw. Sinus ein, erhalten wir die Identitäten*

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

Analog folgen für Cosinushyperbolikus und Sinushyperbolikus die Reihenentwicklungen

$$\cosh z = \sum_{k=0}^{\infty} \frac{z^{2k}}{(2k)!}, \qquad \sinh z = \sum_{k=0}^{\infty} \frac{z^{2k+1}}{(2k+1)!}.$$

Vergleicht man diese Potenzreihen miteinander, so ergibt dies die Beziehung

$$\cos z = \cosh iz, \qquad \sin z = \frac{1}{i}\sinh iz.$$

Korollar 8.2.13.

1. *Für alle $x \in \mathbb{R}$ gilt die Eulersche Formel*

$$e^{ix} = \cos x + i \sin x.$$

2. *Für alle $z \in \mathbb{C}$ gilt die Identität*

$$\cos^2 z + \sin^2 z = 1.$$

3. *Für alle $z, w \in \mathbb{C}$ gelten die folgenden Additionstheoreme*

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

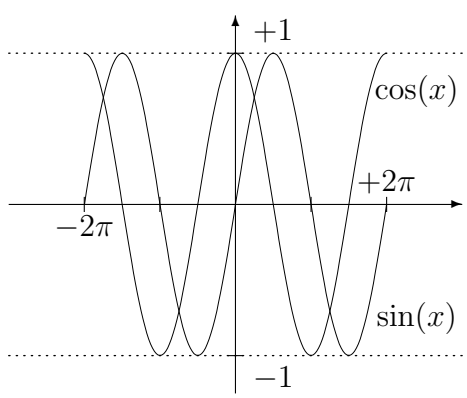
4. *Für $x \in \mathbb{R}$ gilt die Moivresche Formel*

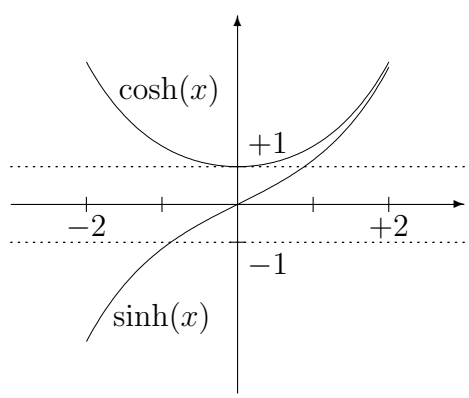
$$(\cos x + i \sin x)^n = \cos nx + i \sin nx.$$

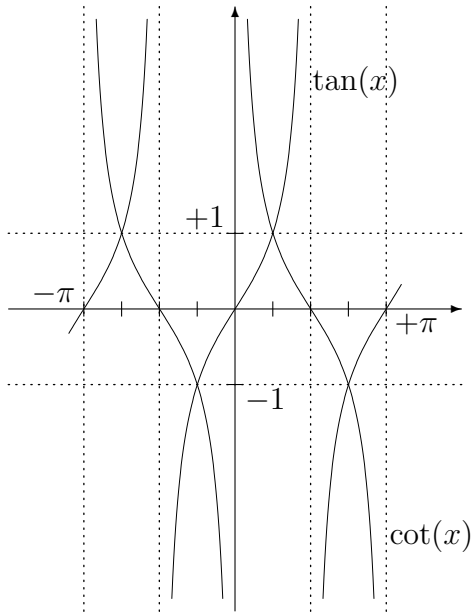
5. *Für alle $x \in \mathbb{R}$ gilt*

$$\begin{aligned} (\sin x)' &= \frac{d}{dx}\sin x = +\cos x, \\ (\cos x)' &= \frac{d}{dx}\cos x = -\sin x. \end{aligned}$$

Beweis. Der Beweis verbleibt dem Leser als Übung. □







Bemerkung 8.2.14. *Aus den ersten zwei Aussagen des Satzes folgt*

$$|e^{ix}|^2 = e^{ix}\overline{e^{ix}} = (\cos x + i \sin x)(\cos x - i \sin x) = \cos^2 x + \sin^2 x = 1,$$

d.h. $|e^{ix}| = 1$ *für alle* $x \in \mathbb{R}$.

Definition 8.2.15.

1. *Die Funktion*

$$f : \begin{cases} [-\frac{\pi}{2}, \frac{\pi}{2}] \to [-1, 1], \\ x \mapsto \sin x, \end{cases}$$

ist streng monoton wachsend und somit bijektiv. Die Umkehrfunktion $f^{-1} : [-1, 1] \to [-\frac{\pi}{2}, \frac{\pi}{2}]$ *ist der* **Arcussinus**

$$f^{-1}(x) = \arcsin x.$$

2. *Die Funktion*

$$f : \begin{cases} [0, \pi] \to [-1, 1], \\ x \mapsto \cos x, \end{cases}$$

ist streng monoton fallend und somit bijektiv. Die Umkehrfunktion $f^{-1} : [-1, 1] \to [0, \pi]$ *ist der* **Arcuscosinus**

$$f^{-1}(x) = \arccos x.$$

3. *Die Funktion*

$$f : \begin{cases} (-\frac{\pi}{2}, \frac{\pi}{2}) \to \mathbb{R}, \\ x \mapsto \tan x, \end{cases}$$

ist streng monoton steigend und somit bijektiv. Die Umkehrfunktion $f^{-1} : \mathbb{R} \to (-\frac{\pi}{2}, \frac{\pi}{2})$ *ist der* **Arcustangens**

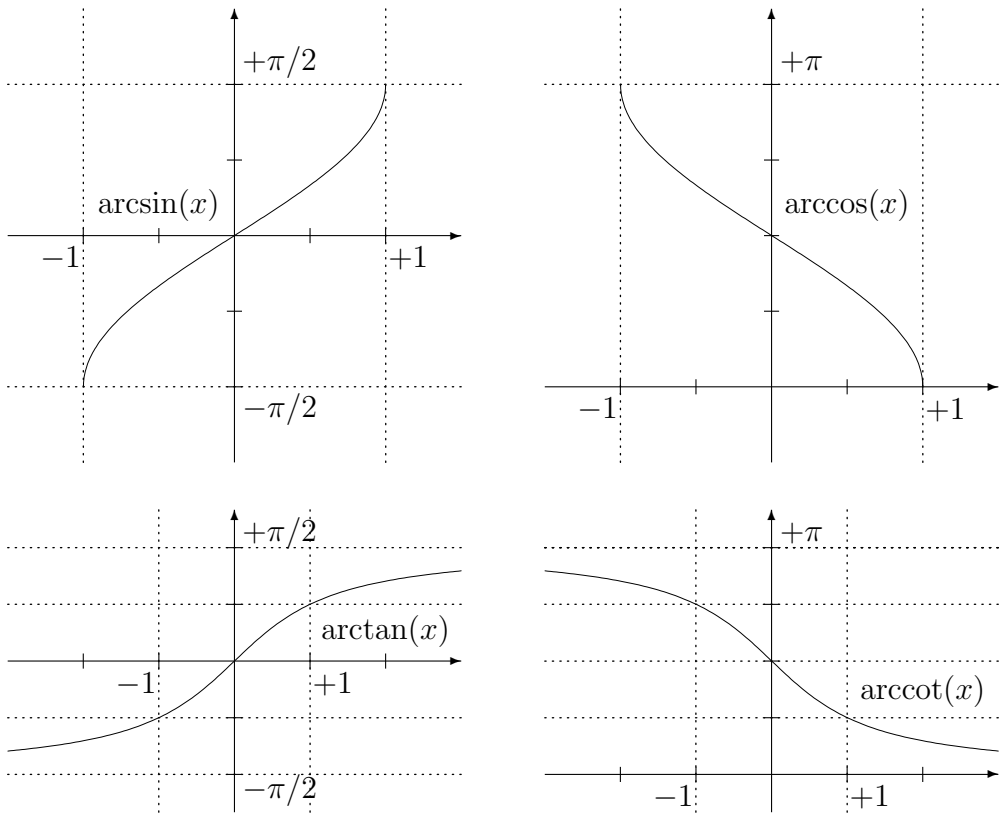
$$f^{-1}(x) = \arctan x.$$

4. *Die Funktion*

$$f : \begin{cases} (0,\pi) \to \mathbb{R}, \\ x \mapsto \cot x, \end{cases}$$

ist streng monoton fallend und somit bijektiv. Die Umkehrfunktion $f^{-1} : \mathbb{R} \to (0,\pi)$ *ist der* **Arcuscotangens**

$$f^{-1}(x) = \operatorname{arccot} x.$$



Satz 8.2.16. *Es gilt für alle* $|x| \leq 1$

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

bzw.

$$\int \frac{dx}{\sqrt{1-x^2}} = \arcsin x + C, \qquad \int \frac{-dx}{\sqrt{1-x^2}} = \arccos x + C,$$

und für alle $x \in \mathbb{R}$

$$(\arctan x)' = \frac{1}{1+x^2}, \qquad (\operatorname{arccot} x)' = -\frac{1}{1+x^2}$$

bzw.

$$\int \frac{dx}{1+x^2} = \arctan x + C, \qquad \int \frac{-dx}{1+x^2} = \operatorname{arccot} x + C,$$

Beweis. Der Beweis verbleibt dem Leser als Übung. □

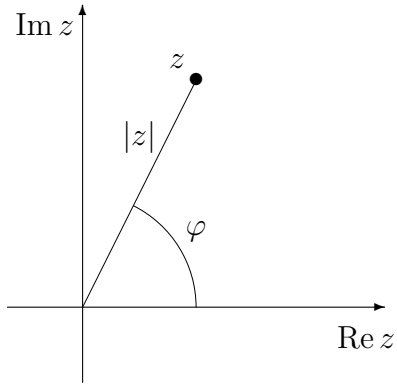
Satz 8.2.17 (Polardarstellung komplexer Zahlen). *Zu jeder komplexen Zahl $z \in \mathbb{C}$ gibt es genau ein $\varphi \in [0, 2\pi]$ mit $z = |z|e^{i\varphi}$.*

Beweis. Sei $z = x + iy$ mit $x, y \in \mathbb{R}$. Ohne Einschränkung der Allgemeinheit sei $|z| = \sqrt{x^2 + y^2} = 1$ mit $x \geq 0$. Damit folgt $0 \leq x \leq 1$ und $y^2 = 1 - x^2$ folgt. Setzen wir nun $\varphi := \arccos x$, dann gilt in der Tat

$$\begin{aligned} |z|e^{i\varphi} &= 1 \cdot e^{i \arccos x} = \cos(\arccos x) + i \sin(\arccos x) \\ &= x + i\sqrt{1 - \cos^2(\arccos x)} = x + i\sqrt{1 - x^2} \\ &= x + iy = z. \end{aligned}$$

Den Beweis der Eindeutigkeit der Polardarstellung verbleibt dem Leser als Übung. □

Geometrisch kann man sich die Aussage dieses Satzes wie folgt veranschaulichen:



Satz 8.2.18 (Fundamentalsatz der Algebra). *Seien $a_0, \ldots, a_n \in \mathbb{C}$ mit $a_n \neq 0$ und $n \in \mathbb{N}$ beliebig. Sei*

$$p(z) = \sum_{k=0}^{n} a_k z^k = a_0 + a_1 z + \ldots + a_n z^n$$

ein komplexes Polynom vom Grade n. Dann existiert ein $z^\star \in \mathbb{C}$ mit $P(z^\star) = 0$.

Beweis. Sei $\alpha := \inf_{z \in \mathbb{C}} |P(z)|$. Wegen

$$|p(z)| = |z^n||a_n + a_{n-1}z^{-1} + \ldots + a_0 z^{-n}|$$

gilt $|p(z)| \to \infty$ für $z \to \infty$. Daraus folgt, dass $R > 0$ und $C > 0$ existieren mit $\alpha \leq |p(z)| \leq C$ für alle $|z| \leq R$ und $|p(z)| > \alpha$ für $|z| > R$.

Nun ist die Funktion $z \mapsto |p(z)|$ stetig auf der kompakten Menge $D = \{z \in \mathbb{C} : |z| \leq R\}$. Nach Satz 5.5.4 nehmen stetige Funktionen auf einer kompakten Menge ihre Extremalwerte an, d.h. es existiert ein $z^\star \in D$ mit

$$|p(z^\star)| = \alpha = \min_{z \in D} |P(z)|.$$

Angenommen, es gilt nun $\alpha = |p(z^\star)| > 0$. Wir benutzen z.B. die Taylorentwicklung von $p(z)$ um das Polynom um $z^\star$ zu entwickeln

$$\begin{aligned} p(z) &= \sum_{k=0}^{n} \frac{p^{(k)}(z^\star)}{k!}(z - z^\star)^k \\ &= p(z^\star) + \sum_{k=l}^{n} p_k (z - z^\star)^k \\ &= p(z^\star) + (z - z^\star)^l \sum_{k=0}^{n-l} p_{k-l}(z - z^\star)^k \end{aligned}$$

mit gewissen Koeffizienten $p_k \in \mathbb{C}$. Dabei ist $l \in \mathbb{N}$ die kleinste Zahl mit $p_l \neq 0$.

Wir betrachten nun das Polynom

$$q(z) = \frac{p(z + z^\star)}{p(z^\star)}.$$

Offensichtlich ist $q(0) = 1$ und $|q(z)| \geq 1$ für alle $z \in D$. Weiter gilt

$$\begin{aligned} q(z) &= \frac{p(z^\star) + z^l \sum_{k=0}^{n-l} p_{k-l} z^k}{p(z^\star)} \\ &= 1 + z^l \sum_{k=0}^{n-l} \frac{p_{k-l}}{p(z^\star)} z^k \\ &=: 1 + z^l \sum_{k=0}^{n-l} q_k z^k \end{aligned}$$

mit gewissen Koeffizienten $q_k \in \mathbb{C}$.

Nach Satz 8.2.17 existiert ein $\varphi \in [0, 2\pi]$ mit $q_0 = -|q_0| e^{i\varphi}$. Dann gilt für $z := re^{-i\varphi/l}$ mit $r < \sqrt[l]{|q_0|}$ dass

$$|1 + q_0 z^l| = \left|1 - |q_0| e^{i\varphi} r^l e^{-i\varphi}\right| = 1 - |q_0| r^l.$$

Wir schätzen nun für dieses spezielle z die Größe $|q(z)|$ ab

$$\begin{aligned}
|q(z)| &= \left|1 + q_l z^l + z^l \sum_{k=1}^{n-l} q_k z^k\right| \\
&\leq 1 - |q_0| r^l + \left|z^l \sum_{k=1}^{n-l} q_k z^k\right| \\
&\leq 1 - |q_0| r^l + r^l \sum_{k=1}^{n-l} |q_k| r^k \\
&= 1 - \{|q_0| - A(r)\} r^l
\end{aligned}$$

mit $A(r) \to 0$ für $r \to 0$. Für $r < \delta$ hinreichend klein gilt $|q_0| - A(r) > 0$ und daher $|q(z)| < 1$ für alle $|z| < \delta$ mit $z \neq z^\star$. Dies ist aber ein Widerspruch zu $|q(z)| \geq 1$. □

Korollar 8.2.19. *Sei $p(z) = \sum_{k=0}^{n} a_k z^k$ mit $a_n \neq 0$. Dann existieren genau n, nicht notwendigerweise verschiedene, Nullstellen $z_1, \ldots, z_n \in \mathbb{C}$.*

Beweis. Wegen Satz 8.2.18 existiert ein $z_1 \in \mathbb{C}$ mit $p(z_1) = 0$. Dann ist $q(z) = p(z)/(z - z_1)$ ein Polynom vom Grad $n - 1$. Wenden wir Satz 8.2.18 auf $q(z)$ an, finden wir ein $z_2 \in \mathbb{C}$ mit $p(z_2) = 0$. Da nun $q(z)/(z - z_2)$ ein Polynom vom Grad $n - 2$ ist, erhalten wir so fortfahrend die Behauptung. □

8.3 Fourier-Reihen

Wir betrachten zunächst *trigonometrische Polynome*. Diese sind Funktionen der Form

$$p_N(x) = \frac{a_0}{2} + \sum_{k=1}^{N} a_k \cos kx + b_k \sin kx$$

mit $x \in \mathbb{R}$ und Koeffizienten $a_k, b_k \in \mathbb{C}$. Diese trigonometrischen Polynome sind 2π-*periodische* Funktionen, d.h. es gilt $p_N(x) = p_N(x + 2\pi)$ für alle $k \in \mathbb{Z}$.

Wir können mit Hilfe der Eulerschen Formel p_N auch folgendermaßen darstellen

$$p_N(x) = \sum_{k=-N}^{N} c_k e^{ikx}$$

mit Koeffizienten $c_k \in \mathbb{C}$, was in den meisten Fällen zweckmäßiger ist. Wegen der Beziehung

$$\int_{-\pi}^{\pi} e^{ikx} dx = \int_{x_0}^{x_0+2\pi} e^{ikx} dx = \begin{cases} 2\pi, & k = 0, \\ 0, & k \in \mathbb{Z} \setminus \{0\}, \end{cases}$$

erhalten wir für $m \in \mathbb{N}$

$$\begin{aligned}\frac{1}{2\pi}\int_{-\pi}^{\pi} p_N(x)e^{-imx}dx &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\sum_{k=-N}^{N} c_k e^{i(k-m)x}\right)dx \\ &= \sum_{k=-N}^{N} c_k \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{i(k-m)x}dx \\ &= \begin{cases} c_m, & |m| \leq N, \\ 0, & \text{sonst.} \end{cases}\end{aligned}$$

Definition 8.3.1. *Sei $(\varphi_n)_{n\in\mathbb{N}}$ eine Folge komplexer Funktionen auf $[a,b]$, so dass*

$$\int_a^b \varphi_k(x)\overline{\varphi_{k'}(x)}dx = 0$$

für alle $k \neq k'$. Dann nennt man $(\varphi_n)_{n\in\mathbb{N}}$ ein **orthogonales Funktionensystem** *auf $[a,b]$. Gilt darüberhinaus auch*

$$\int_a^b |\varphi_k(x)|^2 dx = 1,$$

für alle $k \in \mathbb{N}$, so wird $(\varphi_n)_{n\in\mathbb{N}}$ als **orthonormal** *bezeichnet. Ein orthonormales Funktionensystem heißt kurz* **Orthonormalsystem**.

Bemerkung 8.3.2. *Schreiben wir für zwei komplexwertige Funktionen f und g auf $[a,b]$ abkürzend*

$$\langle f, g\rangle := \int_a^b f(x)\overline{g(x)}dx$$

und benutzten wir das Kroneckersymbol

$$\delta_{k,k'} := \begin{cases} 1, & k = k' \\ 0, & \textit{sonst}, \end{cases}$$

so lässt sich ein Orthonormalssystem kurz durch

$$\langle \varphi_k, \varphi_{k'}\rangle = \delta_{k,k'}$$

charakterisieren.

Es sei angemerkt, dass für die Abbildung $\langle\cdot,\cdot\rangle$ folgende Eigenschaften gelten:

1. $\langle f+g, h\rangle = \langle f, h\rangle + \langle g, h\rangle$,
2. $\langle \alpha f, g\rangle = \alpha\langle f, g\rangle$ *für alle* $\alpha \in \mathbb{C}$,
3. $\langle f, g\rangle = \overline{\langle g, f\rangle}$,

4. $\langle f, f\rangle \geq 0$, *wobei für stetiges* f *genau dann gilt* $\langle f, f\rangle = 0$, *wenn* $f \equiv 0$.

Mit Hilfe dieser Schreibweise liese sich nachfolgend vieles eleganter schreiben, aber um den Leser nicht zu verwirren, verzichten wir lieber darauf.

Zum Beispiel bilden die Funktionen $e^{ikx}/\sqrt{2\pi}$ für $k \in \mathbb{N}$ ein Orthonormalsystem auf $[-\pi, \pi]$, denn es gilt in der Tat

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{ikx}\overline{e^{ik'x}}dx = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{i(k-k')x}dx = \delta_{k,k'}.$$

Das gleiche gilt für die reellen Funktionen

$$\frac{1}{\sqrt{2\pi}}, \frac{\cos x}{\sqrt{\pi}}, \frac{\sin x}{\sqrt{\pi}}, \frac{\cos 2x}{\sqrt{\pi}}, \frac{\sin 2x}{\sqrt{\pi}}, \ldots.$$

Definition 8.3.3. *Ist* $(\varphi_n)_{n\in\mathbb{N}}$ *ein Orthonormalsystem auf* $[a, b]$ *und ist*

$$c_n = \int_a^b f(x)\overline{\varphi_n(x)}dx, \qquad n \in \mathbb{N}, \tag{8.4}$$

so nennt man c_n *den* n*-ten* **Fourier-Koeffizienten** *von* f *bezüglich* $(\varphi_n)_{n\in\mathbb{N}}$. *Wir schreiben*

$$f(x) \sim \sum_{n=1}^{\infty} c_n\varphi_n(x), \tag{8.5}$$

und nennen diese Reihe die **Fourier-Reihe** *von* f *bezüglich* $(\varphi_n)_{n\in\mathbb{N}}$.

Man beachte, dass das Symbol "$\sim$" in (8.5) nichts über die Konvergenz der Fourier-Reihe aussagt, sondern wir meinen lediglich, dass die Koeffizienten durch (8.4) gegeben sind.

Die folgenden Sätze zeigen, dass die Partialsummen der Fourier-Reihe von f eine bestimmte Minimaleigenschaft haben. Wir setzen hier und im verbleibenden Teil dieses Abschnitts $f \in \mathcal{R}([a, b])$ voraus, obwohl diese Voraussetzung abgeschwächt werden kann.

Satz 8.3.4 (Approximationssatz für Fourier-Reihen). *Sei* $(\varphi_n)_{n\in\mathbb{N}}$ *ein Orthonormalsystem auf* $[a, b]$ *und sei*

$$s_n(x) = \sum_{k=1}^{n} c_k\varphi_k(x)$$

die n*-te Partialsumme der Fourier-Reihe von* f. *Dann gilt für jede beliebige Summe*

$$t_n(x) = \sum_{k=1}^{n} \gamma_k\varphi_k(x),$$

die Abschätzung

$$\int_a^b |f - s_n|^2 dx \leq \int_a^b |f - t_n|^2 dx, \tag{8.6}$$

wobei Gleichheit genau dann gilt, wenn

$$c_k = \gamma_k, \qquad k = 1, \ldots, n.$$

Dies bedeutet, unter allen Funktionen t_n ermöglicht s_n die bestmögliche Approximation von f im quadratischen Mittel, d.h. im Sinne von (8.6).

Beweis. Es gilt nach Definition der Fourier-Koeffizienten c_k

$$\int_a^b f(x)\overline{t_n(x)}dx = \int_a^b f(x)\left(\sum_{k=1}^n \overline{\gamma_k}\overline{\varphi_k(x)}\right)dx = \sum_{k=1}^n \overline{\gamma_k}\int_a^b f(x)\overline{\varphi_k(x)}dx = \sum_{k=1}^n c_k\overline{\gamma_k}$$

und

$$\int_a^b |t_n(x)|^2 dx = \int_a^b t_n(x)\overline{t_n(x)}dx = \int_a^b \left(\sum_{k=1}^n \gamma_k\varphi_k(x)\right)\left(\sum_{l=1}^n \overline{\gamma_l}\overline{\varphi_l(x)}\right)dx = \sum_{k=1}^n |\gamma_k|^2,$$

da $(\varphi_n)_{n\in\mathbb{N}}$ orthonormal ist. Somit folgt

$$\begin{aligned}
\int_a^b |f(x) - t_n(x)|^2 dx &= \int_a^b |f(x)|^2 dx - \int_a^b f(x)\overline{t_n(x)}dx - \int_a^b t_n(x)\overline{f(x)}dx + \int_a^b |t_n(x)|^2 dx \\
&= \int_a^b |f(x)|^2 dx - \sum_{k=1}^n c_k\overline{\gamma_k} - \sum_{k=1}^n \gamma_k\overline{c_k} + \sum_{k=1}^n |\gamma_k|^2 \\
&= \int_a^b |f(x)|^2 dx - \sum_{k=1}^n |c_k|^2 + \sum_{k=1}^n |\gamma_k - c_k|^2,
\end{aligned}$$

und dieser Ausdruck ist genau dann minimal, wenn $\gamma_k = c_k$ gilt. □

Korollar 8.3.5. *Seien $(\varphi_n)_{n\in\mathbb{N}}$ ein Orthonormalsystem auf $[a, b]$ und*

$$f(x) \sim \sum_{k=1}^\infty c_k\varphi_k(x).$$

Dann gilt die Besselsche Ungleichung

$$\sum_{k=1}^\infty |c_k|^2 \leq \int_a^b |f(x)|^2 dx.$$

Insbesondere folgt hieraus $c_k \to 0$ für $k \to \infty$.

Beweis. Es gilt

$$\begin{aligned}\int_a^b |f(x)-s_n(x)|^2dx &= \int_a^b |f(x)|^2dx - \int_a^b f(x)\overline{s_n(x)}dx - \int_a^b s_n(x)\overline{f(x)}dx + \int_a^b |s_n(x)|^2dx \\ &= \int_a^b |f(x)|^2dx - \sum_{k=1}^n |c_k|^2,\end{aligned}$$

und daher wegen $\int_a^b |f(x)-s_n(x)|^2dx \geq 0$ die Abschätzung

$$\sum_{k=1}^n |c_k|^2 \leq \int_a^b |f(x)|^2dx$$

für alle $n \in \mathbb{N}$. Hieraus erhalten wir durch den Grenzprozess $n \to \infty$ die Behauptung. □

Von nun an spezifizieren wir uns ausschließlich auf trigonometrische Systeme. Wir betrachten Funktionen f mit der Periode 2π, die Riemann-integrierbar auf $[-\pi,\pi]$ sind. Die Fourier-Reihe von f ist dann die Reihe

$$f(x) \sim \sum_{k=-\infty}^{\infty} c_k e^{ikx}, \qquad x \in \mathbb{R},$$

mit den Koeffizienten

$$c_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-ikx}dx.$$

Die N-te Partialsumme der Fourier-Reihe von f ist gegeben durch

$$s_N(x) = s_N(f;x) = \sum_{k=-N}^{N} c_k e^{ikx}.$$

Die Besselsche Ungleichung lautet nun

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} |s_N(x)|^2dx = \sum_{k=-N}^{N} |c_k|^2 \leq \frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x)|^2dx.$$

Wir sehen uns die Darstellung der Partialsumme s_N nun etwas detailierter an

$$\begin{aligned} s_N(x) &= \sum_{k=-N}^{N} c_k e^{ikx} \\ &= \sum_{k=-N}^{N} \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) e^{-ikt} dt \cdot e^{ikx} \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \sum_{k=-N}^{N} e^{ik(x-t)} dt \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) D_N(x-t) dt \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} D_N(t) f(x-t) dt \end{aligned}$$

mit dem sogenannten *Dirichlet-Kern*

$$D_N(x) := \sum_{k=-N}^{N} e^{ikx}.$$

Wegen

$$(e^{ix} - 1) D_N(x) = \sum_{k=-N}^{N} (e^{i(k+1)x} - e^{ikx}) = \sum_{k=-N+1}^{N+1} e^{ikx} - \sum_{k=-N}^{N} e^{ikx} = e^{i(N+1)x} - e^{-iNx}$$

gilt die Beziehung

$$D_N(x) = \frac{\sin(N + \frac{1}{2})x}{\sin \frac{1}{2} x}.$$

Wir beweisen nun einen Satz über die punktweise Konvergenz von Fourier-Reihen.

Satz 8.3.6. *Sei $f : \mathbb{R} \to \mathbb{C}$ eine auf $\mathbb{R}$ stetige, 2π-periodische Funktion. Existiert für $x \in [-\pi, \pi]$ ein $\delta > 0$ und eine Konstante $L > 0$ mit*

$$|f(x+t) - f(x)| \leq L|t| \tag{8.7}$$

für alle $t \in (-\delta, \delta)$, dann konvergiert die Partialsumme $s_N(x)$ gegen $f(x)$ für $N \to \infty$.

Beweis. Sei

$$g(t) = \frac{f(x-t) - f(x)}{\sin(t/2)},$$

dann gilt wegen

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} D_N(t) dt = \frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{k=-N}^{N} e^{ikt} dt = \frac{1}{2\pi} \sum_{k=-N}^{N} \int_{-\pi}^{\pi} e^{ikt} dt = 1$$

die folgenden Umformungen

$$\begin{aligned}
s_N(x) - f(x) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-t)D_N(t) - f(x)D_N(t)dt \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} g(t)\sin(N+1/2)t\ dt \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} g(t)\big(\cos(t/2)\sin Nt + \sin(t/2)\cos Nt\big)dt \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} \big(g(t)\cos(t/2)\big)\sin Ntdt + \frac{1}{2\pi}\int_{-\pi}^{\pi} \big(g(t)\sin(t/2)\big)\cos Ntdt.
\end{aligned}$$

Aufgrund von (8.7) sind $g(t)\cos(t/2)$ und $g(t)\sin(t/2)$ beschränkt und Riemann-integrierbar. Wegen Korollar 8.3.5 folgt daher

$$\begin{aligned}
\frac{1}{2\pi}\int_{-\pi}^{\pi} \big(g(t)\cos(t/2)\big)\sin Ntdt &\overset{N\to\infty}{\longrightarrow} 0, \\
\frac{1}{2\pi}\int_{-\pi}^{\pi} \big(g(t)\sin(t/2)\big)\cos Ntdt &\overset{N\to\infty}{\longrightarrow} 0,
\end{aligned}$$

das heißt

$$\lim_{N\to\infty} s_N(x) - f(x) = 0.$$

□

Bemerkung 8.3.7. *Die Bedingung* (8.7) *ist äquivalent zur* **Lipschitz-Stetigkeit** *von f auf $(x-\delta, x+\delta)$: Eine Funktion $f : X \to Y$ heißt Lipschitz-stetig auf X, falls eine Konstante $L > 0$ existiert, so dass*

$$d_Y\big(f(x), f(y)\big) \leq L d_X(x,y)$$

für alle $x, y \in X$. Die Lipschitz-Stetigkeit von f impliziert Stetigkeit, ja sogar gleichmäßige Stetigkeit, von f auf X.

Korollar 8.3.8. *Sei $f \in C(\mathbb{R};\mathbb{C})$ eine 2π-periodische Funktion und sei $f(x) = 0$ für alle $x \in (a,b)$, d.h. $f(x) \equiv 0$ auf (a,b). Dann konvergiert für alle $x \in (a,b)$ die Fourier-Reihe von f gegen 0.*

Bemerkung 8.3.9. *Eine weitere Formulierung dieses Korollars lautet:*

Seien $f, g \in C(\mathbb{R};\mathbb{C})$ zwei 2π-periodische Funktionen und sei $f(x) \equiv g(x)$ auf (a,b), dann gilt

$$s_N(f;x) - s_N(g;x) = s_N(f-g;x) \overset{N\to\infty}{\longrightarrow} 0.$$

Dies wird gewöhnlich **Lokalisierungssatz** *genannt. Er zeigt, dass das Verhalten von $\big(s_N(f;x)\big)_{n\in\mathbb{N}}$, was die Konvergenz anbelangt, nur von den Werten von f in einer beliebig kleinen Umgebung von x abhängt. Zwei Fourier-Reihen können sich somit in einem Intervall gleich verhalten, in einem anderen Intervall jedoch kann ihr Verhalten völlig verschieden sein. Wir bemerken, dass aufgrund des Identitätssatzes ein solches Verhalten für Potenzreihen nicht möglich ist.*

Der "natürliche" Raum der Fourier-Reihen ist der Raum $L^2_{\text{per}}([-\pi,\pi])$. Diesen Raum können wir wie folgt definieren.

Definition 8.3.10. *Mit $L^2_{\text{per}}([-\pi,\pi])$ bezeichnen wir den metrischen Raum von 2π-periodischen Funktionen $f:\mathbb{R}\to\mathbb{C}$ versehen mit der Metrik*

$$d(f,g) := \|f-g\|_2 := \int_{-\pi}^{\pi} |f(x)-g(x)|^2 dx,$$

in dem die trigonometrischen Polynome

$$\left\{ p_N(x) = \sum_{k=-N}^{N} c_k e^{ikx} : c_k \in \mathbb{C},\ N \in \mathbb{N} \right\}$$

dicht liegen.

Zum konkreten Verständnis des Raumes $L^2_{\text{per}}([-\pi,\pi])$ benötigt man die Lebesguesche Integrationstheorie, die später eingeführt wird. Wir begnügen uns an dieser Stelle damit, die Existenz des Raumes aus dem *Vervollständigungsprinzip* abzuleiten.

Satz 8.3.11 (Vervollständigungsprinzip). *Sei (X,d) ein metrischer Raum. Dann existiert ein vollständiger metrischer Raum $(\widehat{X},\widehat{d})$ und eine Isometrie, d.h. eine Abbildung $\Phi : X \to \widehat{X}$ mit $\widehat{d}\big(\Phi(x),\Phi(y)\big) = d(x,y)$ derart, dass $\Phi(X)$ dicht in $(\widehat{X},\widehat{d})$ liegt. Die Isometrie Φ wird auch Einbettung genannt.*

Beweis. Seien $(x_n)_{n\in\mathbb{N}}$ und $(y_n)_{n\in\mathbb{N}}$ Cauchy-Folgen in (X,d). Dann definiert die Relation

$$(x_n)_{n\in\mathbb{N}} \sim (y_n)_{n\in\mathbb{N}} \quad\Longleftrightarrow\quad \lim_{n\to\infty} d(x_n,y_n) = 0$$

eine Äquivalenzrelation in (X,d). Wir erhalten nun zu einer Cauchy-Folge $(x_n)_{n\in\mathbb{N}}$ aus (X,d) die Äquivalenzklassen

$$\widehat{x} = \widehat{x}\big((x_n)_{n\in\mathbb{N}}\big) = \big\{(y_n)_{n\in\mathbb{N}} : (y_n)_{n\in\mathbb{N}} \sim (x_n)_{n\in\mathbb{N}}\big\}$$

und $(x_n)_{n\in\mathbb{N}}$ heißt Repräsentant von $\widehat{x}$.

Den Raum aller Äquivalenzklassen bezeichnen wir mit $\widehat{X}$

$$\widehat{X} = \big\{\widehat{x} = \widehat{x}\big((x_n)_{n\in\mathbb{N}}\big) : (x_n)_{n\in\mathbb{N}} \text{ ist Cauchy-Folge in } (X,d)\big\}.$$

Hierauf definieren wir nun die Funktion $\widehat{d} : \widehat{X}\times\widehat{X} \to \mathbb{R}^+$ vermittels

$$\widehat{d}(\widehat{x},\widehat{y}) = \lim_{n=\infty} d(x_n,y_n).$$

Dann lässt sich leicht zeigen, dass $\widehat{d}(\widehat{x}, \widehat{y})$ unabhängig von der Wahl der Repräsentanten $(x_n)_{n\in\mathbb{N}}$ bzw. $(y_n)_{n\in\mathbb{N}}$ ist und dass $\widehat{d}$ eine Metrik auf $\widehat{X}$ definiert. Außerdem ist der metrische Raum $(\widehat{X}, \widehat{d})$ vollständig: Sei $(\widehat{x}_k)_{k\in\mathbb{N}}$ mit

$$\widehat{x}_k = \widehat{x}_k\big((x_{k,n})_{n\in\mathbb{N}}\big)$$

eine Cauchy-Folge aus $(\widehat{X}, \widehat{d})$. Setzen wir

$$\widehat{x} = \widehat{x}\big((x_{n,n})_{n\in\mathbb{N}}\big) \in \widehat{X}$$

so folgt

$$\widehat{d}(\widehat{x}_k, \widehat{x}) \to 0.$$

Folglich ist $(\widehat{X}, \widehat{d})$ vollständig.

Wir definieren nun $\Phi : X \to \widehat{X}$ durch $\Phi(x) = (x_n)_{n\in\mathbb{N}}$ mit $x_n = x$ für alle $n \in \mathbb{N}$, d.h. $\Phi(x)$ ist eine stationäre Folge. Dann ist

$$\widehat{d}\big(\Phi(x), \Phi(y)\big) = d(x, y).$$

Zuletzt müssen wir noch zeigen, dass $\Phi(X)$ dicht in $\widehat{X}$ liegt. Sei

$$\widehat{x} = \widehat{x}\big((x_k)_{n\in\mathbb{N}}\big) \in \widehat{X}$$

beliebig gewählt. Da $(x_k)_{n\in\mathbb{N}}$ eine Cauchy-Folge in (X, d) ist, existiert zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$ mit

$$d(x_n, x_m) < \varepsilon$$

für alle $n \geq m > N$. Für $\Phi(x_m) \in \widehat{X}$ gilt nun

$$\widehat{d}\big(\widehat{x}, \Phi(x_m)\big) < \varepsilon,$$

und somit auch dieser letzte Schritt gezeigt. □

Bemerkung 8.3.12. *Die Einbettung Φ ist bis auf Isometrien eindeutig. Das Vervollständigungsprinzip kann auch verwendet werden, um $\mathbb{R}$ als Vervollständigung von $\mathbb{Q}$ zu definieren. Obwohl das Prinzip sehr allgemein ist, ist es leider sehr unanschaulich.*

Satz 8.3.13 (Parseval). *Seien $f, g \in L^2_{\text{per}}([-\pi, \pi])$ mit den Fourier-Reihen*

$$f \sim \sum_{k\in\mathbb{Z}} f_k e^{ikx}, \qquad g \sim \sum_{k\in\mathbb{Z}} g_k e^{ikx},$$

dann gilt

1. *die Parsevalsche Identität*

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\overline{g(x)}dx = \sum_{k=-\infty}^{\infty} f_k\overline{g_k},$$

2. *die Normäquivalenz*
$$\frac{1}{2\pi}\int_{-\pi}^{\pi}|f(x)|^2dx = \sum_{k=-\infty}^{\infty}|f_k|^2,$$

3. *und die Fourier-Reihe konvergiert in der $L^2_{\mathrm{per}}([-\pi,\pi])$-Norm*
$$\lim_{N\to\infty}\left\|f-\sum_{k=-N}^{N}f_ke^{ikx}\right\|_2 = 0.$$

Beweis. Die Menge aller trigonometrischen Polynome liegt dicht in $L^2_{\mathrm{per}}([-\pi,\pi])$. Daher ist die erste Aussage nur für beliebige trigonometrische Polynome zu zeigen. Seien daher
$$f_N(x) = \sum_{k=-N}^{N}f_ke^{ikx}, \qquad g_N(x) = \sum_{k=-N}^{N}g_ke^{ikx}.$$
Dann gilt
$$\begin{aligned}\frac{1}{2\pi}\int_{-\pi}^{\pi}f_N(x)\overline{g_N(x)}dx &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\sum_{k=-N}^{N}f_ke^{ikx}\cdot\sum_{l=-N}^{N}g_le^{ilx}dx\\ &= \sum_{k=-N}^{N}\sum_{l=-N}^{N}f_k\overline{g_l}\frac{1}{2\pi}\int_{-\pi}^{\pi}e^{ikx}\cdot e^{-ilx}dx\\ &= \sum_{k=-N}^{N}f_k\overline{g_k}.\end{aligned}$$
Für $N\to\infty$ folgt die Behauptung.

Die zweite Aussage folgt direkt aus der ersten $g=f$. Die dritte Aussage folgt aus Satz 8.3.6 aufgrund der Dichtheit der trigonometrischen Polynome im $L^2_{\mathrm{per}}([-\pi,\pi])$. □

Bemerkung 8.3.14. *Der Raum $L^2_{\mathrm{per}}([-\pi,\pi])$ ist ein vollständiger normierter Raum mit*
$$\|f\|_2 = \sqrt{\langle f,f\rangle} = \sqrt{\int_{-\pi}^{\pi}|f(x)|^2dx}$$
mit $\langle\cdot,\cdot\rangle$ aus Bemerkung 8.3.2. Insbesondere gilt nun für alle $f,g,h\in L^2_{\mathrm{per}}([-\pi,\pi])$:

1. $\langle f+g,h\rangle = \langle f,h\rangle + \langle g,h\rangle$,
2. $\langle \alpha f,g\rangle = \alpha\langle f,g\rangle$ *für alle* $\alpha\in\mathbb{C}$,
3. $\langle f,g\rangle = \overline{\langle g,f\rangle}$,

4. $\langle f, f\rangle \geq 0$, *wobei* $\langle f, f\rangle = 0$ *genau für* $f \equiv 0$.

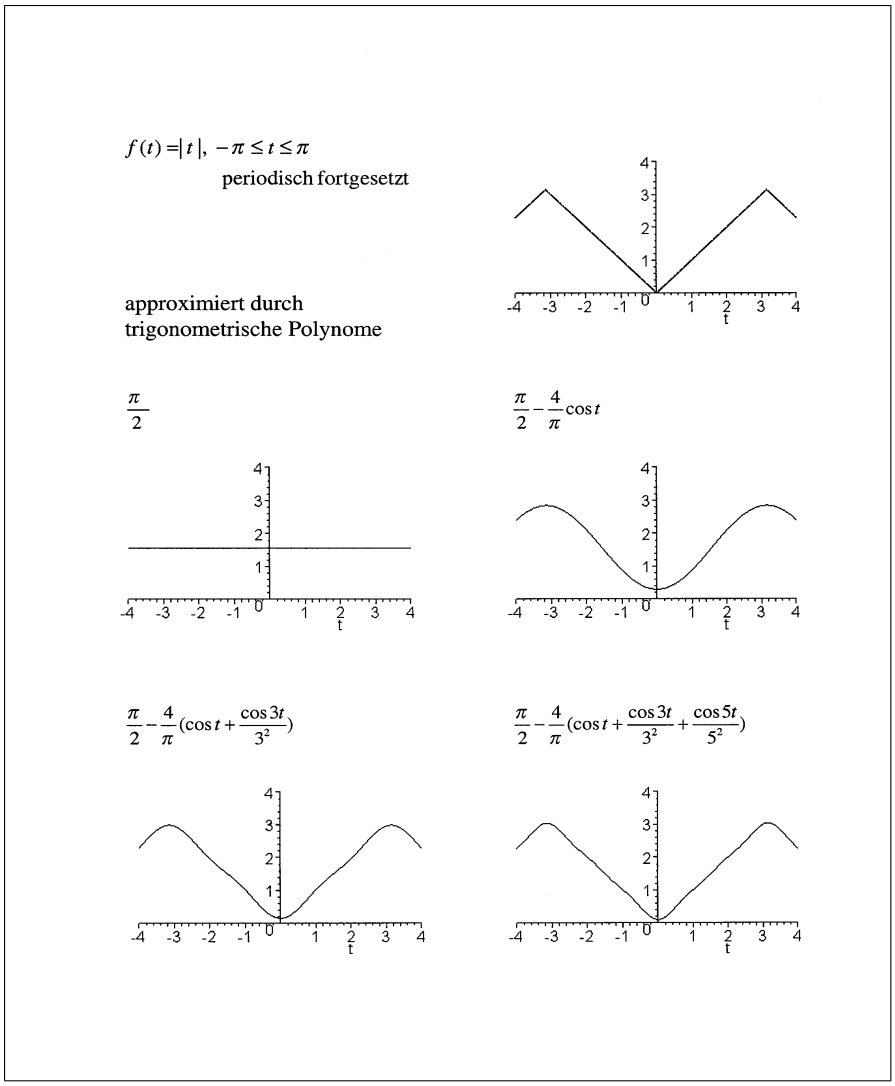
Dies bedeutet, $\langle \cdot, \cdot \rangle$ *definiert ein Skalarprodukt im* $L^2_{\text{per}}([-\pi, \pi])$.

Einen vollständigen normierten Raum nennt man **Banachraum**. *Wird die Norm durch ein Skalarprodukt induziert, so spricht man von einem* **Hilbertraum**. *Der Raum* $L^2_{\text{per}}([-\pi, \pi])$ *ist demanch ein (sehr wichtiger) Vetreter eines Hilbertraumes.*

Man kann auch zeigen, dass alle 2π-periodischen Riemann-integrierbaren dicht in $L^2_{\text{per}}([-\pi, \pi])$ liegt. Die Fourier-Reihen und die damit verbundenen Fragestellungen haben die moderne Mathematik wesentlich geprägt.

Fourier-Reihen, Fourier-Transformationen, und Fourier-Analysis spielen in der E-Technik, in der Maschinendynamik, in der Informations- und Signalverarbeitung eine zentrale Rolle.

Beispiel 8.3.15.



8.4 Gleichgradige Stetigkeit

Definition 8.4.1. *Gegeben sei eine Folge von Funktionen $(f_n)_{n\in\mathbb{N}}$ mit $f_n : D \to \mathbb{R}$ bzw. $f_n : D \to \mathbb{C}$ für $n \in \mathbb{N}$. Die Folge heißt* **punktweise beschränkt**, *falls zu jedem $x \in D$ ein $R = R(x) > 0$ existiert mit $|f_n(x)| < R$ für alle $n \in \mathbb{N}$. Die Folge heißt* **gleichmäßig beschränkt**, *falls ein $R > 0$ existiert, so dass $|f_n(x)| < R$ für alle $x \in D$ und $n \in \mathbb{N}$.*

Beispiel 8.4.2. *Seien $f_n : \mathbb{R} \to \mathbb{R}$ definiert durch*

$$f_n(x) = \frac{x^2}{x^2 + (1 - nx)^2}.$$

Es gilt $f_n(x) < 1$ für alle $x \in \mathbb{R}$, also ist die Folge $(f_n)_{n\in\mathbb{N}}$ gleichmäßig beschränkt auf $\mathbb{R}$.

Weiter ist $f_n \in C(\mathbb{R})$ für jedes $n \in \mathbb{N}$ und es gilt $f_n(x) \to 0$ für $n \to \infty$. Aber für $x_n := 1/n$ gilt

$$\lim_{n\to\infty} f_n(x_n) = \frac{1/n^2}{1/n^2} = 1,$$

d.h. die Folge $(f_n)_{n\in\mathbb{N}}$ konvergiert nicht gleichmäßig auf $\mathbb{R}$. Es gibt, wie man aus obiger Überlegung ersieht, auch keine gleichmäßig konvergente Teilfolge.

Definition 8.4.3. *Es sei (X, d) ein metrischer Raum und $D \subseteq X$. Eine Folge $(f_n)_{n\in\mathbb{N}}$ von Funktionen $f_n : D \to \mathbb{R}$ bzw. $f_n : D \to \mathbb{C}$ heißt* **gleichgradig stetig**, *falls zu jedem $\varepsilon > 0$ ein $\delta > 0$ existiert, mit*

$$|f_n(x) - f_n(y)| < \varepsilon$$

für alle $n \in \mathbb{N}$ und $x, y \in D$ mit $d(x, y) < \delta$.

Bemerkung 8.4.4. *Sei die Folge $(f_n)_{n\in\mathbb{N}}$ gleichgradig stetig, dann ist jede Funktion $f_n : D \to \mathbb{R}$ bzw $f_n : D \to \mathbb{C}$ gleichmäßig stetig. Man beachte aber, dass die Umkehrung im allgemeinen nicht gilt.*

Satz 8.4.5 (Satz von Arzela-Ascoli, Teil 1). *Sei (X, d) ein metrischer Raum und $K \subseteq X$ kompakt. Seien $f_n : K \to \mathbb{R}$ bzw. $f_n : K \to \mathbb{C}$ auf K stetig und $(f_n)_{n\in\mathbb{N}}$ eine gleichmäßig konvergente Folge von Funktionen. Dann ist die Folge $(f_n)_{n\in\mathbb{N}}$ auch gleichgradig stetig.*

Beweis. Sei $\varepsilon > 0$ beliebig. Dann existiert nach dem Cauchy-Kriterium ein $N \in \mathbb{N}$ derart, dass für alle $n > N$ gilt

$$\sup_{x\in K} |f_N(x) - f_n(x)| < \varepsilon/3.$$

Da f_N eine stetige Funktion auf der kompakten Menge K ist, ist f_N gleichmäßig stetig, d.h. es existiert ein $\delta = \delta(N) > 0$ mit

$$|f_N(x) - f_N(y)| < \varepsilon/3$$

für alle $d(x,y) < \delta$. Schließlich folgt für alle $n > N$ und $x, y \in K$ mit $d(x,y) < \delta$ dass

$$\begin{aligned} |f_n(x) - f_n(y)| &\leq |f_n(x) - f_N(x)| + |f_N(x) - f_N(y)| + |f_N(y) - f_n(y)| \\ &\leq \varepsilon/3 + \varepsilon/3 + \varepsilon/3 = \varepsilon. \end{aligned}$$

□

Lemma 8.4.6. *Sei $E = \{x_i : i \in \mathbb{N}\}$ eine abzählbare Menge und $f_n : E \to \mathbb{R}$ punktweise beschränkt. Dann besitzt die Folge $(f_n)_{n\in\mathbb{N}}$ eine Teilfolge $(f_{n_k})_{k\in\mathbb{N}}$, die für jedes $x \in E$ konvergiert, d.h. $f_{n_k}(x)$ konvergiert für jedes $x \in E$ und $k \to \infty$.*

Beweis. Wir wenden das *Diagonalfolgenprinzip* an:

Für $x_1 \in E$ ist $\big(f_k(x_1)\big)_{k\in\mathbb{N}}$ beschränkt. Aufgrund des Satzes von Weierstraß 4.3.14 existiert eine konvergente Teilfolge, die wir mit $\big(f_{1,k}\big)_{k\in\mathbb{N}}$ bezeichen wollen, mit $f_{1,k}(x_1) \to y_1$ für $k \to \infty$.

Da $\big(f_{1,k}(x_2)\big)_{k\in\mathbb{N}}$ ebenfalls beschränkt ist, existiert wiederum eine konvergente Teilfolge, die wir mit mit $(f_{2,k})_{k\in\mathbb{N}}$ bezeichnen, mit $f_{2,k}(x_2) \to y_2$ und insbesondere auch $f_{2,k}(x_1) \to y_1$ für $k \to \infty$.

So fortfahrend erhalten wir Folgen $S_n = (f_{n,k})_{k\in\mathbb{N}}$, die wir im folgenden Schema darstellen:

$$\begin{array}{cccccc} S_1: & f_{1,1} & f_{1,2} & f_{1,3} & f_{1,4} & \cdots \\ S_2: & f_{2,1} & f_{2,2} & f_{2,3} & f_{2,4} & \cdots \\ S_3: & f_{3,1} & f_{3,2} & f_{3,3} & f_{3,4} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{array}$$

Für jedes $n \in \mathbb{N}$ gilt erstens, dass S_n Teilfolge von S_{n-1} ist, und zweitens, dass für alle $1 \leq l \leq n$ gilt $f_{n,k}(x_l) \to y_l$ für $k \to \infty$. Wir setzen $g_k := f_{k,k}$. Dann ist $(g_k)_{k\in\mathbb{N}}$ eine Teilfolge von $(f_n)_{n\in\mathbb{N}}$ und für ein beliebiges $x_n \in E$ gilt

$$\lim_{k\to\infty} g_k(x_n) = \lim_{k\to\infty} f_{k,k}(x_n) = \lim_{k\to\infty} f_{n,k}(x_n) = y_n.$$

□

Satz 8.4.7 (Satz von Arzela-Ascoli, Teil 2). *Sei K kompakt und $(f_n)_{n\in\mathbb{N}}$ eine gleichgradig stetige Folge von Funktionen auf K. Ferner sei für alle $f_n : K \to \mathbb{R}$ bzw $f_n : K \to \mathbb{C}$ stetig und punktweise beschränkt. Dann gelten die folgenden Aussagen:*

1. *$(f_n)_{n\in\mathbb{N}}$ ist gleichmäßig beschränkt auf K.*
2. *Es existiert eine konvergente Teilfolge $(f_{n_k})_{k\in\mathbb{N}}$, welche auf K gleichmäßig konvergiert.*

Beweis. 1. Sei $\varepsilon > 0$ beliebig aber fest. Da $(f_n)_{n\in\mathbb{N}}$ gleichgradig stetig, existiert ein $\delta > 0$ mit

$$|f_n(x) - f_n(y)| < \varepsilon$$

für alle $x, y \in K$ mit $d(x, y) < \delta$ und für alle $n \in \mathbb{N}$.

Da ferner K kompakt ist, existiert zu diesem $\delta > 0$ eine endliche Menge von Punkten $\{x_1, x_2, \ldots, x_k\}$ derart, dass

$$K \subseteq \bigcup_{l=1}^{k} U_\delta(x_l).$$

Aufgrund der punktweisen Beschränktheit von $(f_n)_{n\in\mathbb{N}}$ existieren Zahlen $M_l > 0$, so dass $|f_n(x_l)| < M_l$ für alle $n \in \mathbb{N}$. Setzen wir

$$M := \max\{M_1, M_2, \ldots, M_k\} < \infty,$$

dann gilt

$$|f_n(x_l)| < M$$

für alle $n \in \mathbb{N}$ und $1 \leq l \leq k$.

Schließlich gibt es nun zu beliebigem $y \in K$ einen Punkt x_l mit $d(x_l, y) < \delta$ und daher gilt

$$|f_n(y)| \leq |f_n(y) - f_n(x_l)| + |f_n(x_l)| < \varepsilon + M.$$

2. Wir zeigen zuerst, dass aufgrund der Kompakheit von K eine abzählbare Menge $E = \{x_n : n \in \mathbb{N}\}$ existiert, die dicht in K liegt. Denn zu $j \in \mathbb{N}$ und $r_j = 2^{-j}$ existieren jeweils endlich viele Punkte $x_{j,k}$, $1 \leq k \leq n_j$, derart dass

$$K \subseteq \bigcup_{l=1}^{n_j} U_{r_j}(x_{j,k}).$$

Dann ist zum Beispiel die Menge

$$E = \{x_{j,k} : 1 \leq k \leq n_j,\ j \in \mathbb{N}\}$$

eine in K dichte Teilmenge.

Nach Lemma 8.4.6 besitzt die Folge $(f_n)_{n\in\mathbb{N}}$ eine Teilfolge $(f_{n_k})_{k\in\mathbb{N}}$, so dass für jedes $x \in E$ die Folge $f_{n_k}(x)$ konvergiert für $k \to \infty$. Wir schreiben im weiteren vereinfachend $g_k = f_{n_k}$ und beweisen nun, dass $(g_k)_{k\in\mathbb{N}}$ gleichmäßig auf ganz K konvergiert.

Sei $\varepsilon > 0$ beliebig, dann existiert wie in ersten Teil des Beweises ein $\delta > 0$ mit

$$|f_n(x) - f_n(y)| < \varepsilon$$

für alle $x, y \in K$ mit $d(x, y) < \delta$ und für alle $n \in \mathbb{N}$.

Da E dicht in K ist und K kompakt ist, gibt es zu diesem $\delta > 0$ endlich viele Punkte $y_l \in E$ mit

$$K \subseteq \bigcup_{l=1}^{k} U_\delta(y_l).$$

Da $\big(g_k(y)\big)_{k\in\mathbb{N}}$ für jedes $y \in E$ konvergiert, gibt es ein $N \in \mathbb{N}$, so dass

$$|g_m(y_l) - g_n(y_l)| < \varepsilon$$

für alle $m, n > N$ und für alle $1 \leq l \leq k$.

Schließlich gibt es nun zu beliebigem $y \in K$ einen dieser Punkte y_l mit $d(y_l, y) < \delta$ und daher

$$|g_n(y) - g_n(y_l)| < \varepsilon$$

für alle $n \in \mathbb{N}$. Somit gilt für alle $y \in K$

$$|g_m(y) - g_n(y)| \leq |g_m(y) - g_m(y_l)| + |g_m(y_l) - g_n(y_l)| + |g_n(y_l) - g_n(y)| \leq 3\varepsilon.$$

Damit ist der Beweis vollständig.

□

8.5 Der Approximationssatz von Weierstraß

Satz 8.5.1 (Approximationssatz von Weierstraß). *Sei $f : [a, b] \to \mathbb{R}$ bzw. $f : [a, b] \to \mathbb{C}$ stetig. Dann existiert eine Folge von Polynomen p_n, die auf $[a, b]$ gleichmäßig gegen f konvergiert. Falls f reellwertig ist, können auch die p_n reell gewählt werden.*

Beweis. Ohne Einschränkung der Allgemeinheit sei $[a, b] = [0, 1]$. Wir können annehmen, dass $f(0) = f(1) = 0$. Ist nämlich der Satz in diesem Fall bewiesen, so betrachten wir die Funktion

$$g(x) := f(x) - f(0) - x\big(f(1) - f(0)\big), \qquad 0 \leq x \leq 1.$$

Hierfür gilt $g(0) = 0 = g(1)$ und wenn g sich als Grenzwert einer gleichmäßig konvergenten Folge von Polynomen darstellen lässt, so gilt offensichtlich dasselbe auch für f, da $f - g$ ein Polynom ist.

Wir definieren ferner $f(x) = 0$ außerhalb $[0, 1]$, d.h. $f : [a, b] \to \mathbb{R}$ bzw. $f : [a, b] \to \mathbb{C}$ mit

$$f(x) = \begin{cases} f(x), & x \in [0, 1] \\ 0, & \text{sonst.} \end{cases}$$

Wir definieren für jedes $n \in \mathbb{N}$

$$q_n(x) = c_n(1 - x^2)^n$$

mit

$$c_n = \left(\int_{-1}^{1} (1 - x^2)^n dx \right)^{-1},$$

d.h. es gilt

$$\int_{-1}^{1} q_n(x)dx = 1.$$

Mit Hilfe der Bernoullischen Ungleichung

$$(1-x^2)^n \geq 1 - nx^2, \qquad |x| \leq 1,$$

können wir c_n abschätzen

$$\begin{aligned}\frac{1}{c_n} &= \int_{-1}^{1}(1-x^2)^n dx = 2\int_0^1 (1-x^2)^n dx \geq 2\int_0^{1/\sqrt{n}} (1-x^2)^n dx \\ &\geq 2\int_0^{1/\sqrt{n}} (1-nx^2)dx = \frac{2}{\sqrt{n}} - \frac{2n}{3(\sqrt{n})^3} = \frac{4}{3\sqrt{n}} \geq \frac{1}{\sqrt{n}}\end{aligned}$$

dies bedeutet

$$c_n \leq \sqrt{n} \tag{8.8}$$

Für $0 < \delta \leq |x| \leq 1$ gilt somit

$$q_n(x) \leq c_n(1-\delta^2)^n \leq \sqrt{n}(1-\delta^2)^n,$$

dies bedeutet, q_n konvergiert gleichmäßig gegen 0 auf $\{x : \delta \leq |x| \leq 1\}$.

Wir betrachten nun

$$p_n(x) = \int_{\mathbb{R}} f(x+t)q_n(t)dt = \int_0^1 f(t)q_n(t-x)dt.$$

Offensichtlich ist p_n ein Polynom vom Grade $2n$, insbesondere ist $p_n(x)$ reell, falls f reellwertig ist.

Wir setzen $M := \sup_{x\in[0,1]} |f(x)| < \infty$ und sei $\varepsilon > 0$ beliebig. Da $[0,1] \subseteq \mathbb{R}$ kompakt ist, ist $f : [0,1] \to \mathbb{R}$ gleichmäßig stetig. Daher existiert ein $\delta > 0$ mit

$$|f(x) - f(y)| < \varepsilon/2$$

für alle $|x-y| < \delta$. Wegen $q_n(x) \geq 0$ auf $[0,1]$ folgt für jedes $x \in [0,1]$ die Abschätzung

$$\begin{aligned}|p_n(x) - f(x)| &= \left| \int_{-1}^{1} \big(f(x+t) - f(x)\big) q_n(t)dt \right| \\ &\leq \int_{-1}^{1} |f(x+t) - f(x)| q_n(t)dt \\ &\leq 2M \int_{-1}^{-\delta} q_n(t)dt + \frac{\varepsilon}{2}\int_{-\delta}^{\delta} q_n(t)dt + 2M\int_{\delta}^{1} q_n(t)dt \\ &\leq 4M\sqrt{n}(1-\delta^2)^n + \frac{\varepsilon}{2} \leq \varepsilon\end{aligned}$$

für hinreichend großes n. □

Kapitel 9

Differentiation von Funktionen mehrerer Veränderlicher

9.1 (Partielle) Ableitungen und Richtungsableitungen

Sei $D \subseteq \mathbb{R}^n$ offen und $f : D \to \mathbb{R}$. Mit $\mathbf{e}_i$ bezeichnen wir den Einheitsvektor bezüglich der $i - ten$ Koordinate des $\mathbb{R}^n$. Für $\mathbf{x} = (x_1, \ldots, x_n)^T \in D$ schreiben wir kurz $f(\mathbf{x}) = f\big(x_1, \ldots, x_n)^T\big) = f(x_1, \ldots, x_n)$. Wir betrachten für ein $\mathbf{x} \in D$ die Funktion $g : (-\varepsilon, \varepsilon) \to \mathbb{R}$ definiert durch

$$g(t) := f(x_1, \ldots, x_{i-1}, x_i + t, x_{i+1}, \ldots, x_n).$$

Es gilt

$$g'(t) = \frac{dg}{dt}(t) = \lim_{h \to 0} \frac{g(t+h) - g(t)}{h}.$$

Angenommen, g ist in $t = 0$ differenzierbar, d.h. $g'(0)$ existiert, dann definieren wir

$$g'(0) = \lim_{h \to 0} \frac{f(x_1, \ldots, x_{i-1}, x_i + h, x_{i+1}, \ldots, x_n) - f(x_1, \ldots, x_n)}{h}$$

als die partielle Ableitung von f nach x_i, kurz $\frac{\partial f}{\partial x_i}(\mathbf{x})$ oder auch $f_{x_i}(\mathbf{x})$.

Wir fassen diese Aussage in folgender Definition zusammen.

Definition 9.1.1. *Sei $f : D \to \mathbb{R}^m$ und $\mathbf{x} \in D$, dann heißt f an der Stelle $\mathbf{x}$* **partiell differenzierbar** *nach x_i, falls*

$$\frac{\partial f}{\partial x_i}(\mathbf{x}) := \lim_{h \to 0} \frac{f(\mathbf{x} + h\mathbf{e}_i) - f(\mathbf{x})}{h} = f_{x_i}(\mathbf{x}).$$

existiert.

Beispiel 9.1.2.

1. *Sei $f : \mathbb{R}^2 \to \mathbb{R}$ gegeben durch $f(x_1, x_2) = x_1 e^{x_2}$, dann ist*

$$f_{x_1}(x_1, x_2) = e^{x_2}, \qquad f_{x_2}(x_1, x_2) = x_1 e^{x_2}.$$

2. *Für $f : \mathbb{R}^n \to \mathbb{R}$ definiert durch*

$$f(\mathbf{x}) = \|\mathbf{x}\| = \sqrt{\sum_{k=1}^{n} x_k^2}$$

gilt im Falle $\mathbf{x} \neq \mathbf{0}$

$$f_{x_i}(\mathbf{x}) = \frac{2x_i}{2\sqrt{\sum_{k=1}^{n} x_k^2}} = \frac{x_i}{\|\mathbf{x}\|}.$$

Die Funktion f ist stetig auf $\mathbb{R}^n$, aber nicht partiell differenzierbar in $\mathbf{x}_0 = \mathbf{0}$.

Aus der partiellen Differenzierbarkeit folgt i.a. nicht die Stetigkeit.

Beispiel 9.1.3. *Wir betrachten die Funktion $f : \mathbb{R}^2 \to \mathbb{R}$ mit*

$$f(x, y) = \begin{cases} \frac{xy}{x^2+y^2}, & (x, y)^T \in \mathbb{R}^2 \setminus \{\mathbf{0}\}, \\ 0, & (x, y)^T = \mathbf{0}. \end{cases}$$

Diese Funktion ist in $\mathbf{0}$ nicht stetig, denn es ist

$$f(t, 0) = 0 = f(0, t)$$

für alle $t \in \mathbb{R}$, aber auch

$$f(t, t) = 1/2$$

für alle $t \in \mathbb{R} \setminus \{\mathbf{0}\}$.

Andererseits existieren jedoch die partiellen Ableitung im Nullpunkt

$$f_x(0, 0) = \lim_{h \to 0} \frac{f(h, 0) - f(0, 0)}{h} = 0,$$
$$f_y(0, 0) = \lim_{h \to 0} \frac{f(0, h) - f(0, 0)}{h} = 0.$$

Definition 9.1.4. *Sei $\mathbf{n} \in \mathbb{R}^n$ mit $\|\mathbf{n}\| = 1$ ein Richtungsvektor und $\mathbf{x} \in D$. Dann heißt die Funktion $f : D \to \mathbb{R}^m$ an der Stelle $\mathbf{x}$* **(Gateaux-) differenzierbar in Richtung von $\mathbf{n}$** *oder einfach nur* **differenzierbar in Richtung von $\mathbf{n}$**, *falls*

$$\frac{d}{dt} f(\mathbf{x} + t\mathbf{n})\Big|_{t=0} = \lim_{h \to 0} \frac{f(\mathbf{x} + h\mathbf{n}) - f(\mathbf{x})}{h}$$

existiert. Die Funktion f heißt **Gateaux-differenzierbar in $\mathbf{x} \in D$**, *falls f an der Stelle $\mathbf{x} \in D$ in alle Richtungen $\mathbf{n} \in \mathbb{R}^n$ mit $\|\mathbf{n}\| = 1$ differenzierbar ist.*

Selbst diese Forderung ist nicht hinreichend für die Stetigkeit von f, wie das folgende Beispiel zeigt.

Beispiel 9.1.5. *Wir betrachten die Funktion $f : \mathbb{R}^2 \to \mathbb{R}$ mit*

$$f(x,y) = \begin{cases} \frac{xy^2}{x^2+y^4}, & (x,y)^T \in \mathbb{R}^2 \setminus \{\mathbf{0}\}, \\ 0, & (x,y)^T = \mathbf{0}. \end{cases}$$

Diese Funktion ist in $\mathbf{0}$ nicht stetig, denn es ist

$$f(t,0) = 0 = f(0,t)$$

für alle $t \in \mathbb{R}$, aber auch

$$f(t^2,t) = 1/2$$

für alle $t \in \mathbb{R} \setminus \{\mathbf{0}\}$.

Für eine beliebige Richtung $\mathbf{n} = (n_1, n_2)$ mit $\|\mathbf{n}\| = 1$ gilt

$$\frac{d}{dt} f(\mathbf{0} + t\mathbf{n})\Big|_{t=0} = \frac{d}{dt} \frac{tn_1 \cdot (tn_2)^2}{(tn_1)^2 + (tn_2)^4}\Big|_{t=0} = \frac{d}{dt} \frac{tn_1 n_2^2}{n_1^2 + t^2 n_2^4}\Big|_{t=0} = \begin{cases} 0, & \textit{falls } n_1 = 0, \\ \frac{n_2^2}{n_1}, & \textit{falls } n_1 \neq 0, \end{cases}$$

d.h. f ist im Nullpunkt in jede Richtung differenzierbar.

Definition 9.1.6. *Seien X und Y normierte Vektorräume mit den Normen $\|\cdot\|_X$ und $\|\cdot\|_Y$ über einen Körper $\mathbb{K}$.*

1. *Eine Abbildung $\mathbf{A} : X \to Y$ heißt* **linear**, *falls für alle $\alpha, \beta \in \mathbb{K}$ und $\mathbf{x}, \mathbf{y} \in X$ gilt*

$$\mathbf{A}(\alpha\mathbf{x} + \beta\mathbf{y}) = \alpha\mathbf{A}\mathbf{x} + \beta\mathbf{A}\mathbf{y}.$$

2. *Die Menge aller stetigen, linearen Abbildungen $\mathbf{A} : X \to Y$ bezeichnen wir $\mathcal{L}(X,Y)$.*

3. *Sei $\mathbf{A} \in \mathcal{L}(X,Y)$, dann definieren wir die* **Operatornorm**

$$\|\mathbf{A}\| := \sup_{\|\mathbf{x}\|_X = 1} \|\mathbf{A}\mathbf{x}\|_Y.$$

Proposition 9.1.7. *Sei $X = \mathbb{R}^n$, d.h. $\dim X = n < \infty$, und sei die Abbildung $\mathbf{A} : X \to Y$ linear. Dann gelten die folgenden Aussagen:*

1. *Es gilt $\|\mathbf{A}\| < \infty$.*

2. *Die Abbildung $\mathbf{A} : X \to Y$ ist gleichmäßig stetig und sogar Lipschitz-stetig.*

Beweis. 1. Sei $\mathbf{x} = (x_1, \ldots, x_n)^T \in \mathbb{R}^n$ mit $\|\mathbf{x}\|_X = 1$. Dann gilt $\mathbf{x} = \sum_{i=1}^n x_i \mathbf{e}_i$ mit $|x_i| \leq 1$ für alle $i = 1, \ldots, n$. Aus der Linearität von A folgt nun

$$\|\mathbf{A}\mathbf{x}\|_Y = \left\| \sum_{i=1}^n x_i \mathbf{A}\mathbf{e}_i \right\|_Y \leq \sum_{i=1}^n |x_i| \, \|\mathbf{A}\mathbf{e}_i\|_Y \leq \sum_{i=1}^n \|\mathbf{A}\mathbf{e}_i\|_Y < \infty.$$

2. Sei $\mathbf{x} \in X \setminus \{\mathbf{0}\}$, dann ist $\mathbf{n} = \mathbf{x}/\|\mathbf{x}\|_X \in X$ mit $\|\mathbf{n}\|_X = 1$. Aus $\|\mathbf{An}\|_Y \leq \|\mathbf{A}\|$ folgt $\|\mathbf{Ax}\|_Y/\|\mathbf{x}\|_X \leq \|\mathbf{A}\|$ und somit

$$\|\mathbf{Ax}\|_Y \leq \|\mathbf{A}\| \, \|\mathbf{x}\|_X.$$

□

Bemerkung 9.1.8. *Für die Operatornorm einer linearen Abbildung* $\mathbf{A} : X \to Y$ *gilt*

$$\|\mathbf{A}\| = \sup_{\mathbf{x} \in X \setminus \{\mathbf{0}\}} \frac{\|\mathbf{Ax}\|_Y}{\|\mathbf{x}\|_X}.$$

Proposition 9.1.9.

1. *$\mathcal{L}(X, Y)$ ist ein Vektorraum über den Körper $\mathbb{K}$. Die Operatornorm definiert eine Norm auf $\mathcal{L}(X, Y)$.*

2. *Sei $\mathbf{A} \in \mathcal{L}(X, Y)$ und $\mathbf{B} \in \mathcal{L}(Y, Z)$, dann ist $\mathbf{B} \cdot \mathbf{A} \in \mathcal{L}(X, Z)$ mit*

$$\|\mathbf{B} \cdot \mathbf{A}\| \leq \|\mathbf{B}\| \cdot \|\mathbf{A}\|.$$

Beweis.

1. Der Beweis verbleibt dem Leser als Übung.

2. Sei $\|\mathbf{x}\|_X = 1$, dann gilt

$$\|\mathbf{B} \cdot \mathbf{Ax}\|_Z \leq \|\mathbf{B}\| \, \|\mathbf{Ax}\|_Y \leq \|\mathbf{B}\| \, \|\mathbf{A}\|,$$

dass heißt

$$\|\mathbf{B} \cdot \mathbf{A}\| = \sup_{\|\mathbf{x}\|_X = 1} \|\mathbf{B} \cdot \mathbf{Ax}\|_Z \leq \|\mathbf{B}\| \, \|\mathbf{A}\|.$$

□

Definition 9.1.10. *Sei $D \subseteq X$ offen und $f : D \to Y$. Dann heißt f an der Stelle $\mathbf{x}_0 \in D$* **(Fréchet-) differenzierbar***, falls eine Abbildung $\mathbf{A} \in \mathcal{L}(X, Y)$ existiert, mit*

$$\lim_{\mathbf{x} \to \mathbf{x}_0} \frac{\|f(\mathbf{x}) - f(\mathbf{x}_0) - \mathbf{A}(\mathbf{x} - \mathbf{x}_0)\|_Y}{\|\mathbf{x} - \mathbf{x}_0\|_X} = 0.$$

Die lineare Abbildung $\mathbf{A}$ heißt **Fréchet-Ableitung** *an der Stelle $\mathbf{x}_0$, kurz $\mathbf{A} := f'(\mathbf{x}_0)$.*

Aus dieser Definition erhalten wir offensichtlich folgende Proposition.

Proposition 9.1.11. *Sei $D \subseteq \mathbb{R}^n$ offen und $f : D \to \mathbb{R}^m$. Die Funktion f ist an der Stelle $\mathbf{x}_0 \in D$ differenzierbar, falls $\mathbf{A} = f'(\mathbf{x}_0) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$, d.h. $\mathbf{A}$ ist eine $(m \times n)$-Matrix, und eine Funktion $r : U_\delta(\mathbf{0}) \to \mathbb{R}^m$ mit*

$$\lim_{\mathbf{x} \to \mathbf{0}} r(\mathbf{x}) = \mathbf{0}$$

existiert, so dass

$$f(\mathbf{x}) = f(\mathbf{x}_0) + \mathbf{A}(\mathbf{x} - \mathbf{x}_0) + r(\mathbf{x} - \mathbf{x}_0)\|\mathbf{x} - \mathbf{x}_0\|$$

für alle $\mathbf{x} \in U_\delta(\mathbf{x}_0)$.

Proposition 9.1.12. *Unter den obigen Voraussetzungen sei f an der Stelle $\mathbf{x_0} \in D$ differenzierbar. Dann gelten die folgenden Aussagen:*

1. *Die Ableitung $A = f'(\mathbf{x}_0)$ ist eindeutig.*
2. *Die Funktion f ist stetig an der Stelle $\mathbf{x}_0$.*

Beweis. 1. Seien $\mathbf{A}_1, \mathbf{A}_2 \in \mathbb{R}^{m \times n}$. Setzen wir $\mathbf{B} := \mathbf{A}_1 - \mathbf{A}_2 \in \mathbb{R}^{m \times n}$, dann gilt für beliebiges $\mathbf{x} \in D$ mit $\mathbf{h} := \mathbf{x} - \mathbf{x}_0$ dass

$$\begin{aligned}
\|\mathbf{Bh}\| &= \|f(\mathbf{x}) - f(\mathbf{x}_0) + \mathbf{A}_1\mathbf{h} - \mathbf{A}_2\mathbf{h} - f(\mathbf{x}) + f(\mathbf{x}_0)\| \\
&\leq \|f(\mathbf{x}) - f(\mathbf{x}_0) - \mathbf{A}_1\mathbf{h}\| + \|f(\mathbf{x}) - f(\mathbf{x}_0) - \mathbf{A}_2\mathbf{h}\| \\
&= \|r_1(\mathbf{h})\| \cdot \|\mathbf{h}\| + \|r_2(\mathbf{h})\| \cdot \|\mathbf{h}\|
\end{aligned}$$

und daher

$$\frac{\|\mathbf{Bh}\|}{\|\mathbf{h}\|} = \|r_1(\mathbf{h})\| + \|r_2(\mathbf{h})\| \to 0$$

für $\mathbf{h} \to \mathbf{0}$. Wählen wir nun $\mathbf{h} \neq \mathbf{0}$ beliebig aber fest, so gilt

$$0 = \lim_{t \to 0} \frac{\|\mathbf{B}(t\mathbf{h})\|}{\|t\mathbf{h}\|} = \lim_{t \to 0} \frac{|t| \cdot \|\mathbf{Bh}\|}{|t| \cdot \|\mathbf{h}\|} = \frac{\|\mathbf{Bh}\|}{\|\mathbf{h}\|},$$

d.h. $\|\mathbf{Bh}\| = 0$ für alle $\mathbf{h} \in \mathbb{R}^n$. Somit ist $\mathbf{B} = \mathbf{0}$ die Nullmatrix und daher $\mathbf{A}_1 = \mathbf{A}_2$.

2. Sei $\mathbf{x} \in U_\delta(\mathbf{x}_0)$, dann gilt

$$f(\mathbf{x}) = f(\mathbf{x}_0) + f'(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) + r(\mathbf{x} - \mathbf{x}_0)\|\mathbf{x} - \mathbf{x}_0\| \to f(\mathbf{x}_0)$$

für $\mathbf{x} \to \mathbf{x}_0$. Dies bedeutet

$$\lim_{\mathbf{x} \to \mathbf{x}_0} f(\mathbf{x}) = f(\mathbf{x}_0),$$

d.h. f ist stetig an der Stelle $\mathbf{x}_0$.

□

Definition 9.1.13. *Gegeben sei die Funktion $f : D \subseteq R^n \to \mathbb{R}^m$, d.h. für $\mathbf{x} \in D$ ist $f(\mathbf{x}) = \big(f_1(\mathbf{x}), \ldots, f_m(\mathbf{x})\big)^T$ ein m-dimensionaler, reeller Vektor. Die Matrix*

$$\mathbf{J}_f(\mathbf{x}) = \left[\frac{\partial f_i}{\partial x_j}(\mathbf{x})\right]_{1\le i\le m,\, 1\le j\le n} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{x}) & \frac{\partial f_1}{\partial x_2}(\mathbf{x}) & \ldots & \frac{\partial f_1}{\partial x_n}(\mathbf{x}) \\ \frac{\partial f_2}{\partial x_1}(\mathbf{x}) & \frac{\partial f_2}{\partial x_2}(\mathbf{x}) & \ldots & \frac{\partial f_2}{\partial x_n}(\mathbf{x}) \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(\mathbf{x}) & \frac{\partial f_m}{\partial x_2}(\mathbf{x}) & \ldots & \frac{\partial f_m}{\partial x_n}(\mathbf{x}) \end{bmatrix} \in \mathbb{R}^{m\times n}$$

heißt **Funktionalmatrix** *oder* **Jakobi-Matrix**.

Ist $m = 1$, d.h. $f(x_1, \ldots, x_n) \in \mathbb{R}$, dann ist

$$\mathbf{J}_f(\mathbf{x}) = \big[f_{x_1}(\mathbf{x}), \ldots, f_{x_n}(\mathbf{x})\big]$$

ein Zeilenvektor. Wir nennen den Spaltenvektor

$$\nabla f(\mathbf{x}) := \mathbf{J}_f(\mathbf{x})^T \in \mathbb{R}^m$$

den **Gradienten** *von f an der Stelle $\mathbf{x}$.*

Satz 9.1.14. *Sei $f : D \subseteq R^n \to \mathbb{R}^m$ an der Stelle $\mathbf{x}_0 \in D$ differenzierbar, dann ist f an dieser Stelle partiell nach allen Variablen und in allen Richtungen $\mathbf{n} \in \mathbb{R}^n$ mit $\|\mathbf{n}\| = 1$ differenzierbar. Insbesondere gilt*

$$f'(\mathbf{x}_0) = \mathbf{J}_f(\mathbf{x}_0) = \left[\frac{\partial f_i}{\partial x_j}(\mathbf{x}_0)\right]_{1\le i\le m,\, 1\le j\le n}$$

und

$$\frac{d}{dt} f(\mathbf{x}_0 + t\mathbf{n})\Big|_{t=0} = f'(\mathbf{x}_0)\mathbf{n}.$$

Beweis. Für $1 \le j \le n$ setzen wir $\mathbf{h} := \mathbf{x} - \mathbf{x}_0 := h\mathbf{e}_j$, dann gilt mit (9.1.11)

$$\frac{f(\mathbf{x}_0 + h\mathbf{e}_j) - f(\mathbf{x}_0)}{h} = f'(\mathbf{x}_0)\cdot\mathbf{e}_j + r(h\mathbf{e}_j)\frac{|h|}{h} \to f'(\mathbf{x}_0)\cdot\mathbf{e}_j$$

für $h \to 0$. Dies bedeutet

$$\frac{\partial f_i}{\partial x_j}(\mathbf{x}_0) = \lim_{h\to 0}\frac{f_i(\mathbf{x}_0 + h\mathbf{e}_j) - f_i(\mathbf{x}_0)}{h} = [f'(\mathbf{x}_0)]_{i,j}.$$

Den Beweis bezüglich der Richtungsableitung führt man analog. □

Bemerkung 9.1.15. *Wir betrachten den Spezialfall $m = 1$. Es ist*

$$f'(\mathbf{x}_0) = \big(\nabla f(\mathbf{x}_0)\big)^T$$

und

$$\frac{d}{dt} f(\mathbf{x}_0 + t\mathbf{n})\Big|_{t=0} = \langle \nabla f(\mathbf{x}_0), \mathbf{n}\rangle.$$

für jede Richtung $\mathbf{n} \in \mathbb{R}^n$ *mit* $\|\mathbf{n}\| = 1$. *Insebsondere folgt somit aus der Cauchy-Schwartzschen Ungleichung*

$$\frac{d}{dt} f(\mathbf{x}_0 + t\mathbf{n})\Big|_{t=0} \leq \|\nabla f(\mathbf{x}_0)\|.$$

In dieser Ungleichung erhalten wir Gleichheit genau dann, wenn $\nabla f(\mathbf{x}_0)$ *und* $\mathbf{n}$ *parallel sind, d.h. der Gradient zeigt in die Richtung des* **steilsten Anstieges**.

Fassen wir für $\mathbf{x} \in \mathbb{R}^n$ *und* $y \in \mathbb{R}$ *den Vektor* $(\mathbf{x}, y)$ *als Punkt im* $\mathbb{R}^{n+1}$ *auf, dann definiert die Ebene*

$$E := \left\{ (\mathbf{x}, y)^T \in \mathbb{R}^{n+1} : y(\mathbf{x}) := f(\mathbf{x}_0) + f'(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) \right\}$$

die **Tangentialebene** *im Punkt* $\mathbf{x}_0$ *an den Graphen von* f. *Da diese Ebene die Dimension* n *besitzt, handelt es sich um eine Hyperebene. Speziell im Falle* $n = 2$ *erhalten wir mit* $\mathbf{x} = (x, y)$ *und* $\mathbf{x}_0 = (x_0, y_0)$ *die Tangentialebene*

$$E = \{\mathbf{x} = (x, y, z)^T \in \mathbb{R}^3 : z(x, y) = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0)\}.$$

Satz 9.1.16 (Mittelwertsatz). *Sei* $D \subseteq \mathbb{R}^n$ *offen und sei* $\mathbf{x}_0, \mathbf{x} \in D$ *derart, dass die Verbindungslinie beider Punkte*

$$G := \{t\mathbf{x} + (1 - t)\mathbf{x}_0 : t \in [0, 1]\}$$

in D *enthalten ist. Die Funktion* $\mathbf{f} : D \to \mathbb{R}^m$ *sei differenzierbar auf* G. *Dann existiert eine Zwischenstelle* $\boldsymbol{\xi} := \mathbf{x}_0 + t_0(\mathbf{x} - \mathbf{x}_0)$ *mit* $t_0 \in (0, 1)$, *so dass*

$$\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{x}_0)\| \leq \|\mathbf{f}'(\boldsymbol{\xi})(\mathbf{x} - \mathbf{x}_0)\|. \tag{9.1}$$

Beweis. Wir definieren die Funktion $g : [0, 1] \to \mathbb{R}$ durch

$$g(t) := \langle \mathbf{f}\big(\mathbf{x}_0 + t(\mathbf{x} - \mathbf{x}_0)\big), \mathbf{n} \rangle$$

mit einem beliebigen (normierten) Richtungsvektor $\mathbf{n} \in \mathbb{R}^m$. Die Ableitung dieser Funktion an der Stelle $t_0 \in [0, 1]$ ist gegeben durch

$$\begin{aligned}
g'(t_0) &= \lim_{t \to t_0} \frac{g(t) - g(t_0)}{t - t_0} \\
&= \lim_{t \to t_0} \frac{\langle \mathbf{f}\big(\mathbf{x}_0 + t(\mathbf{x} - \mathbf{x}_0)\big), \mathbf{n} \rangle - \langle \mathbf{f}\big(\mathbf{x}_0 + t_0(\mathbf{x} - \mathbf{x}_0)\big), \mathbf{n} \rangle}{t - t_0} \\
&= \langle \lim_{t \to t_0} \frac{\mathbf{f}\big(\mathbf{x}_0 + t(\mathbf{x} - \mathbf{x}_0)\big) - \mathbf{f}\big(\mathbf{x}_0 + t_0(\mathbf{x} - \mathbf{x}_0)\big)}{t - t_0}, \mathbf{n} \rangle \\
&= \langle \mathbf{f}'\big(\mathbf{x}_0 + t_0(\mathbf{x} - \mathbf{x}_0)\big)(\mathbf{x} - \mathbf{x}_0), \mathbf{n} \rangle.
\end{aligned}$$

Hieraus folgt, g ist in $[0, 1]$ differenzierbar und es gilt mit dem Mittelwertsatz 6.2.4, dass ein $t_0 \in (0, 1)$ existiert, mit

$$\langle \mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{x}_0), \mathbf{n} \rangle = g(1) - g(0) = g'(t_0) = \langle \mathbf{f}'(\boldsymbol{\xi})(\mathbf{x} - \mathbf{x}_0), \mathbf{n} \rangle.$$

Wählen wir nun speziell

$$\mathbf{n} := \frac{\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{x}_0)}{\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{x}_0)\|},$$

dann folgt

$$\frac{\langle \mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{x}_0), \mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{x}_0)\rangle}{\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{x}_0)\|} = \|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{x}_0)\| = \langle \mathbf{f}'(\boldsymbol{\xi})(\mathbf{x} - \mathbf{x}_0), \mathbf{n}\rangle \leq \|\mathbf{f}'(\boldsymbol{\xi})(\mathbf{x} - \mathbf{x}_0)\|.$$

□

Bemerkung 9.1.17. *Die Ungleichung* (9.1) *kann im allgemeinen nicht durch eine Gleichung wie im eindimensionalen Mittelwertsatz 6.2.4 ersetzt werden, wie folgendes Beispiel zeigt.*

Für $f : \mathbb{R} \to \mathbb{R}^2$ *definiert durch*

$$\mathbf{f}(t) = (\cos t, \sin t)^T.$$

gilt

$$\mathbf{f}(0) = \mathbf{f}(2\pi) = (1, 0)^T.$$

Aus

$$\mathbf{f}'(t) = (-\sin t, \cos t)$$

folgt $\|\mathbf{f}'(t)\|^2 = \sin^2 t + \cos^2 t = 1$ *für alle* $t \in \mathbb{R}$. *Damit erhalten wir*

$$0 = \|\mathbf{f}(2\pi) - \mathbf{f}(0)\| < \|\mathbf{f}'(\xi)(2\pi - 0)\| = 2\pi.$$

Satz 9.1.18 (Kettenregel). *Seien* $f : E \subseteq \mathbb{R}^n \to \mathbb{R}$ *und* $g : D \subseteq \mathbb{R}^m \to E$ *mit* D *und* E *offen. Seien* g *in* $\mathbf{x}_0 \in D$ *differenzierbar und* f *in* $g(\mathbf{x}_0)$ *differenzierbar, dann ist* $h = f \circ g : D \to \mathbb{R}$ *ebenfalls differenzierbar mit*

$$h'(\mathbf{x}_0) = f'\big(g(\mathbf{x}_0)\big)g'(\mathbf{x}_0)$$

bzw.

$$\frac{\partial}{\partial x_i} h(\mathbf{x}_0) = \sum_{k=1}^{n} \frac{\partial f}{\partial g_k}\big(g(\mathbf{x}_0)\big) \cdot \frac{\partial g_k}{\partial x_i}(\mathbf{x}_0).$$

Beweis. Es gilt für $\|\mathbf{x} - \mathbf{x}_0\| < \delta$ hinreichend klein

$$g(\mathbf{x}) = g(\mathbf{x}_0) + g'(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) + r_g(\mathbf{x} - \mathbf{x}_0)\|\mathbf{x} - \mathbf{x}_0\|$$

und für $\mathbf{y}_0 = g(\mathbf{x}_0) \in E$ und $\|\mathbf{y} - \mathbf{y}_0\| < \delta$

$$f(\mathbf{y}) = f(\mathbf{y}_0) + f'(\mathbf{y}_0)(\mathbf{y} - \mathbf{y}_0) + r_f(\mathbf{y} - \mathbf{y}_0)\|\mathbf{y} - \mathbf{y}_0\|$$

mit

$$\lim_{\mathbf{x} \to \mathbf{x}_0} r_g(\mathbf{x} - \mathbf{x}_0) = \mathbf{0}, \qquad \lim_{\mathbf{y} \to \mathbf{y}_0} r_f(\mathbf{y} - \mathbf{y}_0) = \mathbf{0}.$$

Wir setzen $\mathbf{y} = g(\mathbf{x})$ und erhalten

$$\begin{aligned} f\big(g(\mathbf{x})\big) &= f\big(g(\mathbf{x}_0)\big) + f'\big(g(\mathbf{x}_0)\big)\big(g(\mathbf{x}) - g(\mathbf{x}_0)\big) + r_f\big(g(\mathbf{x}) - g(\mathbf{x}_0)\big)\|g(\mathbf{x}) - g(\mathbf{x}_0)\| \\ &= f\big(g(\mathbf{x}_0)\big) + f'\big(g(\mathbf{x}_0)\big)g'(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) \\ &\qquad + \left\{ f'\big(g(\mathbf{x}_0)\big)r_g(\mathbf{x} - \mathbf{x}_0) + r_f\big(g(\mathbf{x}) - g(\mathbf{x}_0)\big)\frac{\|g(\mathbf{x}) - g(\mathbf{x}_0)\|}{\|\mathbf{x} - \mathbf{x}_0\|} \right\}\|\mathbf{x} - \mathbf{x}_0\| \\ &:= f\big(g(\mathbf{x}_0)\big) + f'\big(g(\mathbf{x}_0)\big)g'(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) + r(\mathbf{x} - \mathbf{x}_0)\|\mathbf{x} - \mathbf{x}_0\|. \end{aligned}$$

Hierbei gilt

$$\|g(\mathbf{x}) - g(\mathbf{x}_0)\| \le \|g'(\mathbf{x}_0)\|\,\|\mathbf{x} - \mathbf{x}_0\| + |r_g(\mathbf{x} - \mathbf{x}_0)|\,\|\mathbf{x} - \mathbf{x}_0\| \le c\|\mathbf{x} - \mathbf{x}_0\|,$$

und damit insbesondere auch

$$\lim_{\mathbf{x}\to\mathbf{x}_0} r_f\big(g(\mathbf{x}) - g(\mathbf{x}_0)\big) = \mathbf{0}.$$

Beachten wir schließlich noch

$$\big\|f'\big(g(\mathbf{x}_0)\big)r_g(\mathbf{x} - \mathbf{x}_0)\big\| \le \big\|f'\big(g(\mathbf{x}_0)\big)\big\|\,\|r_g(\mathbf{x} - \mathbf{x}_0)\| \overset{\mathbf{x}\to\mathbf{x}_0}{\longrightarrow} 0,$$

so folgt

$$\lim_{\mathbf{x}\to\mathbf{x}_0} r(\mathbf{x} - \mathbf{x}_0) = \mathbf{0}.$$

□

Wir kommen nun zu einem wichtigen hinreichenden Kriterium zur Differenzierbarkeit.

Satz 9.1.19. *Es sei $D \subseteq \mathbb{R}^n$ offen und $f : D \to \mathbb{R}$. Falls zu einem $\mathbf{x}_0 \in D$ eine Umgebung $U_\delta(\mathbf{x}_0) \subseteq D$ existiert, so dass alle partiellen Ableitungen $\frac{\partial f}{\partial x_i} = f_{x_i}$, $i = 1, \ldots, n$, in $U_\delta(\mathbf{x}_0)$ existieren und stetig in $\mathbf{x}_0 \in D$ sind, dann ist f in $\mathbf{x}_0$ differenzierbar mit*

$$f'(\mathbf{x}_0) = \big[f_{x_1}(\mathbf{x}_0), \ldots, f_{x_n}(\mathbf{x}_0)\big].$$

Insbesondere ist somit f stetig an der Stelle $\mathbf{x}_0$.

Beweis. Wir setzen $\mathbf{x} := \mathbf{x}_0 = (x_1, \ldots, x_n)^T$ und sei $\mathbf{h} = (h_1, \ldots, h_n)^T \subseteq U_\delta(\mathbf{0})$, d.h. $\|\mathbf{h}\| < \delta$. Bezeichnen $\mathbf{e}_i$, $i = 1, \ldots, n$, die kanonischen Basisvektoren des $\mathbb{R}^n$, dann ist

$$\mathbf{h} = \sum_{i=1}^{n} h_i \mathbf{e}_i.$$

Wir definieren $\mathbf{h}_0 := \mathbf{0}$ und für $k = 1, \ldots, n$

$$\mathbf{h}_k = \sum_{i=1}^{k} h_i \mathbf{e}_i \subseteq U_\delta(\mathbf{0}).$$

Dann gilt die Identität

$$f(\mathbf{x}+\mathbf{h}) - f(\mathbf{x}) = f(\mathbf{x}+\mathbf{h}_n) - f(\mathbf{x}+\mathbf{h}_0) = \sum_{k=0}^{n-1} \big(f(\mathbf{x}+\mathbf{h}_{k+1}) - f(\mathbf{x}+\mathbf{h}_k)\big).$$

Da

$$\mathbf{x}+\mathbf{h}_{k+1} = \mathbf{x}+\mathbf{h}_k + h_{k+1}\mathbf{e}_{k+1},$$

unterscheiden sich $\mathbf{x}+\mathbf{h}_k$ und $\mathbf{x}+\mathbf{h}_{k+1}$ nur in der $(k+1)$-ten Variablen. Daher gilt

$$f(\mathbf{x}+\mathbf{h}_{k+1}) - f(\mathbf{x}+\mathbf{h}_k) = \frac{\partial f}{\partial x_{k+1}}(\mathbf{x}+\mathbf{h}_k + t_k h_{k+1}\mathbf{e}_{k+1})h_{k+1}$$

für ein gewisses $t_k \in (0,1)$, und insgesamt

$$\begin{aligned}
&f(\mathbf{x}+\mathbf{h}) - f(\mathbf{x}) \\
&\quad = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(\mathbf{x}+\mathbf{h}_{k-1} + t_{k-1}h_k\mathbf{e}_k)h_k \\
&\quad = \sum_{k=1}^{n} \left\{ \frac{\partial f}{\partial x_k}(\mathbf{x})h_k \right\} + \sum_{k=1}^{n} \left\{ \left(\frac{\partial f}{\partial x_k}(\mathbf{x}+\mathbf{h}_{k-1} + t_{k-1}h_k\mathbf{e}_k) - \frac{\partial f}{\partial x_k}(\mathbf{x}) \right) \frac{h_k}{\|\mathbf{h}\|} \right\} \|\mathbf{h}\| \\
&\quad := \big[f_{x_1}(\mathbf{x}), \ldots, f_{x_n}(\mathbf{x})\big]\mathbf{h} + r(\mathbf{h})\|\mathbf{h}\|.
\end{aligned}$$

Hierin gilt aufgrund der Stetigkeit der partiellen Ableitungen $\lim_{\mathbf{h}\to\mathbf{0}} r(\mathbf{h}) = \mathbf{0}$ und da

$$f'(\mathbf{x}) := \big[f_{x_1}(\mathbf{x}), \ldots, f_{x_n}(\mathbf{x})\big] \in \mathcal{L}(\mathbb{R}^n, \mathbb{R})$$

ist damit f in $\mathbf{x}$ differenzierbar. □

Definition 9.1.20. *Wir nennen die Funktion $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ mit D offen stetig differenzierbar in $\mathbf{x}_0 \in D$, falls diese Funktion in einer geeigneten Umgebung $U_\delta(\mathbf{x_0}) \subseteq D$ differenzierbar ist und die Ableitung $f'(\mathbf{x}) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R})$ in $\mathbf{x}_0$ stetig ist. Falls f für alle $\mathbf{x} \in D$ stetig differenzierbar ist, so heißt f stetig differenzierbar auf D. Wir schreiben hierfür kurz $f \in C^1(D, \mathbb{R}) = C^1(D)$.*

Korollar 9.1.21. *Es sei $D \subseteq \mathbb{R}^n$ offen und $f : D \to \mathbb{R}$. Die Funktion f ist genau dann stetig differenzierbar auf D, d.h. $f \in C^1(D)$, falls auf D alle partiellen Ableitungen $\frac{\partial f}{\partial x_i} = f_{x_i}$, $i = 1, \ldots, n$, existieren und stetig sind.*

Beweis. Da D offen, ist jeder Punkt $\mathbf{x} \in D$ ein innerer Punkt von D, d.h. es existiert ein $\delta > 0$ mit $U_\delta(\mathbf{x}) \subseteq D$. Es sind daher die Voraussetzungen von Satz 9.1.19 erfüllt, womit die Differenzierbarkeit von f auf D folgt. Insbesondere gilt nach Satz 9.1.14, dass

$$f'(\mathbf{x}) = \big[f_{x_1}(\mathbf{x}), \ldots, f_{x_n}(\mathbf{x})\big]$$

auf D. Die Stetigkeit der partiellen Ableitungen auf D implizieren nun direkt die Stetigkeit von f' auf D.

Sei f nun stetig differenzierbar auf D. Nach Satz 9.1.14 existieren die partiellen Ableitungen für alle $\mathbf{x} \in D$ und es ist

$$f'(\mathbf{x}) = \big[f_{x_1}(\mathbf{x}), \ldots, f_{x_n}(\mathbf{x})\big].$$

Aus der Stetigkeit von f' auf D folgt wieder direkt die Stetigkeit der partiellen Ableitungen auf D. □

9.2 Höhere (partielle) Ableitungen

Seien $f : D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ und f_{x_i}, $i = 1, \ldots, n$, die partiellen Ableitungen. Falls die Funktionen wieder partiell nach den Variablen x_i abgeleitet werden können, lassen sich die partiellen Ableitungen zweiter oder höherer Ordnung definieren, z.B.

$$\begin{aligned} f_{x_1x_1} &= (f_{x_1})_{x_1} = \frac{\partial^2 f}{\partial x_1^2}, \\ f_{x_1x_2} &= (f_{x_1})_{x_2} = \frac{\partial^2 f}{\partial x_1 \partial x_2}, \\ f_{x_1x_3} &= (f_{x_1})_{x_3} = \frac{\partial^2 f}{\partial x_1 \partial x_3}, \ldots . \end{aligned}$$

Beispiel 9.2.1. *Betrachten wir die Funktion*

$$f(x, y) = xe^y + y^2$$

dann gilt

$$f_x(x, y) = e^y, \qquad f_y(x, y) = xe^y + 2y,$$

und somit

$$f_{xy}(x, y) = f_{yx}(x, y) = e^y.$$

Es zeigt sich, dass unter hinreichenden Voraussetzungen die Reihenfolge der partiellen Ableitungen unwesentlich ist.

Satz 9.2.2 (Satz von Schwarz). *Es sei $f : D \subseteq \mathbb{R}^2 \to \mathbb{R}$ mit D offen. Für $\mathbf{x}_0 = (x_0, y_0) \in D$ seien die Funktionen f, f_x, f_y und f_{xy} in einer geeigneten Umgebung $U_\delta(\mathbf{x}_0)$, definiert und stetig. Dann existiert auch $f_{yx}(\mathbf{x}_0)$ und es gilt*

$$f_{xy}(\mathbf{x}_0) = f_{yx}(\mathbf{x}_0).$$

Beweis. Sei $\mathbf{x}_0 = (x_0, y_0)^T \in D$ und $r > 0$ so gewählt, dass

$$I := [x_0 - r, x_0 + r] \times [y_0 - r, y_0 + r] \subseteq U_\delta(\mathbf{x}_0).$$

Für $|k| < r$ beliebig aber fest, definieren wir

$$g(x) := f(x, y_0) - f(x, y_0 + k),$$

wobei gilt $g \in C^1([x_0 - r, x_0 + r])$. Damit folgt mit dem Hauptsatz der Differential- und Integralrechnung

$$g(x_0 + h) - g(x_0) = \int_{x_0}^{x_0+h} g'(u)du = \int_{x_0}^{x_0+h} \big(f_x(u, y_0) - f_x(u, y_0 + k)\big)du.$$

für alle $|h| < r$. Wegen der Stetigkeit von f_{xy} folgt

$$f_x(u, y_0 + k) - f_x(u, y_0) = \int_{y_0}^{y_0+k} f_{xy}(u, v)dv$$

und somit

$$\begin{aligned} &f(x_0 + h, y_0 + k) - f(x_0 + h, y_0) - f(x_0, y_0 + k) + f(x_0, y_0) \\ &\quad = g(x_0 + h) - g(x_0) \\ &\quad = \int_{x_0}^{x_0+h} \big(f_x(u, y_0) - f_x(u, y_0 + k)\big)du \\ &\quad = \int_{x_0}^{x_0+h} \int_{y_0}^{y_0+k} f_{xy}(u, v)dvdu. \end{aligned}$$

Für eine Nullfolge $(k_j)_{j\in\mathbb{N}}$, d.h. $k_j \to 0$ für $j \to \infty$, mit $|k_j| \leq r$, folgt hieraus

$$\frac{f(x_0 + h, y_0 + k_j) - f(x_0 + h, y_0)}{k_j} - \frac{f(x_0, y_0 + k_j) - f(x_0, y_0)}{k_j} = \int_{x_0}^{x_0+h} F_j(u)du,$$

wobei

$$F_j(u) = \frac{1}{k_j} \int_{y_0}^{y_0+k_j} f_{xy}(u, v)dv.$$

Sei $u \in [x_0 - r, x_0 + r]$, dann gilt

$$\begin{aligned} |F_j(u) - f_{xy}(u, y_0)| &= \left| \frac{1}{k_j} \int_{y_0}^{y_0+k_j} f_{xy}(u, v) - \frac{1}{k_j} \int_{y_0}^{y_0+k_j} f(u, y_0)dv \right| \\ &\leq \frac{1}{k_j} \int_{y_0}^{y_0+k_j} |f_{xy}(u, v) - f_{xy}(u, y_0)|dv. \end{aligned}$$

Da die Funktion f_{xy} stetig auf I ist und I kompakt ist, ist sie dort auch gleichmäßig stetig. Zu jedem $\varepsilon > 0$ existiert daher ein $\delta > 0$, so dass aus $\|\mathbf{x} - \mathbf{x}_0\| < \delta$ folgt $|f_{xy}(\mathbf{x}) - f_{\mathbf{x}_0}(x_0, y_0)| < \varepsilon$. Somit gibt es einen Index $N \in \mathbb{N}$, so dass für alle $j > N$ gilt

$$\frac{1}{k_j} \int_{y_0}^{y_0+k_j} |f_{xy}(u, v) - f_{xy}(u, y_0)|dv \leq \frac{1}{k_j} \cdot \varepsilon \cdot k_j = \varepsilon.$$

Dies bedeutet, dass die Funktionsfolge $(F_j)_{j\in\mathbb{N}}$ auf I gleichmäßig gegen $f_{xy}(\cdot, y_0)$ konvergiert. Somit erhalten wir für $j \to \infty$ dass

$$\begin{aligned}
&f_y(x_0+h, y_0) - f_y(x_0, y_0) \\
&\quad = \lim_{j\to\infty} \frac{f(x_0+h, y_0+k_j) - f(x_0+h, y_0)}{k_j} - \frac{f(x_0, y_0+k_j) - f(x_0, y_0)}{k_j} \\
&\quad = \lim_{j\to\infty} \int_{x_0}^{x_0+h} F_j(u)du \\
&\quad = \int_{x_0}^{x_0+h} f_{xy}(u, y_0)du.
\end{aligned}$$

Folglich gilt aufgrund des Hauptsatzes der Differential- und Integralrechnung die Identität

$$f_{yx}(x_0, y_0) = \lim_{h\to 0} \frac{f_y(x_0+h, y_0) - f_y(x_0, y_0)}{h} = \lim_{h\to 0} \int_{x_0}^{x_0+h} f_{xy}(u, y_0)du = f_{xy}(x_0, y_0).$$

□

Definition 9.2.3. *Ein Vektor $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}_0^n$ heißt ein* **Multi-index**. *Wir definieren*

$$|\boldsymbol{\alpha}| := \alpha_1 + \alpha_2 + \ldots + \alpha_n = \sum_{k=1}^{n} \alpha_k,$$

$$\boldsymbol{\alpha}! := \alpha_1 \cdot \alpha_2 \cdot \ldots \cdot \alpha_n = \prod_{k=1}^{n} \alpha_k.$$

Falls f genügend oft partiell differenzierbar ist, dann schreiben wir

$$D^{\boldsymbol{\alpha}} f := \frac{\partial^{\boldsymbol{\alpha}} f}{\partial \mathbf{x}^{\boldsymbol{\alpha}}} := \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \cdot \frac{\partial^{\alpha_2}}{\partial x_2^{\alpha_2}} \cdot \ldots \cdot \frac{\partial^{\alpha_n}}{\partial x_n^{\alpha_n}} f.$$

Satz 9.2.4 (Satz von Taylor im $\mathbb{R}^n$). *Sei $D \subseteq \mathbb{R}^n$ offen und $f : D \to \mathbb{R}$ mit $f \in C^{m+1}(D, \mathbb{R})$. Seien $\mathbf{x}, \mathbf{x}_0 \in D$, so dass ihre Verbindungslinie*

$$G := \{\mathbf{y} = t\mathbf{x} + (1-t)\mathbf{x}_0 : t \in [0,1]\}$$

in D enthalten ist. Dann gilt

$$f(\mathbf{x}) = T_m(\mathbf{x}, \mathbf{x}_0) + R_m(\mathbf{x}, \mathbf{x}_0)$$

mit dem Taylor-Polynom

$$T_m(\mathbf{x}, \mathbf{x}_0) = \sum_{k=0}^{m} \sum_{|\boldsymbol{\alpha}|=k} \frac{D^{\boldsymbol{\alpha}} f(\mathbf{x}_0)}{\boldsymbol{\alpha}!} (\mathbf{x} - \mathbf{x}_0)^{\boldsymbol{\alpha}}$$

und dem Lagrangeschen Restglied

$$R_m(\mathbf{x}, \mathbf{x}_0) = \int_0^1 \frac{(1-t)^m}{m!} \sum_{|\boldsymbol{\alpha}|=m+1} D^{\boldsymbol{\alpha}} f\big(\mathbf{x}_0 + t(\mathbf{x} - \mathbf{x}_0)\big)(\mathbf{x} - \mathbf{x}_0)^{\boldsymbol{\alpha}} dt.$$

Beweis. Wir setzen $\mathbf{y} := (y_1, \ldots, y_n)^T := \mathbf{x} - \mathbf{x}_0$ und betrachten die Funktionen $g : [0,1] \to \mathbb{R}$ gegeben durch

$$g(t) := f(\mathbf{x}_0 + t\mathbf{y}).$$

Aufgrund der Kettenregel folgt $g \in C^{m+1}([0,1])$ mit

$$\begin{aligned} g'(t) &= \sum_{i=1}^{n} \frac{\partial f(\mathbf{x}_0 + t\mathbf{y})}{\partial x_i} y_i, \\ g''(t) &= \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f(\mathbf{x}_0 + t\mathbf{y})}{\partial x_i \partial x_j} y_i y_j, \end{aligned}$$

und allgemein

$$g^{(k)}(t) = \sum_{|\boldsymbol{\alpha}|=k} \frac{k!}{\boldsymbol{\alpha}!} D^{\boldsymbol{\alpha}} f(\mathbf{x}_0 + t\mathbf{y}) \mathbf{y}^{\boldsymbol{\alpha}}.$$

Mit der Taylor-Formel aus Satz 7.5.4 folgt dann

$$\begin{aligned} g(1) - g(0) &= \sum_{k=1}^{m} \frac{g^{(k)}(0)}{k!} + \int_0^1 \frac{g^{(m+1)}(t)}{m!} (1-t)^m dt \\ &= \sum_{k=1}^{m} \sum_{|\boldsymbol{\alpha}|=k} \frac{D^{\boldsymbol{\alpha}} f(\mathbf{x}_0)}{\boldsymbol{\alpha}!} \mathbf{y}^{\boldsymbol{\alpha}} + \int_0^1 \frac{(1-t)^m}{m!} \Big\{ \sum_{|\boldsymbol{\alpha}|=m+1} D^{\boldsymbol{\alpha}} f(\mathbf{x}_0 + t\mathbf{y}) \mathbf{y}^{\boldsymbol{\alpha}} \Big\} dt \\ &= T_m(\mathbf{x}, \mathbf{x}_0) + R_m(\mathbf{x}, \mathbf{x}_0). \end{aligned}$$

□

Beispiel 9.2.5. *Wir betrachten als wichtige Anwendung das Newton-Verfahren zur Lösung der nichtlinearen Gleichung*

$$f(\mathbf{x}) = \mathbf{0},$$

wobei $f : \mathbb{R}^n \to \mathbb{R}^n$ eine differenzierbare Funktion sei.

Bezeichnen wir die Lösung dieser Gleichung mit $\mathbf{x}^\star$ und ist $\mathbf{x}_0 \in D$ ein Nährerungswert an diese Lösung, dann approximieren wir

$$f(\mathbf{x}^\star) = \mathbf{0} \approx f(\mathbf{x}_0) + f'(\mathbf{x}_0)(\mathbf{x}^\star - \mathbf{x}_0).$$

Falls $\big(f'(\mathbf{x}_0)\big)^{-1} \in \mathbb{R}^{n \times n}$ existiert, so folgt

$$\mathbf{x}^\star \approx \mathbf{x}_0 - \big(f'(\mathbf{x}_0)\big)^{-1} f(\mathbf{x}_0).$$

Setzten wir

$$\mathbf{x}_1 := \mathbf{x}_0 - \big(f'(\mathbf{x})_0\big)^{-1} f(\mathbf{x}_0)$$

so ist $\mathbf{x}_1$ möglicherweise eine bessere Näherung an $\mathbf{x}^\star$, das heißt

$$\|\mathbf{x}^\star - \mathbf{x}_1\| < \|\mathbf{x}^\star - \mathbf{x}_0\|.$$

Dies ist in einer geeigneten Umgebung von $x^\star$ sicherlich der Fall. Daher ist es naheliegend das folgende Iterationsverfahren

$$\mathbf{x}_{k+1} := \mathbf{x_k} - \big(f'(\mathbf{x}_k)\big)^{-1} f(\mathbf{x}_k)$$

zum Auffinden einer Nullstelle zu verwenden. Dies ist das bekannte Newton-Verfahren. Wann und wie schnell dieses Verfahren konvergiert, können wir an dieser Stelle nicht diskutieren.

Bemerkung 9.2.6. *Für $\mathbf{h} = (h_1, h_2, \ldots, h_n)$ und $f : \mathbb{R}^n \to \mathbb{R}$ gilt formal*

$$\langle \mathbf{h}, \nabla \rangle f = \left(\sum_{i=1}^n h_i \frac{\partial}{\partial x_i} \right) f = \sum_{i=1}^n h_i \frac{\partial f}{\partial x_i} = \sum_{|\boldsymbol{\alpha}|=1} \frac{1!}{\boldsymbol{\alpha}!} D^{\boldsymbol{\alpha}} f(\mathbf{x}) \mathbf{h}^{\boldsymbol{\alpha}},$$

$$\langle \mathbf{h}, \nabla \rangle^2 f = \left(\sum_{i=1}^n h_i \frac{\partial}{\partial x_i} \right)^2 f = \sum_{i,j=1}^n h_i h_j \frac{\partial^2 f}{\partial x_i x_j} = \sum_{|\boldsymbol{\alpha}|=2} \frac{2!}{\boldsymbol{\alpha}!} D^{\boldsymbol{\alpha}} f(\mathbf{x}) \mathbf{h}^{\boldsymbol{\alpha}},$$

und allgemein

$$\langle \mathbf{h}, \nabla \rangle^k f(\mathbf{x}) = \sum_{|\boldsymbol{\alpha}|=k} \frac{k!}{\boldsymbol{\alpha}!} D^{\boldsymbol{\alpha}} f(\mathbf{x}) \mathbf{h}^{\boldsymbol{\alpha}}.$$

Daher können wir schreiben

$$T_m(\mathbf{x}, \mathbf{h}) = \sum_{k=1}^m \frac{\langle \mathbf{h}, \nabla \rangle^k f(\mathbf{x})}{k!},$$

$$R_m(\mathbf{x}, \mathbf{h}) = \int_0^1 \frac{(1-t)^m}{m!} \langle \mathbf{h}, \nabla \rangle^{m+1} f(\mathbf{x} + t\mathbf{h}) dt,$$

und für die Taylor-Reihe bekommen wir schließlich zumindest formal

$$f(\mathbf{x} + \mathbf{h}) = e^{\langle \mathbf{h}, \nabla \rangle} f(\mathbf{x}).$$

Beispiel 9.2.7. *Im Falle $m = 2$ lautet die Taylor-Entwicklung*

$$f(\mathbf{x}) = f(\mathbf{x}_0) + f'(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) + \frac{1}{2}(\mathbf{x} - \mathbf{x}_0)^T \mathbf{H}_f(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) + R_2(\mathbf{x}, \mathbf{x}_0)$$

mit der ersten Ableitung

$$f'(\mathbf{x}_0) = \big[f_{x_1}(\mathbf{x}_0), \ldots, f_{x_1}(\mathbf{x}_n) \big] \in \mathbb{R}^n$$

und der **Hesse-Matrix**

$$\mathbf{H}_f(\mathbf{x}_0) = \big[f_{x_i x_j}(\mathbf{x}_0) \big]_{i,j=1}^n = \begin{bmatrix} f_{x_1 x_1}(\mathbf{x}_0) & f_{x_1 x_2}(\mathbf{x}_0) & \ldots & f_{x_1 x_n}(\mathbf{x}_0) \\ f_{x_2 x_1}(\mathbf{x}_0) & f_{x_2 x_2}(\mathbf{x}_0) & \ldots & f_{x_2 x_n}(\mathbf{x}_0) \\ \vdots & \vdots & & \vdots \\ f_{x_n x_1}(\mathbf{x}_0) & f_{x_n x_2}(\mathbf{x}_0) & \ldots & f_{x_n x_n}(\mathbf{x}_0) \end{bmatrix} \in \mathbb{R}^{n \times n}.$$

Für $f \in C^2(D, \mathbb{R})$ ist nach dem Satz von Schwarz die Hesse-Matrix von f symmetrisch

$$\mathbf{H}_f(\mathbf{x}_0) = \mathbf{H}_f^T(\mathbf{x}_0).$$

9.3 Nichtlineare Gleichungen

Definition 9.3.1. *Sei $D \subseteq (X, d)$ eine abgeschlossene Menge und φ eine Selbstabbildung von D, d.h. $\varphi : D \to D$. Ein Punkt $x^\star \in D$ heißt ein* **Fixpunkt** *von φ, falls dieser der Gleichung*

$$x^\star = \varphi(x^\star)$$

genügt.

Satz 9.3.2 (Banachscher Fixpunktsatz). *Sei (X, d) ein vollständiger metrischer Raum und $D \subseteq X$ abgeschlossen. Ferner sei φ eine kontrahierende Selbstabbildung von D, d.h. es existiert ein $L \in (0, 1)$, so dass*

$$d\big(\varphi(x), \varphi(y)\big) \leq L d(x, y)$$

für alle $x, y \in D$. Dann existiert genau ein Fixpunkt $x^\star \in D$ und für jeden Startwert $x_0 \in D$ konvergiert die durch die Iterationsvorschrift

$$x_{n+1} = \varphi(x_n)$$

definierte Folge $(x_n)_{n\in\mathbb{N}_0}$ gegen diesen Fixpunkt. Ferner gelten die folgenden Fehlabschätzungen

$$d(x_n, x^\star) \leq \frac{L}{1-L} d(x_n, x_{n-1}) \leq \frac{L^n}{1-L} d(x_1, x_0).$$

Beweis. Sei $x_0 \in D$ beliebig. Wir zeigen zuerst, dass die Folge $(x_k)_{k\in\mathbb{N}_0}$ eine Cauchy-Folge ist. Da $\varphi(x) \in D$ für beliebige $x \in D$ ist, gilt $x_n \in D$ für alle $n \in \mathbb{N}_0$. Für beliebiges $j \in \mathbb{N}$ folgt die Abschätzung

$$\begin{aligned} d(x_{j+1}, x_j) &= \quad d\big(\varphi(x_j \quad), \varphi(x_{j-1})\big) \leq L \ d(x_j \quad, x_{j-1}) \\ &= L \ d\big(\varphi(x_{j-1}), \varphi(x_{j-2})\big) \leq L^2 d(x_{j-1}, x_{j-2}) \\ &= L^2 d\big(\varphi(x_{j-2}), \varphi(x_{j-3})\big) \leq L^3 d(x_{j-2}, x_{j-3}) \\ &= \ldots \leq L^j d(x_1, x_0). \end{aligned}$$

Eingesetzt in die Dreiecksungleichung

$$d(x_n, x_m) \leq \sum_{j=m}^{n-1} d(x_{j+1}, x_j)$$

liefert dies

$$\begin{aligned} d(x_n, x_m) &\leq \sum_{j=m}^{n-1} L^j d(x_1, x_0) = d(x_1, x_0) \sum_{j=m}^{n-1} L^j = d(x_1, x_0) L^m \sum_{k=0}^{n-m-1} L^k \\ &\leq d(x_1, x_0) \frac{L^m}{1-L} \overset{m\to\infty}{\longrightarrow} 0. \end{aligned}$$

Dies bedeutet, zu jedem $\varepsilon > 0$ existiert ein $N \in \mathbb{N}$, so dass für alle $n \geq m > N$

$$d(x_n, x_m) \leq \varepsilon.$$

Demnach ist $(x_n)_{n\in\mathbb{N}_0}$ eine Cauchy-Folge. Da $D \subseteq X$ selbst vollständig ist, existiert das Grenzelement

$$x^\star = \lim_{n\to\infty} x_n \in \mathbb{D}.$$

Da φ nach Voraussetzung Lipschitz-stetig, also insbesondere stetig, auf D ist, folgt

$$x^\star = \lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} \varphi(x_n) = \varphi\Big(\lim_{n\to\infty} x_n\Big) = \varphi(x^\star),$$

d.h. $x^\star \in D$ ist Fixpunkt von φ. Sei $\xi \in D$ ein weiterer Fixpunkt von φ, dann gilt

$$0 \leq d(\xi, x^\star) = d\big(\varphi(\xi)), \varphi(x^\star)\big) \leq Ld(\xi, x^\star),$$

und daher $d(\xi, x^\star) = 0$, d.h. der Fixpunkt ist eindeutig.

Um die Fehlerabschätzung zu zeigen, betrachten wir

$$0 \leq d(x_n, x^\star) = d\big(\varphi(x_{n-1}), x^\star\big) \leq Ld(x_{n-1}, x_n) + Ld(x_n, x^\star)$$

Hieraus folgt

$$d(x_n, x^\star) \leq \frac{L}{1-L} d(x_n, x_{n-1}) \leq \frac{L^2}{1-L} d(x_{n-2}, x_{n-1}) \leq \ldots \leq \frac{L^n}{1-L} d(x_1, x_0).$$

□

Wir benötigen das folgende Hilfsmittel aus der Linearen Algebra.

Lemma 9.3.3. *Sei $\mathbf{A} \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$, dann ist $\mathbf{A}$ genau dann invertierbar, falls ein $c > 0$ existiert, so dass gilt $\|\mathbf{Ax}\| \geq c\|\mathbf{x}\|$ für alle $\mathbf{x} \in \mathbb{R}^n$. In diesem Fall kann $c = 1/\|\mathbf{A}^{-1}\|$ gewählt werden.*

Beweis. Falls $\|\mathbf{Ax}\| \geq c\|\mathbf{x}\|$ für alle $\mathbf{x} \in \mathbb{R}^n$, dann ist $\mathbf{A} : \mathbb{R}^n \to \mathbb{R}^n$ injektiv. Denn sei $\mathbf{x} \in \ker(\mathbf{A})$, dann ist $\|\mathbf{Ax}\| = 0$ und wegen $\|\mathbf{Ax}\| \geq c\|\mathbf{x}\|$ folgt $\mathbf{x} = \mathbf{0}$. Mit dem Rangsatz

$$\dim\big(\operatorname{im}(\mathbf{A})\big) = n - \dim\big(\ker(\mathbf{A})\big) = n$$

folgt $\mathbf{A} : \mathbb{R}^n \to \mathbb{R}^n$ ist auch surjektiv und somit bijektiv. Daher existiert $\mathbf{A}^{-1} : \mathbb{R}^n \to \mathbb{R}^n$. Umgekehrt folgt aus

$$\|\mathbf{A}^{-1}\mathbf{y}\| \leq \|\mathbf{A}^{-1}\| \cdot \|\mathbf{y}\|.$$

Wir setzten $\mathbf{x} := \mathbf{A}^{-1}\mathbf{y}$ und folgern die Ungleichung

$$\|\mathbf{x}\| \leq \|\mathbf{A}^{-1}\| \, \|\mathbf{Ax}\| = \|\mathbf{Ax}\|/c,$$

mit $c = 1/\|\mathbf{A}^{-1}\|$. □

Satz 9.3.4 (Satz von der inversen Funktion). *Sei $D \subseteq \mathbb{R}^n$ offen und $f : D \to \mathbb{R}^n$ mit $f \in C^1(D, \mathbb{R}^n)$. Im Punkt $\mathbf{x}_0 \in D$ sei $f'(\mathbf{x}_0) \in L(\mathbb{R}^n, \mathbb{R}^n)$ eine nichtsinguläre Matrix, d.h. $\det\big(f'(\mathbf{x}_0)\big) \neq 0$. Dann existieren offene Mengen $U \subseteq D$ und $V \subseteq \mathbb{R}^n$ derart, dass die folgenden Aussagen gelten:*

1. *Die Restriktion von f auf U, d.h. $f\big|_U : U \to V$, ist bijektiv.*
2. *Die Umkehrfunktion $g := \big(f\big|_U\big)^{-1} : V \to U$ ist stetig differenzierbar auf V und für alle $\mathbf{x} \in U$ und $\mathbf{y} = f(\mathbf{x})$ gilt $g'(\mathbf{y}) = \big(f'(\mathbf{x})\big)^{-1}$.*
3. *Es gilt $\det\big(g'(\mathbf{y})\big) \neq 0$ für alle $\mathbf{y} \in V$ und $\det\big(f'(\mathbf{x})\big) \neq 0$ für alle $\mathbf{x} \in U$.*

Beweis. 1. Da $f' : D \to \mathbb{R}^{n\times n}$ stetig in $\mathbf{x}_0 \in D$ ist, gibt es zu jedem $\varepsilon > 0$ ein $\delta > 0$ derart, dass

$$\|f'(\mathbf{x}) - f'(\mathbf{x}_0)\| < \varepsilon$$

für alle $x \in U := U_\delta(\mathbf{x}_0)$. Hierbei sei $\varepsilon > 0$ hinreihend klein gewählt.

Wir wollen nun für ein beliebiges, festes $\mathbf{y} \in V := f(U)$ die Gleichung

$$\mathbf{y} = f(\mathbf{x})$$

iterativ lösen. Dazu definieren wir $\varphi = \varphi_{\mathbf{y}} : U \to \mathbb{R}^n$ vermittels

$$\varphi(\mathbf{x}) := \mathbf{x} - \big(f'(\mathbf{x}_0)\big)^{-1}\big(f(\mathbf{x}) - \mathbf{y}\big),$$

Es gilt $f(\mathbf{x}) = \mathbf{y}$ genau dann, wenn $\mathbf{x}$ Fixpunkt von φ ist.

Aufgrund der Kettenregel ist φ differenzierbar auf U mit

$$\varphi'(\mathbf{x}) = \mathbf{I} - \big(f'(\mathbf{x}_0)\big)^{-1} f'(\mathbf{x}) = \big(f'(\mathbf{x}_0)\big)^{-1}\big(f'(\mathbf{x}_0) - f'(\mathbf{x})\big) \in \mathbb{R}^{n\times n}.$$

Hierbei gilt

$$\|\varphi'(\mathbf{x})\| \leq \Big\|\big(f'(\mathbf{x}_0)\big)^{-1}\Big\| \cdot \|f'(\mathbf{x}_0) - f'(\mathbf{x})\| < M\varepsilon$$

mit $M := \Big\|\big(f'(\mathbf{x}_0)\big)^{-1}\Big\|$. Sei $\varepsilon > 0$ nun so gewählt, dass $L := M\varepsilon < 1$ gilt. Wir betonen an dieser Stelle, dass L nicht von der Wahl von $\mathbf{y} \in V$ abhängt. Der Mittelwertsatz der Differentialrechnung liefert die Abschätzung

$$\|\varphi(\mathbf{x}) - \varphi(\mathbf{x}')\| \leq \|\varphi'(\boldsymbol{\xi})(\mathbf{x} - \mathbf{x}')\| \leq \|\varphi'(\boldsymbol{\xi})\|\,\|\mathbf{x} - \mathbf{x}'\| \leq L\|\mathbf{x} - \mathbf{x}'\|$$

für alle $\mathbf{x}, \mathbf{x}' \in U$, d.h. φ ist kontrahierend auf U.

Wir betrachten die Funktion $f|_U : U \to f(U) = V$, die per Definition surjektiv ist. Sie ist aber auch aus folgendem Grund injektiv, d.h. bijektiv. Seien $\mathbf{x}, \mathbf{x}' \in U$ mit $f(\mathbf{x}) = f(\mathbf{x}')$, so ist

$$\varphi(\mathbf{x}) - \varphi(\mathbf{x}') = \mathbf{x} - \mathbf{x}' - \big(f'(\mathbf{x}_0)\big)^{-1}\big(f(\mathbf{x}) - f(\mathbf{x}')\big) = \mathbf{x} - \mathbf{x}'.$$

Wegen
$$\|\mathbf{x}-\mathbf{x}'\| = \|\varphi(\mathbf{x})-\varphi(\mathbf{x}')\| \le L\|\mathbf{x}-\mathbf{x}'\|$$
und wegen der strengen Ungleichung $L<1$ folgt $\|\mathbf{x}-\mathbf{x}'\|=0$, d.h. $\mathbf{x}=\mathbf{x}'$.

Wir zeigen nun, dass $V \subseteq \mathbb{R}^n$ eine offene Menge ist. Sei $\mathbf{y}' \in V$ beliebig. Wegen der Bijektivität von $f\big|_U$ existiert ein $\mathbf{x}' \in U$ mit $f(\mathbf{x}') = f|_U(\mathbf{x}') = \mathbf{y}'$. Da U offen ist, gibt es ein $r>0$ mit $U_r(\mathbf{x}') \subseteq U$. Wir setzen $0 < r' := r(1-L)/M$ und zeigen, dass $\overline{U_{r'}(\mathbf{y}')} \subseteq V$, d.h. $\mathbf{y}'$ ist innerer Punkt von V. Hierzu sei $\mathbf{y} \in \overline{U_{r'}(\mathbf{y}')}$ beliebig aber fest, dann gilt für $\varphi = \varphi_y$ die Abschätzung
$$\begin{aligned}\|\varphi(\mathbf{x})-\mathbf{x}'\| &\le \|\varphi(\mathbf{x})-\varphi(\mathbf{x}')\| + \|\varphi(\mathbf{x}')-\mathbf{x}'\| \\ &\le L\|\mathbf{x}-\mathbf{x}'\| + \|f'(\mathbf{x}_0)^{-1}\big(f(\mathbf{x}')-\mathbf{y}\big)\| \\ &\le Lr + \|f'(x_0)^{-1}\|\cdot\|\mathbf{y}'-\mathbf{y}\| \\ &= Lr + Mr' = r\end{aligned}$$
für alle $\mathbf{x} \in U_r(\mathbf{x}')$, d.h. $\varphi : \overline{U_r(\mathbf{x}')} \to \overline{U_r(\mathbf{x}')}$ ist eine kontrahierende Selbstabbildung. Nach dem Banachschen Fixpunktsatz existiert genau ein Fixpunkt $\mathbf{x}^\star \in \overline{U_r(\mathbf{x}')} \subseteq U$. Gemäß der Definition von φ erfüllt dieser Punkt $f(\mathbf{x}^\star) = \mathbf{y}$. Folglich ist $\mathbf{y} \in V$, d.h. $\mathbf{y}'$ ist innerer Punkt von V und somit ist V offen. Dies beweist die erste Aussage des Satzes.

2. Nun kommen wir zum Beweis der zweiten Aussage. Da $f|_U : U \to V$ bijektiv ist, folgt zunächst die Existenz der Umkehrfunktion $g : V \to U$.

   Ferner ist g Lipschitz-stetig, d.h. es existiert ein $L'>0$ mit
$$\|g(\mathbf{y})-g(\mathbf{y}')\| \le L'\|\mathbf{y}-\mathbf{y}'\|$$
   für alle $\mathbf{y}=,\mathbf{y}' \in V$. Denn betrachten wir zu beliebigen $\mathbf{y}=f(\mathbf{x}), \mathbf{y}'=f(\mathbf{x}) \in V$ bzw. $\mathbf{x}=g(\mathbf{y}), \mathbf{x}'=g(\mathbf{y}') \in U$ die Abbildungen $\varphi_{\mathbf{y}}$ und $\varphi_{\mathbf{y}'}$, so können wir mit Hilfe der Beziehungen
$$\varphi_{\mathbf{y}}(\mathbf{x}) = \mathbf{x}, \qquad \varphi_{\mathbf{y}'}(\mathbf{x}') = \mathbf{x}',$$
   abschätzen
$$\begin{aligned}\|g(\mathbf{y})-g(\mathbf{y}')\| &= \|\mathbf{x}-\mathbf{x}'\| \\ &= \|\varphi_{\mathbf{y}}(\mathbf{x})-\varphi_{\mathbf{y}'}(\mathbf{x}')\| \\ &\le \|\varphi_{\mathbf{y}}(\mathbf{x})-\varphi_{\mathbf{y}}(\mathbf{x}')\| + \|\varphi_{\mathbf{y}}(\mathbf{x}')-\varphi_{\mathbf{y}'}(\mathbf{x}')\| \\ &\le L\|\mathbf{x}-\mathbf{x}'\| + \|f^{-1}(\mathbf{x}_0)(\mathbf{y}-\mathbf{y}')\| \\ &\le L\|\mathbf{x}-\mathbf{x}'\| + M\|\mathbf{y}-\mathbf{y}'\|.\end{aligned}$$
   Dies bedeutet
$$(1-L)\|\mathbf{x}-\mathbf{x}'\| \le M\|\mathbf{y}-\mathbf{y}'\|$$
   beziehungsweise
$$\|g(\mathbf{y})-g(\mathbf{y}')\| = \|\mathbf{x}-\mathbf{x}'\| \le \frac{M}{1-L}\|\mathbf{y}-\mathbf{y}'\|,$$

d.h. $L' = M/(1-L)$.

Seien $\mathbf{x} \in U$ und $\mathbf{z} \in \mathbb{R}^n$ beliebig, dann gilt wegen der unteren Dreicksungleichung und

$$\|\mathbf{z}\| = \left\|\big(f'(\mathbf{x}_0)\big)^{-1} f'(\mathbf{x}_0)\mathbf{z}\right\| \leq M\|f'(\mathbf{x}_0)\mathbf{z}\|$$

die Abschätzung

$$\begin{aligned}\|f'(\mathbf{x})\mathbf{z}\| &\geq \|f'(\mathbf{x}_0)\mathbf{z}\| - \left\|\big(f'(\mathbf{x}) - f(\mathbf{x}_0)\big)\mathbf{z}\right\| \\ &\geq \|\mathbf{z}\|/M - \|f'(\mathbf{x}) - f'(\mathbf{x}_0)\|\,\|\mathbf{z}\| \\ &\geq (1/M - \varepsilon)\|\mathbf{z}\|.\end{aligned}$$

Für $\varepsilon > 0$ hinreichend klein ist $c := 1/M - \varepsilon$ positiv und infolge Lemma 9.3.3 existiert $\big(f'(\mathbf{x})\big)^{-1}$ für alle $\mathbf{x} \in U$. Insbesondere ist $\big(f'(\mathbf{x})\big)^{-1}$ stetig auf U, d.h. $\Big(f'\big(g(\mathbf{y})\big)\Big)^{-1}$ ist stetig auf V.

Es verbleibt die Identität

$$g'(\mathbf{y}) = \Big(f'\big(g(\mathbf{y})\big)\Big)^{-1} = \big(f'(\mathbf{x})\big)^{-1}$$

für alle $\mathbf{y} = f(\mathbf{x}) \in V$ zu zeigen. Wegen der Lipschitz-Stetigkeit von g folgt für beliebige $\mathbf{y} = f(\mathbf{x}), \mathbf{y}' = f(\mathbf{x}') \in V$, d.h. $\mathbf{x} = g(\mathbf{y}), \mathbf{x}' = g(\mathbf{y}') \in U$, dass

$$\begin{aligned}&\frac{\left\|g(\mathbf{y}') - g(\mathbf{y}) - \big(f(\mathbf{x})\big)^{-1}(\mathbf{y}' - \mathbf{y})\right\|}{\|\mathbf{y}' - \mathbf{y}\|} \\ &\quad= \frac{\left\|\mathbf{x}' - \mathbf{x} - \big(f(\mathbf{x})\big)^{-1}\big(f(\mathbf{x}') - f(\mathbf{x})\big)\right\|}{\|\mathbf{x}' - \mathbf{x}\|} \cdot \frac{\|\mathbf{x}' - \mathbf{x}\|}{\|\mathbf{y}' - \mathbf{y}\|} \\ &\quad= \frac{\left\|\big(f(\mathbf{x})\big)^{-1}\big(f'(\mathbf{x})(\mathbf{x}' - \mathbf{x}) - f(\mathbf{x}') + f(\mathbf{x})\big)\right\|}{\|\mathbf{x}' - \mathbf{x}\|} \cdot \frac{\|g(\mathbf{y}') - g(\mathbf{y})\|}{\|\mathbf{y}' - \mathbf{y}\|} \\ &\quad\leq M\frac{\|f(\mathbf{x}') - f(\mathbf{x}) - f'(\mathbf{x})(\mathbf{x}' - \mathbf{x})\|}{\|\mathbf{x}' - \mathbf{x}\|} \cdot L' \\ &\quad= ML'r(\mathbf{x}' - \mathbf{x}) \xrightarrow{\mathbf{y}' \to \mathbf{y}} 0,\end{aligned}$$

da aus der Differenzierbarkeit von f an der Stelle $\mathbf{x}$ folgt $\lim_{\mathbf{x}' \to \mathbf{x}} r(\mathbf{x}' - \mathbf{x}) = 0$. Aus dieser Abschätzung schließen wir, dass g an der Stelle $\mathbf{y}$ differenzierbar ist mit

$$g'(\mathbf{y}) = \big(f'(\mathbf{x})\big)^{-1} \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n).$$

3. Aus der Existenz von $\big(f'(\mathbf{x})\big)^{-1}$ in Verbindung mit der soeben bewiesenen Eigenschaft folgt unmittelbar die dritte Aussage des Satzes.

□

Bemerkung 9.3.5. *Es sei angemerkt, dass das Iterationsverfahren aus diesem Beweis*

$$\mathbf{x}_{n+1} = \varphi(\mathbf{x}_n) := \mathbf{x}_n - \big(f'(\mathbf{x}_0)\big)^{-1}\big(f(\mathbf{x}_n) - \mathbf{y}\big),$$

das vereinfachte Newton-Verfahren zum Lösen der Gleichung $f(\mathbf{x}) = \mathbf{y}$ entspricht. Im Gegensatz zum Newton-Verfahren wird die Matrix $\big(f'(\mathbf{x}_0)\big)^{-1}$ nur für den Startwert $\mathbf{x}_0$ berechnet. Dies entspricht der approximierten Lösung der Gleichung

$$\mathbf{y} = f(\mathbf{x}) \approx f(\mathbf{x}_0) + f'(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)$$

Beispiel 9.3.6. *Die Aussagen des Satzes sind lokal, d.h. im allgemeinen können wir nicht auf die globale Invertierbarkeit von f schließen.*

Betrachten wir beispielsweise die Funktion $f : [1, \infty] \times \mathbb{R} \to \mathbb{R}^2$ gegeben durch

$$f(r, \varphi) = r \begin{bmatrix} \cos\varphi \\ \sin\varphi \end{bmatrix},$$

so ist f nicht invertierbar aufgrund der Periodizität des Sinus und Cosinus.

Für die Ableitung

$$f'(r, \varphi) = \begin{bmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{bmatrix} \in \mathbb{R}^{2\times 2}$$

gilt $\det f'(r, \varphi) = r \geq 1$, d.h. die Matrix $f'(r, \varphi)$ ist überall invertierbar. Daher existiert lokal die Umkehrfunktion von f.

Wir führen nun die folgende Notation ein. Sei $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{y} \in \mathbb{R}^m$ und $f : D \subseteq \mathbb{R}^{n+m} \to \mathbb{R}^n$. Wir schreiben

$$(\mathbf{x}, \mathbf{y}) := \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} = (x_1, \ldots, x_n, y_1, \ldots, y_m)^T \in \mathbb{R}^{n+m}$$

und

$$f(\mathbf{x}, \mathbf{y}) := f(x_1, \ldots, x_n, y_1, \ldots, y_m).$$

Sei $\mathbf{A} : \mathbb{R}^{n+m} \to \mathbb{R}^n$ linear, dann schreiben wir

$$\mathbf{A} = \big[\mathbf{A}_\mathbf{x}, \mathbf{A}_\mathbf{y}\big] = \left[\begin{array}{cccc|ccc} a_{1,1} & a_{1,2} & \cdots & a_{1,n} & a_{1,n+1} & \cdots & a_{1,n+m} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} & a_{2,n+1} & \cdots & a_{2,n+m} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} & a_{n,n+1} & \cdots & a_{n,n+m} \end{array}\right] \in \mathbb{R}^{n\times(n+m)}$$

und

$$\mathbf{A}\begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} = \big[\mathbf{A}_\mathbf{x}, \mathbf{A}_\mathbf{y}\big]\begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} = \mathbf{A}_\mathbf{x}\mathbf{x} + \mathbf{A}_\mathbf{y}\mathbf{y}.$$

Lemma 9.3.7. *Sei* $\mathbf{A} = [\mathbf{A_x}, \mathbf{A_y}] \in \mathbb{R}^{n\times(n+m)}$ *mit* $\det(\mathbf{A_x}) \neq 0$*, dann existiert zu jedem* $\mathbf{y} \in \mathbb{R}^m$ *ein* $\mathbf{x} \in \mathbb{R}^n$ *mit*

$$\mathbf{A}\begin{bmatrix}\mathbf{x}\\ \mathbf{y}\end{bmatrix} = \mathbf{0}.$$

Hierbei gilt

$$\mathbf{x} = -\mathbf{A_x}^{-1}\mathbf{A_y}\mathbf{y}.$$

Beweis. Sei $(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^{n+m}$ mit

$$\mathbf{A}\begin{bmatrix}\mathbf{x}\\ \mathbf{y}\end{bmatrix} = \mathbf{A_x}\mathbf{x} + \mathbf{A_y}\mathbf{y} = \mathbf{0}.$$

Dann gilt $\mathbf{A_x}\mathbf{x} = -\mathbf{A_y}\mathbf{y}$ und daher ist $\mathbf{x} = -\mathbf{A_x}^{-1}\mathbf{A_y}\mathbf{y}$.. □

Beispiel 9.3.8. *Wir betrachten die Funktion* $f : \mathbb{R}^{1+1} \to \mathbb{R}$ *definiert durch*

$$f(x, y) = x^2 + y^2 - 1.$$

Die Menge $\{(x, y) : f(x, y) = 0\}$ *ist der Einheitskreis. Falls* $x \geq 0$ *vorausgesetzt ist, kann* $f(x, y) = x^2 + y^2 - 1 = 0$ *nach* x *aufgelöst werden gemäß*

$$x = \sqrt{1 - y^2} =: g(y)$$

mit $g : \overline{U_1(0)} \to \mathbb{R}$.

Für lineare Abbildungen ist das Auflösen nach einer oder mehreren Variablen unter den Voraussetzung des Banachschen Fixpunktsatzes möglich.

Satz 9.3.9 (Satz über implizite Funktionen). *Sei* $D \subseteq \mathbb{R}^{n+m}$ *offen und* $f : D \to \mathbb{R}^n$ *mit* $f \in C^1(D, \mathbb{R}^n)$*. Für einen Punkt* $(\mathbf{x}_0, \mathbf{y}_0) \in D$ *gelte* $f(\mathbf{x}_0, \mathbf{y}_0) = \mathbf{0}$ *und*

$$\det\big(f'_{\mathbf{x}}(\mathbf{x}_0, \mathbf{y}_0)\big) = \det\left(\frac{\partial f}{\partial \mathbf{x}}(\mathbf{x}_0, \mathbf{y}_0)\right) \neq 0,$$

d.h. $f'_{\mathbf{x}}(\mathbf{x}_0, \mathbf{y}_0)$ *ist invertierbar. Dann existieren offene Mengen* $U \subseteq D$ *und* $W \subseteq \mathbb{R}^m$ *mit* $(\mathbf{x}_0, \mathbf{y}_0) \in U$ *und* $\mathbf{y}_0 \in W$ *und eine Funktion* $g : W \to \mathbb{R}^n$ *derart, dass* $\big(g(\mathbf{y}), \mathbf{y}\big) \in U$ *und* $f\big(g(\mathbf{y}), \mathbf{y}\big) = \mathbf{0}$ *für alle* $\mathbf{y} \in W$*. Die Funktion* g *ist stetig differenzierbar in* W*, d.h.* $g \in C^1(W, \mathbb{R}^n)$*, mit*

$$g'(\mathbf{y_0}) = -\big(f'_{\mathbf{x}}(\mathbf{x}_0, \mathbf{y}_0)\big)^{-1} f'_{\mathbf{y}}(\mathbf{x}_0, \mathbf{y}_0) = -\Big(f'_{\mathbf{x}}\big(g(\mathbf{y}_0), \mathbf{y}_0\big)\Big)^{-1} f'_{\mathbf{y}}\big(g(\mathbf{y}_0), \mathbf{y}_0\big).$$

Beweis. Wir definieren eine Abbildung $F : D \subseteq \mathbb{R}^{n+m} \to \mathbb{R}^{n+m}$ durch

$$F(\mathbf{x}, \mathbf{y}) = \big(f(\mathbf{x}, \mathbf{y}), \mathbf{y}\big) = \begin{bmatrix} f(\mathbf{x}, \mathbf{y}) \\ \mathbf{y} \end{bmatrix} = \begin{bmatrix} f_1(\mathbf{x}, \mathbf{y}) \\ \vdots \\ f_n(\mathbf{x}, \mathbf{y}) \\ \hline y_1 \\ \vdots \\ y_m \end{bmatrix}.$$

Wegen der Kettenregel ist $F \in C^1(D, \mathbb{R}^{n+m})$ mit

$$F'(\mathbf{x}_0, \mathbf{y}_0) = \begin{bmatrix} f(\mathbf{x}_0, \mathbf{y}_0) \\ \mathbf{0} \quad \mathbf{I} \end{bmatrix} = \begin{bmatrix} f'_{\mathbf{x}}(\mathbf{x}_0, \mathbf{y}_0) & f'_{\mathbf{y}}(\mathbf{x}_0, \mathbf{y}_0) \\ \mathbf{0} & \mathbf{I} \end{bmatrix}$$

$$= \left[\begin{array}{cccc|cccc} f_{1,x_1} & f_{1,x_2} & \dots & f_{1,x_n} & f_{1,y_1} & f_{1,y_2} & \dots & f_{1,y_m} \\ f_{2,x_1} & f_{2,x_1} & \dots & f_{2,x_n} & f_{2,y_1} & f_{2,y_2} & \dots & f_{2,y_m} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ f_{n,x_1} & f_{n,x_1} & \dots & f_{n,x_n} & f_{n,y_1} & f_{n,y_2} & \dots & f_{n,y_m} \\ \hline 0 & 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & 1 & & \\ \vdots & \vdots & & \vdots & \vdots & & \ddots & \\ 0 & 0 & \dots & 0 & 0 & & & 1 \end{array}\right].$$

Da nach Voraussetzung gilt

$$\det \begin{bmatrix} f'_{\mathbf{x}}(\mathbf{x}_0, \mathbf{y}_0) & f'_{\mathbf{y}}(\mathbf{x}_0, \mathbf{y}_0) \\ \mathbf{0} & \mathbf{I} \end{bmatrix} = \det\big(f'_{\mathbf{x}}(\mathbf{x}_0, \mathbf{y}_0)\big) \cdot \det(\mathbf{I}) = \det(f'_{\mathbf{x}}(\mathbf{x}_0, \mathbf{y}_0)) \neq 0,$$

ist F' an der Stelle $(\mathbf{x}_0, \mathbf{y}_0)$ invertierbar. Infolge dem Satz von der inversen Funktion existieren offene Mengen $U \subseteq D$ und $V \subseteq \mathbb{R}^{n+m}$ derart, dass $F : U \to V$ bijektiv ist. Wir setzen

$$W := \{\mathbf{y} \in \mathbb{R}^m : (\mathbf{0}, \mathbf{y}) \in V\} \subseteq \mathbb{R}^m.$$

Für $(\mathbf{x}_0, \mathbf{y}_0) \in U$ gilt $F(\mathbf{x}_0, \mathbf{y}_0) = \big(f(\mathbf{x}_0, \mathbf{y}_0), \mathbf{y}_0\big) = (\mathbf{0}, \mathbf{y}_0) \in V$, d.h. $\mathbf{y}_0 \in W$. Wir bemerken, die Offenheit von V impliziert dass auch $W \subseteq \mathbb{R}^m$ offen ist.

Sei $\mathbf{y} \in W$ beliebig, dann gibt es zu $(\mathbf{0}, \mathbf{y}) \in V$ nach dem Satz von der inversen Funktion ein $\mathbf{z} \in U$ mit $F(\mathbf{z}) = (\mathbf{0}, \mathbf{y}) = \big(f(\mathbf{z}), \mathbf{y}\big)$. Da $\mathbf{z} \in U$, ist $\mathbf{z}$ von der Form $\mathbf{z} = (\mathbf{x}, \mathbf{y})$, d.h. zu $\mathbf{y} \in W$ existiert ein $\mathbf{x} \in \mathbb{R}^n$ mit $(\mathbf{x}, \mathbf{y}) \in U$ und $f(\mathbf{x}, \mathbf{y}) = \mathbf{0}$ sowie $F(\mathbf{x}, \mathbf{y}) = (\mathbf{0}, \mathbf{y})$. Dieses $\mathbf{x}$ ist sogar eindeutig wie man folgendermaßen einsieht: sei $\mathbf{x}' \in \mathbb{R}^n$ mit $(\mathbf{x}', \mathbf{y}) \in U$ ein weiteres Element mit $f(\mathbf{x}', \mathbf{y}) = \mathbf{0}$, dann gilt

$$F(\mathbf{x}', \mathbf{y}) = \big(f(\mathbf{x}', \mathbf{y}), \mathbf{y}\big) = (\mathbf{0}, \mathbf{y}) = \big(f(\mathbf{x}, \mathbf{y}), \mathbf{y}\big) = F(\mathbf{x}, \mathbf{y}).$$

Da $F : U \to V$ injektiv ist, folgt $\mathbf{x}' = \mathbf{x}$.

Wir definieren $g : W \to \mathbb{R}^n$ durch $g(\mathbf{y}) = \mathbf{x}$ für $\mathbf{y} \in W$. Es gilt aufgrund dieser Definition $f\big(g(\mathbf{y}), \mathbf{y}\big) = \mathbf{0}$ für alle $\mathbf{y} \in W$.

Sei $G : V \to U$ die Umkehrfunktion zu F. Nach dem Satz von der inversen Funktion ist $G \in C^1(V, \mathbb{R}^{n+m})$ und $G(\mathbf{0}, \mathbf{y}) = \big(g(\mathbf{y}), \mathbf{y}\big)$ für alle $\mathbf{y} \in W$. Folglich ist auch $g \in C^1(W, \mathbb{R}^n)$. Um die Ableitung von G zu berechnen, definieren wir $h : W \to \mathbb{R}^{n+m}$ mit

$$h(\mathbf{y}) = \big(g(\mathbf{y}), \mathbf{y}\big) = \begin{bmatrix} g(\mathbf{y}) \\ \mathbf{y} \end{bmatrix}.$$

Damit ist $(f \circ h)(\mathbf{y}) = f\big(g(\mathbf{y}), \mathbf{y}\big) = \mathbf{0}$ für alle $\mathbf{y} \in W$, d.h. $f \circ h \equiv \mathbf{0}$, und mit der Kettenregel folgt

$$\mathbf{0} \equiv (f \circ h)'(\mathbf{y}) = f'\big(h(\mathbf{y})\big)h'(\mathbf{y}) = f'\big(g(\mathbf{y}), \mathbf{y}\big) \cdot \begin{bmatrix} g'(\mathbf{y}) \\ \mathbf{I} \end{bmatrix} = f'_{\mathbf{x}}\big(g(\mathbf{y}), \mathbf{y}\big)g'(\mathbf{y}) + f'_{\mathbf{y}}\big(g(\mathbf{y}), \mathbf{y}\big).$$

Hieraus folgt nach Lemma 9.3.7

$$g'(\mathbf{y}) = -\Big(f'_{\mathbf{x}}\big(g(\mathbf{y}), \mathbf{y}\big)\Big)^{-1} f'_{\mathbf{y}}(\mathbf{y}).$$

□

9.4 Extremalwertprobleme

Gegeben sei ein normierter Raum $(V, \|\cdot\|)$. Wir wollen im folgenden die Extremwerte für die auf einer Menge $D \subseteq V$ definierte Funktion $J : D \to \mathbb{R}$ untersuchen.

Definition 9.4.1. *Eine Stelle $u \in D$ heißt* **Minimum** *bzw. ein* **striktes Minimum***, falls für alle $v \in D \setminus \{u\}$ gilt $J(u) \leq J(v)$ bzw. $J(u) < J(v)$.*

Eine Stelle $u \in D$ heißt **lokales Minimum** *oder auch* **relatives Minimum** *von J, falls eine offene Umgebung $U := U_\delta(u)$ existiert, so dass $J(u) \leq J(v)$ für alle $v \in U \cap D$ gilt. Das lokale Mimimum ist strikt, falls sogar $J(u) < J(v)$ gilt.*

Ersetzen wir "$\leq$" bzw. "$<$" durch "$\geq$" bzw. "$>$", so sprechen wir entsprechend von **(strikten) Maxima** *und* **(strikten) lokalen Maxima**.

Satz 9.4.2 (Notwendige Bedingung für lokale Extremstellen). *Sei $D \subseteq V \subseteq \mathbb{R}^n$ offen und sei die Funktion $J : D \to \mathbb{R}$ stetig. Falls die Funktion J in $\mathbf{u} \in D$ ein lokales Extrema besitzt und in $\mathbf{u} \in D$ differenzierbar ist, so gilt*

$$J'(\mathbf{u}) = \mathbf{0}. \tag{9.2}$$

Beweis. Sei $\mathbf{v} \in V$ beliebig mit $\|\mathbf{v}\| = 1$. Dann existiert ein $\delta > 0$, so dass die Funktion $f(t) = J(\mathbf{u} + t\mathbf{v})$ für alle $t \in [-\delta, \delta]$ wohldefiniert ist, das heißt $\mathbf{u} + t\mathbf{v} \in D$.

Aufgrund der Kettenregel ist f an der Stelle $t = 0$ differenzierbar. Es gilt

$$f'(0) = J'(\mathbf{u})\mathbf{v}$$

mit $J'(\mathbf{u}) \in \mathbb{R}^{1\times n}$. Da J an der Stelle ein relatives Extremum besitzt, so ist $t = 0$ eine relative Extremstelle von f. Nach dem Satz von Fermat 6.2.2 gilt

$$f'(0) = 0 = J'(\mathbf{u})\mathbf{v},$$

für jedes $\mathbf{v} \in V$, das heißt

$$J'(\mathbf{u}) = [0, \ldots, 0] \in \mathcal{L}(V, \mathbb{R}).$$

□

Bemerkung 9.4.3. *1. Die Bedingung $J'(\mathbf{u}) = \mathbf{0} \in \mathbb{R}^{1\times n}$ bedeutet für den Gradienten*

$$\nabla J(\mathbf{u}) = \begin{bmatrix} \frac{\partial J(\mathbf{u})}{\partial x_1} \\ \frac{\partial J(\mathbf{u})}{\partial x_2} \\ \vdots \\ \frac{\partial J(\mathbf{u})}{\partial x_n} \end{bmatrix} = \mathbf{0}.$$

Dies sind n Gleichungen für $\mathbf{u} = (u_1, \ldots, u_n)^T \in \mathbb{R}^n$.

2. *Wie wir im Beweis gesehen haben, genügt die Gateaux-Differentierbarkeit von J in allen Richtungen $\mathbf{v} \in V$ mit $\|\mathbf{v}\| = 1$. Die notwendige Bedingung* (9.2) *lautet dann*

$$J'(\mathbf{u})\mathbf{v} = 0 \tag{9.3}$$

für alle $\mathbf{v} \in V$.

3. *Die Gleichung* (9.2) *heißt Eulersche Gleichung. Gleichung* (9.3) *ist die schwache Formulierung der Eulerschen Gleichung.*

4. *Die Stellen $\mathbf{u} \in D$ mit $J'(\mathbf{u}) = \mathbf{0}$ heißen die stationären Punkte von u. Ein stationärer Punkt muss jedoch im allgemeinen keine Extremstelle sein.*

Wir wollen zuerst auf das folgende Resultat aus der linearen Algebra über symmetrische Matrizen aufmerksam machen.

Bemerkung 9.4.4. *Eine symmetrische Matrix $\mathbf{H} \in \mathbb{R}^{n\times n}$ besitzt n nicht notwendigerweise verschiedene reelle Eigenwerte λ_i und eine zugehörige orthonormale Eigenbasis $\boldsymbol{\omega}_i$. Folglich kann jeder Vektor $\boldsymbol{\omega} \in \mathbb{R}^n$ in der Eigenbasis ausgedrückt werden*

$$\boldsymbol{\omega} = \sum_{i=1}^{n} \alpha_i \boldsymbol{\omega}_i, \qquad \alpha_i \in \mathbb{R}$$

und daher gilt

$$\boldsymbol{\omega}^T \mathbf{H} \boldsymbol{\omega} = \sum_{i,j=1}^{n} \lambda_i \alpha_i \alpha_j \boldsymbol{\omega}_j^T \boldsymbol{\omega}_i = \sum_{i=1}^{n} \lambda_i \alpha_i^2.$$

Hieraus folgt sofort, dass die Matrix $\mathbf{H}$ genau dann positiv semidefinit ist, das heißt $\lambda_i \geq 0$ für alle i, falls

$$\boldsymbol{\omega}^T \mathbf{H} \boldsymbol{\omega} \geq 0$$

für alle $\boldsymbol{\omega} \in \mathbb{R}^n$.

Gilt nun $\lambda_i > 0$ für alle i, so gilt offensichtlich sogar

$$\boldsymbol{\omega}^T \mathbf{H} \boldsymbol{\omega} \geq \min_{i=1}^{n} \{\lambda_i\} \|\boldsymbol{\omega}\|^2,$$

das heißt die Matrix $\mathbf{H}$ *ist genau dann positiv definit, falls ein* $\gamma > 0$ *existiert so dass*

$$\boldsymbol{\omega}^T \mathbf{H} \boldsymbol{\omega} \geq \gamma \|\boldsymbol{\omega}\|^2$$

für alle $\boldsymbol{\omega} \in \mathbb{R}^n$.

Entsprechende Resultate lassen sich natürlich auch im negativ (semi-) definiten Fall formulieren.

Satz 9.4.5. *Sei* $D \subseteq \mathbb{R}^n$ *offen und* $J \in C^2(D, \mathbb{R})$. *Desweiteren existiere* $\mathbf{u} \in D$ *mit* $J''(\mathbf{u}) \in \mathbb{R}^{n \times n}$. *Falls* J *an der Stelle* $\mathbf{u} \in D$ *ein lokales Minimum hat, so ist* $\boldsymbol{\omega}^T J''(\mathbf{u}) \boldsymbol{\omega} \geq 0$ *für alle* $\boldsymbol{\omega} \in \mathbb{R}^n$, *mit anderen Worten, die Hessematrix ist in einem lokalen Minimum* $\mathbf{u}$ *positiv semidefinit.*

Beweis. Zu gegebenem $\boldsymbol{\omega}$ ist für hinreichend kleines $t > 0$ sowohl der Punkt $\mathbf{u} + t\boldsymbol{\omega} \in D$ als auch die Verbindungslinie $[\mathbf{u}, \mathbf{u} + t\boldsymbol{\omega}] := \{\mathbf{u} + st\boldsymbol{\omega} : s \in [0, 1]\}$ in D enthalten. Nach dem Taylorschen Satz für die Funktion f definiert durch $f(t) := J(\mathbf{u} + t\boldsymbol{\omega})$ existiert nun zu jedem t eine Zwischenstelle $\lambda \in (0, 1)$, so dass gilt

$$J(\mathbf{u} + t\boldsymbol{\omega}) = J(\mathbf{u}) + tJ'(\mathbf{u})\boldsymbol{\omega} + \frac{1}{2}t^2 \boldsymbol{\omega}^T J''(\mathbf{u} + \lambda t\boldsymbol{\omega})\boldsymbol{\omega}.$$

Da für ein Minimum gilt $J(\mathbf{u} + t\boldsymbol{\omega}) \geq J(\mathbf{u})$ und $J'(\mathbf{u}) = \mathbf{0}$, folgt weiter

$$0 \leq \frac{1}{2}t^2 \boldsymbol{\omega}^T J''(\mathbf{u} + \lambda t\boldsymbol{\omega})\boldsymbol{\omega},$$

das heißt

$$0 \leq \boldsymbol{\omega}^T J''(\mathbf{u} + \lambda t\boldsymbol{\omega})\boldsymbol{\omega}.$$

Lassen wir nun t gegen Null streben, so folgt die Behauptung. □

Proposition 9.4.6. *Sei* $D \subseteq \mathbb{R}^n$ *offen,* $J \in C^2(D, \mathbb{R})$ *und* $J'(\mathbf{u}) = \mathbf{0}$. *Existiert ein* $\gamma > 0$ *derart, dass für alle* $\boldsymbol{\omega} \in \mathbb{R}^n$ *gilt*

$$\boldsymbol{\omega}^T J''(\mathbf{u})\boldsymbol{\omega} \geq \gamma \|\boldsymbol{\omega}\|^2, \tag{9.4}$$

so besitzt J *an der Stelle* $\mathbf{u} \in D$ *ein striktes lokales Minimum. Mit anderen Worten, gilt* $J'(\mathbf{u}) = \mathbf{0}$ *und ist die Hessematrix* $J''(\mathbf{u})$ *positiv definit, so ist die Stelle* $\mathbf{u}$ *ein striktes lokales Minimum.*

Beweis. Sei $\mathbf{0} \neq \boldsymbol{\omega} \in \mathbb{R}^n$ beliebig aber fest. Wegen der Stetigkeit von J'' in der Stelle $\mathbf{u}$ folgt, dass zu gegebenem $\varepsilon > 0$ ein $\delta > 0$ existiert mit

$$\|J''(\mathbf{u}) - J''(\mathbf{u} + t\boldsymbol{\omega})\| < \varepsilon$$

für alle $\|t\boldsymbol{\omega}\| < \delta$. Da J'' eine symmetrische Matrix ist, ist diese Gleichung äquivalent zu

$$\boldsymbol{\omega}^T \big(J''(\mathbf{u}) - J''(\mathbf{u} + t\boldsymbol{\omega})\big)\boldsymbol{\omega} < \varepsilon \|\boldsymbol{\omega}\|^2$$

und daher

$$J''(\mathbf{u} + t\boldsymbol{\omega}) > \boldsymbol{\omega}^T J''(\mathbf{u})\boldsymbol{\omega} - \epsilon \|\boldsymbol{\omega}\|^2 \geq (\gamma - \varepsilon)\|\boldsymbol{\omega}\|^2.$$

Wie im vorhergehenden Beweis, erhalten wir aus dem Taylorschen Satz die Gleichung

$$J(\mathbf{u}+t\boldsymbol{\omega}) = J(\mathbf{u}) + tJ'(\mathbf{u})\boldsymbol{\omega} + \frac{1}{2}t^2\boldsymbol{\omega}^T J''(\mathbf{u}+\lambda t\boldsymbol{\omega})\boldsymbol{\omega}$$

mit $\lambda \in (0,1)$ und t hinreichend klein. Mit $J'(\mathbf{u}) = \mathbf{0}$ ergibt sich schließlich

$$J(\mathbf{u}+t\boldsymbol{\omega}) - J(\mathbf{u}) = \frac{1}{2}t^2\boldsymbol{\omega}^T J''(\mathbf{u}+\lambda t\boldsymbol{\omega})\boldsymbol{\omega} \geq \frac{1}{2}t^2(\gamma-\varepsilon)\|\boldsymbol{\omega}\|^2 > 0,$$

das heißt die Stelle $\mathbf{u}$ ist ein striktes Minimum. □

In der folgenden Proposition übertragen wir die letzten beiden Ergebnisse auf Maxima.

Proposition 9.4.7. *Falls $J \in C^2(D)$ und für alle $\boldsymbol{\omega} \in \mathbb{R}^n$ gilt*

$$\boldsymbol{\omega}^T J''(\mathbf{u})\boldsymbol{\omega} \leq 0$$

dann besitzt J in $\mathbf{u}$ ein lokales Maximum. Gilt sogar

$$\boldsymbol{\omega}^T J''(\mathbf{u})\boldsymbol{\omega} \leq \gamma\|\boldsymbol{\omega}\|^2$$

für ein $\gamma > 0$, so ist das Maximum sogar strikt.

Die nächsten Definitionen und Sätzen beschäftigen sich mit Extremalstellen unter Nebenbedingungen.

Definition 9.4.8. *Sei $D \subseteq \mathbb{R}^{n+m}$ und $\varphi : D \to \mathbb{R}^n$. Für $\mathbf{u} \in \mathbb{R}^{n+m}$ schreiben wir $\mathbf{u} = (\mathbf{x}, \mathbf{y})$ mit $\mathbf{x} \in \mathbb{R}^n$ und $\mathbf{y} \in \mathbb{R}^m$. Durch die Menge*

$$M := \big\{(\mathbf{x}, \mathbf{y}) \in D : \varphi(\mathbf{x}, \mathbf{y}) = \mathbf{0}\big\}$$

ist eine Teilmenge von D gegeben, die wir **zulässige Menge** *nennen.*

Eine Stelle $\mathbf{u} = (\mathbf{x}, \mathbf{y}) \in M$ heißt **lokales Minimum von J unter den Nebenbedingungen** $\varphi(\mathbf{x}, \mathbf{y}) = \mathbf{0}$, *falls eine offene Umgebung*

$$U := U_\delta(\mathbf{u}) \subseteq D$$

um $\mathbf{u}$ existiert und

$$J(\mathbf{u}) \leq J(\mathbf{v})$$

für alle $\mathbf{v} \in U \cap M$ gilt.

Analog wird das **lokale Maximum von J unter den Nebenbedingungen** $\varphi(\mathbf{x}, \mathbf{y}) = \mathbf{0}$ *definiert.*

Satz 9.4.9 (Lagrange-Multiplikatoren). *Sei $D \subseteq \mathbb{R}^{n+m}$ offen und $\varphi : D \to \mathbb{R}^n$ stetig an der Stelle $\mathbf{u} \in M$ stetig differenzierbar mit*

$$\varphi'(\mathbf{u}) = \big[\varphi'_{\mathbf{x}}(\mathbf{u}), \varphi'_{\mathbf{y}}(\mathbf{u})\big].$$

Sei $\det\big(\varphi'_{\mathbf{x}}(\mathbf{u})\big) \neq 0$ und $J : D \to \mathbb{R}$ stetig und in $\mathbf{u} \in M$ ebenfalls differenzierbar. Falls $\mathbf{u}$ ein relatives Extremum von J unter den Nebenbedingungen $\varphi(\mathbf{u}) = \mathbf{0}$ ist, so existiert $\boldsymbol{\lambda} \in \mathbb{R}^n$ derart, dass

$$J'(\mathbf{u}) + \boldsymbol{\lambda}^T \varphi'(\mathbf{u}) = \mathbf{0} \in \mathcal{L}(D, \mathbb{R}). \tag{9.5}$$

Beweis. Aufgrund des Satzes über implizite Funktionen existieren offene Umgebungen $U \subseteq D$ und $W \subseteq \mathbb{R}^m$, und eine stetig differenzierbare Funktion $g : W \to \mathbb{R}^n$ mit $\varphi\big(g(\mathbf{y}), \mathbf{y}\big) = \mathbf{0}$ für alle $\mathbf{y} \in W$. Demnach gilt

$$M \cap U = \big\{\big(g(\mathbf{y}), \mathbf{y}\big) : \mathbf{y} \in W\big\}$$

und mit $\mathbf{u} = \big(g(\mathbf{y}), \mathbf{y}\big)$ folgt

$$g'(\mathbf{y}) = -\big(\varphi'_{\mathbf{x}}(\mathbf{u})\big)^{-1} \varphi'_{\mathbf{y}}(\mathbf{u}).$$

Wir definieren nun die Funktion $G : W \to \mathbb{R}$ durch

$$G(\mathbf{y}) := J\big(g(\mathbf{y}), \mathbf{y}\big)$$

für alle $\mathbf{y} \in W$. Eine Stelle $\mathbf{u} = \big(g(\mathbf{y}), \mathbf{y}\big) \in M$ ist genau dann ein relatives Extremum unter den Nebenbedingungen $\varphi(\mathbf{u}) = \mathbf{0}$, wenn G an der Stelle $\mathbf{y}$ ein relatives Minimum besitzt. In diesem Falle folgt nach Satz 9.4.2, da nach der Kettenregel G stetig differenzierbar auf W ist, dass

$$\mathbf{0} = G'(\mathbf{y}) = J'_{\mathbf{x}}\big(g(\mathbf{y}), \mathbf{y}\big) g'(\mathbf{y}) + J'_{\mathbf{y}}\big(g(\mathbf{y}), \mathbf{y}\big) = J'_{\mathbf{x}}(\mathbf{u}) g'(\mathbf{y}) + J'_{\mathbf{y}}\big(\mathbf{u}\big),$$

und somit

$$\mathbf{0} = J'_{\mathbf{y}}(\mathbf{u}) - J'_{\mathbf{x}}(\mathbf{u})\big(\varphi'_{\mathbf{x}}(\mathbf{u})\big)^{-1} \varphi'_{\mathbf{y}}(\mathbf{u}).$$

Setzen wir

$$\boldsymbol{\lambda}^T := -J'_{\mathbf{x}}(\mathbf{u})'\big(\varphi'_{\mathbf{x}}(\mathbf{u})\big)^{-1},$$

dann ist

$$J'_{\mathbf{y}}(\mathbf{u}) + \boldsymbol{\lambda}^T \varphi'_{\mathbf{y}}(\mathbf{u}) = \mathbf{0}.$$

Andererseits gilt die Identität

$$J'_{\mathbf{x}}(\mathbf{u}) = J'_{\mathbf{x}}(\mathbf{u})\big(\varphi_{\mathbf{x}}(\mathbf{u})\big)^{-1} \varphi'_{\mathbf{x}}(\mathbf{u}) = -\boldsymbol{\lambda}^T \varphi_{\mathbf{x}}(\mathbf{u}).$$

Zusammengefasst ergibt sich schließlich

$$\begin{aligned} J'_{\mathbf{x}}(\mathbf{u}) + \boldsymbol{\lambda}^T \varphi'_{\mathbf{x}}(\mathbf{u}) &= \mathbf{0}, \\ J'_{\mathbf{y}}(\mathbf{u}) + \boldsymbol{\lambda}^T \varphi'_{\mathbf{y}}(\mathbf{u}) &= \mathbf{0}, \end{aligned}$$

das heißt die Behauptung. □

Bemerkung 9.4.10. *Definieren wir die sogenannte Lagrange-Funktion*

$$L : (D \times \mathbb{R}^n) \to \mathbb{R}$$

definiert durch

$$L(\mathbf{u}, \lambda) := J(\mathbf{u}) + \boldsymbol{\lambda}^T \varphi(\mathbf{u}),$$

dann sind die Nebenbedingung $\mathbf{u} \in M$ *und* (9.5) *äquivalent zu* $L'(\mathbf{u}, \boldsymbol{\lambda}) = \mathbf{0}$.

Beispiel 9.4.11. *Sei* $D = \mathbb{R}^2$ *und*

$$J(x, y) = x + y.$$

Die Nebenbedingung $x^2 + y^2 = 1$ *führt zur zulässigen Menge*

$$M = \{(x, y) : \varphi(x, y) = x^2 + y^2 - 1 = 0\}$$

Die zugehörige Lagrange-Funktion lautet

$$L(x, y, \lambda) = x + y + \lambda(x^2 + y^2 - 1).$$

Wegen

$$\begin{aligned} 0 &\stackrel{!}{=} L_x = 1 + 2\lambda x, \\ 0 &\stackrel{!}{=} L_y = 1 + 2\lambda y, \\ 0 &\stackrel{!}{=} L_\lambda = x^2 + y^2 - 1, \end{aligned}$$

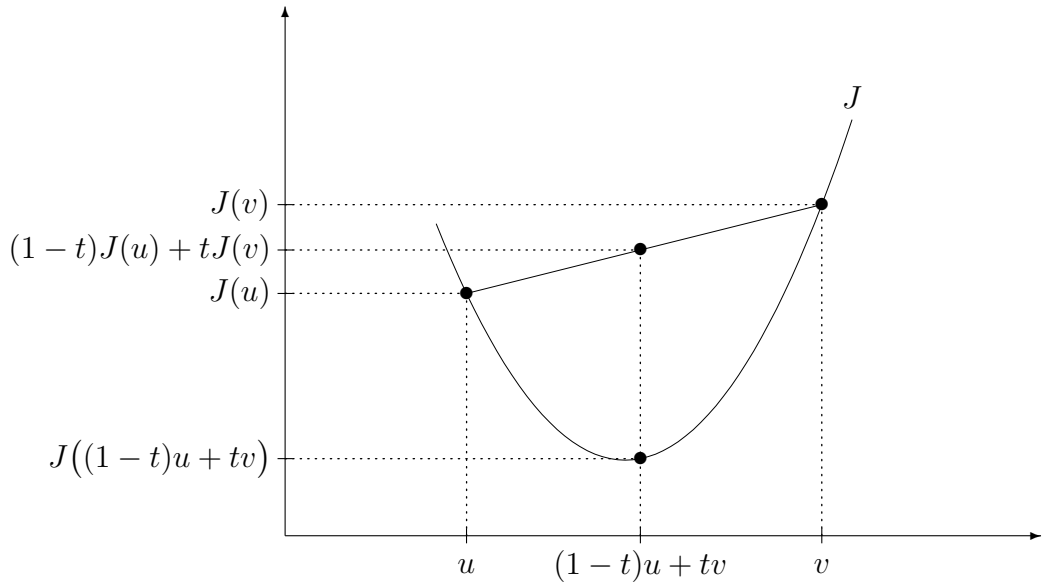
erhalten wir die zwei Punkte $(x_{1/2}, y_{1/2}, \lambda_{1/2}) = (\pm 1/\sqrt{2}, \pm 1/\sqrt{2}, \mp 1/\sqrt{2})$ *als mögliche Kandidaten einer Extremstelle. Da* $J : M \to \mathbb{R}$ *eine stetige Funktion auf der kompakten Menge* M *ist, müssen Maximum und Minimum existieren. Es gilt* $J(x_1, y_1) = \sqrt{2}$ *und* $J(x_2, y_2) = -\sqrt{2}$, *daher ist* (x_1, y_1) *die Maximalstelle und* (x_2, y_2) *die Minimalstelle.*

Wir wollen nun den Begriff der konvexen Menge und konvexen Funktion definieren.

Definition 9.4.12. *Sei* $(V, \|\cdot\|)$ *ein normierter Vektorraum. Die Menge* $K \subseteq V$ *heißt* **konvex**, *falls zu jedem* $\mathbf{u}, \mathbf{v} \in K$ *auch alle* **Konvexkombination** $\{(1-t)\mathbf{u} + t\mathbf{v} : t \in [0, 1]\}$ *in* K *enthalten sind. Eine auf der konvexen Menge* K *definierte Funktion* $J : K \to \mathbb{R}$ *heißt* **konvex**, *falls für alle* $\mathbf{u}, \mathbf{v} \in K$ *und* $t \in [0, 1]$ *gilt*

$$(1-t)J(\mathbf{u}) + tJ(\mathbf{v}) \geq J\big((1-t)\mathbf{u} + t\mathbf{v}\big).$$

Die Funktion J *heißt* **strikt konvex**, *falls strikte Ungleichheit für* $\mathbf{u} \neq \mathbf{v}$ *gilt.*



Bei einer eindimensionalen konvexen Funktion liegt die Verbindungslinie zweier Punkte oberhalb des Graphen.

Satz 9.4.13 (Notwendiges Kriterium für Minima bei konvexen Mengen). *Sei $K \subseteq V$ eine konvexe Menge und $J : K \to \mathbb{R}$. Ist $\mathbf{u} \in K$ ein lokales Minimum von J und J in $\mathbf{u}$ differenzierbar, so gilt für alle $\mathbf{v}$ die Variationsungleichung*

$$J'(\mathbf{u})(\mathbf{v}-\mathbf{u}) \geq 0.$$

Beweis. Betrachten wir zu $\mathbf{v} \in K$ die Funktion $f : [0,1] \to \mathbb{R}$ definiert durch

$$f(t) := J\big((1-t)\mathbf{u} + t\mathbf{v}\big).$$

Da $\mathbf{u}$ ein lokales Minimum ist, gilt

$$0 \leq J\big((1-t)\mathbf{u} + t\mathbf{v}\big) - J(\mathbf{u}) = tJ'(\mathbf{u})(\mathbf{v}-\mathbf{u}) + tr(t)$$

mit $|r(t)| \to 0$ für $t \to 0$. Hieraus folgt

$$0 \leq J'(\mathbf{u})(\mathbf{v}-\mathbf{u}) + r(t),$$

das heißt für $t \to 0$ die Behauptung. □

Bemerkung 9.4.14. *Dieser Satz gilt auch für abgeschlossene konvexe Mengen, die vorigen Sätze dagegen nur für offene Mengen.*

Satz 9.4.15 (Konvexitätskriterien). *Sei $J \in C^1(D, \mathbb{R})$ und $K \subseteq D$ konvex.*

1. *Die Funktion $J : K \to \mathbb{R}$ ist genau dann konvex, wenn für alle $\mathbf{u}, \mathbf{v} \in K$ gilt*

$$J(\mathbf{v}) - J(\mathbf{u}) \geq J'(\mathbf{u})(\mathbf{v}-\mathbf{u}).$$

2. *Die Funktion $J : K \to \mathbb{R}$ ist genau dann strikt konvex, wenn für alle $\mathbf{u}, \mathbf{v} \in K$ gilt*

$$J(\mathbf{v}) - J(\mathbf{u}) > J'(\mathbf{u})(\mathbf{v} - \mathbf{u}),$$

3. *Falls $J \in C^2(D, \mathbb{R})$ mit $D \subseteq \mathbb{R}^n$ ist, so ist $J : K \to \mathbb{R}$ genau dann konvex, wenn für alle $\mathbf{u} \in K$ und für alle $\boldsymbol{\omega} \in \mathbb{R}^n$ gilt*

$$\boldsymbol{\omega}^T J''(\mathbf{u})\boldsymbol{\omega} \geq 0.$$

Beweis. 1. "$\Rightarrow$": Sei $t \in [0, 1]$ und $J : K \to \mathbb{R}$ konvex, dann gilt für beliebige $\mathbf{u}, \mathbf{v} \in K$ die Ungleichung

$$J\big((1-t)\mathbf{u} + t\mathbf{v}\big) \leq (1-t)J(\mathbf{u}) + tJ(\mathbf{v}),$$

das heißt

$$\frac{J\big(\mathbf{u} + t(\mathbf{v} - \mathbf{u})\big) - J(\mathbf{u})}{t} \leq J(\mathbf{v}) - J(\mathbf{u}).$$

Lassen wir nun t gegen Null streben, so ergibt sich

$$J'(\mathbf{u})(\mathbf{v} - \mathbf{u}) \leq J(\mathbf{v}) - J(\mathbf{u}).$$

"$\Leftarrow$": Seien $\mathbf{u}, \mathbf{v} \in K$ beliebig aber fest. Da die Menge K konvex ist, ist für jedes $t \in [0, 1]$ auch $\mathbf{w} := (1-t)\mathbf{u} + t\mathbf{v}$ in K enthalten. Nach Voraussetzung gelten für ein solches $\mathbf{w}$ die Ungleichungen

$$\begin{aligned} J'(\mathbf{w})(\mathbf{v} - \mathbf{w}) &\leq J(\mathbf{v}) - J(\mathbf{w}), \\ J'(\mathbf{w})(\mathbf{u} - \mathbf{w}) &\leq J(\mathbf{u}) - J(\mathbf{w}). \end{aligned}$$

Addieren wir nun das t-fache der ersten Ungleichung zum $(1-t)$-fachen der zweiten, so erhalten wir

$$0 = J'(\mathbf{w})\big((1-t)\mathbf{u} + t\mathbf{v} - \mathbf{w}\big) \leq (1-t)J(\mathbf{u}) + tJ(\mathbf{v}) - J(\mathbf{w}),$$

was gleichbedeutend ist mit

$$J\big((1-t)\mathbf{u} + t\mathbf{v}\big) \leq (1-t)J(\mathbf{u}) + tJ(\mathbf{v}).$$

2. Die Behauptung folgt analog zur vorherigen, indem wir "$\leq$" durch "$<$" ersetzen.

3. Die letzte Aussage verbleibt dem Leser als Übung.

□

Satz 9.4.16. *Sei $K \subseteq \mathbb{R}^n$ konvex und $J : K \to \mathbb{R}$ eine konvexe Funktion, dann gilt:*

1. *Falls $\mathbf{u} \in K$ ein relatives Minimum von J ist, so ist $\mathbf{u}$ ein globales Minimum, das heißt es gilt*

$$J(\mathbf{u}) \leq J(\mathbf{v})$$

für alle $\mathbf{v} \in K$.

2. *Falls $J \in C^1(K, \mathbb{R})$, dann hat J in $\mathbf{u} \in K$ ein globales Minimum, wenn für alle $\mathbf{v} \in K$ gilt*
$$J'(\mathbf{u})(\mathbf{v} - \mathbf{u}) \geq 0.$$

Beweis. 1. Sei $\mathbf{v} \in K$ beliebig, dann gilt für t hinreichend klein
$$0 \leq J\big((1-t)\mathbf{u} + t\mathbf{v}\big) - J(\mathbf{u}) \leq (1-t)J(\mathbf{u}) + tJ(\mathbf{v}) - J(\mathbf{u})$$
das heißt $0 \leq J(\mathbf{v}) - J(\mathbf{u})$ für beliebiges $\mathbf{v} \in K$.

2. Die Voraussetzung impliziert zusammen mit den Konvexitätskriterien
$$0 \leq J'(\mathbf{u})(\mathbf{v} - \mathbf{u}) \leq J(\mathbf{u}) - J(\mathbf{v})$$
für alle $\mathbf{v} \in K$, das heißt $\mathbf{u}$ ist in der Tat ein globales Minimum.

□

Abschließend wollen wir noch folgenden wichtigen Satz beweisen, der uns in einem metrischen Raum (X, d) für eine Funktion $J : X \to \mathbb{R}$ die Existenz eines globalen Minumums liefert ohne die Differenzierbarkeit von J vorauszusetzen.

Satz 9.4.17. *Sei (X, d) ein metrischer Raum und $J : X \to \mathbb{R}$. Falls für alle $\alpha \in \mathbb{R}$ die Mengen*
$$K_\alpha := \{\mathbf{v} \in X : J(\mathbf{v}) \leq \alpha\}$$
kompakt sind, dann existiert ein globales Minumum $\mathbf{u} \in X$.

Beweis. Wir müssen zeigen, dass ein $\mathbf{u} \in X$ existiert, so dass $J(\mathbf{u}) \leq J(\mathbf{v})$ für alle $\mathbf{v} \in X$.

Wir betrachten zunächst den Fall, dass $\alpha^\star := \inf_{\mathbf{v} \in X} J(\mathbf{v}) > -\infty$. Dann existiert eine streng monoton fallende Folge $(\alpha_n)_{n \in \mathbb{N}}$ mit $\lim_{n \to \infty} \alpha_n = \alpha^\star$. Die Mengen
$$K_n := K_{\alpha_n} = \{\mathbf{v} \in X : J(\mathbf{v}) \leq \alpha_n\}$$
sind nichtleer, geschachtelt $K_{n+1} \subseteq K_n$ und aufgrund der Voraussetzung kompakt. Aufgrund des Satzes über zentrierte Systeme kompakter Mengen 3.2.7 ist
$$K^\star := \bigcap_{n=1}^{\infty} K_n \neq \emptyset.$$
Sei $\mathbf{u} \in K^\star$, dann gilt für alle $\alpha > \alpha^\star$, dass ein $n \in \mathbb{N}$ existiert mit $\alpha \geq \alpha_n \geq \alpha^\star$. Wir zeigen, dass nun $J(\mathbf{u}) = \alpha^\star = \inf \alpha_n$ gelten muss.

Hierzu sei zuerst angenommen, dass $J(\mathbf{u}) > \alpha^\star$. Dann existiert ein $n \in \mathbb{N}$ mit $\alpha_n \geq J(\mathbf{u}) > \alpha_{n+1}$, das heißt $\mathbf{u} \in K_n$ und $\mathbf{u} \notin K_{n+1}$. Demnach gilt $\mathbf{u} \notin \bigcap_{n=1}^{\infty} K_n = K^\star$, was im Widerspruch

zu $\mathbf{u} \in K^\star$ steht. Andererseits kann aber auch nicht $J(\mathbf{u}) < \alpha^\star$ gelten, denn dann ist $\alpha^\star \neq \inf_{\mathbf{v}\in X} J(\mathbf{v})$. Folglich gilt also $J(\mathbf{u}) = \alpha^\star$.

Sei nun $\mathbf{v} \in X$ beliebig, dann existiert ein $n \in \mathbb{N}$, so dass $J(\mathbf{v}) > \alpha_n > \alpha^\star = J(\mathbf{u})$, oder es gilt $\mathbf{v} \in K^\star$. Im letzten Fall gilt dann aber $J(\mathbf{v}) = \alpha^\star = J(\mathbf{u})$.

Im Fall $\alpha^\star = -\infty$ betrachten wir eine bestimmt divergente, streng monoton fallende Folge $(\alpha_n)_{n\in\mathbb{N}}$ mit $\alpha_n \to -\infty$ für $n \to \infty$. Die gleiche Argumentation wie oben sichert das

$$K^\star = \bigcap_{n=1}^{\infty} K_n \neq \emptyset.$$

Für beliebiges $\mathbf{u} \in K^\star$ gilt $J(\mathbf{u}) < \alpha_n$ für alle $n \in \mathbb{N}$, was im Widerspruch steht zu $J(\mathbf{u}) \in \mathbb{R}$. $\square$

Kapitel 10

Integration im $\mathbb{R}^n$

10.1 Grundlagen

Wir haben bisher Stetigkeit und Differenzierbarkeit von Funktion nur auf offenen Mengen des $\mathbb{R}^n$ betrachtet. Da wir diese Begriffe nun ebenfalls auf abgeschlossenen Mengen benötigen, treffen wir die folgende Vereinbarung.

Eine Funktion $f : \mathbb{R}^m \to \mathbb{R}^n$ nennen wir stetig bzw. differenzierbar auf der abgeschlossenen Menge $D \subseteq \mathbb{R}^n$, falls eine offene Umgebung E von D existiert, auf der f stetig bzw. differenzierbar ist. Entsprechend sagen wir $f \in C^k(D, \mathbb{R}^n)$, falls $f \in C^k(E, \mathbb{R}^n)$ für eine offene Umgebung E von D.

Definition 10.1.1. *Sei*

$$I^n = \{\mathbf{x} \in \mathbb{R}^n : a_i \leq x_i \leq b_i, 1 \leq i \leq n\} = \bigotimes_{i=1}^{n} [a_i, b_i]$$

eine n-Zelle bzw. ein n-Quader und sei $f : I^n \to \mathbb{R}$ stetig. Für $k = 1, 2, \ldots, n$ seien die Funktionen

$$f_k : I^k \to \mathbb{R}$$

rekursiv gegeben durch

$$\begin{aligned} f_n(x_1, \ldots, x_n) &:= f(x_1, \ldots, x_n), \\ f_{n-1}(x_1, \ldots, x_{n-1}) &:= \int_{a_n}^{b_n} f(x_1, \ldots, x_{n-1}, x_n) dx_n, \end{aligned}$$

und allgemein

$$f_{k-1}(x_1, \ldots, x_{k-1}) := \int_{a_k}^{b_k} f_k(x_1, \ldots, x_{k-1}, x_k) dx_k.$$

Das **Integral** *von f über I^n definieren wir vermittels*

$$\int_{I^n} f(\mathbf{x})d\mathbf{x} = \int_{I^n} f(x_1, \ldots, x_n)dx_1 \cdots dx_n = \int_{a_1}^{b_1} f_1(x_1)dx_1.$$

Die n-Zelle I^n heißt das **Integrationsgebiet**.

Beispiel 10.1.2. *Sei $f(\mathbf{x}) \equiv 1$, dann ist*

$$\begin{aligned}\int_{I^n} 1\ d\mathbf{x} &= (b_n - a_n) \int_{I^{n-1}} 1\ dx_1 \cdots dx_{n-1} \\ &= (b_n - a_n)(b_{n-1} - a_{n-1}) \int_{I^{n-2}} 1\ dx_1 \cdots dx_{n-2} \\ &= \cdots = \prod_{i=1}^{n} (b_i - a_i)\end{aligned}$$

das Volumen der n-Zelle I^n.

Satz 10.1.3. *Sei $I^n \subseteq \mathbb{R}^n$ eine n-Zelle und $f : I^n \to \mathbb{R}$ stetig.*

1. *Es gilt die Ungleichung*
$$\left| \int_{I^n} f(\mathbf{x})\ d\mathbf{x} \right| \leq \int_{I^n} |f(\mathbf{x})|\ d\mathbf{x}.$$
2. *Für jedes $\alpha \in \mathbb{R}$ gilt*
$$\int_{I^n} \alpha f(\mathbf{x})\ d\mathbf{x} = \alpha \int_{I^n} f(\mathbf{x})\ d\mathbf{x}.$$
3. *Ist $g : I^n \to \mathbb{R}$ eine weiter stetige Funktion, so gilt*
$$\int_{I^n} f(\mathbf{x}) + g(\mathbf{x})\ d\mathbf{x} = \int_{I^n} f(\mathbf{x})\ d\mathbf{x} + \int_{I^n} g(\mathbf{x})\ d\mathbf{x}.$$
4. *Für jede n-Zelle $J^n \subseteq I^n$ gilt*
$$\int_{I^n} f(\mathbf{x})\ d\mathbf{x} = \int_{J^n} f(\mathbf{x})\ d\mathbf{x} + \int_{I^n \setminus J^n} f(\mathbf{x})\ d\mathbf{x}.$$

Beweis. Der Beweis verbleibt dem Leser als Übung. □

Satz 10.1.4 (Approximationssatz von Weierstraß im $\mathbb{R}^n$). *Es sei $I^n \subseteq \mathbb{R}^n$ eine n-Zelle. Die Menge aller Polynome*

$$p_l(\mathbf{x}) = \sum_{|\mathbf{k}| \leq l} a_{\mathbf{k}} \mathbf{x}^{\mathbf{k}},$$

$l \in \mathbb{N}$, liegt dicht im Raum $C(I^n, \mathbb{R})$.

Beweis. Wir führen den Beweis des Satzes mit vollständiger Induktion bezüglich der Raumdimension n. Für $n = 1$ entspricht die Aussage Satz 8.5.1 und ist folglich wahr. Wir nehmen im folgenden also an, dass die Aussage für alle Funktionen $f \in C(I^{n-1}, \mathbb{R})$ gilt. Ohne Beschränkung der Allgemeinheit sei hierzu $I^n = [0,1]^n$ angenommen.

Setzen wir für eine beliebige stetige Funktion $f : [0,1]^n \to \mathbb{R}$

$$\begin{aligned} g(x_1, x_2, \ldots, x_n) := f(x_1, x_2, \ldots, x_n) \\ - (1 - x_1) f(0, x_2, x_3, \ldots, x_n) - x_1 f(1, x_2, x_3, \ldots, x_n) \\ - (1 - x_2) f(x_1, 0, x_3, \ldots, x_n) - x_2 f(x_1, 1, x_3, \ldots, x_n) \\ \vdots \\ - (1 - x_n) f(x_1, \ldots, x_{n-2}, x_{n-1}, 0) - x_n f(x_1, \ldots, x_{n-2}, x_{n-1}, 1), \end{aligned}$$

dann gilt $g(\partial I^n) \equiv 0$. Da für alle $i \in \{1, \ldots, n\}$ die Funktionen $f(x_1, \ldots, x_{i-1}, 0, x_{i+1}, \ldots, x_n)$ und $f(x_1, \ldots, x_{i-1}, 0, x_{i+1}, \ldots, x_n)$ Funktionen auf $[0,1]^{n-1}$ sind, können diese nach Induktionsvoraussetzung beliebig genau vermittels Multiploynomen approximiert werden. Zeigen wir nun die Behauptung für g, so gilt diese folglich auch für f.

Vorbereitend betrachten wir für beliebiges $l \in \mathbb{N}$ die Polynome

$$q_l(\mathbf{x}) := (c_l)^n \prod_{i=1}^{n} (1 - x_i^2)^l$$

mit

$$c_l := \left(\int_{-1}^{1} (1 - x^2)^n dx \right)^{-1}.$$

Dann ist q_l ein Polynom vom Grade $(2l)^n$, das auf $[-1,1]^n$ nichtnegativ ist. Ferner gilt aufgrund der Normierung

$$\int_{[-1,1]^n} q_l(\mathbf{x}) d\mathbf{x} = 1$$

und wegen $c_l < \sqrt{l}$ (vgl. (8.8)) folgt offensichtlich

$$q_l(\mathbf{x}) \leq (c_l)^n \leq \sqrt{l^n}$$

für alle $\mathbf{x} \in [-1,1]^n$.

Für $\mathbf{x} \notin [-\delta, \delta]^n$ finden wir zumindest eine Koordinate x_k, so dass $x_k \notin [-\delta, \delta]$. Daher folgt für $\mathbf{x} \in [-1,1]^n \setminus [-\delta, \delta]$ die Abschätzung

$$q_l(\mathbf{x}) = (c_l)^n \underbrace{(1 - x_k^2)^l}_{\leq (1-\delta^2)^l} \prod_{\substack{i \in \{1,\ldots,n\}: \\ i \neq k}} \underbrace{(1 - x_i^2)^l}_{\leq 1} \leq (1 - \delta^2)^l \sqrt{l^n},$$

das heißt q_l konvergiert für jedes $\delta > 0$ gleichmäßig gegen 0 außerhalb von $[-\delta, \delta]^n$.

Nun betrachten wir

$$p_l(\mathbf{x}) = \int_{\mathbb{R}^n} g(\mathbf{x}+\mathbf{t}) q_l(\mathbf{t})\, d\mathbf{t}.$$

Es gilt

$$\begin{aligned}
p_l(\mathbf{x}) &= \int_{\mathbb{R}^n} g(x_1+t_1,\dots,x_n+t_n) q_l(t_1,\dots,t_n)\, dt_n \cdots dt_1 \\
&= \int_{\mathbb{R}^{n-1}} \int_0^1 g(x_1+t_1,\dots,x_{n-1}+t_{n-1},t_n) q_l(t_1,\dots,t_{n-1},t_n-x_n)\, dt_n \cdots dt_1 \\
&= \int_{\mathbb{R}^{n-2}} \int_0^1 \int_0^1 g(x_1+t_1,\dots,x_{n-2}+t_{n-2},t_{n-1},t_n) \\
&\qquad\qquad \cdot q_l(t_1,\dots,t_{n-2},t_{n-1}-x_{n-1},t_n-x_n)\, dt_n \cdots dt_1 \\
&\qquad\qquad \vdots \\
&= \int_{[0,1]^n} g(\mathbf{t}) q_l(\mathbf{t}-\mathbf{x})\, d\mathbf{t},
\end{aligned}$$

dies bedeutet, p_l ist ebenfalls ein Polynom vom Grade $(2l)^n$.

Es sei angemerkt, dass im $\mathbb{R}^n$ der Würfel $[-\delta,\delta]^n$ in der Kugel $K_{\delta\sqrt{n}}(\mathbf{0})$ enthalten ist. Da g stetig ist, können wir daher zu jedem $\varepsilon > 0$ ein $\delta > 0$ so bestimmen, dass

$$|g(\mathbf{x}+\mathbf{t}) - g(\mathbf{x})| < \varepsilon$$

für alle $\mathbf{t} \in [-\delta,\delta]^n$ und $\mathbf{x} \in [-1,1]^n$. Ferner gibt es einen Index N_0, so dass $|q_l(\mathbf{t})| \le \varepsilon$ für alle $\mathbf{t} \in [-1,1]^n \setminus [-\delta,\delta]$ und $l \ge N_0$. Das Maximum von g auf $[-1,1]^n$ bezeichnen wir mit

$$M := \sup_{\mathbf{x}\in[-1,1]^n} g(\mathbf{x}).$$

Wir kommen schließlich zur entscheidenden Abschätzung

$$\begin{aligned}
|p_l(\mathbf{x}) - g(\mathbf{x})| &= \left| \int_{[-1,1]^n} \big(g(\mathbf{x}+\mathbf{t}) - g(\mathbf{x})\big) q_l(\mathbf{t}) d\mathbf{t} \right| \\
&\le \int_{[-1,1]^n} |f(\mathbf{x}+\mathbf{t}) - f(\mathbf{x})| q_l(\mathbf{t}) d\mathbf{t} \\
&\le 2M \int_{[-1,1]^n \setminus [-\delta,\delta]^n} q_l(\mathbf{t}) d\mathbf{t} + \varepsilon \int_{[-\delta,\delta]^n} q_l(\mathbf{t}) d\mathbf{t} \\
&\le 2M 2^n \varepsilon + \varepsilon = \varepsilon(2^{n+1}M + 1),
\end{aligned}$$

woraus die Behauptung folgt. □

Korollar 10.1.5. *Die Menge aller Funktionen vom Typ*

$$f(\mathbf{x}) = \sum_{i=1}^{J} h_{i,1}(x_1) \cdots h_{i,n}(x_n) \tag{10.1}$$

mit $J \in \mathbb{N}$ und $h_{i,k} \in C([a_k,b_k],\mathbb{R})$ liegt dicht in $C(I^n,\mathbb{R})$.

Beweis. Polynome im $\mathbb{R}^n$ sind Funktionen vom Typ (10.1). Da diese nach dem Weierstraßschen Approximationssatz selbst schon dicht in $C(I^n, \mathbb{R})$ liegen, folgt die Behauptung. $\square$

Beispiel 10.1.6. *Die Funktion*

$$f(\mathbf{x}) = e^{\|\mathbf{x}\|^2} = e^{\sum_{i=1}^n x_i^2} = \prod_{i=1}^n e^{x_i^2}$$

kann als Produkt von eindimensionalen Funktionen geschieben werden, wohingegen

$$f(\mathbf{x}) = e^{\|\mathbf{x}\|} = e^{\sqrt{\sum_{i=1}^n x_i^2}}$$

nicht in eindimensionale Funktionen zerfällt.

Satz 10.1.7. *Die Funktion $f : \mathbb{R}^n \to \mathbb{R}$ sei stetig und $I^n \subseteq \mathbb{R}^n$ eine n-Zelle. Die Permutation $\mathbf{T}$ vertausche die Variablen x_i und x_j, das heißt*

$$\mathbf{T}(x_1, \ldots, x_i, \ldots, x_j, \ldots, x_n) = (x_1, \ldots, x_j, \ldots, x_i, \ldots, x_n).$$

Wir setzen $\widehat{f}(\mathbf{x}) := f\big(\mathbf{T}(\mathbf{x})\big)$ und mit $\widehat{I^n} := \mathbf{T}(I_n)$, bezeichnen wir die n-Zelle, die durch entsprechendes Vertauschen der Variablen entsteht. Dann gilt

$$\int_{I^n} f(\mathbf{x})\, d\mathbf{x} = \int_{\widehat{I^n}} \widehat{f}(\mathbf{x})\, d\mathbf{x},$$

mit anderen Worten, die Reihenfolge der Integration ist beliebig.

Beweis. Für jedes $\varepsilon > 0$ existiert nach dem Approximationssatz von Weierstraß 10.1.4 ein Polynom $q_l(\mathbf{x})$ mit

$$|q_l - f| < \varepsilon.$$

Nun ist

$$q_l(\mathbf{x}) = \sum_{i=1}^l \prod_{k=1}^n h_{i,k}(x_k).$$

Bezeichnen wir mit $V(I^n) = \prod_{k=1}^n (b_k - a_k)$ das Volumen der n-Zelle I^n, dann gilt folglich

$$\left| \int_{I^n} q_l(\mathbf{x}) - f(\mathbf{x})\, d\mathbf{x} \right| \leq \int_{I^n} |q_l(\mathbf{x}) - f(\mathbf{x})|\, d\mathbf{x} \leq \varepsilon V(I^n).$$

Daher genügt es, die Aussage für das Polynom q_l zu zeigen.

Da für eine stetige Funktion $h : I^n \to \mathbb{R}$ mit

$$h(\mathbf{x}) = h_1(x_1) \cdot h_2(x_2) \cdots h_n(x_n)$$

die Aussage

$$\begin{aligned}\int_{I^n} h(\mathbf{x})\, d\mathbf{x} &= \int_{a_1}^{b_1} h_1(x_1)dx_1 \cdot \int_{a_2}^{b_2} h_2(x_2)dx_2 \cdots \int_{a_n}^{b_n} h_n(x_n)dx_n \\ &= \int_{\widehat{I^n}} \widehat{h}(\mathbf{x})d\mathbf{x}\end{aligned}$$

gilt, folgt sofort dass

$$\int_{I^n} q_l(\mathbf{x})\, d\mathbf{x} = \int_{\widehat{I^n}} \widehat{q_l}(\mathbf{x})\, d\mathbf{x}.$$

□

Definition 10.1.8. *Im eindimensionalen Fall ist eine Menge $D \subseteq \mathbb{R}$ ein* **Normalbereich***, falls $D = [a, b]$ ein abgeschlossenes Intervall ist. Im höherdimensionalen Fall, das heißt $n > 1$, nennen wir eine kompakte Menge $D \subseteq \mathbb{R}^n$ Normalbereich bezüglich der x_n-Achse, falls ein Normalbereich $E \subseteq \mathbb{R}^{n-1}$ (bezüglich der x_{n-1}-Achse) und zwei stetige Funktionen $G, H : E \to \mathbb{R}$ derart existieren, so dass gilt*

$$D = \left\{\mathbf{x} = \left[\begin{smallmatrix}\mathbf{x}' \\ x_n\end{smallmatrix}\right] \in \mathbb{R}^n : G(\mathbf{x}') \leq x_n \leq H(\mathbf{x}),\ \mathbf{x}' \in E\right\}.$$

Ist $D \subseteq \mathbb{R}^n$ Normalbereich bezüglich aller Achsen, so sprechen wir kurz von einem Normalbereich im $\mathbb{R}^n$.

Beispiel 10.1.9. *1. Der Einheitskreis*

$$D = \{\mathbf{x} \in \mathbb{R}^2 : \|\mathbf{x}\| \leq 1\}$$

ist ein Normalbereich im $\mathbb{R}^2$, denn es gilt sowohl

$$D = \left\{(x, y) \in \mathbb{R}^2 : -\sqrt{1 - x^2} \leq y \leq \sqrt{1 - x^2},\ -1 \leq x \leq 1\right\}.$$

als auch

$$D = \left\{(x, y) \in \mathbb{R}^2 : -\sqrt{1 - y^2} \leq x \leq \sqrt{1 - y^2},\ -1 \leq y \leq 1\right\}.$$

2. Die Menge

$$\triangle^n := \left\{(x_1, \cdots, x_n) \in \mathbb{R}^n : x_i \geq 0\ \forall i \in \{1, \ldots, n\},\ \sum_{i=1}^n x_i \leq 1\right\}$$

ist ein Normalbereich im $\mathbb{R}^n$. Die Menge $\triangle^n$ wird (Referenz-) n-Simplex genannt.

Satz 10.1.10. *Sei $D \subseteq \mathbb{R}^n$ ein Normalbereich, das heißt wir haben*

$$D = \left\{\mathbf{x} = \left[\begin{smallmatrix}\mathbf{x}' \\ x_n\end{smallmatrix}\right] \in \mathbb{R}^n : G(\mathbf{x}') \leq x_n \leq H(\mathbf{x}'),\ \mathbf{x}' \in E\right\}$$

mit einem Normalbereich $E \subseteq \mathbb{R}^{n-1}$ und zwei stetigen Funktionen $G, H : E \to \mathbb{R}$. Ferner sei $I^n \subseteq \mathbb{R}^n$ eine n-Zelle die eine offene Umgebung $U \subseteq D$ von D enthält. Dann ist das Integral einer stetigen Funktion $f : U \to \mathbb{R}$ über D gegeben durch

$$\int_D f(\mathbf{x})\, d\mathbf{x} = \int_E \left[\int_{G(\mathbf{x}')}^{H(\mathbf{x}')} f(\mathbf{x}', x_n)\, dx_n \right] d\mathbf{x}'.$$

Beweis. Für $\delta > 0$ genügend klein und $\mathbf{x}' \in E$ beliebig setzen wir

$$\chi_\delta(\mathbf{x}', x_n) = \begin{cases} 1, & G(\mathbf{x}') \le x_n \le H(\mathbf{x}'), \\ 1 - \frac{x_n - H(\mathbf{x}')}{\delta}, & H(\mathbf{x}') < x_n \le H(\mathbf{x}') + \delta, \\ 1 - \frac{G(\mathbf{x}') - x_n}{\delta}, & G(\mathbf{x}') - \delta \le x_n < G(\mathbf{x}'), \\ 0, & \text{sonst.} \end{cases}$$

Dann gilt

$$\int_{a_n}^{b_n} f(\mathbf{x}', x_n)\chi_\delta(\mathbf{x}', x_n)\, dx_n = \int_{G(\mathbf{x}')}^{H(\mathbf{x}')} f(\mathbf{x}', x_n)\, dx_n + R(\mathbf{x}', \delta)$$

für alle $\mathbf{x}' \in E$, wobei

$$|R(\mathbf{x}', \delta)| \le \int_{G(\mathbf{x}')-\delta}^{G(\mathbf{x}')} f(\mathbf{x}', x_n)\, dx_n + \int_{H(\mathbf{x}')}^{H(\mathbf{x}'+\delta)} f(\mathbf{x}', x_n)\, dx_n \le 2\delta \sup_{\mathbf{x} \in I^n} |f(\mathbf{x})|$$

gleichmäßig bezüglich $\mathbf{x}' \in E$. Dies bedeutet

$$\int_{a_n}^{b_n} f(\mathbf{x}', x_n) \cdot \chi_\delta(\mathbf{x}', x_n)\, dx_n \longrightarrow F(\mathbf{x}') := \int_{G(\mathbf{x}')}^{H(\mathbf{x}')} f(\mathbf{x}', x_n) dx_n$$

gleichmäßig auf E. Die Aussage folgt nun durch Induktion nach n. □

Beispiel 10.1.11. *Sei der Normalbereich $D \subseteq \mathbb{R}^2$ gegeben durch*

$$D = \{(x, y) : x^2 \le y \le x,\ 0 \le x \le 1\}$$

und $f : D \to \mathbb{R}$ mit $f(x, y) = 1$. Dann ist

$$\int_D f(x, y)\, d(x, y) = \int_D 1\, d(x, y) = \int_0^1 \left[\int_{x^2}^x 1\, dy \right] dx = \int_0^1 \left[y\Big|_{x^2}^x \right] dx = \int_0^1 (x - x^2)\, dx = \frac{1}{6}.$$

Es sei angemerkt, das diese Größe dem Flächeninhalt von D entspricht.

Korollar 10.1.12. *Es sei $K \subseteq \mathbb{R}^n$ eine kompakte Menge und $f \in C(K, \mathbb{R})$. Ferner seien Normalbereiche $D_i \subseteq \mathbb{R}^n$, $i = 1, \ldots, N$, gegeben, deren Inneres paarweise disjunkt ist, das heißt*

$$D_i^\circ \cap D_j^\circ = \emptyset, \qquad i \ne j.$$

Falls gilt $K = \bigcup_{i=1}^N D_i \subseteq \mathbb{R}^n$, dann haben wir

$$\int_{K^\circ} f(\mathbf{x})\, d\mathbf{x} = \sum_{i=1}^N \int_{D_i} f(\mathbf{x})\, d\mathbf{x}.$$

Definition 10.1.13. *Sei $E \subseteq \mathbb{R}^n$ und $f : E \to \mathbb{R}$. Die Menge*

$$\operatorname{supp} f := \overline{\{\mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) \neq 0\}}$$

heißt **Träger** *oder* **Support** *der Funktion f.*

Besitzt die Funktion $f \in C(\mathbb{R}^n, \mathbb{R})$ einen kompaktem Träger, das heißt $\operatorname{supp} f$ ist kompakt, dann sei I^n eine n-Zelle die den Träger von f enthält. Wir definieren

$$\int_{\mathbb{R}^n} f(\mathbf{x})\, d\mathbf{x} := \int_{I^n} f(\mathbf{x})\, d\mathbf{x}.$$

Das so definierte Integral ist unabhängig von der Wahl von I^n, vorausgesetzt lediglich, dass I^n den Träger von f enthält. Ist die kompakte Menge $\operatorname{supp} f$ sogar ein Normalbereich, so folgt offensichtlich

$$\int_{\mathbb{R}^n} f(\mathbf{x})\, d\mathbf{x} = \int_{\operatorname{supp} f} f(\mathbf{x})\, d\mathbf{x}.$$

10.2 Die Substitutionsregel im $\mathbb{R}^n$

Um die Substitutionsregel im $\mathbb{R}^n$ beweisen zu können, benötigen wir zuerst einige vorbereitende Überlegungen.

Definition 10.2.1. *Eine Funktion $\mathbf{F} : D \subseteq \mathbb{R}^n \to \mathbb{R}^n$ heißt* **primitiv**, *falls*

$$\mathbf{F}(\mathbf{x}) = \begin{bmatrix} F_1(x_1, \ldots, x_n) \\ \vdots \\ F_n(x_1, \ldots, x_n) \end{bmatrix} = \sum_{\substack{i \in \{1,\ldots,n\}: \\ i \neq m}} x_i \mathbf{e}_i + f(\mathbf{x})\mathbf{e}_m = \begin{bmatrix} x_1 \\ \vdots \\ x_{m-1} \\ f(\mathbf{x}) \\ x_{m+1} \\ \vdots \\ x_n \end{bmatrix}$$

mit $f : D \to \mathbb{R}$.

Wie man leicht überprüft gilt für eine primitive Funktion $\mathbf{F} : D \subseteq \mathbb{R}^n \to \mathbb{R}^n$ die Gleichung

$$\mathbf{F}(\mathbf{x}) = (f(\mathbf{x}) - x_m)\mathbf{e}_m + \sum_{i=1}^{n} x_i \mathbf{e}_i.$$

Aus $\mathbf{F} \in C^1(D, \mathbb{R}^n)$ folgt $f \in C^1(D, \mathbb{R})$, und für $\mathbf{x} \in D$ finden wir

$$\begin{aligned}\mathbf{F}'(\mathbf{x}) &= \mathbf{e}_m\big(f'(\mathbf{x}) - \mathbf{e}_m\big) + \mathbf{I} \\ &= \left[\begin{array}{c|cc} \mathbf{I} & \multicolumn{2}{c}{\mathbf{0}} \\ \hline f_{x_1}(\mathbf{x}) \cdots f_{x_{m-1}}(\mathbf{x_0}) & f_{x_m}(\mathbf{x_0}) & f_{x_{m+1}}(\mathbf{x_0}) \cdots f_{x_n}(\mathbf{x_0}) \\ \hline \multicolumn{2}{c|}{\mathbf{0}} & \mathbf{I} \end{array}\right]\end{aligned}$$

Da offensichtlich gilt

$$\det \mathbf{F}'(\mathbf{x}) = f_{x_m}(\mathbf{x}),$$

ist $\mathbf{F}(\mathbf{x})$ genau dann invertierbar in einer Umgebung von $\mathbf{x}$, falls $f_{x_m}(\mathbf{x}) \neq 0$ ist.

Definition 10.2.2. *Die lineare Abbildung* $\mathbf{P}_m : \mathbb{R}^n \to \mathbb{R}^n$ *mit* $m \leq n$ *definiert durch*

$$\mathbf{P}_m(\mathbf{x}) = \mathbf{P}_m \cdot \left(\sum_{i=1}^{n} x_i \mathbf{e_i}\right) = \mathbf{P}_m \left[x_1, \ldots, x_n\right]^T := \sum_{i=1}^{m} x_i \mathbf{e}_i = \left[x_1, \ldots, x_m, 0, \ldots, 0\right]^T$$

heißt **kanonischer Projektor** *des* $\mathbb{R}^n$ *auf den* $\mathbb{R}^m$.

Die Matrix $\mathbf{P}_m$ des kanonischen Projektors ist gegeben durch

$$\mathbf{P}_m = \begin{bmatrix} \mathbf{I}^{m \times m} & \mathbf{0}^{m \times (n-m)} \\ \mathbf{0}^{(n-m) \times n} & \mathbf{0}^{(n-m) \times (n-m)} \end{bmatrix}.$$

Natürlich ist der kanonische Projektor $\mathbf{P}_n$ die Identität auf dem $\mathbb{R}^n$. Ferner gilt offensichtlich

$$\mathbf{P}_m^2 = \mathbf{P}_m \cdot \mathbf{P}_m = \mathbf{P}_m.$$

Definition 10.2.3. *Die lineare Abbildung* $\mathbf{T}_{i,j} : \mathbb{R}^n \to \mathbb{R}^n$ *definiert durch*

$$\mathbf{T}_{i,j}(\mathbf{x}) = \mathbf{T}_{i,j} \cdot \left[x_1, \ldots, x_i, \ldots, x_j, \ldots, x_n\right]^T = \left[x_1, \ldots, x_j, \ldots, x_i, \ldots, x_n\right]^T,$$

welche die zwei Variablen x_i *und* x_j *miteinander vertauscht, nennen wir einen* **Permutationsoperator**.

Als Matrix $\mathbf{T}_{i,j}$ ist dieser Permutationsoperator wie folgt gegeben

$$\mathbf{T}_{i,j} = \begin{array}{cl} \begin{bmatrix} 1 & & & & & & & & & & \\ & \ddots & & & & & & & & & \\ & & 1 & & & & & & & & \\ & & & 0 & & & & 1 & & & \\ & & & & 1 & & & & & & \\ & & & & & \ddots & & & & & \\ & & & & & & 1 & & & & \\ & & & 1 & & & & 0 & & & \\ & & & & & & & & 1 & & \\ & & & & & & & & & \ddots & \\ & & & & & & & & & & 1 \end{bmatrix} & \begin{array}{l} \\ \\ \\ \leftarrow i\text{-te Zeile} \\ \\ \\ \\ \leftarrow j\text{-te Zeile} \\ \\ \\ \\ \end{array} \end{array}.$$

Selbstverständlich folgt $\mathbf{T}_{i,i} = \mathbf{I}$ und, wie man leicht überprüft, gilt die Beziehung

$$\mathbf{T}_{i,j}^2 = \mathbf{T}_{i,j} \cdot \mathbf{T}_{i,j} = \mathbf{I}.$$

Lemma 10.2.4. *Sei $U \subseteq \mathbb{R}^n$ eine offene Umgebung von $\mathbf{0}$. Die Funktion $\mathbf{F} = [F_1(\mathbf{x}), \ldots, F_n(\mathbf{x})]^T \in C^1(U, \mathbb{R}^n)$ erfülle*

$$\mathbf{F}(\mathbf{0}) = \mathbf{0}$$

und $\det \mathbf{F}'(\mathbf{0}) \neq 0$. Ferner gelte

$$\mathbf{P}_{m-1}\mathbf{F}(\mathbf{x}) = \mathbf{P}_{m-1}\mathbf{x}$$

wobei $\mathbf{P}_m$ die kanonische Projektion auf den $\mathbb{R}^m$ bezeichne.

Dann existieren Umgebungen $V, W \subseteq U$ mit $\mathbf{0} \in V, W$ und eine invertierbare primitive Abbildung $\mathbf{G} \in C^1(V, W)$ mit $\mathbf{G}(\mathbf{0}) = \mathbf{0}$, so so dass für eine Permutaionsmatrix $\mathbf{T}$ die Funktion $\mathbf{H} : W \to \mathbb{R}^n$ definiert durch

$$\mathbf{H}(\mathbf{x}) = \mathbf{T}(\mathbf{F} \circ \mathbf{G}^{-1})(\mathbf{x})$$

die Gleichung

$$\mathbf{P}_m\mathbf{H}(\mathbf{x}) = \mathbf{P}_m\mathbf{x}$$

auf W erfüllt.

Beweis. Aufgrund der Voraussetzung besitzt $\mathbf{F}$ die Darstellung

$$\mathbf{F}(\mathbf{x}) = \sum_{i=1}^{m-1} x_i\mathbf{e}_i + \sum_{i=m}^{n} F_i(\mathbf{x})\mathbf{e}_i.$$

Es gilt

$$\mathbf{F}(\mathbf{x})' = \left[\begin{array}{c|c} \mathbf{I}^{m\times m} & \mathbf{0}^{m\times(n-m)} \\ \hline \multicolumn{2}{c}{F_m'(\mathbf{x})} \\ \multicolumn{2}{c}{\vdots} \\ \multicolumn{2}{c}{F_n'(\mathbf{x})} \end{array}\right].$$

Da $\mathbf{F}'(\mathbf{0})$ invertierbar ist, gilt

$$\mathbf{0} \neq \mathbf{F}'(\mathbf{0})\mathbf{e}_m = \sum_{i=m}^{n} \frac{\partial F_i}{\partial x_m}(\mathbf{0})\mathbf{e}_i.$$

Somit existiert ein Index $k \geq m$ derart, dass

$$\frac{\partial F_k}{\partial x_m}(\mathbf{0}) \neq 0.$$

Es bezeichne $\mathbf{T} = \mathbf{T}_{m,k}$ diejenige Permutationsmatrix, die die Variablen x_m und x_k vertauscht. Setzen wir $\widehat{\mathbf{F}} = (\widehat{F}_1, \ldots, \widehat{F}_n) := \mathbf{TF}$, dann folgt

$$\frac{\partial \widehat{F}_m}{\partial x_m}(\mathbf{0}) \neq 0.$$

Definieren wir nun $\mathbf{G} : U \to \mathbb{R}^n$ mit

$$\mathbf{G}(\mathbf{x}) := \big(F_k(\mathbf{x}) - x_k\big)\mathbf{e}_k + \sum_{i=1}^{n} x_i\mathbf{e}_i,$$

so ist $\mathbf{G}$ primitiv. Folglich ergibt sich

$$\widehat{\mathbf{G}}(\mathbf{x}) := \mathbf{TG}(\mathbf{x}) = \big(\widehat{F}_m(\mathbf{x}) - x_m\big)\mathbf{e}_m + \sum_{i=1}^{n} x_i\mathbf{e}_i$$

und es gilt

$$\det \mathbf{G}'(\mathbf{0}) = \det \widehat{\mathbf{G}}'(\mathbf{0}) = \frac{\partial \widehat{F}_m}{\partial x_n}(\mathbf{0}) \neq 0.$$

Da gilt $\mathbf{G}(\mathbf{0}) = \mathbf{0}$ existieren offene infolge des Satzes über inverse Funktionen Umgebungen $V, W \subseteq U$ mit $\mathbf{0} \in V, W$ und $W \subseteq R^n$ derart, dass $\mathbf{G} : V \to W$ bijektiv ist mit $\mathbf{G}^{-1} \in C^1(W, V)$.

Für $\mathbf{y} \in V$ sei $\mathbf{x} = \mathbf{G}^{-1}(\mathbf{y}) \in V$, dann definieren wir die Funktion $\mathbf{H} : W \to \mathbb{R}^n$ vermittels

$$\mathbf{H}(\mathbf{y}) := \mathbf{T}(\mathbf{F} \circ \mathbf{G}^{-1})(\mathbf{y}) = \mathbf{TF}(\mathbf{x}).$$

Dann gilt

$$\begin{aligned}
\mathbf{P}_m\mathbf{H}(\mathbf{y}) &= \mathbf{P}_m\mathbf{T}\mathbf{F}(\mathbf{x}) \\
&= \mathbf{P}_m\left[\mathbf{P}_{m-1}\mathbf{x} + \widehat{F}_m(\mathbf{x})\mathbf{e}_m + \sum_{i>m}\widehat{F}_i(\mathbf{x})\mathbf{e}_i\right] \\
&= \mathbf{P}_{m-1}\mathbf{x} + F_k(\mathbf{x})\mathbf{e}_m + \mathbf{0} \\
&= \mathbf{P}_m\mathbf{x} + \big(F_k(\mathbf{x}) - x_m\big)\mathbf{e}_m \\
&= \mathbf{P}_m\mathbf{G}(\mathbf{x}),
\end{aligned}$$

dies bedeutet $\mathbf{P}_m\mathbf{H} = \mathbf{P}_m$ auf V. □

Satz 10.2.5. *Sei $\mathbf{0} \in U \subseteq \mathbb{R}^n$ offen. Die Funktion $\mathbf{F} \in C^1(U, \mathbb{R}^n)$ erfülle*

$$\mathbf{F}(\mathbf{0}) = \mathbf{0}$$

und

$$\det \mathbf{F}'(\mathbf{0}) \neq 0.$$

Dann existieren Permutationsmatrizen $\mathbf{T}_1, \dots, \mathbf{T}_{n-1}$ und primitive Abbildungen $\mathbf{G}_1, \dots, \mathbf{G}_n \in C^1(V, \mathbb{R}^n)$ mit einer geeigneten offenen Menge $V \subseteq U$ mit $\mathbf{0} \in V$ mit folgenden Eigenschaften:

1. *die Abbildungen $\mathbf{G}_i$ sind injektiv auf V mit $\mathbf{G}_i(\mathbf{0}) = \mathbf{0}$, $i \in \{1, \dots, n\}$,*
2. *es gilt $\det \mathbf{G}_i'(\mathbf{0}) \neq 0$, $i \in \{1, \dots, n\}$, und*
3. *die Funktion $\mathbf{F}$ besitzt für alle $\mathbf{x} \in V$ die Darstellung*

$$\mathbf{F}(\mathbf{x}) = (\mathbf{T}_1 \cdots \mathbf{T}_{n-1})(\mathbf{G}_n \circ \cdots \circ \mathbf{G}_1)(\mathbf{x}).$$

Beweis. Nach Lemma 10.2.4 gibt es offene Umgebungen $U_1, V_1 \subseteq U$ mit $\mathbf{0} \in U_1, V_1$, und eine invertierbare primitive Funktion $\mathbf{G}_1 \in C^1(V_1, U_1)$ mit $\mathbf{G}_1(\mathbf{0}) = \mathbf{0}$ und $\det \mathbf{G}_1'(\mathbf{0}) \neq 0$ sowie eine Permutationsmatrix $\mathbf{T}_1$, so dass $\mathbf{F}_1 \in C^1(U_1, \mathbb{R}^n)$ definiert durch

$$\mathbf{F}_1(\mathbf{x}) = \mathbf{T}_1(\mathbf{F} \circ \mathbf{G}_1^{-1})(\mathbf{x})$$

die Gleichungen $\mathbf{F}_1(\mathbf{0}) = \mathbf{0}$, $\det \mathbf{F}_1'(\mathbf{0}) \neq 0$ und

$$\mathbf{P}_1\mathbf{F}_1(\mathbf{x}) = \mathbf{x}, \qquad \mathbf{x} \in U_1,$$

erfüllt. Da $\mathbf{T}_1^2 = \mathbf{I}$ folgt die Identität

$$\mathbf{F}(\mathbf{x}) = \mathbf{T}_1(\mathbf{F}_1 \circ \mathbf{G}_1)(\mathbf{x}).$$

Wenden wir nun Lemma 10.2.4 auf $\mathbf{F}_1$ an, so finden wir wieder offene Umgebungen $U_2, V_2 \subseteq U_1$ mit $\mathbf{0} \in U_2, V_2$, und eine invertierbare primitive Funktion $\mathbf{G}_2 \in C^1(V_2, U_2)$ mit $\mathbf{G}_2(\mathbf{0}) = \mathbf{0}$ und $\det \mathbf{G}_2'(\mathbf{0}) \neq 0$ sowie eine Permutationsmatrix $\mathbf{T}_2$, so dass $\mathbf{F}_2 \in C^1(U_2, \mathbb{R}^n)$ definiert durch

$$\mathbf{F}_2(\mathbf{x}) = \mathbf{T}_2(\mathbf{F}_1 \circ \mathbf{G}_2^{-1})(\mathbf{x})$$

die Gleichungen $\mathbf{F}_2(\mathbf{0}) = \mathbf{0}$, $\det \mathbf{F}_2'(\mathbf{0}) \neq 0$ und

$$\mathbf{P}_2\mathbf{F}_2(\mathbf{x}) = \mathbf{x}, \qquad \mathbf{x} \in U_2,$$

erfüllt. Bisher haben wir also auf U_2 die Identiät

$$\mathbf{F}(\mathbf{x}) = \mathbf{T}_1\mathbf{T}_2(\mathbf{F}_2 \circ \mathbf{G}_2 \circ \mathbf{G}_1)(\mathbf{x}).$$

Wiederholtes Anwenden von Lemma 10.2.4 liefert uns mit $\mathbf{G}_n := \mathbf{F}_{n-1}$ also eine Umgebung $V := U_{n-1} \subseteq U$, für die offensichtlich das Behauptete gilt. □

Satz 10.2.6 (Zerlegung der Eins). *Sei $K \subseteq \mathbb{R}^n$ kompakt und $\{V_i\}$ eine endliche Überdeckung von K. Dann existieren $\psi_1, \ldots, \psi_L \in C(K, \mathbb{R})$ mit den folgenden Eigenschaften:*

1. *Zu jedem ψ_i existiert ein V_j mit $\operatorname{supp} \psi_i \subseteq V_j$.*
2. *Für jedes $\mathbf{x} \in K$ gilt*
$$0 \leq \psi_i(\mathbf{x}) \leq 1, \qquad i = 1, \ldots, L,$$
und
$$\sum_{i=1}^{L} \psi_i(\mathbf{x}) = 1.$$

Beweis. Zu jedem $\mathbf{x} \in K$ gibt es einen Index $j \in \{1, \ldots, n\}$ mit $\mathbf{x} \in V_j$. Ferner existiert eine offene Umgebung $U(\mathbf{x})$ mit $\mathbf{x} \in U(\mathbf{x})$ und $\overline{U(\mathbf{x})} \subseteq V_j$. Es folgt

$$K \subseteq \bigcup_{\mathbf{x} \in K} U(\mathbf{x}).$$

Da K kompakt ist, existieren $\mathbf{x}_1, \ldots, \mathbf{x}_L \in K$ mit

$$K \subseteq \bigcup_{i=1}^{L} U(\mathbf{x}_i).$$

Wir definieren zunächst zu jedem $\mathbf{x}_i$ eine Funktion $\varphi_i \in C(\mathbb{R}^n, \mathbb{R})$ mit

$$\varphi_i(\mathbf{x}) = \begin{cases} 1, & \mathbf{x} \in U(\mathbf{x}_i), \\ 0, & \mathbf{x} \notin V_j, \end{cases}$$

und $0 \leq \varphi_i(\mathbf{x}) \leq 1$ für $\mathbf{x} \in V_j \setminus U(\mathbf{x}_i)$. Für jedes $i \in \{1, \ldots, L\}$ setzen wir nun

$$\psi_i(\mathbf{x}) := \frac{\varphi_i(\mathbf{x})}{\sum_{i=1}^{L} \varphi_i(\mathbf{x})}.$$

Da alle Funktionen φ_i auf der Menge K nichtnegativ sind und für beliebiges $\mathbf{x} \in K$ ein Index $i \in \{1, \ldots, L\}$ existiert mit $\mathbf{x} \in U(\mathbf{x}_i)$, ist $\sum_{i=1}^{L} \varphi_i(\mathbf{x}) > 0$ und daher gilt

$$0 \leq \psi_i(\mathbf{x}) \leq 1$$

und

$$\sum_{i=1}^{L} \psi_i(\mathbf{x}) = \frac{\sum_{i=1}^{L} \varphi_i(\mathbf{x})}{\sum_{i=1}^{L} \varphi_i(\mathbf{x})} = 1.$$

Ferner ist $\psi_i \in C(K, \mathbb{R})$ für jedes $i \in \{1, \ldots, L\}$. □

Satz 10.2.7 (Substitutionsregel für n-dimensionale Integration). *Sei $\mathbf{\Phi} : D \to E$ injektiv mit $\det \mathbf{\Phi}'(\mathbf{x}) \neq 0$ auf D. Ist $f : \mathbb{R}^n \to \mathbb{R}$ eine stetige Funktion, deren Träger kompakt ist und in $\Phi(D)$ liegt, dann gilt*

$$\int_{\mathbb{R}^n} f(\mathbf{y}) d\mathbf{y} = \int_{\mathbb{R}^n} f\big(\mathbf{\Phi}(\mathbf{x})\big) |\det \mathbf{\Phi}'(\mathbf{x})| d\mathbf{x}.$$

Beweis.

1. Beachten wir im Falle $n = 1$ die Gleichung

$$|\det \mathbf{\Phi}'(\mathbf{x})| = |\mathbf{\Phi}'(\mathbf{x})|$$

und die Reihenfolge der Grenzen, dann entspricht die Aussage der eindimensionalen Substitutionsregel (Satz 7.4.7).

2. Falls $\mathbf{\Phi}(\mathbf{x}) = \mathbf{T}(\mathbf{x})$ eine Permutation ist, so folgt die Aussage unmittelbar aus Satz 10.1.7.

3. Falls

$$\mathbf{\Phi}(\mathbf{x}) = \mathbf{G}(\mathbf{x}) = [x_1, \ldots, x_{m-1}, G_m(\mathbf{x}), x_{m+1}, \ldots, x_n]^T$$

eine primitive Funktion ist, dann kann man zuerst die Integration über die Variablen x_i, $i \neq m$, ausführen. Übrig bleibt ein eindimensionales Integral und hierfür ist die Aussage bewiesen, denn es gilt

$$|\det \mathbf{G}'(\mathbf{x})| = \left| \frac{\partial G_m}{\partial x_m}(\mathbf{x}) \right|.$$

4. Bezeichnen $\mathbf{\Theta}$ und $\mathbf{\Psi}$ Permutationen oder primitive Abbildungen, dann finden wir aufgrund der Kettenregel

$$\det \mathbf{\Phi}'(\mathbf{x}) = \det(\mathbf{\Theta} \circ \mathbf{\Psi})'(\mathbf{x}) = \det \Big(\mathbf{\Theta}'\big(\mathbf{\Psi}(\mathbf{x})\big) \mathbf{\Psi}'(\mathbf{x}) \Big) = \det \mathbf{\Theta}'\big(\mathbf{\Psi}(\mathbf{x})\big) \det \mathbf{\Psi}'(\mathbf{x}).$$

Mit den letzten beiden Fällen ergibt sich daher die Behauptung

$$\begin{aligned}
\int_{\mathbb{R}^n} f(\mathbf{z})\, d\mathbf{z} &= \int_{\mathbb{R}^n} f\big(\mathbf{\Theta}(\mathbf{y})\big) \big|\det \mathbf{\Theta}'(\mathbf{y})\big|\, d\mathbf{y} \\
&= \int_{\mathbb{R}^n} f\Big(\mathbf{\Theta}\big(\mathbf{\Psi}(\mathbf{x})\big)\Big) \big|\det \mathbf{\Theta}'\big(\mathbf{\Psi}(\mathbf{x})\big)\big| |\det \mathbf{\Psi}'(\mathbf{x})|\, d\mathbf{x} \\
&= \int_{\mathbb{R}^n} f\big(\mathbf{\Phi}(\mathbf{x})\big) |\det \mathbf{\Phi}'(\mathbf{x})|\, d\mathbf{x}.
\end{aligned}$$

5. Für jeden Punkt $\mathbf{x}_0 \in \operatorname{supp} f$ existiert eine geeignete offene Umgebung $U(\mathbf{x}_0)$ in der $\mathbf{\Phi}$ gemäß Satz 10.2.5 die Darstellung

$$\mathbf{\Phi}(\mathbf{x}) - \mathbf{\Phi}(\mathbf{x}_0) = \mathbf{T}_1 \cdots \mathbf{T}_{n-1}(\mathbf{G}_n \circ \cdots \circ \mathbf{G}_1)(\mathbf{x})$$

besitzt, mit anderen Worten

$$\mathbf{\Phi}(\mathbf{x}) = \mathbf{T}_1 \cdots \mathbf{T}_{n-1}(\mathbf{G}_n \circ \cdots \circ \mathbf{G}_1)(\mathbf{x} - \mathbf{x}_0),$$

wobei $\mathbf{T}_1, \ldots, \mathbf{T}_{n-1}$ Permutationsmatrizen und $\mathbf{G}_1, \ldots, \mathbf{G}_n$ primitiven Abbildungen sind.

Da nach Voraussetzung die Menge $K := \operatorname{supp} f$ kompakt ist, existieren $\{\mathbf{x}_1, \ldots, \mathbf{x}_M\} \subseteq K$, so dass $\bigcup_{i=1}^M U(\mathbf{x}_i) \supseteq K$ eine Überdeckung von K ist. Nach Satz 10.2.6 existiert eine zugehörige Zerlegung der Eins $\psi_1, \ldots, \psi_L \in C(\mathbb{R}^n, \mathbb{R})$. Dann gilt $\psi_i \cdot f \in C(\mathbb{R}^n, \mathbb{R})$ mit $\operatorname{supp}(\psi_i \cdot f) \subseteq U(\mathbf{x}_i)$ und $f = \sum_{i=1}^L \psi_i \cdot f$. Somit folgt mit dem bereits Bewiesenen

$$\begin{aligned}
\int_{\mathbb{R}^n} f(\mathbf{y})\, d\mathbf{y} &= \sum_{i=1}^M \int_{\mathbb{R}^n} (\psi_i \cdot f)(\mathbf{y})\, d\mathbf{y} \\
&= \sum_{i=1}^M \int_{\mathbb{R}^n} (\psi_i \cdot f)\big(\mathbf{\Phi}(\mathbf{x})\big) |\det \mathbf{\Phi}'(\mathbf{x})|\, d\mathbf{x} \\
&= \int_{\mathbb{R}^n} f\big(\mathbf{\Phi}(\mathbf{x})\big) |\det \mathbf{\Phi}'(\mathbf{x})| d\mathbf{x}.
\end{aligned}$$

□

Bemerkung 10.2.8. *Die Determinante* $\det \mathbf{\Phi}'(\mathbf{x})$ *heißt auch* **Jakobi-Determinante**.

Beispiel 10.2.9. 1. *Zwei linear unabhängige Vektoren* $\mathbf{u}, \mathbf{v} \in \mathbb{R}^2$ *definieren ein Parallelogramm* E *via*

$$E = \{\alpha \mathbf{u} + \beta \mathbf{v} : 0 \leq \alpha, \beta \leq 1\},$$

dessen Fäche $V(E)$ *gegeben ist durch*

$$V(E) = \int_E d\mathbf{x} = |\det([\mathbf{u}, \mathbf{v}])|.$$

2. *Für drei linear unabhängige Vektoren* $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^3$ *definiert*

$$E = \{\alpha \mathbf{u} + \beta \mathbf{v} + \gamma \mathbf{w} \in \mathbb{R}^3 : 0 \leq \alpha, \beta, \gamma \leq 1\}$$

einen Spat. Sein Volumen $V(E)$ *ist gegeben durch*

$$V(E) = \int_E d\mathbf{x} = |\det([\mathbf{u}, \mathbf{v}, \mathbf{w}])|.$$

3. *Wir betrachten* **Polarkoordinaten** *in* $\mathbb{R}^2$*, das heißt* $\mathbf{\Phi}$ *ist definiert durch*

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} r\cos\phi \\ r\sin\phi \end{bmatrix} = \mathbf{\Phi}(r,\varphi)$$

mit $0 \le r < \infty$ *und* $0 \le \varphi \le 2\pi$*. Wie man leicht nachrechnet gilt*

$$\det \mathbf{\Phi}(r,\varphi) = \begin{vmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{vmatrix} = r$$

für alle $0 \le r < \infty$ *und* $0 \le \varphi \le 2\pi$*. Allerdings ist die Abbildung* $\Phi(r,\phi)$ *nicht mehr injektiv für* $r = 0$*, da* $\Phi(0,\phi) \equiv 0$*. Ferner gilt auch* $\Phi(r,0) = \Phi(r,2\pi)$ *für alle* $0 \le r < \infty$*. Daher können wir Satz 10.2.7 im allgemeinen nur mittels eines Grenzpozesses anwenden.*

Sei

$$D = \{(r,\varphi) : 0 \le r \le R,\ 0 \le \varphi \le 2\pi\}$$

dann ist

$$E := \mathbf{\Phi}(D) = \{(x,y) : x^2 + y^2 \le R^2\}$$

die Kreisscheibe mit Radius R*.*

Nach der Substitutionsregel im $\mathbb{R}^n$ *ergibt sich*

$$\begin{aligned} V(E) = \int_E d\mathbf{x} &= \int_D |\det \mathbf{\Phi}'(r,\varphi)|\ dr\ d\varphi \\ &= \lim_{\varepsilon\to 0} \int_\varepsilon^R \left[\lim_{\delta\to 0} \int_0^{2\pi-\delta} |\det \mathbf{\Phi}'(r,\varphi)\ d\varphi| \right] dr \\ &= \lim_{\varepsilon\to 0} \int_\varepsilon^R \left[\lim_{\delta\to 0} \int_0^{2\pi-\delta} r\ d\varphi \right] dr \\ &= \lim_{\varepsilon\to 0} \int_\varepsilon^R \left[\lim_{\delta\to 0} r\varphi \Big|_0^{2\pi-\delta} \right] dr \\ &= \lim_{\varepsilon\to 0} \int_\varepsilon^R 2\pi r\ dr \\ &= \lim_{\varepsilon\to 0} 2\pi \frac{r^2}{2}\Big|_\varepsilon^R = \pi R^2. \end{aligned}$$

Bemerkung 10.2.10. *Man beachte die Gleichheit*

$$\int_D f\big(\mathbf{\Phi}(\mathbf{x})\big) |\det \mathbf{\Phi}'(\mathbf{x})| d\mathbf{x} = \int_{\mathbf{\Phi}(D)} f(\mathbf{x}) d\mathbf{x}.$$

Kapitel 11

Differentialformen

11.1 Grundlegende Definitionen

Definition 11.1.1. *Sei $E \subseteq \mathbb{R}^n$ eine offene Menge und $D \subseteq \mathbb{R}^k$ eine kompakte Menge. Ist $\mathbf{\Phi} \in C^1(D, E)$, dann nennen wir $(\mathbf{\Phi}, D)$ eine k-***Fläche** *in E.*

Die Menge D nennen wir den **Parameterbereich** *von $\mathbf{\Phi}$. Die Funktion $\mathbf{\Phi}$ wird* **Parametrisierung** *der Menge $\mathbf{\Phi}(D)$ genannt.*

Bemerkung 11.1.2. *Die Parametrisierung einer Menge $\mathbf{\Phi}(D)$ ist nicht eindeutig. Denn betrachten wir eine Selbstabbildung $\mathbf{\Psi} \in C^1(D, D)$, dann ist*

$$(\mathbf{\Phi} \circ \mathbf{\Psi})(D) = \mathbf{\Phi}\big(\mathbf{\Psi}(D)\big) = \mathbf{\Phi}(D).$$

Beispiel 11.1.3. *Ist $\mathbf{\Phi} \in C^1([0,1], \mathbb{R}^3)$, dann definiert die Einsfläche $\big(\mathbf{\Phi}, [0,1]\big)$ eine* **Kurve** *im $\mathbb{R}^3$. Ist $\mathbf{\Phi} \in C(I^2, \mathbb{R}^3)$ mit $I^n = [a_1, b_1] \times [a_2, b_2] \subseteq \mathbb{R}^2$, dann definiert die 2-Fläche $(\mathbf{\Phi}, I_2)$ ein* **gekrümmtes Flächenstück** *im $\mathbb{R}^3$.*

Definition 11.1.4. *Sei $E \subseteq \mathbb{R}^n$ offen. Eine* **Differentialform** *der Ordnung $k \geq 1$ in E oder kurz k-***Form** *in E ist ein Funktional ω, das symbolisch durch die Summe*

$$\omega = \sum_{i_1,\ldots,i_k=1}^{n} a_{i_1\cdots i_k}(\mathbf{x})\; dx_{i_1} \wedge \ldots \wedge dx_{i_k}$$

dargestellt wird, das jeder k-Fläche $(\mathbf{\Phi}, D)$ in E eine Zahl

$$\omega(\mathbf{\Phi}) = \int_{\mathbf{\Phi}} \omega := \int_D \sum a_{i_1\cdots i_k}\big(\mathbf{\Phi}(\mathbf{u})\big) \det\left(\frac{\partial(x_{i_1}, \ldots, x_{i_k})}{\partial(u_1, \ldots, u_k)}(\mathbf{u})\right)\; d\mathbf{u} \tag{11.1}$$

zuordnet. Hierbei bezeichnen $a_{i_1\cdots i_k} : E \to \mathbb{R}$ stetige Funktionen. Gilt $\omega = a(\mathbf{x})$ mit einer stetigen Funktion $a : E \to \mathbb{R}^n$, so sprechen wir von einer **Nullform**.

Bemerkung 11.1.5. *Sind $\Phi_1, \ldots, \Phi_n$ die Komponenten von $\mathbf{\Phi}$ und ist die Funktion $\widehat{\mathbf{\Phi}}$ definiert vermittels*

$$\widehat{\mathbf{\Phi}}(\mathbf{u}) := \begin{bmatrix} \Phi_{i_1}(\mathbf{u}) \\ \vdots \\ \Phi_{i_k}(\mathbf{u}) \end{bmatrix},$$

so ist die in (11.1) *auftretende Determinate gegeben durch*

$$\det\left(\frac{\partial(x_{i_1}, \ldots, x_{i_k})}{\partial(u_1, \ldots, u_k)}(\mathbf{u})\right) = \det\left(\frac{\partial(\Phi_{i_1}, \ldots, \Phi_{i_k})}{\partial(u_1, \ldots, u_k)}(\mathbf{u})\right) = \det \widehat{\mathbf{\Phi}}'(\mathbf{u}).$$

Beispiel 11.1.6. *Es sei $k = n = 1$ und $E = \mathbb{R}$, dann ist $\omega = a(x)dx$ mit $a \in C(E, \mathbb{R})$ eine Einsform in $\mathbb{R}$. Ist $D := [a, b] \subseteq \mathbb{R}$ und ist die Funktion $\Phi : D \to E$ stetig differenzierbar und injektiv, dann gilt*

$$\omega(\Phi) = \int_\Phi a(x)dx = \int_D a\big(\Phi(x)\big)\Phi'(x)dx = \int_{\Phi(a)}^{\Phi(b)} a(y)dy.$$

11.2 Volumen- und Oberflächenintegrale

Wir geben nun eine Reihe von Beispielen an, die in den Anwendungen des mehrdimensionalen Integrationskalküls bedeutend sind und die unsere späteren Untersuchungen motivieren sollen.

Beispiel 11.2.1 (Kurvenintegrale im $\mathbb{R}^3$). *Es sei $k = 1$ und $n = 3$ sowie $D = [a, b]$ und $E = \mathbb{R}^3$. Für eine stetige Funktion $\mathbf{f} = [f_1, f_2, f_3]^T : \mathbb{R}^3 \to \mathbb{R}^3$ definiert ω gegeben durch*

$$\omega := f_1(\mathbf{x})dx + f_2(\mathbf{x})dy + f_3(\mathbf{x})dz$$

eine Einsform in $\mathbb{R}^3$.

Ist die Funktion $\mathbf{\Phi} = [\Phi_1, \Phi_2, \Phi_3]^T : D \to \mathbb{R}^3$ stetig differenzierbar und injektiv, dann ist $\Gamma := \mathbf{\Phi}(D)$ ein **Kurvenstück** *im $\mathbb{R}^3$.*

Durch die Menge

$$T = \{\mathbf{\Phi}(t_0) + \mathbf{\Phi}'(t_0)(t - t_0) : t \in \mathbb{R}\} \subseteq \mathbb{R}^3$$

ist die Tangente T an Γ im Punkt $\mathbf{\Phi}(t_0)$ gegeben. Der Vektor

$$\mathbf{T}(t_0) = \mathbf{\Phi}'(t_0) = \begin{bmatrix} \Phi_1'(t_0) \\ \Phi_2'(t_0) \\ \Phi_3'(t_0) \end{bmatrix}$$

definiert den **Tangentenvektor** *in $\mathbf{\Phi}(t_0)$ an die Kurve Γ. Normieren wir diesen Vektor, erhalten wir den der* **Tangenteneinheitsvektor** *an dieser Stelle*

$$\mathbf{t}(t_0) = \frac{\mathbf{T}(t_0)}{\|\mathbf{T}(t_0)\|}.$$

Wegen

$$\mathbf{x}(t) = \begin{bmatrix} x(t) \\ y(t) \\ z(t) \end{bmatrix} = \begin{bmatrix} \Phi_1(t) \\ \Phi_2(t) \\ \Phi_3(t) \end{bmatrix} = \mathbf{\Phi}(t)$$

gilt

$$\frac{\partial \mathbf{x}(t)}{\partial t} = \begin{bmatrix} \frac{\partial x}{\partial t}(t) \\ \frac{\partial y}{\partial t}(t) \\ \frac{\partial z}{\partial t}(t) \end{bmatrix} = \begin{bmatrix} \frac{\partial \Phi_1}{\partial t}(t) \\ \frac{\partial \Phi_2}{\partial t}(t) \\ \frac{\partial \Phi_3}{\partial t}(t) \end{bmatrix} = \frac{\partial \mathbf{\Phi}(t)}{\partial t},$$

und daher folgt

$$\begin{aligned} \omega(\mathbf{\Phi}) = \int_{\mathbf{\Phi}} \omega &= \int_{\mathbf{\Phi}} f_1(\mathbf{x})dx + f_2(\mathbf{x})dy + f_3(\mathbf{x})dz \\ &= \int_D f_1\big(\mathbf{\Phi}(t)\big)\Phi_1'(t) + f_2\big(\mathbf{\Phi}(t)\big)\Phi_2'(t) + f_3\big(\mathbf{\Phi}(t)\big)\Phi_3'(t)dt \\ &= \int_D \langle \mathbf{f}\big(\mathbf{\Phi}(t)\big), \mathbf{\Phi}'(t)\rangle dt \\ &= \int_D \langle \mathbf{f}\big(\mathbf{\Phi}(t)\big), \mathbf{t}(t)\rangle \|\mathbf{T}(t)\| dt. \end{aligned}$$

Führen wir das **Bogenlängenelement**

$$ds := \|\mathbf{\Phi}'(t)\| dt$$

ein, so lässt sich vereinfacht schreiben

$$\omega(\mathbf{\Phi}) = \int_\Gamma \langle \mathbf{f}(\mathbf{x}), \mathbf{t}(\mathbf{x})\rangle ds.$$

Wählen wir in diesem Beispiel speziell $\mathbf{f}(\mathbf{x}) := \mathbf{t}(\mathbf{x})$ auf Γ, dann ist

$$\omega(\mathbf{\Phi}) = \int_\Gamma \langle \mathbf{t}(\mathbf{x}), \mathbf{t}(\mathbf{x})\rangle ds = \int_D \|\Phi'(t)\| dt = \int_a^b \|\Phi'(t)\| dt$$

die **Bogenlänge** des Kurvenstücks Γ.

Wie wir gesehen haben ist das **Kurvenintegral**

$$\int_\Gamma \langle \mathbf{f}(\mathbf{x}), \mathbf{t}(\mathbf{x})\rangle ds = \int_{\mathbf{\Phi}} f_1(\mathbf{x})dx + f_2(\mathbf{x})dy + f_3(\mathbf{x})dz$$

ein Integral über eine spezielle Einsform.

Solche Kurvenintegrale sind in der Physik von großer Bedeutung. Sei zum Beispiel $\mathbf{f} : \mathbb{R}^3 \to \mathbb{R}^3$ ein Kraftfeld, das heißt $\mathbf{f}(\mathbf{x})$ beschreibt die Krafteinwirkung, die auf einen Massepunkt m an einer Stelle $\mathbf{x} \in \mathbb{R}$ wirkt. Bewegt man diesen Massepunkt entlang einem Kurvenstück Γ, dann ist

$$E = \int_\Gamma \langle \mathbf{f}(\mathbf{x}), \mathbf{t}(\mathbf{x})\rangle ds$$

die aufgewandte Arbeit bzw. Energie, die hierzu benötigt wird.

Beispiel 11.2.2 (spezielle Einsformen im $\mathbb{R}^2$).

1. *Wir betrachten den Fall $k = 1$, $n = 2$ und $E = \mathbb{R}^2$. Die Differentialform*

$$\omega_1 := xdy$$

ist von der Ordnung 1. Betrachten wir $\mathbf{\Phi}(t) : D := [0, 2\pi] \to E$ definiert durch

$$\mathbf{\Phi}(t) = \begin{bmatrix} r\cos t \\ r\sin t \end{bmatrix},$$

dann ist

$$\Gamma = \mathbf{\Phi}(D) = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = r^2\}$$

eine Kreislinie. Aufgrund von

$$\mathbf{\Phi}'(t) = \begin{bmatrix} -r\sin t \\ r\cos t \end{bmatrix}$$

finden wir

$$\omega_1(\mathbf{\Phi}) = \int_{\mathbf{\Phi}} \omega_1 = \int_{\Phi} xdy = \int_0^{2\pi} (r\cos t)^2 dt = r^2 \int_0^{2\pi} \cos^2 t dt = \frac{\pi}{2} r^2$$

Dagegen gilt für die Differentialform der Ordnung 1 gegeben durch

$$\omega_2 := ydx$$

die Gleichung

$$\omega_2(\mathbf{\Phi}) = \int_{\mathbf{\Phi}} \omega_2 = \int_{\Phi} ydx = -\int_0^{2\pi} (r\sin t)^2 dt = -r^2 \int_0^{2\pi} \sin^2 t dt = -\frac{\pi}{2} r^2.$$

Zusammengefasst bedeutet dies, dass die Differentialform

$$\omega = \frac{1}{2}(\omega_1 - \omega_2) = \frac{1}{2}(xdy - ydx)$$

die Größe $\omega(\mathbf{\Phi}) = \pi r^2$ der eingeschlossenen Kreisfläche ergibt.

2. *Wie im ersten Fall sei $k = 1$, $n = 2$ und $E = \mathbb{R}^2$. Ferner sei $\mathbf{\Phi} = [\Phi_1, \Phi_2]^T : D := [a, b] \to E$ beliebig, aber stetig differenzierbar auf $[a, b]$ und injektiv auf $[a, b)$, und $\omega := xdy + ydx$. Dann gilt*

$$\begin{aligned} \omega(\mathbf{\Phi}) &= \int_{\Phi} xdy + ydx \\ &= \int_{\Phi} \Phi_1(t)\Phi_2'(t) + \Phi_2(t)\Phi_1'(t) dt \\ &= \int_a^b \big(\Phi_1(t) \cdot \Phi_2(t)\big)' dt \\ &= \Phi_2(b)\Phi_1(b) - \Phi_1(a)\Phi_2(a). \end{aligned}$$

Gilt $\Phi(a) = \Phi(b)$, das heißt ist $\Phi(D)$ eine geschlossene Kurve, so folgt offensichtlich

$$\int_\Phi xdy + ydx = 0.$$

Es sei an dieser Stelle angemerkt, dass auch für allgemeinere geschlossene doppelpunktfreie Kurven die Zweiform

$$F = \frac{1}{2}\int_{\mathbf{\Phi}} xdy - ydx$$

der von der Kurve eingeschlossene Flächeninhalt ist (dies ist die **Sektorformel**) während

$$0 = \frac{1}{2}\int_{\mathbf{\Phi}} xdy + ydx.$$

gilt.

Beispiel 11.2.3 (Volumenintegrale im $\mathbb{R}^2$). *Es sei nun $k = 2$, $n = 2$ und $E = \mathbb{R}^2$, dann definiert $\omega := dx \wedge dy$ eine Differentialform der Ordnung 2 in $\mathbb{R}^2$. Für $\mathbf{\Phi} : D = [0, R] \times [0, 2\pi] \to \mathbb{R}^2$ definiert durch*

$$\mathbf{\Phi}(r, t) := \begin{bmatrix} r\cos t \\ r\sin t \end{bmatrix}$$

ist $\mathbf{\Phi}(D)$ der Kreis mit Radius R und es gilt

$$\omega(\mathbf{\Phi}) = \int_{\mathbf{\Phi}} dx \wedge dy = \int_0^R \int_0^{2\pi} rdtdr = \pi R^2.$$

Die Zweiform $dx \wedge dy$ ist das **Volumenelement** *im $\mathbb{R}^2$ und liefert den Flächeninhalt von $\Phi(D)$.*

Für $\omega := a(x, y)dx \wedge dy$ gilt

$$\omega(\mathbf{\Phi}) = \int_{\mathbf{\Phi}} a(x, y)dx \wedge dy = \int_D a\big(\mathbf{\Phi}(r, \phi)\big) \det \mathbf{\Phi}'(r, \phi)d(r, \phi).$$

Bis auf das Vorzeichen entspricht dies der Substitutionsregel.

Beispiel 11.2.4 (Oberflächenintegrale im $\mathbb{R}^3$). *Sei $k = 2$, $n = 3$ und $\mathbf{f} = [f_1, f_2, f_3]^T : \mathbb{R}^3 \to \mathbb{R}^3$ stetig. Wir betrachten die folgende Zweiform*

$$\omega = f_1(\mathbf{x})\, dy \wedge dz + f_2(\mathbf{x})\, dz \wedge dx + f_3(\mathbf{x})\, dx \wedge dy.$$

Sei $\mathbf{\Phi} = [\Phi_1, \Phi_2, \Phi_3]^T : D \subseteq \mathbb{R}^2 \to \mathbb{R}^3$ eine Zweifläche, dann ist $\Gamma := \mathbf{\Phi}(D)$ ein Flächenstück im $\mathbb{R}^3$. Schreiben wir $\mathbf{u} = (u, v) \in D$ und $\mathbf{x} = (x, y, z)$, so hat ein Punkt $\mathbf{x} \in \Gamma$ die Darstellung

$$\mathbf{x}(\mathbf{u}) = \begin{bmatrix} x(u, v) \\ y(u, v) \\ z(u, v) \end{bmatrix} = \begin{bmatrix} \Phi_1(u, v) \\ \Phi_2(u, v) \\ \Phi_3(u, v) \end{bmatrix} = \mathbf{\Phi}(\mathbf{u}).$$

Ferner gilt

$$\begin{aligned}
dy \wedge dz &= \det \frac{\partial(y,z)}{\partial(u,v)} \cdot d(u,v) = \begin{vmatrix} \frac{\partial \Phi_2}{\partial u} & \frac{\partial \Phi_2}{\partial v} \\ \frac{\partial \Phi_3}{\partial u} & \frac{\partial \Phi_3}{\partial v} \end{vmatrix} \cdot d(u,v), \\
dz \wedge dx &= \det \frac{\partial(z,x)}{\partial(u,v)} \cdot d(u,v) = \begin{vmatrix} \frac{\partial \Phi_3}{\partial u} & \frac{\partial \Phi_3}{\partial v} \\ \frac{\partial \Phi_1}{\partial u} & \frac{\partial \Phi_1}{\partial v} \end{vmatrix} \cdot d(u,v), \\
dx \wedge dy &= \det \frac{\partial(x,y)}{\partial(u,v)} \cdot d(u,v) = \begin{vmatrix} \frac{\partial \Phi_1}{\partial u} & \frac{\partial \Phi_1}{\partial v} \\ \frac{\partial \Phi_2}{\partial u} & \frac{\partial \Phi_2}{\partial v} \end{vmatrix} \cdot d(u,v),
\end{aligned}$$

und, wie man leicht nachrechnet, gilt die Identität

$$\begin{bmatrix} \det \frac{\partial(y,z)}{\partial(u,v)} \\ \det \frac{\partial(z,x)}{\partial(u,v)} \\ \det \frac{\partial(x,y)}{\partial(u,v)} \end{bmatrix} (u,v) = (\mathbf{\Phi}_u \times \mathbf{\Phi}_v)(u,v) =: \mathbf{N}(u,v).$$

Gemäß Bemerkung 9.1.15 spannen die beiden Vektoren

$$\mathbf{\Phi}_u(u,v) = \frac{\partial \mathbf{\Phi}}{\partial u}(u,v) = \begin{bmatrix} \Phi_{1,u}(u,v) \\ \Phi_{2,u}(u,v) \\ \Phi_{3,u}(u,v) \end{bmatrix}, \qquad \mathbf{\Phi}_v(u,v) = \frac{\partial \mathbf{\Phi}}{\partial v}(u,v) = \begin{bmatrix} \Phi_{1,v}(u,v) \\ \Phi_{2,v}(u,v) \\ \Phi_{3,v}(u,v) \end{bmatrix},$$

die **Tangentialebene** *an die Fläche $\mathbf{\Phi}(D)$ im Punkt $\mathbf{x}(u,v) = \mathbf{\Phi}(u,v)$ auf. Da der Vektor $\mathbf{N}(u,v) = \mathbf{\Phi}_u(u,v) \times \mathbf{\Phi}_v(u,v)$ senkrecht zu dieser Tangentialebene steht, sprechen wir vom* **Normalenvektor** *an das Flächenstück Γ im Punkt $\mathbf{x}(u,v) = \mathbf{\Phi}(u,v)$. Normieren wir diesen Normalenvektor, so erhalten wir den* **Normaleneinheitsvektor** $\mathbf{n} := \mathbf{N}/\|\mathbf{N}\|$. *Da $-\mathbf{n}$ ebenfalls die Norm 1 besitzt und senkrecht auf $\Psi(\boldsymbol{D}$ steht, wird durch unsere Definition eine Orientierung von Γ eingeführt wird. Bis auf die Richtung ist $\mathbf{n}$ unabhängig von der jeweiligen Parametrisierung von Γ.*

Es ergibt sich nun

$$\begin{aligned}
\omega(\mathbf{\Phi}) &= \int_{\mathbf{\Phi}} f_1(\mathbf{x}) dy \wedge dz + f_2(\mathbf{x}) dz \wedge dx + f_3(\mathbf{x}) dx \wedge dy \\
&= \int_D \Big[f_1\big(\mathbf{\Phi}(u,v)\big) \det \frac{\partial(y,z)}{\partial(u,v)} \\
&\qquad\qquad + f_2\big(\mathbf{\Phi}(u,v)\big) \det \frac{\partial(z,y)}{\partial(u,v)} + f_3\big(\mathbf{\Phi}(u,v)\big) \det \frac{\partial(y,x)}{\partial(u,v)} \Big] d(u,v) \\
&= \int_D \langle \mathbf{f}\big(\mathbf{\Phi}(u,v)\big), \mathbf{N}(u,v) \rangle d(u,v) \\
&= \int_D \langle (\mathbf{f} \circ \mathbf{\Phi})(u,v), \mathbf{n}(u,v) \rangle \|\mathbf{N}(u,v)\| d(u,v).
\end{aligned}$$

Führen wir das **Oberflächenmaß** *im $\mathbb{R}^3$ vermittels $do = \|\mathbf{N}(u,v)\| d(u,v)$ ein, dann lässt sich vereinfacht schreiben*

$$\omega(\mathbf{\Phi}) = \int_{\mathbf{\Phi}} f_1(\mathbf{x}) dy \wedge dz + f_2(\mathbf{x}) dz \wedge dx + f_3(\mathbf{x}) dx \wedge dy = \int_\Gamma \langle (\mathbf{f}(\mathbf{x}), \mathbf{n}(\mathbf{x}) \rangle do.$$

Diese Größe wird **Oberflächenfluss** *genannt.*

Schließlich sei angemerkt, dass die spezielle Wahl $\mathbf{f} = \mathbf{n}$*, das heißt*

$$\omega(\mathbf{\Phi}) = \int_\Gamma \langle \mathbf{n}(\mathbf{x}), \mathbf{n}(\mathbf{x}) \rangle do = \int_\Gamma do,$$

den Flächeninhalt des Flächenstückes Γ *liefert.*

Gegeben sei eine k-Fläche im $\mathbb{R}^n$

$$\mathbf{\Phi} = [\Phi_1, \ldots, \Phi_n]^T : D \subseteq \mathbb{R}^k \to \mathbb{R}^n.$$

Für jede Stelle $\mathbf{u} = (u_1, \ldots, u_k) \in D$ wollen wir abkürzend im folgenden schreiben

$$\Phi_{u_i}(\mathbf{u}) := \frac{\partial \mathbf{\Phi}}{\partial u_i}(\mathbf{u}) \in \mathbb{R}^n.$$

Wir definieren für

$$g_{ij}(\mathbf{u}) := \langle \Phi_{u_i}(\mathbf{u}), \Phi_{u_j}(\mathbf{u}) \rangle, \qquad 1 \leq i, j \leq k,$$

die Matrix

$$\mathbf{G}(\mathbf{u}) = \begin{bmatrix} g_{11}(\mathbf{u}) & \cdots & g_{1k}(\mathbf{u}) \\ \vdots & & \vdots \\ g_{k1}(\mathbf{u}) & \cdots & g_{kk}(\mathbf{u}) \end{bmatrix}$$

Dann ist die Matrix $\mathbf{G} = \mathbf{G}(\mathbf{u})$ symmetrisch und, falls $\mathbf{\Phi}$ injektiv ist, auch positiv definit (der Beweis verbleibt dem Leser als Übung). Sie wird **erster Fundamentaltensor der Differentialgeometrie** genannt. Folglich ist

$$g(\mathbf{u}) := \det \mathbf{G}(\mathbf{u})$$

reell und positiv.

Definition 11.2.5. *Sei* $\mathbf{\Phi} : D \subseteq \mathbb{R}^k \to \mathbb{R}^n$ *eine injektive* k*-Fläche im* $\mathbb{R}^n$ *und* $\Gamma := \mathbf{\Phi}(D)$*, wobei* $n \geq k \geq 0$ *beliebig sei. Bezeichne ferner* $g(\mathbf{u}) := \det(\langle \mathbf{\Phi}_{u_i}, \mathbf{\Phi}_{u_j} \rangle)$ *die Determinante des ersten Fundamentaltensors der Differentialgeometrie. Dann nennen wir die Größe*

$$do := \sqrt{g(\mathbf{u})} d\mathbf{u}$$

das **Oberflächenelement** *oder das* **Oberflächenmaß**.

Für $f : \Gamma \to \mathbb{R}$ *definieren wir das* **Oberflächenintegral** *vermittels*

$$\int_\Gamma f(\mathbf{x}) do := \int_D f(\Phi(\mathbf{u})) \sqrt{g(\mathbf{u})} d\mathbf{u}.$$

Im Spezialfall $f \equiv 1$ *definiert das Integral*

$$\int_\Gamma do := \int_D \sqrt{g(\mathbf{u})} d\mathbf{u}$$

den **Flächeninhalt** *der* k*-Fläche* $\Gamma := \mathbf{\Phi}(D)$.

Der Vorteil dieser Definition ist, dass sie für beliebige Dimensionen $n \geq k$ gültig ist. In der Tat ist das unabhängig von der jeweiligen Parametrisierung von Γ.

Satz 11.2.6. *Seien $\mathbf{\Phi} : D \to \mathbb{R}^n$ und $\mathbf{\Psi} : E \to \mathbb{R}^n$ injektive Funktionen mit $\Gamma = \mathbf{\Phi}(D) = \mathbf{\Psi}(E)$. Ferner seien $g(\mathbf{u}) := \det(\langle \mathbf{\Phi}_{u_i}, \mathbf{\Phi}_{u_j} \rangle)$ und $h(\mathbf{v}) := \det(\langle \mathbf{\Psi}_{v_i}, \mathbf{\Psi}_{v_j} \rangle)$ die zugehörigen ersten Fundamentaltensoren der Differentialgeometrie. Dann gilt*

$$\int_\Gamma F(\mathbf{x}) do = \int_D F\big(\mathbf{\Phi}(\mathbf{u})\big) \sqrt{g(\mathbf{u})} d\mathbf{u} = \int_E F\big(\mathbf{\Psi}(\mathbf{v})\big) \sqrt{h(\mathbf{v})} d\mathbf{v}.$$

Beweis. Die Abbildungen $\mathbf{\Phi} : D \to \Gamma$ und $\mathbf{\Psi} : E \to \Gamma$ sind bijektiv. Daher ist die Funktion $\mathbf{\Xi} : D \to E$ gegeben durch $\mathbf{\Xi} := \mathbf{\Psi}^{-1} \circ \mathbf{\Phi}$ ebenfalls bijektiv und nach dem Satz über inverse Funktionen auch stetig differenzierbar. Die Behauptung folgt nun mit Hilfe der Substitutionsregel im $\mathbb{R}^k$. □

Wir wollen im Falle $n = 3$ die zwei Spezialfälle $k = 1$ und $k = 2$ näher betrachten, da gerade diese in den Anwendungen sehr wichtig sind und sich die obige Definition des Oberflächenmaßes wesentlich vereinfacht. Wie wir ebenfalls sehen, stimmt die allgemeine Definition des Oberflächenmaßes mit den bereits betrachteten Beispielen überein.

Wie wir bereits wissen, definiert die Einsfläche $\Phi : [a, b] \to \mathbb{R}^3$ eine Kurve Γ. Wir finden

$$\mathbf{G} = \langle \mathbf{\Phi}', \mathbf{\Phi}' \rangle = \|\mathbf{T}\|^2$$

und folglich gilt

$$do = \|\mathbf{T}(\mathbf{u})\|^2 d(\mathbf{u}).$$

Gemäß Beispiel 11.2.1 wird im Fall $k = 1$ das Oberflächenmaß do auch als Bogenlängenelement ds bezeichnet.

Durch die Zweifläche $\mathbf{\Phi} : \mathbb{R}^2 \to \mathbb{R}^3$ ist ein Flächenstück $\Gamma = \mathbf{\Phi}(D)$ gegeben. Wie man leicht nachrechnet gilt hier

$$g(\mathbf{u}) = \langle \mathbf{\Phi}_u, \mathbf{\Phi}_u \rangle \langle \mathbf{\Phi}_v, \mathbf{\Phi}_v \rangle - \langle \mathbf{\Phi}_u, \mathbf{\Phi}_v \rangle^2.$$

Andererseits ergibt sich aber auch aufgrund der Identität von Lagrange

$$\|\mathbf{N}\|^2 = \langle \mathbf{\Phi}_u \times \mathbf{\Phi}_v, \mathbf{\Phi}_u \times \mathbf{\Phi}_v \rangle = \langle \mathbf{\Phi}_u, \mathbf{\Phi}_u \rangle \langle \mathbf{\Phi}_v, \mathbf{\Phi}_v \rangle - \langle \mathbf{\Phi}_u, \mathbf{\Phi}_v \rangle^2.$$

Dies bedeutet die Gleichung

$$do = \|\mathbf{N}(\mathbf{u})\|^2 d(\mathbf{u}),$$

welche mit Beispiel 11.2.4 übereinstimmt.

Beispiel 11.2.7. *1. Wir wollen die Größe der Oberfläche einer Kugel mit Radius r bestimmen. Dazu verwenden wir* **Kugelkoordinaten***, das heißt es sei*

$$\mathbf{x}(\mathbf{u}) = \begin{bmatrix} x(u,v) \\ y(u,v) \\ z(u,v) \end{bmatrix} = \begin{bmatrix} r \cos u \sin v \\ r \sin u \sin v \\ r \cos v \end{bmatrix} = \mathbf{\Phi}(\mathbf{u}). \tag{11.2}$$

Für $D := \{(u,v) \in \mathbb{R}^2 : u \in [0,2\pi],\ v \in [0,\pi]\}$ *ist dann*

$$\Gamma = \Phi(D) := \{(x,y,z) = \mathbf{\Phi}(u,v) : (u,v) \in D\}$$

die Oberfläche der Kugel.

Es gilt

$$\mathbf{\Phi}_u(u,v) = r \begin{bmatrix} -\sin u \sin v \\ \cos u \sin v \\ 0 \end{bmatrix}, \qquad \mathbf{\Phi}_v(u,v) = r \begin{bmatrix} \cos u \cos v \\ \sin u \cos v \\ -\sin v \end{bmatrix},$$

und somit folgt

$$\sqrt{g(u,v)} = \|\mathbf{\Phi}_u(u,v) \times \mathbf{\Phi}_v(u,v)\| = r^2 \sin v.$$

Damit erhalten wir die sicherlich bekannte Größe der Kugeloberfläche

$$\int_\Gamma do = \int_D \sqrt{g(u,v)} d(u,v) = \int_0^{2\pi} \int_0^\pi r^2 \sin v dv du = r^2 2\pi(-\cos v)\big|_0^\pi = 4\pi r^2.$$

Falls gilt $v \in [\theta_0, \theta_1] \subseteq [0, 2\pi]$*, das heißt* Γ *ist nur ein Stück der Kugeloberfläche, dann erhalten wir analog*

$$\int_\Gamma do = 2\pi r^2 (\cos \Theta_0 - \cos \Theta_1).$$

2. *Es bezeichne* $E \subseteq \mathbb{R}^3$ *die Kugel vom Radius* R *und* $\omega = dx \wedge dy \wedge dz$ *das Volumenelement. Wie oben verwenden wir Kugelkoordinaten gemäß* (11.2)*. Für* $D := \{(r,u,v) \in \mathbb{R}^3 : r \in [0,R],\ u \in [0,2\pi],\ v \in [0,\pi]\}$ *ist dann* $E = \mathbf{\Phi}(D)$*. Es gilt*

$$\det \frac{\partial(x,y,z)}{\partial(r,u,v)} = \begin{vmatrix} \cos u \sin v & \sin u \sin v & \cos v \\ -r \sin u \sin v & r \cos u \sin v & 0 \\ r \cos u \cos v & r \sin u \cos v & -r \sin v \end{vmatrix} = r^2 \sin v$$

und daher folgt

$$\begin{aligned} \int_{\mathbf{\Phi}} \omega &= \int_{\mathbf{\Phi}} dx \wedge dy \wedge dz \\ &= \int_D \det \frac{\partial(x,y,z)}{\partial(r,u,v)} d(r,u,v) \\ &= \int_0^R \int_0^{2\pi} \int_0^\pi r^2 \sin v dv du dr \\ &= \frac{4}{3}\pi R^3. \end{aligned}$$

Neben den bisher verwendeten impliziten Parametrisierungen wollen wir nun die natürliche **explizite Parametrisierung** eines Flächenstückes Γ betrachten. Dies bedeutet, dass das Flächenstück Γ beschrieben ist durch

$$\Gamma = \operatorname{graph} f = \left\{ \begin{bmatrix} x \\ y \\ f(x,y) \end{bmatrix} \in \mathbb{R}^3 : (x,y) \in D \right\}$$

mit einer stetig differenzierbaren Funktion $f : D \to \mathbb{R}$.

Wir parametrisieren Γ nun mit Hilfe einer Zweifläche $\mathbf{\Phi} : D \to \mathbb{R}^3$. Hierzu definieren wir die Parametrisierung $\mathbf{\Phi}$ gemäß

$$\mathbf{x}(\mathbf{u}) = \begin{bmatrix} x(u,v) \\ y(u,v) \\ z(u,v) \end{bmatrix} = \begin{bmatrix} u \\ v \\ f(u,v) \end{bmatrix} = \mathbf{\Phi}(\mathbf{u}).$$

Diese implizite Parametrisierung ist offensichtlich äquivalent zur obigen expliziten Parametrisierung. Wegen

$$\mathbf{\Phi}_u(u,v) = \begin{bmatrix} 1 \\ 0 \\ f_u(u,v) \end{bmatrix}, \qquad \mathbf{\Phi}_v(u,v) = \begin{bmatrix} 0 \\ 1 \\ f_v(u,v) \end{bmatrix},$$

folgern wir, wie man durch Nachrechnen leicht bestätigt, dass

$$\sqrt{g(u,v)} = \|\mathbf{\Phi}_u(u,v) \times \mathbf{\Phi}_v(u,v)\| = \sqrt{1 + |f_u(u,v)|^2 + |f_v(u,v)|^2},$$

dies bedeutet $do = \sqrt{1 + |f_u(u,v)|^2 + |f_v(u,v)|^2} d(u,v)$.

Bemerkung 11.2.8.

1. *Wir haben die Differentialform ω durch ihre Implementierung als Integral zu k-Flächen Φ definiert. Die Differentialform selbst ist unabhängig von der jeweiligen k-Fläche, das heißt, sie ist in diesem Sinne universell.*

2. *Gemäß unserer Definition unterscheiden wir zwischen der k-Fläche $\mathbf{\Phi} : D \to \mathbb{R}^n$, das heißt, der Parametrisierung, welche eine Funktion ist, und ihrer Bildmenge $\Gamma = \mathbf{\Phi}(D)$.*

3. *Die bisher eingeführten Kurven-, Oberflächen- und Volumenintegrale sind unabhängig von der jeweiligen Parametrisierung definiert.*

11.3 Differentialformenkalkül

11.3.1 Elementare Eigenschaften

Definition 11.3.1. *Seien ω und ψ zwei k-Formen über $E \subseteq \mathbb{R}^n$, dann heißen sie gleich (kurz $\omega = \psi$), falls für alle k-Flächen $\mathbf{\Phi} \in C^1(D_{\mathbf{\Phi}}, E)$ gilt*

$$\int_{\mathbf{\Phi}} \omega = \int_{\mathbf{\Phi}} \psi.$$

Insbesondere ist $\omega = 0$, falls für alle k-Flächen $\mathbf{\Phi} \in C^1(D_{\mathbf{\Phi}}, E)$ gilt

$$\int_{\mathbf{\Phi}} \omega = 0.$$

Unter der k-Form $\omega + \psi$ verstehen wir diejenige k-Form, für die für beliebige k-Flächen $\mathbf{\Phi} \in C^1(D_{\mathbf{\Phi}}, E)$ gilt

$$\int_{\mathbf{\Phi}} (\omega + \psi) = \int_{\mathbf{\Phi}} \omega + \int_{\mathbf{\Phi}} \psi.$$

Analog ist für $c \in \mathbb{R}$ die k-Form $c\omega$ definiert durch

$$\int_{\mathbf{\Phi}} c\omega = c \int_{\mathbf{\Phi}} \omega$$

für beliebige k-Flächen $\mathbf{\Phi} \in C^1(D_{\mathbf{\Phi}}, E)$.

Es sei $\omega = f(\mathbf{x})\ dx_{i_1} \wedge \ldots \wedge dx_{i_k}$ eine k-Form und es bezeichne $\overline{\omega}$ die k-Form, die sich aus ω durch Vertauschen des Indexpaares i_m und i_n, $m \neq n$, ergibt. Aus der Tatsache, dass eine Determinante ihr Vorzeichen ändert, wenn zwei ihrer Zeilen vertauscht werden, wird ersichtlich, dass

$$\overline{\omega} = -\omega.$$

Es ergibt sich folglich **Antikommutativität**

$$dx_i \wedge dx_j = (-1) dx_j \wedge dx_i \tag{11.3}$$

und daher der Spezialfall

$$dx_i \wedge dx_i = 0.$$

11.3.2 Normaldarstellungen

Definition 11.3.2. *Ein n-Tupel $I = (i_1, \ldots, i_k)$ mit $1 \leq i_1 < i_2 < \cdots < i_k \leq n$ heißt* **wachsender k-Index**. *Wir schreiben kurz*

$$d\mathbf{x}_I := dx_{i_1} \wedge \ldots \wedge dx_{i_k}.$$

Jede k-Form kann durch Permutationen auf die Gestalt

$$\omega = \sum_I b_i(\mathbf{x}) d\mathbf{x}_I$$

mit wachsenden k-Indizes transformiert werden. Diese Darstellung heißt **Normaldarstellung** *von ω.*

Man überlegt sich leicht, dass die Normaldarstellung einer k-Form eindeutig ist, das heißt, es gibt genau eine Normaldarstellung.

Beispiel 11.3.3. *Die Normaldarstellung der Zweiform*

$$\omega = x dx \wedge dy + dy \wedge dz - dz \wedge dy + dy \wedge dx$$

lautet

$$\omega = (x - 1)\ dx \wedge dy + 2\ dy \wedge dz.$$

Satz 11.3.4. *Sei $\omega = \sum_I b_I(\mathbf{x}) d\mathbf{x}_I$ eine k-Form in Normaldarstellung über $E \subseteq \mathbb{R}^n$. Dann ist $\omega = 0$ genau dann, wenn $b_I(\mathbf{x}) = 0$ gilt für alle $\mathbf{x} \in E$ und wachsenden Indizes I.*

Beweis. Sind alle $b_I(\mathbf{x}) = 0$ für alle $\mathbf{x} \in E$ und Indizes I, dann gilt offensichtlich $\omega = 0$. Sei umgekehrt $\omega = 0$. Angenommen, es existiert ein $\mathbf{x} \in E$ und ein wachsender k-Index I mit $b_I(\mathbf{x}) \neq 0$. Ohne Beschränkung der Allgemeinheit sei $b_I(\mathbf{x}) > 0$. Da $b_I \in C(E, \mathbb{R})$ ist, existiert eine offene Umgebung $U_\delta(\mathbf{x})$, so dass $b(\mathbf{y}) > 0$ für alle $\mathbf{y} \in U_\delta(\mathbf{x})$.

Für genügend kleines $\varepsilon > 0$ definieren wir zu $I = (i_1, \ldots, i_k)$ die k-Zelle

$$D := \{\mathbf{u} = (u_1, \ldots, u_k) : |u_l| < \varepsilon,\ l = 1, \ldots, k\}.$$

Weiterhin sei $\mathbf{\Phi} : D \to E$ eine k-Fläche definiert durch

$$\mathbf{\Phi}(\mathbf{u}) = \mathbf{x} + (u_1 \mathbf{e}_{i_1} + \cdots + u_k \mathbf{e}_{i_k}) = \mathbf{x} + \sum_{l=1}^{k} u_l \mathbf{e}_{i_l} \in \mathbb{R}^n$$

mit $\mathbf{\Phi}(D) \subseteq U_\delta(\mathbf{x})$. Dann gilt

$$\det \frac{\partial(x_{i_1}, \ldots, x_{i_k})}{\partial(u_1, \ldots, u_k)} = \begin{vmatrix} \frac{\partial \Phi_{i_1}}{\partial u_1} & \cdots & \frac{\partial \Phi_{i_1}}{\partial u_k} \\ \vdots & & \vdots \\ \frac{\partial \Phi_{i_k}}{\partial u_1} & \cdots & \frac{\partial \Phi_{i_k}}{\partial u_k} \end{vmatrix} = 1,$$

während für alle anderen wachsenden Indizes $J = (j_1, \ldots, j_k) \neq I$ gilt

$$\det \frac{\partial(x_{j_1}, \ldots, x_{j_k})}{\partial(u_1, \ldots, u_k)} = 0.$$

Folglich erhalten wir

$$0 = \int_{\mathbf{\Phi}} \omega = \int_D b_I\big(\mathbf{\Phi}(\mathbf{u})\big) d\mathbf{u} > 0.$$

Da $\mathbf{\Phi}(\mathbf{u}) \in U_\delta(\mathbf{x})$ ist, liegt ein Widerspruch zu $\int_{\mathbf{\Phi}} \omega = 0$ vor. □

11.3.3 Äußere Multiplikation

Definition 11.3.5. *Seien $d\mathbf{x}_P$ und $d\mathbf{x}_Q$ mit*

$$d\mathbf{x}_P = dx_{i_1} \wedge \ldots \wedge dx_{i_p}, \qquad d\mathbf{x}_Q = dx_{j_1} \wedge \ldots \wedge dx_{j_q}$$

eine p- bzw. q-Form mit wachsenden Indizes P, Q, dann definieren wir das **äußere Produkt** *vermittels*

$$d\mathbf{x}_P \wedge d\mathbf{x}_Q = dx_{i_1} \wedge \ldots \wedge dx_{i_p} \wedge dx_{j_1} \wedge \ldots \wedge dx_{j_q}.$$

Dies ist eine $(p+q)$-Form. Mit $[P, Q]$ bezeichnen wir den wachsenden $(p+q)$-Index aus der Menge $\{i_1, \ldots, i_p, j_1, \ldots, j_q\} \subseteq \mathbb{N}$.

Satz 11.3.6. *Sei α die Anzahl der Paarungen i_r, j_s mit $j_s < i_r$, dann gilt*

$$d\mathbf{x}_P \wedge d\mathbf{x}_Q = (-1)^\alpha d\mathbf{x}_{[P,Q]}.$$

Beweis. Gilt $j_s < i_r$, dann müssen wir die Variablen j_s und i_r vertauschen. Insgesamt benötigen wir α Vertauschungen um die Indizes in aufsteigender Reihenfolge zu ordnen. Jede Vertauschung liefert wegen der Antikommutativität 11.3 den Faktor -1. □

Definition 11.3.7. *Seien ω eine k-Form und λ eine m-Form mit Normaldarstellungen*

$$\omega = \sum_I a_I \mathbf{x}_I, \qquad \lambda = \sum_J b_J \mathbf{x}_J,$$

dann definiert $\omega \wedge \lambda$ die $(k+m)$-Form gegeben durch

$$\omega \wedge \lambda := \sum_I \sum_J a_I b_J d\mathbf{x}_I \wedge d\mathbf{x}_J$$

Satz 11.3.8. *Es sei $E \subseteq \mathbb{R}^n$ eine offene Menge und $k, m, p < n$. Für beliebige k-, m- bzw. p-Formen ω, λ bzw. μ in E gelten*

1. *das Assoziativgesetz*
$$(\omega \wedge \lambda) \wedge \mu = \omega \wedge (\lambda \wedge \mu),$$

2. *das Distributivgesetz*
$$\omega \wedge (\lambda + \mu) = (\omega \wedge \lambda) + (\omega \wedge \mu)$$
vorausgesetzt $p = m$, sowie

3. *das Antikommutativgesetz*
$$\omega \wedge \lambda = (-1)^{k \cdot m}(\lambda \wedge \omega).$$

Beweis. Der Beweis verbleibt dem Leser als Übung. □

Bemerkung 11.3.9. *Ist f eine Nullform und ω eine k-Form, so schreiben wir*

$$f \wedge \omega = f \cdot \omega.$$

Für $\omega = d\mathbf{x}_I$ gilt beispielsweise

$$f \wedge \omega = f d\mathbf{x}_I$$

während für $\omega = \sum_I a_I d\mathbf{x}_I$ gilt

$$f \wedge \omega = \sum_I f a_I d\mathbf{x}_I.$$

11.3.4 Äußere Differentiation

Definition 11.3.10. *Es sei $k < n$ und $E \subseteq \mathbb{R}^n$ offen. Ist $f \in C^1(E, \mathbb{R})$ eine Nullform in E, dann definieren wir die* **äußere Ableitung** *als Einsform in E*

$$df := \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{x}) dx_i = f'(\mathbf{x}) \begin{bmatrix} dx_1 \\ \vdots \\ dx_n \end{bmatrix}.$$

Ist $\omega = \sum_I f_I dx_I$ eine k-Form in E in Normaldarstellung, dann definieren wir die **äußere Ableitung** *$d\omega$ als folgende $(k+1)$-Form in E*

$$d\omega = d\left(\sum_I f_I dx_I \right) := \sum_I df_I \wedge d\mathbf{x}_I.$$

*Gilt $f_I \in C^l(E, \mathbb{R})$ für alle wachsenden Indizes I, dann heißt $\omega = \sum_I f_I d\mathbf{x}_I$ eine k-***Form der Klasse** *C^l in E.*

Satz 11.3.11. *Es sei $E \subseteq \mathbb{R}^n$ eine offene Menge und $k, m < n$.*

1. *Seien $\omega = \sum_I a_I d\mathbf{x}_I$ eine k-Form und $\lambda = \sum_J b_J d\mathbf{x}_J$ eine m-Form in E, dann gilt die Produktregel*
$$d(\omega \wedge \lambda) = d\omega \wedge \lambda + (-1)^k \omega \wedge d\lambda.$$
2. *Sei $\omega = \sum_I a_I d\mathbf{x}_I$ eine k-Form der Klasse C^2 in E, dann gilt*
$$d^2\omega = d(d\omega) = 0.$$

Beweis. 1. Ohne Beschränkung der Allgemeinheit haben ω und λ die Normaldarstellungen $\omega = a_I d\mathbf{x}_I$ und $\lambda = b_J d\mathbf{x}_J$. Es gilt

$$\omega \wedge \lambda = a_I b_J d\mathbf{x}_I \wedge d\mathbf{x}_J$$

und somit folgt

$$d(a_I \cdot b_J) = \sum_{i=1}^{n} \frac{\partial a_I(\mathbf{x})}{\partial x_i} b_J(\mathbf{x}) dx_i + \sum_{j=1}^{n} a_I(\mathbf{x}) \frac{\partial b_J(\mathbf{x})}{\partial x_j} dx_j = b_J \wedge da_J + a_I db_J.$$

Wegen

$$a_I \; db_J \wedge d\mathbf{x}_I = a_I \; db_J \wedge dx_{i_1} \wedge \ldots \wedge dx_{i_k} = (-1)^k \; d\mathbf{x}_I \wedge a_I db_J$$

folgt

$$\begin{aligned} d(\omega \wedge \lambda) &= d(a_I b_J \; d\mathbf{x}_I \wedge d\mathbf{x}_J) \\ &= d(a_I b_J) \wedge d\mathbf{x}_I \wedge d\mathbf{x}_J \\ &= (b_J da_I + a_I db_J) \wedge d\mathbf{x}_I \wedge d\mathbf{x}_J \\ &= b_J da_I \wedge d\mathbf{x}_I \wedge d\mathbf{x}_J + (-1)^k d\mathbf{x}_I \wedge a_I db_J \wedge d\mathbf{x}_J \\ &= da_I \wedge d\mathbf{x}_I \wedge b_J d\mathbf{x}_J + (-1)^k d\mathbf{x}_I \wedge a_I db_J \wedge d\mathbf{x}_J \\ &= d\omega \wedge \lambda + (-1)^k \omega \wedge d\lambda. \end{aligned}$$

2. Sei $\omega = f \in C^2(E, \mathbb{R})$ eine Nullform, dann ist

$$d\omega = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} dx_i = \sum_{i=1}^{n} f_{x_i} dx_i$$

und weiter

$$\begin{aligned} d(d\omega) &= d\left(\sum_{i=1}^{n} f_{x_i} dx_i \right) = \sum_{i=1}^{n} d(f_{x_i}) \wedge dx_i = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial f_{x_i}}{\partial x_j} dx_j \wedge dx_i \\ &= \sum_{i,j=1}^{n} f_{x_i x_j} dx_j \wedge dx_i. \end{aligned}$$

Da $f \in C^2(E, \mathbb{R})$, gilt mit nach dem Satz von Schwarz $f_{x_i x_j} = f_{x_j x_i}$. Aufgrund der Antikommutativität

$$dx_j \wedge dx_i = -dx_i \wedge dx_j$$

gilt daher

$$f_{x_i x_j} dx_j \wedge dx_i = f_{x_j x_i} dx_i \wedge dx_j,$$

und folglich ergibt sich

$$d(d\omega) = \sum_{i,j=1}^{n} f_{x_i x_j} dx_j \wedge dx_i = 0.$$

□

Beispiel 11.3.12. *Sei $f \in C(E, \mathbb{R})$ eine Nullform im $\mathbb{R}^3$. Für stetig differenzierbares $\mathbf{\Phi} = [\Phi_1, \Phi_2, \Phi_3]^T : [a, b] \to \mathbb{R}^3$ mit $\mathbf{\Phi}(a) = \mathbf{\Phi}(b)$ ist $\Gamma = \mathbf{\Phi}([a, b])$ eine geschlossene Kurve. Es gilt*

$$\begin{aligned} \int_{\mathbf{\Phi}} df &= \int_{\mathbf{\Phi}} \sum_{i=1}^{3} \frac{\partial f}{\partial x_i} dx_i = \int_a^b \sum_{i=1}^{3} \frac{\partial f}{\partial x_i} \big(\mathbf{\Phi}(t)\big) \Phi_i'(t) dt = \int_a^b (f \circ \mathbf{\Phi})'(t) dt \\ &= f\big(\Phi(t)\big) - f\big(\Phi(a)\big) = 0. \end{aligned}$$

11.3.5 Substitution

Definition 11.3.13. *Seien $D \subseteq \mathbb{R}^m$ und $E \subseteq \mathbb{R}^n$ offen und $\mathbf{t} = [t_1, \ldots, t_n]^T \in C^1(D, E)$. Desweiteren sei ω die k-Form in E mit der Normaldarstellung*

$$\omega = \sum_I a_I(\mathbf{y}) d\mathbf{y}_I$$

mit $\mathbf{y} = [y_1, \ldots, y_n]^T \in E$. *Für* $\mathbf{x} = [x_1, \ldots, x_m]^T \in D$ *ist* $\mathbf{y} = \mathbf{t}(\mathbf{x}) \in E$, *es ist also* $y_i = t_i(\mathbf{x})$ *für alle* $i = 1, \ldots, n$. *Ferner gilt*

$$dy_i = dt_i = \sum_{j=1}^{m} \frac{\partial t_i}{\partial x_j} dx_j.$$

Offenbar ist jedes dt_i *eine Einsform in* D.

Die Abbildung $\mathbf{t}$ *überführt die* k*-Form* ω *in* E *in eine* k*-Form* $\omega_{\mathbf{t}}$ *in* D, *die wie folgt definiert ist*

$$\omega_{\mathbf{t}} := \sum_I a_I\big(\mathbf{t}(\mathbf{x})\big) dt_I = \sum_I a_I\big(\mathbf{t}(\mathbf{x})\big)\ dt_{i_1} \wedge \ldots \wedge dt_{i_k}.$$

Die Abbildung $\omega \mapsto \omega_{\mathbf{t}}$ *nennen wir* **Substitution**. *Die* k*-Form* $\omega_{\mathbf{t}}$ *wird auch* **Pullback** *von* ω *genannt.*

Satz 11.3.14. *Sei* D, E *und* $\mathbf{t}$ *wie oben definiert und* ω *und* λ k*-Formen bzw.* l*-Formen in* E. *Dann gilt*

1. *die Beziehung*
$$(\omega \wedge \lambda)_{\mathbf{t}} = \omega_{\mathbf{t}} \wedge \lambda_{\mathbf{t}},$$

2. *im Falle* $k = l$ *die Gleichung*
$$(\omega + \lambda)_{\mathbf{t}} = \omega_{\mathbf{t}} + \lambda_{\mathbf{t}}$$
und

3. $d(\omega_{\mathbf{t}}) = (d\omega)_{\mathbf{t}}$, *falls* $\mathbf{t} \in C^2(D, E)$ *und* ω *der Klasse* C^1 *angehört.*

Beweis. 1. Die Behauptung ist offensichtlich, wenn man eingesehen hat, dass

$$(dy_{i_1} \wedge \ldots \wedge dy_{i_k})_{\mathbf{t}} = dt_{i_1} \wedge \ldots \wedge dt_{i_k} \tag{11.4}$$

auch dann gilt, wenn $(i_1, \ldots, i_k)$ nicht wächst. Denn auf jeder Seite von (11.4) ist dieselbe Anzahl von Minuszeichen nötig, um eine wachsende Anordnung zu erzeugen.

2. Diese Aussage ergibt sich unmittelbar aus der Definition der Substitution.

3. Ist $\omega = f$ eine Nullform in E der Klasse C^1, dann ist

$$f_{\mathbf{t}}(\mathbf{x}) = f\big(\mathbf{t}(\mathbf{x})\big), \qquad df = \sum_{i=1}^{n} \frac{\partial f}{\partial y_i}(\mathbf{y}) dy_i.$$

Nach der Kettenregel folgt

$$\begin{aligned} d(f_{\mathbf{t}}) &= \sum_{j=1}^{m} \frac{\partial f_{\mathbf{t}}}{\partial x_j}(\mathbf{x}) dx_j \\ &= \sum_{j=1}^{m} \sum_{i=1}^{n} \frac{\partial f}{\partial y_i}\big(\mathbf{t}(\mathbf{x})\big) \frac{\partial t_i}{\partial x_j}(\mathbf{x}) dx_j \\ &= \sum_{i=1}^{n} \frac{\partial f}{\partial y_i}\big(\mathbf{t}(\mathbf{x})\big) \cdot \sum_{j=1}^{m} \frac{\partial t_i}{\partial x_j}(\mathbf{x}) dx_j \\ &= \sum_{i=1}^{n} \frac{\partial f}{\partial y_i}\big(\mathbf{t}(\mathbf{x})\big) \cdot dt_i = (df)_{\mathbf{t}}. \end{aligned}$$

Ist $d\mathbf{y}_I = dy_{i_1} \wedge \ldots \wedge dy_{i_k}$, dann ist $(d\mathbf{y}_I)_{\mathbf{t}} = dt_{i_1} \wedge \ldots \wedge dt_{i_k}$. Gemäß der zweiten Aussage von Satz 11.3.11 folgt nun $d\big((d\mathbf{y}_I)_{\mathbf{t}}\big) = 0$, da $\mathbf{t} \in C^2(D, E)$.

Sei nun $\omega = f(\mathbf{y}) d\mathbf{y}_I$, dann ist $\omega_{\mathbf{t}} = f_{\mathbf{t}}(\mathbf{x})(d\mathbf{y}_I)_{\mathbf{t}}$ und mit der obigen Überlegung ergibt sich

$$\begin{aligned} d(\omega_{\mathbf{t}}) &= d(f_{\mathbf{t}}) \wedge (d\mathbf{y}_I)_{\mathbf{t}} + (-1) f_{\mathbf{t}} \wedge d\big((d\mathbf{y}_I)_{\mathbf{t}}\big) \\ &= (df)_{\mathbf{t}} \wedge (d\mathbf{y}_I)_{\mathbf{t}} \\ &= (df \wedge d\mathbf{y}_I)_{\mathbf{t}} = (d\omega)_{\mathbf{t}}. \end{aligned}$$

Aus der zweiten Aussage folgt nun direkt die Behauptung für den allgemeinen Fall von ω.

□

Satz 11.3.15. *Seien $D \subseteq \mathbb{R}^m$, $E \subseteq \mathbb{R}^n$ und $F \subseteq \mathbb{R}^p$ offene Mengen, sowie $\mathbf{t} : D \to E$ und $\mathbf{s} : E \to F$ zwei C^1-Abbildungen. Für eine gegebene k-Form ω in F ist $\omega_{\mathbf{s}}$ eine k-Form in E und $\omega_{\mathbf{s} \circ \mathbf{t}}$ eine k-Form in D, und es gilt*

$$(\omega_{\mathbf{s}})_{\mathbf{t}} = \omega_{\mathbf{s} \circ \mathbf{t}}. \tag{11.5}$$

Beweis. Sind ω, λ Formen in F, so folgt mit Satz 11.3.14

$$\big((\omega \wedge \lambda)_{\mathbf{s}}\big)_{\mathbf{t}} = (\omega_{\mathbf{s}} \wedge \lambda_{\mathbf{s}})_{\mathbf{t}} = (\omega_{\mathbf{s}})_{\mathbf{t}} \wedge (\lambda_{\mathbf{s}})_{\mathbf{t}}$$

und

$$(\omega \wedge \lambda)_{\mathbf{s} \circ \mathbf{t}} = \omega_{\mathbf{s} \circ \mathbf{t}} \wedge \lambda_{\mathbf{s} \circ \mathbf{t}}.$$

Somit gilt die Behauptung (11.5) für $\omega \wedge \lambda$, wenn sie für ω und λ Gültigkeit hat. Da jede Form aus Nullformen und Einsformen durch Addition und Multiplikation aufgebaut werden kann, und da (11.5) für Nullformen trivial ist, genügt es, (11.5) für $\omega = dz_q$, $q = 1, \ldots, p$, zu beweisen.

Seien $\mathbf{x} \in D$, $\mathbf{y} = \mathbf{t}(\mathbf{x}) \in F$ und $\mathbf{z} = \mathbf{s}(\mathbf{y}) \in G$ mit $\mathbf{y} = [y_1, \ldots, y_m]^T = [t_1(\mathbf{x}), \ldots, t_m(\mathbf{x})]^T$ und $\mathbf{z} = [z_1, \ldots, z_p]^T = [s_1(\mathbf{y}), \ldots, s_p(\mathbf{y})]^T$, dann gilt $\mathbf{z} = \mathbf{s}(\mathbf{y}) = \mathbf{s}\big(\mathbf{t}(\mathbf{x})\big) = (\mathbf{s} \circ \mathbf{t})(\mathbf{x}) = [r_1(\mathbf{x}), \ldots, r_p(\mathbf{x})]^T$.

Für $\omega = dz_q$ gilt

$$\omega_{\mathbf{s}} = ds_q = \sum_{j=1}^{n} \frac{\partial s_q}{\partial y_j}(\mathbf{y}) dy_j.$$

Daraus ergibt sich mit der Kettenregel

$$\begin{aligned}
(\omega_{\mathbf{s}})_{\mathbf{t}} &= \sum_{j=1}^{n} \frac{\partial s_q}{\partial y_j}\big(\mathbf{t}(\mathbf{x})\big) dt_j \\
&= \sum_{j=1}^{n} \frac{\partial s_q}{\partial y_j}\big(\mathbf{t}(\mathbf{x})\big) \sum_{i=1}^{m} \frac{\partial t_i}{\partial x_i}(\mathbf{x}) dx_i \\
&= \sum_{i=1}^{m} \sum_{j=1}^{n} \frac{\partial s_q}{\partial y_i} \cdot \frac{\partial t_j}{\partial x_j} dx_i \\
&= \sum_{i=1}^{m} \frac{\partial r_q}{\partial x_i}(\mathbf{x}) dx_i \\
&= dr_q = \omega_{\mathbf{s}\circ\mathbf{t}},
\end{aligned}$$

das heißt, die Behauptung. □

Wir wollen an dieser Stelle die allgemeine Definition der Determinante wiederholen. Hierzu sei $(q_1, \ldots, q_k)$ ein k-Tupel ganzer Zahlen aus der Menge $\{1, 2, \ldots, k\}$. Ferner bezeichne

$$s(q_1, \ldots, q_k) := \prod_{1 \le m < n \le k} \operatorname{sgn}(q_n - q_m) \tag{11.6}$$

die *Signatur* bzw. *Charakteristik* des k-Tupels $(q_1, \ldots, q_k)$. Ist das k-Tupel $(q_1, \ldots, q_k)$ eine Permutation des k-Tupels $(1, \ldots, k)$, so gilt $s(q_1, \ldots, q_k) = (-1)^N$, wenn N die Anzahl der Vertauschungen zählt, die benötigt werden, um von $(q_1, \ldots, q_k)$ zu $(1, \ldots, k)$ zu gelangen.

Zu einer gegebenen $(k \times k)$-Matrix $\mathbf{A} = [a_{ij}]_{i,j=1}^{k}$ ist die Determinante definiert durch

$$\det \mathbf{A} = \sum_{q_1, \ldots, q_k = 1}^{k} s(q_1, \ldots, q_k) a_{1,q_1} a_{1,q_2} \cdots a_{k,q_k}.$$

Dies ist der *Leibnitzsche Entwicklungssatz für Determinaten*.

Satz 11.3.16. *Sei ω eine k-Form in der offenen Menge $E \subseteq \mathbb{R}^n$ und $\mathbf{\Phi} : D \subseteq \mathbb{R}^k \to E$ eine k-Fläche in E ($k \le n$). Desweiteren sei $\mathbf{\Psi} : D \to \mathbb{R}^k$, definiert durch $\mathbf{\Psi}(\mathbf{u}) = \mathbf{u}$ für alle $\mathbf{u} \in D$ eine k-Fläche in $\mathbb{R}^k$, dann gilt*

$$\int_{\mathbf{\Phi}} \omega = \int_{\mathbf{\Psi}} \omega_{\mathbf{\Phi}}.$$

Beweis. Ohne Beschränkung der Allgemeinheit sei $\omega = f(\mathbf{x})dx_{i_1} \wedge \ldots \wedge dx_{i_k}$ mit $1 \leq i_1 < \cdots < i_k \leq n$. Für $\mathbf{u} \in D$ ist $\mathbf{x} = \mathbf{\Phi}(\mathbf{u}) = [\Phi_1(\mathbf{u}), \ldots, \Phi_k(\mathbf{u})]^T \in E$. Nach Definition gilt

$$\omega_{\mathbf{\Phi}} = f\big(\mathbf{\Phi}(\mathbf{u})\big)d\Phi_{i_1} \wedge \ldots \wedge d\Phi_{i_k}.$$

Definieren wir uns

$$J(\mathbf{u}) := \det \frac{\partial(x_{i_1}, \ldots, x_{i_k})}{\partial(u_1, \ldots, u_k)} = \det \frac{\partial(\Phi_{i_1}, \ldots, \Phi_{i_k})}{\partial(u_1, \ldots, u_k)},$$

so folgt der Satz, wenn wir zeigen können, dass

$$d\Phi_{i_1} \wedge \ldots \wedge d\Phi_{i_k} = J(\mathbf{u})du_1 \wedge \ldots \wedge du_k, \tag{11.7}$$

denn diese Gleichung impliziert

$$\begin{aligned}\int_{\mathbf{\Phi}} \omega &= \int_D f\big(\mathbf{\Phi}(\mathbf{u})\big)J(\mathbf{u})d\mathbf{u} \\ &= \int_{\mathbf{\Psi}} (f \circ \mathbf{\Phi})(\mathbf{u})J(\mathbf{u})du_1 \wedge \ldots \wedge du_k \\ &= \int_{\mathbf{\Psi}} (f \circ \mathbf{\Phi})(\mathbf{u})d\Phi_{i_1} \wedge \ldots \wedge \Phi_{i_k} = \int_{\mathbf{\Psi}} \omega_{\mathbf{\Phi}}.\end{aligned}$$

Nun zum Beweis von (11.7). Für alle $p \in \{1, \ldots, k\}$ gilt

$$d\Phi_p = \sum_{q=1}^{k} \frac{\partial \Phi_p}{\partial u_q} du_q,$$

woraus sich das äußere Produkt

$$d\Phi_{i_1} \wedge \ldots \wedge d\Phi_{i_k} = \sum_{q_1,\ldots,q_k=1}^{k} \frac{\partial \Phi_{i_1}}{\partial u_{q_1}} \cdot \frac{\partial \Phi_{i_2}}{\partial u_{q_2}} \cdots \frac{\partial \Phi_{i_k}}{\partial u_{q_k}}\, du_{q_1} \wedge \ldots \wedge du_{q_k}$$

ergibt. Es gilt

$$du_{q_1} \wedge \ldots \wedge du_{q_k} = s(q_1, \ldots, q_k)\; du_1 \wedge \ldots \wedge du_k$$

und weiter

$$d\Phi_{i_1} \wedge \ldots \wedge d\Phi_{i_k} = \sum_{q_1,\ldots,q_k=1}^{k} \frac{\partial \Phi_{i_1}}{\partial u_{q_1}} \cdot \frac{\partial \Phi_{i_2}}{\partial u_{q_2}} \cdots \frac{\partial \Phi_{i_k}}{\partial u_{q_k}} s(q_1, \ldots, q_k)\; du_1 \wedge \ldots \wedge du_k.$$

Nun gilt aber nach dem Leibnitzschen Entwicklungssatz für Determinaten

$$J(\mathbf{u}) = \sum_{q_1,\ldots,q_k=1}^{k} \frac{\partial \Phi_{i_1}}{\partial u_{q_1}} \cdot \frac{\partial \Phi_{i_2}}{\partial u_{q_2}} \cdots \frac{\partial \Phi_{i_k}}{\partial u_{q_k}} s(q_1, \ldots, q_k).$$

Damit erhalten wir die Aussage (11.7) und schließlich die Behauptung. □

Satz 11.3.17. *Seien $E \subseteq \mathbb{R}^m$ und $F \subseteq \mathbb{R}^n$ offene Mengen und $\mathbf{t} \in C^1(E, F)$. Für eine eine k-Form ω in F und eine k-Fläche $\mathbf{\Phi} : D \to E$ in E gilt*

$$\int_{\mathbf{t}\circ\mathbf{\Phi}} \omega = \int_{\mathbf{\Phi}} \omega_{\mathbf{t}}.$$

Beweis. Es ist $\mathbf{t} \circ \mathbf{\Phi} \in C^1(D, F)$ eine k-Fläche in F und $\omega_{\mathbf{t}}$ eine k-Form in E. Mit der Identität $\mathbf{\Psi} : D \to D$ definiert durch $\mathbf{\Psi}(\mathbf{u}) = \mathbf{u}$ gilt nach den letzten beiden Sätzen

$$\int_{\mathbf{t}\circ\mathbf{\Phi}} \omega = \int_{\mathbf{\Psi}} \omega_{\mathbf{t}\circ\mathbf{\Phi}} = \int_{\mathbf{\Psi}} (\omega_{\mathbf{t}})_{\mathbf{\Phi}} = \int_{\mathbf{\Phi}} \omega_{\mathbf{t}}.$$

□

11.4 Simplizes und Ketten

Definition 11.4.1. *Seien X und Y Vektorräume. Eine Abbildung $\mathbf{f} : X \to Y$ heißt* **affin** *genau dann, wenn $\mathbf{f} - \mathbf{f}(\mathbf{0})$ linear ist, das heißt*

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{0}) + \mathbf{A}\mathbf{x}$$

mit $\mathbf{A} \in \mathcal{L}(X, Y)$.

Eine affine Abbildung $\mathbf{f} : \mathbb{R}^k \to \mathbb{R}^n$ ist folglich eindeutig gegeben durch $\mathbf{f}(\mathbf{0})$ und $\mathbf{f}(\mathbf{e}_i)$, $i = 1, \ldots, k$. Wie üblich bezeichne dabei die Menge $\{\mathbf{e}_1, \ldots, \mathbf{e}_k\}$ die Standardbasis im $\mathbb{R}^k$.

Definition 11.4.2. *Das* **Referenzsimplex** $\triangle^k \subseteq \mathbb{R}^k$ *ist die Menge*

$$\triangle^k = \{\mathbf{u} = \sum_{i=1}^k \alpha_i \mathbf{e}_i \in \mathbb{R}^k : \alpha_i \geq 0,\ \sum_{i=1}^k \alpha_i \leq 1\}.$$

Seien $\mathbf{p}_0, \ldots, \mathbf{p}_k$ Punkte von $\mathbb{R}^n$. Das **orientierte affine k-Simplex** $\sigma = [\mathbf{p}_0, \mathbf{p}_1, \ldots, \mathbf{p}_k] \subseteq \mathbb{R}^k$ *sei definiert als die k-Fläche im $\mathbb{R}^n$, die durch die affine Abbildung*

$$\sigma\left(\sum_{i=1}^k \alpha_i \mathbf{e}_i\right) = \mathbf{p}_0 + \sum_{i=1}^k \alpha_i (\mathbf{p}_i - \mathbf{p}_0)$$

aus dem Referenzsimplex $\triangle^k$ entsteht.

Die affinie Abbildung $\sigma : \triangle^k \to \sigma = [\mathbf{p}_0, \ldots, \mathbf{p}_k]$ ist festgelegt durch $\sigma(\mathbf{0}) = \mathbf{p}_0$ und $\sigma(\mathbf{e}_i) = \mathbf{p}_i$, wobei $\mathbf{p}_i$, $i = 1, \ldots, k$, die Ecken des k-Simplex σ sind. Folglich ist

$$\sigma(\mathbf{u}) = \mathbf{p}_0 + \mathbf{A}\mathbf{u}$$

mit der Matrix

$$\mathbf{A} = \left[\ \mathbf{p}_1 - \mathbf{p}_0 \mid \mathbf{p}_2 - \mathbf{p}_0 \mid \cdots \mid \mathbf{p}_k - \mathbf{p}_0\ \right] \in \mathbb{R}^{n\times k}.$$

Dabei ist natürlich die Reihenfolge der Eckpunkte $\mathbf{p}_0, \ldots, \mathbf{p}_k$ wichtig. Daher nennen wir das k-Simplex σ orientiert, um hervorzuheben, dass die Reihenfolge der Ecken berücksichtigt wird. Dies rechtfertigt unsere Notation $\sigma = [\mathbf{p}_0, \ldots, \mathbf{p}_k]$.

Eine Permutation der Reihenfolge der Ecken ergibt ein geändertes k-Simplex $\overline{\sigma}$ mit einer anderen Orientierung. Bezeichnet geäß (11.6) $s(i_0, \ldots, i_k)$ die Signatur der Permutation $(i_0, \ldots, i_k)$ des k-Tuples $(0, \ldots, k)$, so schreiben wir für

$$\overline{\sigma} = [\mathbf{p}_{i_0}, \mathbf{p}_{i_1}, \ldots, \mathbf{p}_{i_k}]$$

einfach kurz $\overline{\sigma} := s(i_0, \ldots, i_k) \cdot \sigma$. Demnach ist $\overline{\sigma} = \pm\sigma$, je nachdem ob $s = 1$ oder $s = -1$ gilt.

Gilt $k = n$, so ist die Orientierung berechenbar. Denn sind $\mathbf{p}_1 - \mathbf{p}_0, \ldots, \mathbf{p}_n - \mathbf{p}_0$ linear unabhängig, so folgt $\det \mathbf{A} \neq 0$. Die Orientierung von σ ist folglich $\operatorname{sgn}(\det \mathbf{A})$. Das Referenzsimplex $\triangle^k \subseteq \mathbb{R}^k$ ist immer positiv orientiert.

Definition 11.4.3. *Das* **Nullsimplex** *definieren wir als einen Punkt $\mathbf{p}_0 \in \mathbb{R}^n$ zusammen mit einem Vorzeichen. Wir schreiben $\sigma = +\mathbf{p}_0$ oder $\sigma = -\mathbf{p}_0$. Ist $\sigma = \epsilon\mathbf{p}_0$ mit $\epsilon \pm 1$, und f eine Nullform im $\mathbb{R}^n$, das heißt $f \in C(\mathbb{R}^n, \mathbb{R})$, so definieren wir*

$$\int_\sigma f = \epsilon f(\mathbf{p}_0).$$

Satz 11.4.4. *Sei σ ein affiner, orientierter k-Simplex in einer offenen Menge $E \subseteq \mathbb{R}^n$ und $\overline{\sigma} = \epsilon\sigma$ ($\epsilon = \pm 1$), dann gilt*

$$\int_{\overline{\sigma}} \omega = \epsilon \int_\sigma \omega$$

für jede k-Form ω in E.

Beweis. Wir zeigen zuerst

$$\int_\sigma \omega = \int_{\triangle^k} \sum a_{i_1,\ldots,i_k}(\mathbf{p}_0 + \mathbf{A}\mathbf{u}) \det \frac{\partial(x_{i_1}, \ldots, x_{i_k})}{\partial(u_1, \ldots, u_k)} d\mathbf{u}.$$

Für $k = 0$ stimmt diese Gleichung mit der obigen Definition überein. Sei also $k > 0$ und $\sigma = [\mathbf{p}_0, \ldots, \mathbf{p}_k] = \mathbf{p}_0 + \mathbf{A}\mathbf{u}$ mit $A = [\mathbf{p}_1 - \mathbf{p}_0, \ldots, \mathbf{p}_k - \mathbf{p}_0] \in \mathbb{R}^{n\times n}$.

Wir nehmen an, dass für ein $1 \leq j \leq k$ das k-Simplex $\overline{\sigma}$ aus dem k-Simplex σ durch Vertauschen von $\mathbf{p}_0$ und $\mathbf{p}_j$ entsteht, das heißt

$$\overline{\sigma} = [\mathbf{p}_j, \mathbf{p}_1, \ldots, \mathbf{p}_{j-1}, \mathbf{p}_0, \mathbf{p}_{j+1}, \ldots, \mathbf{p}_k] = \mathbf{p}_j + \mathbf{B}\mathbf{u}$$

mit

$$\mathbf{B} = [\mathbf{p}_1 - \mathbf{p}_j, \mathbf{p}_0 - \mathbf{p}_j, \ldots, \mathbf{p}_{j-1} - \mathbf{p}_j, \mathbf{p}_0 - \mathbf{p}_j, \mathbf{p}_{j+1} - \mathbf{p}_j, \ldots, \mathbf{p}_k - \mathbf{p}_j].$$

Schreiben wir $\mathbf{x}_i := \mathbf{A}\mathbf{e}_i$, so gilt

$$\mathbf{B} = [\mathbf{x}_1 - \mathbf{x}_j, \ldots, \mathbf{x}_{j-1} - \mathbf{x}_j, -\mathbf{x}_j, \mathbf{x}_{j+1} - \mathbf{x}_j, \ldots, \mathbf{x}_k - \mathbf{x}_j].$$

Subtrahieren wir jetzt die j-te Spalte von $\mathbf{B}$ von den anderen Spalte, so wird keine der Determinanten von (11.1) verändert und wir erhalten die neuen Spalten

$$[\mathbf{x}_1, \ldots, \mathbf{x}_{j-1}, -\mathbf{x}_j, \mathbf{x}_{j+1}, \ldots, \mathbf{x}_k].$$

Diese Matrix unterscheidet sich nur durch das Vorzeichen der j-tem Spalte von $\mathbf{A}$. Also gilt die Behauptung in diesem Fall.

Nun sei $1 \leq i < j \leq k$ und $\overline{\sigma}$ entstehe aus σ durch Vertauschen von $\mathbf{p}_i$ und $\mathbf{p}_j$, das heißt $\overline{\sigma} = \mathbf{p}_0 + \mathbf{C}\mathbf{u}$ mit

$$\begin{aligned}\mathbf{C} = [&\mathbf{p}_1 - \mathbf{p}_0, \ldots, \mathbf{p}_{i-1} - \mathbf{p}_0, \mathbf{p}_j - \mathbf{p}_0, \mathbf{p}_{i+1} - \mathbf{p}_0, \\ &\ldots, \mathbf{p}_{j-1} - \mathbf{p}_0, \mathbf{p}_i - \mathbf{p}_0, \mathbf{p}_{j+1} - \mathbf{p}_0, \ldots, \mathbf{p}_k - \mathbf{p}_0].\end{aligned}$$

Es sind folglich im Vergleich zu A zwei Spalten vertauscht, das heißt $\epsilon = -1$, und folglich ergibt sich die Behauptung.

Der allgemeine Fall folgt nun aufgrund der Tatsache, dass sich jede Permutation aus den obigen Spezialfällen erzeugen lässt. Benötigen wir N Vertauschungen, um $(i_0, \ldots, i_k)$ überzuführen in $(1, \ldots, k)$, so gilt $\epsilon = (-1)^N = s(i_0, \ldots, i_k)$. □

Definition 11.4.5. *Eine* **affine k-Kette** Γ *in einer offenen Menge $E \subseteq \mathbb{R}^n$ ist eine Famile endlich vieller orientierter affiner k-Simplizes $\sigma_1, \ldots, \sigma_r$ in E. Diese müssen nicht notwendig voneinander verschieden sein, das heißt, ein Simplex darf in Γ mehrfach vorkommen.*

Ist Γ eine solche affinie k-Kette und ist ω eine k-Form in E, dann definieren wir

$$\int_\Gamma \omega = \sum_{i=1}^{r} \int_{\sigma_i} \omega. \tag{11.8}$$

Bemerkung 11.4.6. *Mann kann eine k-Fäche $\mathbf{\Phi}$ in E als eine Funktion ansehen, deren Definitionsbereich die Familie aller k-Formen in E ist und die der Differentialform ω die Zahl $\int_{\mathbf{\Phi}} \omega$ zuordnet. Da reellwertige Funktionen addiert werden können, liegt es nahe, die Schreibweise*

$$\Gamma = \sigma_1 + \sigma_2 + \ldots + \sigma_r = \sum_{i=1}^{r} \sigma_i \tag{11.9}$$

zu verwenden, um die Tatsache auszudrücken, dass (11.8) *für jede k-Form ω in E gilt.*

Wir hatten σ als eine affine Abbildung des Referenzsimplex $\triangle^k$ definiert. Wir weisen an dieser Stelle ausdrücklich darauf hin, dass die Summe auf der linken Seite von (11.9) **nicht** *als Funktion $\sigma : \triangle^k \to \mathbb{R}^n$ definiert durch*

$$\sigma(\mathbf{u}) = \sum_{i=1}^{r} \sigma_i(\mathbf{u})$$

zu verstehen ist.

Ist zum Beispiel $\sigma_1 = -\sigma_2$, das heißt σ_1 und σ_2 bilden den selben Bildbereich, sind aber entgegengesetzt orientiert, und ist $\Gamma = \sigma_1 + \sigma_2$, dann ist $\int_\Gamma \omega = 0$ für alle ω, was wir durch die Schreibweise $\Gamma = 0$ bzw. $\sigma_1 + \sigma_2 = 0$ ausdrücken. Dies bedeutet aber nicht, dass $\sigma_1(\mathbf{u}) + \sigma_2(\mathbf{u})$ identisch dem Nullvektor im $\mathbb{R}^n$ ist.

Definition 11.4.7. *Sei σ ein affines orientierter k-Simplex und $k \geq 1$. Dann ist die $(k-1)$-Kette*

$$\partial\sigma = \sum_{j=0}^{k} (-1)^j [\mathbf{p}_0, \ldots, \mathbf{p}_{j-1}, \mathbf{p}_{j+1}, \ldots, \mathbf{p}_k]$$

der **Rand** *von σ.*

Beispiel 11.4.8. *Für die Punkte $\mathbf{p}_0, \mathbf{p}_1, \mathbf{p}_2 \in \mathbb{R}^2$ definiert das Zweisimplex $\sigma = [\mathbf{p}_0, \mathbf{p}_1, \mathbf{p}_2]$ eine Abbildung*

$$\sigma : \triangle^2 = \triangle([\begin{smallmatrix}0\\0\end{smallmatrix}], [\begin{smallmatrix}1\\0\end{smallmatrix}], [\begin{smallmatrix}0\\1\end{smallmatrix}]) \to \triangle(\mathbf{p}_0, \mathbf{p}_1, \mathbf{p}_2) \subseteq \mathbb{R}^2.$$

Sein Rand ist gegeben durch die Einskette

$$\partial\sigma = [\mathbf{p}_1, \mathbf{p}_2] - [\mathbf{p}_0, \mathbf{p}_2] + [\mathbf{p}_0, \mathbf{p}_1] = [\mathbf{p}_1, \mathbf{p}_2] + [\mathbf{p}_2, \mathbf{p}_0] + [\mathbf{p}_0, \mathbf{p}_1].$$

Definition 11.4.9. *Seien $E \subseteq \mathbb{R}^m$ und $F \subseteq \mathbb{R}^n$ offene Mengen und $\mathbf{T} \in C^2(E, F)$. Ferner sei $\sigma = [\mathbf{p}_0, \ldots, \mathbf{p}_k]$ ein orientiertes affines k-Simplex in E, das heißt es gilt $\mathbf{p}_0, \ldots, \mathbf{p}_k \in E$ und*

$$\sigma(\mathbf{u}) = \sigma(u_1, \ldots, u_k) = \mathbf{p}_0 + \sum_{i_1}^{k} u_i(\mathbf{p}_i - \mathbf{p}_0)$$

für alle

$$\mathbf{u} \in \triangle^k = \left\{ \mathbf{u} = [u_1, \ldots, u_k]^T \in \mathbb{R}^k : u_i \geq 0,\ i = 1, \ldots, k,\ \sum_{i=1}^{k} u_i \leq 1 \right\}.$$

Dann ist $\mathbf{\Phi} = \mathbf{T} \circ \sigma$ eine k-Fläche in F mit Parameterbereich $\triangle^k$. Wir schreiben hierfür auch auch kurz $\mathbf{\Phi} = \mathbf{T}\sigma$ und nennen $\mathbf{\Phi}$ ein **orientiertes k-Simplex der Klasse** *C^2 bzw. ein* **differenzierbares Simplex**.

*Eine endliche Familie $\mathbf{\Psi}$ von orientierten Simplizes $\mathbf{\Phi}_1, \ldots, \mathbf{\Phi}_N$ heißt eine k-***Kette der Klasse** *C^2. Ist ω ein k-Form in F, dann definieren wir*

$$\int_{\mathbf{\Psi}} \omega = \sum_{i_1}^{N} \int_{\mathbf{\Phi}_i} \omega$$

und verwenden formal die Schreibweise $\mathbf{\Psi} = \sum_{i=1}^{N} \mathbf{\Phi}_i$.

Ist $\Gamma = \sum_{i=1}^{N} \sigma_i$ *eine affine* k*-Kette und ist für jedes* $i = 1, \ldots, n$ *das Simplex* $\mathbf{\Phi}_i = \mathbf{T} \circ \sigma_i$ *differenzierbar, so schreiben wir*

$$\mathbf{\Psi} = \mathbf{T} \circ \Gamma = \sum_{i=1}^{N} \mathbf{\Phi}_i = \sum_{i=1}^{N} \mathbf{T} \circ \sigma_i.$$

Der Rand $\partial\mathbf{\Phi}$ *des orientierten* k*-Simplex* $\mathbf{\Phi} = \mathbf{T} \circ \sigma$ *sei als die* $(k-1)$*-Kette*

$$\partial\mathbf{\Phi} = \mathbf{T} \circ \partial\sigma$$

definiert.

Beispiel 11.4.10. *1. Schreiben wir abkürzend*

$$\mathbf{1} := \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \mathbf{e}_1 + \mathbf{e}_2,$$

dann sind

$$\sigma_1 = [\mathbf{0}, \mathbf{e}_1, \mathbf{e}_2], \qquad \sigma_2 = [\mathbf{1}, \mathbf{e}_2, \mathbf{e}_1]$$

affine orientierte 2*-Simplizes, wobei* σ_1 *dem orientierten Referenzsimplex entspricht. Für* $\mathbf{u} = (u_1, u_2) \in \triangle^2$ *gilt*

$$\sigma_2(\mathbf{u}) = \begin{bmatrix} 1 \\ 1 \end{bmatrix} + u_1 \left(\begin{bmatrix} 0 \\ 1 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right) + u_2 \left(\begin{bmatrix} 1 \\ 0 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right) = \begin{bmatrix} 1 \\ 1 \end{bmatrix} - \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}.$$

Wir bemerken, dass $\sigma_1 + \sigma_2 = [0,1]^2 =: \Omega$.

Per Definition ist der Rand von σ_1 *bzw.* σ_2 *eine Einskette. Wir finden*

$$\begin{aligned} \partial\sigma_1 &:= [\mathbf{e}_1, \mathbf{e}_2] + (-1)^1[\mathbf{0}, \mathbf{e}_2] + (-1)^2[\mathbf{0}, \mathbf{e}_1] = +[\mathbf{e}_1, \mathbf{e}_2] + [\mathbf{e}_2, \mathbf{0}] + [\mathbf{0}, \mathbf{e}_1], \\ \partial\sigma_2 &:= [\mathbf{e}_2, \mathbf{e}_1] + (-1)^1[\mathbf{1}, \mathbf{e}_1] + (-1)^2[\mathbf{1}, \mathbf{e}_2] = -[\mathbf{e}_1, \mathbf{e}_2] + [\mathbf{e}_1, \mathbf{1}] + [\mathbf{1}, \mathbf{e}_2]. \end{aligned}$$

Somit folgt nach obiger Konvention

$$\partial\Omega = \partial\sigma_1 + \partial\sigma_2 = [\mathbf{0}, \mathbf{e}_1] + [\mathbf{e}_1, \mathbf{1}] + [\mathbf{1}, \mathbf{e}_2] + [\mathbf{e}_2, \mathbf{0}].$$

2. Sei $\mathbf{\Phi} : \Omega = [0, \pi] \times [0, 2\pi] \to \mathbb{R}^3$ *definiert durch*

$$\mathbf{\Phi}(u, v) := \begin{bmatrix} \cos v \sin u \\ \sin v \sin u \\ \cos u \end{bmatrix},$$

dann ist $\mathbf{\Phi}(\Omega)$ ist die Oberfläche der Einheitskugel. Dann ist der Rand von $\mathbf{\Phi}$ die Einskette $\partial\mathbf{\Phi} = \boldsymbol{\gamma}_1 + \boldsymbol{\gamma}_2 + \boldsymbol{\gamma}_3 + \boldsymbol{\gamma}_4$ mit

$$\begin{aligned}
\boldsymbol{\gamma}_1(t) &= \begin{bmatrix} \sin t \\ 0 \\ \cos t \end{bmatrix} && \textit{für } t = u \in [0,\pi],\ v = 0,\\
\boldsymbol{\gamma}_2(t) &= \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} && \textit{für } t = v \in [0,2\pi],\ u = \pi,\\
\boldsymbol{\gamma}_3(t) &= \begin{bmatrix} \sin(\pi - t) \\ 0 \\ \cos(\pi - t) \end{bmatrix} && \textit{für } t = \pi - u \in [0,\pi],\ v = 2\pi,\\
\boldsymbol{\gamma}_4(t) &= \begin{bmatrix} 0 \\ 0 \\ +1 \end{bmatrix} && \textit{für } t = 2\pi - v \in [0,2\pi],\ u = 0.
\end{aligned}$$

Wir sehen, dass hierbei insbesondere $\boldsymbol{\gamma}_3(t) = \boldsymbol{\gamma}_1(\pi - t)$ gilt. Ferner sind $\boldsymbol{\gamma}_2$ und $\boldsymbol{\gamma}_4$ konstant, so dass $\boldsymbol{\gamma}_2'(t) = \boldsymbol{\gamma}_4'(t) = \mathbf{0}$.

Ist ω eine beliebige Einsform im $\mathbb{R}^3$, dann gilt

$$\int_{\boldsymbol{\gamma}_3} \omega = -\int_{\boldsymbol{\gamma}_1} \omega.$$

Somit folgt

$$\int_{\partial\mathbf{\Phi}} \omega = \int_{\boldsymbol{\gamma}_1} \omega + \int_{\boldsymbol{\gamma}_2} \omega + \int_{\boldsymbol{\gamma}_3} \omega + \int_{\boldsymbol{\gamma}_4} \omega = \int_{\boldsymbol{\gamma}_1} \omega + \int_{\boldsymbol{\gamma}_3} \omega = 0$$

für alle Einsformen im $\mathbb{R}^3$.

11.5 Der Satz von Stokes

Wir beweisen nun den Satz von Stokes, den wir als ultimative n-dimensionale Version des Hauptsatzes der Integralrechnung ansehen können, denn im eindimensionalen Fall entspricht er dem Hauptsatz der Integralrechnung.

Satz 11.5.1 (Satz von Stokes). *Sei $\mathbf{\Psi}$ eine k-Kette der Klasse C^2 in der offenen Menge $V \subseteq \mathbb{R}^m$. Sei weiterhin ω eine $(k-1)$-Form der Klasse C^2 in V. Dann gilt*

$$\int_{\partial\mathbf{\Psi}} \omega = \int_{\mathbf{\Psi}} d\omega.$$

Beweis. Da gilt $\mathbf{\Psi} = \sum_{i=1}^{N} \mathbf{\Phi}_i$ mit differenzierbaren orientierten k-Simplizes $\mathbf{\Phi}_i$ und da $\partial\mathbf{\Psi} = \sum_{i=1}^{N} \partial\mathbf{\Phi}_i$ eine $(k-1)$-Kette ist, genügt es, den Fall $N = 1$ zu betrachten, das heißt $\mathbf{\Psi} = \mathbf{\Phi}$.

Sei $\sigma = [\mathbf{0}, \mathbf{e}_1, \ldots, \mathbf{e}_n]$ das Referenzsimplex, das heißt

$$\sigma : \triangle_k \to \triangle_k, \quad \sigma(\mathbf{u}) = \mathbf{u}.$$

für alle

$$\mathbf{u} \in \triangle^k := \left\{ (u_1, \ldots, u_k)^T \in \mathbb{R}^k : u_i \geq 0, i = 1, \ldots, k, \ \sum_{i=1}^{k} u_i \leq 1 \right\}.$$

Da $\triangle^k \subseteq \mathbb{R}^k$ abgeschlossen ist und $V \subseteq \mathbb{R}^m$ offen ist, gibt es wegen $\mathbf{\Phi} \in C^2(\triangle^k, V)$ eine offene Menge $D \subseteq \mathbb{R}^k$ mit $\triangle^k \subseteq D$ und $\mathbf{\Phi} \in C^2(D, V)$. Folglich existiert eine Abbildung $\mathbf{t} \in C^2(D, V)$ mit $\mathbf{\Phi} = \mathbf{t} \circ \sigma$. Mit den Sätzen 11.3.14 und 11.3.17 gilt dann

$$\int_{\mathbf{\Phi}} d\omega = \int_{\mathbf{t}\circ\sigma} d\omega = \int_{\sigma} (d\omega)_{\mathbf{t}} = \int_{\sigma} d\omega_{\mathbf{t}}.$$

Wegen $\partial\mathbf{\Phi} = \partial(\mathbf{t} \circ \sigma) = \mathbf{t}(\partial\sigma)$ ist

$$\int_{\partial\mathbf{\Phi}} \omega = \int_{\mathbf{t}(\partial\sigma)} \omega = \int_{\partial\sigma} \omega_{\mathbf{t}}.$$

Da $\omega_{\mathbf{t}}$ eine $(k-1)$-Form in E ist, müssen wir folglich lediglich die Beziehung

$$\int_{\partial\sigma} \lambda = \int_{\sigma} d\lambda$$

für das Referenzsimplex $\sigma = [\mathbf{0}, \mathbf{e}_1, \ldots, \mathbf{e}_n]$ und beliebige $(k-1)$-Formen λ zeigen.

Für $f \in C^1(D, \mathbb{R})$ sei ohne Beschränkung der Allgemeinheit

$$\lambda = f(\mathbf{x})\, dx_1 \wedge \ldots \wedge dx_{r-1} \wedge dx_{r+1} \wedge \ldots \wedge dx_k,$$

das heißt dx_r fehlt für ein festes $r \in \{1, \ldots, k\}$. Der positiv orientierte Rand des Referenzsimplex σ ist per Definition die $(k-1)$-Kette

$$\partial\sigma = [\mathbf{e}_1, \ldots, \mathbf{e}_k] + \sum_{i=1}^{k} (-1)^i \tau_i$$

mit

$$\tau_i = [\mathbf{0}, \mathbf{e}_1, \ldots, \mathbf{e}_{i-1}, \mathbf{e}_{i+1}, \ldots, \mathbf{e}_k].$$

Definieren wir

$$\tau_0 := [\mathbf{e}_r, \mathbf{e}_1, \ldots, \mathbf{e}_{r-1}, \mathbf{e}_{r+1}, \ldots, \mathbf{e}_k],$$

dann gilt $[\mathbf{e}_1, \ldots, \mathbf{e}_k] = (-1)^{r-1}\tau_0$ und daher

$$\partial\sigma = (-1)^{r-1}\tau_0 + \sum_{i=1}^{k}(-1)^k\tau_i.$$

Dabei besitzt jedes τ_i den Parameterbereich $\triangle^{k-1}$. Für gegebenes $\mathbf{u} = [u_1, \ldots, u_{k-1}]^T \in \triangle^{k-1}$ ist $\mathbf{x} = \tau_0(\mathbf{u}) \in \mathbb{R}^k$ gegeben durch

$$\mathbf{x} = \mathbf{e}_r + \sum_{i=1}^{r-1} u_i(\mathbf{e}_i - \mathbf{e}_r) + \sum_{i=r+1}^{k} u_{i-1}[\mathbf{e}_i - \mathbf{e}_r],$$

beziehungsweise

$$x_j = \begin{cases} u_j, & j = 1, \ldots, r-1, \\ 1 - (u_1 + \ldots + u_{k-1}), & j = r, \\ u_{j-1}, & j = r+1, \ldots, k. \end{cases} \tag{11.10}$$

Für $i = 1, \ldots, k$ ist $\mathbf{x} = \tau_i(\mathbf{u})$ gegeben durch

$$\mathbf{x} = \sum_{l=1}^{i-1} u_l\mathbf{e}_l + \sum_{i+1}^{k} u_{k-1}\mathbf{e}_l,$$

beziehungsweise

$$x_j = \begin{cases} u_j, & j = 1, \ldots, i-1, \\ 0, & j = i, \\ u_{j-1}, & j = i+1, \ldots, k. \end{cases} \tag{11.11}$$

Folglich gilt für $i = 0, 1, \ldots, k$

$$J_{\tau_i}(\mathbf{u}) := \det \frac{\partial(x_1, \ldots, x_{r-1}, x_{r+1}, \ldots, x_k)}{\partial(u_1, \ldots, u_k)}(\mathbf{u}) = \begin{cases} 1, & i = 0, \\ 1, & i = r, \\ 0, & \text{sonst}, \end{cases}$$

da wir in den ersten beiden Fällen die Determinante der Einheitsmatrix bilden während wir im letzten Fall die Determinante einer Matrix mit Nullzeile bilden. Es gilt folglich

$$\int_{\tau_i} \lambda = \int_{\triangle^{k-1}} f\big(\tau_i(\mathbf{u})\big)d\mathbf{u} = 0$$

für $i \neq 0$ und $i \neq r$, und daher folgt

$$\begin{aligned}
\int_{\partial\sigma} \lambda &= (-1)^{r-1} \int_{\tau_0} \lambda + \sum_{i=1}^{k} (-1)^i \int_{\tau_i} \lambda \\
&= (-1)^{r-1} \int_{\tau_0} \lambda + (-1)^r \int_{\tau_r} \lambda \\
&= (-1)^{r-1} \left[\int_{\tau_0} \lambda - \int_{\tau_r} \lambda \right] \\
&= (-1)^{r-1} \int_{\triangle^{k-1}} \big[f\big(\tau_0(\mathbf{u})\big) J_{\tau_0}(\mathbf{u}) - f\big(\tau_r(\mathbf{u})\big) J_{\tau_r}(\mathbf{u}) \big] d\mathbf{u} \\
&= (-1)^{r-1} \int_{\triangle^{k-1}} \big[f\big(\tau_0(\mathbf{u})\big) - f\big(\tau_r(\mathbf{u})\big) \big] d\mathbf{u}. \qquad (11.12)
\end{aligned}$$

Andererseits gilt

$$\begin{aligned}
d\lambda &= \sum_{i=1}^{k} \frac{\partial f}{\partial x_i}\, dx_i \wedge dx_1 \wedge \ldots \wedge dx_{r-1} \wedge dx_{r+1} \wedge \ldots \wedge dx_k \\
&= \frac{\partial f}{\partial x_r}\, dx_r \wedge dx_1 \wedge \ldots \wedge dx_{r-1} \wedge dx_{r+1} \wedge \ldots \wedge dx_k \\
&= (-1)^{r-1} \frac{\partial f}{\partial x_r}\, dx_1 \wedge \ldots \wedge dx_{r-1} \wedge dx_r \wedge dx_{r+1} \wedge \ldots \wedge dx_k \\
&= (-1)^{r-1} \frac{\partial f}{\partial x_r} d\mathbf{x},
\end{aligned}$$

das heißt

$$\int_\sigma d\lambda = (-1)^{r-1} \int_{\triangle^k} \frac{\partial f}{\partial x_r} d\mathbf{x}.$$

Wir setzen

$$\mathbf{u} = [u_1, \ldots, u_{k-1}]^T = [x_1, \ldots, x_{r-1}, x_{r+1}, \ldots, x_k]^T \in \triangle^{k-1}$$

und berechnen letzteres Integral, indem wir bezüglich x_r über das Intervall

$$[0, 1 - (x_1 + \cdots + x_{r-1} + x_{r+1} + \cdots + x_k)] = \left[0, 1 - \sum_{i=1}^{k} u_i\right]$$

integrieren,

$$\begin{aligned}
\int_\sigma d\lambda &= \int_{\triangle^{k-1}} \left[\int_0^{1-\sum_{i=1}^k u_i} \frac{\partial f}{\partial x_r}(u_1, \ldots, u_{r-1}, x_r, u_r, \ldots, u_{k-1}) dx_r \right] d\mathbf{u} \\
&= \int_{\triangle^{k-1}} \big[f\big(u_1, \ldots, u_{r-1}, 1 - (u_1 + \cdots + u_{k-1}), u_r, \ldots, u_{k-1}\big) \\
&\qquad\qquad - f(u_1, \ldots, u_{r-1}, 0, u_r, \ldots, u_{k-1}) \big] d\mathbf{u}.
\end{aligned}$$

Gemäß (11.10) und (11.11) ergibt sich hieraus

$$\int_\sigma d\lambda = \int_{\triangle^{k-1}} \big[f\big(\tau_0(\mathbf{u})\big) - f\big(\tau_r(\mathbf{u})\big)\big]\, d\mathbf{u}.$$

Vergleichen wir diesen ausdruck mit (11.12) erhalten wir

$$\int_{\partial\sigma} \lambda = \int_\sigma d\lambda$$

und damit die Behauptung. □

11.6 Anwendungen des Satzes von Stokes

Definition 11.6.1. *Sei $D \subseteq \mathbb{R}^3$ eine offene Menge und $\mathbf{f} = [f_1, f_2, f_3]^T : D \to \mathbb{R}^3$ ein Vektorfeld. Zu $\mathbf{F}$ definieren wir die* **Rotation** *von $\mathbf{f}$ vermittels*

$$\operatorname{rot} \mathbf{f} := \nabla \times \mathbf{f} = \mathbf{e}_1 \left\{ \frac{\partial f_3}{\partial x_2} - \frac{\partial f_2}{\partial x_3} \right\} + \mathbf{e}_2 \left\{ \frac{\partial f_1}{\partial x_3} - \frac{\partial f_3}{\partial x_1} \right\} + \mathbf{e}_3 \left\{ \frac{\partial f_2}{\partial x_1} - \frac{\partial f_1}{\partial x_2} \right\}$$

und die **Divergenz** *von $\mathbf{f}$ vermittels*

$$\operatorname{div} \mathbf{f} := \nabla \cdot \mathbf{f} = \frac{\partial f_1}{\partial x_1} + \frac{\partial f_2}{\partial x_2} + \frac{\partial f_3}{\partial x_3}.$$

Satz 11.6.2 (Integralsatz von Stokes). *Sei $E \subseteq \mathbb{R}^3$ eine offene Menge und $\mathbf{f} : C^1(E, \mathbb{R}^3)$ ein gegebenes Vektorfeld. Sei $\mathbf{\Phi} : D \subseteq \mathbb{R}^2 \to \mathbb{R}^3$ eine Zweifläche der Klasse C^2 in E. Wir setzen $\Gamma := \mathbf{\Phi}(D)$, wobei $\partial\Gamma$ den Rand von Γ bezeichne. Es gilt*

$$\int_\Gamma \langle \operatorname{rot} \mathbf{f}, \mathbf{n} \rangle \; do = \int_{\partial\Gamma} \langle \mathbf{f}, \mathbf{t} \rangle \; ds.$$

Beweis. Für $\mathbf{f} = [f_1, f_2, f_3]^T \in C^1(E, \mathbb{R}^3)$ definieren wir in E die Einsform

$$\lambda_{\mathbf{f}} := f_1(\mathbf{x})dx + f_2(\mathbf{x})dy + f_3(\mathbf{x})dz.$$

Ferner definieren wir in E für $\mathbf{g} = [g_1, g_2, g_3]^T := \operatorname{rot} \mathbf{f} \in C(E, \mathbb{R}^3)$ die Zweiform

$$\omega_{\mathbf{g}} := g_1(\mathbf{x}) \; dy \wedge dz + g_2(\mathbf{x}) \; dz \wedge dx + g_3(\mathbf{x}) \; dx \wedge dy.$$

Wie man leicht nachprüft gilt dann $d\lambda_{\mathbf{f}} = \omega_{\mathbf{g}}$ und somit ergibt sich einerseits

$$\int_{\mathbf{\Phi}} \omega_{\mathbf{g}} = \int_\Gamma \langle \mathbf{g}, \mathbf{n} \rangle do.$$

Andereseits gilt aber auch

$$\int_{\partial\mathbf{\Phi}} \lambda_{\mathbf{f}} = \int_{\partial\Gamma} \langle \mathbf{f}, \mathbf{t} \rangle ds.$$

Nach dem Satz von Stokes folgt daher

$$\int_{\Gamma} \langle \operatorname{rot} \mathbf{f}, \mathbf{n} \rangle ds = \int_{\mathbf{\Phi}} \omega_{\mathbf{g}} = \int_{\mathbf{\Phi}} d\lambda_{\mathbf{f}} = \int_{\partial\mathbf{\Phi}} \lambda_{\mathbf{f}} = \int_{\partial\Gamma} \langle \mathbf{f}, \mathbf{t} \rangle ds.$$

□

Satz 11.6.3 (Integralsatz von Gauß). *Es sei $E \subseteq \mathbb{R}^3$ offen und $\Omega \subseteq E$ eine abgeschlossene Menge mit orientierbaren Rand $\Gamma := \partial\Omega$. Für $\mathbf{f} \in C^1(E, \mathbb{R})$ gilt dann*

$$\int_{\Omega} \operatorname{div} \mathbf{f} \; d\mathbf{x} = \int_{\Gamma} \langle \mathbf{f}, \mathbf{n} \rangle \; do$$

Beweis. Wir schreiben $\mathbf{f} = [f_1, f_2, f_3]^T$ und betrachten die Zweiform

$$\omega_{\mathbf{f}} := f_1(\mathbf{x}) \; dy \wedge dz + f_2(\mathbf{x}) \; dz \wedge dx + f_3(\mathbf{x}) \; dx \wedge dy.$$

Es gilt

$$\begin{aligned} d\omega_{\mathbf{f}} &= \frac{\partial f_1}{\partial x} \; dx \wedge dy \wedge dz + \frac{\partial f_2}{\partial y} \; dy \wedge dz \wedge dx + \frac{\partial f_3}{\partial z} \; dz \wedge dx \wedge dy \\ &= \left[\frac{\partial f_1}{\partial x} + (-1)^2 \frac{\partial f_2}{\partial y} + (-1)^2 \frac{\partial f_3}{\partial z} \right] dx \wedge dy \wedge dz \\ &= \operatorname{div} \mathbf{f} \; dx \wedge dy \wedge dz. \end{aligned}$$

Nach dem Satz von Stokes folgt für $\mathbf{\Phi} : \Omega \to \mathbb{R}^3$ schließlich

$$\int_{\partial\Omega} \langle \mathbf{f}, \mathbf{n} \rangle do = \int_{\partial\mathbf{\Phi}} \omega_{\mathbf{f}} = \int_{\mathbf{\Phi}} d\omega_{\mathbf{f}} = \int_{\Omega} \operatorname{div} \mathbf{f} d\mathbf{x}.$$

□

Satz 11.6.4 (Integralsatz von Green). *Es sei $E \subseteq \mathbb{R}^3$ offen und $\Omega \subseteq E$ eine abgeschlossene Menge mit orientierbaren Rand $\Gamma := \partial\Omega$. Für $f, g \in C^1(E, \mathbb{R})$ gilt dann die Vektorgleichung*

$$\int_{\Omega} (\nabla g \cdot f + g \cdot \nabla f) d\mathbf{x} = \int_{\Gamma} fg \; \mathbf{n} \; do.$$

Beweis. Wir beweisen die obige Gleichung komponentenweise. Für die Zweiformen

$$\boldsymbol{\lambda} = f(\mathbf{x}) g(\mathbf{x}) \begin{bmatrix} dy \wedge dz \\ dz \wedge dx \\ dx \wedge dy \end{bmatrix}$$

gilt

$$d\boldsymbol{\lambda} = (g \nabla f + f \nabla g)(dx \wedge dy \wedge dz).$$

Somit folgt die Aussage sofort mit Hilfe des Satzes von Stokes. □

Es sei an dieser Stelle angemerkt, dass die beiden letzten Sätze analog im $\mathbb{R}^2$ gelten.

Definition 11.6.5. *Sei $E \subseteq \mathbb{R}^n$ eine offene Menge, dann heißt E* **sternförmig**, *falls ein $\mathbf{x}_0 \in E$ existiert, so dass für alle $\mathbf{x} \in E$ auch die Verbindungstrecke von $\mathbf{x}$ und $\mathbf{x}_0$ ganz in E liegt, mit anderen Worten*

$$\{\mathbf{y} = t\mathbf{x} + (1-t)\mathbf{x}_0 : t \in [0,1]\} \subseteq E.$$

Den Punkt $\mathbf{x}_0 \in E$ nennen wir **Sternpunkt**.

Beispiel 11.6.6. *Sei E konvex, dann ist E auch sternförmig.*

Lemma 11.6.7. *Sei $E \subseteq \mathbb{R}^n$ eine offene und sternförmige Menge. Für $f \in C^1(E, \mathbb{R})$ gelte mit einem festen $p \in \{1, \ldots, n\}$, dass*

$$\frac{\partial f}{\partial x_j}(\mathbf{x}) = 0, \qquad p < j \leq n, \tag{11.13}$$

für alle $\mathbf{x} \in E$. Dann existiert eine Funktion $g \in C^1(E, \mathbb{R})$ mit

$$f(\mathbf{x}) = \frac{\partial g}{\partial x_p}(\mathbf{x}).$$

Beweis. Wir schreiben $\mathbf{x} = (\mathbf{x}', x_p, \mathbf{x}'') \in E$ mit $\mathbf{x}' = (x_1, \ldots, x_{p-1})$ und $\mathbf{x}'' = (x_{p+1}, \ldots, x_n)$. Den Sternpunkt von E bezeichnen wir mit $\mathbf{x}_0 = (\mathbf{x}_0', x_{p,0}, \mathbf{x}_0'')$.

Betrachten wir die Mengen

$$\begin{aligned} U &:= \{(\mathbf{x}', x_p) \in \mathbb{R}^p : \text{es existiert } \mathbf{x}'' \in \mathbb{R}^{n-p} \text{ mit } \mathbf{x} = (\mathbf{x}', x_p, \mathbf{x}'') \in E\}, \\ V &:= \{\mathbf{x}'' \in \mathbb{R}^{n-p} : \text{es existiert } (\mathbf{x}', x_p) \in \mathbb{R}^p \text{ mit } \mathbf{x} = (\mathbf{x}', x_p, \mathbf{x}'') \in E\}, \end{aligned}$$

so stellen wir fest, dass beide sternförmig sind mit den Sternpunkten $(\mathbf{x}_0', x_{p,0}) \in U$ bzw. $\mathbf{x}_0'' \in V$. Denn sei beispielsweise $\mathbf{x}'' \in V$ gegeben, dann existiert $(\mathbf{x}', x_p) \in \mathbb{R}^p$ mit $\mathbf{x} = (\mathbf{x}', x_p, \mathbf{x}'') \in E$ und damit liegt auch die Verbindungsgerade von $\mathbf{x}$ und $\mathbf{x}_0$ in E, das heißt

$$t\mathbf{x}_0 + (1-t)\mathbf{x} \in E, \qquad t \in [0,1],$$

und folglich auch $t\mathbf{x}_0'' + (1-t)\mathbf{x}'' \in U$ für alle $t \in [0,1]$.

Da V sternförmig ist, folgt aus dem Mittelwertsatz aufgrund der Voraussetzung (11.13), dass die Funktion f nicht von $\mathbf{x}''$ abhängt. Dies bedeutet, es existiert eine Funktion $\varphi \in C^1(U, \mathbb{R})$ mit

$$f(\mathbf{x}) = \varphi(\mathbf{x}', x_p)$$

für alle $\mathbf{x} = (\mathbf{x}', x_p, \mathbf{x}'') \in E$. Definieren wir nun $g \in C^1(E, \mathbb{R})$ vermittels

$$g(\mathbf{x}', x_p, \mathbf{x}'') := \int_{x_{p,0}}^{x_p} \varphi(\mathbf{x}', t)dt,$$

dann sehen wir, dass $\frac{\partial g}{\partial x_p}(\mathbf{x}', x_p, \mathbf{x}'') = \varphi(\mathbf{x}', x_p)$. $\square$

Definition 11.6.8. *Sei ω eine k-Form in einer offenen Menge $E \subseteq \mathbb{R}^n$. Falls ω in der Klasse C^1 liegt und $d\omega = 0$ gilt, dann heißt ω* **geschlossene** *k-Form. Die k-Form ω heißt* **exakt**, *falls eine $(k-1)$-Form λ aus der Klasse C^1 in E existiert, so dass $\omega = d\lambda$ gilt.*

Satz 11.6.9. *Sei ω eine k-Form in einer offenen Menge $E \subseteq \mathbb{R}^n$. Falls die k-Form ω exakt ist, so ist ω auch geschlossen.*

Beweis. Sei $\omega = d\lambda$ eine exakte k-Form, so folgt aus der zweiten Aussage von Satz 11.3.11

$$d\omega = d(d\lambda) = d^2\lambda = 0.$$

□

Satz 11.6.10 (Lemma von Poincaré). *Sei $E \subseteq \mathbb{R}^n$ eine offene und konvexe Menge und sei ω eine geschlossene k-Form der Klasse C^1 in E. Dann ist ω exakt, das heißt, es existiert eine $(k-1)$-Form λ in E der Klasse C^1 mit $\omega = d\lambda$.*

Beweis. Für festes $p \in \{1, \ldots, n\}$ bezeichne M_p die Menge aller k-Formen der Klasse C^1 in E, in deren Normaldarstellung

$$\omega = \sum_I f_I(\mathbf{x}) d\mathbf{x}_I \tag{11.14}$$

die Größen $dx_{p+1}, \ldots, dx_n$ nicht auftreten. Mit anderen Worten, die wachsendenden k-Indizes I sind Tupel von Zahlen aus der Menge $\{1, \ldots, p\}$.

Wir führen den Beweis durch Induktion nach p. Sei zunächst $\omega \in Y_1$, das heißt $\omega = f(\mathbf{x})dx_1$. Nach Voraussetzung gilt

$$d\omega = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(\mathbf{x})\ dx_j \wedge dx_1 = 0$$

und folglich $\frac{\partial f}{\partial x_j}(\mathbf{x}) = 0$ für alle $\mathbf{x} \in E$ und $1 < j \leq n$. Nach Lemma 11.6.7 existiert ein $g \in C^1(E, \mathbb{R})$, so dass für alle $\mathbf{x} \in E$ gilt $\frac{\partial g}{\partial x_1}(\mathbf{x}) = f(\mathbf{x})$ und $\frac{\partial g}{\partial x_j}(\mathbf{x}) = 0$ für $1 < j \leq n$. Somit erüllt die Nullform $\lambda = g$ die Gleichung

$$d\lambda = \frac{\partial g}{\partial x_j}(\mathbf{x})dx_1 = f(\mathbf{x})dx_1 = \omega.$$

Sei nun $p > 1$ und die Induktionsvoraussetzung laute: Jede geschlossene k-Form aus M_{p-1} ist exakt in E.

Wir wählen nun ein $\omega \in M_p$ mit $d\omega = 0$, dann gilt gemäß (11.14)

$$d\omega = \sum_I \sum_{j=1}^{n} \frac{\partial f_I}{\partial x_j}(\mathbf{x})\ dx_j \wedge d\mathbf{x}_I$$

Wir betrachten nun ein festgewähltes $j \in \{p, \ldots, n\}$. Da jeder wachsende k-Index I in der Summe in (11.14) genau einmal vorkommt und dieser nur Zahlen aus der Menge $\{1, \ldots, p\}$ enthält, gilt $dx_j \wedge d\mathbf{x}_I \neq 0$ und wir schließen

$$\frac{\partial f_I}{\partial x_j}(\mathbf{x}) = 0, \qquad \mathbf{x} \in E, \qquad p < j \leq n. \tag{11.15}$$

Wir fassen nun diejenigen Glieder aus (11.14) zusammen, die dx_p enthalten zusammen und schreiben ω um in der Form

$$\omega = \alpha + \sum_J f_{(J,p)}\, d\mathbf{x}_J \wedge dx_p,$$

wobei $\alpha \in M_{p-1}$ gilt und jeder wachsende $(k-1)$-Index J ein Tupel von Zahlen aus der Menge $\{1, \ldots, p-1\}$ ist. Nach Lemma 11.6.7 und (11.15) existieren zu $f_{(J,p)}$ Funktionen $g_{(J,p)}$ mit

$$f_{(J,p)}(\mathbf{x}) = \frac{\partial g_{(J,p)}}{\partial x_p}(\mathbf{x}), \qquad \mathbf{x} \in E,$$

und

$$\frac{\partial g_{(J,p)}}{\partial x_j}(\mathbf{x}) = 0, \qquad \mathbf{x} \in E, \qquad p < j \leq n.$$

Definieren wir nun die $(k-1)$-Form

$$\beta := \sum_J g_{(J,p)} d\mathbf{x}_J$$

und setzen $\gamma := \omega - (-1)^{k-1} d\beta$, dann gilt

$$\begin{aligned}
\gamma &= \omega - (-1)^{k-1} \sum_J \sum_{j=1}^{n} \frac{\partial g_{(J,p)}}{\partial x_j}\, dx_j \wedge d\mathbf{x}_J \\
&= \omega - \sum_J \sum_{j=1}^{p} \frac{\partial g_{(J,p)}}{\partial x_j}\, d\mathbf{x}_J \wedge dx_j \\
&= \alpha + \sum_J f_{(J,p)}\, d\mathbf{x}_J \wedge dx_p - \sum_J \sum_{j=1}^{p} \frac{\partial g_{(J,p)}}{\partial x_j}\, d\mathbf{x}_J \wedge dx_j \\
&= \alpha - \sum_J \sum_{j=1}^{p-1} \frac{\partial g_{(J,p)}}{\partial x_j} d\mathbf{x}_J \wedge dx_j.
\end{aligned}$$

Dies bedeutet $\gamma \in M_{p-1}$ und wegen $d\omega = 0$ und $d^2\beta = 0$ ist $d\gamma = 0$. Also ist γ eine geschlossene k-Form in M_{p-1} und nach der Induktionsvorschrift exakt, das heißt, es existiert eine $(k-1)$-Form λ mit $\gamma = d\lambda$. Eingesetzt in die Definition von γ liefert dies

$$\omega = d\lambda + (-1)^{k-1} d\beta = d\big(\lambda + (-1)^{k-1}\beta\big)$$

womit wir gezeigt haben, dass $\omega \in M_p$ exakt ist. $\square$

Beispiel 11.6.11. *Sei $E = \mathbb{R}^2 \setminus \{\mathbf{0}\}$ und $\omega := (xdy - ydx)/(x^2 + y^2)$ eine Einsform, dann gilt $d\omega = 0$, das heißt ω ist geschlossen.*

Für ein festes $r > 0$ sei $\boldsymbol{\gamma} : [0, 2\pi] \to \mathbb{R}^2$ definiert durch mit $\boldsymbol{\gamma}(t) = [r\cos t, r\sin t]^T$. Dann ist $\boldsymbol{\gamma}([0, 2\pi])$ eine geschlossene Kurve und es gilt

$$\int_{\boldsymbol{\gamma}} \omega = \int_0^{2\pi} \frac{r\cos t \cdot r\cos t + r(-\sin t) \cdot r(-\sin t)}{r^2} dt = \int_0^{2\pi} 1dt = 2\pi \neq 0.$$

Proposition 11.6.12. *Sei die Menge $E \subseteq \mathbb{R}^3$ offen und konvex, und sei $\mathbf{f} \in C^1(E, \mathbb{R}^3$. Gilt*

$$\operatorname{rot} \mathbf{f}(\mathbf{x}) = 0$$

für alle $\mathbf{x} \in E$, dann existiert eine Funktion $g \in C^2(E, \mathbb{R})$ mit

$$\mathbf{f}(\mathbf{x}) = \nabla g(\mathbf{x}).$$

Gilt

$$\operatorname{div} \mathbf{f}(\mathbf{x}) = 0$$

für alle $\mathbf{x} \in E$, dann existiert eine Funktion $\mathbf{h} \in C^2(E, \mathbb{R})$ mit

$$\mathbf{f}(\mathbf{x}) = \operatorname{rot} \mathbf{h}(\mathbf{x}).$$

Beweis. Schreiben wir $\mathbf{f} = [f_1, f_2, f_3]^T$, dann folgt der Beweis unmittelbar aus dem Lemma von Poincaré mit

$$\omega_{\mathbf{f}} = f_1(\mathbf{x})\, dy \wedge dz + f_2(\mathbf{x})\, dz \wedge dx + f_3(\mathbf{x})\, dx \wedge dy$$

beziehungsweise

$$\lambda_{\mathbf{f}} := f_1(\mathbf{x})dx + f_2(\mathbf{x})dy + f_3(\mathbf{x})dz.$$

□

Korollar 11.6.13. *Sei $E \subseteq \mathbb{R}^n$ eine offene Menge in der jede geschlossene k-Form auch exakt ist. Sei $U \subseteq \mathbb{R}^n$ ebenfalls offen und $\mathbf{t} \in C^2(E, U)$ bijektiv mit $\mathbf{t}^{-1} \in C^2(U, E)$. Dann ist jede geschlossene k-Form ω in U ebenfalls exakt.*

Beweis. Sei ω eine geschlossene k-Form in U, das heißt es gilt $d\omega = 0$, und $\omega_{\mathbf{t}}$ die zugehörige Pullback-k-Form in E. Dann gilt

$$d\omega_{\mathbf{t}} = (d\omega)_{\mathbf{t}} = 0,$$

das heißt, $\omega_{\mathbf{t}}$ ist geschlossen und es existiert ein λ mit $\omega_{\mathbf{t}} = d\lambda$. Daher gilt

$$\omega = \omega_{\mathbf{t} \circ \mathbf{t}^{-1}} = (\omega_{\mathbf{t}})_{\mathbf{t}^{-1}} = (d\lambda)_{\mathbf{t}^{-1}} = d\lambda_{\mathbf{t}^{-1}}$$

mit dem Pullback $\lambda_{\mathbf{t}^{-1}}$, der eine $(k-1)$-Form in U definiert. Folglich ist ω ist exakt. □

Kapitel 12

Maßtheorie

12.1 Ringe und Algebren von Mengen

Definition 12.1.1. *Sei $X \neq 0$ eine beliebige Menge und $\mathcal{M}$ ein nichtleeres System von Teilmengen der Menge X, das heißt $\mathcal{M} \subseteq \mathcal{P}(X)$.*

1. *Das Mengensystem $\mathcal{M}$ heißt* **Semiring** *oder auch* **zerlegbares Mengensystem** *in X, wenn gilt*
   (a) $\emptyset \in \mathcal{M}$,
   (b) *aus $A, B \in \mathcal{M}$ folgt $A \cap B \in \mathcal{M}$, und*
   (c) *zu jedem Paar $A, B \in \mathcal{M}$ existieren paarweise disjunkte Mengen $A_1, \ldots, A_m \in \mathcal{M}$ mit $A \setminus B = \bigcup_{i=1}^m A_i$.*
2. *Das Mengensystem $\mathcal{M}$ heißt* **Ring** *in X, wenn für alle $A, B \in \mathcal{M}$ gilt*
   (a) $A \cup B \in \mathcal{M}$ *und*
   (b) $A \setminus B \in \mathcal{M}$.
3. *Das Mengensystem $\mathcal{M}$ heißt* **Algebra** *in X, wenn gilt*
   (a) $A \cup B \in \mathcal{M}$ *für alle $A, B \in \mathcal{M}$ und*
   (b) *aus $A \in \mathcal{M}$ folgt $X \setminus A \in \mathcal{M}$.*
4. *Ein Ring bzw. eine Algebra $\mathcal{M}$ heißt σ-***Ring** *bzw. σ-***Algebra***, wenn aus $A_n \in \mathcal{M}$, $n \in \mathbb{N}$, folgt*
$$\bigcup_{n=1}^{\infty} A_n \in \mathcal{M}. \tag{12.1}$$
5. *Das Mengensystem $\mathcal{M}$ heißt* **monoton***, wenn gilt*

(a) aus $A_1 \subseteq A_2 \subseteq A_3 \subseteq \ldots$ *mit* $A_n \in M$, $n \in \mathbb{N}$, *folgt*

$$\bigcup_{n=1}^{\infty} A_n \in \mathcal{M}$$

und

(b) aus $A_1 \supseteq A_2 \supseteq A_3 \supseteq \ldots$ *mit* $A_n \in \mathcal{M}$, $n \in \mathbb{N}$, *folgt*

$$\bigcap_{n=1}^{\infty} A_n \in \mathcal{M}.$$

Beispiel 12.1.2. *1. Es sei* $X = \mathbb{R}$, *dann ist*

$$\mathcal{M} = \{(a, b] : a < b \text{ beliebig reell}\} \cup \{\emptyset\}$$

ein Semiring.

2. Sei X *eine (unendliche) Menge und*

$$\mathcal{M} = \{A \subseteq X : A \text{ oder } X \setminus A \text{ endlich}\},$$

so ist M *eine Algebra.*

3. Sei X *eine (überabzählbare) Menge, dann ist*

$$\mathcal{M} = \{A \subseteq X : A \text{ oder } X \setminus A \text{ höchstens abzählbar}\}$$

ist eine σ*-Algebra.*

Dabei zeigt man die Beziehung (12.1) *wie folgt. Seien* $A_n \in \mathcal{M}$, $n \in \mathbb{N}$, *beliebige Mengen aus* $\mathcal{M}$. *Sind alle* A_n *höchstens abzählbar, so folgt, dass auch* $\bigcup_{n=1}^{\infty} A_n$ *höchstens abzählbar ist. Sei nun etwa* A_1 *überabzählbar, so ist* $X \setminus A_1$ *höchstens abzählbar. Aus*

$$A := X \setminus \bigcup_{n=1}^{\infty} A_n = \bigcap_{n=1}^{\infty} (X \setminus A_n) \subseteq X \setminus A_1.$$

ergibt sich, dass A *höchstens abzählbar ist und weiterhin* $X \setminus A = \bigcup_{n=1}^{\infty} A_n \in \mathcal{M}$.

Satz 12.1.3. *Es sei* T *eine Indexmenge, dann ist der Durchschnitt beliebig vieler Ringe* $\mathcal{R}_t$, $t \in T$, *(*σ*-Ringe, Algebren,* σ*-Algebren, monotone Systeme) in* X *wieder ein solches System.*

Beweis. Das Mengensystem $\mathcal{R}_t$ bezeichne für jedes $t \in T$ einen Ring. Für $A, B \in \mathcal{R}$ gilt auch $A, B \in \mathcal{R}_t$ für alle $t \in T$. Daher ist aber auch $A \cup B, A \setminus B \in \mathcal{R}_t$ für alle $t \in T$ und folglich $A \cup B, A \setminus B \in \mathcal{R}$.

Betreff den anderen algebraischen Strukturen führt man den Beweis analog. □

Definition 12.1.4. *Seien X eine beliebige nichtleere Menge und $\mathcal{M} \subseteq \mathcal{P}(X)$ ein beliebiges nichtleeres Mengensystem. Die Durchschnittsmenge aller Ringe (σ-Ringe, Algebren, σ-Algebren, monotonen Systeme), die $\mathcal{M}$ umfassen, heißt der (die, das) von $\mathcal{M}$* **erzeugte Ring** *$\mathcal{R}(\mathcal{M})$ (σ-Ring $\mathcal{R}_\sigma(\mathcal{M})$, Algebra $\mathcal{A}(\mathcal{M})$, σ-Algebra $\mathcal{A}_\sigma(\mathcal{M})$, monotone System $m(\mathcal{M})$).*

Bemerkung 12.1.5.

1. *Da die Potenzmenge $\mathcal{P}(X)$ jedes Mengensystem $\mathcal{M}$ von Teilmengen X umfasst und $\mathcal{P}(X)$ ein σ-Ring, eine Algebra, eine σ-Algebra und ein monotones System ist, ist die obige Definition korrekt.*

2. *Der von $\mathcal{M}$ erzeute Ring $\mathcal{R}(\mathcal{M})$ ist bezüglich der Inklusion der kleinste Ring, der das Mengensystem M umfasst.*

3. *Die zu $\mathcal{M}$ erzeute σ-Algebra $\mathcal{A}_\sigma(\mathcal{M})$ wird auch* Borelsche Erweiterung *$\sigma(\mathcal{M})$ von $\mathcal{M}$ genannt.*

Lemma 12.1.6. *Seien $\mathcal{M}_1$ und $\mathcal{M}_2$ nichtleere Mengensysteme von Teilmengen von X mit*

$$\mathcal{M}_1 \subseteq \mathcal{R}(\mathcal{M}_2), \qquad \mathcal{M}_2 \subseteq \mathcal{R}(\mathcal{M}_1).$$

Dann gilt

$$\mathcal{R}(\mathcal{M}_2) = \mathcal{R}(\mathcal{M}_2)$$

Beweis. Aus $\mathcal{M}_1 \subseteq \mathcal{R}(\mathcal{M}_2)$ folgt sofort $\mathcal{R}(\mathcal{M}_1) \subseteq \mathcal{R}(\mathcal{M}_2)$. Anolog findet man $\mathcal{R}(\mathcal{M}_2) \subseteq \mathcal{R}(\mathcal{M}_1)$. □

Satz 12.1.7. *Der von einem Semiring $\mathcal{S}$ erzeugte Ring $\mathcal{R}(\mathcal{S})$ besteht aus dem System $\mathcal{V}$ aller endlichen Vereinigungen paarweise disjunkter Mengen aus $\mathcal{S}$.*

Beweis. Für $i = 1, \ldots, m$ gelte $A_i \in \mathcal{S}$ mit $A_i \cap A_j = \emptyset$ für $i \neq j$. Weil $\mathcal{S} \subseteq \mathcal{R}(\mathcal{S})$, ist $A = \bigcup_{i=1}^m A_i \in \mathcal{R}(\mathcal{S})$, das heißt $\mathcal{V} \subseteq \mathcal{R}(\mathcal{S})$.

Andererseits folgt mit Hilfe von Lemma 12.1.6 aus $\mathcal{S} \subseteq \mathcal{V} \subseteq \mathcal{R}(\mathcal{S})$ dass $\mathcal{R}(\mathcal{S}) = \mathcal{R}(\mathcal{V})$. Wir zeigen, dass $\mathcal{V}$ ein Ring ist, so dass gilt $\mathcal{V} = \mathcal{R}(\mathcal{V}) = \mathcal{R}(\mathcal{S})$. Für $k = 1, \ldots, n$ seien $B_k \in \mathcal{S}$ mit $B_k \cap B_l = \emptyset$ für $k \neq l$ und $B = \bigcup_{k=1}^n B_k$. Es gilt

$$A \cap B = \left(\bigcup_{i=1}^m A_i\right) \cap \left(\bigcup_{k=1}^n B_k\right) = \bigcup_{i=1}^m \bigcup_{k=1}^n (A_i \cap B_k).$$

Da die Mengen $A_i \cap B_k \in \mathcal{S}$ paarweise disjunkt sind, folgt $A \cap B \in \mathcal{V}$. Wir zeigen nun, dass auch $A \setminus B \in \mathcal{V}$. Für festes k sind die Mengen $A_i \setminus B_k$ paarweise disjunkt und lassen sich als eine endliche Vereinigung von Mengen aus $\mathcal{S}$ schreiben. Somit folgt

$$A \setminus B_k = \bigcup_{i=1}^m (A_i \setminus B_k) \in \mathcal{V}$$

und weiter

$$A \setminus B = A \setminus \bigcup_{k=1}^{n} B_k = \bigcap_{k=1}^{n} (A \setminus B_k) \in \mathcal{V}$$

nach dem bereits Bewiesenen, das heißt $A \setminus B \in \mathcal{V}$.

Zu zeigen bleibt schließlich noch, dass $A \cup B \in \mathcal{V}$. Dies folgt aber sofort aus der Gleichung

$$A \cup B = (A \setminus B) \cup B,$$

wenn wir bedenken, dass die Mengen $A \setminus B$ und B paarweise disjunkt sind. □

Satz 12.1.8. *Ist $\mathcal{A}$ eine Algebra, so gilt*

$$m(\mathcal{A}) = \sigma(\mathcal{A}).$$

Beweis. Der Beweis verbleibt dem Leser als Übung. □

12.2 Mengenfunktionen

Es bezeichne $\overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\} \cup \{-\infty\}$ das erweiterte System aller Zahlen, wobei wir folgendes festlegen.

1. Für $x \in \mathbb{R}$, sei $x \pm \infty = \pm\infty$ und $\frac{x}{\pm\infty} = 0$.
2. Für $x \in \overline{\mathbb{R}} \setminus \{0\}$ sei $x \cdot (\pm\infty) = \pm \operatorname{sgn}(x) \cdot \infty$ und $0 \cdot (\pm\infty) = 0$.
3. Wir definieren $|+\infty| = |-\infty| = +\infty$.
4. Es sei $\pm\infty + (\pm\infty) = \pm\infty$ und $\pm\infty - (\mp\infty) = \pm\infty$.
5. Es sei $(\pm\infty) \cdot (\pm\infty) = +\infty$ und $(\pm\infty) \cdot (\mp\infty) = -\infty$.

Definition 12.2.1. *Sei $\mathcal{M}$ ein nichtleeres Mengensystem von Teilmengen der Menge $X \neq \emptyset$.*

1. *Eine eindeutige Abbildung $\varphi : \mathcal{M} \to \overline{\mathbb{R}}$ heißt eine auf $\mathcal{M}$ definierte* **Mengenfunktion**.
2. *Die Mengenfunktion $\varphi : \mathcal{M} \to \overline{\mathbb{R}}$ heißt* **additiv** *auf $\mathcal{M}$, wenn zum Wertevorat von φ höchstens eines der uneigentlichen Elemente $\pm\infty$ gehört und wenn aus $A, B, A \cup B \in \mathcal{M}$ mit $A \cap B = \emptyset$ stets folgt*
$$\varphi(A \cup B) = \varphi(A) + \varphi(B). \tag{12.2}$$
3. *Die additive Mengenfunktion $\varphi : \mathcal{M} \to \overline{\mathbb{R}}$ heißt* **volladditiv**, *wenn aus $A_n \in \mathcal{M}$, $n \in \mathbb{N}$, mit $A_i \cap A_j = \emptyset$ für $i \neq j$ und $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{M}$ folgt*
$$\varphi\Big(\bigcup_{n\in\mathbb{N}} A_n\Big) = \sum_{n\in\mathbb{N}} \varphi(A_n). \tag{12.3}$$

4. *Die additive Mengenfunktion* $\varphi : \mathcal{M} \to \overline{\mathbb{R}}$ *heißt* **subvolladditiv**, *wenn aus* $A, A_n \in \mathcal{M}$, $n \in \mathbb{N}$, *mit* $A \subseteq \bigcup_{n\in\mathbb{N}} A_n$ *folgt*

$$\varphi(A) \leq \sum_{n\in\mathbb{N}} \varphi(A_n).$$

Bemerkung 12.2.2. 1. *Wenn die Reihe in* (12.3) *in* $\mathbb{R}$ *konvergiert, dann ist sie auch absolut konvergent. Denn die linke Seite darf beliebig umgeordnet werden ohne die zugrundeliegende Menge zu verändern. Daher darf auch die rechte beliebig umgeordnet werden ohne ihren Wert zu verändern, was nach dem Riemannschen Umordnungssatz die absolute Konvergenz bedeutet.*

2. *Wir betrachten im folgenden nur Mengenfunktionen* $\not\equiv \pm\infty$.

Beispiel 12.2.3. 1. *Sei* X *eine unendliche Menge und* $\mathcal{M} = \mathcal{P}(X)$, *so ist die Mengenfunktion*

$$\varphi(A) = \begin{cases} 0, & \text{wenn } A \text{ endlich}, \\ +\infty & \text{wenn } A \text{ unendlich}, \end{cases}$$

additiv, aber nicht volladditiv oder subvolladditiv.

2. *Sei* $X = \mathbb{R}$ *und* $\mathcal{M} = \{(n, n+1] : n \in \mathbb{N}_0\}$. *Die Mengenfunktion*

$$\varphi\big((n, n+1]\big) = \begin{cases} 0, & \text{wenn } n \text{ gerade}, \\ 1, & \text{wenn } n \text{ ungerade}, \end{cases}$$

ist trivialerweise additiv, da die Vereinigung aus $\mathcal{M}$ *herausführt.*

3. *Sei* $X = \mathbb{R}^n$ *und* $\mathcal{M} = \mathcal{P}(X)$. *Die höchstens abzählbar Menge* $\widetilde{X} = \{x_1, x_2, \ldots\} \subseteq \mathbb{R}^n$ *besitze keinen Häufungspunkt in* $\mathbb{R}^n$. *Ferner sei* $f : \widetilde{X} \to \mathbb{R}_+$ *eine beliebige Funktion. Dann ist die Mengenfunktion*

$$\varphi(A) := \sum_{x_i \in A} f(x_i)$$

volladditiv.

4. *Sei* $X = \mathbb{R}$ *und*

$$\mathcal{M} = \{(a, b] : a < b \text{ beliebig reell}\} \cup \{\emptyset\}$$

Weiter sei $h : \mathbb{R} \to \mathbb{R}$ *monoton wachsend. Setzen wir*

$$\varphi\big((a, b]\big) = h(b+) - h(a+), \qquad \varphi(\emptyset) = 0,$$

dann ist φ *volladditiv.*

Beispielsweise gilt $\varphi\big((a, b]\big) = b - a$ *für* $h(x) = x$ *während*

$$\varphi(x) = \begin{cases} 1, & \text{wenn } 0 \in (a, b], \\ 0, & \text{sonst}, \end{cases}$$

falls

$$h(x) = \begin{cases} 1, & x \geq 0, \\ 0, & x < 0. \end{cases}$$

5. *Sei X überabzählbar und $\mathcal{M} = \{A \subseteq X : A$ oder $X \setminus A$ höchsten abzählbar$\}$, so ist*

$$\varphi(A) = \begin{cases} 0, & \text{falls } A \text{ höchstens abzählbar}, \\ 1, & \text{falls } X \setminus A \text{ höchstens abzählbar}, \end{cases}$$

vollادditiv. Denn seien $A_n \in \mathcal{M}$, $n \in \mathbb{N}$, mit $A_i \cup A_j = \emptyset$ für $i \neq j$. Da $\mathcal{M}$ eine σ-Algebra ist, folgt

$$A = \bigcup_{n=1}^{\infty} A_n \in \mathcal{M}.$$

Es ergeben sich zwei Fälle. Erstens, dass alle A_n höchstens abzählbar sind. Dann folgt aber, dass A auch nur höchstens abzählbar ist und somit gilt $\varphi(A) = 0 = \sum_{n\in\mathbb{N}} \varphi(A_n)$. Zweitens, dass ohne Beschränkung der Allgemeinheit A_1 überabzählbar ist. Folglich ist A überabzählbar und daher $X \setminus A$ abzählbar. Da für $n \geq 2$ gilt $A_n \subseteq X \setminus A_1$, sind für $n \geq 2$ alle Mengen A_n abzählbar. Daraus ergibt sich

$$\varphi(A) = 1 = \varphi(A_1) + \sum_{n=2}^{\infty} \varphi(A_n).$$

Satz 12.2.4 (Eigenschaften von Mengenfunktionen). *Sei $\mathcal{R}$ ein Mengenring und $\varphi : \mathcal{R} \to \overline{\mathbb{R}}$ additiv.*

1. *Für $i = 1, \ldots, n$ gelte $A_i \in \mathcal{R}$ mit $A_i \cap A_j = \emptyset$ für $i \neq j$. Dann folgt*

$$\varphi\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \varphi(A_i).$$

2. *Für zwei Mengen $A, B \in \mathcal{R}$ mit $B \subseteq A$ und $|\varphi(B)| < \infty$ gilt*

$$\varphi(A \setminus B) = \varphi(A) - \varphi(B).$$

3. *Wenn ein $A \in \mathcal{R}$ mit $|\varphi(A)| < \infty$ existiert, so gilt $\varphi(\emptyset) = 0$.*

4. *Für $A, B \in \mathcal{R}$ gilt*

$$\varphi(A \cup B) + \varphi(A \cap B) = \varphi(A) + \varphi(B)$$

5. *Seien $A, B \in \mathcal{R}$ mit $B \subseteq A$ und $\varphi(C) \geq 0$ für alle $C \in \mathcal{R}$, dann folgt*

$$\varphi(B) \leq \varphi(A).$$

6. *Seien $A, B \in \mathcal{R}$ beliebig und $\varphi(C) \geq 0$ für alle $C \in \mathcal{R}$, dann gilt*

$$\varphi(A \cup B) \leq \varphi(A) + \varphi(B).$$

7. *Gilt $\varphi(C) \geq 0$ für alle $C \in \mathcal{R}$, so ist φ genau dann volladditiv, wenn φ subvolladditiv ist.*

Beweis. Der Beweis verbleibt dem Leser als Übung. □

Satz 12.2.5. *Sei $\varphi : \mathcal{R} \to \mathbb{R}$ eine additive Mengenfunktion auf dem Ring $\mathcal{R}$. Dann ist φ genau dann volladditiv, falls aus $A_n \in \mathcal{R}$, $n \in \mathbb{N}$, mit*

$$A_1 \subseteq A_2 \subseteq \cdots \subseteq A_n \subseteq A_{n+1} \subseteq \cdots$$

und $A := \bigcup_{n=1}^{\infty} A_n \in \mathcal{R}$ folgt

$$\lim_{n\to\infty} \varphi(A_n) = \varphi(A).$$

Beweis. "$\Rightarrow$" Sei φ volladditiv. Definieren wir

$$B_i = \begin{cases} A_1, & i = 1, \\ A_i \setminus A_{i-1}, & i > 1, \end{cases}$$

so sind die Mengen B_i paarweise disjunkt. Ferner gilt

$$A_n = \bigcup_{i=1}^{n} B_i.$$

für alle $n \in \mathbb{N}$ und

$$A = \bigcup_{i=1}^{\infty} B_i.$$

Somit folgt

$$\varphi(A) = \varphi\left(\bigcup_{i\in\mathbb{N}} B_i\right) = \sum_{i=1}^{\infty} \varphi(B_i) = \lim_{n\to\infty} \sum_{i=1}^{n} \varphi(B_i) = \lim_{n\to\infty} \varphi(A_n).$$

"$\Leftarrow$": Es gelte

$$\lim_{n\to\infty} \varphi(A_n) = \varphi(A).$$

Für beliebige paarweise disjunkte Mengen $B_i \in \mathcal{R}$, $i \in \mathbb{N}$, definieren wir

$$A_n = \bigcup_{i=1}^{n} B_i, \quad n \in \mathbb{N},$$

und erhalten

$$A_1 \subseteq A_2 \subseteq \cdots \subseteq A_n \subseteq A_{n+1} \subseteq \cdots$$

und

$$A = \bigcup_{i=1}^{\infty} A_i = \bigcup_{i=1}^{\infty} B_i \in \mathcal{R}.$$

Damit folgt

$$\varphi(A) = \varphi\left(\bigcup_{i=1}^{\infty} A_i\right) = \lim_{n\to\infty} \varphi(A_n) = \lim_{n\to\infty} \varphi\left(\bigcup_{i=1}^{n} B_i\right) = \lim_{n\to\infty} \sum_{i=1}^{n} \varphi(B_i) = \sum_{i=1}^{\infty} \varphi(B_i),$$

dies bedeutet, φ ist volladditiv. □

Satz 12.2.6. *Sei φ eine auf dem Ring $\mathcal{R}$ volladditive Mengenfunktion. Gilt für die Mengen $B_i \in \mathcal{R}$, $i \in \mathbb{N}$, mit*

$$B_1 \supseteq B_2 \supseteq \ldots \supseteq B_i \supseteq B_{i+1} \supseteq \ldots,$$

dass $|\varphi(B_1)| < \infty$ und

$$B = \bigcap_{i=1}^{\infty} B_i \in \mathcal{R},$$

dann folgt

$$\varphi(B) = \lim_{n\to\infty} \varphi(B_n).$$

Beweis. Für jedes $i \in \mathbb{N}$ setzen wir $A_i := B_1 \setminus B_i$. Es gilt dann $A_i \in \mathcal{R}$ und

$$A_1 \subseteq A_2 \subseteq \ldots \subseteq A_n \subseteq A_{n+1} \subseteq \ldots.$$

Mit der Definition $A := B_1 \setminus B = B_1 \setminus \left(\bigcap_{i=1}^{\infty} B_i\right) \in \mathcal{R}$ folgt wegen $B \subseteq B_1$ einerseits

$$\varphi(A) = \varphi(B_1) - \varphi(B).$$

Andereseits gilt aber nach Satz 12.2.5 auch

$$\varphi(A) = \lim_{n\to\infty} \varphi(A_n) = \lim_{n\to\infty} \varphi(B_1 \setminus B_n) = \varphi(B_1) - \lim_{n\to\infty} \varphi(B_n),$$

das heißt $\lim_{n\to\infty} \varphi(B_n) = \varphi(B)$. □

Satz 12.2.7. *Sei φ eine endliche und additive Mengenfunktion auf dem Ring $\mathcal{R}$. Gilt für jede Folge von Mengen $B_i \in \mathcal{R}$, $i \in \mathbb{N}$, mit*

$$B_1 \supseteq B_2 \supseteq \ldots \supseteq B_i \supseteq B_{i+1} \supseteq \ldots$$

und

$$\bigcap_{n=1}^{\infty} B_n = \emptyset,$$

dass $\lim_{n\to\infty} \varphi(B_n) = 0$, so ist φ volladditiv.

Beweis. Seien $A_i \in \mathcal{R}$, $i \in \mathbb{N}$, paarweise disjunkte Mengen und $A := \bigcup_{i=1}^{\infty} A_i \in \mathcal{R}$. Definieren wir

$$B_n := A \setminus \bigcup_{i=1}^{n} A_i = \bigcup_{i=n+1}^{\infty} A_i,$$

für jedes $n \in \mathbb{N}$, dann gilt

$$B_1 \supseteq B_2 \supseteq \ldots \supseteq B_i \supseteq B_{i+1} \supseteq \ldots$$

und

$$\bigcap_{n=1}^{\infty} B_n = A \setminus \left(\bigcup_{i=1}^{n} A_n \right) = \emptyset.$$

Aufgrund der Voraussetzung schließen wir

$$\begin{aligned} 0 = \lim_{n\to\infty} \varphi(B_n) &= \lim_{n\to\infty} \left\{ \varphi(A) \setminus \left(\bigcup_{i=1}^{n} A_n \right) \right\} = \lim_{n\to\infty} \left\{ \varphi(A) - \sum_{i=1}^{n} A_i \right\} \\ &= \varphi(A) - \sum_{i=1}^{\infty} \varphi(A_i), \end{aligned}$$

das heißt die Behauptung. □

12.3 Inhalt, Maß und Konstuktion von Maßen

Definition 12.3.1. *Eine additive Mengenfunktion φ auf dem Ring $\mathcal{R}$ heißt* **Inhalt**, *falls gilt $\varphi(\emptyset) = 0$ und $\varphi(A) \geq 0$ für alle $A \in \mathcal{R}$. Ein volladditiver Inhalt φ, definiert auf einer σ-Algebra $\mathcal{A}$, heißt* **Maß**.

Bemerkung 12.3.2. *Für die Maßtheorie ist die algebraische Struktur σ-Algebra und die Definition eines Maßes auf dieser von zentraler Bedeutung.*

Definition 12.3.3. *Seien $\mathcal{M}_1$ und $\mathcal{M}_2$ zwei Mengensysteme in X mit $\mathcal{M}_1 \subseteq \mathcal{M}_2$ und seien $\varphi_i : \mathcal{M}_i \to \overline{\mathbb{R}}$, $i = 1, 2$, Mengenfunktionen. Die mengenfunktion φ_2 heißt* **Fortsetzung** *von φ_1 auf $\mathcal{M}_2$, falls $\varphi_2(A) = \varphi_1(A)$ für alle $A \in \mathcal{M}_1$. In diesem Falle nennen wir φ_1 auch die* **Einschränkung** *von φ_2 auf $\mathcal{M}_2$.*

Satz 12.3.4. *Sei $\mathcal{S}$ ein Semiring. Jede additive Mengenfunktion $\varphi : \mathcal{S} \to \overline{\mathbb{R}}$ lässt sich in einer eindeutigen Art und Weise fortsetzen zu einer additiven Mengenfunktion $\widetilde{\varphi}$ auf $\mathcal{R}(\mathcal{S})$. Ist φ subvolladditiv und nichtnegativ mit $\varphi(\emptyset) = 0$, so definiert $\widetilde{\varphi}$ sogar einen subvolladditiven Inhalt auf $\mathcal{R}(\mathcal{S})$.*

Beweis. Sei $A \in \mathcal{R}(\mathcal{S})$ beliebig. Nach Satz 12.1.7 existieren paarweise disjunkte Mengen $S_1, \ldots, S_n \in \mathcal{S}$ mit $A = \bigcup_{i=1}^n S_i$. Wir definieren nun die Mengenfunktion $\widetilde{\varphi}$ vermittels

$$\widetilde{\varphi}(A) := \sum_{i=1}^n \varphi(S_i).$$

Zuerst stellen wir fest, dass $\widetilde{\varphi}(A)$ nicht von der Zerlegung $A = \bigcup_{i=1}^n S_i$ abhängt. Denn seien $S_1', \ldots S_m' \in \mathcal{S}$ mit $A = \bigcup_{k=1}^n S_k'$ paarweise disjunkte Mengen, dann gilt

$$S_k = S_k \cap A = S_k \cap \left(\bigcup_{i=1}^m S_i' \right) = \bigcup_{i=1}^m (S_k \cap S_i'), \qquad 1 \leq k \leq n,$$

und folglich

$$\begin{aligned} \sum_{k=1}^n \varphi(S_k) &= \sum_{k=1}^n \varphi\left(\bigcup_{i=1}^m (S_k \cap S_i') \right) = \sum_{k=1}^n \sum_{i=1}^m \varphi(S_k \cap S_i') = \sum_{i=1}^m \varphi\left(\bigcup_{k=1}^n (S_k \cap S_i') \right) \\ &= \sum_{i=1}^m \varphi(A \cap S_i') = \sum_{i=1}^m \varphi(S_i') = \widetilde{\varphi}(A). \end{aligned}$$

Trivialerweise ist $\widetilde{\varphi}$ eine Fortsetzung von φ auf $\mathcal{R}(\mathcal{S})$.

Wir zeigen nun, dass $\widetilde{\varphi}$ ebenfalls additiv ist. Seien hierzu $A, B \in \mathcal{R}(\mathcal{S})$ mit $A \cap B = \emptyset$. Dann existieren paarweise disjunkte Mengen $S_1, \ldots, S_n \in \mathcal{S}$ und $R_1, \ldots, R_m \in \mathcal{S}$, so dass

$$A = \bigcup_{i=1}^n S_i, \qquad B = \bigcup_{j=1}^m R_j.$$

Wegen $A \cap B = \emptyset$ gilt auch $S_i \cap R_j = \emptyset$ für alle i und j und daher ergibt sich

$$A \cup B = \sum_{i=1}^n \varphi(S_i) + \sum_{j=1}^m \varphi(R_j) = \widetilde{\varphi}(A) + \widetilde{\varphi}(B),$$

das heißt $\widetilde{\varphi}$ ist additiv.

Zum Beweis der Eindeutigkeit von $\widetilde{\varphi}$ sei $\psi : \mathcal{R}(\mathcal{S}) \to \overline{\mathbb{R}}$ ebenfalls eine additive Fortsetzung von φ. Dann gilt aber für beliebige $A = \bigcup_{i=1}^n S_i \in \mathcal{R}(\mathcal{S})$ aufgrund der Addititivität

$$\psi(A) = \psi\left(\bigcup_{i=1}^n S_i \right) = \sum_{i=1}^n \varphi(S_i) = \widetilde{\varphi}(A).$$

Es verbleibt, den zweiten Teil des Satzes zu beweisen. Aus $\varphi(\emptyset) = 0$ folgt natürlich sofort $\widetilde{\varphi}(\emptyset) = 0$. Ferner impliziert $\varphi(A) \geq 0$ für alle $A \in \mathcal{S}$ auch $\widetilde{\varphi}(A) \geq 0$ für alle $A \in \mathcal{R}(\mathcal{S})$. Zum Beweis der Subvolladditivität von $\widetilde{\varphi}$ seien zunächst $A, A_n \in \mathcal{R}(\mathcal{S})$ ($n \in \mathbb{N}$) mit $A = \bigcup_{n=1}^\infty A_n$, das heißt

$A_n \subseteq A$ für alle $n \in \mathbb{N}$. Zu A bzw. A_n gibt es paarweise disjunkte Mengen $S_1, \ldots, S_m \in \mathcal{S}$ bzw. $S_1^{(n)}, \ldots, S_{m^{(n)}}^{(n)} \in \mathcal{S}$ mit

$$A = \bigcup_{i=1}^{m} S_i, \qquad A_n = \bigcup_{i=1}^{m^{(n)}} S_i^{(n)}.$$

Nun sind die Mengen $S_i \cap S_j^{(n)} \in \mathcal{S}$ für alle $1 \leq i \leq m$ und $1 \leq j \leq m^{(n)}$ ebenfalls paarweise disjunkt mit

$$A_n = A_n \cap A = \bigcup_{i=1}^{m} \bigcup_{j=1}^{m^{(n)}} (S_j^{(n)} \cap S_i).$$

Da für jedes $1 \leq i \leq m$ gilt

$$S_i = \bigcup_{n=1}^{\infty} \bigcup_{j=1}^{m^{(n)}} (S_j^{(n)} \cap S_i),$$

liefert die Subvolladditivität von φ auf $\mathcal{S}$ die Ungleichung

$$\varphi(S_i) \leq \sum_{n=1}^{\infty} \sum_{j=1}^{m^{(n)}} \varphi(S_j^{(n)} \cap S_i).$$

Folglich erhalten wir

$$\widetilde{\varphi}(A) = \sum_{i=1}^{m} \varphi(S_i) \leq \sum_{i=1}^{m} \sum_{n=1}^{\infty} \sum_{j=1}^{m^{(n)}} \varphi(S_j^{(n)} \cap S_i) = \sum_{n=1}^{\infty} \sum_{j=1}^{m^{(n)}} \sum_{i=1}^{m} \varphi(S_j^{(n)} \cap S_i) = \sum_{n=1}^{\infty} \widetilde{\varphi}(A_n).$$

Es gelte nun im Gegensatz zu oben nur $A \subseteq \bigcup_{n=1}^{\infty} A_n$. Mit dem eben Gezeigten führt uns dann die Zerlegung $B_n := A_n \cap A$ und $C_n := A_n \setminus A$ zum Ziel

$$\widetilde{\varphi}(A) \leq \sum_{n=1}^{\infty} \widetilde{\varphi}(B_n) \leq \sum_{n=1}^{\infty} \big(\widetilde{\varphi}(B_n) + \widetilde{\varphi}(C_n)\big) = \sum_{n=1}^{\infty} \widetilde{\varphi}(A_n).$$

Damit ist der Beweis des Satzes vollständig. □

Beispiel 12.3.5. *1. Für festes $n \in \mathbb{N}$ definiert die Menge*

$$\mathcal{S} = \left\{ I^n = \bigotimes_{k=1}^{n} (a_k, b_k] \right\}$$

einen Semiring. Die Mengenfunktion

$$\mu(I^n) := \prod_{k=1}^{n} (b_k - a_k)$$

ist ein Inhalt auf $\mathcal{R}(\mathcal{S})$.

2. *$\mathcal{M} = \mathcal{P}(\mathbb{R})$, dann ist φ gegeben durch*

$$\varphi = \begin{cases} 1, & 0 \in A \subseteq \mathbb{R}, \\ 0, & 0 \notin A \subseteq R, \end{cases}$$

ein Maß. Dieses Maß nennt man auch das Dirac-Maß konzentriert im Punkt 0 oder mit dem Trägerpunkt 0.

Definition 12.3.6. *Sei $\mu^\star : \mathcal{P}(X) \to \overline{\mathbb{R}}$ nichtnegativ, dann heißt $\mu^\star$* **äußeres Maß***, falls folgende Aussagen gelten:*

1. *Es gilt $\mu^\star(\emptyset) = 0$.*
2. *Die Mengenfunktion $\mu^\star$ ist monoton, das heißt aus $A \subseteq B \subseteq X$ folgt $\mu^\star(A) \leq \mu^\star(B)$.*
3. *Für $A_n \subseteq X$, $n \in \mathbb{N}$, erfüllt die Mengenfunktion $\mu^\star$ die Ungleichung*

$$\mu^\star\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mu^\star(A_n). \tag{12.4}$$

Beispiel 12.3.7. *Sei $X = [0,1]$, dann setzen wir für $A \subseteq X$*

$$\mu^\star(A) = \begin{cases} \sup A, & A \neq \emptyset, \\ 0, & A = \emptyset, \end{cases}.$$

Die Mengenfunktion $\mu^\star$ ist ein äußeres Maß, denn die Gültigkeit der ersten beiden Aussagen der Definition ist direkt einsichtig, während man die dritte Aussage wie folgt zeigt. Seien die Mengen $A_n \subseteq X$, $n \in \mathbb{N}$, beliebig und $A := \bigcup_{n\in\mathbb{N}} A_n$. Es gelte $s := \sum_{n=1}^{\infty} \sup A_n < \infty$, ansonsten sind wir wegen $\sup A \leq \sup X = 1$ fertig. Nun folgt aus der Definition des Supremums, dass zu jedem $\varepsilon > 0$ ein $x_\varepsilon \in A$ existiert mit

$$\sup A \leq x_\varepsilon + \varepsilon.$$

Folglich existiert ein n_ε mit $x_\varepsilon \in A_{n_\varepsilon}$ und daher gilt

$$\sup A \leq \sup A_{n_\varepsilon} + \varepsilon.$$

Für $\varepsilon \to 0$ ergibt sich die Behauptung

$$\mu^\star(A) = \sup A \leq \sum_{n\in\mathbb{N}} \sup A_n = \sum_{n\in\mathbb{N}} \mu^\star(A_n).$$

Satz 12.3.8. *Es bezeichne μ den Inhalt auf dem Ring $\mathcal{R} \subseteq \mathcal{P}(X)$. Zu jeder Menge $E \subseteq X$ bezeichne*

$$\mathcal{M}_E := \left\{ \{U_n : n \in \mathbb{N}\} : E \subseteq \bigcup_{n=1}^{\infty} U_n \text{ und } U_n \in \mathcal{R} \text{ für alle } n \in \mathbb{N} \right\}. \tag{12.5}$$

die Menge aller Überdeckungen von E. Dann definiert die Mengenfunktion

$$\mu^\star(E) = \begin{cases} \inf\left\{ \sum_{n=1}^\infty \mu(U_n) : \{U_n : n \in \mathbb{N}\} \in \mathcal{M}_E \right\}, & \textit{falls } \mathcal{M}_E \neq \emptyset, \\ +\infty, & \textit{falls } \mathcal{M}_E = \emptyset, \end{cases} \tag{12.6}$$

ein äußeres Maß.

Beweis. Da μ Inhalt vom Ring $\mathcal{R}$ ist, folgt $\mu^\star(E) \geq 0$ und $\mu^\star(\emptyset) = 0$. Wir zeigen nun die Monotonie. Dazu seien $E_1 \subseteq E_2 \subseteq X$ beliebig gewählt. Offensichtlich gilt $\mathcal{M}_{E_1} \subseteq \mathcal{M}_{E_2}$, denn jede Überdeckung von E_2 ist wegen $E_1 \subseteq E_2$ auch Überdeckung von E_1. Folglich gilt

$$\mu^\star(E_1) \leq \mu^\star(E_2).$$

Um (12.4) zu zeigen, betrachten wir beliebige Mengen $E_k \subseteq X$, $k \in \mathbb{N}$. Falls ein Index $l \in \mathbb{N}$ existiert mit $\mu^\star(E_l) = \infty$, dann ist die Aussage trivial. Sei also $\mu^\star(E_k) < \infty$ für alle $k \in \mathbb{N}$. Zu jedem $\varepsilon > 0$ existieren Überdeckungen $\{U_n^{(k)} : n \in \mathbb{N}\} \in \mathcal{M}_{E_k}$ von E_k so dass

$$\sum_{n=1}^\infty \mu(U_n^{(k)}) < \mu^\star(E_k) + \frac{\varepsilon}{2^k}$$

für alle $k \in \mathbb{N}$. Da jedoch $\{U_n^{(k)} : n, k \in \mathbb{N}\}$ eine Überdeckung von $\bigcup_{k=1}^\infty E_k$ ist, folgt

$$\mu^\star\left(\bigcup_{k=1}^\infty E_k\right) \leq \sum_{k=1}^\infty \sum_{n=1}^\infty \mu(U_n^{(k)}) < \sum_{k=1}^\infty \left\{ \mu^\star(E_k) + \frac{\varepsilon}{2^k} \right\} = \varepsilon + \sum_{k=1}^\infty \mu^\star(E_k).$$

Durch den Grenzübergang $\varepsilon \to 0$ erhalten wir hieraus (12.4). □

Satz 12.3.9. *Unter den Voraussetzungen des vorigen Satzes ist das durch* (12.5) *und* (12.6) *definierte äußere Maß $\mu^\star$ genau dann Fortsetzung von μ auf $\mathcal{P}(X)$, falls μ auf $\mathcal{R}$ subvolladditiv ist.*

Beweis. "$\Rightarrow$": Es gelte $\mu^\star(A) = \mu(A)$ für $A \in \mathcal{R}$. Für beliebige $A_n \in \mathcal{R}$, $n \in \mathbb{N}$, mit $A = \bigcup_{n=1}^\infty A_n \in \mathcal{R}$ gilt

$$\mu(A) = \mu^\star(A) = \mu^\star\left(\bigcup_{n=1}^\infty A_n\right) \leq \sum_{k=1}^\infty \mu^\star(A_n) = \sum_{n=1}^\infty \mu^\star(A_n),$$

das heißt μ ist subvolladditiv.

"$\Leftarrow$": Sei μ subvolladditiv und $A \in \mathcal{R}$ beliebig. Definieren wir $A_1 = A$ und $A_n = \emptyset$ für $n > 1$, so gilt $A = \bigcup_{n\in\mathbb{N}} A_n$ und $\{A_n : n \in \mathbb{N}\}$ ist eine Überdeckung von A. Somit gilt

$$\mu^\star(A) \leq \sum_{n=1}^\infty \mu(A_n) = \mu(A) + \sum_{n=2}^\infty \mu(\emptyset) = \mu(A).$$

Umgekehrt existiert zu jedem $\varepsilon > 0$ eine Überdeckung $\{A_n : n \in \mathbb{N}\} \in \mathcal{M}_A$ von A mit

$$\mu^\star(A) < \sum_{n=1}^{\infty} \mu(A_n) + \varepsilon.$$

Folglich ist

$$\mu(A) \leq \mu\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mu(A_n) < \mu^\star(A) + \varepsilon.$$

Der Grenzübergang $\varepsilon \to 0$ liefert schließlich zusammen mit dem vorigen Resultat $\mu(A) = \mu^\star(A)$. □

Satz 12.3.10. *Es gelten die Voraussetzungen von Satz 12.3.8, zusätzlich sei jedoch μ subvolladditiv auf dem Ring $\mathcal{R}$. Dann gilt für jedes $A \subseteq \mathcal{R}$ und beliebigen $E \subseteq X$*

$$\mu^\star(E) = \mu^\star(A \cap E) + \mu^\star(E \setminus A).$$

Beweis. Gemäß dem vorigen Satz ist das äußere Maß $\mu^\star$ eine Fortsetzung von μ auf $\mathcal{P}(X)$. Aufgrund der Identität $E = (A \cap E) \cup (E \setminus A)$ folgt daher

$$\mu^\star(E) \leq \mu^\star(A \cap E) + \mu^\star(E \setminus A).$$

Falls $\mu^\star(E) = \infty$ ist, dann gilt trivialerweise auch $\mu^\star(A \cap E) + \mu^\star(E \setminus A) \leq \mu^\star(E)$. Sei also angenommen, dass $\mu^\star(E) < \infty$. Zu jedem $\varepsilon > 0$ existiert dann eine Überdeckung $\{U_n : n \in \mathbb{N}\} \in \mathcal{M}_E$ von E mit

$$\mu^\star(E) + \varepsilon > \sum_{n=1}^{\infty} \mu(U_n).$$

Definieren wir

$$V_n := A \cap U_n, \qquad W_n := U_n \setminus A, \qquad n \in \mathbb{N},$$

dann gilt $V_n, W_n \in \mathcal{R}$ sowie $U_n = V_n \cup W_n$ und $V_n \cap W_n = \emptyset$ für alle $n \in \mathbb{N}$. Ferner gilt

$$A \cap E \subseteq A \cap \bigcup_{n=1}^{\infty} U_n = \bigcup_{n=1}^{\infty} V_n, \qquad E \setminus A \subseteq \left(\bigcup_{n=1}^{\infty} U_n\right) \setminus A = \bigcup_{n=1}^{\infty} W_n,$$

sowie $\mu(U_n) = \mu(V_n) + \mu(W_n)$ wegen der Additivität des Inhaltes μ. Fassen wir zusammen, so ergibt sich

$$\begin{aligned}\mu^\star(E) + \varepsilon &\geq \sum_{n=1}^{\infty} \mu(U_n) = \sum_{n=1}^{\infty} \mu(V_n) + \sum_{n=1}^{\infty} \mu(W_n) \geq \mu\left(\bigcup_{n=1}^{\infty} V_n\right) + \mu\left(\bigcup_{n=1}^{\infty} W_n\right) \\ &\geq \mu^\star(E \cap A) + \mu^\star(E \setminus A).\end{aligned}$$

□

Definition 12.3.11 (Caratheodory). *Sei $\mu^\star$ äußeres Maß auf X. Dann heißt die Teilmenge $B \subseteq X$ $\mu^\star$-**messbar**, falls $\mu^\star(E) = \mu^\star(E \cap B) + \mu^\star(E \setminus B)$ für alle $E \subseteq X$ gilt. Die Menge aller $\mu^\star$-messbaren Mengen in X bezeichnen wir mit $\mathcal{A}_{\mu^\star}(X)$.*

Proposition 12.3.12. *Sei $A \subseteq X$ mit $\mu^\star(A) = 0$, dann ist $A \in \mathcal{A}_{\mu^\star}(X)$, das heißt A ist $\mu^\star$-messbar.*

Beweis. Sei $E \subseteq X$ beliebig, dann ist $E \cap A \subseteq A$ und wegen der Monotonie gilt

$$0 \leq \mu^\star(E \cap A) \leq \mu^\star(A) = 0,$$

dies bedeudet $\mu^\star(E \cap A) = 0$. Wegen $E \setminus A \subseteq E$ folgt aus (12.4)

$$\mu^\star(E) \leq \mu^\star(E \cap A) + \mu^\star(E \setminus A) \leq 0 + \mu^\star(E),$$

beziehungsweise

$$\mu^\star(E) = \mu^\star(E \cap A) + \mu^\star(E \setminus A).$$

□

Proposition 12.3.13. *Gilt $A \in \mathcal{A}_{\mu^\star}(X)$, dann ist die Komplementärmenge $X \setminus A$ ebenfalls $\mu^\star$-messbar.*

Beweis. Für beliebiges $E \subseteq X$ gilt $E \setminus A = E \cap (X \setminus A)$ und $E \cap A = E \setminus (X \setminus A)$. Damit folgt aber schon die Behauptung gemäß

$$\mu^\star(E) = \mu^\star(E \cap A) + \mu^\star(E \setminus A) = \mu^\star\big(E \setminus (X \setminus A)\big) + \mu^\star\big(E \cap (X \setminus A)\big).$$

□

Satz 12.3.14. *Sei $\mu^\star$ ein äußeres Maß auf X. Dann ist $\mathcal{A}_{\mu^\star}(X)$ eine σ-Algebra und $\mu^\star$ ein Maß auf $\mathcal{A}_{\mu^\star}(X)$*

Beweis. Aus den beiden vorhergehenden Propositionen folgt, dass $\emptyset$ und X in $\mathcal{A}_{\mu^\star}(X)$ enthalten sind. Wir zeigen nun, dass aus $A, B \in \mathcal{A}_{\mu^\star}(X)$ auch $A \cap B \in \mathcal{A}_{\mu^\star}(X)$ folgt. Einfachheitshalber schreiben wir kurz $\overline{A}$ anstelle von $X \setminus A$ für das Komplement einer Menge $A \subseteq X$. Sei $E \subseteq X$ beliebig gewählt, dann gilt

$$\mu^\star\big(E \cap (A \cap B)\big) + \mu^\star\big(E \cap \overline{(A \cap B)}\big) = \mu^\star(E \cap A \cap B) + \mu^\star\big(E \cap \overline{(A \cap B)}) \cap (A \cup \overline{A})\big).$$

Aufgrund der $\mu^\star$-Messbarkeit von A gilt nun

$$\begin{aligned}\mu^\star\big(E \cap \overline{(A \cap B)}) \cap (A \cup \overline{A})\big) &= \mu^\star\big(E \cap (\overline{A} \cup \overline{B}) \cap A\big) + \mu^\star\big(E \cap (\overline{A} \cup \overline{B}) \cap \overline{A}\big) \\ &= \mu^\star\big(E \cap A \cap \overline{B}\big) + \mu^\star\big(E \cap \overline{A}\big),\end{aligned}$$

und folglich

$$\mu^\star\big(E \cap (A \cap B)\big) + \mu^\star\big(E \cap \overline{(A \cap B)}\big) = \mu^\star(E \cap A \cap B) + \mu^\star\big(E \cap A \cap \overline{B}\big) + \mu^\star\big(E \cap \overline{A}\big).$$

Wegen $B \in \mathcal{A}_{\mu^\star}(X)$ gilt jedoch

$$\mu^\star(E \cap A) = \mu^\star(E \cap A \cap B) + \mu^\star\big(E \cap A \cap \overline{B}\big)$$

und insgesamt

$$\mu^\star\big(E \cap (A \cap B)\big) + \mu^\star\big(E \cap \overline{(A \cap B)}\big) = \mu^\star(E \cap A) + \mu^\star\big(E \cap \overline{A}\big) = \mu^\star(E),$$

mit anderen Worten, es gilt $A \cap B \in \mathcal{A}_{\mu^\star}(X)$. Da aus $A \in \mathcal{A}_{\mu^\star}(X)$ gemäß Proposition 12.3.13 auch $\overline{A} \in \mathcal{A}_{\mu^\star}(X)$ folgt, liefert die Identität

$$A \cup B = \overline{\overline{A} \cap \overline{B}}$$

dass $A, B \in \mathcal{A}_{\mu^\star}(X)$ auch $A \cup B \in \mathcal{A}_{\mu^\star}(X)$ impliziert. Somit ist $\mathcal{A}_{\mu^\star}(X)$ eine Algebra.

Wir zeigen nun, dass $\mathcal{A}_{\mu^\star}(X)$ eine σ-Algebra ist. Seien zunächst $A_i \in \mathcal{A}_{\mu^\star}(X)$, $i \in \mathbb{N}$, paarweise disjunkte Mengen. Dann folgt aus der Messbarkeit mit Hilfe der vollständigen Induktion $\bigcup_{i=1}^n A_i \in \mathcal{A}_{\mu^\star}(X)$ und

$$\mu^\star\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n \mu^\star(A_i).$$

Sei nun $E \subseteq X$ beliebig, dann gilt aufgrund der Monotonie für alle $n \in \mathbb{N}$

$$\begin{aligned}
\mu^\star(E) &= \mu^\star\left(E \cap \bigcup_{i=1}^n A_i\right) + \mu^\star\left(E \cap \overline{\bigcup_{i=1}^n A_i}\right) \\
&\geq \mu^\star\left(\bigcup_{i=1}^n (E \cap A_i)\right) + \mu^\star\left(E \cap \overline{\bigcup_{i=1}^\infty A_i}\right) \\
&= \sum_{i=1}^n \mu^\star(E \cap A_i) + \mu^\star\left(E \cap \overline{\bigcup_{i=1}^\infty A_i}\right).
\end{aligned}$$

Für $n \to \infty$ folgt hieraus

$$\mu^\star(E) \geq \sum_{i=1}^\infty \mu^\star(E \cap A_i) + \mu^\star\left(E \cap \overline{\bigcup_{i=1}^\infty A_i}\right).$$

Andererseits impliziert aber (12.4) auch

$$\mu^\star(E) \leq \mu^\star\left(E \cap \bigcup_{i=1}^\infty A_i\right) + \mu^\star\left(E \cap \overline{\bigcup_{i=1}^\infty A_i}\right) \leq \sum_{i=1}^\infty \mu^\star(E \cap A_i) + \mu^\star\left(E \cap \overline{\bigcup_{i=1}^\infty A_i}\right).$$

Dies bedeutet

$$\mu^\star(E) = \mu^\star\left(E \cap \bigcup_{i=1}^{\infty} A_i\right) + \mu^\star\left(E \cap \overline{\bigcup_{i=1}^{\infty} A_i}\right)$$

bzw. $\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}_{\mu^\star}(X)$.

Seien nun für $i \in \mathbb{N}$ die Mengen $A_i \in \mathcal{A}_{\mu^\star}(X)$ beliebig. Definieren wir die Mengen

$$B_i := \begin{cases} A_1, & i = 1, \\ A_i \setminus \bigcup_{j=1}^{i-1} A_j, & i > 1, \end{cases}$$

dann gilt $B_i \in \mathcal{A}_{\mu^\star}(X)$ und $\bigcup_{i=1}^{n} A_i = \bigcup_{i=1}^{n} B_i$ für alle $i \in \mathbb{N}$ sowie $B_i \cap B_j = \emptyset$ für $i \neq j$. Nach dem oben bewiesenen ergibt sich $\bigcup_{i=1}^{\infty} C_i \in \mathcal{A}_{\mu^\star}(X)$ und folglich auch $\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}_{\mu^\star}(X)$. Somit ist $\mathcal{A}_{\mu^\star}(X)$ ist eine σ-Algebra.

Wir zeigen nun, dass $\mu^\star$ ein Maß auf der σ-Algebra $\mathcal{A}_{\mu^\star}(X)$ ist. Seien für $i \in \mathbb{N}$ die Mengen $A_i \in \mathcal{A}_{\mu^\star}(X)$ paarweise disjunkt. Wegen $\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}_{\mu^\star}(X)$ folgt aus der Gleichung

$$\mu^\star\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mu^\star(A_i)$$

für $n \to \infty$ auch

$$\mu^\star\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu^\star(A_i).$$

Somit ist $\mu^\star$ volladditiv und damit ein Maß auf $\mathcal{A}_{\mu^\star}(X)$. □

Satz 12.3.15 (Fortsetzungssatz von Hahn). *Sei φ eine nichtnegative additive Mengenfunktion auf einen Semiring $\mathcal{S}$. Das durch* (12.5) *und* (12.6) *eingeführte Maß $\mu^\star$ ist genau dann eine Fortsetzung von φ auf die σ-Algebra $\mathcal{A}_{\mu^\star}(X)$, wenn φ subvolladditiv auf $\mathcal{S}$ ist.*

Beweis. Aus den Sätzen 12.3.4 und 12.3.9 folgt, dass φ genau dann als äußers Maß auf $\mathcal{P}(X)$ fortgesetzt werden kann, wenn φ auf $\mathcal{S}$ subvolladditiv ist. Nach Satz 12.3.10 ist $\mathcal{R}(\mathcal{S}) \subseteq \mathcal{A}_{\mu^\star}(X)$ und $\mu^\star$ definiert nach Satz 12.3.14 sogar ein Maß auf $\mathcal{A}_{\mu^\star}(X)$. □

Definition 12.3.16. *Sei $\mathcal{M} \subseteq \mathcal{P}(X)$, dann heißt die Mengenfunktion $\varphi : \mathcal{M} \to \overline{\mathbb{R}}$* **$\sigma$-endlich**, *falls zu jedem $A \in \mathcal{M}$ abzählbar viele paarweise disjunkte Mengen A_n existieren mit $|\varphi(A_n)| < \infty$ für alle $n \in \mathbb{N}$ und*

$$A = \bigcup_{n=1}^{\infty} A_n.$$

Satz 12.3.17 (Ergänzung zum Fortsetzungssatz von Hahn). *Sei φ eine nichtnegative, subvolladditive und σ-endliche Mengenfunktion auf einem Semiring $\mathcal{S} \subseteq \mathcal{P}(X)$. Ferner existiere zu X eine abzählbare Überdeckung aus Mengen von $\mathcal{S}$, dass heißt $\mathcal{M}_X \neq \emptyset$. Dann ist $\mu^\star$ definiert vermittels* (12.5) *und* (12.6) *ein σ-endliches Maß auf der σ-Algebra $\mathcal{A}_{\mu^\star}(X)$.*

Beweis. Aus dem Fortsetzungssatz von Hahn folgt, dass $\mu^\star$ eine Fortsetzung von φ auf $\mathcal{A}_{\mu^\star}(X)$ ist. Ist φ σ-endlich auf $\mathcal{S}$, so ist $\mu^\star$ offensichtlich auch σ-endlich auf $\mathcal{R}(\mathcal{S})$. Es sei nun $A \in \mathcal{A}_{\mu^\star}(X)$ beliebig. Wir müssen zeigen, dass paarweise disjunkte Mengen $A_n \in \mathcal{A}_{\mu^\star}(X)$ mit $\mu^\star(A_n) < \infty$, $n \in \mathbb{N}$, existieren, so dass $A = \bigcup_{i=1}^\infty A_n$ und $\mu^\star(A) = \sum_{i=1}^\infty \mu^\star(A_n)$. Nun gibt es aber nach Voraussetzung zu X eine Überdeckung $\{U_n : n \in \mathbb{N}\}$ aus Mengen von $\mathcal{S} \subseteq \mathcal{R}(\mathcal{S})$, das heißt es gilt $X = \bigcup_{n=1}^\infty U_n$. Ohne Beschränkung der Allgemeinheit dürfen wir die Mengen U_n, $n \in \mathbb{N}$, als paarweise disjunkt annehmen, denn sonst betrachten die Überdeckung

$$V_n = \begin{cases} U_1, & n = 1, \\ U_n \setminus \bigcup_{i=1}^{n-1} V_i, & n > 1. \end{cases}$$

Wegen der σ-Endlichkeit von $\mathcal{R}(\mathcal{S})$ lässt sich jedes U_n darstellen als abzählbare Vereinigung paarweise disjunkter Mengen $V_i^{(n)}$ mit $\mu(V_i^{(n)}) < \infty$ und $U_n = \bigcup_{i=1}^\infty V_i^{(n)}$ beziehungsweise

$$X = \bigcup_{n=1}^\infty \bigcup_{i=1}^\infty V_i^{(n)}.$$

Damit erhalten wir aber auch

$$A = \bigcup_{n=1}^\infty \bigcup_{i=1}^\infty (V_i^{(n)} \cap A)$$

und folglich die Behauptung wegen $\mu^\star(V_i^{(n)} \cap A) \leq \mu^\star(V_i^{(n)})$. □

Definition 12.3.18. *Ein Maß $\mu^\star$ auf einer σ-Algebra $\mathcal{A}$ heißt* **vollständig**, *falls aus $\mu^\star(A) = 0$ für $A \in \mathcal{A}$ folgt, dass jede Teilmenge $B \subseteq A$ ebenfalls in $\mathcal{A}$ enthalten ist (und folglich $\mu^\star(B) = 0$ gilt).*

Satz 12.3.19. *Sei $\mathcal{A}_{\mu^\star}(X) \subseteq \mathcal{P}(X)$ eine σ-Algebra mit dem Maß $\mu^\star$, dann ist das Mengensystem*

$$\overline{\mathcal{A}_{\mu^\star}(X)} = \{A \cup N : A \in \mathcal{A}_{\mu^\star}(X) \text{ und } N \subseteq X \text{ ist Teilmenge einer Menge vom Maße 0}\}$$

eine vollständige σ-Algebra mit dem Maß

$$\overline{\mu^\star}(A \cup N) = \mu^\star(A).$$

Dabei wird $\overline{\mu^\star}$ Vervollständigung von $\mu^\star$ genannt.

Beweis. Man zeigt leicht das $\overline{\mathcal{A}_{\mu^\star}(X)}$ eine σ-Algebra ist. Sei $A \in \overline{\mathcal{A}_{\mu^\star}(X)}$ mit $\mu^\star(A) = 0$. Für beliebiges $B \subseteq A$ gilt dann gilt wegen der Monotonie

$$0 \leq \mu^\star(B) \leq \mu^\star(A) = 0,$$

dies bedeutet $\mu^\star(B) = 0$. Hieraus folgt mit Proposition 12.3.13, dass B $\mu^\star$-messbar ist, dies bedeutet $B \in \overline{\mathcal{A}_{\mu^\star}(X)}$. □

Definition 12.3.20. *Sei $X = \mathbb{R}^n$ und sei $I^n := \{\mathbf{x} = (x_1, \ldots, x_n) : a_i < x_i \leq b_i,\ i = 1, \ldots, n\}$. Dann ist $\mathcal{S} = \{I_k\}$ ein Semiring. Der von $\mathcal{S}$ erzeugte Ring*

$$\mathcal{E} = \mathcal{R}(\mathcal{S})$$

heißt **Ring der Elementarmengen**. *Wir definieren den Inhalt von I^n durch*

$$\varphi(I^n) = \prod_{i=1}^{n}(b_i - a_i)$$

als subvolladditiven Inhalt auf den Semiring.

Die gemäß dem Fortsetzungssatz von Hahn und dessen Ergänzung definierte σ-Algebra aller $\mu^\star$-messbaren Mengen $L = \mathcal{A}_{\mu^\star}(X) = \overline{\mathcal{A}_{\mu^\star}(X)}$ heißt σ-Algebra der Lebesgue-messbaren Mengen. Eine Menge $A \in L$ heißt **Lebesgue-messbar** *und das Maß $\mu^\star$ heißt das* **Lebesgue-Maß**.

Satz 12.3.21. *Bezeichne $\mu^\star$ das Lebesgue-Maß. Ist $A \subseteq \mathbb{R}^n$ Lebesgue-messbar, dann existieren zu jedem $\varepsilon > 0$ eine abgeschlossene Menge F und eine offene Menge G mit $F \subseteq A \subseteq G$, so dass $\mu^\star(A \setminus F) < \varepsilon$ und $\mu^\star(G \setminus A) < \varepsilon$.*

Beweis. Ohne Einschränkung der Allgemeinheit sei A eine beschränkte Menge. Denn anderenfalls kann man A zerlegen in eine abzählbare Vereinigung $A = \bigcup_{k=1}^{\infty} A_k$ mit beschränkten Mengen A_k. Da daher A beschränkt angenommen werden kann, gilt

$$\mu^\star(A) < \infty.$$

Nach der Definition von $\mu^\star$ existiert zu jedem $\varepsilon > 0$ eine abzählbare Überdeckung $E_k \in \mathcal{E}$, $k \in \mathbb{N}$, mit $A \subseteq \bigcup_{k=1}^{\infty} E_i$ und

$$\varepsilon + \mu^\star(A) \geq \sum_{k=1}^{\infty} \mu^\star(E_k).$$

Wegen $E_k \in \mathcal{E}$ existiert zu jedem $E_k = \bigcup_{l=1}^{n_k} I_l^{(k)}$ mit paarweise disjunkten Zellen $I_l^{(k)}$ und daher

$$\varepsilon + \mu^\star(A) \geq \sum_{k=1}^{\infty} \sum_{l=1}^{n_k} \mu^\star(I_l^{(k)}).$$

Da hierin die Summe über abzählbar viele Zellen gebildet wird, dürfen wir vereinfachend schreiben

$$\varepsilon + \mu^\star(A) \geq \sum_{k=1}^{\infty} I_k.$$

Es existieren nun offene Zellen

$$J_k := \{\mathbf{x} \in \mathbb{R}^n : a_i < x_i < b_i + \delta_k,\ i = 1, \ldots, n\}$$

mit $\delta_k > 0$ geeignet, so dass

$$\mu^\star(J_k) \leq \mu^\star(I_k) + \frac{\varepsilon}{2^k}, \qquad k \in \mathbb{N}.$$

Setzen wir $G := \bigcup_{k\in\mathbb{N}} J_k$, dann ist G offen und es gilt

$$\begin{aligned}\mu^\star(G) = \mu^\star\left(\bigcup_{k=1}^{\infty} J_k\right) &\leq \sum_{l=1}^{\infty} \mu^\star(J_k) \leq \varepsilon \sum_{k=1}^{\infty} \frac{1}{2^k} + \sum_{k=1}^{\infty} \mu^\star(I_k) \leq \varepsilon + \sum_{k=1}^{\infty} \mu^\star(I_k) \\ &\leq \mu^\star(A) + 2\varepsilon.\end{aligned}$$

Wegen $\mu^\star(G) = \mu^\star(A) + \mu^\star(G \setminus A)$ folgt hieraus der erste Teil der Behauptung

$$\mu^\star(G \setminus A) = \mu^\star(G) - \mu^\star(A) < 2\varepsilon.$$

Der zweite Teil der Behauptung ergibt sich wie folgt. Ist A Lebesgue-messbar, dann ist $X \setminus A$ ebenfalls Lebesgue-messbar und es existiert eine offene Menge G mit $X \setminus A \subseteq G$ und $\mu^\star\big(G \setminus (X \setminus A)\big) = \mu^\star(G \cap A) < \varepsilon$. Nun ist $F := X \setminus G \subseteq A$ abgeschlossen und $A \setminus F = G \cap A$, dies bedeutet $\mu^\star(A \setminus F) < \varepsilon$. □

Definition 12.3.22. *Die Vereinigung F_σ abzählbar vieler abgeschlossener Mengen F_n, $n \in \mathbb{N}$, bzw. der Schnitt G_δ abzählbar vieler offener Mengen G_n, $n \in \mathbb{N}$, heißt* **Borel-Menge**. *Die Familie $\mathcal{B}(X)$ aller Borel-Mengen ist damit die kleinste σ-Algebra, die alle offenen Teilmengen von X enthält.*

Korollar 12.3.23. *Sei $A \subset \mathbb{R}^n$ Lebesgue-messbar, dann existieren Borel-Mengen F_σ und G_δ, so dass*

$$\mu^\star(F_\sigma \setminus A), \qquad \mu^\star(G_\delta \setminus A) = 0.$$

Beweis. Gemäß dem letzen Satz existieren für jedes $n \in \mathbb{N}$ eine abgeschlossen Menge $F_n \supseteq A$ mit $\mu^\star(A \setminus F_n) < 1/n$ und eine offene Menge $G_n \subseteq A$ mit $\mu^\star(G_n \setminus A) < 1/n$. Setzen wir

$$G_\delta := \bigcap_{n=1}^{\infty} G_n, \qquad F_\sigma := \bigcup_{n=1}^{\infty} F_n,$$

dann gilt

$$\mu^\star(G_\delta \setminus A) = \lim_{n\to\infty} \mu^\star\left(\bigcap_{i=1}^{n} G_n \setminus A\right) \leq \lim_{n\to\infty} \mu^\star(G_n \setminus A) = 0.$$

und analog $\mu^\star(A \setminus F_\sigma) = 0$. □

Bemerkung 12.3.24.

1. *Wir haben gesehen, dass jede Lebesgue-messbare Menge A zerlegt werden kann in die disjunkte Vereinigung einer Borel-Menge B und einer Nullmenge N vom Maß 0, das heißt $A = B \cup N$. Umgekehrt ist jede Borel-Menge $\mu^\star$-messbar.*

2. *Abzählbare Teilmengen $N \subset \mathbb{R}^n$ haben das Lebesgue-Maß 0. Es gibt aber auch überabzählbare Nullmengen, beispielsweise die Cantor-Menge.*

12.4 Messbare Funktionen

12.4.1 Definitionen und Eigenschaften

Definition 12.4.1. *Sei X eine beliebige Menge und $\mathcal{M} \subseteq \mathcal{P}(X)$ eine σ-Algebra mit dem Maß μ. Dann heißt $(X, \mathcal{M})$* **messbarer Raum** *über X. Desweiteren heißt die Menge $A \subseteq X$ genau dann* **messbar**, *falls sie in der σ-Algebra enthalten ist, d.h. $A \in \mathcal{M}$. Das Tripel $(X, \mathcal{M}, \mu)$ heißt* **Maßraum**. *Wir schreiben auch kurz $\mathcal{M}$ statt $(X, \mathcal{M})$ und $(\mathcal{M}, \mu)$ statt $(X, \mathcal{M}, \mu)$.*

Im folgenden werden wir immer einen Maßraum $(X, \mathcal{M}, \mu)$ zugrundelegen, auch wenn dies nicht immer explizit angegeben wird.

Definition 12.4.2 (Messbare Funktionen).

1. *Seien $(X, \mathcal{M})$ und $(Y, \mathcal{N})$ zwei messbare Räume, dann heißt eine Funktion*

$$f : X \to Y$$

   genau dann messbar, falls für alle $B \in \mathcal{N}$ gilt $f^{-1}(B) \in \mathcal{M}$, dies bedeutet, das Urbild jeder in Y messbaren Menge ist in X messbar.

2. *Sei $f : X \to \overline{\mathbb{R}}$, dann heißt f messbar, falls die Menge*

$$X(f \geq a) := \{x \in X : f(x) \geq a\}$$

   für jedes $a \in \mathbb{R}$ messbar ist.

Bemerkung 12.4.3. *In dieser Definition impliziert der erste Punkt den zweiten, falls wir $\mathbb{R}$ mit dem Lebesgue-Maß versehen.*

Beispiel 12.4.4. *Die Funktion*

$$f(x) = \begin{cases} 1/x, & x \neq 0, \\ \infty, & x = 0, \end{cases}$$

ist messbar bezüglich der σ-Algebra der Lebesgue-messbaren Mengen.

Satz 12.4.5. *Sei $f : \mathbb{R}^n \to \mathbb{R}$ stetig und $X = \mathbb{R}^n$. Sei weiterhin $\mathcal{M}$ die σ-Algebra der Lebesgue-messbaren Mengen, dann ist f messbar.*

Beweis. Der einfache Beweis verbleibt dem Leser als Übung. □

Satz 12.4.6. *Sei $(X, \mathcal{M})$ ein messbarer Raum und $f : X \to \mathbb{R}$, dann ist f genau dann messbar, falls eine der folgenden Voraussetzungen erfüllt ist:*

1. *Die Menge $X(f > a) = \{x \in X : f(x) > a\}$ ist für alle $a \in \mathbb{R}$ messbar.*

2. *Die Menge* $X(f \geq a) = \{x \in X : f(x) \geq a\}$ *ist für alle* $a \in \mathbb{R}$ *messbar.*

3. *Die Menge* $X(f < a) = \{x \in X : f(x) < a\}$ *ist für alle* $a \in \mathbb{R}$ *messbar.*

4. *Die Menge* $X(f \leq a) = \{x \in X : f(x) \leq a\}$ *ist für alle* $a \in \mathbb{R}$ *messbar.*

Beweis. Der Beweis verbleibt dem Leser als Übung. □

Satz 12.4.7. *Sei* f *messbar, dann ist auch die Funktion* $|f|$ *messbar.*

Beweis. Wir betrachten die Menge

$$X(|f| > a) = X(f > a) \cup X(f < -a).$$

Da f messbar ist, sind die beiden letzten Mengen messbar, das heißt sie sind in der σ-Algebra $\mathcal{A}$ enthalten. Da $\mathcal{M}$ eine σ-Algebra ist, ist auch $X(|f| > a) \in \mathcal{M}$. □

Bemerkung 12.4.8. *Die Umkehrung dieses Satzes gilt im allgemeinen nicht wie folgendes Gegenbeispiel zeigt: Für* $A \subseteq X$ *mit* $A \notin \mathcal{M}$*, d.h.* A *ist nicht messbar, definieren wir*

$$f(x) := \begin{cases} +1, & x \notin A, \\ -1, & x \in A. \end{cases}$$

Dann gilt $|f(x)| = 1$*, also ist die Menge* $X(|f| > a)$ *messbar, jedoch ist* $X(f > 0) = A$ *nach Voraussetzung nicht messbar.*

Satz 12.4.9. *Seien für alle* $n \in \mathbb{N}$ *die Funktionen* $f_n : X \to \mathbb{R}$ *messbar. Dann sind auch die Funktionen* $g, h : X \to \mathbb{R}$ *definiert durch*

$$g(x) := \sup_{n \in \mathbb{N}} f_n(x), \qquad h(x) := \inf_{n \in \mathbb{N}} f_n(x),$$

sowie die Funktionen $p, q : X \to \mathbb{R}$ *definiert durch*

$$p(x) := \limsup_{n \to \infty} f_n(x), \qquad q(x) := \liminf_{n \to \infty} f_n(x),$$

messbar.

Beweis. Es gilt

$$X(g > a) = \{x \in X : g(x) > a\} = \bigcap_{n=1}^{\infty} X(f_n > a).$$

Aus der Messbarkeit der Funktionen f_n folgt $X(f_n > a) \in \mathcal{M}$. Da $\mathcal{M}$ eine σ-Algebra ist, ist der abzählbare Schnitt $\bigcap_{n=1}^{\infty} X(f_n > a)$ ebenfalls in $\mathcal{M}$. Folglich ist g messbar. Analog zeigt man, dass h messbar ist.

Aus

$$p(x) = \limsup_{n\to\infty} f_n(x) = \inf_{n\in\mathbb{N}} \sup_{k\geq n} f_k(x)$$

folgt die Messbarkeit von p analog zum obigen Beweis aus der Messbarkeit von g und h. Analog zeigt man für

$$q(x) = \sup_{n\in\mathbb{N}} \inf_{n\geq k} f_k(x)$$

die Messbarkeit. □

Lemma 12.4.10. *Jede offene Teilmenge $G \subseteq \mathbb{R}^n$ lässt sich schreiben als abzählbare Vereinigungen offener n-Zellen*

$$I^n := \{\mathbf{x} \in \mathbb{R}^n : a_i < x_i < b_i\} = \bigotimes_{i=1}^{n} (a_i, b_i).$$

Beweis. Ist $G \subseteq X$ offen, so existiert zu jedem Punkt $\mathbf{x} = (x_1, \ldots, x_n) \in G$ ein $r_i \in \mathbb{Q}$ und $m \in \mathbb{N}$, so dass die n-Zelle

$$I^n_{\mathbf{x}} = \left\{\mathbf{y} \in \mathbb{R}^n : r_i - \frac{1}{m} < y_i < r_i + \frac{1}{m}\right\} \subseteq G$$

den Punkt $\mathbf{x}$ enthält. Dabei ist die Gesamtheit aller n-Zellen $I^n_{\mathbf{x}}$ abzählbar. Wegen $I^n_{\mathbf{x}} \subseteq G$ gilt einerseits $\bigcup_{\mathbf{x}\in G} I^n_{\mathbf{x}} \subseteq G$, und andererseits folgt aus $\mathbf{x} \in I^n_{\mathbf{x}}$ auch $G \subseteq \bigcup_{\mathbf{x}\in G} I^n_{\mathbf{x}}$, dies bedeutet wegen der Abzählbarkeit

$$G = \bigcup_{\mathbf{x}\in G} I^n_{\mathbf{x}} \cup \bigcup_{i=1}^{n} I^n_{\mathbf{x}_i}$$

mit gewissen $\mathbf{x}_i \in G$. □

Satz 12.4.11. *Seien $(X, \mathcal{M}, \mu)$ ein Maßraum und $f, g : X \to \mathbb{R}$ messbare Funktionen. Ist die Funktion $F : \mathbb{R}^2 \to R$ stetig, dann ist die Funktion $h : X \to \mathbb{R}$ definiert vermittels $h(x) := F\big(f(x), g(x)\big)$ ebenfalls messbar.*

Beweis. Sei $a \in \mathbb{R}$ beliebig und

$$G_a := \{(u, v) \in \mathbb{R}^2 : F(u, v) > a\}, \qquad H_a := X(h > a) = \{x \in X : h(x) > a\}.$$

Da das Intervall (a, ∞) offen ist, folgt aus der Stetigkeit von $F : \mathbb{R}^2 \to R$, dass $G_a = F^{-1}\big((a, \infty)\big)$ ebenfalls offen ist. Nach dem oben stehenden Lemma existieren folglich Zweizellen

$$I^2_k = (a_k, b_k) \otimes (c_k, d_k) = \{(u, v) : a_k < u < b_k,\ c_k < v < d_k\}$$

mit $G = \bigcup_{k=1}^{\infty} I^2_k$.

Nun ist $\big(f(x), g(x)\big) \in I_k$ genau dann, wenn $a_k < f(x) < b_k$ und $c_k < g(x) < d_k$. Hieraus folgt

$$\begin{aligned} H_a &= \big\{x \in X : \big(f(x), g(x)\big) \in G_a\big\} \\ &= \bigcup_{k=1}^{\infty} \big\{x \in X : \big(f(x), g(x)\big) \in I_k\big\} \\ &= \bigcup_{k=1}^{\infty} \{x \in X : a_k < f(x) < b_k,\ c_k < g(x) < d_k\} \\ &= \bigcup_{k=1}^{\infty} \big[X\big(f(x) > a_k\big) \cap X\big(f(x) < b_k\big) \cap X\big(g(x) > c_k\big) \cap X\big(g(x) < d_k\big)\big] \in \mathcal{M}. \end{aligned}$$

□

Korollar 12.4.12. *Seien $f, g : X \to \mathbb{R}$ messbare Funktionen, dann sind die Funktionen $f + g$, $f \cdot g$ und $c \cdot f$ für jedes $c \in \mathbb{R}$, sowie*

$$f_+(x) = \max\{f(x), 0\}, \qquad f_-(x) = -\min\{f(x), 0\}, \tag{12.7}$$

ebenfalls messbar.

Beweis. Die Behauptung ist im Falle $f \cdot g$ und $c \cdot f$ trivial, während die Messbarkeit von $f + g$ folgt aus dem letzen Satz mit $F(u, v) = u + v$. Die Messbarkeit von $f_+(x)$ bzw. $f_-(x)$ ergibt sich unmittelbar aus der Identität $f_+(x) = \frac{1}{2}\big(f(x) + |f(x)|\big)$ bzw. $f_-(x) = \frac{1}{2}\big(|f(x)| - f(x)\big)$. □

Bemerkung 12.4.13. *Es gilt $f_+(x), f_-(x) \geq 0$ und $f(x) = f_+(x) - f_-(x)$ für alle $x \in X$. Die Funktion f_+ heißt* **positiver Anteil** *von f während f_-* **negativer Anteil** *von f gennant wird.*

Definition 12.4.14. *Eine Funktion $\varphi : X \to \mathbb{R}$ heißt* **Treppenfunktion***, falls ihre Bildmenge*

$$\mathrm{Im}(\varphi) = \{y \in \mathbb{R} : \textit{es existieren } x \in X \textit{ mit } \varphi(x) = y\}$$

endlich ist.

Sei $E \subseteq X$, dann heißt die spezielle Treppenfunktion $\chi_E : X \to \mathbb{R}$ mit

$$\chi_E(x) = \begin{cases} 1, & x \in E, \\ 0, & x \notin E, \end{cases}$$

die **charakteristische Funktion** *von E.*

Proposition 12.4.15. *Sei $\varphi : X \to \mathbb{R}$ eine Treppenfunktion mit den paarweise verschiedenen Funktionswerten $c_i \in \mathbb{R}$, $i = 1, \ldots, n$, dann existieren Teilmengen $E_i \subseteq \mathbb{R}$ mit*

$$E_i \cap E_j = \emptyset, \qquad i \neq j,$$

und

$$\varphi(x) = \sum_{i=1}^{n} c_i \chi_{E_i}(x), \qquad x \in X.$$

Beweis. Setzen wir $E_i := \{x \in X : \varphi(x) = c_i\}$, $i = 1, \ldots, n$, dann erhalten wir sofort die Behauptung. □

Satz 12.4.16. *Eine Treppenfunktion $\varphi : X \to \mathbb{R}$ mit*

$$\varphi(x) = \sum_{i=1}^{n} c_i \chi_{E_i}(x)$$

ist genau dann eine messbare Funktion, falls für alle $i = 1, \ldots, n$ die Mengen E_i messbar sind, d.h. $E_i \in \mathcal{M}$.

Beweis. Ohne Einschränkung der Allgemeinheit gelte $c_i < c_{i+1}$ für alle $i = 1, \ldots, n$. Sei $\alpha \in \mathbb{R}$ beliebig gewählt, dann gibt es einen Index $i \in \mathbb{N}$, so dass $c_i \leq \alpha < c_{i+1}$. Folglich ist

$$X(\varphi < \alpha) = \bigcup_{i=1}^{j} E_i$$

genau dann messbar, falls für alle $i = 1, \ldots, j$ gilt $E_i \in \mathcal{M}$. □

Satz 12.4.17. *Sei $f : X \to \overline{\mathbb{R}}$ eine messbare Funktion, dann existiert eine Folge von messbaren Treppenfunktionen φ_n, $n \in \mathbb{N}$, derart, dass $\lim_{n\to\infty} \varphi_n(x) = f(x)$ für alle $x \in X$. Falls gilt $f(x) \geq 0$ für ein $x \in X$, dann ist die Folge $\big(\varphi_n(x)\big)$ nicht monoton fallend. Ist f beschränkt, so konvergiert die Folge (φ_n) sogar gleichmäßig gegen f.*

Beweis. Für jedes $n \in \mathbb{N}$ definieren wir

$$E_n^{(i)} := \left\{x \in X : \frac{i-1}{2^n} \leq f(x) < \frac{i}{2^n}\right\}, \qquad i = 1, \ldots, n \cdot 2^n,$$

und

$$F_n = \{x \in X : f(x) \geq n\}$$

Aufgrund der Zerlegung $f = f_+ - f_-$ mit $f_+, f_- \geq 0$, können wir nun ohne Beschränkung der Allgemeinheit annehmen, dass $f \geq 0$. Wegen der Messbarkeit von f sind dann die Mengen $E_n^{(i)}$ und F_n alle messbar und für ein festes $n \in \mathbb{N}$ gilt

$$X = F_n \cup \bigcup_{i=1}^{n\cdot 2^n} E_n^{(i)}.$$

Wir definieren die Treppenfunktion φ_n gemäß

$$\varphi_n(x) := \sum_{i=1}^{n\cdot 2^n} \frac{i-1}{2^n} \chi_{E_i}(x) + n\chi_{F_n}(x).$$

Nach dem vorhergehenden Satz sind die Funktionen φ_n für alle $\in \mathbb{N}$ messbar.

Wir zeigen nun die punktweise Konvergenz. Sei $x \in X$ mit $f(x) = \infty$, dann ist $x \in F_n$ für alle $n \in \mathbb{N}$ und es gilt $\lim_{n\to\infty} \varphi_n(x) = \infty$. Sei $x \in X$ mit $f(x) < \infty$, dann existiert für alle $n \in \mathbb{N}$ ein Index $1 \le i \le n \cdot 2^n$ mit

$$f(x) - \frac{i-1}{2^n} \le \frac{1}{2^n}, \qquad \frac{i}{2^n} - f(x) \le \frac{1}{2^n}. \tag{12.8}$$

Hieraus folgt schließlich die punktweise Konvergenz

$$|f(x) - \varphi_n(x)| \le \frac{1}{2^n} \overset{n\to\infty}{\longrightarrow} 0.$$

Die Monotonie der Folge $\big(\varphi_n(x)\big)$ für festes $x \in X$ ist klar.

Sei f nun beschränkt und $\varepsilon > 0$ beliebig. Wir wählen $n \in \mathbb{N}$ so, dass sowohl $f < n$ als auch $1/2^n < \varepsilon$ gilt. Wir bemerken, dass dann $F = \emptyset$ gilt. Folglich existiert zu jedem $x \in X$ ein Index $1 \le i \le n \cdot 2^n$ mit $x \in E_i$ und es folgt

$$|f(x) - \varphi_n(x)| \le \frac{1}{2^n} < \varepsilon,$$

das heißt φ_n konvergiert gleichmäßig gegen f. □

Definition 12.4.18. *Eine Aussage $A(x)$, die von $x \in E \in \mathcal{M}$ abhängt, heißt* **fast überall auf** E *oder* **für fast alle** $x \in E$ *gültig, falls eine Teilmenge $B \subseteq \mathcal{M}$ existiert mit $\mu(B) = 0$, so dass die Aussage $A(x)$ gültig ist für alle $x \in E \setminus B$.*

12.4.2 Konvergenzsätze

Satz 12.4.19. *Sei das Maß μ vollständig und die Funktionen $f_n : X \to \overline{\mathbb{R}}$, $n \in \mathbb{N}$, messbar. Gilt für $f : X \to \mathbb{R}$ dass $\lim_{n\to\infty} f_n(x) = f(x)$ fast überall auf X, so ist f messbar.*

Beweis. Sei $A = \{x \in X : \lim_{n\to\infty} f_n(x) \neq f(x)\}$. Dann gilt nach Voraussetzung $\mu(A) = 0$ und $A \in \mathcal{M}$. Da μ vollständig ist, ist jede Teilmenge $B \subseteq A$ messbar mit $\mu(B) = 0$. Damit ist aber $f : A \to \overline{\mathbb{R}}$ messbar. Da gilt $\lim_{n\to\infty} f_n(x) = f(x)$ für alle $x \in X \setminus A$, folgt aus Satz 12.4.9 dass f auch messbar ist auf $X \setminus A$. □

Bemerkung 12.4.20. *Man beachte, dass die Vollständigkeit des Maßes μ im allgemeinen erforderlich ist.*

Definition 12.4.21. *Die Funktionen $f_n : X \to \overline{\mathbb{R}}$, $n \in \mathbb{N}$, und $f : X \to \overline{\mathbb{R}}$ seien messbar und fast überall endlich. Falls für jedes $\varepsilon > 0$ gilt*

$$\lim_{n\to\infty} \mu\big(X(|f_n - f| < \varepsilon)\big) = 0,$$

so heißt die Folge (f_n) **konvergent gegen** f **auf** X **dem Maße nach** *bzw.* **konvergent auf** X **bezüglich des Maßes** μ*, kurz $f_n \Rightarrow f$ auf X.*

Satz 12.4.22 (Lebesgue). *Sei $(X, \mathcal{M}, \mu)$ ein Maßraum mit $\mu(X) < \infty$ und μ ein vollständiges Maß. Desweiteren seien für alle $n \in \mathbb{N}$ die Funtionen $f_n : X \to \overline{\mathbb{R}}$ und $f : X \to \overline{\mathbb{R}}$ messbar und fast überall endlich. Falls $f_n(x) \to f(x)$ für fast alle $x \in X$ konvergiert, so gilt $f_n \Rightarrow f$, das heißt die Folge (f_n) konvergiert bezüglich des Maßes μ gegen f.*

Beweis. Ohne Beschränkung der Allgemeinheit sei $f \geq 0$. Nach Satz 12.4.19 ist f messbar. Wir setzen $A := \{x \in X : f(x) = \infty\}$ und $A_n := \{x \in X : f_n(x) = \infty\}$ für alle $n \in \mathbb{N}$. Ferner sei $B := \{x \in X : \lim_{n\to\infty} f_n(x) \neq f(x)\}$. Dann gilt für

$$Q := \bigcup_{n=1}^{\infty} A_n \cup A \cup B$$

offenbar $\mu(Q) = 0$. Für alle $n \in \mathbb{N}$ definieren wir nun

$$E_n(\varepsilon) := \{x \in X : |f_n(x) - f(x)| \geq \varepsilon\}, \qquad R_n(\varepsilon) := \bigcup_{k=n}^{\infty} E_k(\varepsilon),$$

und setzen

$$M(\varepsilon) := \bigcap_{n=1}^{\infty} R_n(\varepsilon).$$

Die Mengen M, $E_n(\varepsilon)$ und $R_n(\varepsilon)$, $n \in \mathbb{N}$, sind messbar und es gilt $R_1(\varepsilon) \supseteq R_2(\varepsilon) \supseteq \cdots$. Die Folge $\mu\big(R_n(\varepsilon)\big)$ ist somit monoton fallend bezüglich n und wegen $R_n(\varepsilon) \subseteq X$ ist R_n für alle $n \in \mathbb{N}$ beschränkt. Folglich gilt nach Satz 12.2.6

$$\lim_{n\to\infty} \mu\big(R_n(\varepsilon)\big) = \mu\big(M(\varepsilon)\big).$$

Wir zeigen nun $M \subseteq Q$. Wir betrachten hierzu ein beliebiges festes $x_0 \notin Q$. Nach Konstruktion gilt $f(x_0) < \infty$ und $f_n(x_0) < \infty$ für alle $n \in \mathbb{N}$ sowie $\lim_{n\to\infty} f_n(x_0) = f(x_0)$. Zu jedem $\varepsilon > 0$ existiert nun ein Index $n_0 \in \mathbb{N}$ mit $x_0 \in R_{n_0}(\varepsilon)$ und $x_0 \notin R_{n_0+1}(\varepsilon)$, dies bedeutet $x_0 \notin M(\varepsilon)$ bzw. $M(\varepsilon) \subseteq Q$. Da μ vollständig und $M(\varepsilon) \in \mathcal{M}$, folgt $\mu(M(\varepsilon)) = \mu(Q) = 0$. Damit haben wir

$$\lim_{n\to\infty} \mu\big(R_n(\varepsilon)\big) = \mu\big(M(\varepsilon)\big) = 0$$

und da $E_n \subseteq R_n$ ergibt sich

$$\lim_{n\to\infty} \mu\big(E_n(\varepsilon)\big) = 0.$$

□

Bemerkung 12.4.23. *1. Die Umkehrung der Aussage dieses Satzes gilt im allgemeinen nicht, wie folgendes Gegenbeispiel zeigt. Sei $X = [0, 1)$ versehen mit dem Lebesgue-Maß μ. Für $1 \leq i \leq k$ und $k \in \mathbb{N}$ setzen wir*

$$E_k^{(i)} = \left[\frac{i-1}{k}, \frac{i}{k}\right), \qquad f_k^{(i)}(x) := \chi_{E_k^{(i)}}(x),$$

Für $0 < \varepsilon < 1$ *gilt dann für alle* $1 \leq i \leq k$

$$\mu\big(X(|f_k^{(i)}| \geq \varepsilon)\big) = \mu\left(\left[\frac{i-1}{k}, \frac{i}{k}\right)\right) = \frac{1}{k} \overset{k\to\infty}{\longrightarrow} \infty.$$

und folglich $f_k^{(i)} \Rightarrow 0$ *für* $k \to \infty$. *Für ein beliebiges* $x_0 \in [0,1)$ *existiert zu jedem* $k \in \mathbb{N}$ *jedoch ein Index* i_k *mit* $x_0 \in E_k^{(i)}$. *Wir schließen* $f_k^{(i_k)}(x_0) = 1$ *für alle* i_k, *dies bedeutet* $f_k^{(i_k)}(x_0) \to 1$ *für* $k \to \infty$.

2. *Die Vorraussetzung, dass das Maß* $\mu(X) < \infty$ *ist, ist ebenfalls wesentlich. Denn sei* $X = [0,\infty)$ *versehen mit dem Lebesgue-Maß und* $f_n(x) := (x + \frac{1}{n})^2$ *und* $f(x) = x^2$, *dann gilt* $\lim_{n\to\infty} f_n(x) = f(x)$ *für alle* $x \in X$, *aber auch*

$$\mu\big(X(|f - f_n| \geq \varepsilon)\big) = \mu\left(\left\{x \in X : \frac{2}{n}x + \frac{1}{n^2} \geq \varepsilon\right\}\right) = \infty.$$

3. *Falls* μ *nicht vollständig ist, muss* f *als messbar vorausgesetzt werden.*

12.5 Das Lebesgue-Integral

Definition 12.5.1. *Sei* $(X, \mathcal{M}, \mu)$ *ein Maßraum und sei* $f : E \subseteq \mathcal{M} \to \overline{\mathbb{R}}$ *eine messbare nichtnegative Funktion. Desweiteren sei*

$$T_f = \left\{\varphi = \sum_{i=1}^{n} c_i \chi_{E_i} : n \in \mathbb{N},\ c_i \in \mathbb{R}_0^+,\ E_i \subseteq \mathcal{M} \text{ mit } \varphi \leq f\right\} \tag{12.9}$$

die Menge aller Treppenfunktionen, die punktweise kleiner oder gleich der Funktion f *sind. Dann heißt*

$$\int_E f d\mu = \sup_{\varphi \in T_f} \sum_{i=1}^{n} c_i \mu(E \cap E_i)$$

das **Lebesgue-Integral** *der Funktion* f *auf* E.

Falls f beschränkt ist, das heißt $-\infty < m := \inf_{x\in E} f(x)$ und $M := \sup_{x\in E} f(x) < \infty$, dann ist die folgende Definition äquivalent zur vorherigen.

Zur Zerlegung $Z = \{y_1, \ldots, y_n\}$ des Intervalls $[m, M]$ mit $m = y_1 < y_2 < \ldots < y_n = M$ definieren wir die Mengen

$$E_i := \{x \in X : y_i \leq f(x) < y_{i+1}\}, \qquad i = 1, \ldots, n-1,$$

und

$$E_n := \{x \in X : f(x) = y_n = M\}.$$

Dann gilt für das Lebesgue-Integral

$$\int_E f d\mu = \sup \sum_{i=1}^{n} y_i \mu(E_i),$$

wobei das Supremum über alle endlichen Zerlegungen zu nehmen ist.

Bemerkung 12.5.2. *Das Lebesgue-Integral einer Funktion $f : X \to \overline{\mathbb{R}}$ mit $f \geq 0$ kann den Wert $+\infty$ annehmen.*

Definition 12.5.3. *Sei $f : E \to \overline{\mathbb{R}}$ messbar, dann definieren wir*

$$\int_E f d\mu := \int_E f_+ d\mu - \int_E f_- d\mu,$$

wobei $f_\pm$ definiert ist gemäß (12.7)*, falls $\int_E f_+ d\mu < \infty$ oder $\int_E f_- d\mu < \infty$. Falls beide Integrale endlich sind, das heißt $\int_E f_+ d\mu < \infty$ und $\int_E f_- d\mu < \infty$, dann nennen wir f* **summierbar** *auf E bezüglich μ.*

Die Menge der summierbaren Funktionen auf E bezüglich μ bezeichnen wir mit $L(E, \mu)$.

Beispiel 12.5.4. *Sei $X = \mathbb{R}$ und $E = [0, 1]$ versehen mit dem Lebesgue-Maß μ. Für*

$$f(x) = \begin{cases} 1, & x \in \mathbb{Q}, \\ 0 & x \notin \mathbb{Q}, \end{cases}$$

gilt

$$\int_E f d\mu = 0.$$

Denn sei $Z = \{y_1, \ldots, y_n\}$ mit $0 = y_1 < y_2 < \ldots < y_n = 1$ ($n \geq 2$) eine Zerlegung von $[0, 1]$, dann ist

$$E_i = \{x \in [0, 1] : y_i \leq f(x) < y_{i+1}\} = \emptyset$$

für $i = 2, \ldots, n-1$ während

$$E_1 = \{x \in E : 0 \leq f(x) < y_1\} = [0, 1] \setminus \mathbb{Q}, \qquad E_n = \{x \in E : f(x) = 1\} = [0, 1] \cap \mathbb{Q}.$$

Insbesondere ist E_n abzählbar, das heißt $\mu(E_n) = 0$, und folglich

$$\mu(E_1) = \mu([0, 1]) - \mu(\mathbb{Q}) = 1.$$

Somit ergibt sich für jede Zerlegung

$$\sum_{i=1}^{n} y_i \mu(E_1) = y_1 \mu(E_1) + y_n \mu(E_n) = 0 \cdot 1 + 1 \cdot 0 = 0.$$

Proposition 12.5.5. *Sei $\varphi_n = \sum_{i=1}^n c_i \chi_{E_i} : E \to \mathbb{R}$ eine Treppenfunktion, dann gilt*

$$\int_E \varphi_n d\mu = \sum_{i=1}^n c_i \mu(E \cap E_i).$$

Satz 12.5.6 (Eigenschaften des Lebesgue-Integrals).

1. *Gilt $m \leq f(x) \leq M$ für alle $x \in E$, dann folgt*

$$m\mu(E) \leq \int_E f d\mu \leq M\mu(E).$$

2. *Seien $f, g \in L(E, \mu)$ und $c \in \mathbb{R}$, dann ist $f + g, cf \in L(E, \mu)$ und es gilt*

$$\int_E (f+g) d\mu = \int_E f d\mu + \int_E g d\mu, \qquad \int_E cf d\mu = c \int_E f d\mu.$$

3. *Sei $A \subseteq E$ messbar, dann folgt für $f \geq 0$*

$$\int_A f d\mu \leq \int_E f d\mu.$$

4. *Aus $0 \leq f(x) \leq g(x)$ für alle $x \in E$ folgt*

$$0 \leq \int_E f d\mu \leq \int_E g d\mu.$$

5. *Ist $f \geq 0$, so gilt $\int_E f d\mu = 0$ genau dann, wenn $f(x) = 0$ fast überall auf E.*

Beweis. Wir beweisen nur die fünfte Aussage. Die leichten Beweise der anderen Aussagen verbleiben dem Leser als Übung.

Wir definieren $F := \{x \in E : f(x) > 0\}$ und für $n \in \mathbb{N}$

$$E_n := \left\{ x \in E : f(x) > \frac{1}{n} \right\}.$$

Dann gilt $E_n \subseteq E_{n+1}$ für alle $n \in \mathbb{N}$ und $F = \bigcup_{n=1^\infty} E_n$. Wegen der Monotonie von μ folgen hieraus die Beziehungen $\mu(E_n) \leq \mu(E_{n+1}) \leq F$ für alle $n \in \mathbb{N}$ und $\lim_{n\to\infty} \mu(E_n) = \mu(F)$. Nehmen wir $\mu(F) > 0$ an, so existiert zu jedem $\varepsilon > 0$ ein $n \in \mathbb{N}$ mit $\mu(E_n) > \varepsilon$. Aufgrund von $f(x) > 1/n$ für alle $x \in E_n$ folgt hieraus

$$0 = \int_E f d\mu \geq \int_{E_n} f d\mu \geq \frac{1}{n}\mu(E_n) \geq \frac{1}{n}\varepsilon > 0.$$

Dieser Widerspruch bedeutet, dass $\mu(E_0) = 0$ gelten muss. Die Umkehrung der Aussage ist trivial, so dass die Äquivalenz bewiesen ist. □

Satz 12.5.7. *Sei $(X, \mathcal{M}, \mu)$ ein Maßraum.*

1. *Ist $f : X \to \overline{\mathbb{R}}$ messbar mit $f \geq 0$ auf X, dann ist die Mengenfunktion $\nu : \mathcal{M} \to \overline{\mathbb{R}}$ definiert durch*
$$\nu(A) := \int_A f d\mu, \qquad A \in \mathcal{M},$$
volladditiv.

2. *Falls $f \in L(X, \mu)$, dann ist $\nu(A)$ ebenfalls volladditiv.*

Beweis. 1. Nach Voraussetzung ist $\nu(A) \geq 0$ für alle $A \in \mathcal{M}$. Sei $A \in \mathcal{M}$ beliebig und seien A_n, $n \in \mathbb{N}$, paarweise disjunkt mit $A = \bigcup_{n=1}^{\infty} A_n$. Für alle $\varphi \in T_f$ gemäß (12.9) ist

$$\begin{aligned}\int_A \varphi d\mu &= \int_A \Big(\sum_{i=1}^{n} c_i \chi_{E_i}(x)\Big) d\mu = \sum_{i=1}^{n} c_i \mu(A \cap E_i) = \sum_{i=1}^{n} c_i \mu\Big(\bigcup_{m=1}^{\infty} (A_m \cap E_i)\Big) \\ &= \sum_{m=1}^{\infty} \sum_{i=1}^{n} c_i \mu(A_m \cap E_i) = \sum_{m=1}^{\infty} \int_{A_m} \varphi d\mu \leq \sum_{m=1}^{\infty} \int_{A_m} f d\mu. \end{aligned} \tag{12.10}$$

Nehmen wir das Supremum über alle $\varphi \in T_f$, so erhalten wir

$$\nu(A) = \int_A f d\mu \leq \sum_{m=1}^{\infty} \int_{A_m} f d\mu \leq \sum_{m=1}^{\infty} \nu(A_m),$$

dies bedeutet dass ν subvolladditiv ist.

Wir zeigen nun, dass aus $A_1 \cap A_2 = \emptyset$ folgt

$$\nu(A_1 \cup A_2) = \int_{A_1 \cup A_2} f d\mu = \int_{A_1} f d\mu + \int_{A_2} f d\mu = \nu(A_1) + \nu(A_2).$$

Die obige Überlegung impliziert $\nu(A_1 \cup A_2) \leq \nu(A_1) + \nu(A_2)$. Umgekehrt existieren aber auch zu jedem $\varepsilon > 0$ zwei Treppenfunktionen $\varphi_1, \varphi_2 \in T_f$ mit

$$\int_{A_1} f d\mu = \int_{A_1} \varphi_1 d\mu + \varepsilon, \qquad \int_{A_2} f d\mu = \int_{A_2} \varphi_2 d\mu + \varepsilon.$$

Nun sei $\varphi(x) = \max\{\varphi_1(x), \varphi_2(x)\}$, dann ist $\varphi \in T_f$ messbar. Nach (12.10) gilt

$$\begin{aligned}\nu(A_1) + \nu(A_2) &\leq \int_{A_1} \varphi d\mu + \int_{A_2} \varphi d\mu + 2\varepsilon = \int_{A_1 \cup A_2} \varphi d\mu + 2\varepsilon \\ &\leq \int_{A_1 \cup A_2} f d\mu + 2\varepsilon = \nu(A_1 \cup A_2) + 2\varepsilon.\end{aligned}$$

Für $\varepsilon \to 0$ bedeutet dies

$$\nu(A_1) + \nu(A_2) \leq \nu(A_1 \cup A_2)$$

beziehungsweise zusammengefasst

$$\nu(A_1 \cup A_2) = \nu(A_1) + \nu(A_2)$$

Aufgrund der siebten Aussage von Satz 12.2.4 folgt hieraus die Behauptung.

2. Wir zerlegen $f = f_+ - f_-$, dann gilt wegen $f \in L(X, \mu)$ $\int_E f_\pm d\mu < \infty$ für alle $E \in \mathcal{M}$. Somit folgt die Behauptung aus der vorherigen Aussage. □

Korollar 12.5.8. *Seien $A, B \in \mathcal{M}$ mit $B \subseteq A$ und $\mu(A \setminus B) = 0$ und $f : A \to \overline{\mathbb{R}}$ eine messbaren Funktion. Gilt für eines der Integrale $\int_A f d\mu < \infty$ oder $\int_B f d\mu < \infty$, so gilt dies auch für das zweite Integral und es folgt*

$$\int_A f d\mu = \int_B f d\mu < \infty.$$

Beweis. Sei ohne Einschränkung der Allgemeinheit $f \geq 0$, dann folgt aus $\mu(A \setminus B) = 0$ die Behauptung

$$\int_A f d\mu = \int_B f d\mu + \int_{A \setminus B} f d\mu = \int_B f d\mu.$$

□

Definition 12.5.9. *Seien $E \in \mathcal{M}$ und $f, g : E \to \overline{\mathbb{R}}$ messbar, dann schreiben wir $f \overset{E}{\sim} g$, falls $\mu(\{x \in E : f(x) \neq g(x)\}) = 0$, das heißt f und g stimmen fast überall auf E überein.*

Bemerkung 12.5.10.

1. *Die Relation "$\overset{E}{\sim}$" ist eine Äquivalentrelation*

2. *Falls $f, g \in L(X, \mu)$ und $f \overset{E}{\sim} g$, so gilt*

$$\int_E f d\mu = \int_E g d\mu.$$

3. *Falls $f \in L(X, \mu)$, dann ist f fast überall endlich.*

Satz 12.5.11. *Seien $g, h : E \to \overline{\mathbb{R}}$ messbar, $E \in \mathcal{M}$ mit $|g(x)| \leq h(x)$ für alle $x \in E$. Falls $h \in L(E, \mu)$, so ist auch $g \in L(E, \mu)$ und es gilt*

$$\int_E g d\mu \leq \int_E h d\mu.$$

Beweis. Sei ohne Einschränkung der Allgemeinheit $g \geq 0$, dann folgt die Behauptung sofort aus $T_g \subseteq T_h$. □

Satz 12.5.12. *Es gilt $f \in L(E, \mu)$ genau dann, wenn $|f| \in L(E, \mu)$.*

Beweis. Wir zerlegen $f = f_+ - f_-$, dann gilt

$$|f|(x) = f_+(x) + f_-(x) = \begin{cases} f_+(x), & x \in A, \\ f_-(x), & x \in E \setminus A, \end{cases}$$

mit $A := \{x \in X : f(x) \geq 0\}$. Aufgrund von Satz 12.5.7 folgt, dass

$$\int_E |f| d\mu = \int_A f_+ d\mu + \int_{E \setminus A} f_- d\mu = \int_E f_+ d\mu + \int_E f_- d\mu$$

genau dann definiert und endlich ist, falls $f_\pm \in L(X, \mu)$ bzw. $f \in L(X, \mu)$. □

12.6 Grenzwertsätze für Integrale

Satz 12.6.1 (Satz von Bepo Levi über monotone Konvergenz). *Sei $E \in \mathcal{M}$ und für alle $n \in \mathbb{N}$ seien $f_n : E \to \overline{\mathbb{R}}$ messbare Funktionen mit $0 \leq f_1(x) \leq f_2(x) \leq \ldots$ für alle $x \in E$. Dann gilt*

$$\lim_{n\to\infty} \int_E f_n d\mu = \int_E \lim_{n\to\infty} f_n d\mu.$$

Beweis. Sei ohne Einschränkung der Allgemeinheit $\int_E f_n d\mu \leq C < \infty$ für alle $n \in \mathbb{N}$. Die Funktion $f : E \to \overline{\mathbb{R}}$ definiert vermittels $f(x) := \lim_{n\to\infty} f_n(x)$ für alle $x \in E$ erfüllt $f \geq 0$ und ist nach Satz 12.4.9 messbar. Setzen wir $\alpha_n := \int_E f_n d\mu$ für alle $n \in \mathbb{N}$, so gilt $0 \leq \alpha_1 \leq \alpha_2 \leq \ldots \leq C\mu(E)$. Demnach existiert $\alpha := \lim_{n\to\infty} \alpha_n$ und ist endlich. Wegen $f_n(x) \leq f(x)$ für alle $x \in E$ folgt

$$\alpha \leq \int_E f d\mu.$$

Um die Ungleichung andersherum zu zeigen, definieren wir für beliebige aber feste $c \in (0, 1)$ und $\varphi \in T_f$ die Mengen

$$E_n = E_n(\varphi, T_f) := \{x \in E : f_n(x) > c\varphi(x)\}, \qquad n \in \mathbb{N}.$$

Es gilt $E_n \in \mathcal{M}$ und $E_n \subseteq E_{n+1} \subseteq E$ für alle $n \in \mathbb{N}$. Wegen $f \geq \varphi$ auf E, existiert zu jedem $x \in E$ ein $n \in \mathbb{N}$ mit $f_n(x) \geq c\varphi(x)$. Hieraus folgt

$$E = \bigcup_{n=1}^{\infty} E_n$$

und wir schließen

$$\int_E f_n d\mu \geq \int_{E_n} f_n d\mu \geq c \int_{E_n} \varphi d\mu$$

für alle $\varphi \in T_f$ und $c \in (0,1)$. Für $c \to 1$ ergibt sich

$$\int_E f_n d\mu \geq \int_{E_n} f_n d\mu \geq \int_{E_n} \varphi d\mu.$$

Bilden wir nun das Supremum über alle $\varphi \in T_f$, so folgt

$$\lim_{n\to\infty} \int_E f_n d\mu \geq \int_E f d\mu.$$

□

Satz 12.6.2 (Satz von Lebesgue über monotone Konvergenz). *Sei $E \in \mathcal{M}$ und für alle $n \in \mathbb{N}$ seien $f_n : E \to \overline{\mathbb{R}}$ messbare Funktionen mit $f_1(x) \leq f_2(x) \leq \ldots$ für alle $x \in E$. Gibt es ein $m \in \mathbb{N}$ mit $f_m \in L(E,\mu)$, dann ist $f(x) := \lim_{n\to\infty} f_n(x) \in L(E,\mu)$ und es gilt*

$$\int_E f d\mu = \lim_{n\to\infty} \int_E f_n d\mu.$$

Beweis. Ohne Beschränkung der Allgemeinheit sei $m = 1$, das heißt $f_1 \in L(E,\mu)$, dann ist f_1 fast überall endlich auf E. Es sei f_1 sogar endlich angenommen, da wir zu jeder fast überall endlichen Funktion $g \in L(E,\mu)$ eine überall endliche Funktion $\widetilde{g} \in L(E,\mu)$ finden können mit $\widetilde{g} \overset{E}{\sim} g$. Wir definieren die Funktionen $g_n(x) := f_n(x) - f_1(x)$ für alle $x \in E$. Für alle $n \in \mathbb{N}$ ist g_n messbar mit $0 \leq g_n(x) \leq g_{n+1}(x)$, $x \in E$. Sei $g(x) := \lim_{n\to\infty} g_n(x)$, dann folgt wegen $g(x) = f(x) - f_1(x)$ aus dem Satz von Beppo Levi

$$\begin{aligned}
\int_E \big(f - f_1\big) d\mu &= \int_E g d\mu \\
&= \int_E \lim_{n\to\infty} g_n d\mu \\
&= \int_E \lim_{n\to\infty} (f_n - f_1) d\mu \\
&= \lim_{n\to\infty} \left(\int_E f_n d\mu - \int_E f_1 d\mu \right),
\end{aligned}$$

das heißt die Behauptung. □

Korollar 12.6.3. *Sei $E \in \mathcal{M}$ und für alle $n \in \mathbb{N}$ seien $f_n : E \to \overline{\mathbb{R}}$ messbare Funktionen mit $f_1(x) \leq f_2(x) \leq \ldots$ für alle $x \in E$. Falls $f_n \in L(E,\mu)$ für alle $n \in \mathbb{N}$, so ist $\int_E f_n d\mu < C$ für alle $n \in \mathbb{N}$ und es gilt $f \in L(E,\mu)$.*

Satz 12.6.4 (Lemma von Fatou). *Sei $E \in \mathcal{M}$ und für alle $n \in \mathbb{N}$ seien $f_n : E \to \overline{\mathbb{R}}$ messbare Funktionen.*

1. *Falls $f_n \geq 0$ für alle $n \in \mathbb{N}$ gilt*

$$\int_E \liminf_{n\to\infty} f_n d\mu \leq \liminf_{n\to\infty} \int_E f_n d\mu.$$

2. *falls* $f_n \leq 0$ *für alle* $n \in \mathbb{N}$ *gilt*

$$\int_E \limsup_{n\to\infty} f_n d\mu \geq \limsup_{n\to\infty} \int_E f_n d\mu.$$

Beweis. Wir zeigen nur die erste Aussage da der Beweis der zweiten vollkommen analog geführt wird. Definieren wir $g_n(x) := \inf_{k\geq n} f_k(x)$ für alle $n \in \mathbb{N}$, dann sind die Funktionen g_n messbar und es gilt $0 \leq g_1 \leq g_2 \leq \ldots$. Ferner schließen wir $g_n \leq f_n$ für alle $n \in \mathbb{N}$ und somit

$$\int_E g_n d\mu \leq \int_E f_n d\mu$$

beziehungsweise für $n \to \infty$

$$\lim_{n\to\infty} \int_E g_n d\mu \leq \liminf_{n\to\infty} \int_E f_n d\mu.$$

Wenden wir nun auf die Folge (g_n) den Satz von Beppo Levi an, so ergibt sich wegen

$$\liminf_{n\to\infty} f_n(x) = \lim_{n\to\infty} g_n(x)$$

die Behauptung

$$\int_E \liminf_{n\to\infty} f_n d\mu = \int_E \lim_{n\to\infty} g_n d\mu = \lim_{n\to\infty} \int_E g_n d\mu \leq \liminf_{n\to\infty} \int_E f_n d\mu.$$

□

Satz 12.6.5 (Satz von Lebesgue über dominante Konvergenz). *Seien* (f_n) *eine Folge messbarer Funktionen* $f_n : E \to \overline{\mathbb{R}}$ *mit* $f_n \to f$ *auf* E. *Falls eine Funktion* $g \in L(E, \mu)$ *existiert, so dass für alle* $n \in \mathbb{N}$ *gilt* $|f_n(x)| \leq g(x)$ *für alle* $x \in E$, *dann gilt* $f \in L(E, \mu)$ *mit*

$$\int_E f d\mu \leq \int_E |f| d\mu \leq \int_E g d\mu$$

und

$$\int_E f d\mu = \lim_{n\to\infty} \int_E f_n d\mu.$$

Beweis. Offenbar gilt wegen $0 \leq |f_n| \leq g$ für alle $n \in \mathbb{N}$ auch $0 \leq |f| \leq g$, so dass aus Satz 12.5.11 folgt $f, f_n \in L(E, \mu)$. Nach dem Lemma von Fatou gelten

$$\limsup_{n\to\infty} \int_E (f_n - g) d\mu \leq \int_E \limsup_{n\to\infty} (f_n - g) d\mu = \int_E (f - g) d\mu$$

und umgekehrt

$$\int_E (f + g) d\mu = \int_E \liminf_{n\to\infty} (f_n + g) d\mu \leq \liminf_{n\to\infty} \int_E (f_n + g) d\mu = \int_E g d\mu + \liminf_{n\to\infty} \int_E f_n d\mu.$$

Zusammengefasst folgt hieraus

$$\limsup_{n\to\infty} \int_E f_n d\mu \leq \int_E f d\mu \leq \liminf_{n\to\infty} \int_E f_n d\mu$$

bzw.

$$\int_E f_n d\mu = \lim_{n\to\infty} \int_E f_n d\mu.$$

□

12.7 Vergleich des Lebesgue-Integrals mit dem Riemann-Integral

Sei $(\mathbb{R}, \mathcal{M}, \mu)$ ein Lebesgue'scher Maßraum und $E = [a, b]$, $a < b$, endlich. Für $f : [a, b] \to \mathbb{R}$ bezeichne $(R) \int_a^b f dx$ das Riemann-Integral und $(L) \int_a^b f dx$ das Lebesgue-Integral von f.

Satz 12.7.1. *1. Wenn f auf $[a, b]$ Riemann-integrierbar ist, so ist $f \in L([a, b], \mu)$ und*

$$(L) \int_a^b f dx = (R) \int_a^b f dx.$$

2. Sei f beschränkt, dann ist f genau dann Riemann-integrierbar, wenn f fast überall stetig auf $[a, b]$ ist.

Beweis. 1. Sei $(Z_k)_{k=1}^{\infty}$ eine Folge von Zerlegungen des Intervalls $[a, b]$ mit $Z_{k+1} \supseteq Z_k$ und $\Delta(Z_k) \to 0$ für $k \to \infty$. Da f Riemann-integrierbar ist, ist f beschränkt auf $[a, b]$. Wir definieren für alle $k \in \mathbb{N}$ die Treppenfunktionen U_k und L_k vermittels $U_k(a) = L_k(a) = f(a)$ und

$$\left.\begin{aligned} U_k(x) &= M_i = \sup_{x\in[x_{i-1},x_i]} f(x) \\ L_k(x) &= m_i = \inf_{x\in[x_{i-1},x_i]} f(x) \end{aligned}\right\} \quad \text{für } x \in (x_{i-1}, x_i].$$

Es gilt dann

$$\begin{aligned} (L) \int_a^b U_k dx &= \sum_{i=1}^{k} M_i(x_i - x_{i-1}) = S(Z_k, f), \\ (L) \int_a^b L_k dx &= \sum_{i=1}^{k} m_i(x_i - x_{i-1}) = s(P_k, f). \end{aligned} \tag{12.11}$$

Wegen $Z_{k+1} \supseteq Z_k$ erhalten wir

$$U_1(x) \geq U_2(x) \geq \ldots \geq f(x) \geq \ldots \geq L_2(x) \geq L_1(x)$$

für alle $x \in [a, b]$. Wir setzen

$$U(x) := \lim_{k\to\infty} U_k(x), \qquad L(x) := \lim_{k\to\infty} L_k(x).$$

Ist f Riemann-integrierbar, so folgt aus $\Delta(Z_k) \to 0$, dass

$$\lim_{k\to\infty} s(Z_k, f) = (R)\int_a^b f dx = \lim_{k\to\infty} S(Z_k, f). \tag{12.12}$$

Da U_k und L_k messbare Treppenfunktionen sind, folgt aus dem Satz von Lebesgue über dominante Konvergenz dass

$$\lim_{k\to\infty} (L)\int_a^b U_k dx = (L)\int_a^b U dx, \qquad \lim_{k\to\infty} (L)\int_a^b L_k dx = (L)\int_a^b L dx. \tag{12.13}$$

Hieraus folgt mit (12.11) und (12.12)

$$(L)\int_a^b L dx = (R)\int_a^b f dx = (L)\int_a^b U dx.$$

Weiter ergibt sich wegen $L(x) \le f(x) \le U(x)$ und $(L)\int_a^b (U-L)dx = 0$ dass $U(x) = f(x) = L(x)$ fast überall auf $[a,b]$, das heißt $f \in L([a,b],\mu)$ und

$$(L)\int_a^b f dx = (R)\int_a^b f dx.$$

2. Sei $Z = \bigcup_{k=1}^\infty Z_k$, dann gilt $\mu(Z) = 0$, da Z abzählbar ist. Wir zeigen zuerst, dass f genau dann stetig im Punkt $x_0 \in [a,b] \setminus Z$ ist, falls $U(x_0) = L(x_0)$ gilt.

Sei f also zunächst stetig in x_0 angenommen, dann existiert für alle $\varepsilon > 0$ ein $\delta > 0$ mit $|f(x) - f(x_0)| < \varepsilon/2$ für alle x mit $|x - x_0| < \delta$. Wegen $\Delta(Z_k) \to 0$ existiert ein $K \in \mathbb{N}$, so dass $\Delta(Z_k) < \delta$ für alle $k \ge K$. Folglich gilt

$$U_k(x_0) - L_k(x_0) = M_i - m_i = \big(M_i - f(x_0)\big) + \big(f(x_0) - m_i\big) < \varepsilon/2 + \varepsilon/2 = \varepsilon,$$

mit anderen Worten $U(x_0) = L(x_0)$.

Sei umgekehrt nun $L(x_0) = U(x_0)$ und $\varepsilon > 0$ beliebig. Es gilt dann $L(x_0) = f(x_0) = U(x_0)$ und daher existiert ein $K \in \mathbb{N}$ mit

$$U_K(x_0) - f(x_0) \le \varepsilon/2, \qquad f(x_0) - L_K(x_0) \le \varepsilon/2,$$

Setzen wir $\delta = \min\{|x_0 - x| : x \in Z_K\}$, dann gilt für alle $|x - x_0| < \delta$

$$L_K(x_0) \le f(x) \le U_K(x_0).$$

Daraus folgt $|f(x) - f(x_0)| < \varepsilon$ und folglich ist f stetig in x_0.

Wenn f Riemann-integrierbar ist, dann folgt aus dem ersten Teil des Satzes, dass $U(x) = L(x)$ fast überall auf $[a,b]$ gilt, das heißt f ist fast überall stetig auf $[a,b]$. Ist umgekehrt f fast überall stetig auf $[a,b]$, dann gilt $U(x) = f(x) = L(x)$ fast überall auf $[a,b]$ und folglich

$$(L)\int_a^b U dx = (L)\int_a^b L dx,$$

woraus sich nach (12.11) und (12.13) $S(P_k, f) - s(P_k, f) \overset{k\to\infty}{\longrightarrow} 0$ ergibt, das heißt f ist Riemann-integrierbar. $\square$

Beispiel 12.7.2. *1. Sei* $[a, b] = [0, 1]$ *und*

$$f(x) = \begin{cases} x^2, & x \text{ irrational}, \\ \sin x, & x \text{ rational}. \end{cases}$$

Dann gilt

$$(L)\int_0^1 f dx = (L)\int_{[0,1]\cap\mathbb{Q}} f dx + \int_{[0,1]\setminus\mathbb{Q}} f dx = (R)\int_0^1 x^2 dx = 1/3.$$

2. Sei die Funktion $f : \mathbb{R}^+ \to \mathbb{R}$ *definiert durch* $f(x) = \sin x/x$. *Es gilt*

$$\int_0^\infty \frac{\sin x}{x} dx = \pi/2,$$

aber wegen $(L)\int_0^\infty |\sin x/x| dx = +\infty$ *existiert das Integral* $(L)\int_0^\infty \sin x/x dx$ *nicht.*

3. Sei $f \geq 0$ *auf* $[0, \infty)$ *mit* $f \in R([0, A])$ *für alle* $A > 0$. *Setzen wir*

$$f_n(x) := \begin{cases} f(x), & 0 \leq x \leq n, \\ 0, & n < x, \end{cases}$$

dann gilt $0 \leq f_1 \leq f_2 \leq \ldots$ *und*

$$(L)\int_{[0,\infty)} f_n dx = (L)\int_0^n f dx = (R)\int_0^n f dx < \infty.$$

Mit der Tatsache, dass $f_n \to f$ *für* $n \to \infty$ *gilt, folgt aus dem Satz von Beppo Levi*

$$\begin{aligned} (L)\int_{[0,\infty)} f dx &= (L)\int_{[0,\infty)} \lim_{n\to\infty} f_n dx \\ &= \lim_{n\to\infty} (L)\int_{[0,\infty)} f_n dx \\ &= \lim_{n\to\infty} (R)\int_0^n f dx \\ &= \lim_0^\infty f dx. \end{aligned}$$

das heißt $f \in L([0, \infty), \mu)$.

12.8 L^p-Räume

Definition 12.8.1. *Sei $f \in L(X, \mu)$. Wir bezeichnen mit $[f] := \{g : g \overset{X}{\sim} f\}$ die Restklassen zu f bezüglich der Äquivalenzrelation "$\overset{E}{\sim}$". Wir definieren für $1 \leq p \leq \infty$ vermittels*

$$\|f\|_p := \left(\int_X |f|^p(x) d\mu \right)^{1/p}$$

die Funktionenräume

$$L^p(X, \mu) := \{[f] : \|f\|_p < \infty\}.$$

Bemerkung 12.8.2. *1. Es gilt $[g] = [f]$ genau dann, wenn $f = g$ fast überall auf X, d.h. $\mu(\{x \in X : f(x) \neq g(x)\}) = 0$.*

2. Das Supremum aller Häufungspunkte der Menge $\{f(x) \in \overline{\mathbb{R}} : x \in X\}$ bezeichnen wir mit $\lim \operatorname{ess}_{x \in X} f(x)$. Es wird das **essentielle Supremum** *genannt. Mit diesem gilt dann die Beziehung*

$$\|f\|_\infty = \lim_{x \in X} \operatorname{ess} |f(x)| \mu(X).$$

Lemma 12.8.3. *Seien $a, b > 0$, dann gilt für $1 < p, q < \infty$ mit*

$$\frac{1}{p} + \frac{1}{q} = 1$$

die Ungleichung

$$a^{1/p} \cdot b^{1/q} \leq \frac{a}{p} + \frac{b}{q}.$$

Beweis. Sei ohne Beschränkung der Allgemeinheit $t := a/b \geq 1$. Wir können die obige Ungleichung vermittels Division durch b umformen in

$$t^{1/p} \leq \frac{t}{p} + \frac{1}{q}.$$

Setzen wir

$$g(t) := t^{1/p}, \qquad h(t) := \frac{t}{p} + \frac{1}{q},$$

so müssen wir zeigen $g(t) \leq h(t)$ für alle $t \geq 1$. Für $t = 1$ folgt $g(1) = 1 = h(1)$. Weiter gilt für alle $t \geq 1$ dass $t^{-1/q} \leq 1$ und somit

$$g'(t) = \frac{1}{p} \cdot t^{1/p-1} = \frac{1}{p} t^{-1/q} \leq h'(t) = \frac{1}{p}$$

Hieraus folgt in der Tat $g(t) \leq h(t)$ für alle $t \geq 1$, das heißt die Behauptung □

Satz 12.8.4. *Es sei $1 \leq p, q \leq \infty$ mit $1/p + 1/q = 1$, wobei $p = 1$ und $q = \infty$ ebenfalls zugelassen sei.*

1. *Seien $f, g : X \to \overline{\mathbb{R}}$ messbar mit $\|f\|_p, \|g\|_q < \infty$, dann gilt die Hölder-Ungleichung*

$$\int_X |f \cdot g| d\mu \leq \|f\|_p \cdot \|g\|_q.$$

Insbesondere folgt $f \cdot g \in L(X, \mu)$.

2. *Seien $\|f\|_p, \|g\|_p \leq \infty$, dann ist $|f+g| \in L^p(X, \mu)$ und es gilt die Minkowski-Ungleichung*

$$\|f + g\|_p \leq \|f\|_p + \|g\|_q.$$

Beweis. 1. Zuerst betrachten wir den Fall $1 < p, q < \infty$. Ohne Beschränkung der Allgemeinheit seien $\|f\|_p, \|g\|_q \neq 0$, denn anderenfalls ist $f(x) \cdot g(x) = 0$ fast überall auf X. Wir setzen

$$a := \big(|f(x)|/\|f\|_p\big)^p, \qquad b := \big(|g(x)|^q/\|g\|_q\big)^q,$$

dann folgt aus dem vorhergehenden Lemma

$$0 \leq a^{1/p} \cdot b^{1/q} = \frac{|f(x)|}{\|f\|_p} \cdot \frac{|g(x)|}{\|g\|_q} \leq \frac{1}{p}\left(\frac{|f(x)|}{\|f\|_p}\right)^p + \frac{1}{q}\left(\frac{|g(x)|}{\|g\|_q}\right)^q$$

für alle $x \in X$. Wir integrieren nun über X

$$\begin{aligned}
\frac{1}{\|f\|_p \cdot \|g\|_q} \int_X |f \cdot g| d\mu &\leq \frac{1}{p\|f\|_p^p} \int_X |f|^p d\mu + \frac{1}{q\|q\|_q^q} \int_X |g|^q d\mu \\
&\leq \frac{1}{p\|f\|_p^p} \|f\|_p^p + \frac{1}{q\|q\|_q^q} \|g\|_q^q \\
&= \frac{1}{p} + \frac{1}{q} = 1.
\end{aligned}$$

Hieraus folgt die Hölder-Ungleichung.

Im Falle $p = 1$ und $q = \infty$ ergibt sich die Hölder-Ungleichung aus der Abschätzung

$$\int_X |f \cdot g| d\mu \leq \lim_{x \in X} \operatorname{ess} |f(x)| \int_X |g| d\mu = \|f\|_\infty \|g\|_1.$$

2. Es sei $p > 1$, dann gilt für alle $x \in X$ die Abschätzung

$$|f(x) + g(x)|^p \leq |f(x)| \cdot |(f + g)(x)|^{p-1} + |g(x)| \cdot |(f + g)(x)|^{p-1}.$$

Integration über X liefert

$$\|f + g\|_p^p = \int_X |f + g|^p d\mu \leq \int_X |f| \cdot |f + g|^{p-1} d\mu + \int_X |g| \cdot |f + g|^{p-1} d\mu$$

Hieraus ergibt sich für $q = p/(p-1)$ mit Hilfe der Hölder-Ungleichung

$$\begin{aligned}\|f+g\|_p^p &\leq \|f\|_p \cdot \|(f+g)^{p-1}\|_q + \|g\|_p \cdot \|(f+g)^{p-1}\|_q \\ &= (\|f\|_p + \|g\|_p)\left(\int_X |f+g|^{(p-1)q} d\mu\right)^{1/q} \\ &= (\|f\|_p + \|g\|_p)\left(\int_X |f+g|^p d\mu\right)^{1/q} \\ &= (\|f\|_p + \|g\|_p)\|f+g\|_p^{p/q} \\ &= (\|f\|_p + \|g\|_p)\|f+g\|_p^{p-1},\end{aligned}$$

das heißt $\|f+g\|_p \leq \|f\|_p + \|g\|_p$. Der Fall $p = 1$ und $q = \infty$ verbleibt dem Leser als Übung. □

Satz 12.8.5. *Für $1 \leq p \leq \infty$ bildet die Menge $L^p(X, \mu)$ einen linearen Raum, d.h. einen Vektorraum, mit der Norm $\|\cdot\|_p$.*

Beweis. Seien $[f], [g] \in L^p(X, \mu)$, das heißt $\|f\|_p, \|g\|_p < \infty$, dann folgt mit der Minkowski-Ingleichung $\|f+g\|_p \leq \|f\|_p + \|g\|_p < \infty$, das heißt $[f+g] \in L^p(X, \mu)$.

Für $[f] \in alleL^p(X, \mu)$, das heißt $\|f\|_p < \infty$, und für $\alpha \in \mathbb{R}$ gilt

$$\|\alpha \cdot f\|_p^p = \int_X |\alpha \cdot f|^p d\mu = |\alpha|^p \|f\|_p^p, \tag{12.14}$$

dies bedeutet $\|\alpha \cdot f\|_p = |\alpha| \|f\|_p < \infty$ und somit $[\alpha \cdot f] \in L^p(X, \mu)$. Damit ist $L^p(X, \mu)$ ein Vektorraum.

Wir müssen noch zeigen, dass $\|\cdot\|_p$ eine Norm definiert. Offensichtlich gilt $\|f\|_p \geq 0$ für alle $f \in L^p(X, \mu)$, wobei aus $\|f\|_p = 0$ folgt $\int_X |f|^p d\mu = 0$, das heißt $|f(x)| = 0$ bzw. $f(x) = 0$ fast überall auf X und folglich $f \overset{X}{\sim} 0$. Die Homogenität von Homogenität $\|\cdot\|_p$ haben wir in (12.14) gezeigt. Schließlich entspricht die Dreiecksungleichung der Minkowski-Ungleichung. □

Satz 12.8.6. *Die normierten Räume $L^p(X, \mu)$, $1 \leq p < \infty$, sind vollständig und folglich Banachräume.*

Beweis. Sei $f_n \in L^p(X, \mu)$, $n \in \mathbb{N}$, eine Cauchy-Folge in $L^p(X, \mu)$, das heißt zu jedem $n \in \mathbb{N}$ existiert ein $N \in \mathbb{N}$ mit

$$\|f_k - f_m\|_p^p \leq \frac{1}{2^{2n}} =: \varepsilon$$

für alle $k, m \geq N$. Wir definieren via $g_n := f_N$ die Teilfolge $(g_n)_{n\in\mathbb{N}}$ von $(f_n)_{n\in\mathbb{N}}$. Die Behauptung folgt, wenn wir zeigen $g := \lim_{n\to\infty} g_n \in L^p(X, \mu)$. Für $Y_n = \{x \in X : |g_{n+1}(x) - g_n(x)|^p \geq 2^{-n}\}$ gilt

$$\begin{aligned}\frac{\mu(Y_n)}{2^n} &= \int_{Y_n} \frac{1}{2^n} d\mu \leq \int_{Y_n} |g_{n+1}(x) - g_n(x)|^p d\mu \leq \int_X |g_{n+1}(x) - g_n(x)|^p d\mu \\ &= \|g_{n+1} - g_n\|_p^p \leq \frac{1}{2^{2n}},\end{aligned}$$

das heißt $\mu(Y_n) \leq 2^n$. Setzen wir nun $Z_n = \bigcup_{k=n}^{\infty} Y_k$, so folgt $Z_1 \supseteq Z_2 \supseteq \ldots$ und weiter

$$\mu(Z_n) = \mu\left(\bigcup_{k=n}^{\infty} Y_k\right) \leq \sum_{k=n}^{\infty} \mu(Y_k) = \sum_{k=n}^{\infty} 2^k \leq 2^{n-1}.$$

Somit ergibt sich für $Z = \bigcap_{n=1}^{\infty} Z_n$ dass

$$\mu(Z) = \lim_{n\to\infty} \mu(Z_n) = 0,$$

dies bedeutet, Z ist eine Menge vom Maß Null. Wir setzen nun

$$g(x) := \begin{cases} g_1(x) + \sum_{n=1}^{\infty} \big(g_{n+1}(x) - g_n(x)\big), & x \in X \setminus Z, \\ 0, & x \in Z. \end{cases}$$

Aufgrund der Tatsache, dass für alle $x \in X \setminus Z_N$ gilt

$$|g_{k+1}(x) - g_k(x)|^p \leq \frac{1}{2^k}, \qquad k \geq N,$$

folgt aus der Identität

$$g_n(x) = g_1(x) + \sum_{n=1}^{m} \big(g_{n+1}(x) - g_n(x)\big),$$

dass

$$\sup_{x\in X\setminus Z_n} |g(x) - g_n(x)|^p \leq \sum_{n=m}^{\infty} |g_{n+1} - g_n|^p \leq \sum_{n=m}^{\infty} \frac{1}{2^n} = \frac{1}{2^{m-1}} \overset{m\to\infty}{\longrightarrow} 0.$$

Folglich konvergiert die Folge $(g_n)_{n\in\mathbb{N}}$ gleichmäßig und absolut gegen g auf $X \setminus Z_n$, das heißt $g_n \to g$ bezüglich des Maßes μ und auch fast überall. Folglich ist g messbar. Aus $\|g_k - g_m\|_p^p \leq \frac{1}{2^{2n}}$ ergibt sich schließlich

$$\|g - g_m\|_p \leq \limsup_{k\to\infty} \|g_k - g_m\|_p \leq \frac{1}{2^{2n/p}} < \infty,$$

und folglich $\int_X |g - g_m|^p d\mu \leq 2^{-2n/p}$, und mit dem Lebesgue'schen Satz über dominante Konvergenz folgt $|g - g_m|^p \in L(X,\mu)$ beziehungsweise $[g - g_m] \in L^p(X,\mu)$. Wegen $g_m \in L^p(X,\mu)$ folgt hieraus nach Satz 12.8.5, dass $g \in L^p(X,\mu)$.

Der Fall $p = \infty$ verbleibt als Übung. □

Satz 12.8.7. *Im Falle $p = 2$ induziert die Abbildung $\langle\cdot,\cdot\rangle : L^2(X,\mu) \times L^2(X,\mu) \to \mathbb{R}$ gegeben durch*

$$\langle f, g\rangle := \int_X fg d\mu, \qquad f, g \in L^2(X,\mu),$$

ein Skalarprodukt. Es gilt

$$\|f\|_2 = \sqrt{\langle f, f\rangle}.$$

Der Raum $L^2(X,\mu)$ versehen mit $\langle\cdot,\cdot\rangle$ bildet einen Hilbertraum.

12.9 Integration in Produkträumen

12.9.1 Produktmaße

Definition 12.9.1. *Seien $(X_1, \mathcal{M}_1, \mu_1)$ und $(X_2, \mathcal{M}_2, \mu_2)$ zwei Maßräume. Die kleinste σ-Algebra in $X := X_1 \times X_2$, die die Menge*

$$\mathcal{M}_1 \times \mathcal{M}_2 = \{A_1 \times A_2 : A_1 \in \mathcal{M}_1,\ A_2 \in \mathcal{M}_2\}$$

enthält, definieren wir als das **Produkt** $\mathcal{M}_1 \otimes \mathcal{M}_2$ *der beiden σ-Algebren.*

Lemma 12.9.2. *Das Mengensystem $\mathcal{M}_1 \times \mathcal{M}_2$ ist ein Semiring und $\mathcal{R}(\mathcal{M}_1 \times \mathcal{M}_2)$ ist eine Algebra.*

Beweis. Wegen $\emptyset \in \mathcal{M}_1$ und $\emptyset \in \mathcal{M}_2$ gilt auch $\emptyset = \emptyset \times \emptyset \in \mathcal{M}_1 \times \mathcal{M}_2$. Weiter gilt für alle $A = A_1 \times A_2, B = B_1 \times B_2 \in \mathcal{M}_1 \times \mathcal{M}_2$ auch

$$A \cap B = (A_1 \times A_2) \cap (B_1 \times B_2) = (A_1 \cap B_1) \times (A_2 \cap B_2) \in \mathcal{M}_1 \times \mathcal{M}_2$$

Schließlich folgt aus

$$A \setminus B = (A_1 \times A_2) \setminus (B_1 \times B_2) = \big(\underbrace{(A_1 \setminus B_1) \times A_2}_{\in \mathcal{M}_1 \times \mathcal{M}_2}\big) \cup \big(\underbrace{(A_1 \cap B_1) \times (A_2 \setminus B_2)}_{\in \mathcal{M}_1 \times \mathcal{M}_2}\big)$$

wegen

$$\big((A_1 \setminus B_1) \times A_2\big) \cap \big((A_1 \cap B_1) \times (A_2 \setminus B_2)\big) = \emptyset,$$

dass wir $A \setminus B$ in eine endliche Vereinigung paarweise disjunkter Mengen aus $\mathcal{M}_1 \times \mathcal{M}_2$ zerlegen können. Folglich ist $\mathcal{M}_1 \times \mathcal{M}_2$ ein Semiring.

Um zu zeigen, dass $\mathcal{R}(\mathcal{M}_1 \times \mathcal{M}_2)$ eine Algebra ist, genügt es zu zeigen, dass $X_1 \times X_2 \in \mathcal{R}(\mathcal{M}_1 \times \mathcal{M}_2)$ gilt. Wegen $X_1 \in \mathcal{M}_1$ und $X_2 \in \mathcal{M}_2$ folgt dies aber sofort aus $X = X_1 \times X_2 \in \mathcal{M}_1 \times \mathcal{M}_2$. □

Lemma 12.9.3. *Seien $(X_1, \mathcal{M}_1, \mu_1)$ und $(X_2, \mathcal{M}_2, \mu_2)$ zwei Maßräume mit σ-endlichen Maßen. Dann ist $\varphi : \mathcal{M}_1 \times \mathcal{M}_2 \to \overline{\mathbb{R}}$ definiert durch*

$$\varphi(A_1 \times A_2) := \mu(A_1) \cdot \mu(A_2)$$

eine nichtnegative, additive und subvolladditive Mengenfunktionen mit $\varphi(\emptyset) = 0$, falls wir im obigen Produkt vereinbaren $0 \cdot \infty = \infty \cdot 0 = \infty$.

Beweis. Es gilt

$$\varphi(\emptyset) = \varphi(\emptyset \times \emptyset) = \mu_1(\emptyset) \cdot \mu_2(\emptyset) = 0.$$

Im nächsten Schritt zeigen wir die Additivität. Seien hierzu $A = A_1 \times A_2, B = B_1 \times B_2 \in \mathcal{M}_1 \times \mathcal{M}_2$ mit $B \subseteq A$, das heißt es gelte $B_1 \subseteq A_1$ und $B_2 \subseteq A_2$, dann folgt

$$\begin{aligned}\varphi(A) &= \mu_1(A_1)\mu_2(A_2)\\ &= \mu_1(B_1)\mu_2(A_2) + \mu_1(A_1 \setminus B_1)\mu_2(A_2)\\ &= \mu_1(B_1)\mu_2(B_2) + \mu_1(B_1)\mu_2(A_2 \setminus B_2) + \mu_1(A_1 \setminus B_1)\mu_2(A_2)\\ &= \varphi(B) + \varphi\big(B_1 \times (A_2 \setminus B_2)\big) + \varphi\big((A_1 \setminus B_1) \times A_2\big).\end{aligned}$$

Sind $B = B_1 \times B_2$ und $C = C_1 \times C_2$ disjunkt mit $B \cup C \in \mathcal{M}_1 \times \mathcal{M}_2$. Dann gilt entweder $B_1 = C_1$, $B_2 \cap C_2 = \emptyset$ oder $B_1 \cap C_1 = \emptyset$, $B_2 = C_2$. Nehmen wir ohne Beschränkung der Allgemeinheit zweiteres an und setzen $A := B \cup C = (B_1 \cup C_1) \times C_2$, dann folgt aus obiger Gleichung

$$\varphi(A) = \varphi(B) + \varphi\big(C_1 \times \emptyset\big) + \varphi\big(C_1 \times C_2\big) = \varphi(B) + \varphi(C),$$

dies bedeutet

$$\varphi(B \cup C) = \varphi(B) + \varphi(C).$$

Es verbleibt, die Subvolladditivität zu zeigen. Seien $A^{(n)} = A_1^{(n)} \times A_2^{(n)}, \quad A_i^{(n)} \in \mathcal{M}_i, \quad n \in \mathbb{N}$. Desweiteren seien

$$\chi(A^{(n)}) : X_1 \times X_2 \to \mathbb{R}, \quad (x_1, x_2) \mapsto \chi(x_1, x_2)$$

die charakteristische Funktionen von $A^{(n)}$ und $x_i \mapsto \chi_{A_i^{(n)}}(x_i)$ entsprechend.
Dann gilt:

$$\chi_{A^{(n)}}(x_1, x_2) = \chi_{A_1^{(n)}}(x_1) \cdot \chi_{A_2^{(n)}}(x_2)$$

Sei $A = \bigcup_{n=1}^{\infty} A^{(n)}$, dann gilt:

$$\chi_A(x_1, x_2) \leq \sum_{n=1}^{\infty} \chi_{A^{(n)}}(x_1, x_2) = \sum_{n=1}^{\infty} \chi_{A_1^{(n)}}(x_1) \cdot \chi_{A_2^{(n)}}(x_2)$$

Die Integration bezüglich μ_2 liefert dann

$$\begin{aligned}\int_{A_2^{(n)}} \chi_A(x_1, x_2) d\mu_2 &= \chi_{A_1^{(n)}}(x_1) \cdot \int_{(} A_2^{(n)})(x_2) d\mu_2\\ &= \chi_{A_1^{(n)}}(x_1) \cdot \mu_2(A_2^{(n)}) \text{ und}\\ \int_{A_1^{(n)}} (\int_{A_2^{(n)}} \chi_{A^{(n)}} d\mu_2) d\mu_2 &= \mu_1(A_1^{(n)})\mu_2(A_2^{(n)}) \text{ und}\\ \mu(A) = \int_A \chi_A d\mu &\leq \sum_{n=1}^{\infty} \int_{A^{(n)}} \chi_{A^{(n)}} d\mu_1 d\mu_2\\ &\leq \sum_{n=1}^{\infty} \mu_1(A_1^{(n)})\mu_2(A_2^{(n)}) = \sum_{n=1}^{\infty} \mu(A^{(n)})\end{aligned}$$

□

Satz 12.9.4. *Seien μ_i σ-endlich, dann existiert ein eindeutig bestimmtes Maß*

$$\mu : \mathcal{M}_1 \otimes \mathcal{M}_2 \to \overline{\mathbb{R}}$$

auf der σ-Algebra $\mathcal{M}_1 \otimes \mathcal{M}_2$.
Dies gilt auch für die Vollständigkeit.

Beweis. Der Beweis ist eine unmittelbare Folgerung aus den Ergebnissen und dem Satz von Hahn. □

Beispiel 12.9.5. *Sei $X_1 = \mathbb{R}^n, \quad X_2 = \mathbb{R}^m$ versehen mit dem Lebesquie-Maß μ_1, μ_2, dann ist das Produktmaß μ mit $x_1 \times x_2 = \mathbb{R}^{n+m}$ ebenfalls das Lebesquie-Maß auf $\mathbb{R}^{n+m}$ und das Produkt der σ-Algebren die Mengen aller Lebesquie-messbaren Mengen.*

Beweis. $A_1 = \prod_{i=1}^n [a_i, b_i] \in \mathcal{M}_1, \quad A_2 = \prod_{i=1}^m [a_i, b_i] \in \mathcal{M}_2$ und $A = \prod_{i=1}^{n+m} [a_i, b_i] = A_1 \times A_2$ □

2.) Funktionen auf Produktnormen

Definition 12.9.6. *Sei $A \subseteq X_1 \times X_2$, dann definieren wir die sogenannten Schnitte von A wie folgt:*

$$A(x_1) := \{x_2 \in X_2 : (x_1, x_2) \in A\} \subseteq X_2$$
$$A(x_2) := \{x_1 \in X_1 : (x_1, x_2) \in A\}$$

Satz 12.9.7. *Sei $A \in \mathcal{M}_1 \otimes \mathcal{M}_2$, dann gilt:*

$$A(x_1) \in \mathcal{M}_2 \text{ und } A(x_2) \in \mathcal{M}_1$$

Beweis. Wir betrachten die Menge $N = \{A \subseteq X_1 \times X_2 : A(x_1) \in \mathcal{M}_2 \wedge A(x_2) \in calM_1\}$.
Zu zeigen ist, dass N eine σ-Algebra, die $\mathcal{M}_1 \otimes \mathcal{M}_2$ enthält.
Sei $A \subseteq X_1 \times X_2, \quad x_1 \in X_1$ beliebig aber fest. Desweiteren sei $A \in N$, dann gilt:

$$(X \setminus A)(x_1) = X_2 \setminus A(x_1) \in \mathcal{M}_2$$

Seien $A^{(n)} \subseteq X_1 \times X_2, \quad A^{(n)} \in N$, dann gilt:

$$(\bigcup_{n=1}^{\infty} A^{(n)})(x_i) = \bigcup_{n=1}^{\infty} A^{(n)}(x_1) \in \mathcal{M}_2,$$

d.h. die Komplementenbildung und die abzählbare Vereinigung führen nicht aus N heraus, d.h. N ist eine σ-Algebra und $\mathcal{M}_1 \otimes \mathcal{M}_2 \subseteq N$. □

Satz 12.9.8. *Sei $A \in \mathcal{M}_1 \otimes \mathcal{M}_2$ und $f : X_1 \times X_2 \to \mathbb{R}$ eine messbare Funktion, dann sind für alle $x_1 \in X_1$ die Funktion $f(x_1, \cdot) : X_2 \to \overline{\mathbb{R}}$ ebenfalls messbar.*

Beweis. Sei $a > 0$ beliebig, dann ist

$$\begin{aligned} A(f > a)(x_2) &= \{(x_1, x_2) : f(x_1, x_2) > a, x_2 \text{ beliebig}\} \\ = A(x_2)(f > a) &= \{(x_1, x_2) : x_2 \text{ fest }, f(x_1, x_2) > a\} \end{aligned}$$

Da $A(f > a)$ messbar und nach Satz $A(f > a)(x_2)$ ebenfalls messbar ist, ist

$$A(x_2)(f > a) = A(x_2)(f(\cdot, x_2) > a)$$

ebenfalls messbar, d.h. $f(\cdot, x_2)$ ist messbar. □

Satz 12.9.9. *Seien μ_i endliche Maße auf X_i.*
Für $A \in \mathcal{M}_1 \otimes \mathcal{M}_2$ definieren wir:

$$\left.\begin{array}{lcl} \eta_a(x_1) & := & \mu_2(A(x_1)) \\ \xi_a(x_2) & := & \mu_1(A(x_2)) \end{array}\right\} \textit{messbar}$$

und es gilt:

$$\int_{x_2} \xi_A(x_2) d\mu_2 = \int_{x_1} \eta_A(x_1) d\mu_1$$

Beweis. Sei $A = A_1 \times A_2, \quad A \in \mathcal{M}_i, \quad i = 1, 2$, dann ist

$$\xi_A(x_2) = \mu_1(A(x_2)) = \mu(A_1)\chi_A(x_2)$$

eine Treppenfunktion und messbar.
Analog η_A und es gilt:

$$\int_{x_2} \xi_A(x_2) d\mu_2 = \mu_1(A_1) \int_{x_2} \chi_A d\mu = \mu_1(A_1)\mu_2(A_2)$$

Da das Maß μ volltständig ist, gilt die Gleichheit auf $r(\mathcal{M}_1 \otimes \mathcal{M}_2)$.
Fobini: Sei $f \in \mathcal{L}(X_1 \times X_2, \mu) \Rightarrow f(\cdot, x_2) \in \mathcal{L}(X_1, \mu) \quad f(x_1, \cdot)$ und

$$\int_{A_2} (\int_{A_1} f d\mu_1) d\mu_2 \quad (*) = \int_{A_1} (\int_{A_2} f d\mu_1) d\mu_2 \quad (*) = \int_{A_1 \times A_2} f d\mu$$

□

Bemerkung 12.9.10. *Sei $f \in \mathcal{L}(X_1 \times X_2, \mu)$.*
Die Summierbarkeit in $A_1 \times A_2$ impliziert die Existenz der iterierten Integrale $(*)$.
Die Umkehrung gilt im Allgemeinen nicht.

Satz 12.9.11 (Zusatz zu Satz). *Sind die $\mu_i \quad (i = 1, 2)$ σ-endlich, dann bleibt Satz in folgender Form richtig:*

1. *Sei*

$$D(\xi_A) := \{x_2 \in X_2 : \mu_1(A(x_2)) < +\infty\} \text{ und}$$
$$D(\eta_A) := \{x_1 \in X_1 : \mu_2(A(x_1)) < +\infty\}$$

Dann ist ξ_A (bzw. η_A) fast überall auf $D(\xi_A)$ (bzw. $D(\eta_A)$) definiert und messbar.

2. $$\xi_A \in \mathcal{L}(D(\xi_A), \mu_2) \Leftrightarrow \eta_A \in \mathcal{L}(D(\eta_A), \mu_1)$$

Die Gleichung aus Satz gilt genau dann, wenn eine dieser Bedingungen erfüllt ist.

Satz 12.9.12. *Seien μ_i σ-endliche Maße über $\mathcal{M}_i \quad (i = 1, 2)$, dann gilt für $A \in \mathcal{M}_1 \otimes \mathcal{M}_2$:*

$$\mu(A) = (\mu_1 \times \mu_2)(A) = \begin{cases} \int_{X_1} \eta_A(x_1) d\mu_1 = \int_{X_2} \xi_A(x_2) d\mu_2 & ; \quad \text{wenn } \eta_A \in \mathcal{L}(D(\eta_A), \mu) \\ \infty & ; \quad \text{sonst} \end{cases}$$

Beweis. ohne Beweis □

Satz 12.9.13. *Sei $A \in \mathcal{M}_1 \otimes \mathcal{M}_2$ mit $\mu(A) = (\mu_1 \times \mu_2)(A) < +\infty$ und $\chi_A(x_1, x_2)$ die charakteristische Funktion von A. Dann gilt:*

1. $$\chi_A(\cdot, x_2) \in \mathcal{L}(X_1, \mu_1) \text{ für fast alle } x_2 \in X_2 \text{ bezüglich } \mu_2$$
$$\chi_A(x_1, \cdot) \in \mathcal{L}(X_2, \mu_2) \text{ für fast alle } x_1 \in X_1 \text{ bezüglich } \mu_1$$

2. $$\begin{aligned} \int_{X_1 \times X_2} \chi_A d(\mu_1 \times \mu_2) &= \int_{X_2} \{\int_{X_1} \chi_A(\cdot, x_2) d\mu_1\} d\mu_2 \\ &= \int_{X_1} \{\int_{X_2} \chi_A(x_1, \cdot) d\mu_2\} d\mu_1 \end{aligned}$$

Beweis. Wegen Satz gilt:
$(\mu_1 \times \mu_2)(A) < +\infty \Leftrightarrow \xi_A \in \mathcal{L}(\cdot, \mu_2) \Leftrightarrow \eta_A \in \mathcal{L}(\cdot, \mu_1)$
$\Rightarrow \xi_A$ ist für fast alle $x_2 \in X_2$ und η_A für fast alle $x_1 \in X_1$ definiert.
Wegen der Definition von ξ_A und η_A gilt:

$$\xi_A(x_2) = \mu_1(A(x_2)) = \int_{X_1} \chi_A(\cdot, x_2) d\mu_1 \text{ und}$$
$$\eta_A(x_1) = \mu_2(A(x_1)) = \int_{X_2} \chi_A(x_1, \cdot) d\mu_2$$

$\Rightarrow$ Behauptung (1) und aus Satz folgt:

$$\int_{X_1\times X_2} \chi_A d(\mu_1\times\mu_2) = (\mu_1\times\mu_2)(A) = \int_{X_2} \xi_A(x_2)d\mu_2$$
$$= \int_{X_2}\{\int_{X_1}\chi_A(\cdot,x_2)d\mu_1\}d\mu_2$$

Analog geht der Beweis für die andere Gleichung □

Satz 12.9.14 (Fubini). *Seien* $\mu_i \quad (i=1,2)$ *auf* $\mathcal{M}_i$ *und* $f\in\mathcal{L}(X_1\times X_2,\mu_1\times\mu_2)$. *Dann gilt:*

1. $f(\cdot,x_2)\in\mathcal{L}(X_1,\mu_1)$ *für fast alle* $x_2\in X_2$ *bezüglich* μ_2 *und* $\int_{X_1} f(\cdot,x_2)d\mu_1\in\mathcal{L}(X_2,\mu_2)$ *analog*
   $f(x_1,\cdot)\in\mathcal{L}(X_2,\mu_2)$ *für fast alle* $x_1\in X_1$ *bezüglich* μ_1 *und* $\int_{X_2} f(x_1,\cdot)d\mu_2\in\mathcal{L}(X_1,\mu_1)$
2. $\int_{X_1\times X_2} fd(\mu_1\times\mu_2) = \int_{X_2}\left\{\int_{X_1} f(\cdot,x_2)d\mu_1\right\}d\mu_2 = \int_{X_1}\left\{\int_{X_2} f(x_1,\cdot)d\mu_2\right\}d\mu_1$

Beweis. Sei $f\geq 0$. Nach Satz existieren monoton wachsende Folgen $\{\varphi_n\}$ nichtnegativer Treppenfunktionen mit $\varphi_n\to f$.
Da

$$\varphi_n = \sum_{\text{endlich}} c_i^{(n)}\chi_{E_i}$$

gilt nach Satz die Behauptung für jeden Summanden von φ_n und damit auch für φ_n selbst, also z.B.

$$\int_{X_1\times X_2}\varphi_n d(\mu_1\times\mu_2) = \int_{X_2}\underbrace{\left\{\int_{X_1}\varphi_n(\cdot,x_2)d\mu_1\right\}}_{\Phi_n}d\mu_2 \quad (*)$$

Wegen

$$\int_{X_2}\Phi_n(x_2)d\mu_2 = \int_{X_1\times X_2}\varphi_n d(\mu_1\times\mu_2) \leq \int_{X_1\times X_2} fd(\mu_1\times\mu_2) < +\infty$$

folgt aus Satz die Folgerung

$$\lim_{n\to\infty}\Phi_n\in\mathcal{L}(X_2,\mu_2) \quad (**) \Rightarrow \lim_{n\to\infty}\Phi_n \text{ ist fast überall endlich auf } X_2$$

und

$$\lim_{n\to\infty}\int_{X_2}\Phi_n(x_2)d\mu_2 = \int_{X_2}(\lim_{n\to\infty}\Phi_n)d\mu_2$$

Da $\lim\limits_{n\to\infty}\int_{X_1}\varphi_n(\cdot,x_2)d\mu_1 < +\infty$ für fast alle $x_2\in X_2$

$\Rightarrow \lim\limits_{n\to\infty}\varphi_n(\cdot,x_2) = f(\cdot,x_2)\in\mathcal{L}(X_1,\mu_1)$ für fast alle $x_2\in X_2$

$\Rightarrow \lim\limits_{n\to\infty}\int_{X_1}\varphi_n(\cdot,x_2)d\mu_1 = \int_{X_1} f(\cdot,x_2)d\mu_1 \quad (***)$ für fast alle $x_2\in X_2$

Nun lassen wir in $(*)$ $\quad n \to \infty$ laufen:

$$\begin{aligned}\int_{X_1\times X_2} f d(\mu_1\times\mu_2) &= \lim_{n\to\infty}\int_{X_1\times X_2}\varphi_n d(\mu_1\times\mu_2) = \lim_{n\to\infty}\int_{X_2}\Phi_n(x_2)d\mu_2\\ &= \int_{X_2}\left\{\lim_{n\to\infty}\int_{X_1}\varphi_n(\cdot,x_2)d\mu_1\right\}d\mu_2\\ &= \int_{X_2}\left\{\int_{X_1} f(\cdot,x_2)d\mu_1\right\}d\mu_2\end{aligned}$$

und aus $(**)$, $(***) \Rightarrow \int_{X_1} f(\cdot,x_2)d\mu_1 \in \mathcal{L}(X_2,\mu_2)$
Die andere Beziehung kann man analog zeigen. □

Satz 12.9.15. *Sind die Maße μ_i σ-endlich und vollständig, dann gilt Satz auch für $f \in \mathcal{L}(X_1 \times X_2, (\mu_1 \times \mu_2))$*

Bemerkung 12.9.16. *1. Ist $f : X_1 \times X_2 \to \mathbb{R}^1$, $\quad \mathcal{M}_1 \otimes \mathcal{M}_2$-messbar und ist $|f(\cdot,x_2)| \in \mathcal{L}(X_1,\mu_1)$ für fast alle $x_2 \in X_2$ sowie*

$$\int_{X_1}|f(\cdot,x_2)|d\mu_1 \in \mathcal{L}(X_2,\mu_2),$$

so ist $f \in \mathcal{L}(X_1 \times X_2, \mu_1 \times mu_2)$.

2. Aus der Existenz des Doppelintegrals folgt die Existenz der iterierten Integrale und deren Gleichheit. Die Umkehrung gilt im allgemeinen nicht.

12.10 Zusatz:Zerlegung von Mengenfunktionen

Definition 12.10.1. *Sei $(X,\mathcal{M})$ ein meßbarer Raum. Das Mengensystem $\mathcal{M}^E$ bezeichne das System aller meßbaren Teilmengen von E.*

Satz 12.10.2. *Sei η eine beliebige volladditive Mengenfunktion auf $\mathcal{M}$ mit $\eta(\emptyset) = 0$, dann ist die Mengenfunktion*

$$\eta^+(E) = \sup_{e\in\mathcal{M}^E}\eta(e), \qquad E \in \mathcal{M},$$

ein auf $\mathcal{M}$ definiertes Maß. Ist zusätzlich $\eta(A) < \infty$ für alle $A \in \mathcal{M}$, so ist das Maß η^+ endlich.

Beweis. 1. Da $\emptyset \subseteq E$ für alle $E \in \mathcal{M}$ folgt $0 = \eta(\emptyset) \le \eta^+(E)$, das heißt η^+ ist eine nichtnegative Mengenfunktion.

2. Wir zeigen nun die Volladditivität von η^+: Sei $E = \bigcup_{n=1}^\infty E_n$ mit $E_n \in \mathcal{M}$ und $E_i \cap E_j = \emptyset$ für alle $i \ne j$. Wir betrachten eine beliebige Menge $e \in \mathcal{M}^E$. Setzen wir $e_n := E_n \cap e \subseteq E_n$, so folgt $e = \bigcup_{n=1}^\infty e_n$ und $e_i \cap e_j = \emptyset$ für alle $i \ne j$. Dies bedeutet

$$\eta(e) = \sum_{n=1}^\infty \eta(e_n) \le \sum_{n=1}^\infty \eta^+(E_n).$$

Da $e \in \mathcal{M}^E$ beliebig gewählt war, folgt hieraus

$$\eta^+(E) \leq \sum_{n=1}^{\infty} \eta^+(E_n). \tag{12.15}$$

Wir müssen nun auch die entgegengesetzte Ungleichung zeigen. Wenn $\eta^+(E) = \infty$ ist, dann folgt aus (12.15) die Behauptung. Sei also $\eta^+(E_n) < \infty$. Angenommen, es existiert ein $m \in \mathbb{N}$ mit $\eta^+(E_m) = \infty$. Aus der Definition von η^+ folgt, dass dann für jedes $n \in \mathbb{N}$ ein $e_n \in \mathcal{M}^{E_m}$ existiert mit $\eta(e_n) > n$. Wegen $e_n \subseteq E_m \subseteq E$ folgt $e_n \in \mathcal{M}^E$ für alle $n \in \mathbb{N}$ und damit $\eta^+(E) = \infty$, was im Widerspruch zur Annahme steht. Folglich muss $\eta^+(E_n) < \infty$ für alle $n \in \mathbb{N}$ gelten.

Sei nun $\varepsilon > 0$ beliebig. Aus der Definition von η^+ folgt, dass für jedes $n \in \mathbb{N}$ ein $e_n \in \mathcal{M}^{E_n}$ existiert mit $\eta(e_n) > \eta^+(E_n) - \varepsilon/2^n$. Wir setzen $e := \bigcup_{n=1}^{\infty} e_n \subseteq E$, das heißt $e \in \mathcal{M}^E$, und wegen $E_i \cap E_j = \emptyset$ für $i \neq j$ folgt auch $e_i \cap e_j = \emptyset$ für $i \neq j$. Folglich gilt

$$\eta^+(E) \geq \eta(e) = \sum_{n=1}^{\infty} \eta(e_n) > \sum_{n=1}^{\infty} \eta^+(E_n) - \varepsilon,$$

dies bedeutet für $\varepsilon \to 0$ dass

$$\eta^+(E) \geq \sum_{n=1}^{\infty} \eta^+(E_n).$$

Zusammen mit (12.15) folgt die Behauptung.

3. Wir zeigen zuletzt, dass unter der Voraussetzung $\eta(A) < \infty$ für alle $A \in \mathcal{M}$ folgt η^+ ist σ-endlich. Dazu setzen wir für $E \in \mathcal{M}$

$$\mathcal{M}_n^E := \{e \in \mathcal{M}^E : \eta(e) > n\}.$$

Sei nun $E \in \mathcal{M}$ mit $\eta^+(E) = \infty$, dann folgt $\mathcal{M}_n^E \neq \emptyset$ für alle $n \in \mathbb{N}$. Wir zeigen, dass bei festem n ein $e \in \mathcal{M}_n^E$ existiert mit $\eta^+(e) = \infty$.

Dazu nehmen wir an, dass eine solche Menge nicht existiert, dies bedeutet für alle $e \in \mathcal{M}_n^E$ gilt $\eta^+(e) < \infty$. Sei etwa E_1 eine solche Menge. Wir setzen $\overline{E_1} := E \setminus E_1$, das heißt $\infty = \eta^+(E) = \eta^+(E_1) + \eta^+(\overline{E_1})$ und folglich $\eta^+(\overline{E_1}) = \infty$. Wegen $\overline{E_1} \subseteq E$ gilt für alle $e \in \mathcal{M}_n^{\overline{E_1}}$ dass $\eta^+(e) < \infty$. Sei etwa E_2 eine solche Menge, dann ergibt sich wie eben für $\overline{E_2} := E_1 \setminus E_2$ dass $\eta^+(\overline{E_2}) = \infty$. Ferner gilt $E_1 \cap E_2 = \emptyset$. Setzen wir dieses Verfahren fort, liefert es uns eine Folge $\{E_k\} \subseteq \mathcal{M}$ mit $E_i \cap E_j = \emptyset$ für $i \neq j$ und $\eta(E_k) > n$ für alle $k = 1, 2, \ldots$. Daraus folgt wegen der Volladditivität von η

$$\eta\left(\bigcup_{k=1}^{\infty} E_k\right) = \sum_{k=1}^{\infty} \eta(E_k) > \sum_{k=1}^{\infty} n = \infty,$$

was im Widerspruch steht zu $\eta(A) < \infty$ für alle $A \in \mathcal{M}$. Folglich schließen wir, dass für jedes feste $n \in \mathbb{N}$ ein $e \in \mathcal{M}_n^E$ existiert mit $\eta^+(e) = \infty$.

Um die σ-Endlichkeit von η^+ zu zeigen, genügt es zu zeigen, dass $\eta^+(X) < \infty$ ist. Dazu nehmen wir an, dass $\eta^+(X) = \infty$ gilt. Dann existiert aber ein $A_1 \in \mathcal{M}$ mit $\eta(A_1) > 1$ und $\eta^+(A_1) = \infty$. Folglich existiert ein $A_2 \in \mathcal{M}$ mit $A_2 \subseteq A_1$ $\eta(A_2) > 2$ und $\eta^+(A_2) = \infty$. Dieses Verfahren fortgesetzt liefert eine Folge $\{A_n\} \subseteq \mathcal{M}$ mit $A_{n+1} \subseteq A_n$, $\eta(A_n) > n$ und $\eta^+(A_n) = \infty$ für alle $n \in \mathbb{N}$. Nun setzen wir $A := \bigcap_{n=1}^{\infty} A_n \in \mathcal{M}$. Da $n < \eta(A_n) < \infty$ gilt, ist $\eta(A_n)$ endlich für alle $n \in \mathbb{N}$. Dann folgt aus Satz 12.2.6, dass $\lim_{n\to\infty} \eta(A_n) = \eta(A) = \infty$ ist. Dies ist ein Widerspruch zur Annahme und daher gilt

$$\eta^+(X) < \infty$$

□

Proposition 12.10.3 (Jordanscher Zerlegungssatz). *Jede auf einer σ-Algebra $\mathcal{M}$ definierte volladditive Mengenfunktion η läßt sich als Differenz zweier Maße darstellen, wobei eines der beiden Maße endlich ist.*

Beweis. Wir setzen

$$\eta^-(E) := \eta^+(E) - \eta(E), \qquad E \in \mathcal{M}.$$

Dann ist η^- volladditiv und wegen $\eta(E) \leq \eta^+(E)$ nichtnegativ, also ist η^- eine Maß. Da η nur einen der beiden uneigentlichen Werte $\pm\infty$ annimmt, folgt die letzte Behauptung aus dem vorhergehenden Satz. □

Definition 12.10.4. *Die Funktionen η^+, η^- bzw. $|\eta| := \eta^+ + \eta^-$ heißen* **positive Variation**, **negative Variation** *bzw.* **Totalvariation** *oder* **Betrag** *von η.*

Bemerkung 12.10.5.

1. *Die negative Variation η^- kann auch folgerndermaßen definiert werden*

$$\eta^-(E) := -\inf_{e\in\mathcal{M}^E} \eta(e), \qquad E \in \mathcal{M}.$$

2. *Es gilt $|\eta(E)| \leq |\eta|(E)$ für alle $E \in \mathcal{M}$.*

Satz 12.10.6 (Hahn). *Sei $(X, \mathcal{M})$ ein meßbarer Raum und η eine volladditive Mengenfunktion auf $\mathcal{M}$. Dann existieren Mengen $X^+, X^- \in \mathcal{M}$ so dass*

$$X^+ \cap X^- = \emptyset, \qquad X^+ \cup X^- = X,$$

und

$$\eta^+(E) = \eta(E \cap X^+), \qquad \eta^-(E) = \eta(E \cap X^-),$$

für alle $E \in \mathcal{M}$.

Beweis. Da η nur einen der uneigentlichen Werte $\pm\infty$ annehmen kann, sei der Bestimmtheit halber $\eta(A) < +\infty$ für alle $A \in \mathcal{M}$ angenommen. Nach Satz 12.10.2 ist dann $\eta^+(X) < +\infty$. Deswegen existiert für alle $k = \mathbb{N}$ ein $E_k \in \mathcal{M}$ mit $\eta(E_k) > \eta^+(X) - 1/2^k$. Setzen wir $A_n := \bigcup_{k=n}^{\infty} E_k$, dann folgt $A_n \supseteq E_n$ und

$$\eta^+(A_n) \geq \eta^+(E_n) \geq \eta(E_n) > \eta^+(X) - 1/2^k.$$

Sei nun $X^+ = \bigcap_{n=1}^{\infty} A_n$, dann gilt einerseits wegen $A_n \supseteq A_{n+1}$ und $\eta^+(A_n) < \infty$ dass

$$\eta^+(X^+) = \lim_{n\to\infty} \eta^+(A_n) \geq \eta^+(X).$$

Andererseits gilt aber aufgrund von $\eta^+ \geq 0$ und $X^+ \subseteq X$ dass

$$\eta^+(X^+) \leq \eta^+(X).$$

Beide Abschätzungen zusammengefasst ergeben

$$\eta^+(X^+) = \eta^+(X).$$

Weiterhin folgt aus

$$\eta^-(E_k) = \eta^+(E_k) - \eta(E_k) \leq \eta^+(X) - \left(\eta^+(X) - \frac{1}{2^k}\right) = \frac{1}{2^k}$$

dass

$$0 \leq \eta^-(A_n) \leq \sum_{k=n}^{\infty} \eta^-(E_k) \leq \sum_{k=n}^{\infty} \frac{1}{2^k} < \infty.$$

Wegen $A_n \supseteq A_{n+1}$ und $\eta^-(A_n) < \infty$ folgt hieraus

$$\eta^-(X^+) = \lim_{n\to\infty} \eta^-(A_n) = 0,$$

dies impliziert

$$\eta^+(X) = \eta^+(X^+) = \eta^+(X^+) - \eta^-(X^+) = \eta(X^+).$$

Sei nun $E \in \mathcal{M}$ beliebig und $\overline{E} = X \setminus E$. Dann gilt

$$\eta(E \cap X^+) \leq \eta^+(E), \qquad \eta(\overline{E} \cap X^+) \leq \eta^+(\overline{E}).$$

Die Annahme $\eta(E \cap X^+) < \eta^+(E)$ führt wegen

$$\begin{aligned}\eta^+(E) + \eta^+(\overline{E}) &> \eta(E \cap X^+) + \eta(\overline{E} \cap X^+) = \eta\big((E \cap X^+) \cup (\overline{E} \cap X^+)\big)\\ &= \eta(X^+ \cap X) = \eta(X^+) = \eta^+(X)\\ &= \eta^+(E) + \eta^+(\overline{E})\end{aligned}$$

zum Widerspruch. Mit anderen Worten, es gilt

$$\eta(E \cap X^+) = \eta^+(E)$$

für alle $E \in \mathcal{M}$.

Wir setzen nun $X^- := X \setminus X^+$ und zeigen

$$\eta^-(E) = -\eta(E \cap X^-).$$

Es gilt

$$\eta^-(E) = \eta^+(E) - \eta(E) = \eta(E \cap X^+) - \eta(E).$$

Da $\eta(E \cap X^+) < \infty$ und $E \cap X^+ \subseteq E$ folgt

$$\eta^-(E) = -\eta\big(E - (E \cap X^+)\big) = -\eta\big(E \cap (X - X^+)\big) = -\eta(E \cap X^-).$$

□

Proposition 12.10.7. *Für alle $A \in \mathcal{M}$ gilt*

$$\eta(A \cap X^+) \geq 0, \qquad \eta(A \cap X^-) \leq 0,$$

und

$$\begin{aligned} \eta(E) &\geq 0, \qquad E \in \mathcal{M}^{X^+}, \\ \eta(E) &\leq 0, \qquad E \in \mathcal{M}^{X^-}. \end{aligned}$$

Nun wollen wir einen weiteren Zerlegungssatz vorbereiten.

Definition 12.10.8. *Sei $(X, \mathcal{M}, \mu)$ ein Maßraum. Eine auf $\mathcal{M}$ definierte Mengenfunktion η heißt absolut stetig bezüglich des Maßes μ, wenn für alle $\varepsilon > 0$ ein $\delta > 0$ existiert so dass $|\eta(E)| < \varepsilon$ für alle $E \in \mathcal{M}$ mit $\mu(E) < \delta$.*

Satz 12.10.9. *Wenn μ ein σ-endliches Maß ist und η eine additive, bezüglich μ absolut stetige Mengenfunktion auf $\mathcal{M}$ ist, dann ist η volladditiv auf $\mathcal{M}$.*

Beweis. Sei $E = \bigcup_{k=1}^{\infty} E_k$ mit $E_k \in \mathcal{M}$ und $E_k \cap E_l = \emptyset$ für alle $k \neq l$. Setzen wir $A_n := \bigcup_{k=n+1}^{\infty} E_k$, so folgt $E = A_n \cup \bigcup_{k=1}^{n} E_k$, das heißt

$$\eta(E) = \eta(A_n) + \sum_{k=1}^{n} \eta(E_k).$$

Wegen $\mu(E) = \sum_{k=1}^{\infty} \mu(E_k) < \infty$ ergibt sich hieraus

$$\mu(A_n) = \sum_{k=n+1}^{\infty} \eta(E_k) \overset{n\to\infty}{\longrightarrow} 0.$$

Die Absolutstetigkeit von η liefert $\eta(A_n) \overset{n\to\infty}{\longrightarrow} 0$, das heißt

$$\eta(E) = \lim_{n\to\infty} \sum_{k=1}^{n} \eta(E_k) = \sum_{k=1}^{\infty} \eta(E_k).$$

□

Satz 12.10.10. *Sei η eine volladditive Mengenfunktion auf $\mathcal{M}$. Wenn η absolut stetig ist bezüglich μ, dann sind es auch die Variationen η^+, η^- und $|\eta|$.*

Beweis. Es genügt zu zeigen, dass η^+ absolut stetig ist. Sei also $\varepsilon > 0$ beliebig aber fest. Da η absolut stetig ist, existiert ein $\delta > 0$ mit $|\eta(E)| < \varepsilon$ für alle $E \in \mathcal{M}$ mit $\mu(E) < \delta$. Dies impliziert $\mu(e) < \delta$ und damit $|\eta(e)| < \varepsilon$ für alle $e \in \mathcal{M}^E$. Hieraus ergibt sich

$$0 \leq \eta^+(E) \leq \sup_{e \in \mathcal{M}^E} \eta(e) \leq \varepsilon$$

für alle $E \in \mathcal{M}$ mit $\mu(E) < \delta$. □

Bemerkung 12.10.11. *Wenn $|\eta|$ oder η^+ und η^- absolut stetig bezüglich μ sind, so auch η.*

Satz 12.10.12. *Sei η eine auf $\mathcal{M}$ endliche und volladditive Mengenfunktion. η ist genau dann absolut stetig bezüglich μ, wenn für alle $E \in \mathcal{M}$ aus $\mu(E) = 0$ stets $\eta(E) = 0$ folgt.*

Beweis. "$\Rightarrow$": Sei η absolut stetig und $\varepsilon > 0$ beliebig, dann folgt für alle $E \in \mathcal{M}$ mit $\mu(E) = 0 < \delta$ dass $|\eta(E)| < \varepsilon$. Für $\varepsilon \to 0$ ergibt sich daraus $\eta(E) = 0$.

"$\Leftarrow$": Wegen Proposition 12.10.3 und dem vorhergehenden Satz können wir ohne Beschränkung der Allgemeinheit η als nichtnegativ annehmen.

Angenommen, η ist nicht absolut stetig. Das heißt, es existiert ein $\varepsilon > 0$, so dass für alle $k \in \mathbb{N}$ ein $E_k \in \mathcal{M}$ existiert mit $\mu(E_k) < 1/2^k$ und $\eta(E_k) \geq \varepsilon$.

Für $n \in \mathbb{N}$ setzen wir $A_n := \bigcup_{k=n}^{\infty} E_k$ und $A := \bigcap_{n=1}^{\infty} A_n$. Dann gilt $A_n \supseteq A_{n+1}$ für alle $n \in \mathbb{N}$ und folglich

$$0 \leq \mu(A_n) \leq \sum_{k=n}^{\infty} \mu(E_k) < \sum_{k=n}^{\infty} \frac{1}{2^k} < \infty.$$

Hieraus ergibt sich

$$\mu(A) = \lim_{n \to \infty} \mu(A_n) = 0,$$

dies bedeutet nach Voraussetzung $\eta(A) = 0$.

Andererseits folgt aber aus $E_n \subseteq A_n$ für alle $n \in \mathbb{N}$ dass $\eta(A_n) \geq \eta(E_n) \geq \varepsilon$. Da η endlich ist, ergibt sich

$$\eta(A) = \lim_{n \to \infty} \eta(A_n) \geq \varepsilon > 0.$$

Dies ist ein Widerspruch. □

Proposition 12.10.13. *Sei $f \in L(X, \mu)$. Dann ist die Mengenfunktion*

$$\eta(E) = \int_E f d\mu, \qquad E \in \mathcal{M},$$

auf $\mathcal{M}$ absolut stetig bezüglich μ.

Beweis. η ist endlich und nach Satz 12.5.7 volladditiv. Weiterhin gilt $\int_E f d\mu = 0$ genau dann, wenn $\mu(E) = 0$. □

Definition 12.10.14. *Sei $(X, \mathcal{M}, \mu)$ ein Maßraum. Die auf $\mathcal{M}$ volladditive Mengenfunktion η heißt* **singulär** *bezüglich μ, wenn ein $N \in \mathcal{M}$ existiert mit $\mu(N) = 0$ und*

$$\eta\big(A \cap (X \setminus N)\big) = 0$$

für alle $A \in \mathcal{M}$.

Bemerkung 12.10.15.

1. *Man sagt, dass η auf N konzentriert ist, wenn $\eta\big(A \cap (X \setminus N)\big) = 0$ ist.*

2. *Wenn η ein Maß auf $\mathcal{M}$ ist, so ist auch μ bezüglich η singulär, da offenbar μ auf $X \setminus N$ konzentriert ist.*

3. *Wenn η gleichzeitig absolut stetig und singulär bezüglich μ ist, dann ist $\eta \equiv 0$.*

Satz 12.10.16 (Lebesgue'scher Zerlegungssatz). *Sei $(X, \mathcal{M}, \mu)$ ein Maßraum mit $\mu(X) < \infty$. Jede auf $\mathcal{M}$ endliche, volladditive Mengenfunktion η lässt sich eindeutig als Summe einer bezüglich μ absolut stetigen Mengenfunktion η_a und einer singulären Funktion η_s darstellen, dies bedeutet*

$$\eta(A) = \eta_a(A) + \eta_s(A)$$

für alle $A \in \mathcal{M}$.

Beweis. Wir zeigen zunächst die Eindeutigkeit einer solchen Darstellung. Sei etwa $\eta = \eta_a' + \eta_s'$ eine weitere Darstellung, dann folgt $\eta_a - \eta_a' = \eta_s - \eta_s' = 0$, das heißt $\eta_a = \eta_a'$ und $\eta_s = \eta_s'$.

Nun konstruieren wir die Mengenfunktionen η_a und η_s. Ohne Beschränkung der Algemeinheit sei η ein Maß. Dann ist $\varphi = \mu + \eta$ ebenfalls ein Maß auf $\mathcal{M}$. Wir betrachten $H := L_2(X, \mathcal{M}, \varphi)$, einen Hilbertraum mit dem Skalarprodukt

$$\langle f, g \rangle = \int_X f g d\varphi.$$

Wir definieren ein Funktional $L : H \to \overline{\mathbb{R}}$ durch

$$L(f) = \int_X f d\eta.$$

Dieses Funktional ist offenbar linear und mit einer geeigneten Konstante $c \in \mathbb{R}$ gilt

$$|L(f)|^2 = |\int_X f d\eta|^2 \leq \mu(X) \cdot \int_X |f|^2 d\eta \leq c \cdot \int_X |f|^2 d\varphi \leq c \cdot \|f\|_H^2.$$

Folglich ist L ein lineares, stetiges Funktionel über H. Nach dem Darstellungssatz von Riesz existiert ein $h \in H$ mit $L(f) = \langle f, h \rangle$ für alle $f \in H$, das heißt

$$\int_X f d\eta = \int_X f h d\varphi$$

für alle $f \in H$. Hieraus folgt

$$\int_X f(1-h)d\eta = \int_X f d\eta - \int_X f h d\eta = \int_X f h d\varphi - \int_X f \cdot h d\eta = \int_X f h d\mu. \tag{12.16}$$

Setzen wir $f := \chi_A$ mit beliebigem $A \in \mathcal{M}$, so ist

$$\eta(A) = \int_A d\eta = \int_A h d\varphi$$

und daher

$$\int_A f(1-h)d\eta = \int_A f h d\mu. \tag{12.17}$$

Im nächsten Schritt zeigen wir nun $0 \leq h \leq 1$ fast überall auf X bezüglich φ. Für $A := X(h > 1)$ ergibt sich

$$\eta(A) = \int_A h d\varphi \geq \int_A 1 d\varphi = \varphi(A) = \eta(A) + \mu(A),$$

dies bedeutet wegen der Nichtnegativität von μ, dass gilt $\mu(A) = 0$ und weiter

$$\eta(A) = \int_A h(d\mu + d\eta) = \int_A h d\mu.$$

Wir wollen nun $\eta(A) > 0$ annehmen und setzen $A_n := X(h \geq 1 + 1/n)$ für alle $n \in \mathbb{N}$. Dann gilt $A_1 \subseteq A_2 \subseteq \ldots$ und $A = \bigcup_{n=1}^{\infty} A_n$, woraus folgt

$$\lim_{n \to \infty} \eta(A_n) = \eta(A) > 0.$$

Somit existiert ein $m \in \mathbb{N}$ mit $\eta(A_m) > 0$. Wegen $A \subseteq A_m$ ergibt sich $0 \leq \mu(A_m) \leq \mu(A) = 0$, das heißt

$$\eta(A_n) = \int_{A_n} h d\varphi = \int_{A_n} h d\eta \geq (1 + 1/n)\eta(A_n)$$

Dies ist aber ein Widerspruch, weshalb insgesamt folgt

$$\eta(A) = \mu(A) = \varphi(A) = 0.$$

Sei nun $B = X(h < 0)$, dann gilt

$$0 \leq \eta(B) = \int_B h d\varphi \leq 0 \cdot \varphi(B) = 0,$$

das heißt $h = 0$ fast überall auf B bezüglich φ. Mit anderen Worten, $\varphi(B) = 0$ oder $0 \leq h \leq 1$ fast überall auf B bezüglich φ. Für alle $n \in \mathbb{N}$ gilt folglich

$$\int_X (h^n)^2 d\varphi \leq \varphi(X) < \infty,$$

das heißt $h^n \in L_2(X, \varphi)$.

Sei nun $N = X(h = 1)$, dann liefert (12.16) die Gleichung

$$\mu(N) = \int_N d\mu = \int_N h d\mu = \int_X \chi_N h d\mu = \int_X \chi_N (1-h) d\eta = \int_N (1-h) d\eta = 0.$$

Definieren wir nun $\eta_s(A) := \eta(A \cap N)$ für alle $A \in \mathcal{M}$, dann ist η_s ist ein Maß, das wegen

$$\eta_s\big(A \cap (X \setminus N)\big) = \eta\big(A \cap N \cap (X \setminus N)\big) = \eta(A) = 0$$

bezüglich μ singulär ist. Weiter setzen wir

$$f_n = \begin{cases} \sum_{k=0}^n h^k & \text{auf } X \setminus N, \\ 0 & \text{auf } N, \end{cases}$$

und erhalten aus (12.16)

$$0 \leq \int_X f_n h d\mu = \int_X f_n (1-h) d\eta = \int_X (1 - h^{n+1}) d\eta \leq \eta(X) < \infty.$$

Insbesondere gilt $f_1 h \leq f_2 h \leq \ldots$.

Setzen wir $g(x) = \lim_{n\to\infty}(f_n h)(x)$, so folgt aus Satz 12.6.2 $g \in L(X, \mu)$. Weiter folgt mit (12.17) für beliebige $A \in \mathcal{M}$, dass

$$\int_{A\setminus N} f_n h d\mu = \int_{A\setminus N} f_n (1-h) d\eta = \int_{A\setminus N} (1 - h^{n+1}) d\eta.$$

Da $h^{n+1} \to 0$ auf $A \setminus N$, ergibt sich für $n \to \infty$ hieraus

$$\int_{A\setminus N} g d\mu = \int_{A\setminus N} d\eta = \eta(A \setminus N).$$

Wegen $\mu(N) = 0$ ist

$$\eta(A \setminus N) = \int_A g d\mu =: \eta_a(A)$$

absolut stetig bezüglich μ und folglich gilt

$$\eta(A) = \eta(A \setminus N) + \eta(A \cap N) = \eta_a(A) + \eta_s(A).$$

□

Bemerkung 12.10.17. *Der Satz bleibt richtig, wenn μ und η σ-endlich sind.*

Satz 12.10.18 (Radon-Nikodym). *Sei $(X, \mathcal{M}, \mu)$ ein Maßraum mit dem σ-endlichem Maß μ und η eine auf $\mathcal{M}$ endliche, volladditive, bezüglich μ absolut stetige Mengenfunktion. Dann existiert ein $f \in L(X, \mu)$, so daß gilt*

$$\eta(E) = \int_E f d\mu$$

für alle $E \in \mathcal{M}$. Dabei ist f bis auf Äquivalenz eindeutig bestimmt.

Beweis. Nach dem Lebegueschen Zerlegungssatz existiert ein $f \in L(X, \mu)$ mit

$$\eta(E) = \int_E f d\mu + \eta_s(E).$$

Da η selbst schon absolut stetig ist, folgt $\eta_s(E) = 0$ für alle $E \in \mathcal{M}$. Die Eindeutigkeit fon f sieht man wie folgt ein. Sei etwa $g \in L(X, \mu)$ ebenfalls eine Funktion mit

$$\eta(E) = \int_E g d\mu,$$

dann folgt

$$\int_E (f - g) d\mu = 0$$

für alle $E \in \mathcal{M}$, dies bedeutet $f - g = 0$ fast überall auf E. □

Bemerkung 12.10.19.

1. *Der Satz gilt auch für eine σ-endliche Mengenfunktion η.*
2. *Die Bedingung $f \in L(X, \mu)$ impliziert die σ-Endlichkeit von η.*
3. *Die Funktion f kann so gewählt werden, dass alle Funktionswerte $f(x)$ endlich sind.*

Definition 12.10.20. *Sei f messbar auf dem Maßraum $(X, \mathcal{M}, \mu)$ und es existiere das Integral $\int_X f d\mu$. Die Mengenfunktion*

$$\eta(A) = \int_A f d\mu, \qquad A \in \mathcal{M},$$

heißt **unbestimmtes Integral** *der Funktion f und f heißt* **Ableitung** *von η nach μ, kurz*

$$f = \frac{d\eta}{d\mu}.$$

Satz 12.10.21 (Variablensubstitution). *Sei $(X, \mathcal{M})$ ein messbarer Raum mit den σ-endlichen Maßen μ und ν. Wenn ν bezüglich μ absolut stetig ist, so gilt für jede auf X messbare Funktion f die Gleichung*

$$\int_A f d\nu = \int_A f \frac{d\nu}{d\mu} d\mu, \qquad A \in \mathcal{M},$$

sofern eines der Integrale existiert.

Beweis. Da ν bezüglich μ absolut stetig ist, so existiert nach dem Satz von Radon-Nikodym eine messbare Funktion $g \geq 0$ mit $\frac{d\nu}{d\mu} = g$.

Ist $f \geq 0$ eine messbare Treppenfunktion, etwa

$$f(x) = \sum_{i=1}^{n} c_i \chi_{E_i}(x),$$

dann gilt

$$\begin{aligned}\int_A f \frac{d\nu}{d\mu} d\mu = \int_A f g d\mu = \sum_{i=1}^{m} \int_{E_i \cap A} f g d\mu = \sum_{i=1}^{m} c_i \int_{E_i \cap A} g d\mu = \sum_{i=1}^{m} c_i \nu(E_i \cap A) \\ = \int_A f d\nu,\end{aligned}$$

das heißt die Behauptung.

Sei nun $f \geq 0$ beliebig. Dann existieren messbare, nichtnegative und monoton wachsende Treppenfunktionen $\{f_n\}_{n \in \mathbb{N}}$ mit $f_n(x) \to f(x)$ für alle $x \in X$. Aus dem vorher gezeigten folgt

$$\int_A f_n g d\mu = \int_A f_n d\nu$$

und da $g \geq 0$ folgt weiter $0 \leq f_n g \leq f_{n+1} g$ für alle $n \in \mathbb{N}$. Nach dem Satz von Bepo-Levi gilt dann

$$\int_A f g d\mu = \lim_{n \to \infty} \int_A f_n g d\mu = \lim_{n \to \infty} \int_A f_n d\nu.$$

Für beliebiges Funktionen f wenden folgt die Behauptung durch Betrachten von $f^{\pm}$ an. $\square$